Uncertainty Associated with the Measurement of Airborne Sound Insulation in the Field

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy

By

William Austin Whitfield

School of Architecture, University of Liverpool

November 2013

Dedication: For Janet, Lottie & Bill. I got there in the end

Acknowledgements

I would like to thank my Supervisor Professor Barry Gibbs, who throughout nearly seven years of supervision has provided support, advice and guidance both personally and professionally. It has been a privilege Sir.

Secondly I would like to thank my colleagues at the Acoustics Research Unit, in particular Dr. Gary Seiffert and Dr Carl Hopkins for advice and useful discussions on a wide range of subjects.

Thanks must also go to the noise.co.uk test team, you know who you are. In particular the organisational skills of Steve McKeever who played a large part in making the experiments run smoothly.

Finally, I would like to thank my wife, Janet and daughters Kate and Lauren who I hope to see more often! Thank you for your support.

i. Abstract

Understanding uncertainty is an important part of any scientific measurement process and the ability to evaluate and understand uncertainty is a requirement of the International Standards for quality control.

The basic uncertainties relating to the measurement of airborne sound insulation in the field can be assessed using the methods in BS5725. However, identifying the components that contribute to the total variability is beyond the scope of the standard and more detailed information requires a more advanced approach.

Recent developments in the "Guide to the expression of uncertainty in measurement" (GUM) suggest an approach can be used where identification of the input variables and their likely contribution will result in a solution that can be modelled providing enough information is available. However, recent research on uncertainty in sound insulation using GUM has identified problems involving the correlation between frequency bands, which leads to an overestimate of uncertainty.

An empirical approach is used in this thesis, which incorporates advanced analysis of variance (ANOVA) and a specific model called Gauge Repeatability and Reproducibility (GRR). It enables the components of variance in the measurement system to be partitioned and provides an estimate of their contribution. In addition, ANOVA highlights any interaction between factors. In the GRR, carried out on a lightweight timber floor and a heavyweight concrete floor, significant interaction was detected between the operator and part.

Good agreement is obtained in the repeatability and reproducibility calculated for each construction and the samples are combined with measurements of test elements that provide a wider range of sound insulation performance. The uncertainty associated with the instrumentation, operator, interaction and part are calculated in each case.

It is shown that the interaction component is important and should be contained in any approach evaluating uncertainty. Further evidence reveals that the total uncertainty in the measurement process is dependent on the construction being measured.

Contents

ii. List of Figures

iii. List of Tables

1 Introduction

"They [There] are many apparent discrepancies on the published data on sound-insulation. These discrepancies may not be real but they are none-the-less responsible for a great deal of unfortunate, and unnecessary, confusion amongst architects, builders and even acoustical engineers. In the absence of satisfactory data, the inquirer may have doubts concerning the reliability of all published data on sound-insulation."

Vern O Knudsen 1929 [1].

1.1 Statement of the problem

Current Building Regulation requirements for England, Wales and Scotland [2, 3] feature a pre-completion test, to demonstrate compliance with the required performance standards for sound insulation. A test sample is obtained for the walls and floors that form a separating element between dwellings. Failure to achieve the minimum standard requires additional work to improve the failed surface. Questions are invariably raised when a test is a fail and a common one concerns the reliability of the test. The point raised by Knudsen [1] illustrates succinctly the problem when evaluating sound insulation. Why does it vary so much and what causes the variability? Apart from the implied confusion surrounding the definition of sound insulation, he notes there is doubt about the published sound insulation data. This doubt or "uncertainty" about the data is warranted as the variability is genuine and is called measurement uncertainty. For laboratory sound insulation tests there are British and International Standards that specify ways to quantify the measurement uncertainty and to partition it into that which occurs within the laboratory, or repeatability and that which occurs between laboratories, normally termed reproducibility. Accreditation bodies such as the United Kingdom Accreditation body (UKAS) require accredited measurement laboratories to attempt to; identify all significant components of uncertainty, make a reasonable attempt to estimate the uncertainty and to ensure that reported results do not give a false impression of uncertainty.

The focus of this thesis is to provide a quantitative assessment of the factors that contribute to the uncertainty in the measurement of airborne sound insulation in the field.

1.2 Aims of the research

In the United Kingdom the United Kingdom's Accreditation Service (UKAS) require an uncertainty evaluation for all measurements undertaken by certified laboratories in line with the international quality management standard BS EN ISO17025: 2005 [4]. In this case the specific application of interest is the field measurement of airborne sound insulation and this research focuses on two of the most common forms of floor construction that are regularly constructed and tested in Britain in order to understand the causes of uncertainty due to the construction of the floor and the measurement process itself.

One of the main aims of this research is to obtain quantitative estimates of the components of variance associated with airborne sound insulation testing in the field in order to construct an uncertainty budget that satisfies the requirements of UKAS and the International Standards.

Four areas of influence are investigated in this thesis: the measurement uncertainty provided by the instrumentation or test kit; the contribution of the operator; the contribution of the part being measured and the influence of interaction, if any, between them. These components of variance are determined for each of the floor constructions measured and compared in order to understand how the measurement process and the part being measured contribute to the total variability seen in the field testing of sound insulation.

1.3 Thesis overview

The remainder of this chapter summarises the chapter contents and discusses the literature concerning measurement uncertainty in sound insulation.

Chapter 2 introduces the types of error found in measurement and the basic mathematical theory for calculating measurement uncertainty. It primarily considers two procedures for evaluating the uncertainty in the measurement process, BS5725-2 [5] and GUM [6] and discusses their basic characteristics.

Chapter 3 describes the sources of variability in the measurement process. It constructs a cause and effect diagram that combines the classification of factors from BS5725 [7] with a list of input variables, based on GUM. The a priori requirements for the input variables are considered and the uncertainty information currently available for each.

The benefits of the modelling and empirical approaches are discussed. The GUM is found to be unsuitable because there is correlation between third octave bands which invalidates its use in predicting uncertainty for single figure values. Additional research shows that, without detailed knowledge of the input values, GUM significantly overestimates the total uncertainty. The empirical approach of BS5725 is favoured though improvements to the design of experiment (DOE) are shown to be required.

Chapter 4 examines limitations of using BS5725 for comparative testing in the field. Its inability to prioritise and quantify improvements in the measurement process is discussed. An enhancement of the empirical approach is proposed, which focuses specifically on the use of analysis of variance (ANOVA) and the recent precedent for its use in acoustical research.

Chapter 5 presents a detailed look at ANOVA starting with the simplest general linear model and concluding in advanced ANOVA with specific reference to the Gauge Repeatability and Reproducibility model (GRR) used and developed by the Automobile Industry Action Group (AIAG). The design of experiment (DOE) requirements of the two-way balanced crossed ANOVA design are discussed and the optimum number of operators, parts and replicates are identified.

Chapter 6 deals with the practicalities of the experimental design, the resources available, time limitations and the site specific information involved in the GRR survey.

Chapter 7 presents the data from the GRR and basic descriptive statistical analysis, to verify it is representative and consistent, and to identify any outliers.

Additional checks that the measurements align with predicted values provide a degree of confidence that the data is consistent with data obtained in the field. They also highlight areas where the test environment is imperfect and the data shifts from the theoretical values. External influences, such as background noise, are highlighted. Comparisons are made between the results obtained and that of researchers, to compare the effect of workmanship on a floor's performance. No significant anomalies are observed, which again gives confidence that the field test data provides a reliable, representative sample for further analysis and assessment.

Chapter 8 presents the results from the ANOVA, after each data set is assessed for normality. The focus of the ANOVA is on the variability due to the instrument, operator and the part and any interaction there may be between these factors. A frequency analysis is carried out for the timber and concrete floors. The standard floor construction types only provide a small measurement range so, following GRR design guidance, an additional test sample is introduced, which incorporates values at the limits of the normal measurement range. The "linear" data is combined with the timber and concrete GRR samples to give a single data set to test the measurement system capabilities. The combined sample is assessed to identify the measurement systems aggregate variability across construction types and range.

Chapter 9 takes the repeatability and reproducibility information from the ANOVA and compares it with the repeatability (r) & reproducibility (R) values in the current and the new proposed International Standards [8, 9]. Observations are made with respect to the new definitions of uncertainty and improvements suggested.

Chapter 10 applies a simple cross-check for the test of significance of the factors, termed a Latin Square. It is concluded that a Latin Square test provides a quick test of the significance of the sources of variability in the measurement system, prior to carrying out a full GRR.

Chapter 11 completes the assessment of uncertainty in the measurement system by developing confidence intervals for the point estimates of variance. Two methods are used to calculate the confidence intervals and comparisons are made for both timber and concrete data. Alternative methods of calculating confidence intervals, using specialist software and computer simulation, are also identified.

Chapter 12 looks at some historical data and applies the principles of summation in quadrature to determine the contribution of the part in a concrete floor survey, conducted by Parkin et al in the late 1950's.

Chapter 13 contains the conclusions and gives suggestions for topics for further research.

1.4 Literature review

This section discusses the relevant literature on uncertainty in sound insulation measurement, with particular reference to the historical development of measurement techniques and standards for sound insulation. The survey provides an understanding of the way the minimum sound insulation performance standards were developed and identifies anomalies in measurement and calculation procedures. The measurement and calculation procedure has implications for determining the components of variance and for evaluating the uncertainty in the measurement process.

In addition the review identifies previous key findings, which forms the basis of this thesis work.

1.4.1 Sound insulation standards

The origins of the reasonable standard for sound insulation are concealed within the sociological and scientific research, undertaken by several key individuals in the 1940's and 1950's. Much of the research was intended to "review existing scientific information", "make recommendations for further research" or be of "interest to designers"[10]. Though the early studies were not specifically carried out to develop a standard, they supplied the social information on what was acceptable, setting the agenda for further research through the 1950's and up to the introduction of the first Building Regulation Document in 1965[11].

Early research was carried out under the "Post War Building Studies" initiative, during the war with the co-operation of professional bodies and institutions, set up to advise the Government and specifically the Ministry of Works, with regard to post war construction and rebuilding plans. The report identified inadequate types of construction e.g. lightweight 9 inch walls, also desirable types of construction, which afforded "reasonable" levels of sound insulation. It also gave the first clear recommendation for a measurable level of sound insulation, for airborne and impact sound insulation. The airborne sound insulation index was proposed as a simple uncorrected arithmetic average of sound insulation between the sixteen third octave bands (100Hz – 3200Hz) and the rating limit set at 55dB. For impact sound insulation, the perceived loudness of the sound in the receiving room was preferred. The equal loudness contours of 15 Phons was the limit for impact sound on a bare concrete floor and 20 Phons for a bare timber floor.

A large scale sound insulation survey, on flats only, was carried out in 1952/53 by the Building Research Station, which was reported in Research Paper No27, authored by Gray et al[12] and published in 1958. The survey was extended by Parkin et al [13] and linked to objective measurements of sound insulation across a variety of separating constructions and provided the foundation for the early British grading system. The survey was meticulous and contained a comprehensive list of information: the element construction specification and mass the layout, shape and size of the rooms under test. The study focused on the different type of construction of floors and walls and their average performance rather than the measurement uncertainty in the data collection. However, the published results of the survey included basic statistical information such as standard deviation for the sound insulation performance of the single figure values and frequency data.

A sample of test data for a simple concrete floor is detailed in Table 1-1, with the construction shown in Figure 1-1. In this example eight measurements were carried out on rooms of similar size. Each survey standard deviation is shown in Figure 1-2 and the variance across the 29 floors measured is shown in Figure 1-3.

c/s	A1	A2	A3	A4
100	33	35	34	36
125	33	35	31	36
160	37	35	38	34
200	34	34	37	37
250	39	42	40	35
320	39	42	43	40
400	44	45	45	40
500	43	43	45	43
640	47	47	45	46
800	51	49	47	47
1000	52	50	52	50
1250	57	54	54	56
1600	58	59	59	56
2000	60	57	60	60
2500	63	58	60	61
3200	65	63	63	64
D	**47**	**47**	**47**	**46**
Area	17.2m^2	17.2m^2	14.9m^2	14.9m^2
Volume	44.6m^3	44.6m^3	38.5m^3	38.5m^3

Table 1-1: Sample of Concrete Floor Test Results – after Parkin et al[13]

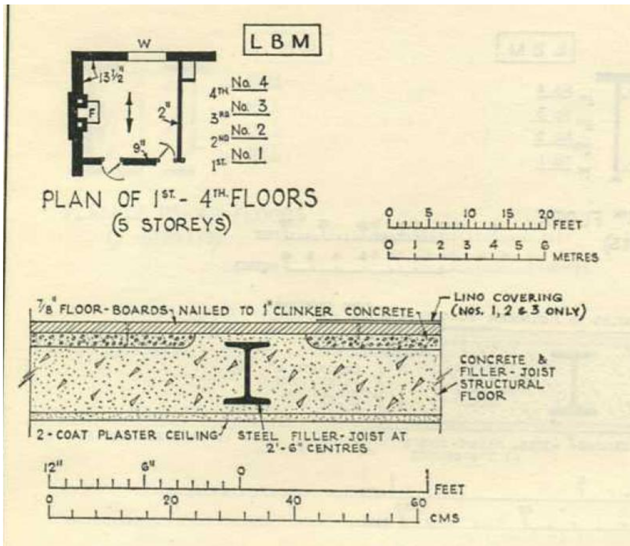


Figure 1-1: Simple Concrete Floor after Parkin et al [13]

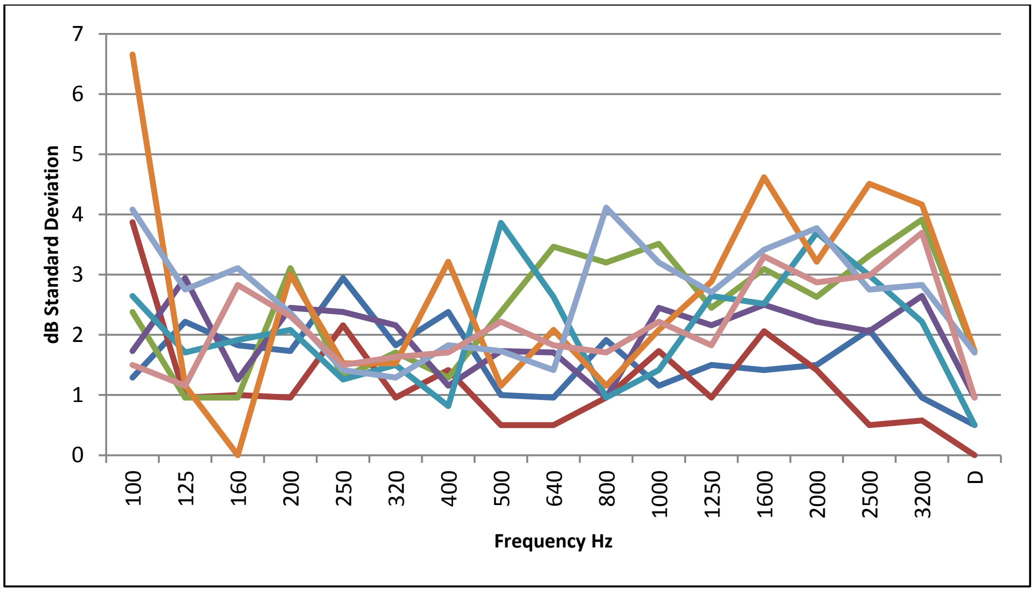


Figure 1-2: Parkin, Concrete Floor Tests (1960) 8 Surveys of concrete floors - 29 floor airborne tests total: standard deviations of survey results [13].

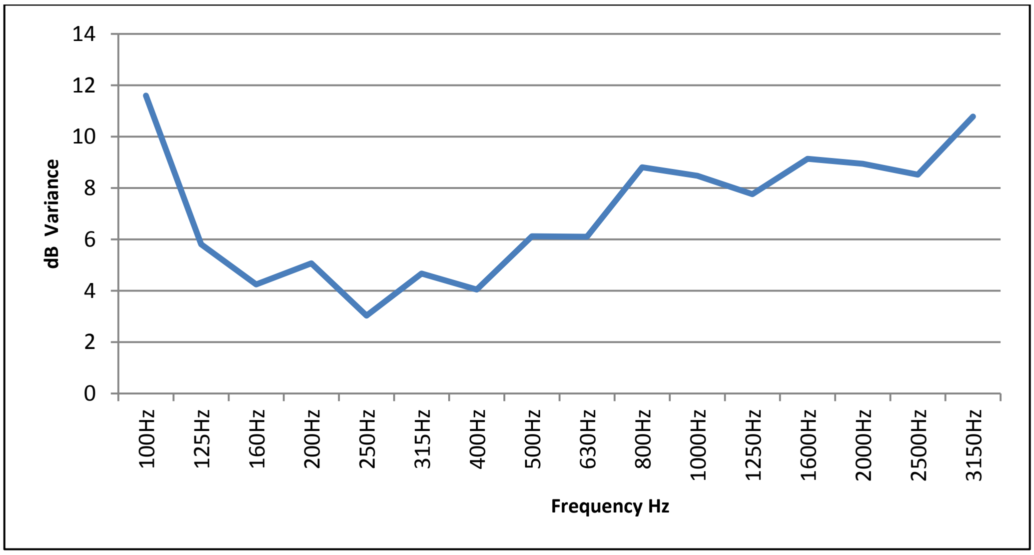


Figure 1-3: Parkin Concrete Floor - combined sample variance of level difference D - 29 floor airborne tests total

The results of this survey, informed later studies, with two digests, Nos 88 & 89 in 1956, by the Building Research Station entitled, "Sound Insulation in Dwellings – I and II, [14],[15]. The digests provided the "House Standard" grade of sound insulation for walls in houses and both of the reference curves for Grade I and Grade II walls and floors in Flats. Documented evidence of how the sound insulation standards for residential dwellings were developed is reported by Parkin et al [16]. It describes in detail the research that culminated in the Digests being released. Several items are worthy of note as they signpost both the basis for the sound insulation standard and the development of the calculation procedure.

The rationalization for the 8dB "deputation" level. Originally this was a classification which identified separating elements likely to lead to complaints. It was based on the experience on some of the housing projects that featured sound insulation 8dB below the Grade II standard. This is a feature which resurfaces in future sound insulation testing reporting (e.g. ISO 717[17], also earlier versions) and British Standard document BS5821, where third octave values are recorded if they fall below the amended standard curve by more than 8dB.

The difference between the airborne sound insulation grading classifications for floors and walls is highlighted in this document as 1dB. This corresponds with the "guideline values", which feature later in the Approved Document E of the Building Regulations (1985)[18] , minimum individual values of 49dB $D_{nT,w}$ for walls and 48dB $D_{nT,w}$ for floors for airborne sound insulation (and mean values of 53dB and 52dB respectively). This "historical" difference for airborne performance was removed in 2003.

In summary, the development work from 1941 – 1960 produced the data, procedures and limits required to enable the standard classification curves for airborne and impact sound insulation to be developed and implemented. It also gave confidence that the curves provided reasonable sound insulation for residential dwellings. Walls and floors, at the lower standard (Grade II) represented an even balance between satisfaction and dissatisfaction.

The original concept in the 1956 digests, was that a separating element would fail if any frequency band fell below the reference curve line. This was too restrictive because a single transgression of this parameter at one third octave band would cause the element to fail the test and measured insulation curves of a wall or floor rarely follow the reference curve spectra exactly. Sound insulation performance, which in the majority of frequency bands achieves or exceeds the curves, should not be penalised by shortfalls in a few frequency bands. A tolerance was proposed, which was detailed in the revised digest No88 published in 1964 [19], where an adverse deviation from the grade curve was allowed. The tolerance for compliance with a particular grade was 23dB over 16 third octave bands. This was incorporated for both airborne and impact sound insulation values and formed the basis for the assessment of sound insulation using the comparison against a curve.

The standard of sound insulation based on the calculation of the aggregate adverse deviation were defined in the first edition of the Building Regulations for England and Wales (1965)[11]. The Building Regulations 1965, which came into operation on 1st February 1966, replaced the local building byelaws in England and Wales with the exception of the GLC area, formerly under the administration of the London County Council (LCC), where LCC Building Byelaws continued in force.

The next significant development came with the 1984 Building Act [4] which consolidated Building Regulations under one piece of legislation. This resulted in the introduction of the Building Regulations 1985 [18]. This revision introduced the concept of Approved Document E which focussed on the “Resistance to the passage of sound” and formalised a move from Grading Curves, using aggregate adverse deviations, to ISO Standard Method of Assessment e.g. the weighted standardised sound level difference ($D_{nT,w}$) using the shifting curve method, which is used today.

The Approved Document was revised in 1992[20] and finally resulted in the current Building Regulation Document Approved Document E (2003)[2] and the introduction of pre-completion testing and the spectrum adaptation term for airborne tests (C_{tr}). This document was reissued in 2004 with the inclusion of the “Robust Details”.

1.4.2 Test Procedures

The sound insulation test procedure in Britain was first formalised and documented in British Standard BS2750: 1956[21]. In this document, the airborne and impact test method was standardised to allow comparison of data between testing organisations, a desire Vern Knudsen had alluded to in the opening quotation in this document almost 30 years previously [1].

The desire to create a standardised environment is detailed in the foreword to the BS2750 document:

“The purpose of these recommendations is to define methods for measurement of sound transmission in buildings and for the expression of results both for field and laboratory measurements in this sphere so that data obtained by different workers can be directly compared.”[21]

This standard did not have international recognition at the time of publication but it formed the basis of definitions of indices D_N, R and L_N and proved to be influential, as these indices appeared in the international standards, which followed afterwards.

The first international standard for measuring sound insulation was ISO/R 140:1960: “Field and laboratory measurements of airborne & impact sound transmission”. The standard detailed the method by which reliable acoustic data could be collected. A complementary standard, for analysing the data collected, was ISO 717/R: 1968: “Rating of sound insulation for dwellings”.

ISO 140 was revised and updated in 1978 to form ISO 140: 1978 Parts 4 & 7 Measurement of sound insulation in buildings and of building Elements”.

BS2750 was updated in 1980 and formally linked to the International standard ISO140: 1978. It was also merged in 1984 with a British document, how to present and analyse data. BS 5821: Parts 1 & 2:1984. "Methods for rating the sound insulation in buildings and of building elements", was identical to ISO 717: Parts 1 & 2- 1982 [and was replaced formally by ISO 717-1 & 2: 1997].

There was a convergence of the International and British Standard documents, during the late 1970’s and early 1980’s. This probably was a recognition of the fact that the discipline of sound insulation was a common practical science, which, even if individual countries differed because of cultural differences, on the level of sound insulation which was “reasonable” in the home, they could standardise the measurement, analysis and reporting procedures.

The international standards are referred to within the Building Regulations for England and Wales, with minor amendments in the averaging procedure.

1.4.2.1 Airborne sound insulation field test

The sound insulation test procedure used in this study is carried out using a UKAS work instruction detailed in the Appendix to the study, see the sound insulation measurement procedure in paragraph 14.3: which follows both the British and International Standard test procedures and, in order to comply with the Building Regulations for England and Wales, Approved Document E (rev 2004) guidance where the requirements differ.

For airborne sound insulation, sound pressure levels are sampled on each side of the separating element which can be a wall or a floor; in our study we are testing separating floors. A sound source is placed in the source room which in order to comply with the British and International Standard [22]and Building Regulations [2] is usually the largest of the pair of rooms chosen. A high level of noise is generated in the source room and an average sound pressure level is sampled across the space using either a rotating boom microphone or by placing the microphone at a number of fixed locations across the room area. In our study fixed microphone positions were used at a minimum of 5 positions across the room and with a minimum averaging time of 6 seconds in each position so a mean and standard deviation of the measurements could be determined. It is important to take care to maintain stated distances between loudspeaker and room boundaries, loudspeaker and microphone, the separation between microphone positions and between the microphone position and room boundaries. In all a 30 second sample is taken. The sound level meter is removed to the receiver room on the other side of the separating test element and the average sound pressure level is measured in the same way. The loudspeaker is moved to a different location in the source room and the process is repeated. A minimum of two loudspeaker positions are required in total to complete the test.

The test procedure can be illustrated pictorially, see Figure 1-4:

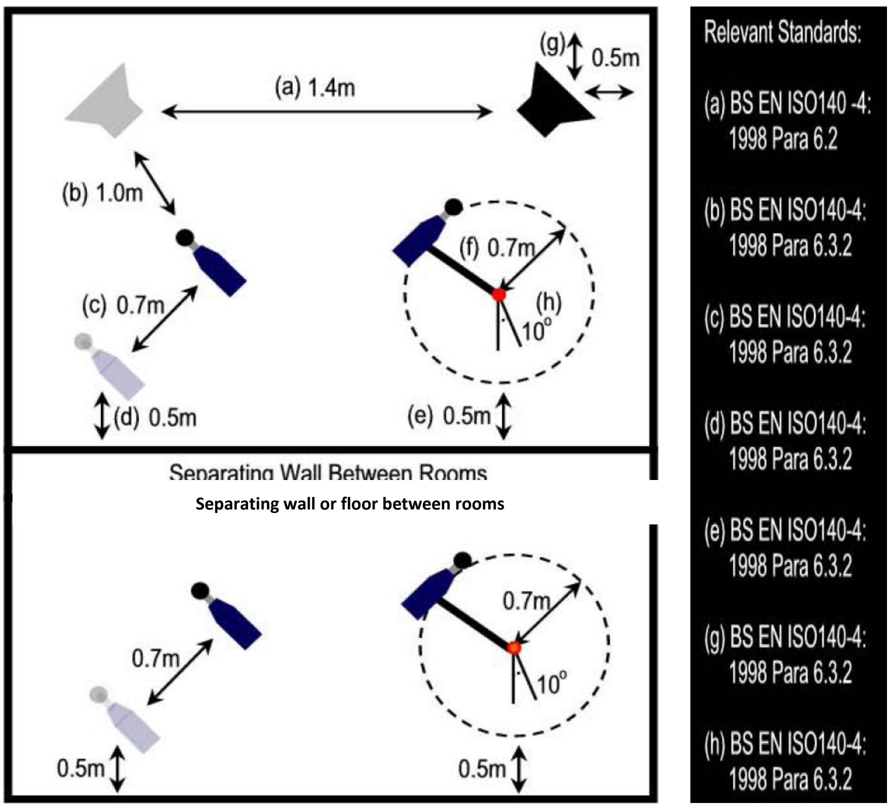


Figure 1-4: Airborne sound insulation test procedure

In addition to the source and receiver room sound pressure levels a background noise level is taken prior to the survey start in order to ensure that the sound pressure levels measured are those from the loudspeaker source and not part of the background noise prevailing on site.

A measurement of the receiver room reverberation time is also carried out to correct for room effects on the sound field. In the UKAS works procedure used the reverberation time is sampled in at least three locations in the receiver room with a minimum of two measurements made at each position giving a minimum total of 6 individual measurements in order to obtain an average level. In our study the interrupted source method is used where a signal is generated by the loudspeaker which is switched off and the decay measured by the sound level meter. The position of the loudspeaker and the microphone in the room and in relation to each other is prescribed in the relevant standards BS EN ISO 140 Part 4: 1998 [22], BS EN ISO 354: 2003 [23]and is illustrated in Figure 1-5:

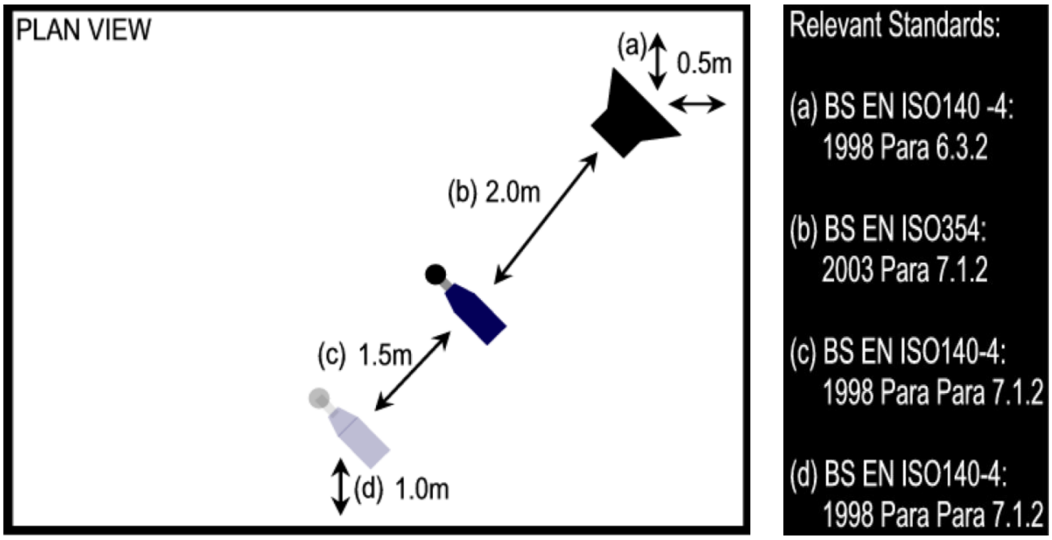


Figure 1-5: Reverberation time measurement – airborne test procedure

1.4.3 Calculation of Measurement Uncertainty

1.4.3.1 BS5725

Most of the relevant literature on measurement uncertainty in sound insulation centres on inter-laboratory studies. Quality management standards [4] require laboratories to assess measurement uncertainty and attempt to: identify all significant components;

make a reasonable estimate of the size of the total uncertainty and its variability; ensure that the reported results do not give a false estimate of uncertainty. The main reason for assessing the measurement uncertainty in laboratories is to ensure that there is no competitive advantage, favouring one laboratory over another. The normal method of assessing measurement uncertainty is by carrying out inter-laboratory studies which follow an empirical method detailed in BS5725 [5, 7, 24-26].

There are several inter-laboratory and round robin studies, which identify the repeatability and reproducibility components [27-35]. They use the BS5725 assessment process to determine the repeatability (r) and reproducibility (R) compare the measurements obtained by each participant laboratory and with the reference values in ISO 140-2 [8]. It is important to note that the ISO r & R reference values are the product of several inter-laboratory studies, where a chosen element was reconstructed or remounted in each laboratory, measured, and the r and R values pooled. These uncertainty reference levels are being updated and in some additional cases redefined in the draft standard ISO/CD 12999-1 [9].

Most of the participants in inter-laboratory studies are national or commercial testing laboratories; each with their own in-house test facilities. The samples selected vary from lightweight partitions, for example see Farina et al [27], to heavyweight walls e.g. Luxemburg et al [35]. The samples are ideally reconstructed from readily available homogeneous materials or transported between and re-mounted in each laboratory.

There are studies on field testing of sound insulation. These focus on existing buildings [31, 36, 37]. The studies are informative, but only comparable if they follow the BS5725 methodology. Closer inspection reveals deviations and inconsistencies in the test procedure, which can lead to discrepancies in the results. An example is the Delta study by Hoffmeyer et al [31], which undertook field measurements of separating walls between a pair of terraced houses. The reproducibility obtained showed good agreement with the reference values in ISO140-2. Further scrutiny shows that, to reduce the uncertainty caused by differences in test equipment, the 5 participating test laboratories in some instances used the same test kit. More importantly, due to time constraint, it was not possible to repeat all measurements. This meant that repeatability is not included in the reproducibility value as required by BS5725. It will therefore underestimate the value of “R” and is probably the reason the reproducibility was lower than the reference values. It is also noted that the reproducibility is calculated for each room measured. The single test specimens are therefore identical, not only of similar construction. The reproducibility therefore does not incorporate the variability due to the reconstruction or remounting of the part. It therefore will underestimate the true reproducibility. Comparison with the reference values in ISO140-2 is erroneous. See Lang and also Hall [36, 38] where this also occurs.

Similar situations, where the test specimen is identical also occur in other research studies [39]. Their impact on the reproducibility may be acknowledged but often it is ignored, either because it is thought to be insignificant but also perhaps because it is not understood. It demonstrates that care is needed when attempting to draw direct comparisons between research on uncertainty.

1.4.3.2 Guide to the expression of uncertainty in measurement (GUM)

An alternative to the empirical method described in BS5725 is described in the guide to the expression of uncertainty in measurement (GUM) [6].

The method is based on modelling the uncertainties by constructing a combined budget which contains all input variables likely to contribute to the uncertainty in the measurement process. This method is considered in detail and leads to the development of a comprehensive list of factors likely to contribute to the total uncertainty in measuring airborne sound insulation in the field. These factors are often referred to as “input variables”.

1.4.3.3 Input variables

Informative research is cited if it describes variability in sound insulation performance of a construction, or if it estimates the variability in any measured component.

Where the research undertaken follows BS5725, the information obtained is limited to the terms defined in the standard. Repeatability is associated only with the instrumentation. What remains, referred to as the “between laboratory” variability, accounts for the rest of the variability in the measurement process.

The variables which contribute to this are many and may be difficult to quantify individually. An example of how the effects of one of these variables relating to the mounting conditions of the test specimen is explored is by Schmitz et al [28]. In addition to calculating the r & R in an inter-laboratory study the mounting conditions of the specimen under test were investigated. The input variable related to the damping effect makes up the total reproducibility and its importance, though measurable, is based on its magnitude and its predictability. In this example Schmitz et al conclude that the influence of the total loss factor may be limited and will likely vary due to the specimen undergoing measurement. Wittstock et al [40] recommends the use of data without correcting for total loss factor. Flanking transmission is also considered by Cocchi et al [41] and Mahn [42], though it is also realised that when accounting for uncertainty in measurement, the variability due to this can be minimised by selecting a common construction across the test sample. This is not considered further in this thesis.

Other input variables can be identified, although only a few have been the focus of research and for some, their contribution to uncertainty is demonstrably small or can be minimised. An example is metrological conditions on site. The influence of temperature on measurement was highlighted by Scholes [43] and together with barometric pressure was the subject of recent research by Wittstock et al [44]. Humidity effects are provided in manufacturers information for the microphone, for example see the B&K handbook [45]. It is noted that the sound insulation value obtained may be affected by metrological conditions but their effects in this study can be minimised by ensuring the measurements are over a short time period, while the conditions are stable.

Others relate to the acoustics of the space and include room effects e.g. spatial variation in sound pressure level and discrete versus continuous sampling in the space; see Schroeder[46], Waterhouse [47-51], Lubman [49, 52-54] and Craik [55]. The uncertainty due to these influences is relatively large though predictable. Predicting the expected variability of sound pressure level is useful in assessing the consistency and reliability of the data obtained on site. It also helps indicate where site test conditions affect the measurement process.

The surface area of the test element and the room volume also contribute to the measured sound insulation. Theoretical formulae are provided in the International Standards ISO 140-4 [22] that can be used to determine the expected difference due to these factors. It is also possible to constrain the variability of these factors by testing similar room sizes. Other examples of how construction on site can affect the sound insulation performance of a test construction are detailed by Sewell [56] who investigated the effects of a step or stagger on the performance of a construction. Though this does not apply to the survey samples in this thesis it indicates the limit of variability that can be expected from this construction feature. As previously stated the construction of the part being measured is also variable and, given a suitable test sample, can be calculated. This variability was referred to as “workmanship” and was calculated for a simple concrete floor by Craik et al [57, 58] the results of which inform this study.

Goydke et al in 2003 [59] assembled a number of input variables using the GUM approach and carried out a Monte-Carlo simulation, to calculate the uncertainty in sound insulation measurement. Wittstock also produced a model using GUM in 2005 [60] to predict measurement uncertainty but concluded that additional work was required to investigate the correlation effects between adjacent third octave bands. This additional work was completed by Wittstock in 2007 [61]. To apply an accurate estimate of uncertainty to any measurement an assumption about the independence of that measurement must be made. In the case of sound insulation measurement, it is known that the adjacent third octave bands are not independent and they have an unknown degree of correlation. It may be possible to estimate an upper limit for the correlation effects by assuming no correlation and apply a simulation process to determine the uncertainty and a correlation of 1 between third octave bands and apply the calculation techniques developed by Wittstock [60,8]. The correlation effects examined by Wittstock raise questions over the usefulness of GUM because correlation effects between third octave bands were found to dominate the measurement uncertainty of the single figure ratings. This evidence, together with its apparent tendency to significantly overestimate the measurement uncertainty shown by Lyn et al [62], suggests that GUM does not provide a suitable framework to assess uncertainty in sound insulation testing in the field.

1.4.3.4 Analysis of variance (ANOVA)

Available research has not identified or addressed the major components that make up reproducibility. BS5725 does not provide a solution for this and recent evidence shows the modelling method used in GUM appears unsuitable.

An alternative approach used in this thesis uses an empirical approach that incorporates analysis of variance (ANOVA) to identify the components of variance. ANOVA has previously been used to good effect in acoustical research, two good examples of which are a laboratory sound insulation study by Taibo et al [63] and a round robin study on the measurement of absorption coefficients by Davern et al [64, 65]. The results demonstrate the strengths of the technique and offer insights into the contributions to variance which allows informed decisions to be made on improvements to the measurement process.

The main advantages of ANOVA are listed by Deldossi et al [66] and include the ability to determine the contribution of the operator and part and operator. Measurement system analysis has been developed as a specialist area of statistics by the Automobile Industries Action Group (AIAG) and is used in industry as a quality control tool. The ANOVA method used is called a Gauge Repeatability and Reproducibility study (GRR) and the appropriate one, for the purpose of this thesis study, is described by Burdick et al [67] as a Balanced Two Factor Crossed random model with interaction. It is this model and additional information provided by Montgomery[68-70] , Borror [71] and Burdick[72, 73] which forms the analytical framework, to separate out and quantify the components of variance in sound insulation measurement and their confidence intervals for timber and concrete floors.

1.5 Summary

This chapter introduced the motivation and outlined the aims of the research. The literature review identifies how standards for reasonable sound insulation and for target values were created. It discusses the changes in the assessment method which led to the development of the shifting curve method of calculating sound insulation which moves away from a simple mathematical calculation process and means the error analysis becomes more complex.

Previous research concerning measurement uncertainty shows that the normal method of assessment provided by GUM does not offer a suitable framework for investigating the components of variability, and other statistical methods are required to partition the reproducibility.

As there are many sources that contribute to the variability in the measurement of airborne sound insulation in the field some of which are difficult to isolate, an empirical approach is proposed. Some of the factors are fundamental to uncertainty and must be included. Others are confounding factors and must be minimised or blocked. The design of the experiment is therefore key, in order to provide a data set consistent with field measurement conditions and to maximise the quality of the statistical information obtained. This allows the variability due to the measurement system and the construction of the part being measured to be isolated and quantified.

Suitable statistical analysis techniques have been identified, which provide methods of calculating both the point estimates of variance and their confidence intervals, for a sound insulation measurement system. The use of ANOVA and in particular the GRR design of experiment is central to the identification of the contributory factors to uncertainty in measurement and their interaction, all of which are addressed in this thesis.

2 Error and Calculation of Measurement Uncertainty

"Statistics is the science of problem solving in the presence of variability." [Mason 2003]

2.1 Introduction

There is no fixed method or procedure for calculating the component parts of uncertainty in sound insulation measurement. This chapter considers the existing practical guidance, for calculating and interpreting measurement uncertainty, both quantitatively and qualitatively. It draws on present methods of metrology for the physical sciences, identifies the basic methods of evaluation, it then provides an overview of the two assessment frameworks and highlights the main differences and potential benefits, with respect to this research.

The basic concepts in measurement uncertainty are described, with examples where appropriate, for clarity.

2.2 Basic concepts

To understand uncertainty, it is essential to understand the key components that determine the nature of the errors involved in the measurement process and their influence on the overall accuracy and precision of measurement. The standard method of calculation of uncertainty is then described, within the context of the most commonly used standards.

2.3 Accuracy and precision

The concept of accuracy and precision can be represented graphically; see Figure 2-1 The target value is the "true" value of the quantity being measured.

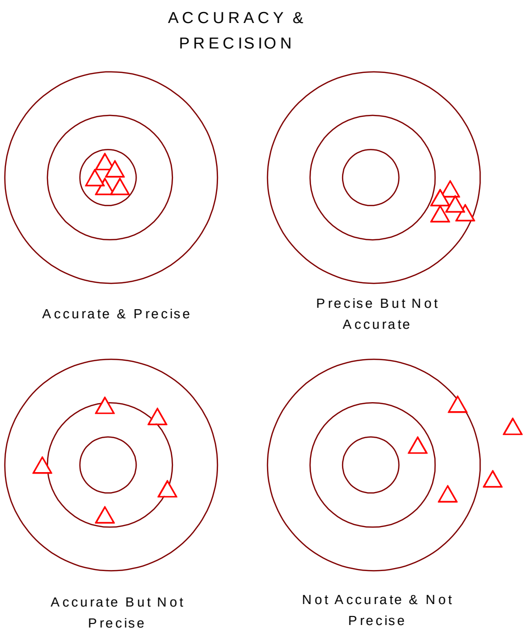


Figure 2-1: Schematic of types of error

An accurate measurement is one where the results of an experiment agree closely with the true or accepted value. A precise measurement is one where the distribution of the results is small. The target diagrams show accuracy and precision in the various combinations e.g. that it is possible to be precise but not accurate. Both accuracy and precision can be influenced by errors in the measurement process.

2.4 Types of error

The two main types of error in a measurement process are random and systematic. Each is considered in more detail below. A third, human error, is also considered. Their influence on data sets is summarised as

- Random Errors – influence precision
- Systematic errors – influence accuracy
- Human errors – generate outliers.

NB: Outliers can be defined and identified for removal by statistical tests such as Grubb's Test, or Cochran's test introduced in paragraph 7.3.2.2 of BS5725 [5], amongst others. If they are attributed to a known fault or error in the measurement process, they can be removed by visual inspection.

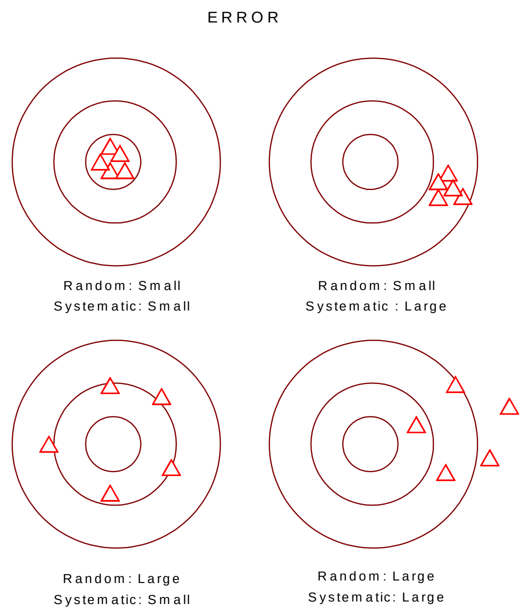


Figure 2-2: Error represented as random and systematic

In most field measurements of sound insulation we do not know the true value and can only comment that the random errors are large or small. Without the target it is not possible to comment on the systematic errors as there is no reference or "true value".

2.4.1 Random errors

In studies of uncertainty, the experimental methods employed generally focus on repeated measurements which reveal the *random* errors. These represent the natural variability expected from the experimental process, and are a requirement of any statistical approach to the assessment of uncertainty.

2.4.2 Systematic errors

By definition, this type of error is not random and adds a bias to the results, usually in a consistent manner. Systematic errors are sometimes difficult to detect. It is important therefore, to try to anticipate possible sources of systematic error in any experimental process, even if it is beyond the control of the experimentalist. In sound insulation testing, it can be attributed to a poorly designed measurement approach such as incorrect calibration of the instrumentation. However, it will be demonstrated that it occurs in the field, even if all the correct procedures are followed.

Systematic errors can be indigenous to the measurement instrumentation, which are bounded by physical limitations of performance or operational specifications. They can also be inherent in the measurement method itself or they can be a feature of the data analysis techniques required by the international standards.

2.4.3 Human errors

This type of error is one most commonly associated with outliers in the data set, the other being equipment failure. It is almost always difficult to detect post experiment, though if noted during a measurement procedure, can be more easily attributable. The measurement uncertainty calculation standards do not explicitly deal with human error, though some provide a systematic application of statistical tests for the detection of outliers [5].

2.5 Calculating uncertainty

There are several methods and relevant reference texts which highlight the different methods of calculating and reporting measurement uncertainty in scientific measurement [45, 74, 75]. The simplest explain how measurement uncertainty can be calculated using basic information such as a mathematical formulae or relationships. Others require a significant amount of knowledge about the specific components of uncertainty likely to contribute to the overall uncertainty e.g. GUM[6].

2.5.1 Uncertainty theory

2.5.1.1 Uncertainty in Single variable functions

For single variable functions, the usual way to estimate the uncertainty is to refer to tables for the common mathematical functions. For example:

Function Z (A)	$\frac{dZ}{dA}$	Uncertainty
10^A	$10^A \ln(10)$	$\alpha_Z = 10^A \ln(10)\ \alpha_A$

Table 2-1: Example from Hughes & Hase [45] p39 section 4.1.3

Worked example: $Z = 10^A$

A has been measured; $A = 2.3 \pm 0.1$

What is the value of Z and its error?

Best estimate of Z is the mean $Z_{mean} = 10^{2.3} = 199.5$

The error is calculated using the functional approach:

$\alpha_z + = 10^{2.3+0.1} - 10^{2.3} = 51.7$

$\alpha_z - = 10^{2.3-0.1} - 10^{2.3} = 41.0$

The best estimate of Z lies within the range: $158 \leq Z_{mean} \leq 251$

Using the calculus approach above $\alpha_z = Z \ln(10)\, \alpha_A = 199.5 \times \ln(10) \times 0.1 = 45.9$

The calculus approach assumes symmetry and the approximation to the error is

$Z = (2.0 \pm 0.5) \times 10^2$

It is rare that single variables are the only source of uncertainty and techniques are required to determine the combined uncertainty. Most physical quantities cannot be directly measured in one step; they are more commonly determined by two or more steps, see section 7.1 p97 Kirkup et al [75] involving multiple variables.

2.5.1.2 Uncertainty in multi-variable functions

The classical case of error analysis for physical quantities considers the combination of errors, leading to the propagation of uncertainty.

When measured quantities *x* and *y* are added or subtracted the uncertainties add, when measured quantities *x* and *y* are multiplied or divided, the fractional quantities add.

From Taylor [74] if the measurements of *x* and *y* are independent and random their uncertainties are said to be *added in quadrature*, that is they are squared, summed and the square root taken[74]:

$$\delta q = \sqrt{(\delta x)^2 + (\delta y)^2} \qquad \textit{2-1}$$

The calculation relies on the partial uncertainties being determined for each of the variables. This holds for a measurement process that involves several steps or different types of measurement. It also forms the basis for calculating uncertainty in the international standards. If it is suspected the uncertainties are not independent the recognised convention would be to use the ordinary sum. In any event the ordinary sum of independent uncertainties will form the bound for the upper limit of uncertainty: From Taylor again:

$$\delta q \ \leq \ \left|\frac{\partial q}{\partial x}\right| \partial x + \cdots + \left|\frac{\partial q}{\partial z}\right| \partial z. \qquad \textit{2-2}$$

The process of squaring exaggerates the importance of larger values. If a component of uncertainty is 5 times another, e.g. 5% as opposed to 1% of fractional uncertainty, by adding in quadrature it becomes 25 times that of the other and so dominant that one can generally neglect the influence from the minor component. This can prove helpful in prioritising uncertainties and in calculating the importance of input values. The challenge therefore, is obtaining a measure of the fractional uncertainties thought to be influential, a priori. For example, what would be the contribution to the measurement uncertainty, if barometric pressure is 1038mB, for a sound pressure level measurement rather than 1040mB? What contribution does the measurement position in the room make?

2.5.2 GUM: 2008

The most commonly adopted method of expressing uncertainty in metrology is given in the Guide to the Expression of Uncertainty in Measurement [6](GUM) : current edition revised in 2008 published by the Joint Committee for the Guides in Metrology Working Group 1 (JCGM/WG1). The GUM is internationally recognised and has an equivalent in the UK which has been refined into an information document. This document is from the United Kingdoms' Accreditation Service (UKAS) and is referred to as M3003 [76]. It follows the GUM principles when creating an uncertainty budget for a measurement system. Both documents form the cornerstone of error analysis in metrology and provide the essential reference documentation to measurement laboratories. Each is the corollary of the other and, therefore reference is made only to the GUM in this thesis.

"The 'Guide to the Expression of Uncertainty in Measurement' (GUM) provides general rules for evaluating and expressing uncertainty in measurement that are intended to be applicable to a wide range of measurements and for use within standardization, calibration, laboratory accreditation and measurement services" [77]

The GUM is based on a mathematical model where the functional relationship used to define the uncertainty is defined as:

$$Y = f(X_1, X_{2,\ldots\ldots} X_N) \qquad \textit{2-3}$$

Where Y = The measurand (not measured directly), is determined by N directly measured quantities X_1, X_2,......, X_N

NB: It is noted that the GUM uses lower case letters for estimates of values and upper case for true or actual values.

The quantities X_1, X_2 are usually determined as part of a mathematical formula or relationship, which is defined, but this need not always be the case in the application of the GUM. The functional relationship has no limit on the number of input quantities which contribute to the total uncertainty. This is a significant limitation to the practical use of the GUM to determine an uncertainty budget. It is not always apparent that all the quantities up to X_N have been taken into account, or if some have been included which should not be. Any gaps in knowledge of the components or size of their contribution will affect the accurate implementation of this model.

In essence the GUM is based on collecting a combined sum of the expected components of uncertainty in measurement. This is called the "*Combined standard uncertainty*" and is defined in the GUM in paragraph 2.3.4 as

"standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of

terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities".

The components of uncertainty are classified "Type A" or "Type B". Type A uncertainties are estimated by repeated measurements of a parameter and by considering individual measurement differences from the mean. Type B uncertainties are obtained for example in a calibration certificate or other external reference. It is sometimes known as the "bought in uncertainty" and can be added directly into the uncertainty budget. In the case of the calibration certificate, the measurements made by the calibration laboratory to determine the uncertainly of a measurement instrument would be "Type A" uncertainties to the laboratory but would become "Type B" uncertainties to the user of the instrument in creating the uncertainty budget for the measurement process.

Once the combined standard uncertainty of measurement is determined; from GUM:

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial_{\mathrm{f}}}{\partial_{\mathrm{xi}}}\right)^2 u^2(\mathrm{x_i}) \tag{2-4}$$

The normal practice is to calculate the expanded uncertainty (by defining an appropriate interval using a coverage factor) thereby expressing the uncertainty as a confidence interval within which the true value lies. This is defined in section 0.1 of GUM:

"When reporting the result of a measurement of a physical quantity, it is obligatory that some quantitative indication of the quality of the result be given so that those who use it can assess its reliability. Without such an indication, measurement results cannot be compared, either among themselves or with reference values given in a specification or standard."

The GUM is the latest global method for calculating measurement uncertainty but there have been other methods which featured a more empirical approach e.g. BS5725.

2.5.3 BS5725 1991 – Accuracy (trueness and precision) of measurement methods and results

The current version of this British Standard predates the GUM by 14-17 years depending on which part of the standard is referred to. The approach is based on a random effects model. From BS5725-1 [5]:

$$Y = m + B + e \qquad \text{2-5}$$

where:

m = general mean (expectation)

B = laboratory component of bias under repeatability conditions;

e = random error

Equation 2-5 focuses on a balanced uniform-level experiment [5] also known as a round robin. This is a standard that lends itself to a "Design of Experiment", abbreviated to DOE [5] approach to statistical estimations of uncertainty.

The approach to calculating the uncertainly, comprised of "repeatability" and "reproducibility", is based on the assessment of a directly measured value. This is the variability of the final sum, which is directly determined, rather than from the sum of its component parts. In the context of this thesis study, the measured value is the single figure value of airborne sound insulation e.g. $D_{nT,w}$ or the third octave band descriptor D_{nT}.

Repeatability and reproducibility are calculated using analysis of variance (ANOVA) techniques. The repeatability (r) used in the standard is defined (in paragraph 3.3.5) as:

observation conditions where independent test/measurement results are obtained with the same method on identical test/measurement items in the same test or measuring facility by the same operator using the same equipment within short intervals of time. [78]

The reproducibility (R) used in the standard is defined (in paragraph 3.3.11) as:

Observation conditions where independent test/measurement results are obtained with the same method on identical test/measurement items in different test or measurement facilities with different operators using different equipment.[78]

The process described in BS5725-2: 1994 is based on a statistical approach to uncertainty with the "trueness" referring to closeness of the arithmetic mean to the true value. The "precision" referring to the closeness of the individual results. The earlier "Target Diagram" examples give an indication of the characteristics of trueness and precision in this circumstance, i.e. Precision can be calculated without knowing the true result as it only requires a comparison between results whilst trueness requires a reference value in order to offer meaningful information about the measurement process.

BS5725 is perhaps the prominent Standard for cross comparison of measurement systems and inter-laboratory studies. It contrasts with the approach recommended by the GUM in that its format offers flexibility in how the analysis can be applied to fit common practical situations. Examples of how the method can be expanded are given in BS5725-6: 1994 [26] e.g. determining repeatability and reproducibility limits to be used in examining the test results obtained by a standard measurement method or describing how to assess whether a laboratory is able to use a standard measurement method in a satisfactory way.

2.5.3.1 Comparison of BS5725 and GUM

The two main procedures for identifying the uncertainty in a measurement system are based on significantly differing approaches to the quantification of uncertainty. One is reliant on modelling the uncertainty (GUM) the other is based on an empirical method (BS5725).

The differences are detailed in a review by Deldossi et al and are suitably summarised in Table 1 from that source. See Table 2-2.

Table 2-2: comparison of GUM & BS5725 - after Deldossi et al[66].

Scope	BS 5725 Accuracy: trueness and precision	GUM Measurement uncertainties
Measurand	Y is measurable	Y may be not directly measured
Data obtained by	Statistical experimental design	NO statistical experimental design
Model	$Y = \mu + B + E$	$Y = f(X_1, X_2,\ldots,X_m)$
Variability of the measurand	Repeatibility and reproducibility variance	Combined standard uncertainties $u_c^2(y)$
Confidence interval	NO: only the definition of repeatibility and reproducibility limit	NO: only the definition of an interval, $y \pm U_p$, that could reasonably be attributed to the measurand.

It is assumed that the measurand is directly measured in BS5725, which contrasts with the GUM that allows for situations where it is not directly measurable. This is important as sound insulation is not directly measurable. However, BS5725 can include single values, which are the outcome of a calculation from a set of observations. [See section 1.2 from BS5725-2: 1994]

The sound insulation values are the outcomes of sets of observations. The only issue in this instance is that although the sound insulation measurement is on a continuous scale, because of the way they are calculated, $D_{nT,w}$ and $D_{nT,w} + C_{tr}$ are reported in 1dB steps. What this could be described as, is a discrete value based function. It could be argued instead that the values reported are just rounded down to zero decimal places and therefore fall within the requirement. D and D_{nT} values on the other hand are measured on a continuous scale and are rounded to 1 decimal place and comply with the whole statement.

Uncertainties in BS5725 are computed statistically using ANOVA techniques, based on a finite dataset. In GUM they are determined using quadrature summation in of all the Type A and Type B input values. The main point is that the summation of the individual input values may not contain all the uncertainties, leading to approximations being made by experienced practitioners.

Possibly the most important consideration is that in BS5725, it is possible to extend the ANOVA experiment to investigate interactions between factors and extend the design of experiment to incorporate other factors. The GUM has no simple way of achieving this, given individual input value uncertainties, and does not explain how to obtain this information. In fact the statistical term “interaction” does not appear in the GUM document. Zero interaction between input values is a broad and possibly erroneous assumption, unless there is certainty that there is independence between factors. If the present research is to be truly informative about identifying the components of variance and identifying their relationships and characteristics, then the inability to calculate and identify interactions between factors is a significant challenge.

2.6 Conclusion

This chapter considered current commonly used methods for calculating uncertainty inherent in a measurement system, and compared the two approaches with respect to the current research requirements: the GUM and BS5725.

If the sources of uncertainty can be identified, the contribution from these input variables can be quantified and a combined quadrature sum results. The GUM approach is useful if the measurand is not measurable or the input variables have bought in uncertainty only (Type B). The combined uncertainty must include all component parts to ensure the uncertainty in the measurement process is not underestimated. In addition it is attractive if the alternative empirical approach is cost prohibitive due to the sample being measured, geographical location of the laboratories or time involved in the process.

The empirical approach combines a specifically designed experiment with a statistical assessment method (ANOVA) to reveal the individual repeatability and reproducibility of the measurement system. It is the more common approach as these defined statistical quantities (r & R) are regularly referred to in other sound insulation standards e.g. BS EN ISO 140-2 [8].

BS5725 gives additional information to GUM, in that the design of the experiment allows one user (laboratory) to comment on what another user (laboratory) expects from the same measurement method on similar parts. It also offers a deeper understanding of the inter-relationship between the major factors.

In its present form and in the context of this research study, BS5725 does not offer sufficient information on the potential components of variance likely to be encountered in field measurement of sound insulation. In order to decide on the approach required the potential sources of variability in field measurement of sound insulation are described in the next chapter.

3 Sources of variability

3.1 Introduction

In Chapter 2, two approaches for the calculation of measurement uncertainty were considered. They offer different approaches to the identification and quantification of uncertainty.

Any study into measurement uncertainty needs to consider a suitable method of identifying the components that contribute to the uncertainty and the size of their contribution.

The identification of individual components of variance can be complicated and the estimation of their specific contribution problematical. Having demonstrated in Chapter 2 how uncertainty can be calculated from simple mathematical relationships and physical measurement studies, the aim of this chapter is to take the two methodologies and identify the major components of variance in sound insulation testing and develop a way to quantify their contribution.

Three key aspects of uncertainty that are developed in this chapter are: consolidating the general headings of uncertainty used in BS5725; identifying the individual GUM input values, which could be listed under those headings; and pinpointing where information is missing or incomplete.

A way forward is proposed, of extending BS5725 ANOVA techniques, within a modified DOE framework, which focuses on the measurements system's contribution to the variability in collected data and which isolates the sources of variability in the measurement system.

3.2 Sources of variability – field measurement of sound insulation

BS5725 and GUM provide separate frameworks for determining measurement uncertainty. The former can be referred to as a "top down" the latter as a "bottom up" approach, Ellison et al [79] propose alternative terms: "holistic" and "deconstructive", respectively.

Both methods should yield the same result. To demonstrate the compatibility between the methods, the factors, as defined in paragraph 0.3 of BS5725-1 [7] will form the section headings in the next section. They will be mapped to the ”Imported input values” as defined in paragraph F.2.3 of GUM [80].

3.3 BS5725 framework factors

The “top down” approach of BS5725 uses an empirical DOE and statistical techniques to separate the factors contributing to variability. The general classification of factors which are likely to affect the trueness and precision of a measurement method e.g. the measurement of sound insulation in the field, are listed according to BS5725-1: 1994: operator; equipment used; calibration of equipment; environment; time elapsed between measurements.

A notable omission from this list is the part or specimen to be measured. The standard disregards the variation between the test specimens because in the design of experiment (DOE), the specimens are assumed to be identical, see paragraph 4.3 of BS5725-1 [7]. This allows only the variability associated with the measurement method and between laboratories, to be considered. It is known that apparently identical separating elements, in this case floors, can perform differently and the differences are non-trivial. It has been shown by Craik [57] that there would be variability even for a nominally identical floor and test scenario. Therefore the quality control of the construction process or “workmanship”, as termed by Craik, should be included. Identification of part variability is core to this research and the ability of the measurement system to distinguish between parts is one of the parameters by which it should be assessed.

The DOE method, proposed in BS5725 requires modification if it is to include the part as a component of variability. The modified method would then be able to consider non-similar performing specimens and allow assessment of the measurement systems capability across the full range of practical measurement conditions. In a similar way the GUM would have to take this part variability into account, in order to ensure that the measurement process was able to be assessed independent of test specimen.

The second omission from the list is that BS5725-1 does not consider the measurement method as a potential source of variability. Section 4.1 of BS5725-1 describes the use of a “standard measurement method”.

This highlights the importance of a coherent practical method to minimise the potential for variability between measurement bodies. It also corresponds closely with the UKAS Laboratory requirement for a working method, for measuring sound insulation in the field. From paragraph 4.2.2 of BS5725-1 [7]:

“Pronounced differences in the within - laboratory variance or between the laboratory means may indicate that the standard measurement method is not yet sufficiently detailed and can possibly be improved.”

It is therefore important that participants understand, but more importantly use the agreed measurement procedures consistently. If the measurement method differences can be minimised, it will highlight the contribution from the other factors associated with the measurement system. The variability of the method will largely be attributable to the operator, based on their interpretation of the instructions and their ability to carry them out consistently. A standard method is required to minimise the variability associated with this factor, but this does not mean that the contribution will always be small.

It is desirable if the time elapsed between measurements could be minimised, thereby negating the requirement for this as a factor. This is in line with the recommendation of the standard, see paragraph 4.4.1 of BS5725-1 [7].

In light of the comments above and if the equipment used and the calibration of the equipment could be combined into one term, called “The instrument” that would leave us with the following DOE framework headings for consideration.

a) The operator
b) The instrument
c) The environment
d) The part being measured

3.4 Development of GUM

The simple DOE Model can be used as a basis on which to overlay the GUM approach. More detail can be added, for the specific case(s) investigated. This additional detail is included in the model by introducing the GUM input variables likely to contribute to the combined uncertainty under the respective headings. These general uncertainty headings and the sub-set of associated input variables can be illustrated graphically in Figure 3-1 based on a simple cause & effect diagram. See Fig 1 of Ellison et al [79]:

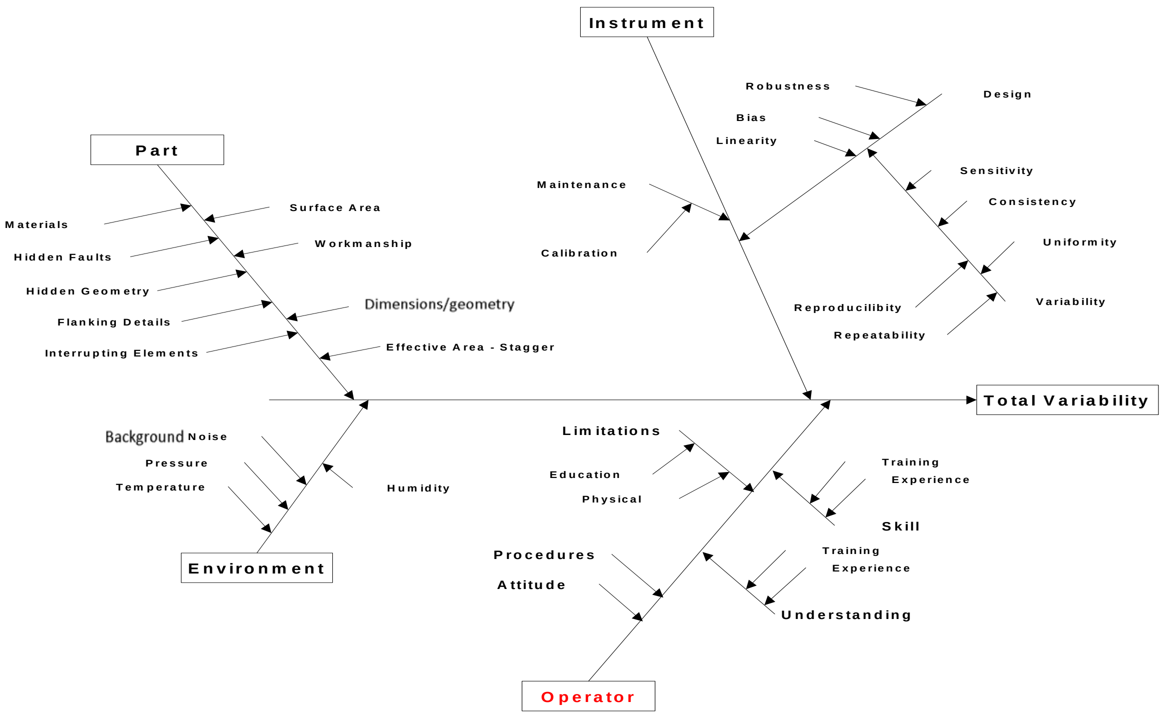


Figure 3-1: Cause & effect diagram for sound insulation testing in the field – Uncertainty headings & individual input variables

If the GUM framework is to be used the quantification of each of these input variables is required.

3.4.1 Input variables

Laboratories do not have sufficient time or resources to investigate measurement uncertainty. It would be impractical to take each input variable listed above and attempt to quantify the individual contribution to the uncertainty budget. This would, in effect be attempting to create a purely “Type A” evaluation of standard uncertainty.

The GUM offers practical guidance in compiling a budget using external data sources or “Type B” evaluations to compliment the available “Type A” statistical uncertainty data. These are called “imported input values”. This allows the uncertainty associated with an input quantity to be obtained from elsewhere, perhaps from an independent study or research paper. The imported values may differ in the way the statement of uncertainty is described or presented. For example these can be given as standard deviations or variances supplied with or without confidence limits. Judgement should be applied as to the reliability of the uncertainties obtained.

It would be up to the user of such information to determine how it may be best applied and this may involve an informed “best guess” or approximation of the uncertainty contribution to the budget. Additional GUM assumptions may also prove restrictive:

3.4.1.1 Independence

The GUM assumes independence between input quantities and does not provide a procedure to determine correlations if they exist. However, it does give a method for adding the uncertainties if the correlations between input quantities are known see paragraph 5.2.2 GUM [6]. Sound insulation testing may not have complete independence between input variables. Wittstock [61] has found that correlation effects are present between third octave bands, which means this assumption may not apply.

3.4.1.2 Normality

GUM assumes normality, however measured sound insulation data does not always follow the Gaussian shape. The Probability Density Function (PDF) of a sound reduction index is assumed by Mahn [81] to be Log-Normal and the PDF of the mean square sound pressure in a reverberant room has been described by Waterhouse [82] and Bodlund [83] as having a gamma distribution. It may be that transformations of these distributions may be Gaussian but this may not always be as straightforward or practical if corrections are applied e.g. correction for background sound pressure level.

Modelling non-normal input variables is recognised as a suitable alternative and the Monte Carlo method, also described in GUM, can be used to assess the uncertainty of the input value [84]. However, as there are so many input variables in the measurement system, which may diverge from normality and independence, the modelling of their uncertainties may not be practical or offer an accurate approximation. Both Lyn et al [62] and Deldossi et al [66] have highlighted that identification of all such sources can be problematic.

3.4.1.3 Other factors

It has been suggested by Mahn [85] that when the GUM reported uncertainty is large, the normal rule of propagation of uncertainties appears to break down. This is of concern as it is expected that the standard deviation of results at low frequency will naturally be large due to low modal density in standard sized residential rooms.

Work by Bessac et al [86] used GUM for estimating measurement uncertainty in sound power measurement in a reverberant room. They found that the uncertainty was specific to the laboratory and the source being measured. They showed that the "device", or in our case the "part", offers uncertainty and also the test method used. They conclude that the method was best used to calculate the hierarchy of uncertainty, enabling identification of where improvements can be made leading to lower uncertainties. The GUM therefore could be said to be a tool to offer a micro view of uncertainty specific to the individual laboratory and the source being measured rather than a macro view for all laboratories and sources.

3.4.1.4 Incomplete information

There are likely to be situations where there is little or no information available about certain input variables. Comparative work on empirical v modelling methods of estimating measurement uncertainty has highlighted issues, reliably identifying all the measurement input variables in primary sampling. Even if all the input components are identified there are further issues if the variability quantum is unknown. The GUM modelling method has been criticised by Lyn et al [62] who have identified estimates of uncertainty that are 6 times larger (136% v 22.5%) when a GUM modelling approach has been compared to empirical results.

Incomplete information is considered in more detail in the next section. It is important, as it informs the assessment method selected and the design of the experiment selected.

3.5 Quantitative Assessment of Uncertainty

3.5.1 Introduction

The main DOE headings are: instrument; operator; part, environment. Each of these headings is considered with the associated input variables.

3.5.2 Instrument

The term instrument covers the sound insulation test kit used in the measurement of airborne sound insulation in the field. It comprises: sound level meter, loudspeaker, wireless transmitter and receiver and the microphone calibrator. The input variables likely to contribute to the instrument uncertainty are listed in Table 3-1 , together with information on the associated uncertainty quantum or likely source.

3.5.2.1 Informative references

Table 3-1: Input variables - Instrument variability

Heading: Instrument Input Variable	Reference Source	Possible Uncertainty Information	Direct Uncertainty Component: σ Empirical/Statistical
Sound Level Meter	Calibration Certificate	Empirical (Lab only)	Yes: third octave Band
Sound Level Meter Calibrator	Calibration Certificate	Empirical (Lab only)	Yes: third octave Band
Loudspeaker	None	No	No
Wireless transmitter & receiver	None	No	No

Under the GUM the measurement instrumentation would be included under the Type "B" bought in uncertainty, through the calibration certification. This calibration certificate will, for the purposes of this research, include: frequency response of the microphone (free field microphone); signal processing part of the electronics (Analoguedigital conversion); filters, response (time) & bandwidth (Assumed as constant percentage, third octave band).

The Type B bought in uncertainty would, under normal circumstances, only include the sound level meter and the associated calibrator. Regular calibration of these two items for a UKAS Accredited Acoustic laboratory is a standard requirement and the sound level meter is calibrated before and after field tests. In the case of a laboratory which is UKAS qualified for testing sound insulation in the field, the sound pressure level meter has the sound pressure level (and filters) and the reverberation time calibrated and certificates provided by an external calibration laboratory. Therefore the bought in uncertainty Type B is relatively straightforward, though relatively expensive to acquire.

However, it is not the case nor is it expected that any other part of the measurement system e.g. the wireless kit or loudspeaker would have any certification verifying the measurement uncertainty of the equipment. If this information was thought to be important, more work would be required to determine its contribution as an input variable e.g. statistical analysis to provide Type “A” information. This input variable information is missing from the measurement systems contribution to the total sum. In a BS5725 assessment the uncertainty of the instrumentation is described by the “Repeatability” “r”. In both respects it is relatively simple to incorporate into an uncertainty calculation.

3.5.3 The Operator

When considering the GUM approach the “operator” component represents the surveyor or acoustic engineer responsible for the test, and comprises multiple input variables, likely to have an impact on the combined uncertainty budget. Some of these apply to the ability or competence of the operator. It is difficult to see how a parametric study could isolate the uncertainty attributable to training, understanding, skill and attitude, though they must all in some way play a part. The input variables likely to contribute to the operator’s uncertainty are listed in Table 3-2:

3.5.3.1 Informative references

Table 3-2: Input variables – Operator variability

Heading: Operator Input Variable	Reference Source	Possible Uncertainty Information	Direct Uncertainty Component: σ Empirical/Statistical
Test Method	None	No	No
Procedures	None	No	No
Attitude	None	No	No
Understanding	None	No	No
Skill	None	No	No
Limitations	None	No	No
Physical Limitations	Hopkins – Moving Microphone – arm length: number of discrete measurements	Yes	No
Test Method	Waterman, Lubman, Chien	Yes	No

Waterhouse & Lubman [49] and Lubman et al [54] highlight the importance of statistical independence of the sample measurement positions. The two sampling routines allowed by the measurement standard BS EN ISO 140-4 [22] are: static microphone and; moving microphone.

In a series of papers by Waterhouse et al [49] , Lubman et al [54] and Chien et al [87] it is shown that sampling by discrete microphone positions, at specified minimum distances apart, relative to wavelength, minimises correlation between positions and shows a discrete average can always be obtained that is better than from continuously recording and averaging along a straight line in a room, a line tracing the path of a rotating boom microphone, or at random positions across the surface of a circle of similar radius. Given the radius "*r*" of the circular path, it is possible to calculate the number of equivalent uncorrelated samples to which it relates.

Hopkins[88] proposes a hand-held moving microphone technique, to give the best sampling paths, and calculates the maximum number of equivalent discrete positions given the anthropometric limitations of the human arm and body. Hopkins discusses the difficulty in estimating the effect of self noise, caused by the operator moving the microphone i.e. walking across a room, clothes rustling, knee, ankle, elbow and shoulder joints cracking. Another potential problem with the moving microphone method, is that there is no way to calculate the standard deviation of results across the space, in either the source or receiver room, and so the variability of the measurement is inferred by the number of independent positions rather than calculated as a standard deviation. In this research, the moving microphone method was avoided and a discrete sampling regime adopted.

There is no other identifiable research that highlights operator input variables. With no other Type “A” or “B” uncertainty quantification available there is a significant gap in the GUM combined uncertainty budget for the operator. Conversely, using the DOE approach, the operator (or laboratory in BS5725) can be added in the identification of input variables likely to contribute to the uncertainty budget. The statistical framework of BS5725 and in particular ANOVA, lends itself to assessing the contribution from operators as a part of the total “Reproducibility” component.

3.5.4 The part

The “part” in this context is the specimen to be measured, in this study it is a floor and its associated flanking construction as well as the rooms in which the measurements take place. A list of input variables associated with the part are listed in Table 3-3 .This includes room elements which are not always considered, boxed-in elements and services, window reveals, inbuilt wardrobes and cupboards. Ideally, if the part is to be controlled, the volume of the room, its shape and size should also be controlled, in order that any variability disclosed by the measurement procedure and interactions with other factors can more easily be identified and apportioned.

3.5.4.1 Informative references

Table 3-3: Input Variables the Part.

The Part: Input Variable	Reference Source	Possible Uncertainty Information	Direct Uncertainty Component: σ Empirical/Statistical
Materials	None	None	No
Spatial variation: Sound Pressure in Room	Craik et al [Lubman[52], Schroeder[46]]	*3-1: From Craik [55]* $$s(dB) = \left[\frac{\sum L_i^2 - (\sum L_i)^2/n}{n-1}\right]^{1/2}$$ *3-2: From Craik[55]* $$\frac{s}{m} = \sqrt{\left[\sum \frac{10^{(L_i - L_t)/5}}{(1 + 0.145\Delta f_i T_i)}\right]}$$	Yes
Hidden Faults - Holes	Fothergill[89]	None: Mentions degradation of SI performance at high frequency	No
Hidden Geometry	None	None	No
Flanking Details	Craik[84] Van Zyl Erasmus et al[42] Mahn[90]	None: Mahn gives a value for expanded uncertainty but advices more work is required on determining the effect of workmanship, especially on more complex constructions.	No
Interrupting Elements	Fothergill et al - BRE[91]	Empirical	No
Surface Area	ISO 140[22]	Theoretical Formula *3-3:From para 3.5[22]* $$R' = D + 10lg\frac{S}{A}dB$$ D is level difference S is area of separating element A is the equivalent sound absorption area in receiving room	No
Room Volume	ISO 140[22]	Theoretical Formula *3-4: From para 3.5[22]* $$A = \frac{0.16V}{T}$$ A is the equivalent absorption area in square m. V is the receiving room volume in cubic m. T is the reverberation time in the receiving room in seconds	
Workmanship	Craik[57, 58]	Empirical	Yes
Effective Area: Step & Stagger	Sewell[92]	Theoretical Formula: Research	No

Depending on the specific test situation on site there may be other factors which could be included that contribute to the uncertainty of the measurement for example, pipework penetrations and airborne transmission paths such as along corridors in adjacent rooms.

Ideally the GUM and a DOE method should yield similar total variability. If BS5725 is to be followed, the test specimens should be identical and therefore the variability to be measured will be due to the repeatability (within laboratory variance) and the reproducibility (between laboratory variance) or at least the impact of the variability of the parts will be minimised. However, in field measurement of sound insulation of separating elements, the elements are known to vary even if the test specimens are identical[57]. This is due to variation in the sound field in the room, dimensional factors, construction or design. If the nominally identical "parts" or test specimens are a factor in the uncertainty in BS5725 then it is appropriate that a GUM assessment is required of this element of variability as an input variable.

Assuming the construction of the floor between all residential flats is identical there is still variability due to the physical dimensions of the part (floor) and the room (volume and the geometry). This can have a significant impact on the total variability because of the way sound insulation values are calculated. There could be workmanship issues, not necessarily with respect to the separating element, but with associated building elements or flanking details.

Even nominally identical test room pairs with the same shape, size and volume will yield different sound insulation performances as described by the random error. Research into this difference tells us not all of the difference is due to the measurement system alone[57, 58]. The quantification of the measurement uncertainty due to the part is considered in the final design of the experiment.

The input variables listed under the heading "part" can be classified under four main headings:

Spatial variation: Sampling within the rooms assumed to be a simple box shape;
Dependent on size;
Dimensional: (Surface area, room volume, effective area - Stagger);
Construction (Workmanship, hidden faults);
Design: (Hidden geometry, materials, interrupting elements, flanking details).

This research concentrates on separating floors between residential flats. The most common lightweight and heavyweight floor types are considered which typify the diverse methods of construction used today.

3.5.4.1.1 Spatial Variation

A major component of any measurement of sound insulation is the sound pressure level measurement in a reverberant room. The variability of a sound field in a reverberant space is described by Waterhouse [82] who determined the mean of several measurements of the sound field at uncorrelated points in a rectangular room, using the Gamma Probability distribution. Lubman [52] showed that where a room had irregular boundaries or contained objects, a statistical approach was repeatable and predictable. His work on intensity levels in a room shows that knowledge of the statistical distribution within a room makes it possible to determine the number of samples required to achieve a mean value with a specific level of precision. His later work [54] extends this to describe three situations: a laboratory rotating boom microphone; measurements along a line; measurements on the surface of a disk. This allows the user to rank order the effectiveness of various spatial averaging strategies and generates confidence limits for the sample mean [54].

For sound insulation testing the estimation of the standard deviation of sound pressure levels in source and receiver rooms is also well documented, both for diffuse fields and where direct sound from the source may be considered important [93]. Assuming the sound pressure levels are obtained by sampling according to the international Standards, and using the stationary microphone positions, which are remote from corners and room boundaries, and are in the reverberant field, the spatial variation of the mean square pressure is described by a gamma distribution [52],[46],[82]. The standard deviation of the sound pressure level generally is reliably predictable above the region where the modal overlap has a factor of three, commonly known as the Schroeder frequency (fs.)[94, 95]:

$$\mathrm{f_s} = 2000\sqrt{\frac{T}{V}}; \qquad \textit{3-5}$$

Schroeder showed that the normalised variance in a diffuse field above fs can be calculated using the bandwidth (*B*) and reverberation time (*T*) relationship: from Schroeder [46]:

$$\varepsilon^2(p^2) = \frac{1}{1+0.145BT} \qquad \textit{3-6}$$

The standard deviation is:

$$\sigma\ (dB) = \frac{5.57}{\sqrt{1+0.238BT}} \tag{3-7}$$

Craik developed a formula to predict the standard deviation of the sound pressure level in dB [55], where:

$$\sigma\ (dB) = \frac{4.43}{-0.22+\sqrt{1+0.319N}}: \tag{3-8}$$

where N = number of modes in the frequency band

If the direct sound field is considered important, approximation has been proposed by Michelsen [93] which allows for the distance d_{min} between microphone and source to be included. NB: the minimum distance in the international standards is 1m from source to microphone position.

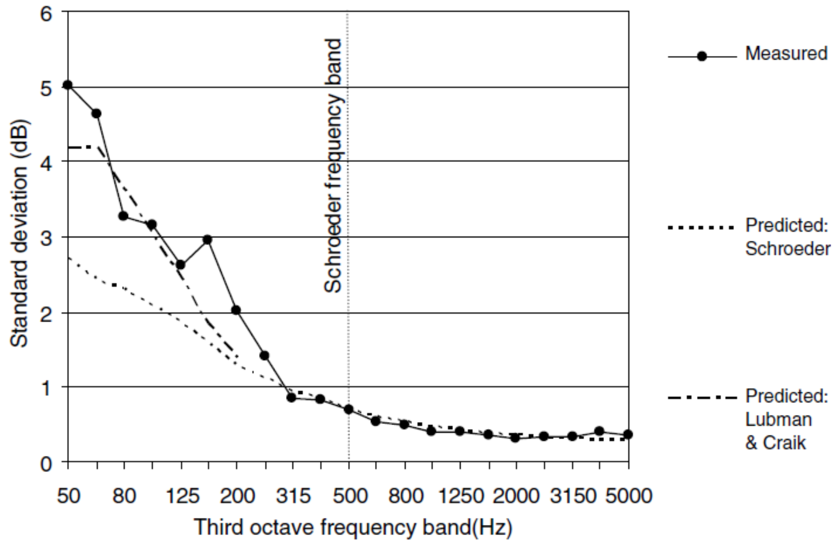
$$\varepsilon^2(p^2) = \frac{1}{1+0.145BT} + \left[\frac{\sqrt{\frac{A}{16\pi}}}{160^2\ d_{min}} \cdot \left(\frac{S_T\sqrt{A}}{V}\right)^3\right]: \tag{3-9}$$

with standard deviation:

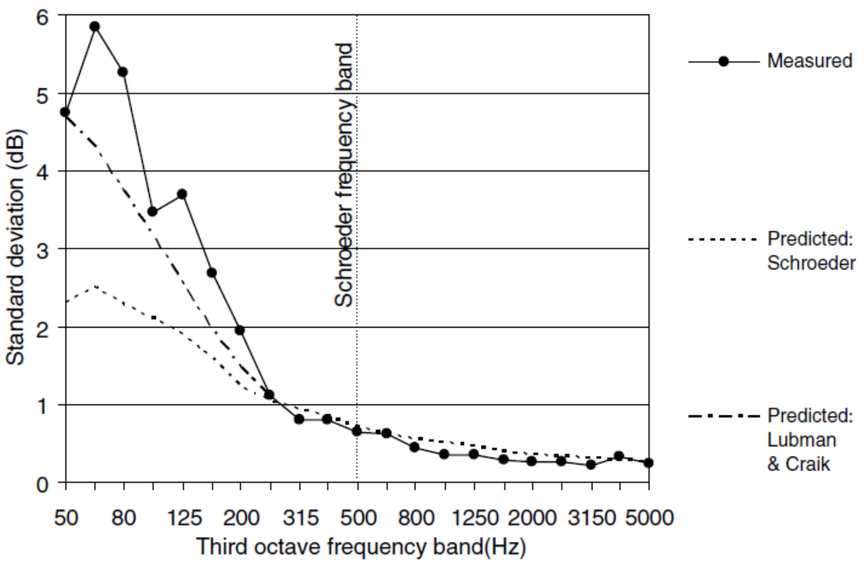
$$\sigma\ (dB) \approx\ 4.34\sqrt{\varepsilon^2(p^2)} \tag{3-10}$$

Comparison of field measurements with the predictions, using the standard theoretical formulae, are given by Hopkins et al [96]. Although the probability distributions in the source and receiver rooms are not identical the formulae are shown to provide relatively good fits with the measured data across the frequency range of interest, 100 -3150Hz, see Figure 3-2.

C. Hopkins, P. Turner / Applied Acoustics 66 (2005) 1339–1382 1369



Construction B – measured and predicted standard deviation of the sound pressure level in the source room.



Construction B – measured and predicted standard deviation of the sound pressure level in the receiving room.

Figure 3-2: After Hopkins et al [96]

The standard deviation at low frequency is large compared with that at high frequency. Unfortunately the accuracy of the prediction method may not give the confidence needed. GUM would require a complex combinations of uncertainties from these input values since it is necessary to consider each third octave band individually. As the individual contributions are so different across the full frequency range it may not produce a high degree of confidence in the combined uncertainty results.

Recently a study by Weise [97] revisits measurement uncertainty, using a comparison of diffuse and modal conditions with a different number of measurement points. He shows that when using only a few measurement positions (1, 3, 5 or 6) in a non-diffuse field, the measurement uncertainty is predicted to be significantly higher and GUM should be used with caution. The probability density distributions of a modal field with few measuring points deviates from a classic Gaussian bell shape. The asymmetry of the distribution means they are less predictable under GUM. This observation is reinforced by Mahn [4] who uses a GUM approach for predictions carried out under BS EN 12354 and finds a problematic situation where certain third octave bands have relatively large unexpected values of expanded uncertainty. Importantly, where Mahn has calculated these uncertainties, using practical guidance from the GUM and found that certain frequencies have confidence intervals that are significantly higher relative to the adjacent third octave bands, the GUM does not give explicit instructions for the estimation of uncertainty when the uncertainty is large compared to the results. His illustration for this is detailed in Figure 3-3 and shows that the expanded uncertainties in the 250Hz and 2000Hz bands is significantly higher than other calculated values. Mahn also notes that the standard laws of uncertainty may not apply accurately to these regions.

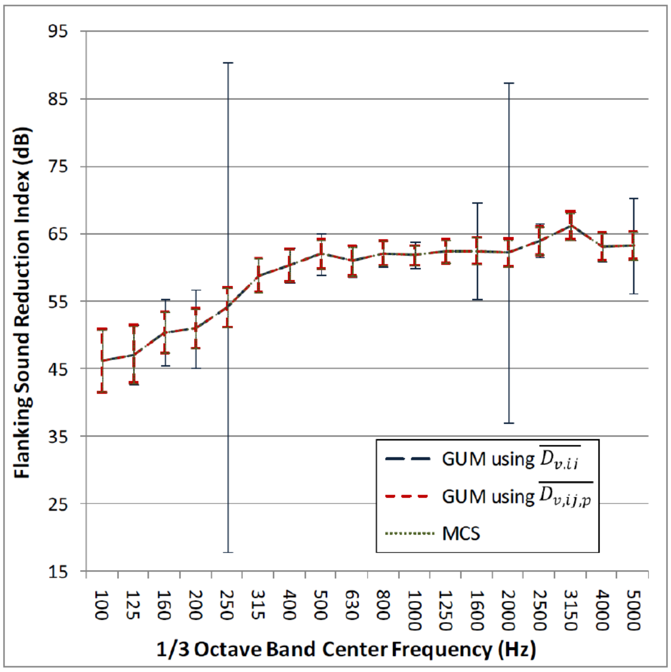


Figure 3-3: from Mahn [4, 85]

Where: $\overline{D_{v,ij}}$ and $\overline{D_{v,ij,p}}$ are the Direction Averaged Velocity Level Differences given by:

$$\overline{D_{v,ij}} = \frac{1}{2}\left[\overline{D_{v,ii}} - \overline{D_{v,ji}}\right]$$

$$\overline{D_{v,ij,p}} = \frac{1}{2}\left[\overline{L_{v,ii}} + \overline{L_{v,jj}} - \overline{L_{v,ij}} - \overline{L_{v,ji}}\right]$$

Given the research in this area of uncertainty [52],[46],[82] it might be assumed that the measurement of sound pressure levels provides a reliable imported input variable to the GUM. However, this may not apply across the full frequency range of interest. Mahn and Weise have highlighted this problem. In Mahn's case the research is focused on uncertainty in structural noise transmission whereas Weise's research is specifically for sound insulation. Mahn highlights the frequency related anomalies, discovered when breaking down the overall predicted uncertainty into the individual component third octave bands. Weise also shows high levels of variability accrue in small rooms at low frequencies with low modal overlaps. Therefore measuring sound in a lightly damped room, with few measurement positions, causes more unpredictably at low frequency than it does at high frequency and this may be a contributory factor that needs to be considered when examining the measurement uncertainty related to the part being measured.

There remains a requirement that the expanded uncertainty, calculated across all 16 third octave bands, is reliable and not only in the diffuse field region. There is a need to avoid situations where frequency bands, with high levels of uncertainty, have their influence exaggerated by the calculation method or are adversely affected by systematic errors from either the measurement instrumentation, test method or calculation procedure. This strengthens the case for DOE as the method of choice, to assess the contribution from the components of variance, because the uncertainty calculated across the frequency range, reflects the test sample selected. The test sample mix can be controlled to isolate variables and block confounding factors. The uncertainty determined for the test sample then is a true reflection of the variability associated with the measurement process and the specific floor construction methods used.

3.5.4.1.2 Dimension

Given a floor construction, a cause of variability is the size of the test element and the shape and volumes of the rooms, particularly the receiver room. These affect the sound insulation performance of all separating elements, particularly at low frequencies. They are based on assumed diffuse conditions and describe the relationship between the sound reduction index "R" and the standardised level difference D_{nT} defined in Table 2 of BS EN ISO 717-1: 1997 [17]. From Annex C2.BS8233: 1999 [37] the definition of the relationship between D_{nT} the surface area and volume of the room is:

$$D_{nT} = R - 10lg_{10}\,\frac{3S}{V}\ \text{dB} \qquad \textit{3-11:}$$

where: S is the area of the separating wall or floor in the field in square metres (m2); V is the volume of the receiving room in the field in cubic metres (m3).

Assuming diffuse conditions, doubling the room volume while retaining the same area of a floor or wall will increase the sound insulation performance by 3dB. From this functional relationship, the measurement uncertainty can be calculated and the input variable contribution included in the combined uncertainty calculation of the GUM.

Steps (vertical displacement) and staggers (horizontal displacement) also affect the performance of a separating wall or floor. Although the relationship determined by Sewell [56] is not confirmed empirically, an improvement in performance of about 3dB for a 300mm offset is indicated [56]. Investigations by Torjussen [98] of room offset surface areas indicated that a stagger gives minimal improvements.

In general, it is desirable to control the dimensional aspects associated with a room test specimen as far as is practicably possible. In a GUM framework, if the dimensional aspect were to be constrained it would be assumed that there would be close to zero impact on the measurement uncertainty. At present, no research is reported on this effect.

3.5.4.1.3 Construction

If the measured sound insulation of apparently identical floors differs, the layman may assume that this must be due to workmanship. In pre-completion sound insulation testing, there are many cases of constructions failing to achieve a given standard when other, apparently identical constructions pass. The causes can be a multiplicity of factors i.e. components of variance, other than workmanship which add to the variability of the measured value about the mean.

Workmanship, in this context, was defined by Craik [57] and it represented the difference between the variability associated with the instrumentation uncertainty and the variability due to the total uncertainty recorded. He concluded, that for a simple concrete floor, in rooms of identical dimensions and flanking details, the variation was 1.5 - 2dB across third octave bands. This was due to "workmanship" I.e. the variability left after allowing for instrumentation uncertainty.

This conclusion offers a lower limit contribution to the total variability in the measurement process. However, it does not include contemporary common floor constructions, nor does the method allow for a more detailed assessment of the other components of variance. It does however provide a starting point to this thesis study and also allows for a reasonable estimate for the imported input variable due to workmanship under the GUM assessment framework.

3.5.4.1.4 Hidden Faults

Hidden faults can have a detrimental effect on insulation. The degradation in performance that small holes can make was investigated by Fothergill [89] for solid masonry walls. The impact that holes and cracks have is generally centred on degradation of high frequency performance, although it was shown that this did not appear to detract significantly from the single figure overall wall sound insulation value ($D_{nT,w}$). It is difficult to see how an uncertainty budget incorporates individual faults, however common (or more probably uncommon), unless the budget was to be used for a specific construction type.

Again, there is impact on uncertainty at high frequencies (>400Hz). It is unlikely that for all GUM input variables, the relevant uncertainty at all third octave bands will be identical, available, or even calculable. Single figure sound insulation values may only be required or a limited range of third octave bands. This will further limit the imported input value uncertainty information and add to the complexity of the assessment. Again, it may require an "expert" in sound insulation measurement, to act by intuition to compensate for this shortfall in knowledge, if detailed third octave band data are to be assessed.

Hidden faults must be set to one side when considering an uncertainty budget, as they are generally specific to the part and not the measurement of the part. This would be difficult to do if the GUM approach, using individual input variables, is used as they are part of the quantified uncertainty and each type of hidden fault. In a DOE, using an ANOVA technique, the uncertainty due to the hidden fault would be attributable to the "part" component of variance and can be separated, to leave the uncertainty of the measurement laboratory or of instrumentation intact. In the GUM approach, this is not possible and is assumed to be present in the overall budget permanently. This may lead to an overestimate of total uncertainty, or if neglected it may give underestimates.

3.5.4.1.5 Design

Insulation can also be affected by building components other than the direct element. If there are interrupting elements, such as windows in external walls, they can have a positive impact on the sound insulation performance of, more usually, the separating wall. This method of improving the sound insulation is well known and is documented in the Building Regulations Approved Document E 2003. The flanking construction is important in determining the field performance of sound insulation. Sewell et al [92] showed how flanking contribution, due to different internal leaf constructions, affected the performance of a concrete floor where the end flats in a test block had a lightweight internal leaf of the external wall which was different from flats in the middle of the block.

3.5.5 Environment

Environmental factors can have an adverse affect on the stability of measurement process. Background noise is always an unknown factor and is specific to the test environment on site. It is possible to calculate its influence with relatively high degree of accuracy in the laboratory, where it may be low and stable. In field testing, it is uncontrolled and the assumption is that it should be minimised and the calculation procedure designed to include a correction. Meteorological conditions on site can vary significantly over time, much more so than the internal conditions experienced in a test laboratory. It is useful to consider these effects and their contribution to measurement uncertainty when testing sound insulation in the field.

3.5.5.1 Informative references

The factors that come under this heading are detailed in Table 3-4 and involve all components related to meteorological conditions, as well as the properties involved with the acoustics of the room.

Table 3-4: Input Variables - The Environment

The Environment: Input Variable	Reference Source	Possible Uncertainty Information	Direct Uncertainty Component: σ Empirical/Statistical
Background Noise	ISO 140 [22]	No; different allowance or mathematical corrections made. Corrections vary depending on signal to background level.	No
Barometric Pressure & Temperature	Wittstock et al Scholes[43]	Theoretical *3-12: after Wittstock & Bethke[44]* $L_{n,N} = L_{n,meas} - 20lg\left[\frac{B_{meas}}{B_N}\sqrt{\frac{T_N}{T_{meas}}}\right] dB$	No
Humidity	None	None	No

3.5.5.1.1 Temperature and barometric pressure:

Temperature affects repeatability of sound insulation measurements [43]. A series of measurements were designed to quantify the accuracy and precision of sound insulation measurements due to the loudspeaker and microphone positions using several operators. A 3.5^{o} C difference causes modal changes at low frequencies large enough to affect the overall third octave band sound pressure level e.g. 100Hz, by up to 5dB in a 40m^3 room.

Recent work by Wittstock et al [44] on absolute temperature and absolute static pressure effects lead to a normalisation of a sound reduction index and impact sound pressure level for both static pressure and temperature. The formulae for these follow a similar format and are detailed below.

$$R_N = R_{meas} - 20lg\left[\frac{B_{meas}}{B_N}\sqrt{\frac{T_N}{T_{meas}}}\right]dB \qquad \textit{3-13}$$

$$L_{n,N} = L_{n,meas} - 20lg\left[\frac{B_{meas}}{B_N}\sqrt{\frac{T_N}{T_{meas}}}\right]dB \qquad \textit{3-14}$$

Where:

R_N= normalized sound reduction index; R_{meas}= measured sound reduction index; $L_{n,N}$= normalized impact sound ressure level; $L_{n,meas}$= measured impact sound pressure level; B_N = reference static pressure; B_{meas} = measured static pressure; T_N = reference temperature; T_{meas}= measured temperature.

It therefore is possible to assess the expected temperature and pressure variation and calculate the uncertainty expected for both these parameters.

3.5.5.1.2 Humidity

In general the expected effect of changes in relative humidity on the measurement uncertainty is small. For example, the layer of quartz on some microphones absorbs moisture and leads to an increase in sensitivity of the microphone which is typically 0.4dB/100% relative humidity[45]. Humidity has an adverse effect when moving from a cold to a warm environment and may cause condensate build up on the microphone. This is easily rectified in the measurement procedure by allowing the instrumentation to adjust to room temperature.

3.6 Conclusion

The “bottom up” GUM method is an unsuitable method of calculating the measurement uncertainty relating to the field testing of airborne sound insulation. There is a paucity of suitable input values and assembling an uncertainty budget would have a significant amount of missing information. The gaps in the combined uncertainty budget thus would require “expert” knowledge. As Lyn et al have shown [62] this method has been found to be unreliable, generally leading to a significant overestimate of total uncertainty.

There are also issues with the requirement for independence, where individual input variables are known to be correlated and it is also clear that some individual input values have non-normal PDFs. Modelling the uncertainty of individual variables affected in this way may not be practical or offer an accurate approximation.

With a restricted data set, GUM may only be able to provide a “micro” view of the expended uncertainty in the measurement system. If significant uncertainty factors were discovered later in the process, there may also be less incentive to identify less intuitive contributors; their influence would be assumed negligible.

A better, more promising approach is a “top down” empirical approach similar to that described in BS5725. Unfortunately the BS5725 framework does not appear to cover all main uncertainty headings and their related components of variance that we wish to investigate in this study. An extension of the BS5725 analysis of variance (ANOVA) is required and the design of the experiment (DOE) should take into account the potential for variability caused by the part as well as identify any effects between factors.

The main advantages of ANOVA are given by Deldossi et al [66]: the randomization of the trials in the design guarantees the independence of the random effects and the error component; the factorial experimental design gives the possibility to estimate, if they exist, interactions among factors.

The ANOVA DOE and its statistical model is examined in detail in the next chapter.

4 Statistical analysis: design of experiment

4.1 Introduction

In Chapter 3 it was argued that the GUM is a relatively complex way of determining overall uncertainty in the process of measuring airborne sound insulation in the field, given the number of potential factors likely to affect the result and the reliance of expert knowledge to estimate the contribution from unknown input variables. This chapter looks at the most commonly used methods of assessing uncertainty in acoustic measurement and refers by example to some key studies involved in determining both repeatability and reproducibility in acoustic measurement. In addition some conclusions are drawn about the apparent difficulty some of the study authors have in explaining high levels of "between laboratory" variance (Reproducibility) using the DOE method from BS5725.

The limitations of the most commonly used methods of assessing measurement uncertainty are examined and an alternative approach to the ANOVA experimental design is suggested.

4.2 Historical use of ANOVA DOE

There are many examples where national and international laboratories have contributed to the understanding of measurement uncertainty and in particular the defined characteristics of repeatability and reproducibility by partaking in "Round Robin" studies. This is where a number of laboratories use their own facilities, operators, instrumentation and measurement procedures to measure the physical characteristics of a test specimen. Less commonly there are other round robin studies that have been done to reveal the magnitude of measurement uncertainty outside the laboratory environment by comparing acoustic measurements made in the field.

4.2.1 Laboratory inter-comparison

The inter-laboratory studies are relatively common in acoustics and they have as their benchmark the repeatability and reproducibility levels for both airborne and impact sound insulation testing detailed in Table A.1 and A.2 ISO 140-2[8] see Figure 4-1 and Figure 4-2.

Table A.1 — Repeatability values for laboratory tests

Third-octave band centre frequency Hz	Repeatability values r for airborne sound reduction index according to ISO 140-3[a] dB	Repeatability values r for normalized impact sound pressure level according to ISO 140-6[b] dB	Repeatability values r for the reduction of impact sound pressure level according to ISO 140-8[c] dB
100	4,5	3	1,5
125	4	2,5	1,5
160	3,5	2	1,5
200	3,5	2	1,5
250	2,5	2	1,5
315	2,5	2	1
400	2	2	1
500	2	1,5	1
630	1,5	1,5	1
800	1,5	1,5	1,5
1 000	1,5	1,5	1,5
1 250	1,5	1,5	1,5
1 600	1,5	1,5	2
2 000	1,5	1,5	3
2 500	1,5	1,5	3
3 150	1,5	1,5	3

[a] The repeatability values r are based on the weighted average of the following results from:

a) An inter-laboratory test conducted in 1976 involving eight laboratories in Germany, using as test specimens a wooden construction resembling acoustically a double-glazed window. Six complete tests were taken in every laboratory.

b) An inter-laboratory test conducted in 1983 involving five laboratories in Scandinavia using as test specimens a glazing in a staggered test opening. Six complete tests were taken in every laboratory.

c) An inter-laboratory test conducted in 1982 to 1985 involving seven laboratories in Belgium and in the Netherlands using as test specimens a lightweight partition and two brick walls with a mass per unit area of 225 kg/m^2 and 450 kg/m^2. These results were averaged.

d) An inter-laboratory test conducted in 1985/1986 as a BCR-project involving eight laboratories using as test specimens double-glazing (6/16/6). Six complete tests were taken in every laboratory.

[b] The repeatability values r are based on 34 tests conducted in periods in 1956/1957 involving one laboratory in Germany, using as test specimen a 12 cm concrete slab with floating floor.

[c] The repeatability values r are based on 21 tests conducted in 1983 involving four laboratories in Scandinavia, using a loosely installed flexible PVC-floor covering of category I (small specimens) having a weighted impact sound improvement index ΔL_w of about 14 dB.

Figure 4-1: Table A.1 Repeatability Values for laboratory tests – ISO 140 Part 2: 1993

Table A.2 — Reproducibility values for laboratory tests

Third-octave band centre frequency Hz	Reproducibility value R for airborne sound reduction index according to ISO 140-3[a] dB	Reproducibility value R for normalized impact sound pressure level according to ISO 140-6[b] dB	Reproducibility value R for the reduction of impact sound pressure level according to ISO 140-8[c] dB
100	9	5	2,5
125	8,5	4	2,5
160	6	3	2,5
200	5,5	3	2,5
250	5,5	3	2
315	4,5	3	1,5
400	4,5	3	1,5
500	4	2,5	1,5
630	3,5	2,5	1,5
800	3	2,5	2
1 000	2,5	2,5	3
1 250	3	2,5	6
1 600	3,5	2,5	9
2 000	3,5	2,5	11
2 500	3,5	2,5	11,5
3 150	3,5	2,5	8

[a] The reproducibility values R are based on the arithmetic mean of the following results from:

a) An inter-laboratory test conducted in 1983 involving five laboratories in Scandinavia using as test specimens a glazing in a staggered test opening. Six complete tests were taken in every laboratory.

b) An inter-laboratory test conducted in 1982 to 1985 involving seven laboratories in Belgium and in the Netherlands using as test specimens a lightweight partition and two brick walls with a mass per unit area of 225 kg/m^2 and 450 kg/m^2. These results were averaged.

c) An inter-laboratory test conducted in 1985/1986 as a BCR-project involving eight laboratories using as test specimens double-glazing (6/16/6). Six complete tests were taken in every laboratory.

[b] The reproducibility values R are based on comparison tests conducted between 1960 and 1980 involving 50 different testing teams using as test specimen a 14 cm concrete slab with floating floors of varying construction, installed in a test facility with flanking transmission, in Braunschweig (Germany). This is the best available estimate for the reproducibility values R of this test method.

[c] The reproducibility values R are based on 21 tests conducted in 1983 involving four laboratories in Scandinavia, using a loosely installed flexible PVC-floor covering of category I (small specimens) having a weighted impact sound improvement index ΔL_w of about 14 dB.

Figure 4-2: Reproducibility values for laboratory tests - ISO 140 part 2: 1993

It is important to the testing industry that when testing a generic specimen the difference in measurement results between testing bodies is small and relatively stable because commercially there may be a perceived advantage in testing at one laboratory over another. The understanding and quantification of this uncertainty is the driving force behind inter-comparison checks and several large studies have been carried out [30, 34, 35] traditionally following the experimental design detailed in BS5725. These inter-comparison checks often highlight significant differences between testing organisations, particularly with respect to reproducibility and the resulting curves compare poorly with the reproducibility values for laboratory tests from ISO140-2. An example of this is detailed in Fausti, et al [32] where at some frequencies the reproducibility levels (R) are 7 – 8dB above the ISO curve; see Figure 4-3 below.

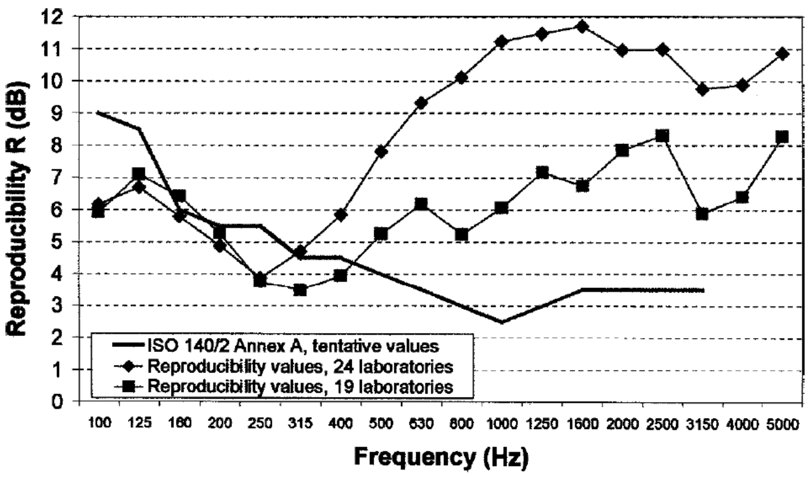


Figure 4-3: from Fausti, et al [32]

There are two elements of interest that stem from this finding. The first is that the testing bodies need some form of regulatory control. A study published by Weise et al [40], proposes criteria for new laboratories which uses the historical round robin test data as a benchmark to regulate the amount of variability allowed in any measurement process.

The second, and more importantly with respect to this research, is that the root causes of the dispersion of reproducibility results in Figure 4-3 remain unclear, even after attempts to eliminate outliers. This forces the testing bodies and their working groups to look ever more closely at their test facilities, test specimen construction and reporting procedures[99] without a specific focus on what is contributing to the large levels of variability.

The analysis methods used in these round robin studies following BS5725 DOE do not provide information on the components of variability that make up the reproducibility values and it has been shown by Meier et al [28, 29] that there is significant variability in third octave band levels of up to 3dB introduced into the measurement process, due to the differences in damping and the effects of flanking between laboratories.

4.2.2 Field sound insulation testing: inter-comparison

Research by Lang [36] provides a partial contrast to the laboratory round robin studies because it was undertaken in the field using the test methods in BS5725. Measurements and assessments were made independently of a separating floor in a pair of 70m^3 rooms in an office building and in a school. There were 11 participants who measured the airborne sound insulation performance of a separating floor in an office building. The results of these separating floor tests were collated, analysed and illustrated graphically. See Figure 4-4 below:

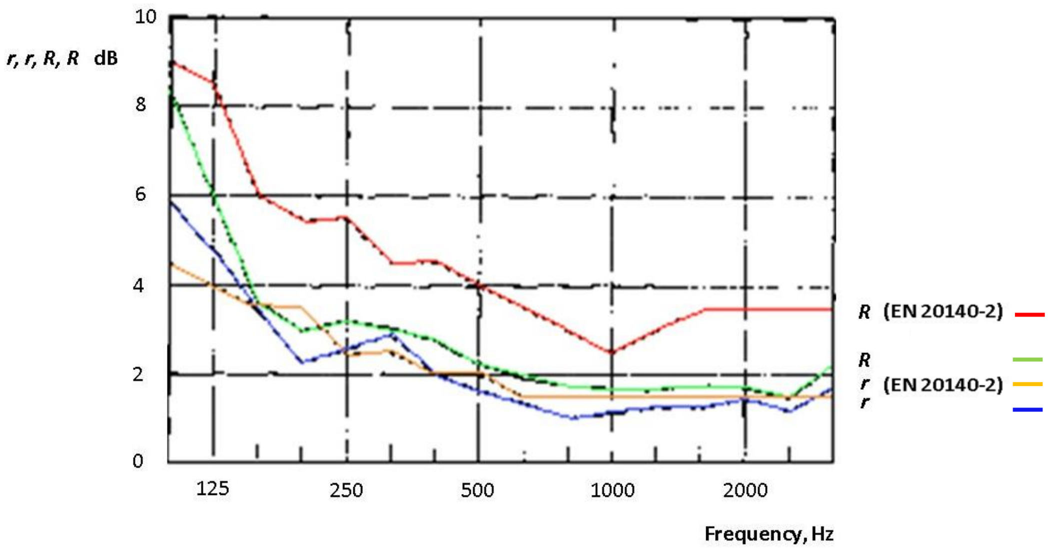


Figure 4-4: after Lang [36]

There were 13 participants in the sound insulation measurements in a school but unfortunately the published paper appears only to graph the repeatability data for this test.

Lang concludes that apart from some issues with background noise influence at higher frequencies the repeatability and reproducibility figures compared well with the tables in the ISO document. Based on basic statistical assumptions[5, 100], if all organisations were to use a similar test method (which they did) the repeatability components would be similar and could be pooled, so for repeatability this conclusion would normally be expected.

By contrast, the reproducibility values in Lang's study do not feature all the same variability constituents of a normal inter-laboratory study.

The reproducibility figures in tables A.1 and A.2 in ISO140-2 were created from several laboratory round robin studies where different test specimens ranging from glazing elements to heavy block walls were supplied, reconstructed, or remounted, in each laboratory which also, as shown by Schmitz et al [28], have differing damping and flanking components. This, as discussed earlier, has the potential to increase the level of total variability significantly across the frequency range as well as the variability in reconstructing the part. Both of Lang's tests were carried out on a single separating floor. The variability, due to the reconstruction of the test specimen in different situations and in rooms with differing flanking and damping components, is not present in Lang's data set. The" part to part" variability must therefore be nil. Because of this the direct comparison of this field test reproducibility data with the ISO 140-2 table is not a fair one. It could easily be added to the DOE by having the participants test the floor in more than one pair of identical rooms and extending the statistical analysis to allow the floor to be one of the "effects" scrutinized.

This is noteworthy because it is the intention of this research to specifically look at and identify the contribution of the part variability, amongst other things, to the overall variance in the calculation of measurement uncertainty and this will incorporate identifying individual "effects" such as the specimen or part to part effect which are likely to contribute to the overall variability in the measurement process. In addition the construction type of the floor being measured is addressed in this study as the total uncertainty measured may also depend on the construction being measured. The individual component effects can then be isolated by use of analysis of variance techniques and any interactions between factors investigated. The way that these techniques have evolved and been applied in acoustic measurement in inter-laboratory comparisons is discussed below.

4.2.3 Analysis of variance - ANOVA

4.2.3.1 Background

The inter-laboratory study protocols generally focus on the two simple terms "repeatability" which is the within laboratory variance see the general formula 4-1 and the "reproducibility" which is the between laboratory variance see 4-2. These are described by Mandel [100] who contributed significantly to the early work in this area of statistics and also to the international standards for inter-laboratory measurement studies (BS5725) and describes the within laboratory variance as:

$$s^2 = \frac{SS}{DF} = \sum(x - \bar{x})^2 \ /(n-1) \qquad \textit{4-1}$$

Where SS stands for sum of squares and DF stands for Degrees of freedom. Mandel denotes S^2_1, S^2_2, S^2_3 .... S^2_p for ***p*** laboratories, making ***n*** estimates on ***q*** samples.

It is assumed that given a generic test procedure, this “within” laboratory variance can be effectively pooled for all laboratories and that it will be close to the “true” variance of the test method σ^2. The pooled variance is given by Mandel as:

$$S^2_{pooled} = \sum_i SS_i \ / \sum_i DF_i \ . \qquad \textit{4-2}$$

$$E\left(s^2_{pooled}\right) = \sigma^2 \qquad \textit{4-3}$$

where E = the estimate of the true variance.

This assumption is adopted in the current British Standard BS5725-1 see 4-4:

For the between laboratory variance

$$S^2_{xi} = \sum_i (\overline{x}_i - \overline{x})^2/(p-1) \qquad \textit{4-4}$$

where $\overline{x}$ is the average of $\overline{x_i}$, and is an unbiased estimate of σ_L^2 (variability between laboratories)

Mandel defines the estimate of the pooled within laboratory variance and the between laboratory variance as[100]:

$$\hat{\sigma}^2 = s^2_{pooled} \qquad \textit{4-5}$$

$$\hat{\sigma}^2_L = s^2_{\hat{x}_i} - s^2_{pooled}/n$$

The caret indicates a sample estimate.

This is replicated in a rearranged form in the British Standard[7] to reflect the definitions of repeatability and reproducibility thus: see BS5725-1 [7]

$$\sigma_r = \sqrt{var(e)} \qquad \text{4-6}$$

For the standard deviation for repeatability (pooled data) and after BS5725-1 [7]:

$$\sigma_R = \sqrt{\sigma_L^2 + \sigma_r^2} \qquad \text{4-7}$$

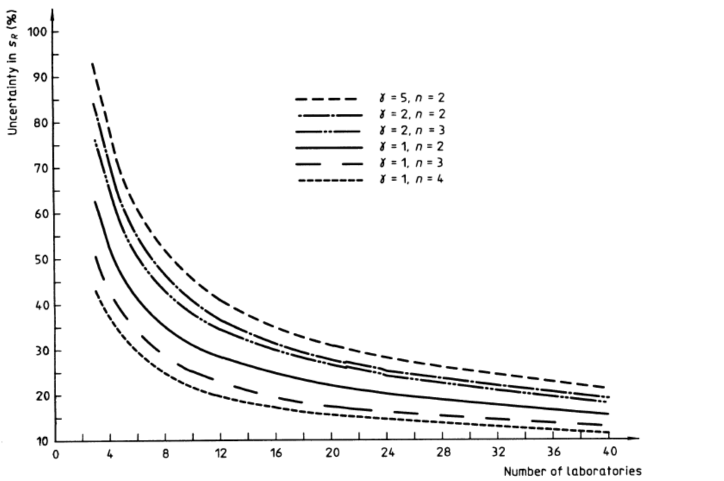
For the standard deviation for reproducibility.

It is worth observing the reproducibility defined here for BS5725, [4-7], combines` the between laboratory variance and the within laboratory variance.

4.2.3.2 Robust statistics - sampling

Mandel[100] comments on the need for statistical robustness in the sampling regime and specifically the number of organisations that should be participating in the inter-laboratory experiment. The sampling design and statistical robustness relates to the degrees of freedom. He uses as his example 13 laboratories testing 8 specimens 4 times each and highlights later in his paper that for reasonably robust statistical experiment it is common to require 30 degrees of freedom implying 31 participating laboratories with a caveat that this number may not be practical or realistic! In this case some estimates are based on more information than others. A sample containing 31 laboratories is based on more information than a sample size of 5 (as we have 5 operators with test kits). The degrees of freedom of an estimate is the number of independent pieces of information on which the estimate is based, hence there would be 31 laboratories if there was 30 degrees of freedom as one piece of information would have to be fixed in order to produce the estimate of the population mean. In general, the degrees of freedom for an estimate is equal to the number of values minus the number of parameters estimated to arrive at the estimate in question.

The uncertainty related to the calculation of reproducibility is also covered in a simple chart in BS5725-1 see Figure 4-5: after BS5725-1 Figure B.2 [7]. The precision of this confidence interval can be expressed as a probability, in this case 95%. The percentage uncertainty in the standard deviation is a function of the number of participating laboratories, the number or replicates (n) and the ratio of reproducibility against repeatability, S_R/S_r (ɣ).



The amount by which s_R can be expected to differ from the true value within a probability level of 95 %

Figure 4-5: after BS5725-1 Figure B.2 [7]

With low numbers of participating laboratories the uncertainty rises significantly and improvements rapidly follow diminishing returns as the numbers of laboratories increase. This holds true for all cases:

The rule of thumb is that more data is better although the effort involved in achieving a single sound insulation test result is a key factor and will obviously limit the number of samples carried out in the field in a day which will in turn limit the number of participating organizations which can be involved.

A balanced approach will be necessary to accommodate this constraint and more investigation will be needed on the minimum practical sample size.

Mandel [100] details the limitations of repeatability and reproducibility and suggests there may be other components of variability of interest e.g. comparison of test materials provided by different manufacturers or the behaviour of individual laboratories.

He concludes that where improvements in precision in the measurement process are a defined objective it is **<u>only</u>** the between laboratory variance that will be required to change. He identifies the factors that make up that effect as, the measurement methods and detection of outliers. It is notable that he does not specifically mention the influence of other factors such as the operator carrying out the tests, though it is clear that in field testing there are a number of factors whose effects can be segregated and individually assessed. Constructing the components of variance in this manner is a key part of this research and a prime reason for adopting the ANOVA process in inter-laboratory studies. There are a few examples of the use of this technique in acoustics which demonstrate this potential.

4.2.3.3 ANOVA – current examples

Analysis of Variance, (ANOVA) as a tool has previously been used in acoustic studies but is not a particularly common feature in acoustic data analysis.

Two good examples of ANOVA used in this context are a laboratory sound insulation study by Taibo et al [63] and a round robin study on the measurement of absorption coefficients by Davern et al [64, 65]. The former study published in 1983 utilizes a DOE where the test specimen is measured multiple times and is remounted for each operator but it is severely constrained by the fact the measurements are made by only two operators in the same laboratory. This means that there is statistically only 1 degree of freedom in the calculation of the reproducibility which is a significant contrast to the recommendations proposed by Mandel [100]. We also know that it is not strictly in accordance with the reproducibility definition, because there is only one laboratory involved in the test regime. Taibo et al emphasize their restricted data set and explain the reasons as the data collection difficulties of measuring sound insulation.

The latter example is of particular interest because it is based on a simple inter-lab round robin comparison study of absorption coefficient measurement but it extends the ANOVA to cover factors other than r & R. It combines the contributions from 13 testing organisations with a total of 21 different test chambers and is detailed in Report 1 dated 1980. In 1985 the same data set was reanalysed to separate out the components of variance in the measured data. This proved to be very informative and useful because it identified three systematic differences which were; between rooms, between methods and between specimens and one random uncertainty inherent in the basic test method. These are detailed in Figure 4-6 with their magnitudes:

systematic differences between methods - 0.0036,
systematic differences between rooms - 0.0021,
random uncertainties inherent in basic method - 0.0011,
systematic differences between specimens - 0.0004.

Figure 4-6: after Davern & Dubot [65]

The total inter-laboratory variance averaged over all frequencies was:

$S^2(\alpha_{so})_R$ = 0.0072

Figure 4-7: Result for inter-laboratory variance: Daven & Dubot [65]

The authors concluded that if the new international standard for the measurement of absorption (ISO354 [23]) was to be adopted, a significant reduction in the second largest contributor, the room difference component (0.0021) could be anticipated but none of the other variance components, which formed approximately two third of the total variance, would be expected to change. The overall impact would therefore be limited. This is a very good example of how the additional information, provided by an analysis of variance assessment can assist in predicting the impact of changes to the measurement process by isolating the qualitative components and quantifying their impact. ANOVA breaks down the individual components and, more importantly, quantifies the magnitude of their contribution to the total variance in the data. Contrast that with the high levels of variability in the reproducibility discovered by Fausti et al [32] and later attributed to the flanking via the frame that forms the test aperture see Smith et al [101] and there is a strong case to incorporate these techniques into any DOE trying to determine the contributors to measurement uncertainty.

4.2.3.4 ANOVA - DOE

ANOVA use in acoustics for inter-laboratory studies using BS5725 and comparison with the guideline value of r & R is relatively common, although the components contributing to reproducibility may not always be appreciated. It has also been shown that there are isolated instances where ANOVA DOE has been extended to provide valuable information on the components of variance and their magnitude though it is recognised that the experimental techniques required are onerous due to the time constraints on data collection and number of participants required.

Further investigations have revealed that there are several specialised statistical models based on ANOVA that focus on the assessment of measurement systems. In particular the Automobile Industries Action Group (AIAG)[102] has developed a system manual for using ANOVA to assess the components of r & R in addition to the components of variance in the measurement system.

A suitable selection of ANOVA models for this purpose is also identified by Burdick et al[67]. They are called Gauge Repeatability and Reproducibility studies (often abbreviated to Gauge R&R or GRR). The aims of a GRR reflect the aims of the thesis in that they are designed to:

(i) determine the amount of variability that is due to the measurement system;

(ii) Isolate the sources of variability in the measurement system.

In particular they allow the variance due to the part to be assessed and also any interaction to be determined between factors.

They are used extensively for quality control purposes in the manufacturing sector to assess whether the measurement system is able to discriminate between good and bad products or specimens without significantly high misclassification rates.

It is therefore appropriate to consider ANOVA for the assessment of measurement uncertainty in sound insulation measurement.

4.2.3.5 ANOVA – GUM

It is worth noting that ANOVA is mentioned in GUM see section H.5 [6] and also H.5.3 but it references ANOVA as a special method used in the certification of reference materials by inter-laboratory testing. It identifies a simple ANOVA design and provides a numerical example but dismisses the procedure as impractical because in most measurement situations it is only possible to evaluate a few components of uncertainty using ANOVA methods. It reiterates that as many input variables as possible must be included in the model and, from paragraph 4.3.1 , using scientific judgement and other measures evaluate the input variables using:

- previous measurement data;
- experience with or general knowledge of the behaviour and properties of relevant materials and instruments;
- manufacturer's specifications;
- data provided in calibration and other certificates;
- uncertainties assigned to reference data taken from handbooks.

It may not be possible to identify all the input variables and as mentioned before the assembly of a model where the specific contributions of the variables are not accurately represented has led to modelled uncertainties being significantly greater by a factor of 6 than those determined empirically [62].

4.3 Conclusion

This chapter has looked at the assessment methods used in inter-laboratory comparison studies and has identified that most follow the DOE and simple statistical analysis contained within BS5725 focussing on the calculation of r & R.

It has noted that although repeatability results follow expectations some of the laboratory studies have obtained results for reproducibility that are outside the expected range. Those that are lower can be explained if the reproducibility is calculated on measurements carried out on a single test specimen e.g. see Lang. Others, where the unexpected levels cannot easily be explained by the elimination of outliers require further information on the part to part variability of the construction being tested.

In addition, research by Weise et al [40] into the causes of relatively high levels of between laboratory variability shows that there is a significant contribution from the damping and flanking components which differ between test facilities and also there is a variability attached to the test specimen where different ones were supplied, reconstructed or remounted in each facility.

Current analysis techniques which determine the r & R from these round robin studies do not provide the level of information that makes a considered response to the causes of the high levels of variability between laboratories possible. It also means that the attempts to improve reproducibility cannot be efficiently targeted and organisations are left with a relatively large list of possible improvement actions without the ability to prioritise the list.

Further investigations have uncovered a more advanced statistical technique which has been demonstrated with minor modifications can be used on a round robin study data set to draw out robust statistical evidence about the quality and quantities of the factors affecting the total variance in the measurement system. The analysis of variance (ANOVA) will be used in this research, modified to allow efficient designs of the experiment without compromising the statistical robustness of the results it supplies.

The next chapter will look at the basic statistical theory behind ANOVA and examine the modified ANOVA model known as a Gauge Repeatability and Reproducibility study. It will show how it can be used to assess measurement systems efficiently and designed to cover the field measurement of the most common types of floor construction.

5 ANOVA

''A successful gauge capability study, one that provides good estimates of the variation in the measurement process and identifies the factors that are most influential to that variation requires more than just an accurate statistical analysis."

Montgomery & Runger.

5.1 Introduction

In Chapter 4 the limitations of the standard inter-laboratory test method, as recommended in BS5725, were highlighted together with a critique of current research. An alternative was suggested where the reproducibility was partitioned into the contributory components. This allows the qualitative and quantitative variability of reproducibility to be expressed.

This approach is based on the analysis of variance (ANOVA). This chapter examines the general linear model, in particular, where this can be applied, the factors that are chosen and how their effects are determined. It identifies an associated field of statistics which developed around measurement system analysis (MSA) and studies of measurement gauge repeatability and reproducibility (GRR). The standard protocols for this treatment of uncertainty, in commercial and industrial metrology, are highlighted.

An ANOVA model, based on GRR, is proposed for the determination of measurement uncertainty in field testing of sound insulation.

5.2 Analysis of Variance (ANOVA)

The basic requirement of this research is that the measurement uncertainty, inherent in the measurement of sound insulation, can be separated into factors that better describe and quantify the contributing components of variance. This is the qualitative and quantitative approach referred to in earlier chapters. It is known statistically as a general linear model.

As no first principle or mathematical functional model exists to determine sound insulation the model is empirical, relying on an understanding of the process and objectives to design a suitable experiment.

It is helpful at this stage to examine the method, in particular the definition of the model and format in which the results are reported. The ANOVA method considers the variation in a measurement (or response variable) and attributes the observed variation to either assignable causes or to random variation. Random variation in this context is analogous to random error as discussed in Chapter 2.

The assignable causes are user defined fixed factors, selected to determine their influence. It is usually assumed that the factor effects on the response variable (measurement) are constant. The complexity of the multifactor experiment will depend on the number of factors selected.

The simplest ANOVA model, the one-way ANOVA, deals with one fixed factor and is defined as:

$$Y_{ij} = \mu_Y + P_i + E_{ij} \qquad \textit{5-1}$$

Where;

μ_Y = population mean

P_i = parts

E_{ij} = random error

i = 1,......p : number of parts, j = 1,.......,r : number of repetions, µ is a constant and P_i and E_{ij} are jointly independent normal random variables with means of zero and variances σ_P^2 and σ_E^2, respectively. The ANOVA model for 5-1 is detailed in Table 5-1.

Table 5-1: ANOVA for Balanced one factor Random Model: in 5-1

Source of variation	Degrees of freedom	Mean square	Expected mean square
Parts (Treatment)	P – 1	S_P^2	$\theta_P = \sigma_E^2 + r\sigma_P^2$
Replicates	P(r - 1)	S_E^2	$\theta_E = \sigma_E^2$

where:

Parts are the specimens being measured;

Replicates are the number of times (repetitions) each part is measured.

The model could be used, for example, to determine the variability between parts measured by a single operator where p = 10 parts and replicates, r = 2 repeat readings; i.e. 2 measurements of each of 10 parts by a single operator;

In addition to S_P^2 and S_E^2 , the variance of the deviations from the grand mean (mean value of all the measurements) can be calculated from the data set where:

$$S_{total}^2 = \frac{1}{pr-1}\sum_{i=1}^{p}\sum_{j=1}^{r}(y_{ij}-\bar{\bar{y}})^2 \tag{5-2}$$

S_{total}^2 is divided into the total variance within sample S_E^2 and the total variance between samples S_P^2. The relationship is determined from the estimated total variance:

$$(pr-1)S_{total}^2 = (p-1)rS_{\bar{y}}^2 + p(r-1)S_E^2 \tag{5-3}$$

This highlights the key reason for employing ANOVA in this research, as the total variance is made up of two quantifiable components.

The results for the ANOVA analysis are generally represented in two tabular formats. One is a summary table, with the calculated variances and the respective degrees of freedom; the other denotes their sums of squares, mean squares and degrees of freedom. Examples of both are detailed in Table 5-2 & Table 5-3:

Table 5-2: Calculated variances - One way ANOVA Table

Source	df	$\hat{\sigma}_y^2$	F
Treatment	p – 1	$nS_{\bar{y}}^2$	$F = \frac{ns_{\bar{y}}^2}{s_E^2}$
Error	p(r – 1)	S_E^2	
Total	pr - 1	S_{total}^2	

Table 5-3: Mean squares - one-way ANOVA Table

Source	df	SS	MS	F
Treatment	p – 1	$SS_{treatment}$	$\frac{SS_{treatment}}{df_{treatment}}$	$\frac{MS_{treatment}}{MS_{error}}$
Error	p(r – 1)	SS_{error}	$\frac{SS_{error}}{df_{error}}$	
Total	pr - 1	SS_{total}	$\frac{SS_{total}}{df_{total}}$	

Table 5-3 is the analysis of variance (ANOVA) table for the simplest model, a single factor experiment. The same format applies for more complex analyses.

It is possible to look at numerous factors or treatments in order to identify other variability contributions provided the experiment is designed accordingly. For measurement system analysis (MSA) we need to look at a special case of ANOVA, or advanced ANOVA, that will determine the components of reproducibility.

5.2.1 Advanced ANOVA

Advanced ANOVA has been developed specifically for measurement system analysis. It is sometimes referred to as a gauge repeatability and reproducibility test or GRR. It has different qualities from inter – laboratory studies which focus on bias or offset of measurement, offering commercial advantages. Its aims are to consider measurement system variation to determine if it is fit to discriminate the parts it is measuring efficiently, within an acceptable tolerance range. The variability observed when measuring ideally should be due to the variability in the parts being measured, and not the variability associated with the measurement system. If the variability associated with the measurement system is demonstrably relatively small, then it can be considered acceptable. If the variability is too large then it must be improved e.g. by modifying the measurement procedure, upgrading the instrumentation or retraining the operators.

The focus of this study is on the differentiation of the components of variance and their size i.e. contribution. This is not confined to the assessment of the suitability of the measurement system. However there are general rules to assess the suitability of measurement systems, see Wheeler[103] who gives a 4 tier class structure for process monitors and the Automotive Industry standard based on a comparison of precision to tolerance ratio [102, 104-106] .

In a GRR study the "repeatability" is associated with the variability of the measurement instrument (or gauge). The variability due to the biases between operators is known as the "reproducibility". In BS5725 DOE the reproducibility contains both the "between" laboratory variability and the "within" laboratory variability. The ANOVA definition of reproducibility does not contain the within laboratory variability. This difference is taken into account and will be shown when the results from the two methods are compared later in the thesis.

Early evaluations of ANOVA are reported by Mandel[100] and Tsai[107] and provide a good example of how these techniques can be used effectively in measurement system analysis. In addition a handbook for Measurement System Analysis (MSA) has been developed by the Automotive Industry Action Group [102, 104-106] for their quality control procedures and there has been an increasing amount of research into the statistical analysis of measurement system uncertainty from the mid 1990s.

A key contribution by Montgomery is reported in [68, 69] as the development of GRR has culminated in a monograph on the subject, including its special applications [67]. Recent research in this developing area of statistics informs this research on achieving an accurate and reliable estimate of the variability in the measurement process due to the part, operator and instrument. The approach in this thesis is outlined in the following.

5.2.2 Basic GRR Experiment Design Considerations

The GRR relies on a number of gauge "operators" to measure a number of test specimens a repeated number of times. The manner in which the data is collected is addressed, in order to enhance the statistical robustness of the data, improve the validity and efficiency of the study and minimise the effect of systematic bias.

Montgomery[108] lists three main elements in DOE for GRR:

Replication: is the repetition of the experiment and involves a refresh of the measurement procedure. E.g. not only pressing the record button on the sound level meter in the same position. It is expected that the setup of the meter in the room and the loudspeaker position and sound pressure level will alter as it would if you were measuring the same wall at a different time on the same day. It defines the experimental error and it represents the repeatability of the measurement instrumentation.

Randomization: is the basis of any robust statistical experiment. It means that each operator should measure the parts in a random order. Randomising reduces systematic bias in the experiment data.

Blocking: is a technique used to minimise nuisance factor variability. For example, to reduce the variability due to room dimensions, element surface area and room volume affecting the total variability the same room shape and volume is selected for the test pairs. Metrological conditions can also be blocked, if the testing can take place over a relatively short timescale.

The preferred GRR contains a "balanced", "crossed" design. That is, every level of one factor is run with every level of another factor (crossed) and each measurement is repeated the same number of times (balanced), e.g. every part in the test sample is measured by every operator the same number of times.

5.2.2.1 Selecting sample size, replicates & operators

The MSA application in industry covers a wide range of measurement devices and techniques. It is usually relatively straightforward to collect a representative test sample and to repeat the measurement with a reasonable number of operatives. What dictates the practicality of the GRR experiment is the availability of operatives and the time they take to complete a measurement. Replicates and a usable sample are the least constraining variables in the measurement process.

Vardeman et al [109] have commented on the relatively small test sample basis run by the majority of GRR studies, which commonly restrict their setup to 2 or 3 operators [110, 111]. This is also insensitive to increases in the part or replicate numbers and will generally result in relatively wide confidence intervals for the operator variance component (σ_0^2).

Burdick et al [67, 72] offer a practical view, recommending at least six operators in any GRR study with random operators. Between 10 to 20 parts with 2 – 3 replicates will produce reasonable confidence intervals for most of the other variance components. It is noted that in the case of the sample of parts selected for measurement, there are distinct advantages to selecting more parts with fewer replicates, and being aware of the implications of restricting the parts sample to one part of a production run or, in this research, construction type.

Firstly, the gauge (or measurement system) being assessed may provide less variable results on a particular "standard" part which is representative only of a middle range of performance than at the extremes of possible measurement and at the very least some examination of this must be included in any GRR assessment of a measurement system. This is called a "Linearity Test" and in our case would include examples of test elements which cover examples in addition to the test samples of the lightweight and heavyweight floors and which feature sound insulation performances lower and higher than the timber and concrete floor test samples. The reason this test is important is because the measurement process might respond in a non-linear way when measuring very high or very low levels of sound insulation. The selection of parts used in this research will span the range of performance expected in the wider population. Under normal circumstances, the variability of a separating floor will only be 5 – 6dB when measured on a single site. Additional low performing building elements will be sampled and high performing situations manufactured to ensure the high and low ranges are included. However, it is beyond the scope of this research to identify the upper and lower ranges, where the model will be expected to break down. If non-linearity is discovered in the data set, it will be identified.

Secondly, the variability of the measurement system may be affected by the construction being measured. The sound field in a room of equivalent size may have an adverse effect on the measurement process if the building is of heavy construction (concrete) rather than lightweight construction (timber). Consideration is given to ensuring the parts being measured are representative of the range of constructions commonly encountered.

Finally, with fewer replicates (2 or 3), there is less tendency for the operators to shortcut the measurement procedure. It is very important if the process of repeatability is to be representative of the true repeatability that the “fixturing” of the test, the setup of the equipment in a sound insulation measurement, is carefully repeated.

5.2.3 The model

The classical gauge repeatability and reproducibility study is a balanced two-factor crossed random effects model with interaction:

$$Y_{ijk} = \mu + O_i + P_j + (OP)_{ij} + E_{k(ij)} \quad \text{5-4}$$

where *i* = 1, 2,....., *p : j* = 1, 2, ..., *o* : *k* = 1, 2, ..., *r*

and; p = number of parts, o = number of operators and r = number of repetitions

and O_i, P_j, $(OP)_{ij}$, and $R_{k(ij)}$ are random variables representing the effects of the operator, parts, operator by part interaction and the replications on the measurement and μ is an overall mean.

It is assumed that the operator, parts, operator by part interaction and the replication are independent random effects, that are normally distributed with zero means and variances of σ_O^2, σ_p^2, σ_{Op}^2 and σ_R^2 , respectively that are assumed constant.

Assuming the mean to be zero implies the measurement system is unbiased i.e. the variability in the measurement system is due to precision only. Assuming the variances to be constant indicates the variability of the system does not change with the magnitude of measurement.

As in the one-way model the description of the two-way ANOVA model, its mean squares and means, the covariance structure and distributional information is detailed in Table 5-4, Table 5-5, Table 5-6, Table 5-7 & Table 5-8.

ANOVA for the balanced two-factor crossed random model [67] see two way model defined by equation 5-4.

Table 5-4: ANOVA for two-way model: 5-4

Source of variation	Degrees of freedom	Mean square	Expected mean square
Parts (P)	P – 1	S_P^2	$\theta_P = \sigma_E^2 + r\sigma_{PO}^2 + or\sigma_P^2$
Operators (O)	O – 1	S_O^2	$\theta_O = \sigma_E^2 + r\sigma_{PO}^2 + pr\sigma_O^2$
P x O	(p – 1)(o – 1)	S_{OP}^2	$\theta_{PO} = \sigma_E^2 + r\sigma_{PO}^2$
Replicates	po(r – 1)	S_E^2	$\theta_E = \sigma_E^2$

Mean squares and means for the balanced two-factor crossed random model [67]

Table 5-5: Mean squares & means for two-way model: 5-4

Statistic	Definition
S_P^2	$\frac{or\Sigma_i\,(\bar{Y}_{i**} - \bar{Y}_{***})^2}{(p-1)}$
S_O^2	$\frac{pr\Sigma_j\,(\bar{Y}_{*j*} - \bar{Y}_{***})^2}{(o-1)}$
S_{PO}^2	$\frac{r\Sigma_i\,\Sigma_j\,(\bar{Y}_{ij*} - \bar{Y}_{i**} - \bar{Y}_{*j*} + \bar{Y}_{***})^2}{(p-1)(o-1)}$
S_E^2	$\frac{\Sigma_i\,\Sigma_j\,\Sigma_k\,(\bar{Y}_{ijk} - \bar{Y}_{ij*})^2}{po(r-1)}$
$\bar{Y}_{i**}$	$\frac{\Sigma_j\,\Sigma_k\,Y_{ijk}}{or}$
$\bar{Y}_{*j*}$	$\frac{\Sigma_i\,\Sigma_k\,Y_{ijk}}{pr}$
$\bar{Y}_{ij*}$	$\frac{\Sigma_k\,Y_{ijk}}{r}$
$\bar{Y}_{***}$	$\frac{\Sigma_i\,\Sigma_j\,\Sigma_k\,Y_{ijk}}{por}$

Distributional results for balanced two-factor crossed random model [67].

Table 5-6: Distributional results for two-way model: 5-4

Result	
1	$\bar{Y}_{***}$, S_P^2, S_O^2, S_{PO}^2 *and* S_E^2 *are jointly independent*
2	(p - 1)S_P^2 / θ_p is a chi squared random variable with p – 1 degrees of freedom
3	(o - 1)S_o^2 / θ_o is a chi squared random variable with o – 1 degrees of freedom
4	(p - 1)(o - 1)S_{po}^2 / θ_{po} is a chi squared random variable with (p – 1)(o – 1) degrees of freedom
5	po(r - 1)S_E^2 / θ_E is a chi squared random variable with po(r – 1) degrees of freedom
6	$\bar{Y}_{***}$ is a normal random variable with a mean μ_Y and variance $\frac{\theta_p + \theta_o - \theta_{po}}{por}$

Covariance structure for the balanced two-factor crossed random model [67]

Table 5-7: Covariance structure for two-way model: 5-4

Condition	Covariance ($Y_{ijk}, Y_{i'j'k'}$)
i = *i´*, *j* = *j´*, *k* ≠ *k´*, (same part and same operator)	$\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2$
i = *i´*, *j* ≠ *j´* (same part with different operator)	σ_P^2
i ≠ *i´*, *j* = *j´* (same operator with different parts)	σ_O^2
i ≠ *i´*, *j* ≠ *j´* (different parts and operators)	0

Gauge R&R parameters and point estimators for the balanced two-factor crossed random model with interaction are defined in Burdick et al [67] .

Note: definition of the measurement system (Gauge) two-way model with interaction. Variation of the measurement system is due to all sources except parts defined as:

$$\gamma_M = \sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2 \quad (5\text{-}5)$$

Table 5-8: Point estimators for balanced two-factor crossed random model with interaction: 5-4

Gauge R&R Notation	Model Representation	Point Estimator
μ_Y	μ_Y	$\hat{\mu}_Y$
γ_P	σ_P^2	$\hat{\gamma}_P = \frac{S_P^2 - S_{PO}^2}{or}$
γ_M	$\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2$	$\hat{\gamma}_M = \frac{S_O^2 + (p-1)S_{PO}^2 + p(r-1)S_E^2}{pr}$
γ_R	$\frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}$	$\hat{\gamma}_R = \frac{\hat{\gamma}_P}{\hat{\gamma}_M}$

5.2.3.1 Point estimates

Inferential statistics point to descriptions of total populations using smaller samples. Ideally the statistic calculated from the sample is a reasonable approximation of the population parameter. The point estimates detailed above are based on a reasonable assumption that for example given a sample, $\bar{x}$ (sample mean) is a reasonable estimate for the μ (population mean). Normally one would like to say something about the confidence with which this statement is made. This will be covered in more detail in Chapter 11.

The assumption that the model assumes random effects is subtle in its distinction and is worth noting. It is relatively straightforward to expect the parts to be identified as a random sample, as there are likely to be many parts available. This is less apparent with respect to operators, because they may be the only ones available and there is a logical argument they should be a fixed effect. A useful explanation is offered by Montgomery et al [68]. The assumption is that the operators are representative and also are treated as a randomly selected sample from a population. If the operators are the sum total to be used, the inference concerns only these individuals and there is no wider population They are classed as “fixed” and the experiment defined as a balanced two-factor crossed mixed model.

5.2.3.2 Interaction

Interaction is a joint factor effect. The factors in this instance, for a GRR study, are operator and part, where the effect on one factor depends on the other. When interactions occur the factors involved cannot be evaluated individually. As an example an interaction occurs where a drug is given in combination with or shortly after another drug. This alters the effect of one or both drugs. In this research study it may occur where a room type, shape or size means that an operator sets up the speaker and microphone positions in a fixed way, which is different from that of another operator. It affects the measurement and applies to some parts and not to others. One of the challenges of the DOE is to arrive at a test sample which minimises the risk of this occurring.

The knowledge that there is interaction between factors can be more useful than knowledge of the main effect itself and it is strength of the ANOVA DOE that these effects can be included in the assessment. Presently, they are undetectable using the methods in BS5725 and GUM. Indeed if there is significant interaction between factors, increasing the sample size in a traditional study will not improve the analysis results, as might be expected. Instead the size of the reproducibility will be underestimated [68].

The definition of reproducibility in the GRR is covered in Burdick et al and incorporates the interaction term and is:

$$\sigma^2_{reproducibility} = \sigma^2_O + \sigma^2_{PO} \quad \text{5-6}$$

It is noted by Montgomery et al [68] that in certain circumstances an anomaly can occur with the interaction term which is possibly the only drawback of this method. If the repeatability mean square is larger than the mean square for operator by part interaction, a negative estimate of σ^2_{op} is obtained when the repeatability is subtracted from the operator by part interaction. This is clearly nonsense in statistical terms (for which this method has attracted some criticism) so to circumvent this problem, a test is conducted on the significance of the interaction term (p-Value). If the operator by part interaction is not statistically significant common practice dictates that the variance component for the interaction is set to zero thus avoiding a negative variance term. Searle [112] argues that this will result in biased estimates of the other model parameters. This may be, but it is realised that the estimated operator variance component is set to zero because the error mean square is larger than the operator mean square. This will occur in studies where the measurement variation is primarily due to repeatability. The importance of this is reduced because, as seen in the international standard estimates of r and R [8], it is relatively uncommon to find r > R in acoustic measurement and, consequently, as repeatability is not expected to be dominant in this research, its potential impact is reduced.

5.3 Conclusion

This chapter has looked at the basic ANOVA model and described how a specialist technique has evolved due to the demands of industry for detailed information on variability of measurement systems and its effect on total variability.

Measurement system analysis has evolved into a specialist area of statistics and has developed gauge repeatability and reproducibility techniques and models where there is a desire to consider not only the measurement uncertainty but also the contributions of the factors of operators and the parts together with their interaction. The ability to assess the interaction terms is a key strength of this ANOVA technique over traditional inter-laboratory methods. It is expected that it will provide additional information without compromising or adding bias to the other variance components. This technique also has advantages over GUM, in identifying and quantifying the main components of variance in the measurement process, especially when blocking specific input variables. Empirical rigour replaces the intuitive quantification of input variables selected from the long list of potential influences. In this research empirical testing can be used to great advantage when comparing the influence of the parts (separating floors), and in particular the construction of the parts (timber/concrete), on the variability in the measurement process.

Careful note has been made of the DOE requirements of GRR in order to optimise the value of the results. The appropriate number of operator, part and replicates has been identified and the assumptions for the model stated.

The next chapter will look at implementing the GRR study for testing sound insulation in the field. The sampling routine will cover the field measurement of the most common types of floor construction and, in order to fulfil the linearity requirements, identify additional test elements which are likely to present performances at the extreme end of the measurement process.

6 Design of Experiment - GRR

6.1 Introduction

In chapter 5 the ANOVA model for a GRR was defined. This enables a clear statement of objectives for the study: determining the variance from instrumentation, operator and part. The approach is based on that of Montgomery [108]. Some of the particular difficulties with respect to large scale field testing of sound insulation are identified and solutions proposed.

The field survey plan, method of measurement, data collection and choice of factors, levels and ranges are described.

6.2 Experimental format

The balanced two-factor crossed random model, with interaction, was the selected GRR for this survey. In essence, a number of operators measure a chosen number of parts a number of times and everyone measures the same parts. There is potential, with careful design, to extract the variability in the measurement process, and importantly the variability due to the part itself.

This section uses the experimental design list proposed by Montgomery [108] as a basis. The approach involves blocking of certain factors, to constrain the specimens measured in order to standardise the results and improve data quality for evaluation by reducing nuisance factors e.g. making measurements over a short duration to reduce the effects of temperature and barometric pressure on the results.

It also includes a brief description of the measurement process undertaken by the operators, their qualifications, procedures, gauges, equipment and software, used to assign a value to the measurand of interest [113].

6.2.1 Choice of test sample construction

In order to carry out the GRR study and ensure that maximum information is collected about the variability of measurement in the field, it was decided to focus on the two main types of separating floor construction common in England and Wales. This was because it was not clear, a priori, if and how, different sound insulation performances affect the statistical behaviour.

One test site had a lightweight timber separating floor and the other a heavyweight concrete construction. They represented the most popular types of construction in residential dwellings. Each test site formed a separate GRR study. It is noted that in measurement system analysis it is usually the 'gauge' or measurement system that is the focus of attention, rather than the part. This was still the case, however, as information on typical part variability in the construction industry is sparse this is a desirable by- product of the GRR experiment.

6.2.1.1 Construction

A heavyweight concrete floor and a lightweight timber floor were selected for study. The sites for each floor type were selected because they employed floor designs that were classed as "Robust Details" [114]. Robust Details were introduced in England and Wales as a means of avoiding the pre-completion testing of sound insulation at the end of a project. The reason for choosing sites which employed a Robust Detail floor construction was because each floor design is prescriptive and the construction would be known in detail and because of the additional quality control responsibilities placed on the builder by the Robust Detail scheme, workmanship was expected to be reasonable.

The timber floor was a robust detail reference E-FT-3, the concrete floor was a robust detail reference E-FC-5. The basic construction details are listed in and illustrated in the Robust Details section drawings. The concrete and timber floor constructions are detailed in Table 6-1 and are represented in the section drawings in Figure 6-1.

Table 6-1: Test floor construction descriptions E-FT-3 (Timber) & E-FC-5 (Concrete) [114]

Component	Timber (E-FT-3)	Concrete (E- FC-5)
Floating Floor	18 mm (min) t&g flooring board gypsum-based board nominal 13.5 kg/m2 FFT-1™ resilient composite deep battens battens may have the resilient layer at the top or the bottom mineral wool quilt laid between battens – 25mm (min) 10-36 kg/m3	65mm (min) cement: sand or 40mm (min) proprietary screed of nominal 80 kg/m2 mass per unit area YELOfon® HD10+ with E-strip perimeter edging and J-strip tape for jointing
Structural	Floor Deck - 18mm thick (min) wood based board, density min 600 kg/m3. 253mm (min) metal web joists	- 150mm (min) precast concrete floor plank - 300 kg/m2 (min) mass per unit area
Ceiling	CT1 Ceiling – Two layers of gypsum-based board, composed of 19mm (nominal 13.5 kg/m2) fixed with 32mm screws, and 12.5mm (nominal 10 kg/m2) fixed with 42 mm screws Resilient Bar - 16mm (min) metal resilient ceiling bars mounted at right angles to the joists at 400mm centres (bars must achieve a minimum laboratory performance of rd Rw+Ctr=17dB and rd Lw=16dB) 100mm (min) mineral wool quilt insulation (10–36 kg/m3) between joists	CT0 – Metal ceiling system - 150mm void To be used for 150mm (min) depth concrete planks • any metal ceiling system providing 150mm (min) ceiling void • one layer of nominal 8 kg/m2 gypsum-based board

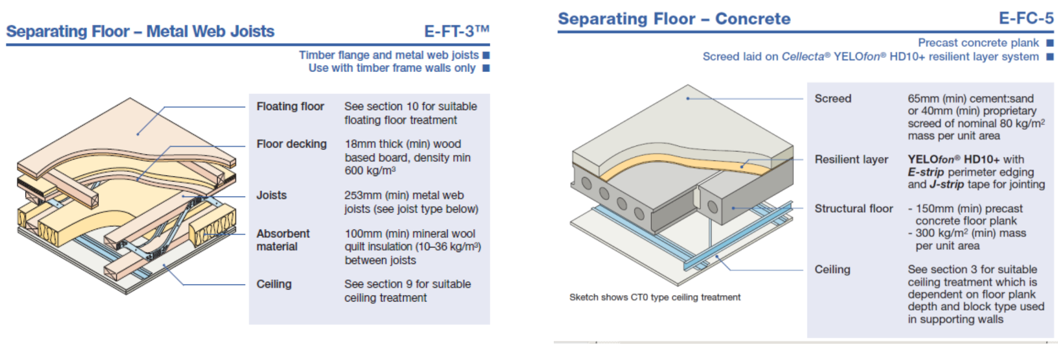


Figure 6-1: E-FT-3 (Timber) & E-FC-5 (Concrete) Floor Constructions[114].

The proposed test specimens represent common "lightweight" and "heavyweight" construction. These two generic construction types form the majority of floors constructed in new build apartments today and have different frequency based sound insulation performance over the measurement range considered in this thesis from 100 to 3150Hz.

In general the heavyweight construction performs better at low frequency, then the sound insulation performance for both floor types increases with frequency to 1KHz, above 1KHz the lightweight floor performs significantly better with the heavyweight floor performance falling with frequency from 1kHz towards 3.15kHz. The average frequency performance comparison is shown in Figure 6-2.

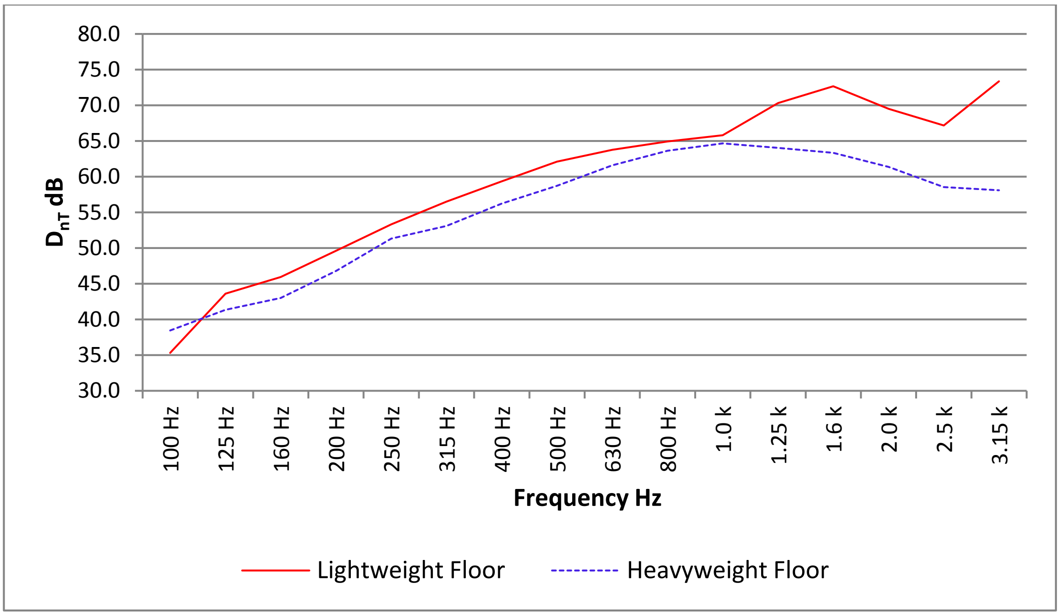


Figure 6-2: Lightweight timber v heavyweight concrete floor D_{nT} values between 100-3150Hz.

6.2.1.2 Level and range

In addition to the frequency behaviour, selecting a random test sample from the standard residential housing market constructions will only provide airborne sound insulation performance centring on an average performance of approximately 50dB $D_{nT,w} + C_{tr}$ with a predicted range of 45 – 55dB $D_{nT,w} + C_{tr}$.

According to Montgomery [68], if the measurement system capability is to be tested properly it should have a test sample which spans the full range of values likely in the field To accommodate this, four additional test elements were selected, at the concrete floor site, to significantly increase the range of the test sample. Their airborne sound insulation was predicted to be between 30 – 60dB $D_{nT,w} + C_{tr}$.

6.2.1.3 Blocking

To minimise unwanted factors and attempt to standardise results by 'part', the test samples were "blocked" with respect to meteorological conditions (e.g. temperature, humidity and barometric pressure) and where possible room size.

The metrological conditions were blocked by ensuring that all the testing took place over a short time period nominally two to three days. This meant the weather showed minimal variation and internal conditions in the test rooms were within ±4C and ±20%RH for the duration of the experiment.

In addition the room size was fixed on the timber floor site (33m^3). The reasons for this were; firstly, by fixing the room sizes, the variability of the performance of the floor cannot be due to dimensional changes or differences in room volume. It is then due to the variability of the onsite construction so by blocking room size the part to part variation will reflect that provided by the method of construction. See the influence of "workmanship" by Craik et al [57, 58]. Additionally if the room size chosen is similar to Craik et al a comparison can be made with their results; secondly if similar sized rooms could be identified on both GRR study sites, it would hopefully minimise the impact of variability in sound pressure levels in rooms especially at low frequency due to room dimensions identified by Maluski et al [115, 116]. By standardising the room sizes, there could be a direct comparison with the variability associated with the construction of the floors; thirdly, a small room provides a non-diffuse field at low frequencies over a significant part of the frequency range. The predictably high level of measurement variability expected at low frequencies associated with sound pressure level measurement in a non-diffuse field offers an opportunity to assess the proportion of variability due to the part, operator or instrument. A GRR assessment of the smallest regularly sampled room would then provide a "worst case" estimate for the measurement uncertainty in field testing of airborne sound insulation. Unfortunately it was not possible to fully control the room size factor in the concrete GRR as a suitable site was not available. Small room sizes were used but they were not identical and the volumes ranged from 19 – 25m^3.

6.2.2 Selection and measurement of the response variable & test method

6.2.2.1 Response variable

The response variable in this case is the airborne sound insulation value, otherwise known in metrology as the "measurand". The International standard test procedure for airborne sound insulation measurement and calculation[17, 22] clearly demonstrates that the response variable for sound insulation is not directly measurable and relies on several factors: calibration of the instrumentation, a steady, flat response sound source level, several measurements to be taken in each source and receive room, averaging of room levels (ISO standard requires logarithmic averaging [17], Approved Document E 2003 [2] of the Building Regulations requires arithmetic averaging. In this study arithmetic averaging was carried out in line with Building Regulation requirements for England and Wales) and corrections for both reverberation time and where appropriate background noise were applied.

The Building Regulation airborne performance standard for new build separating floors is (45dB $D_{nT,w}$ + C_{tr}) and is the relevant single figure value that will be used to assess the uncertainty in sound insulation, where:

$D_{nT,w}$ is the weighted standardised level difference. See Table 2 in [90] and:

C_{tr} is the spectrum adaptation term calculated using Spectrum No2 from [90] for urban traffic noise.

In the statistical assessment the measured value (Measurand) whether it is $D_{nT,w}$ or $D_{nT,w}$ + C_{tr} is known as the "response variable".

This study extends the measurement uncertainty evaluation beyond just the single figure value to incorporate the full measured frequency range (100Hz to 3.15KHz) using the standardised level difference (D_{nT}) [90] in each of the 16 third octave bands. An assessment of the commonly used single figure airborne value [3] $D_{nT,w}$ will also be carried out.

6.2.2.2 Test Method: Airborne Sound Insulation Test

In order that the response variable is accurately represented, all operators need to be trained to perform the required test measurement in accordance with a standard test procedure. The sound insulation test method used follows the requirements of the international standard EN ISO 140-4: 1998 [22] and uses the single static microphone method. This is preferred to the moving microphone method as the individual measurements made spatially across the room provide additional information on the measurement uncertainty present in the sound field.

In this study all operators were United Kingdom Accreditation Service (UKAS) trained and accredited and the instrumentation has traceable UKAS calibration. The test procedures were those applied by UKAS Testing Laboratory No2694 and were fully compliant with the requirements of the International Standards [17, 22]. The "Airborne Sound Insulation Test Procedure"[117] is in the Appendix.

6.2.2.3 Choice of GRR experimental design

Five UKAS accredited operators were selected as the maximum possible due to equipment and operator availability. Under normal GRR studies, all operators use the same gauge. This study, aimed to replicate typical field tests, where different operators, using their own equipment, would measure the same sample of parts. Each operator used his own test kit, and this introduced the additional variability. The 5 test kits were the same make and include sound level meters, calibrators, loudspeakers and wireless transmitters/receivers. A list of the typical sound insulation test kit is detailed in the Appendix in 14.2.

The selection of 6 floor elements, with three repetitions, was set by the time constraints on site. Careful consideration was given to the length of time it would take each operator to complete one test (nominally estimated at between 20 – 40 minutes) and it was considered that 6 was the maximum number of parts that could be tested by all 5 operators in a day on the same site.

For reasons previously discussed the testing was scheduled over a relatively short time period, nominally 2 - 3 consecutive days, to minimise the influence of metrological conditions on the tests. The number of operators, test sample and repetitions complied closely with the recommended robust GRR design suggested by Burdick et al [67].

The test designs are summarised in Table 6-2, Table 6-3 & Table 6-4

Table 6-2: Testing Schedule - Lightweight Timber Floor

Test Site: Timber Floor	Separating Element Floor: Timber	Floor Type : E-FT-3
Operators	Parts	Repetitions
5	6	3

Table 6-3: Testing Schedule - Heavyweight Concrete Floor Tests

Test Site: Concrete Floor	Separating Element Floor: Concrete	Floor Type : E-FC-4
Operators	Parts	Repetitions
5	6	3

Table 6-4: Testing Schedule - Linear Tests – Concrete Floor Site

Test Site: Concrete Floor	Separating Element: Various	Additional Testing - Linear
Operators	Parts	Repetitions
5	4	3

6.2.3 Preparing the test site

The test site selected for the timber floor was in the Midlands and the concrete floor site was in the South of England. The linear testing took place on the concrete floor site.

Written permission was requested from the home builders and a full list of site condition requirements were forwarded to the site agents. The sites were ready for test, with free access over the measurement period. Projects that were nearing completion were selected to minimise the potential adverse influence of building work inside the apartment blocks. The sites were essentially complete, with doors and windows fitted, but the rooms were without soft furnishings (carpets & curtains). Vacant possession was granted for both sites for all test days.

The timber floor GRR testing took place over three days on 27/4/09, 30/4/09 and 1/5/09

The concrete floor and additional construction GRR testing took place between 29/6/10 and 30/6/10.

6.2.3.1 The operators

All operators were fully briefed, prior to arrival on site, to ensure that a consistent, standard approach was followed prescriptively.

Completion of one full test set (all 5 operators testing one set of 6 floors each) took between 10 – 14 hours, depending on the availability of rooms and the speed of the testers. Breaks were taken through the day for lunch and refreshments.

Each operator arrived each morning and setup independently. They were instructed to test rooms promptly, as they became available during the day. Because each test took a different amount of time, rooms became available randomly during the day. This was intentional to develop a random sample for the operators in order to avoid bias in the measurement process. It was also highly efficient, allowing each operator who completed one test to take the next room available.

To avoid replication, where a part is measured consecutive times without changing the setup of the measurement kit, testers were instructed to test each room fully before moving on to the next, thereby ensuring that the equipment had to be removed from the test room and reset for each test. In essence replication, requiring removing and resetting the measurement equipment from the test rooms, better represents the inherent noise in the standard test process and is the way chosen to increase the number of runs in this study.

6.2.3.2 Test rooms

The basic room information for the timber GRR is in Table 6-5 for ease of reference. All the test rooms were matched pairs i.e. same size shape and volume, and all rooms stacked vertically. Room details for the five room pairs used in the concrete GRR are contained in Appendix 14.7.3. The room pairs are tabulated with their respective details and all the rooms have volumes of between 19 – 25m^3.

Table 6-5: Test Room Information Summary

Test Site	Source/Receiver Room Dimensions	Source/Receiver room Volume
Timber GRR (6 Tests)	2.4m H x 3.05m W x 4.8m L	33m^3
Concrete GRR (6 Tests)	Varies see Room Dimensions 14.7.3	19-25m^3

See example timber floor layout plan in Figure 6-3.

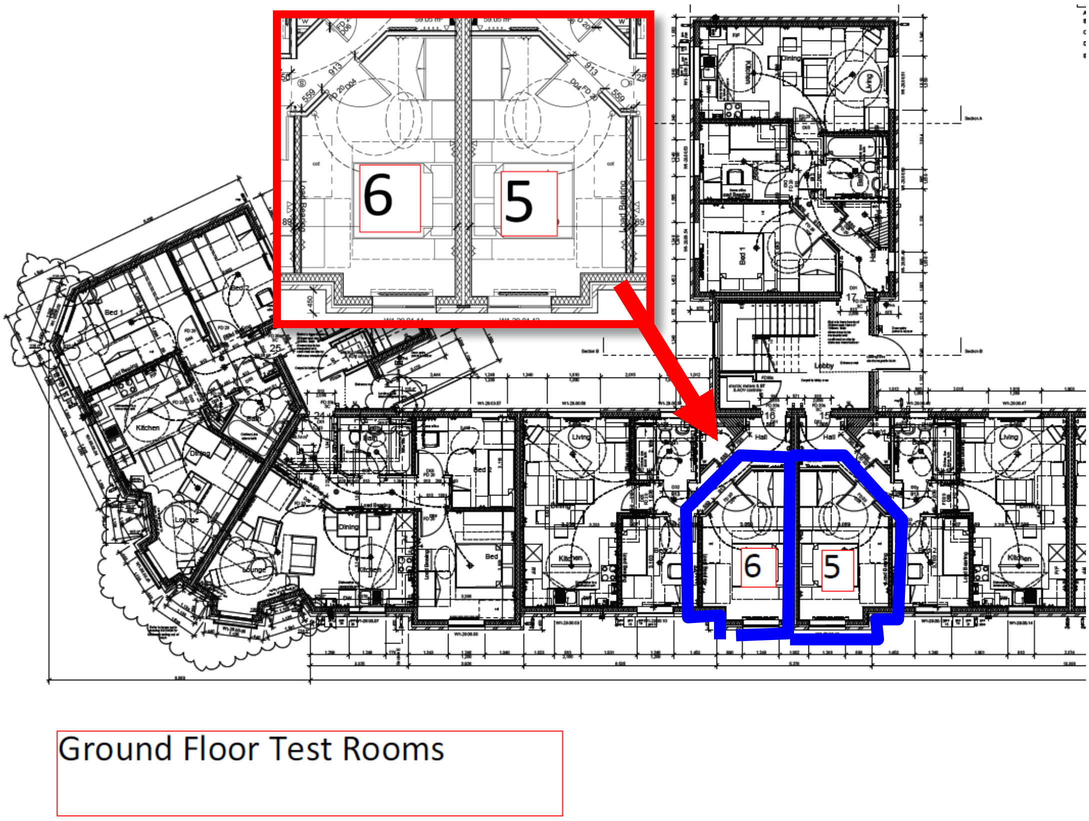


Figure 6-3: Site layout showing typical test rooms on ground floor of flats(blue) on Midlands test site with matched pairs of rooms on 1st floor above (red).

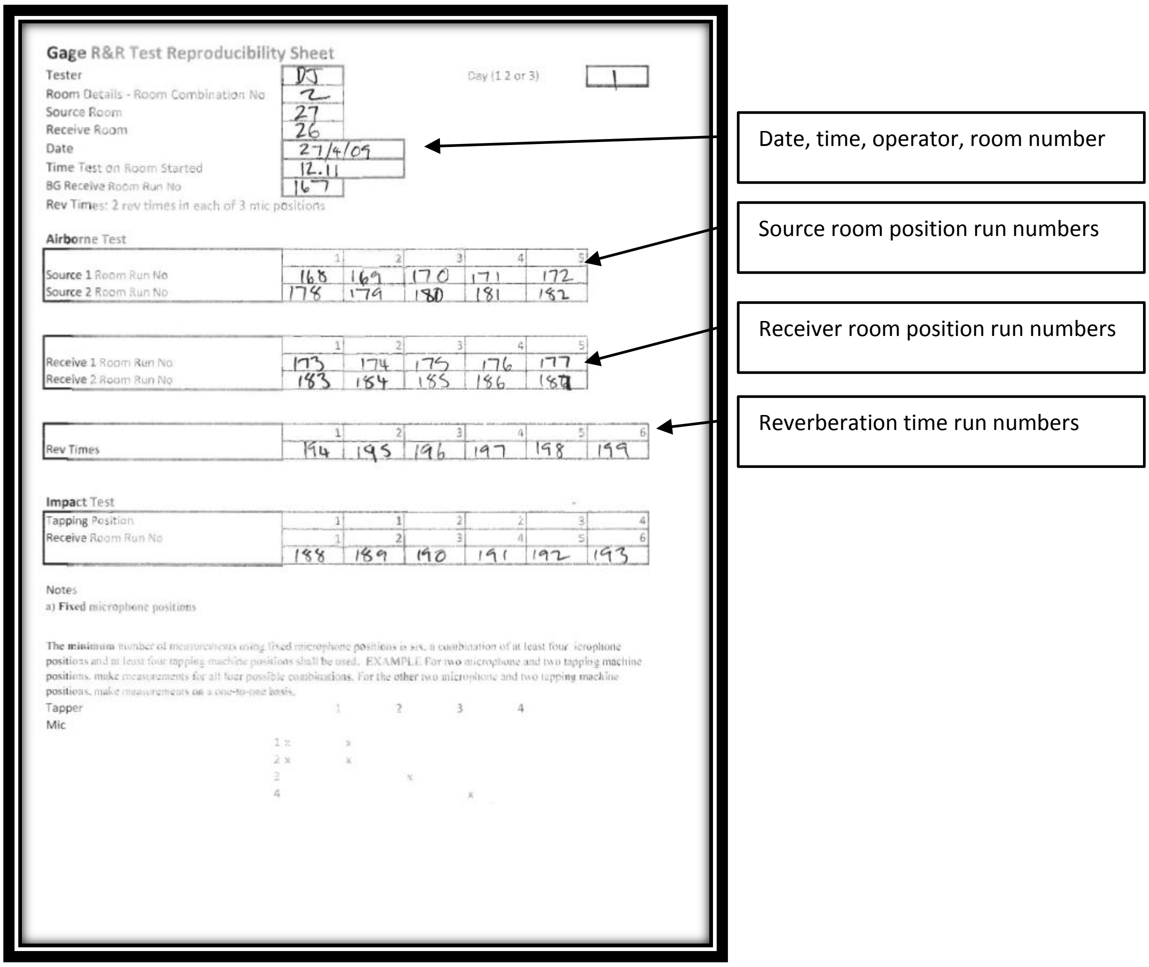
Care was taken to select room locations to avoid noise being transmitted between operators who were testing simultaneously. Due to time constraints it was also important to ensure all test pairs could be in use simultaneously.

6.2.3.3 Measurement procedure

The field test for airborne sound insulation requires several measurements to be carried out for each test that record average sound pressure levels in the source and receiver rooms, reverberation times in the receiver rooms and background sopund pressure levels. The measurement carried out by the trained operator follows a UKAS work instruction that lists the test procedure needed to fulfil the internaltional standard and Building Regulation requirements prescriptively. The work instruction is detailed in the Appendix to this document in section 14.3 .

6.2.3.4 Recording the data

Each test that was conducted was noted on a site record sheet with run numbers and room details. These were collected at the end of each day and a sample test sheet is shown in Figure 6-4 with a full size sheet provided in the appendix.



Gage R&R Test Reproducibility Sheet

Tester: DJ
Room Details - Room Combination No: 2
Source Room: 27
Receive Room: 26
Date: 27/4/09
Time Test on Room Started: 12.11
BG Receive Room Run No: 167
Rev Times: 2 rev times in each of 3 mic positions

Day (1 2 or 3): 1

Airborne Test

	1	2	3	4	5
Source 1 Room Run No	168	169	170	171	172
Source 2 Room Run No	178	179	180	181	182

	1	2	3	4	5
Receive 1 Room Run No	173	174	175	176	177
Receive 2 Room Run No	183	184	185	186	187

	1	2	3	4	5	6
Rev Times	194	195	196	197	198	199

Impact Test

Tapping Position	1	1	2	2	3	4
Receive Room Run No	1	2	3	4	5	6
	188	189	190	191	192	193

Notes
a) Fixed microphone positions

The minimum number of measurements using fixed microphone positions is six, a combination of at least four icrophone positions and at least four tapping machine positions shall be used. EXAMPLE For two microphone and two tapping machine positions, make measurements for all four possible combinations. For the other two microphone and two tapping machine positions, make measurements on a one-to-one basis.

Tapper 1 2 3 4
Mic
1 x x
2 x x
3 x
4 x

Figure 6-4: Example Site Test Record Sheet - Timber Floor Site

The sound level meters were downloaded daily to a central file store, to preserve the electronic data record. No data was analysed during the survey process, so that the testing was carried out “blind” by reducing feedback to the operators and minimising the potential for bias.

On completion, each data record was exported and saved to a single spreadsheet for each operator, prior to analysis. The files were in an excel spreadsheet format suitable for importing into a proprietary statistical software package (Minitab v15) for the ANOVA.

The data manipulation and analysis are described in more detail in chapter 7 and also listed in 14.1 Appendix 1 – GRR Data .

6.3 Conclusion

This chapter has considered the practicalities of the experimental design, with reference to recommendations from Montgomery et al , the resources available and the time limitations. The proposed GRR survey plan includes numbers of operators, parts and replicates, to provide a robust data set. It is proposed that in order to improve the reliability of the data and reduce the unwanted factors the testing will take place over a short time period, nominally 2 – 3 days so, given stable weather the influence of meteorological conditions will be minimised.

In addition, by blocking room size the part to part variation will reflect that provided by the method of construction. This information augments the expected qualitative and quantitative results for the measurement system. This standardisation of room size also allows a meaningful comparison to be made between the lightweight and heavyweight floors, and possibly identifies secondary influences of the construction. In addition, the small room size replicates the worst case condition with respect to a non-diffuse field and allows a retrospective comparison with previous work of others [58].

The next chapter describes the data analysis and preliminary results, for the GRR experiments for timber and concrete floors and additional elements measured for gauge linearity.

7 Results

7.1 Introduction

This chapter presents preliminary comparisons of the data from the GRR experiment. The comparisons and checks were designed to establish the reliability of the data and identify what influences the variability of field test data and the differences between theoretical performance, laboratory performance and that measured in the field.

To familiarise the reader with the data, results are first considered using simple descriptive statistical techniques. This allows comparisons of data sets and to obtain the overall perspective onthe mean values. Three field measurements are examined: source room sound pressure level, receiver room sound pressure level and the reverberation time, and the variability associated with each. Data from the source and receiver rooms are compared with theoretical values of sound pressure level measurement. Deviations from the theoretical values are identified, thus avoiding spurious or systematic variability.

To examine the impact of a complex, multi component design, the lightweight timber floor data is used to obtain the apparent sound reduction index, for comparison with previous results of the simple concrete floor, by Craik et al [57].

Using the calculation procedures in BS5725 the timber floor and concrete floor data are analysed with respect to current international standards for repeatability and reproducibility; again anomalies are highlighted and the timber and concrete floor samples compared.

7.2 Initial Data Analysis

It is useful in the first instance to consider basic descriptive statistical analyses in order to provide an overview of the differences in floor performance overall and then the variability across the frequency range.

7.2.1 Single Figure Values

The Timber floor and concrete floor test single figure sound insulations (90 tests in total) are summarised in Table 7-1 & Table 7-2 with respect to the current Building Regulations[2] for airborne sound insulation performance $D_{nT,w} + C_{tr}$:

Table 7-1: Mean and standard deviation Single figure descriptors for timber floor tests

Timber	**$D_{nT,w}$**	**$D_{nT,w} + C_{tr}$**
Mean	62.6	53.2
s.d.	1.0	2.3

Table 7-2: Mean and standard deviation Single figure descriptors for concrete floor tests

Concrete	**$D_{nT,w}$**	**$D_{nT,w} + C_{tr}$**
Mean	59.0	53.0
s.d.	1.9	1.5

The arithmetic averages of the $D_{nT,w} + C_{tr}$ value of the concrete floors and timber floors are within 0.2dB. However the standard deviation is notably higher for timber floors than for the concrete floors. This was despite there being some variability in the room sizes for the concrete floor tests which would have normally been expected to contribute additional variability to the single figure value.

The reverse is true for the $D_{nT,w}$, where the standard deviation of the timber floor is significantly lower than that of the concrete element. There is a direct mathematical relationship between the apparent sound reduction index of the floor, the field test single figure value $D_{nT,w}$ and the room size. The lower variability in the timber floor tests for $D_{nT,w}$ could be partly due to the fact that the room size was fixed.

The spectrum adaptation term C_{tr} is based on a weighted low frequency performance with the resultant applied to the $D_{nT,w}$ in order to get the value of the correction term (a negative number).

The likely cause of the higher variability in $D_{nT,w} + C_{tr}$ for the timber floor case stems from the low frequency performance affecting the spectrum adaptation term. These weightings at low frequency increase sensitivity in measured sound insulation at 100Hz

– 160Hz. This is where the performance of the timber floors are likely to be lower than for concrete, and result in a greater negative spectrum adaptation term.

The relationship between the spectrum adaptation term and the single figure value $D_{nT,w}$ is further evidence that GUM [6] is an inappropriate method for estimating uncertainty in sound insulation measurement.

This is because it is not possible to treat these two component parts of the single figure value as individual inputs for addition in *quadrature*, (i.e. squared, summed and the square root taken) because they are not independent. In addition, there is indirect evidence of interaction between the two components $D_{nT,w}$ and C_{tr}. The C_{tr} term is not the same magnitude for every $D_{nT,w}$ value (again due to the variability in the low frequency performance of the separating elements tested). If it were, this would provide a single number offset for each $D_{nT,w}$ and would not affect the magnitude of the standard deviation. The C_{tr} in the timber floor tests interacts with the single value in a way that gives more variability for the combined value $D_{nT,w}$ + C_{tr} e.g. the C_{tr} term is bigger for lower values of $D_{nT,w}$ and smaller for higher values, leading to a greater spread of results about the mean. This observation suggests that there is interaction between these measurand components and the independence requirement of GUM is violated.

7.2.2 Frequency Data

The third octave band frequency data for both the timber and concrete GRR experiments can be compared in the same way as for the single figure values. The mean sound insulation for the third octave bands 100 – 3150Hz is detailed in the Table 7-3 & Table 7-4 for *D* and D_{nT} respectively. The difference between the two floor types is also shown.

Table 7-3: Level Difference D – (90 test sample): difference between timber & concrete

D dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
Timber (Mean)	32.5	39.7	41.9	44.8	47.8	50.6	53.1	56.1	58.3	59.9	61.1	65.5	67.8	65.3	63.2	68.8
Concrete (Mean)	37.1	41.4	42.2	45.8	50.2	53.3	56.9	59.0	61.0	62.8	63.8	63.3	62.8	61.4	58.8	58.0
Difference	-4.7	-1.7	-0.3	-1.0	-2.4	-2.7	-3.8	-2.9	-2.7	-2.9	-2.8	2.1	5.0	3.9	4.4	10.8

The data shown is a simple subtraction of source and receiver room sound pressure levels, after correction for background noise, without correction for reverberation time.

Using this measure, the concrete floor gives greater sound insulation across the frequency range, until 1.25 KHz when the performance of the timber floor is greater.

Overall, see Figure 7-1, the low frequency performance of the concrete floor is slightly better than the timber floor. It is of similar spectrum shape between 125Hz – 1KHz. Above 1 KHz the timber floor performance is greater.

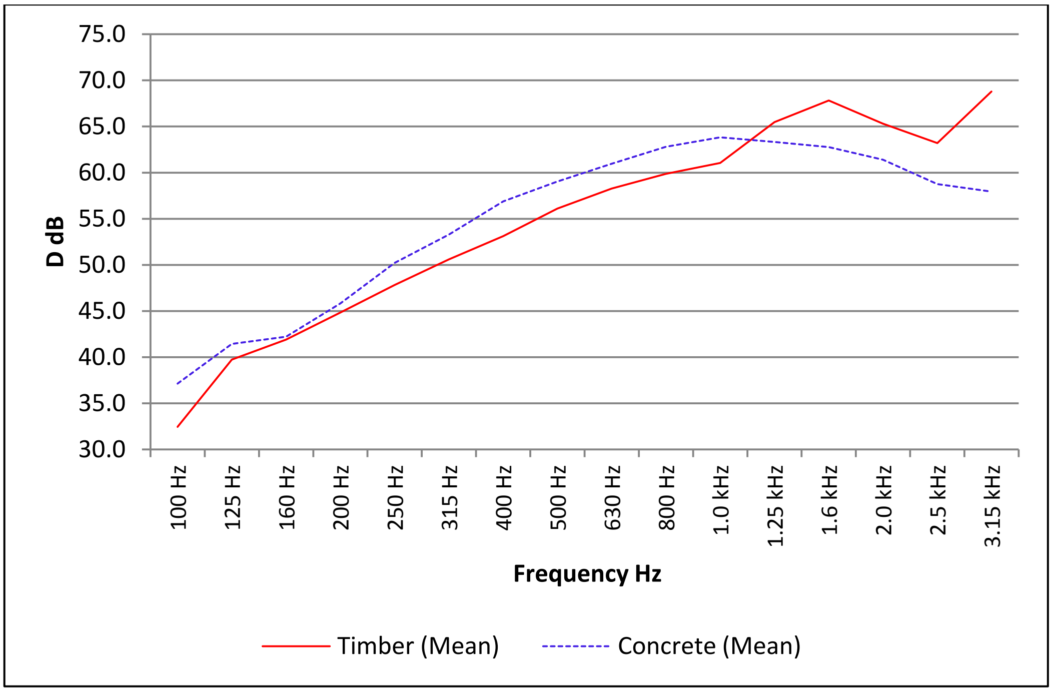


Figure 7-1: Mean Difference values timber and concrete floor samples 100-3150Hz

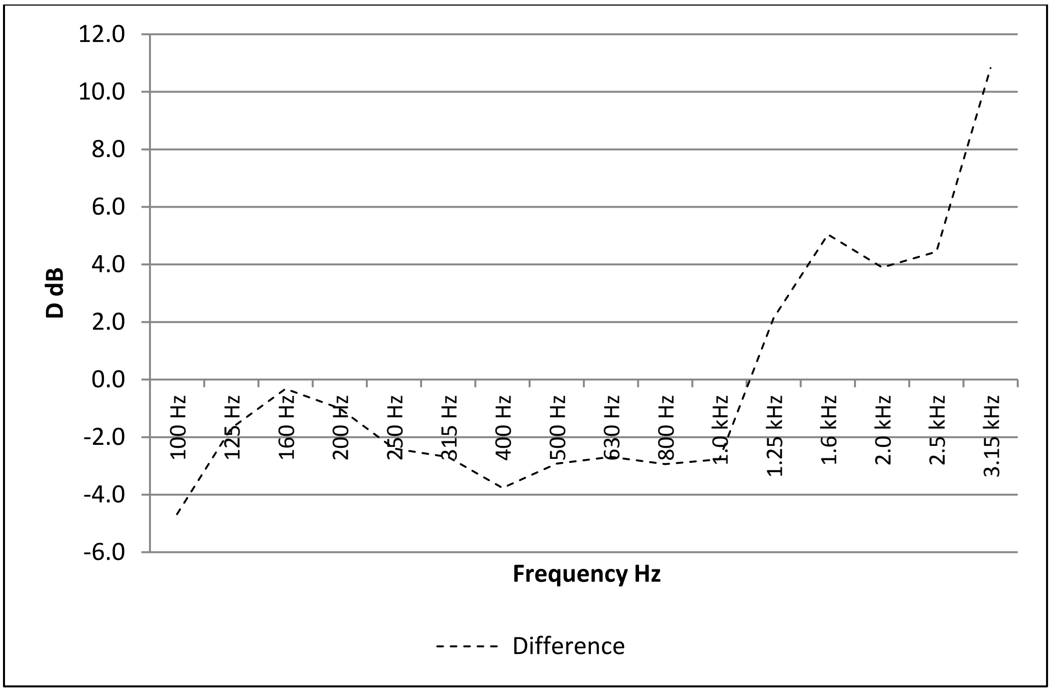


Figure 7-2: Difference in D levels between timber and concrete floor samples 100-3150Hz

Table 7-4 includes the reverberation time correction.

Table 7-4: Standardised Level Difference D_{nT} – (90 test sample): difference between timber & concrete

DnT dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
Timber (Mean)	35.3	43.6	46.0	49.6	53.3	56.5	59.4	62.1	63.8	64.9	65.8	70.3	72.7	69.5	67.2	73.4
Concrete (Mean)	38.5	41.4	43.0	46.8	51.3	53.1	56.2	58.7	61.6	63.6	64.7	64.0	63.3	61.4	58.5	58.1
Difference	-3.1	2.3	2.9	2.8	2.0	3.4	3.1	3.4	2.2	1.3	1.1	6.3	9.3	8.1	8.6	15.3

Here the timber separating floor performs slightly better than the concrete floor across the frequency range and significantly better from 1.25 kHz to 3.15 kHz. The sound insulation of the timber floor is relatively good at high frequencies and the concrete floor is only better at 100Hz. The D_{nT} data set replicates the full calculation procedure of the international standard sound insulation test and will form the basic data set for advanced ANOVA in the GRR study. The third octave band D_{nT} values are graphically illustrated in Figure 7-3.

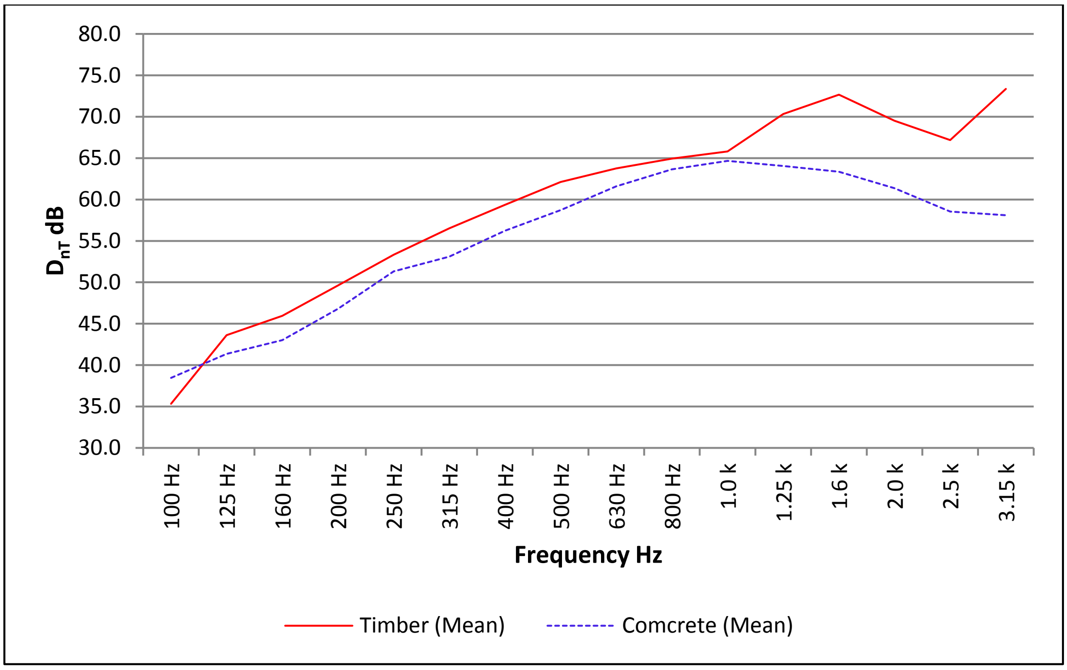


Figure 7-3: Mean D_{nT} values timber and concrete floor samples 100-3150Hz

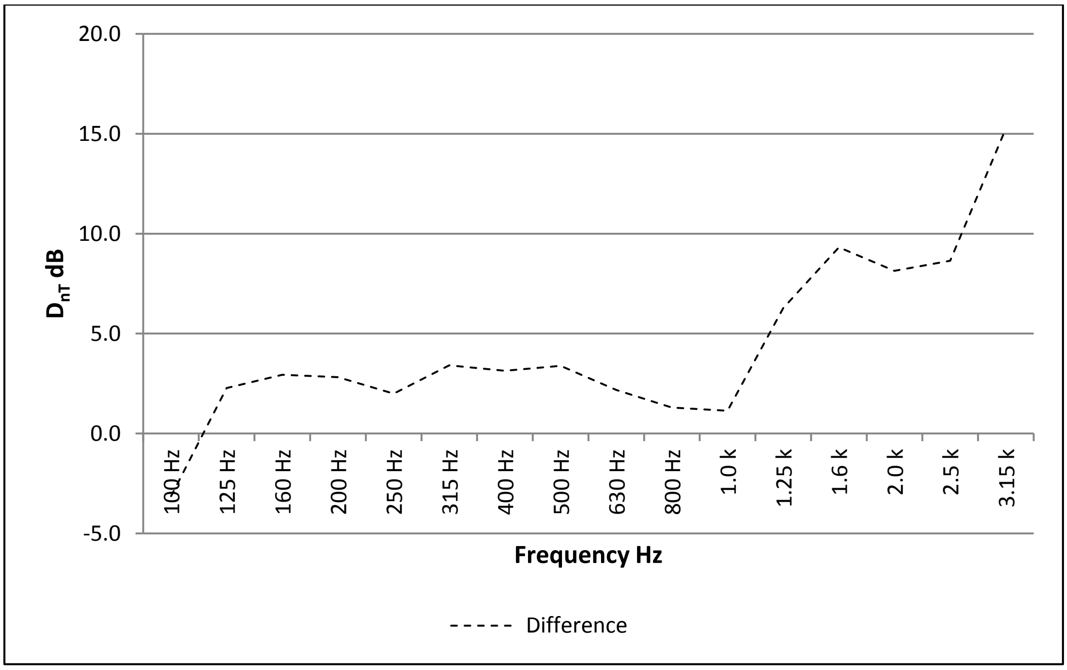


Figure 7-4: Difference in D_{nT} values between timber and concrete floor samples 100-3150Hz.

7.2.3 Reverberation time

The reverberation time is defined as the time it takes sound to decay by 60dB [118]

The quantification of reverberation time is part of the measurement process as it influences the sound field in the receiver room and therefore can affect the measured sound pressure level and impact on the overall measurement uncertainty of the measurement process.

It is the sound insulation of the separating element which is key not the contribution resulting from the reverberation of sound in the room. The calculation process requires an allowance be made for the room’s effect on the receiver room sound pressure level to ensure that the sound insulation of the separating element remains, as far as practicable, a function of the construction and not the room in which the measurements are made. For this reason a correction term is applied to normalise the effect based on a standard reverberation time in residential rooms, taken as 0.5 seconds. It is helpful to review the reverberation time data and the variability associated with its measurement.

It is the intention to assess measurement uncertainty of the sound pressure levels corrected for reverberation time (D_{nT}), it is useful to examine the variation in reverberation time (measured in seconds) and its impact on the correction term. As the experiments were designed to take place over a relatively short period, i.e. such that the internal room environmental factors are relatively stable (temperature, humidity, barometric pressure), any changes in reverberation time are expected to be due to measurement and part variability in the rooms sampled.

A comparison of the pooled mean reverberation times for the lightweight and heavyweight floors across the frequency range 100Hz – 3.15 kHz is given in Figure 7-5 with error bars, which represent one standard deviation about the mean. To highlight the variability associated with the measurement of the reverberation time the reverberation times are compared between measurements taken in the same rooms as well as reverberation times taken within the room. the standard deviation of reverberation time measurement is calculated in each case and error bars are fixed on the mean values. The data show that the reverberation time varies more between rooms than it does within rooms. This is shown in Figure 7-6 and as expected shows that the within room standard deviations are lower across the frequency range than the between room standard deviations for both floor constructions.

In these experiments the lightweight timber floor receiver rooms offer a much higher reverberation time across the 100Hz – 3150Hz frequency bands (between 1.0 – 2.0 seconds) compared to the heavy concrete floor construction receiver rooms (0.4 - 0.6 seconds). The standard deviation of the timber floor room reverberation time measurements measured in seconds is also significantly higher than the concrete floor rooms.

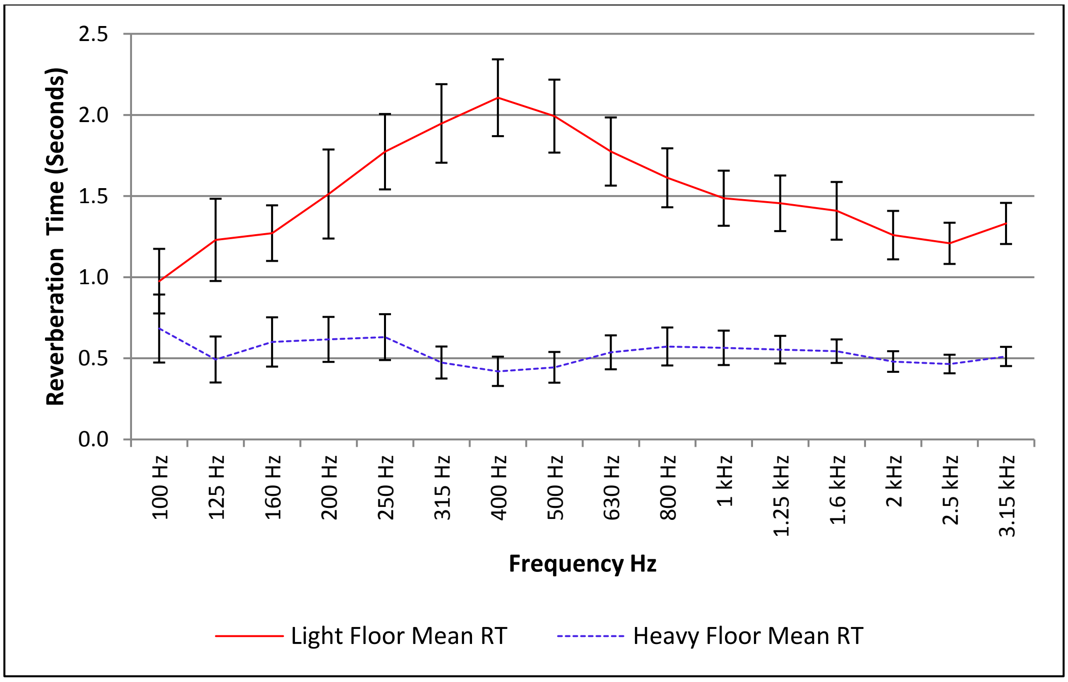


Figure 7-5: Mean Reverberation Time Measurements - Heavy v Light Floor Constructions with error bars to show standard deviation

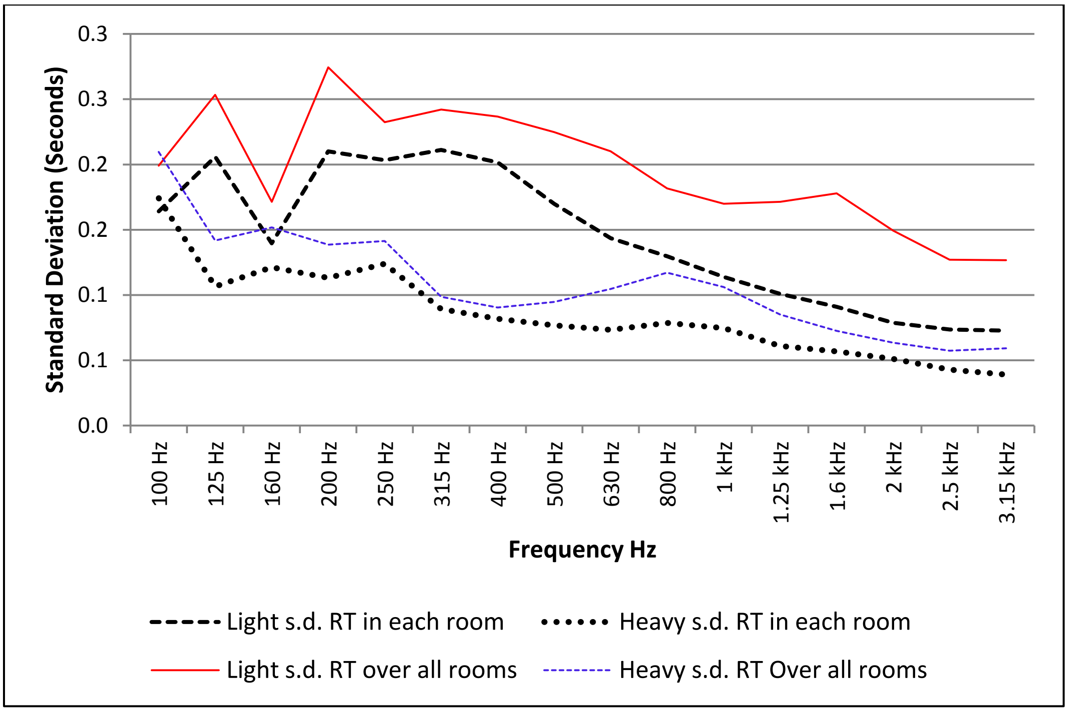


Figure 7-6: Heavy v Light Floor comparison - total test of all rooms and within room standard deviation.

The reverberation time correction is in decibels, the uncertainty of which can be quantified. The mean correction terms and their standard deviation are detailed in Table 7-5 and the standard deviations shown graphically in Figure 7-7:

Table 7-5: Reverberation time correction mean values for Lightweight Timber floor and Heavyweight Concrete Floor (dB)

$10Lg(T/T_0)dB$	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
Mean Correction Light	2.8	3.8	4.0	4.7	5.5	5.9	6.2	6.0	5.5	5.1	4.7	4.6	4.5	4.0	3.8	4.2
Mean Correction Heavy	1.1	- 0.3	0.7	0.8	0.9	- 0.3	- 0.9	- 0.6	0.2	0.5	0.4	0.4	0.3	- 0.2	- 0.4	0.1
Light correction(s.d.)	0.9	1.0	0.6	0.8	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.4
Heavy Correction(s.d.)	1.4	1.4	1.2	1.0	1.0	0.9	1.0	1.0	0.9	0.9	0.8	0.7	0.6	0.6	0.5	0.5

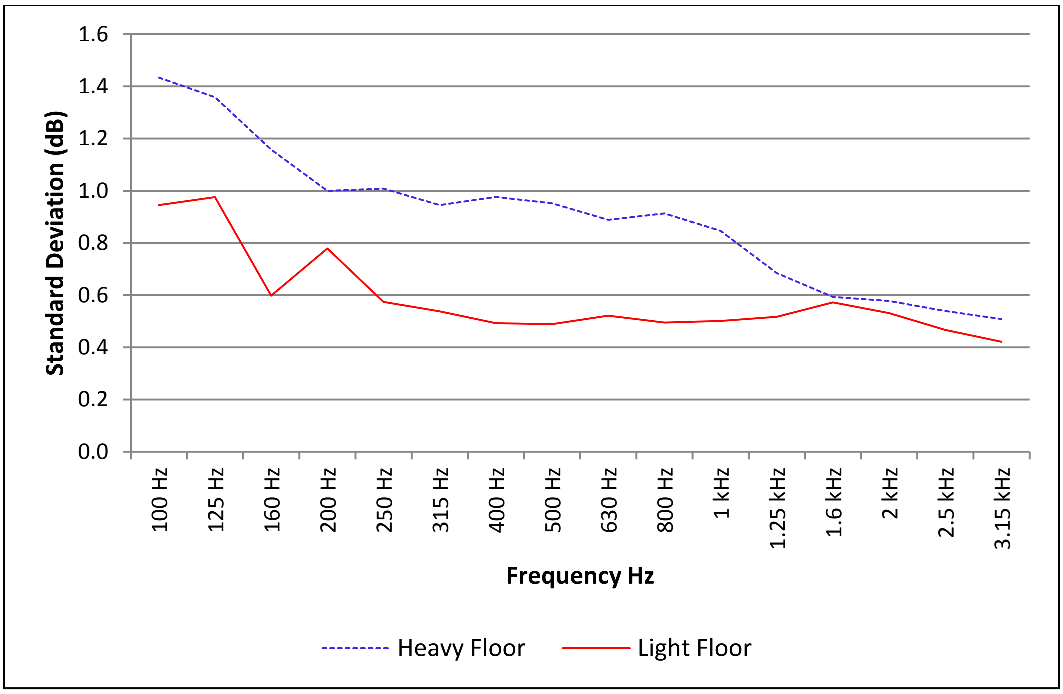


Figure 7-7: Reverberation Time Correction Applied to the RT measurements in seconds to determine standard deviation in dB.

The standard deviations of the correction term for the concrete floor tests are higher than those for the timber floor, across the full frequency range. This is not what would have been expected, given that the correction term in decibels is counter to when standard deviations are in seconds.

The explanation for this apparent anomaly is because the transformation is logarithmic and it normalises to 0.5 seconds. The transformation can result in negative corrections in some third octave bands. This affects the pooled mean value of the transformed data and also the standard deviation of the sample around the mean, which is the case here.

This demonstrates the difficulty in intuitively estimating the impact of the reverberation time input variable on the overall uncertainty budget and it reinforces the need to be inclusive of all factors in uncertainty calculations, before proceeding with the assessment, rather than treating them as individual input variables to be independently estimated and summed. This is another reason why the GUM [6] method and also M3003 [76] cannot be used for determining the uncertainty, since it relies on combining identifiable uncertainties from input variables, this would be impossible to achieve for each of the frequencies across all third octave bands.

The overall importance of the reverberation time correction variability can be easily assessed when compared with the source and receiver sound pressure levels variability. This is shown for both lightweight and heavyweight floors, in Figure 7-8 & Figure 7-9 respectively.

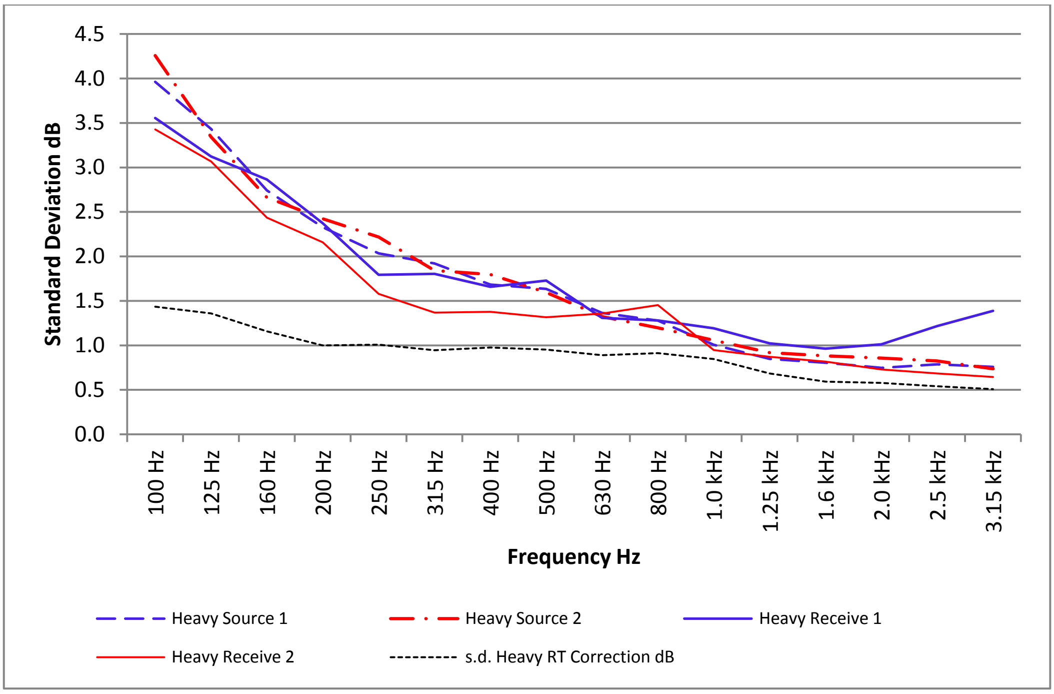


Figure 7-8: Heavy concrete floor - standard deviation values, source room Lp, receive room Lp and RT

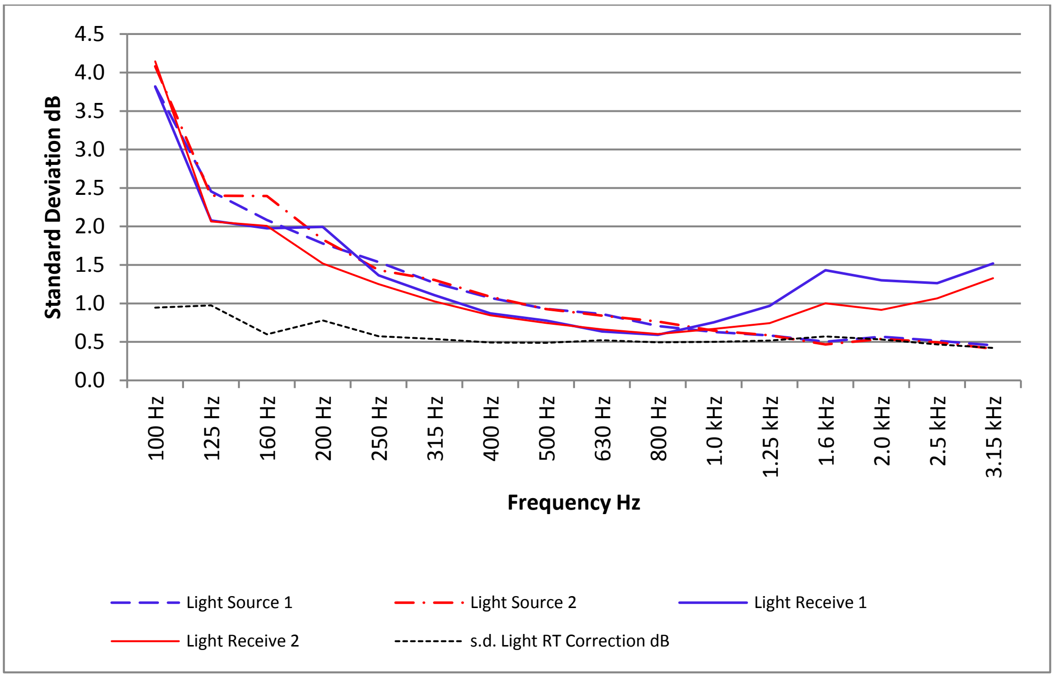


Figure 7-9: Light timber floor - standard deviation values, source room Lp, receive room Lp and RT

In both cases, the reverberation time has the lowest standard deviation of the measured components, especially at low frequency, where it is significantly less than the measured source and receive room sound pressure level variability. In addition it does not appear to be affected by background noise in the higher frequency range, as is the case with the receive side sound pressure levels. The standard deviation for reverberation time correction falls to around 0.5dB between 1.25KHz and 3.15KHz. In this part of the frequency range the receive room sound pressure level variability is dominant and will, based on the summation in quadrature principle, have the greatest influence on the overall measurement uncertainty. In any event the reverberation time measurement correction term variability is relatively low and consequently will be the least influential measured component in the calculation of the third octave band D_{nT} sound insulation values.

7.2.4 Source and receive room sound pressure levels

In order to provide more information on the variability of the single figure values and third octave band levels (D_{nT}), the measured sound pressure levels, in the source and receiver rooms, were considered separately. In this situation it is helpful to use pooled standard deviations, where it is assumed that test series are of the same precision, although their means may differ. It is assumed therefore that there exists a single underlying standard deviation σ of which the pooled standard deviation Sp is a better estimate than the individual standard deviations S_1, S_2....S_k

The pooled variability for all testers is calculated from the arithmetic mean of the within laboratory variances, see repeatability variance, Para 5.1.3.3 of [7]:

$$\sigma_r^2 = \overline{var\ (e)} \qquad 7\text{-}1$$

The first consideration is the loudspeaker positions. The source and receiver room sound pressure levels, for two loudspeaker positions (1 & 2) are shown for concrete (Heavyweight) floor in Figure 7-10 & Figure 7-11 and for timber (Lightweight) floor in Figure 7-12 & Figure 7-13.

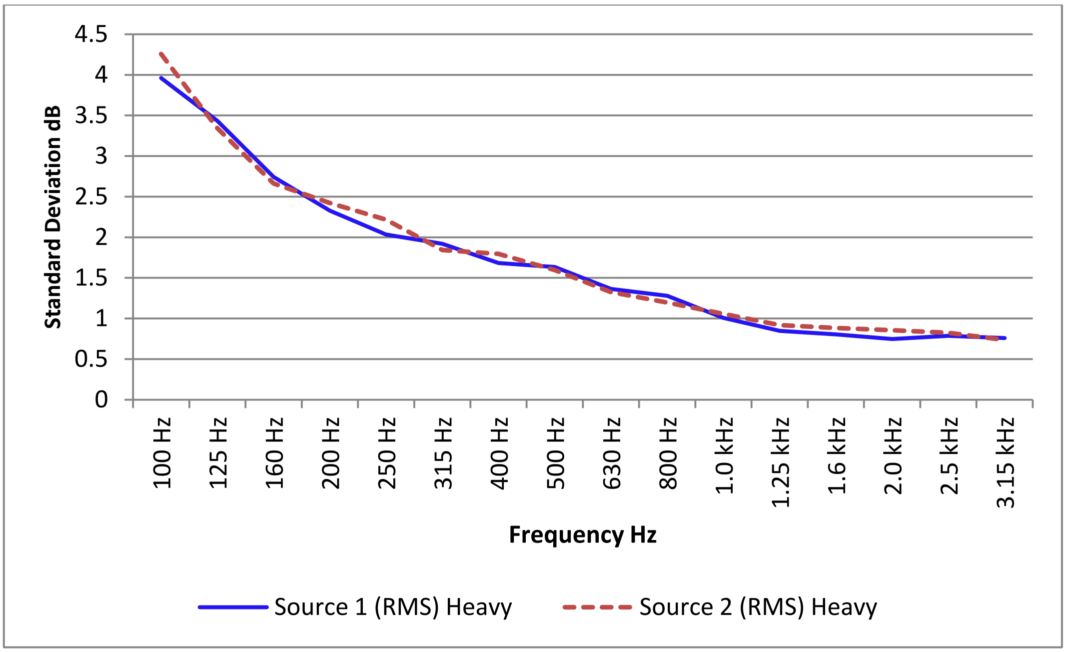


Figure 7-10: Source Room Position Standard Deviations Concrete (Heavyweight) Floor

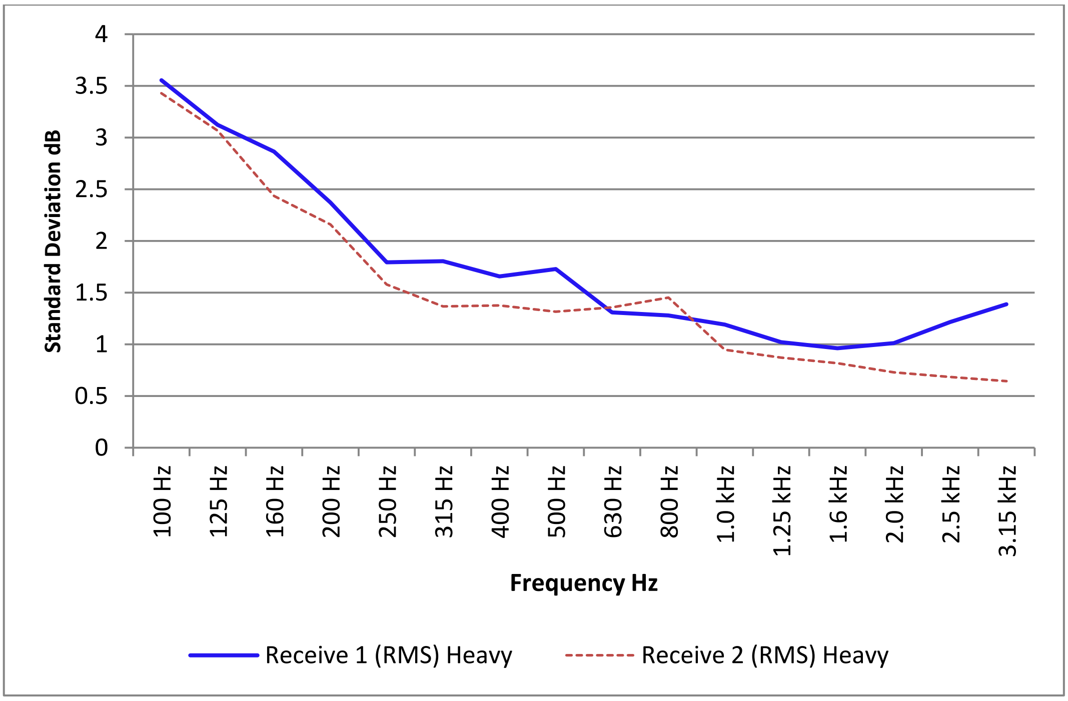


Figure 7-11: Receive Room Position Standard Deviations Concrete (Heavyweight) Floor

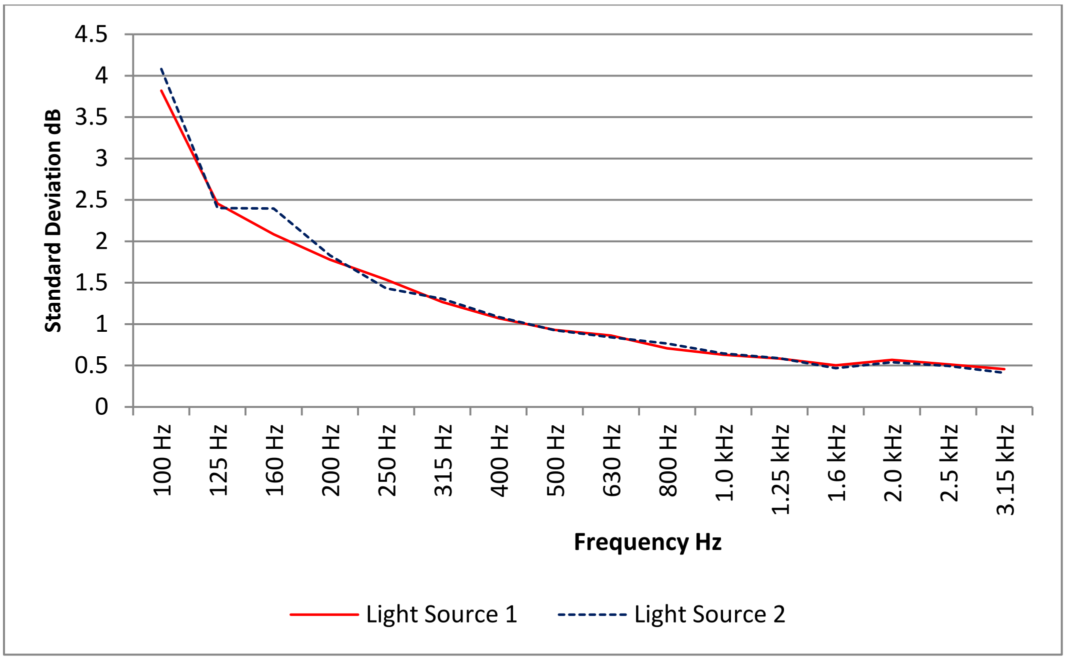


Figure 7-12: Source Room Position Standard Deviations Timber (Lightweight) Floor

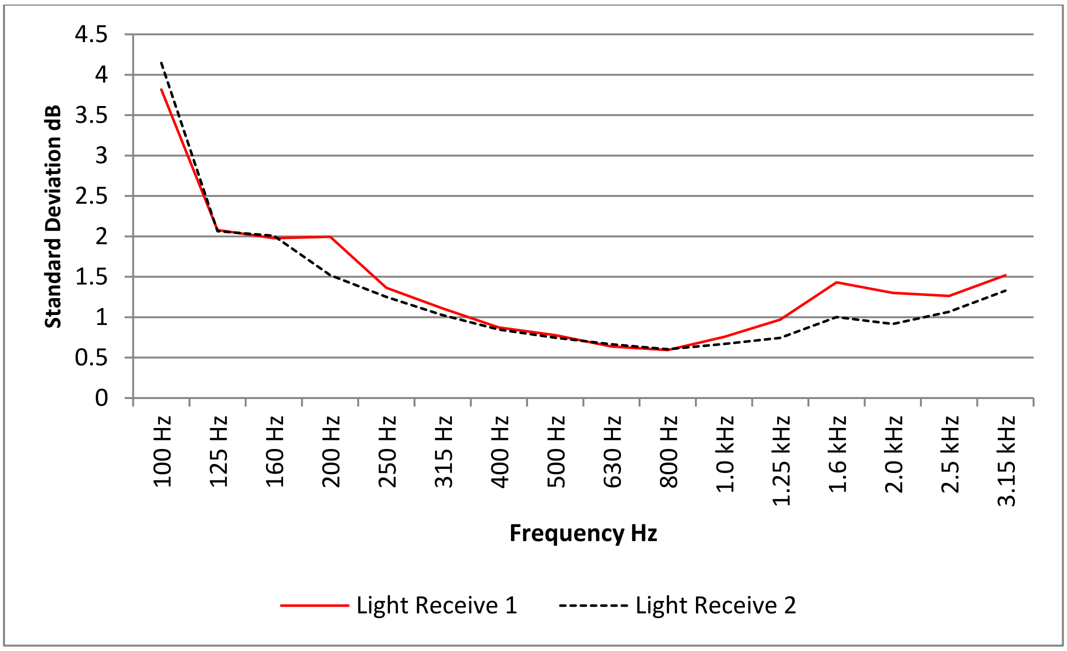


Figure 7-13: Receive Room Position Standard Deviations Timber (Lightweight) Floor

The curves for the loud speaker positions are similar and sometimes nominally identical in the timber and concrete tests. The concrete and timber floor tests show a consistent reducing standard deviation with increase in frequency.

The receiver room curves exhibit greater differences for both the timber and concrete tests, when compared with the source room curves.

The receiver curves, both for timber and concrete floors, both show an up-turning curve at higher frequency. In the timber floor test both loudspeaker position data are relatively consistent and show an upturn at 630Hz. For concrete the receiver room data is less consistent over the loudspeaker positions with one position showing a decline and the other exhibiting an upturn at 1250Hz. The concrete data is more variable in the mid range than the timber receiver room data.

Comparing rooms, the receiver room sound pressure levels exhibit higher standard deviations than the source room data which is expected due to the presence of background noise. The similarities between measurements of both loudspeaker positions are encouraging in that the measurement procedure appears to offer reliable data for both floor types.

In order to ensure that these data are reliable, it is now compared with theoretical standard deviations for both source and receiver room data.

7.2.5 Source and receiver room L_p standard deviations

Initial comparisons in Figure 7-14 & Figure 7-15 show the source room standard deviations follow each other closely and exhibit the same standard deviation curve shape between 100-3150Hz with the lightweight timber floor source room having slightly lower (0.4 – 1.0dB) standard deviations across the frequency range than the heavyweight concrete floor source room. The receiver room sound pressure level data standard deviations are different for the timber and concrete tests. The timber floor has lower values below 1000Hz when there is an upturn and the timber floor test data standard deviations are higher than the concrete case, which tends to indicate there may have been factors affecting the measurement of sound pressure level in the timber test receiver rooms. This effect may have been due to external influences on the measured data in the receiver room and the possible causes are investigated later in this chapter.

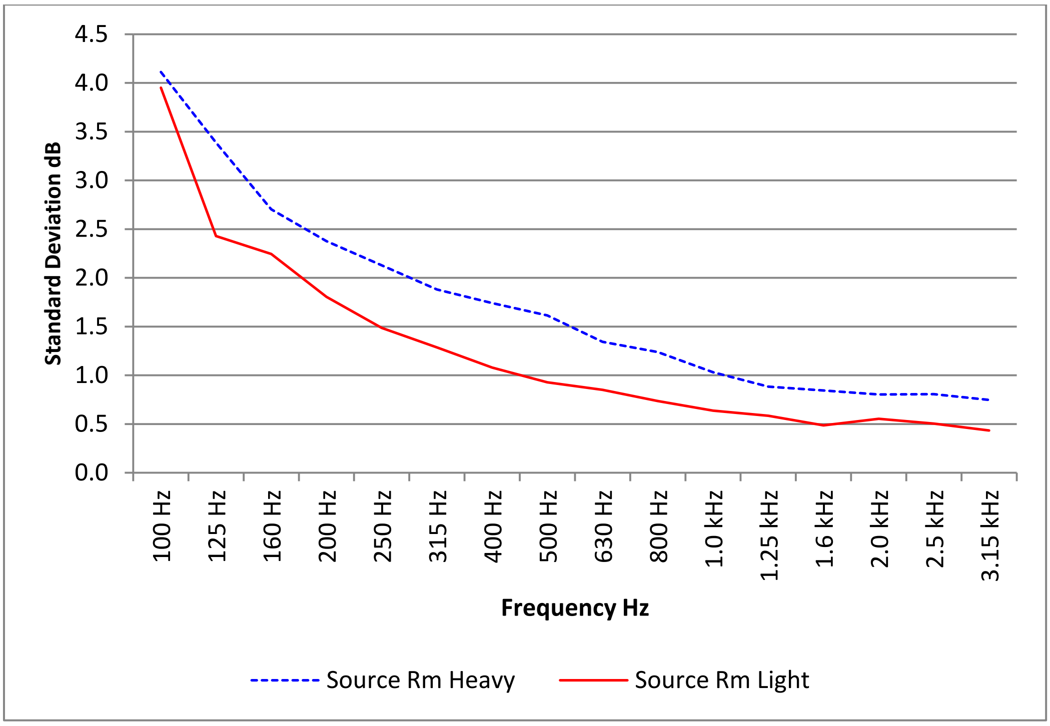


Figure 7-14: Source Room Standard Deviation Comparison Light v Heavy Construction – 180 samples

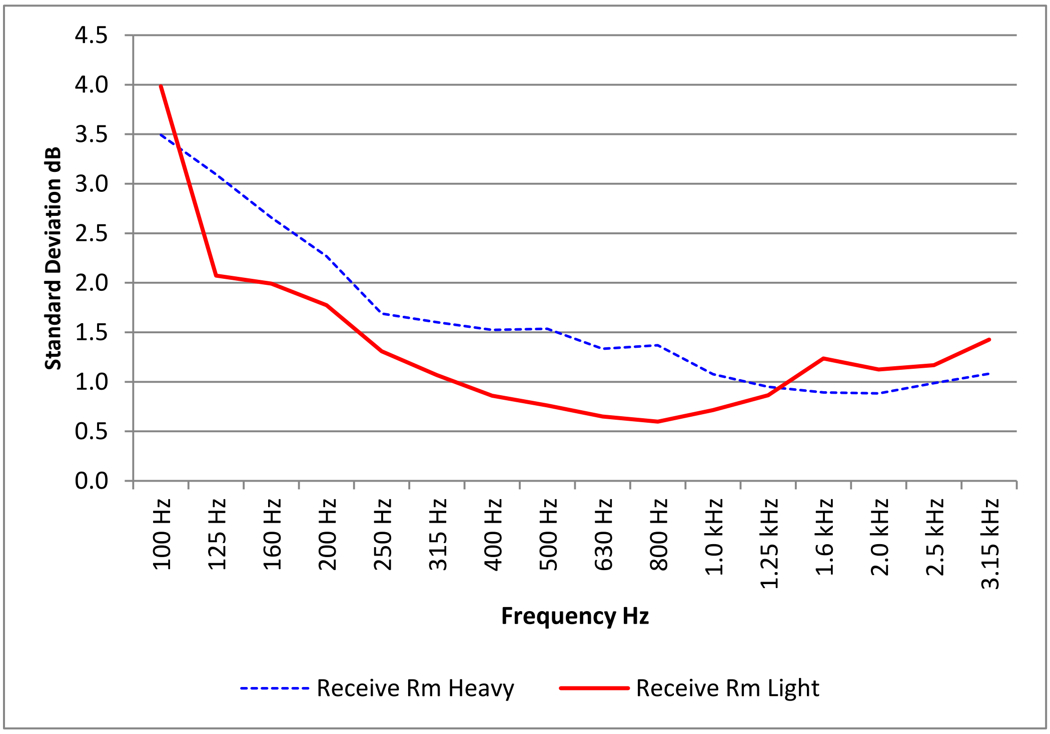


Figure 7-15: Receiver Room Standard Deviation Comparison Light v Heavy Construction – 180 samples

The theoretical variability of sound pressure level in rooms was first discussed by Schroeder [46] for a multi modal space with broadband excitation, see 7-2:

$$\sigma_{SPL} = \frac{5.57}{\sqrt{1+0.238BT}}: \qquad \textit{7-2}$$

where BT is the bandwidth, B and the reverberation time T.

This predicts the standard deviation above the Schroeder frequency f_s and Hopkins [95] indicates it gives reasonable estimates above $0.5f_s$ (p72).

Lubman [119] proposed an equation to calculate this which was further developed by Craik [55], as follows in 7-3 see also in Hopkins et al [96] which applies to the region $0.2f_s$ to $0.5f_s$.

$$\sigma(dB) = \frac{4.43}{-0.22+\sqrt{1+0.319\,N}} \qquad \textit{7-3}$$

N = number of modes in the frequency band

These can be calculated from the product of the filter bandwidth, B and the statistical modal density n(f), from Dah-You [120]. See also Hopkins et al [96].

$$n(f) = \frac{4\pi f^2 V}{c_0^3} + \frac{\pi f S\prime}{2c_0^2} + \frac{L\prime}{8c_0}: \qquad \textit{7-4}$$

where:

f = the band centre frequency (Hz); V = volume of the room (m3); S’ is the total surface area of the space (m2); L’ = the total length of all edges (m); c_0 = speed of sound m/s: assumed to be 340.3 m/s which corresponded to a temperature of 15°C inside test rooms.

This is detailed for the lightweight case, which because of the similarities, is representative of both constructions: see Figure 7-16.

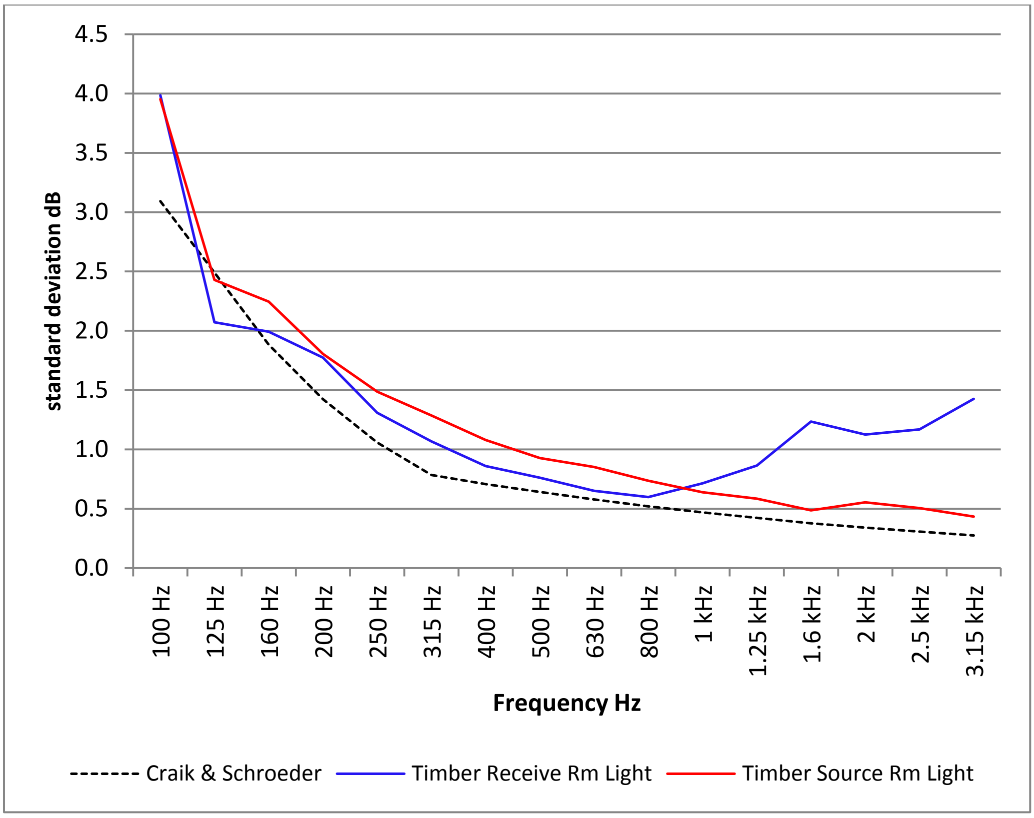


Figure 7-16: Timber Lightweight floor: average standard deviation for all testers of Source & Receiver room measurements comparison with theoretical level 100-3150Hz: after Schroeder & Craik.

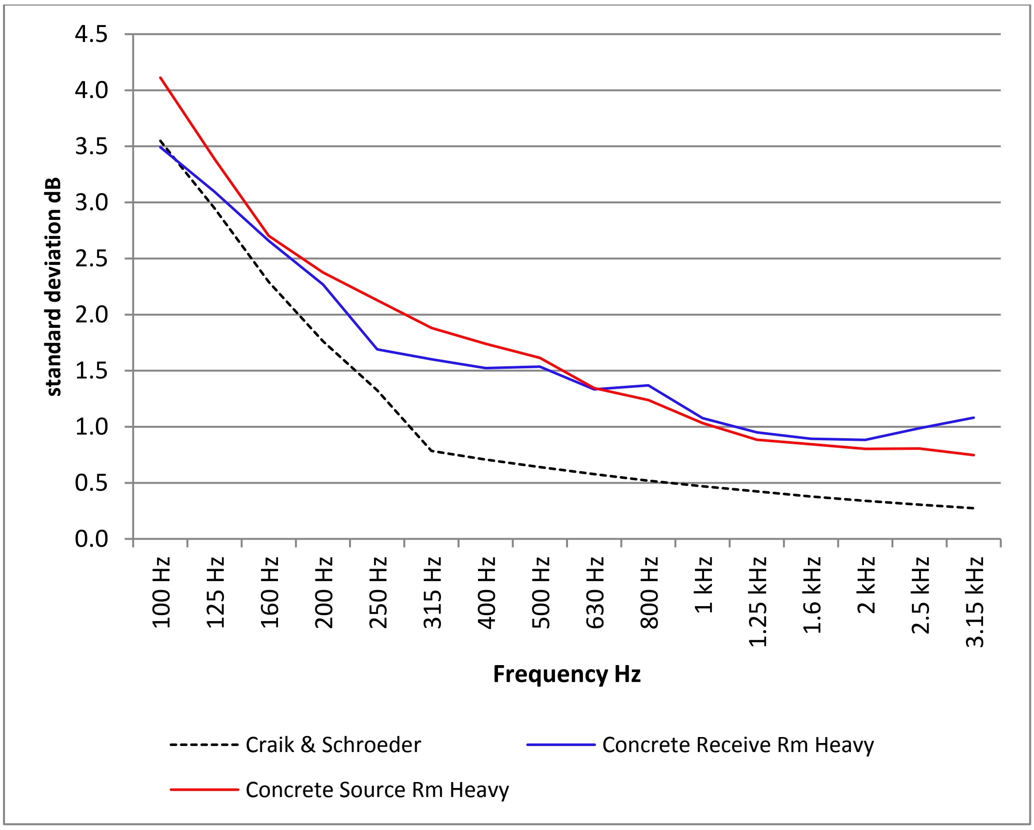


Figure 7-17: Concrete Heavyweight floor: average standard deviation for all testers of Source & Receiver room measurements comparison with theoretical level 100-3150Hz: after Schroeder & Craik.

Given the room sizes for the concrete and timber GRR the Schroeder frequency ranges between the 250 and 315Hz third octave bands Given this 0.2fs is approximately 50 – 60Hz and 0.5fs is 125 – 160Hz which indicates where the Craik and Schroeder formulae provide reasonable estimates of the standard deviation of the sound field.

This can be compared with a test case in a laboratory environment (Hopkins et al [96]) previously detailed in Chapter 3**Error! Reference source not found.** & Figure 3-2.

The predicted variability and that measured in the laboratory is generally lower than that measured in the field

For the timber floor test data the source room levels are 2dB higher than the predicted levels at 100Hz and follow a downward trend which has a standard deviation between 0.5dB and 1dB higher than the predicted levels indicate. The receiver room levels are also 2dB higher at 100Hz falling 1dB below the predicted standard deviation at 125Hz and then are between 0.2 and 0.7dB above the predicted standard deviations between 160 – 800Hz. After 800Hz the standard deviation in the receiver room increases with frequency which is likely to be due to the influence of background noise in this region.

For the concrete floor source and receiver room data the standard deviation generally falls with an increase in frequency. For the source room the standard deviation is 1dB higher than the predicted level between 100-160Hz, it still reduces with increasing frequency but is 2dB higher when it reaches 400Hz. The difference between measured and predicted standard deviation in the source room reduced gradually to be 1dB higher than the predicted level at 3150Hz; this feature can not be explained definitively at present. The receiver room is generally below or similar to the source room standard deviation between 100 – 2000Hz. It achieves the predicted level at 100Hz and is approximately 1dB higher than the predicted level upto 250Hz.. Similar to the source room data it is approximately 2dB higher than the prediction between 315 – 800Hz. This is attributed to the influence of background noise in the mid frequency range on this site.

For both timber and concrete floor experiments the variability of sound pressure level was expected to be higher then the predicted standard deviations as the measurement conditions on site are never likely to be ideal and there is always likely for some of the “pooled” data to be adversely affected by site conditions. In addition the data may contain “outliers” which are difficult to detect against this backdrop. A slight increase in the standard deviation of the source and receiver room sound pressure levels over the theoretical predictions and laboratory comparison would therefore be expected as will a consequential increase in the calculated standard measure of uncertainty r and R.

The comparison of the measured data with the predicted levels demonstrates that the field testing situation has more variability which are likely to be due to site conditions. One of the reasons the measured levels in the receiver room are affected is background noise on site. It is worth examining the contribution is has on the timber and concrete floor receiver room levels as it appears there may be systematic effects which affect both floor construction types to some degree.

7.2.6 Influence of background noise on measured level

Paragraph 6.2 of BS EN ISO140 part 4 [22]: 1998, states :

"*The sound power should be sufficiently high for the sound pressure level in the receiving room to be at least 10 dB higher than the background noise level in any frequency band. If this is not fulfilled, corrections shall be applied as shown in 6.6*."

If the difference is smaller than 10 but greater than 6 the correction is a logarithmic subtraction see 7-5:

$$L = 10\ Lg\left(10^{\frac{L_{sb}}{10}} - 10^{\frac{L_b}{10}}\right) dB \qquad \textit{7-5}$$

where:

L = the adjusted signal level in decibels; L_{sb} = the level of signal and background noise combined; L_b = the background noise level in decibels.

If the difference in level is less than or equal to 6dB in any frequency band a blanket correction of 1.3dB is used corresponding to a difference of 6dB.

Where background noise corrections have to be applied there will be an error introduced into the calculation, which is either a log correction or a fixed value. This is because the instantaneous background noise level in the receiver room varies over time and may be different when the receiver room level is recorded to when the background noise level was recorded. It is clear that if the background noise is relatively high compared to the measured level in the receiver room, then a blanket correction will be applied leading to a clear systematic error in the affected frequency bands.

It is possible to identify three situations in Table 7-6, which individually or more likely, in combination, can allow the background noise correction to systematically influence the variability of the data on a field test. They are:

Table 7-6: Background Noise Correction Scenarios

Cause	**Reason**	**Effect**
1.Insufficient sound power level	Caused by the limitations of the loudspeaker	Not enough sound is transmitted to give a sufficiently high sound pressure level in the receiver room.
2.High levels of attenuation	Caused by the sound insulation performance of the construction under test	Not enough sound is transmitted to give a sufficiently high sound pressure level in the receiver room.
3.High levels of background noise	Caused by site activity in or around the test building or high prevailing ambient noise levels outside receiver room.	Background noise dominates or is at a level that prevents accurate measurement of the sound transmitted through the test construction. adversely influences the total sound pressure level in the receiver room.

The first two indirectly affect the background correction and the measurement uncertainty even if the prevailing background level is relatively low. The third is where background noise itself directly affects the measurement/calculation process. High levels of background noise often are obvious to the operator when arriving on a noisy site.

The background noise influence is investigated for both the lightweight and heavyweight floor tests.

It is essential to understand that the background noise level fluctuates on site. The levels recorded prior to each test are a snapshot of the level at that time and as illustrated by one of the tests on the concrete floor test they can vary between operators by over 20dB in the same room, see Figure 7-18 for the same room recorded by different operators during the experimental process.

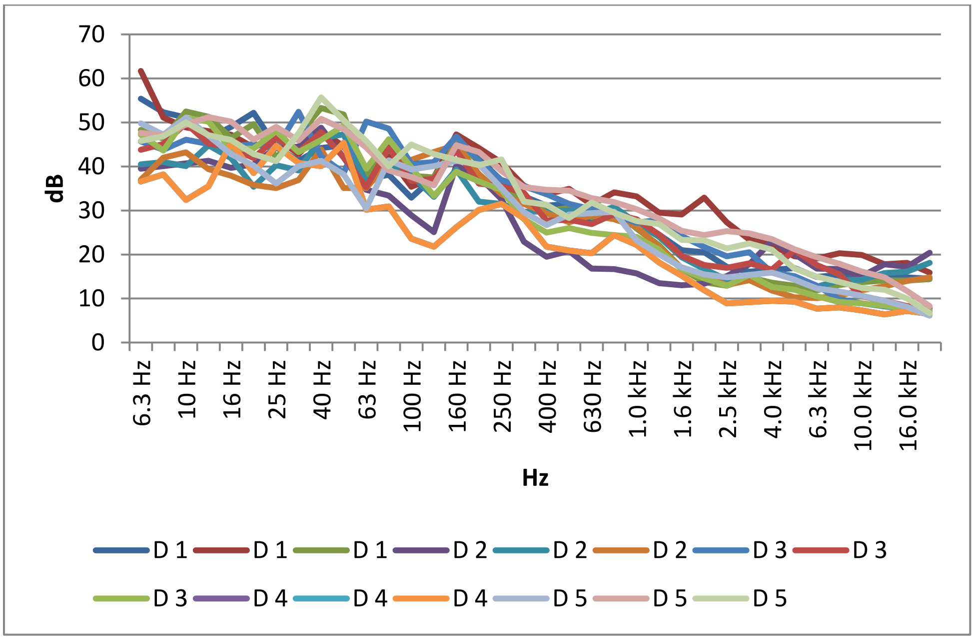
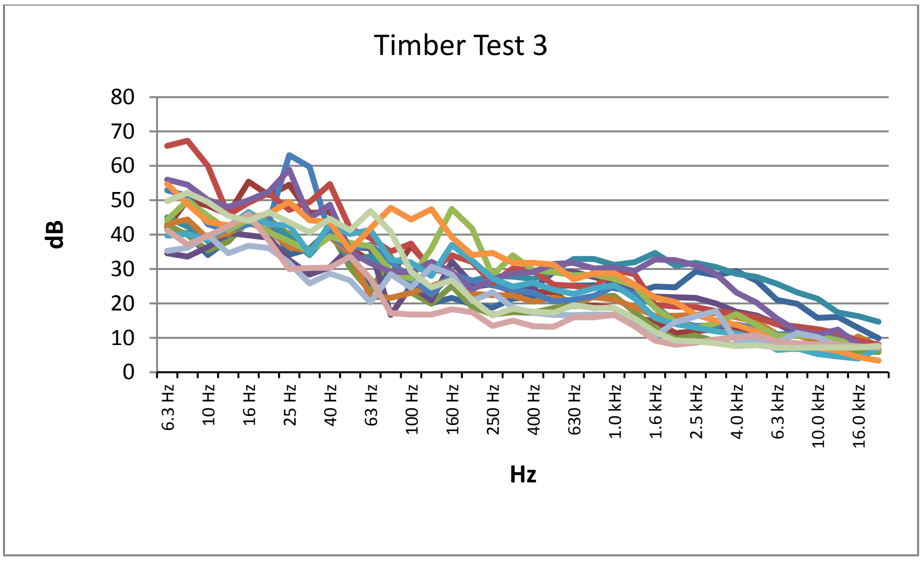
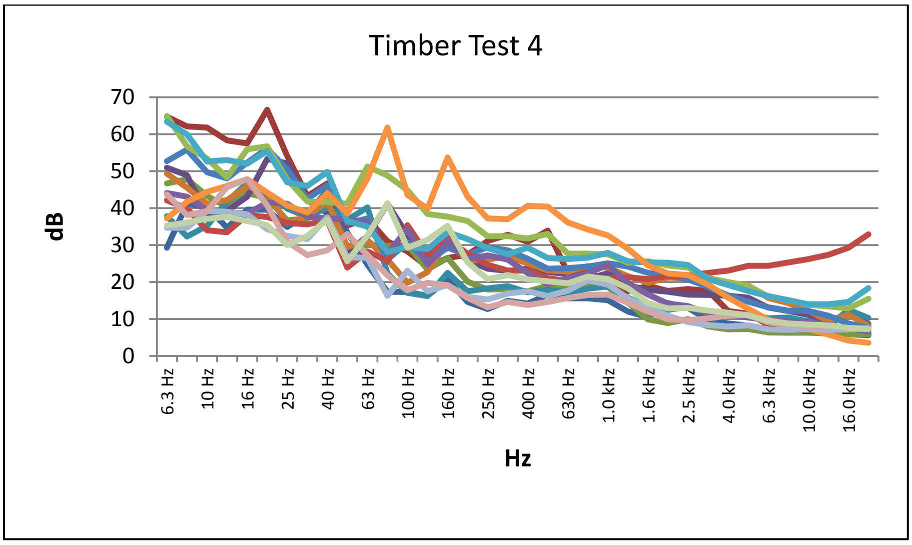


Figure 7-18: Background Sound Pressure Levels - Test room D - Concrete Floor

Several examples of the background noise levels in the timber floor test rooms are represented in Figure 7-19.





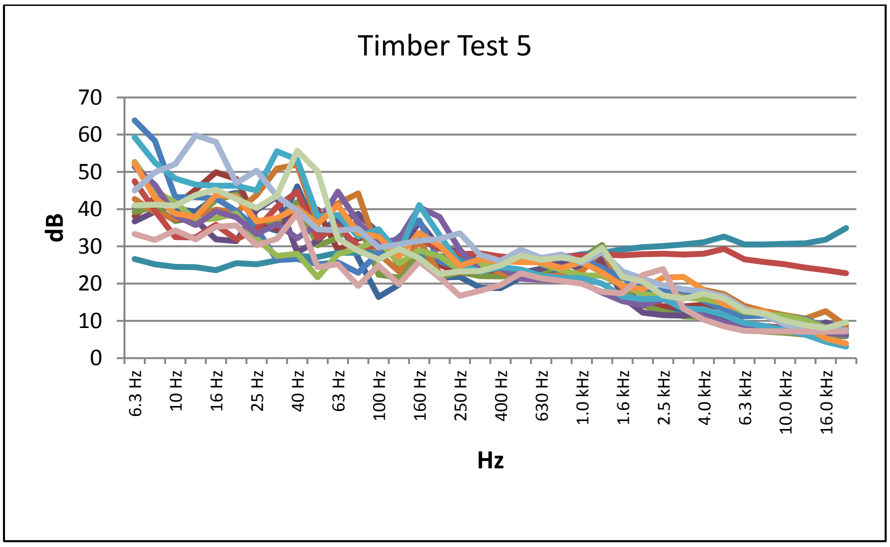


Figure 7-19: Background Sound Pressure levels timber test site variability

The variability in a room is >20dB in some frequencies, over the duration of the experiment.

The field sound insulation test procedure does not allow assessment of the data on site. It is therefore difficult to know with confidence that the background levels the operator has measured in the receiver room at the start of the test will have an adverse impact or not once the data is analysed. Apart from obvious intrusive transients e.g. if someone slams a door or shouts inside the building, it is difficult to know what data to discard and repeat. As the data from the site survey presents a realistic representation of the typical site conditions the pragmatic solution is to retain the data set recorded to reflect the field test situation.

In order to illustrate the impact of the background correction, the frequencies where corrections are made to the measured level in the receiver room are highlighted. This is detailed in Figure 7-20 and Figure 7-21 for both the timber and concrete floor tests. The first pair of rooms tested in the timber and concrete GRR are colour coded for all operators.

The correction applied is represented in Table 7-7:

In this case the corrections are:

Table 7-7: Colour coding for background noise correction during field tests.

No Correction required	>10dB Difference between BG & Lp
Log Correction	Lp between 10dB – 6dB above BG
1.3dB	Lp < 6dB above BG (But greater than BG)
1.3dB	Lp < BG

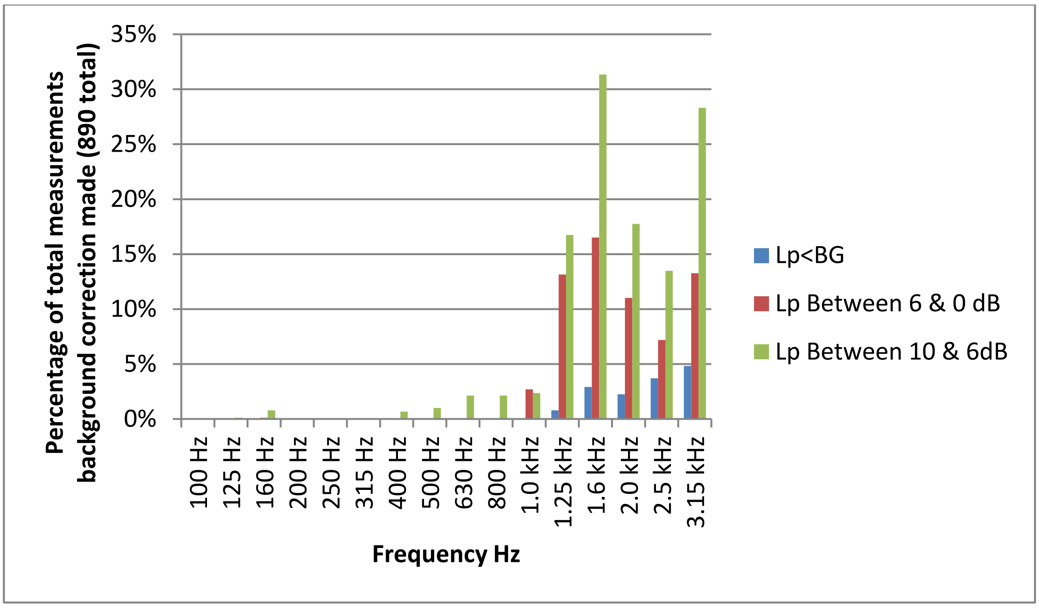


Figure 7-20: Correction for background noise as percentage of receiver side measurements made - Timber floor GRR.

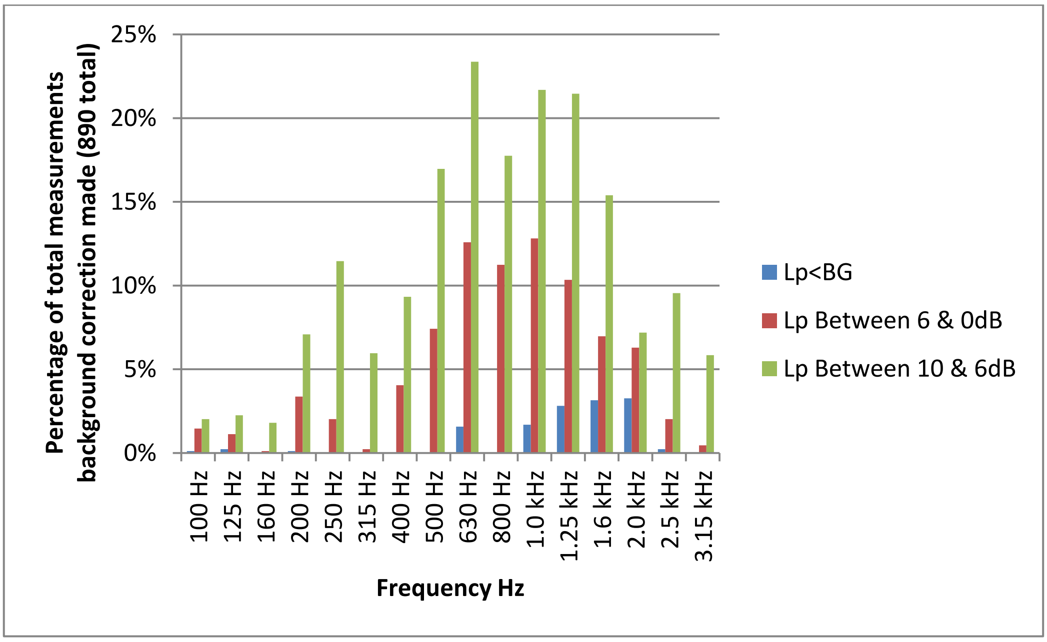


Figure 7-21: Correction for background noise as percentage of receiver side measurements made - Concrete floor GRR.

From the graphical illustration, there is a significant difference between the size and frequency bands where the background noise corrections were applied in each sound insulation test. The timber floor has significant corrections in the frequencies between 1250Hz and 3150Hz. The background noise does not appear to have any significant effects below this frequency range. The concrete floor has the majority of the corrections in the mid range between 400Hz -2000Hz. This would indicate that the background noise corrections for the timber floor would be more likely to influence the higher frequencies, whereas the background noise in the concrete experiment was more influential across the mid range, see Figure 7-22, which shows the region of influence from background noise corrections for the timber and concrete receiver side sound pressure level measurements

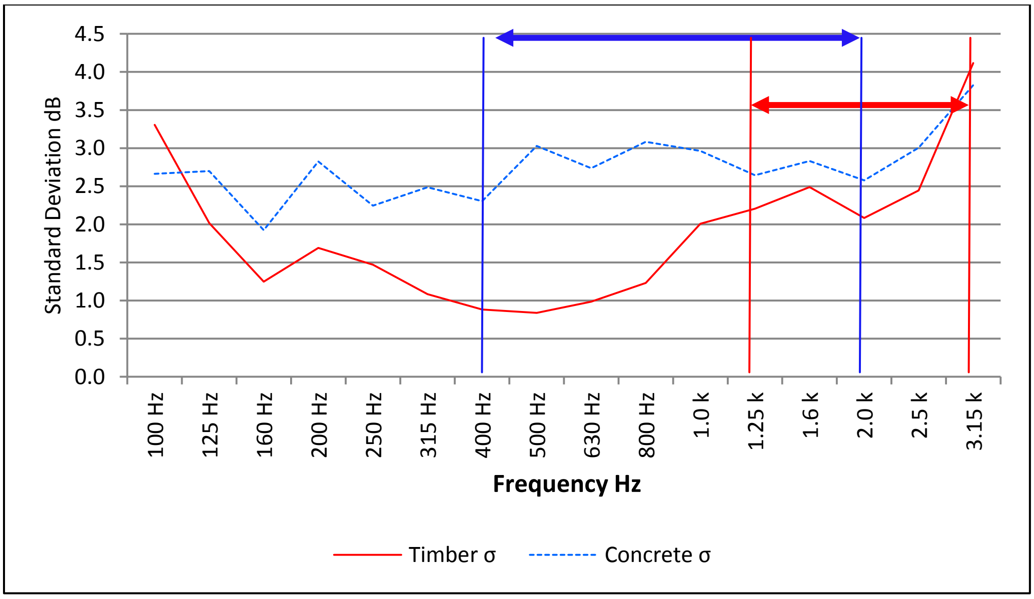


Figure 7-22: Light (Timber) v Heavy (Concrete) Construction D_{nT} - background noise correction region shown

Key	Description
	background corrections in Heavyweight Concrete Tests mainly between 400Hz – 2000Hz
	background corrections in Lightweight Timber Tests mainly between 1250Hz – 3150Hz

7.2.7 Standardised level difference (D_{nT})

The standard deviations of the calculated D_{nT} third octave band values for the timber (σ_T) and concrete (σ_C) floor GRR experiments are detailed in Table 7-8. They contain the influences of background noise and reverberation time corrections previously discussed together with the measurement and the part to part contributions to total variability:

Table 7-8: Standardised Level Difference (D_{nT}) – 90 Test sample

DnT dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
σ_T	3.3	2.0	1.2	1.7	1.5	1.1	0.9	0.8	1.0	1.2	2.0	2.2	2.5	2.1	2.4	4.1
σ_C	2.7	2.7	1.9	2.8	2.2	2.5	2.3	3.0	2.7	3.1	3.0	2.6	2.8	2.6	3.0	3.8

where:

σ_T = standard deviation of timber floor

σ_C = standard deviation of concrete floor

This is shown graphically in Figure 7-23.

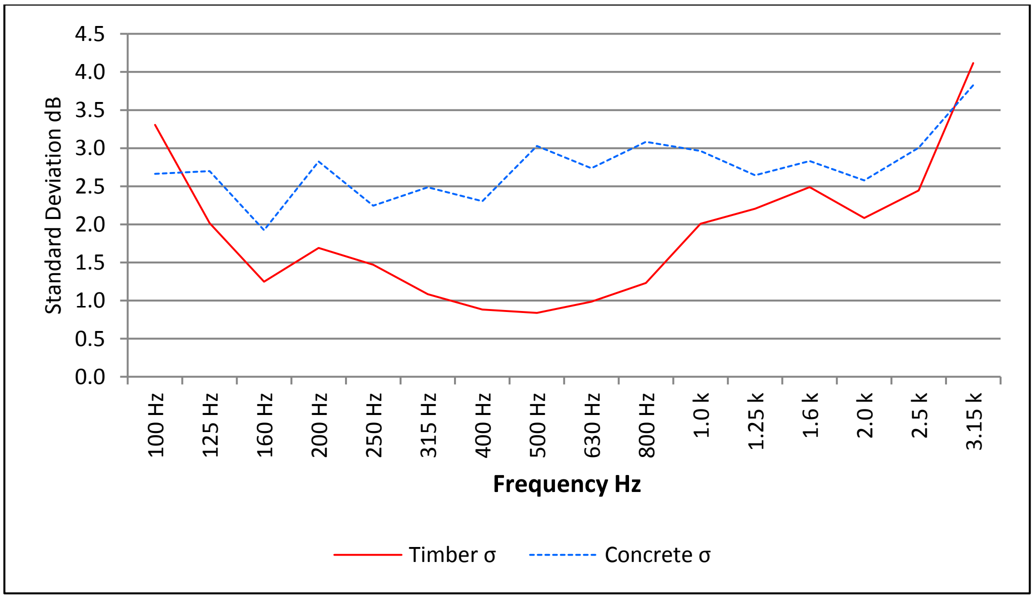


Figure 7-23: Standard deviation of D_{nT} data for timber and concrete floors.

The relative variability of the calculated D_{nT} levels is much lower for the timber floor across virtually the whole of the frequency range. It is comparatively low <1.5dB in the mid range between 250Hz – 800Hz with a minimum value reached at 500Hz (0.8dB).

The standard deviation of the calculated D_{nT} levels for the concrete floors is greater than that of the timber floor from 100Hz and 3150Hz and is between 2 – 3dB higher for the frequency range 125Hz to 2500Hz.

One of the reasons the concrete floor calculated D_{nT} values exhibit relatively high variability across the frequency range may be due to the fact that room sizes varied slightly on site and as a consequence could also be compounded by the reverberation time influence which as demonstrated above has significantly greater variability for the concrete floor test.

The source and receiver room levels and their variability have been assessed and the influence of the background noise and reverberation time corrections has been considered. The data appear to offer a reasonable representation of field test data without any identifiable outliers.

The upturned shape at high frequency associated with the effect of background noise is not unique to this study. In an inter-laboratory study carried out by Muellner [33] see Figure 7-24 the extended range of frequencies above 3.15KHz shows a significant increase in the variability of the reproducibility term based on measurements taken on a timber floor. This is likely due to the combination of the three factors previously highlighted in Table 7-6.

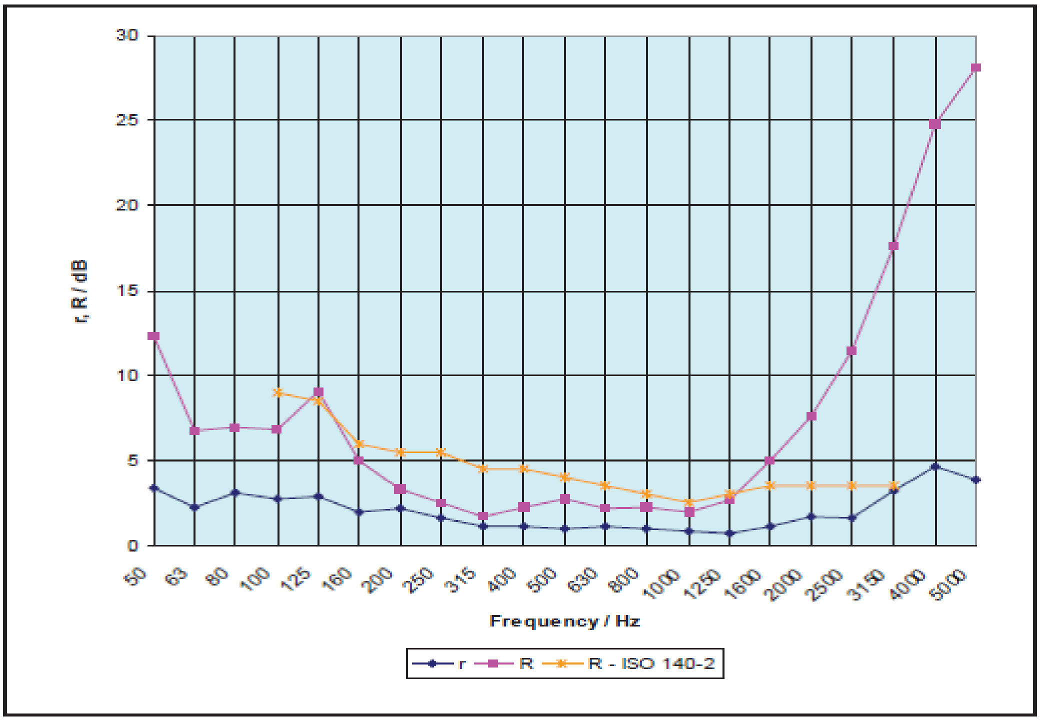


Figure 7-24: Repeatability and Reproducibility of D_{nT} levels of Timber Separating Floor with Wooden Joist Ceiling: after Meuller (fig 5) 2011: compared to ISO140-2 uncertainty values.

7.3 BS5725 (r & R)

An analysis, according to BS5725, was conducted of the concrete and timber floor data, to determine the International Standard repeatability and reproducibility levels, for comparison with the variability guideline values in Table A1 and A2 for laboratories in ISO 140-2: 1993 [8].

7.3.1 ISO 140-2

The guideline values for third octave band data between 100 – 3150Hz for repeatability are produced in Table A.1 and A.2, previously detailed in Figure 4-1 and 4-2. The values are detailed in Table 7-9 for ease of reference.

Table 7-9: Repeatability & Reproducibility values for laboratory tests (airborne sound insulation) ISO 140-2: 1993.

dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
ISO140-2 r	4.5	4	3.5	3.5	2.5	2.5	2	2	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
ISO140-2 R	9	8.5	6	5.5	5.5	4.5	4.5	4	3.5	3	2.5	3	3.5	3.5	3.5	3.5

Field test data is also provided in ISO140-2 the reproducibility values are tabulated in Section A.3 of the standard.

Table 7-10: Reproducibility values for field tests from Table A.3 in ISO140 Part 2 1991

Table A.3 — Reproducibility values for field tests

Third-octave band centre frequency Hz	Reproducibility values *R* for airborne sound reduction index according to ISO 140-4[a] dB	Reproducibility values *R* for normalized impact sound pressure level according to ISO 140-7[b] dB
100	9	7
125	8,5	6
160	6	5,5
200	5,5	5
250	5,5	5
315	4,5	5
400	4,5	5
500	4	5
630	3,5	5
800	3	5
1 000	2,5	4
1 250	3	4
1 600	3,5	5
2 000	3,5	6
2 500	3,5	7
3 150	3,5	7

[a] The reproducibility values *R* have been taken from the corresponding values in Table A.2. The values have been obtained under good acoustical conditions and can be worse in small rooms or in rooms with significant flanking transmission.
[b] The reproducibility values *R* are based on test results from inter-laboratory tests conducted in 1978 involving seven laboratories from the United Kingdom, using as test specimen a wood joist floor construction in one location.

NB. No field test values for repeatability are provided. For reproducibility the field test values are produced in Table A.3[8]. See Table 7-10:

The single number values for sound insulation are not explicitly given though a range is stated in Annex B2 that 1dB is normally achievable for repeatability, and reproducibility will be in the range of between 1 – 3dB.

Using the calculation procedures in BS5725-2 the GRR data for concrete and timber floors can be processed and compared to the guideline values from ISO 140-2. It is essential to do the analysis this way rather than use the ANOVA reproducibility data direct because under BS5725 reproducibility variance is defined as:

$$S_R^2 = S_R^2 + S_r^2 \qquad \textit{7-6:}$$

where S_R^2 is the reproducibility variance and S_R^2 is the between laboratory variance and S_r^2 is the within laboratory variance. The reproducibility variance incorporates the between laboratory variance ***plus*** the repeatability variance. In the GRR ANOVA the reproducibility is defined as the operator variance (or between Lab variance alone).

We would expect that any comparison between field test data and laboratory standard uncertainties, such as the comparison with the laboratory values for repeatability, would show higher levels as there are numerous influential components that are uncontrolled in field testing situations, background noise being the easiest one to define and probably the most influential based on our test evidence.

7.3.2 Concrete GRR

The concrete floor GRR r & R data are graphically illustrated in Figure 7-25 and Figure 7-26 with the ISO 140-2 figures overlaid:

Table 7-11: Concrete Floor (r & R) 90 Test Sample

BS5725 dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
R	4.7	5.3	3.9	3.3	3.7	3.0	2.9	3.2	3.5	4.0	2.6	2.5	2.5	2.6	2.3	2.2
R	6.7	7.3	5.2	4.9	4.7	3.5	3.8	3.8	4.3	4.4	4.0	3.7	3.0	2.9	2.7	2.4
ISO140-2 r	4.5	4	3.5	3.5	2.5	2.5	2	2	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
ISO140-2 R	9	8.5	6	5.5	5.5	4.5	4.5	4	3.5	3	2.5	3	3.5	3.5	3.5	3.5

The concrete floor reproducibility is lower than the ISO curve apart from the frequency range 630-1250Hz.

The field test repeatability data show an expected increase over the ISO140-2 repeatability values for laboratory tests. The repeatability data is above the ISO curve across most of the frequency range and it was higher at low frequency and declined as frequency increased. The repeatability results also had a peak in the data in the mid frequency range 630-800Hz similar to the reproducibility data. It is likely that this was due to some construction site background noise effects from plant and machinery in that particular frequency range.

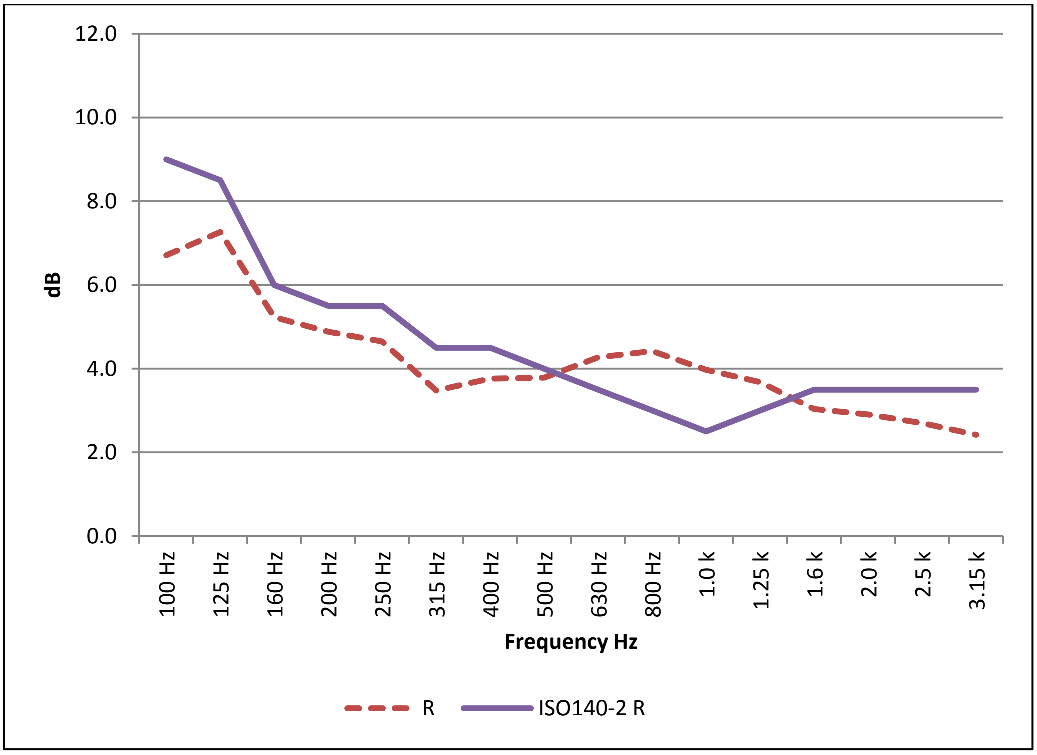


Figure 7-25: Heavyweight floor (Concrete) Reproducibility (R) compared with guideline values from ISO140-2

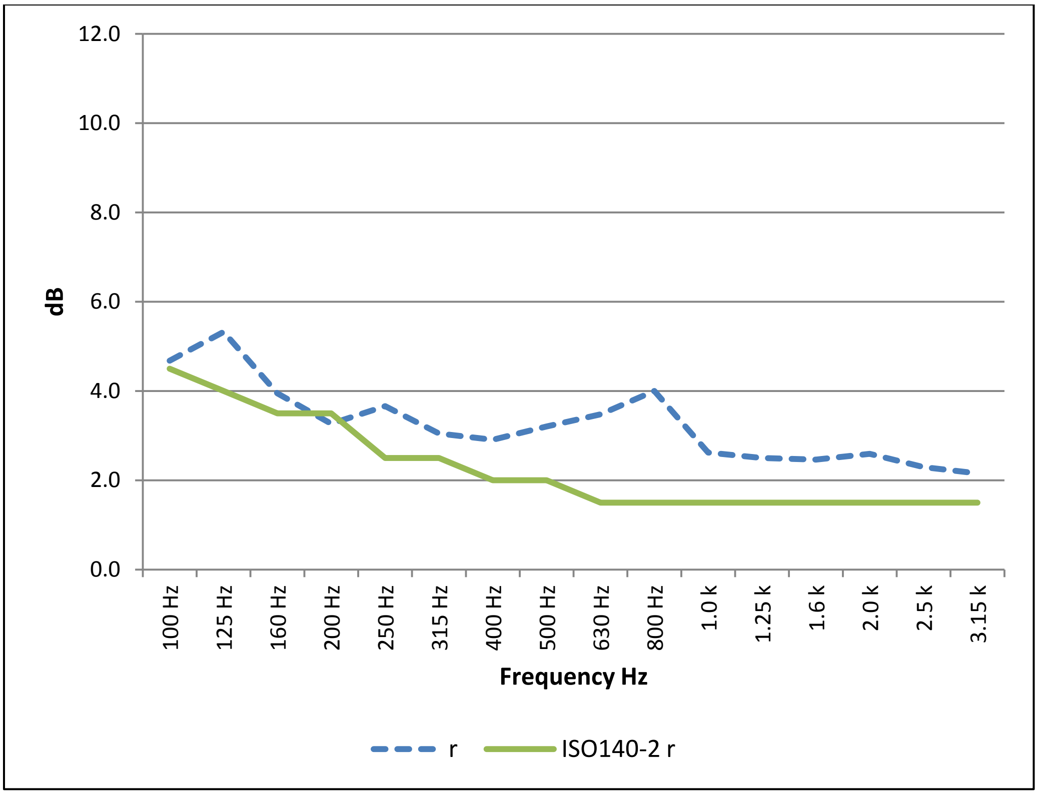


Figure 7-26: Heavyweight floor (Concrete) Repeatability (r) compared with guideline values from ISO140-2

7.3.3 Timber GRR

The timber floor GRR r & R data are detailed in Table 7-12 and are graphically illustrated in Figure 7-27 and Figure 7-28:

Table 7-12: Timber Floor (r & R) 90 Test Sample

BS5725 dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
R	5.4	3.9	2.9	3.5	2.6	2.0	1.5	1.6	1.6	1.7	2.3	2.5	3.2	3.1	2.9	4.2
R	9.9	6.0	2.8	3.8	3.5	2.8	1.9	2.0	2.1	1.8	2.8	4.1	6.9	5.5	5.4	8.4
ISO140-2 r	4.5	4	3.5	3.5	2.5	2.5	2	2	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
ISO140-2 R	9	8.5	6	5.5	5.5	4.5	4.5	4	3.5	3	2.5	3	3.5	3.5	3.5	3.5

The timber floor has relatively good agreement with the ISO 140 guideline values for repeatability although the repeatability levels are generally slightly higher than the ISO curve after 630Hz. The reproducibility also has good agreement apart from the high frequency region after 1000Hz where background noise affected site measurements. At the higher frequency end of the spectrum increased performance of the element under test and the limitations of the loudspeaker to emit sufficient sound pressure in the source room are also considered to be factors in the increase in standard deviation.

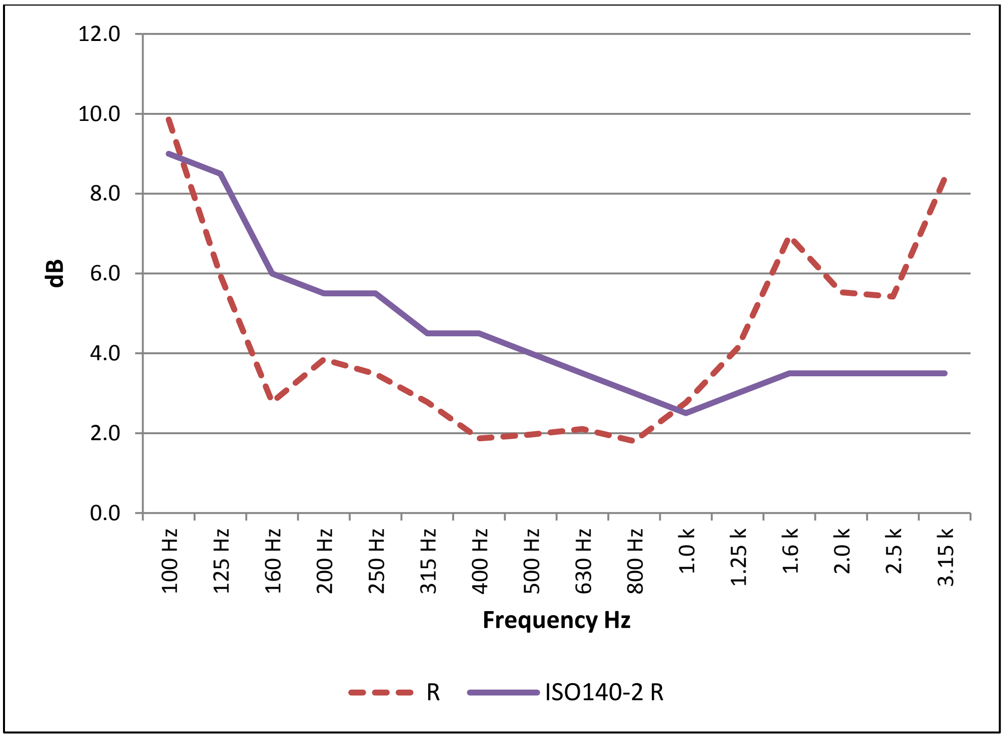


Figure 7-27: Lightweight floor (Timber) Reproducibility (R) compared with guideline values from ISO140-2

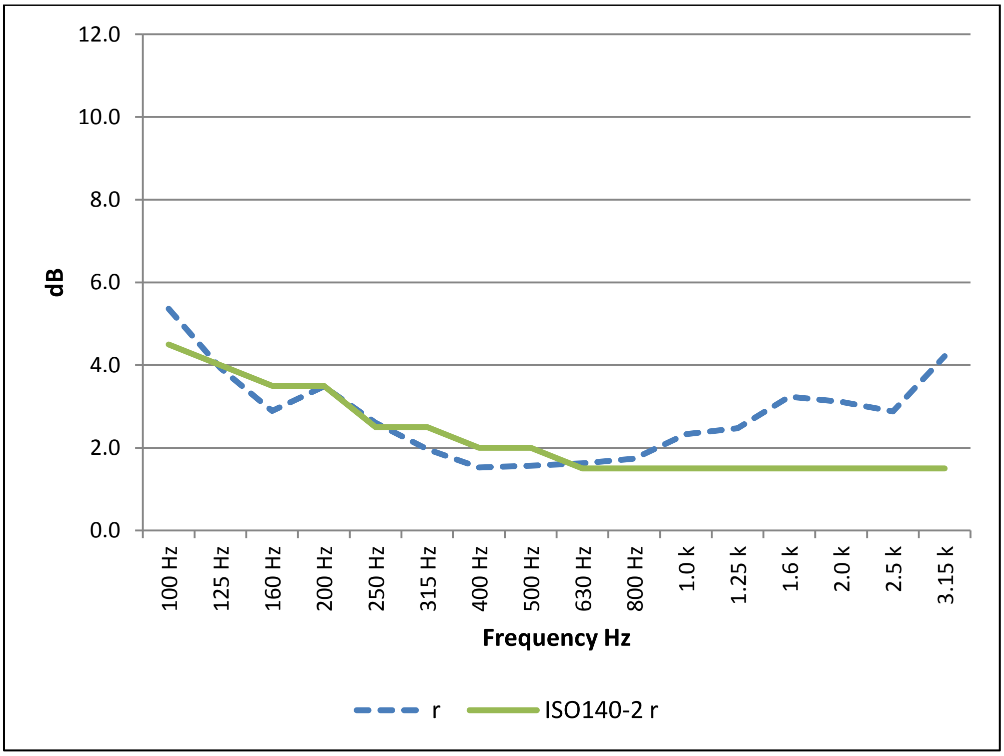


Figure 7-28: Lightweight floor (Timber) Repeatability (r) compared with guideline values from ISO140-2

Apart from the previously identified systematic errors, where the background noise effects are strongest, the above comparison appears to confirm that the field test data offers reasonable data sets for further analysis.

7.3.4 Specimen sample variability

Previous research by Craik et al [57] considered the construction of the separating element, in their case a simple solid concrete cast in situ floor. They determined the part to part variability.

A comparison can now be made between Craik's results and the lightweight timber floor data. In their study multiple measurements of a concrete floor construction (pre-cast slab 125mm thick with 50mm bonded screed) in an existing residential building were recorded to demonstrate that apparently identical floors can give significantly different results. The test method employed used a series of similar room pairs in a student accommodation block that were measured in turn with a "control" floor measured every other test. The repeated "control" test sample was used to demonstrate the measurement variability due to the instrumentation. They concluded the difference between the instrumentation variability and the total variability must be due to "workmanship". A comparison with their data was undertaken by Whitfield et al [121] for a lightweight timber floor with multiple parts. In this case the apparent sound reduction index was calculated to provide a direct comparison between field test data sets.

The Whitfield's data allowed a direct comparison, since all the rooms were of identical shape, size and volume. In addition, the volumes were similar to those tested by Craik i.e. 27m^3 as opposed to Craik's room volume which varied only slightly between 20.7 – 21.5m$^{3.}$. Craik used the same measurement system and operator for all tests; 5 independent test systems and operators were used in the timber GRR experiment. The experimental differences are illustrated in Figure 7-29.

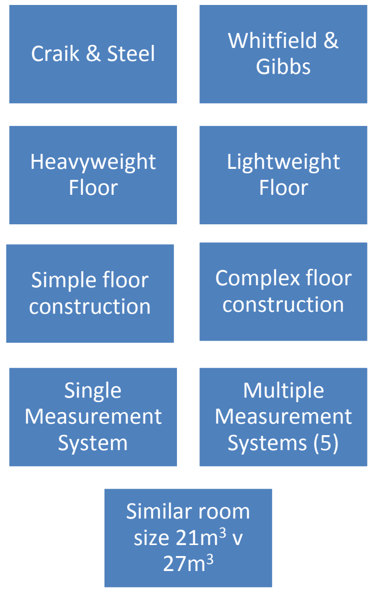


Figure 7-29: Field Measurement Studies for Comparison

In Figure 7-29 the approach for the lightweight timber floor 5 test systems data compares well with Craik's. In this situation, as previously discussed, the pooled system variability from the timber floor experiment offers an aggregated result for all similar test instrumentation. It is expected therefore that the variability due to the instrumentation will be in close agreement.

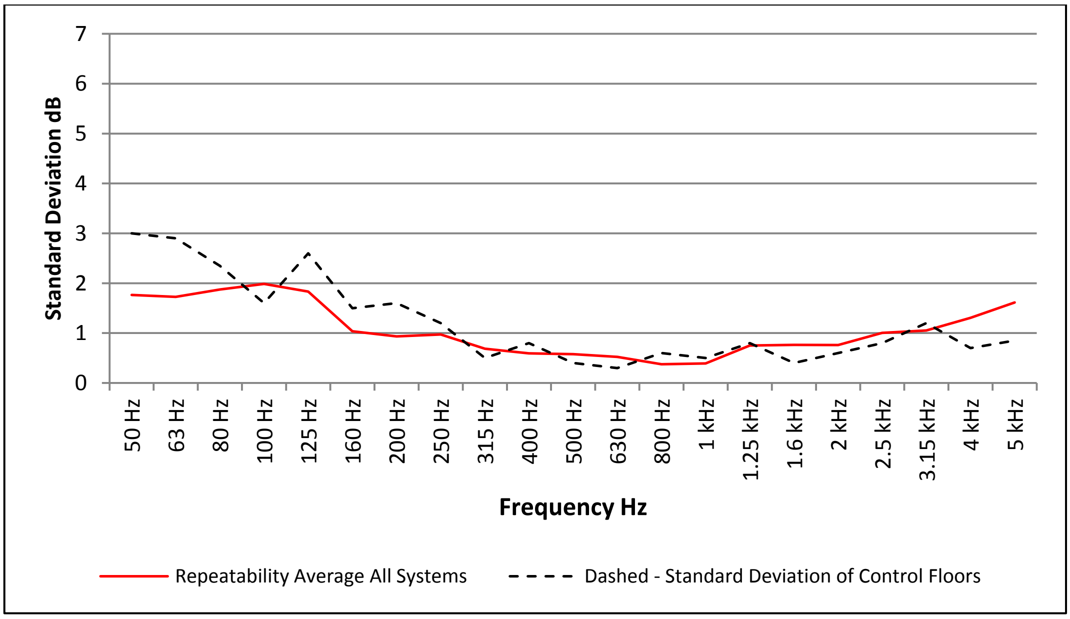


Figure 7-30: Repeatability data compared to standard deviation of Craik's control floors. Apparent SRI - R'dB

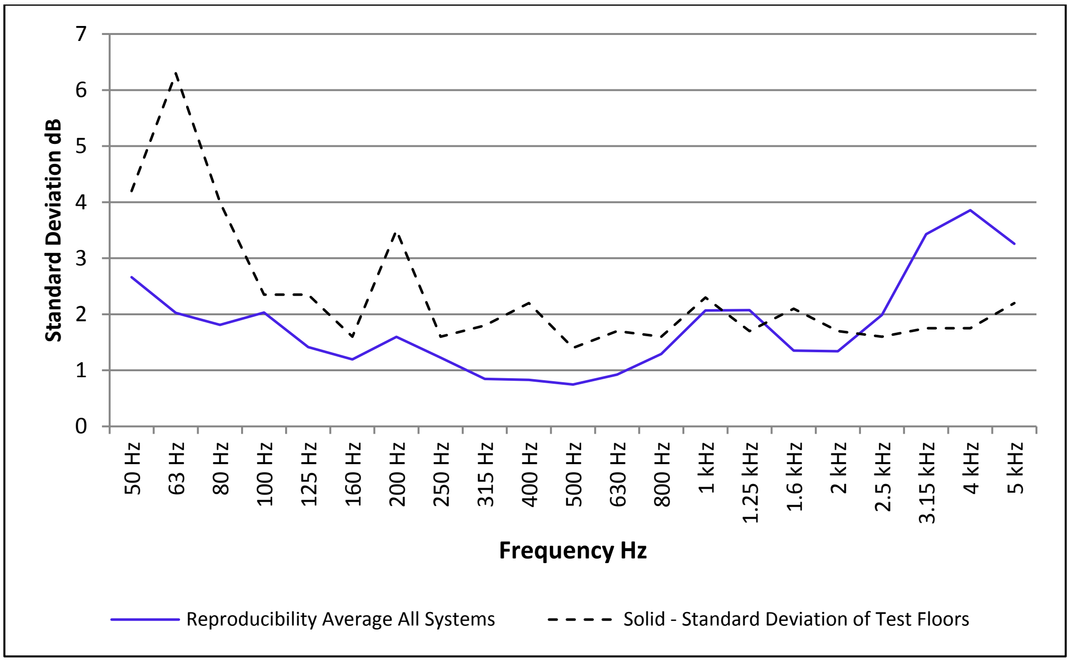


Figure 7-31: Reproducibility data compared to standard deviation of Craik's test floors. Apparent SRI R'dB

Craik subtracts the variability, due to the measurement system, from the total variability in the measurement process and implies the residual variability must therefore be due to the floor construction.

In Figure 7-31 the average reproducibility data from the lightweight timber floor is lower than the control floor tests of Craik for the majority of frequency bands up to 2KHz, above which the background noise influences the results.

The results over the Building Regulation test range of 100-3150Hz showed that the average standard deviation of the simple concrete floor was 1.7dB whilst the multiple component timber floors were 1.3dB. The difference may be partly attributable to construction. It is noted that interaction if any, between factors, cannot be determined using the statistical techniques employed, so far.

The conclusion [121] was that either the workmanship on the timber floor was good leading to lower variability or the floor construction, although complex, does not have a significant bearing on the result. It was noted that Craik's rooms were slightly smaller than the timber floor rooms which could have theoretically led to slightly greater variability at lower frequencies thereby skewing the overall result and that he only had one measurement system which, although reassuringly similar to the repeatability variability for the timber floor tests, may have added to a skewed result in one of the "test floors" raising the overall variability. In any event, with room volumes held fixed (and similar) and both experiments carried out using the international standards test method there does not appear to be any significant evidence in this experiment to suggest that the floor's contribution to the overall uncertainty is based on construction complexity.

7.4 Conclusion: Basic Statistical Comparison

Preliminary comparisons yield useful information about how the timber floor field performance compares with the concrete floor. Consideration of the mean and standard deviations of the single figure values show the floors have broadly similar sound insulation albeit with different variability for the values $D_{nT,w}$ and $D_{nT,w} + C_{tr}$. In addition the spectrum shapes are seen to be similar between the 100 - 1000Hz frequency band. The timber floor performs better at high frequency above 1000Hz.

D and D_{nT} were compared and the influence of the measured reverberation time was investigated, together with the variability of the reverberation time correction term applied to the receiver room sound pressure level. The influence of the reverberation time was counter intuitive because the higher standard deviations for the timber reverberation time data sample measured in seconds resulted in a correction term measured in decibels which had a lower standard deviation overall. This showed that the correction term depended more on the magnitude of the reverberation time than the variability of the measured reverberation time in seconds and its effect would be difficult to classify as an individual input variable.

The measured sound pressure levels in the source and receiver rooms were compared for two loudspeaker positions in the timber and concrete floor tests. There appear not to be significant differences in standard deviation caused by the loud speaker positions but it was apparent that there was increased variability in the receiver rooms and that there appeared to be some systematic uncertainties influencing the receiver room results particularly at high frequency.

The increased variability of sound pressure data in the receiver room at high frequency is due to the background noise. This occurs due to high background noise levels on site and the limitations of the loudspeaker sound power.

The concrete floor exhibited higher standard deviations across most of the frequency range. Some of this increased variability was undoubtedly due to the contribution of background noise in the mid range frequencies and some due to non-identical room sizes. Comparison of the timber floor data with Craik's simple concrete floor showed the timber floor exhibited a lower variability even though it was relatively complex in comparison to the cast in situ concrete construction.

Both timber and concrete test data were analysed according to BS5725 and compared to the repeatability and reproducibility values for laboratory measurement in ISO`140-2. The field test data exhibited only marginally more variability compared with the standard r & r curves. This echoed the comparison of the source and receiver room sound pressure data when compared to the theoretical predicted variability and gives confidence in the field test data and provides a realistic sample for further analysis and assessment.

The difference we see between the timber and concrete samples are worth noting but are not of significant concern as the ANOVA assessments' ability to quantify the contributions of the individual variance components or interactions between factors is not affected. This analysis is carried out in the next chapter.

8 Discussion of Results – GRR ANOVA

8.1 Introduction

The basic analyses in Chapter 7 allowed an overview of the data collected. Both concrete and timber floors were seen to be very similar in performance across the majority of the frequency range. No significant anomalies were observed and apart from some unexpected effects and uncertainty contributions from reverberation time corrections and background noise the data appears to be representative of airborne sound insulation measurements collected in the field for the most common forms of separating floor construction.

This chapter deals with the GRR ANOVA analysis of the same data. The analysis focuses on the proportion of variability due to the instrument, the operator and the part being measured and any interaction there may be between these factors. Comparisons are made between the different floor constructions and an in depth study is made of the variability in sound insulation results at different frequencies.

The linearity test data is added to form a combined sample in order to modify the intentionally constrained timber and concrete floor GRR and to allow the universal DOE guidance for Gauge studies [68] to be followed.

8.2 Analysis

Initial ANOVA calculations were undertaken using a spreadsheet constructed for the task. As the amount of input data for the full GRR was substantial, the ANOVA employed a statistical software package where the analysis could be automated [122]. The software was programmable to tailor numerical & graphical representations of the data.

The tabular output of the statistical software is reproduced in Figure 8-1. It is a familiar ANOVA formatted output and is used to explain the terminology in this chapter:

8.2.1 Two-Way ANOVA Table with Interaction

The information can be described as follows:

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.022	5.0044	3.5774	0.018
Engineer	4	324.822	81.2056	58.0500	0.000
Test Scenari * Engineer	20	27.978	1.3989	0.9611	0.519
Repeatability	60	87.333	1.4556		
Total	89	465.156			

Alpha to remove interaction term = 0.25

Degrees of freedom

Sum of squares

Mean squares

F statistic

P - value

Tests for the main effects of the two factors: P – value <0.05 there is evidence that both these factors are statistically significant

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.022	5.0044	3.4720	0.007
Engineer	4	324.822	81.2056	56.3384	0.000
Repeatability	80	115.311	1.4414		
Total	89	465.156			

Figure 8-1: ANOVA Table of Results for GRR showing "with interaction" and "without interaction" cases

Definitions (term used in results table)

(DF) = Degrees of freedom for each of the source factors or groups analysed

(SS) = Sum of squares; expresses the total variation attributed to each factor, they can be considered as:

SS_O: variation around operators' mean (Between groups "Engineer" in our ANOVA table Figure 8-1)

SS_P: variation around parts' mean (Between groups "Test Scenari" in our ANOVA table Figure 8-1);

SS_{O*P}: Variation around operators' and parts' mean (Within groups)

SS_E: variation around the measurement instrument

(MS) = Mean squares; sum of squares divided by degrees of freedom. (NB: MS are always variances)

MS_O: operator variance

MS_P: part variance

MS_{O*P}: variance between parts and operators

MS_E: instrument variance

(F) = F Statistic – Used in test of variance: F is a ratio of sample variances which the expected value = 1 indicates no difference. It can be classed as the proportion of between-group variation divided by within-group variation. If the F statistic is larger than the critical F value then the variation between groups is statistically significant.

(P) = P – value: in this case the critical p-value is 0.05 (95% level). If $p<0.05$ then reject Ho; there is statistically significant evidence of an overall effect from the factor.

NB: "Gage" is the American spelling of "Gauge"

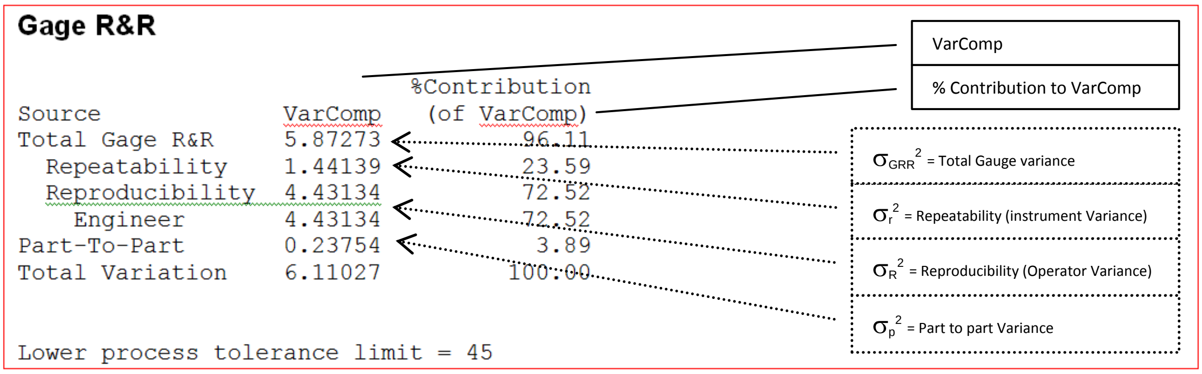


Figure 8-2: Software printout : GRR components of variance quantities (decibels)

Definitions:

VarComp – (term Used in results table) = It is the variance component for each source, in our case measured in decibels: NB: if the p-value of the operator*part interaction (or Test Scenari * Engineer) is greater than 0.25 (it is 0.519 in Figure 8-1) the interaction term is dropped from the calculation and a reduced model is used without interaction.

%Contribution (of VarComp) : this is simply the proportional contribution of a particular source to the total variation in the model.

The results from the ANOVA table software outputs and graphs are summarised in this chapter for brevity and to assist in comparisons.

8.3 GRR Results

The GRR data is presented chronologically. That is:

1. Timber – Lightweight Floor GRR
2. Concrete – Heavyweight Floor GRR
3. Linear Test Sample – Extended range GRR.

In the case of the main timber and concrete floors, the data is presented individually. Iin the case of the linear test, which is added to modify the test sample in order to accord with classical GRR DOE, the test samples are combined to form one large GRR set.

In each case, a preliminary data assessment was carried out using visual and numerical check for normality.

An advantage of ANOVA is that it is relatively resilient to non - normality (see Montgomery's discussion in Section 3.4.1 [70]). However since the field test data was obtained in non-ideal environments visual inspection of the residuals identified deviations from normality.

One form of normal probability plot presents a straight line upon which the sample data can be superimposed. The plots feature each value vs. the percentage of values in the sample that are less than or equal to it, along a fitted distribution line (middle blue line). If the sample follows the straight line closely it can be concluded that the data sample comes from a normal population. It is usual to prefer the normal probability plot to the histogram graph of data as it is usually easier to detect deviation from a straight line than assess a distribution shape against a bell curve. For similar reasons, it is customarily used, where data samples are less than 200 (90 in our case). An example of the residual normal probability plot and the probability plot of data with confidence limits drawn is shown for the 1KHz frequency band for timber in Figure 8-3. In this investigation the probability plot shows the 95% confidence limits.

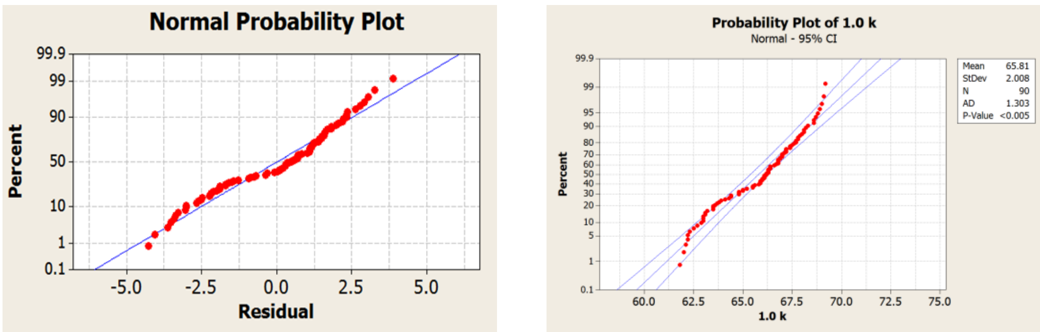
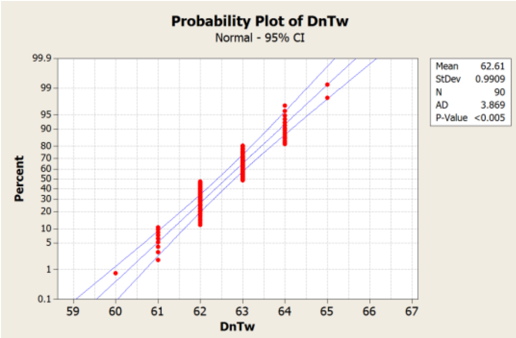
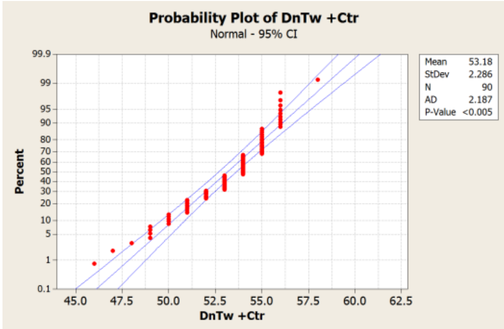


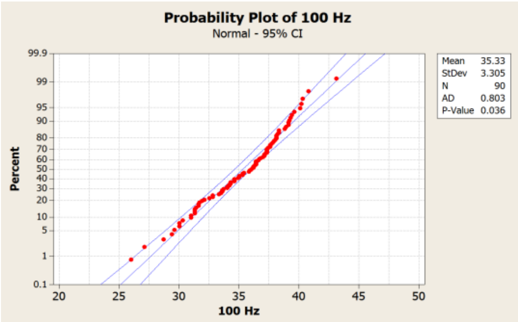
Figure 8-3: Normal Probability Plot of Residuals and Normal Probability Plot of data for 1kHz band for Timber

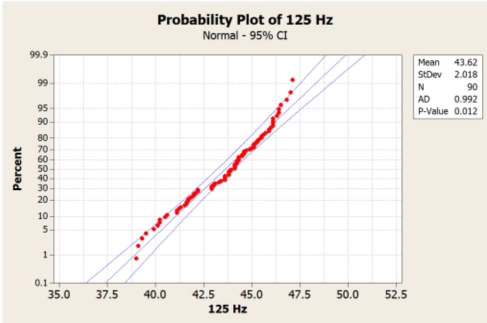
In this example the distribution shape for 1kHz is typical with the tendency of the normal probability plot to bend down at the left and upwards on the right indicating the tails of the error distribution are thinner than would be expected in a normal distribution. In this case the error distribution is said to be "approximately" normal.

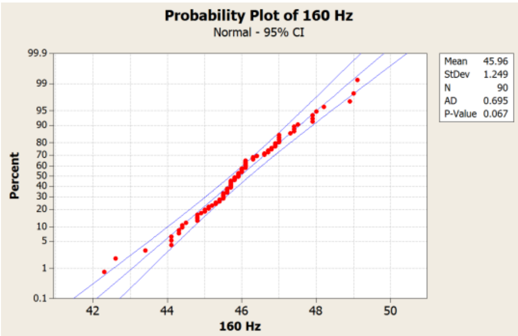
Some probability plots of the single value and frequency data for timber and concrete floors are given in Figure 8-4 & Figure 8-5:

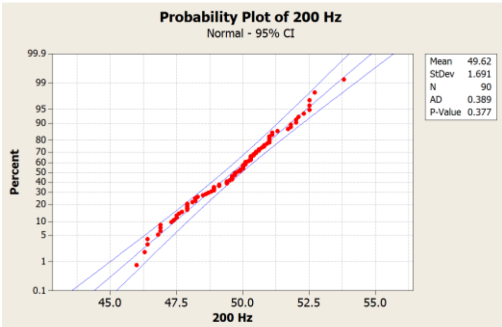

Probability Plot of DnTw
Normal - 95% CI
Mean 62.61
StDev 0.9909
N 90
AD 3.869
P-Value <0.005
Percent
DnTw

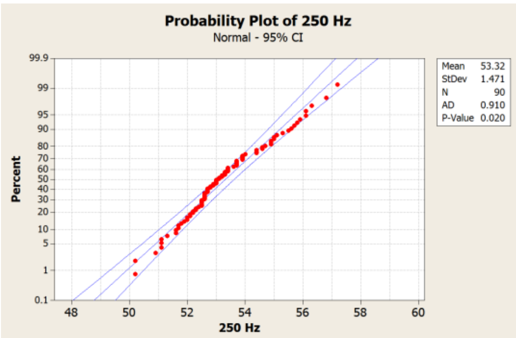

Probability Plot of DnTw +Ctr
Normal - 95% CI
Mean 53.18
StDev 2.286
N 90
AD 2.187
P-Value <0.005
Percent
DnTw +Ctr

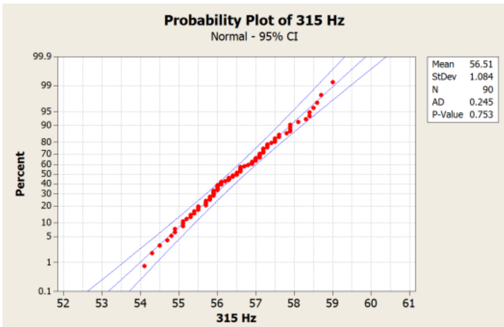

Probability Plot of 100 Hz
Normal - 95% CI
Mean 35.33
StDev 3.305
N 90
AD 0.803
P-Value 0.036
Percent
100 Hz

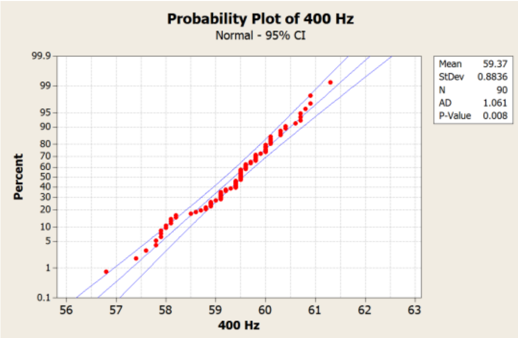

Probability Plot of 125 Hz
Normal - 95% CI
Mean 43.62
StDev 2.018
N 90
AD 0.992
P-Value 0.012
Percent
125 Hz


Probability Plot of 160 Hz
Normal - 95% CI
Mean 45.96
StDev 1.249
N 90
AD 0.695
P-Value 0.067
Percent
160 Hz

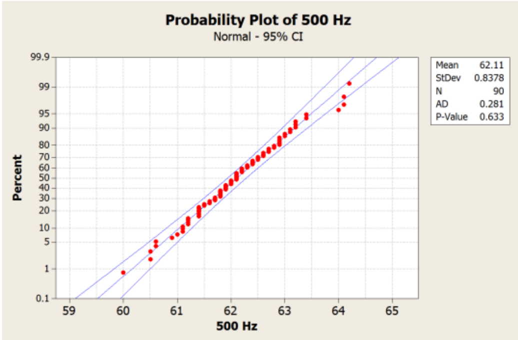

Probability Plot of 200 Hz
Normal - 95% CI
Mean 49.62
StDev 1.691
N 90
AD 0.389
P-Value 0.377
Percent
200 Hz


Probability Plot of 250 Hz
Normal - 95% CI
Mean 53.32
StDev 1.471
N 90
AD 0.910
P-Value 0.020
Percent
250 Hz

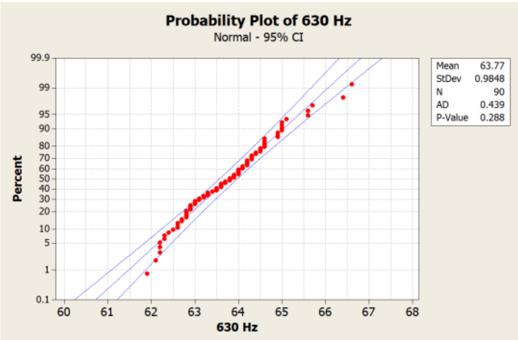

Probability Plot of 315 Hz
Normal - 95% CI
Mean 56.51
StDev 1.084
N 90
AD 0.245
P-Value 0.753
Percent
315 Hz



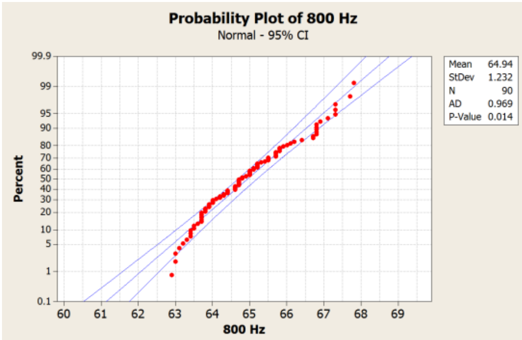
Probability Plot of 400 Hz
Normal - 95% CI
Mean 59.37
StDev 0.8836
N 90
AD 1.061
P-Value 0.008
Percent
400 Hz



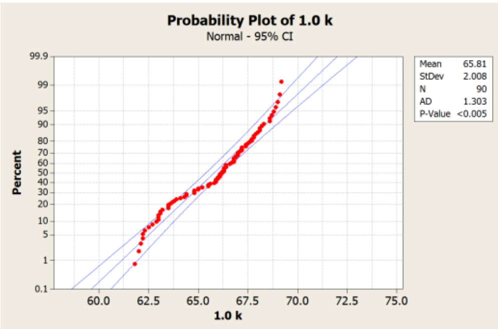
Probability Plot of 500 Hz
Normal - 95% CI
Mean 62.11
StDev 0.8378
N 90
AD 0.281
P-Value 0.633
Percent
500 Hz



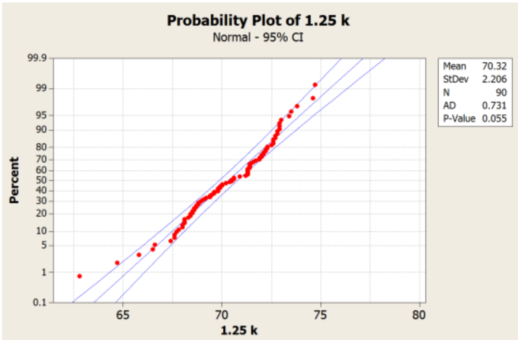
Probability Plot of 630 Hz
Normal - 95% CI
Mean 63.77
StDev 0.9848
N 90
AD 0.439
P-Value 0.288
Percent
630 Hz



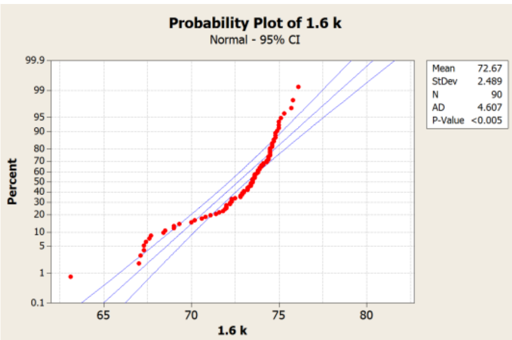
Probability Plot of 800 Hz
Normal - 95% CI
Mean 64.94
StDev 1.232
N 90
AD 0.969
P-Value 0.014
Percent
800 Hz



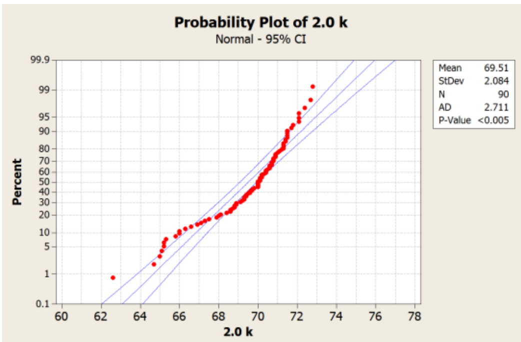
Probability Plot of 1.0 k
Normal - 95% CI
Mean 65.81
StDev 2.008
N 90
AD 1.303
P-Value <0.005
Percent
1.0 k



Probability Plot of 1.25 k
Normal - 95% CI
Mean 70.32
StDev 2.206
N 90
AD 0.731
P-Value 0.055
Percent
1.25 k



Probability Plot of 1.6 k
Normal - 95% CI
Mean 72.67
StDev 2.489
N 90
AD 4.607
P-Value <0.005
Percent
1.6 k



Probability Plot of 2.0 k
Normal - 95% CI
Mean 69.51
StDev 2.084
N 90
AD 2.711
P-Value <0.005
Percent
2.0 k

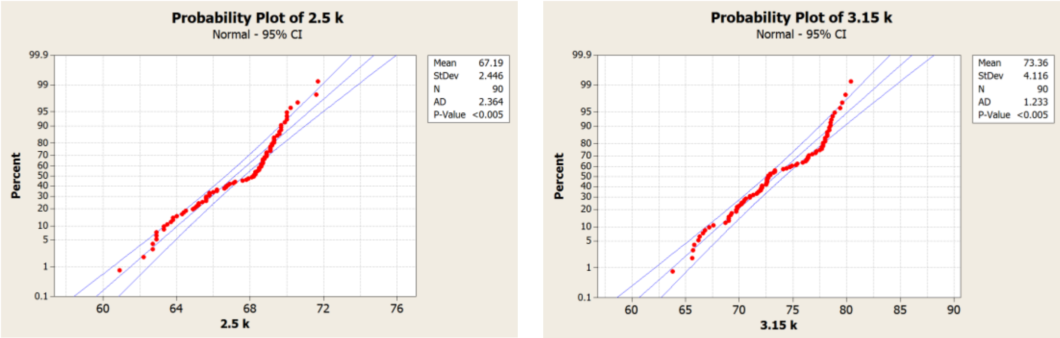
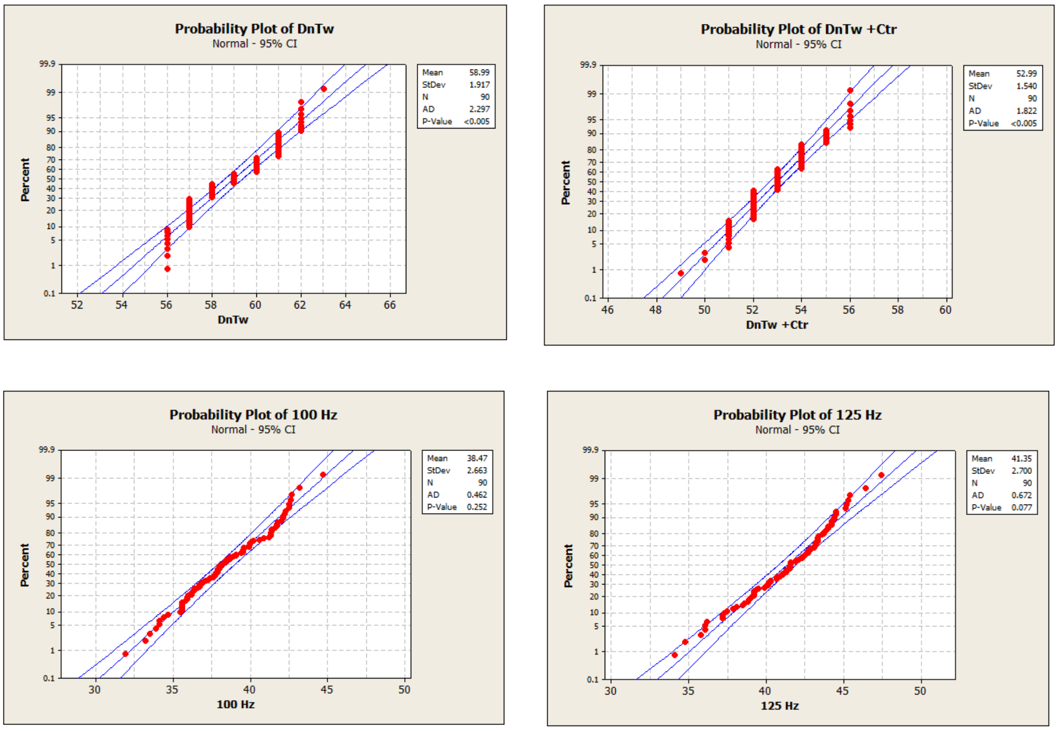
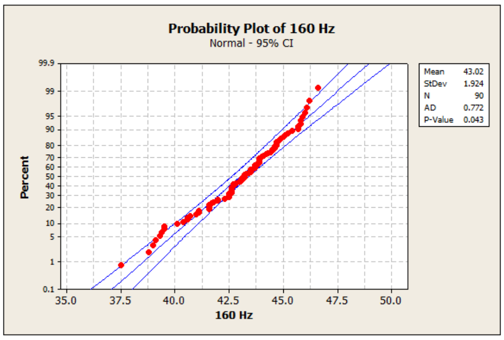
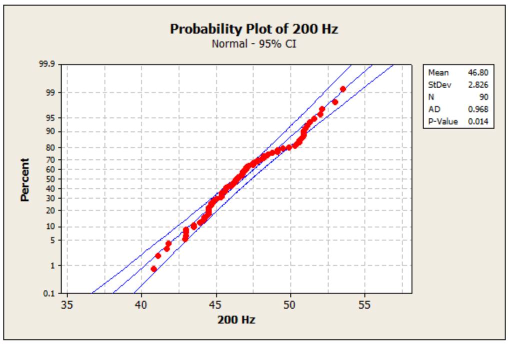


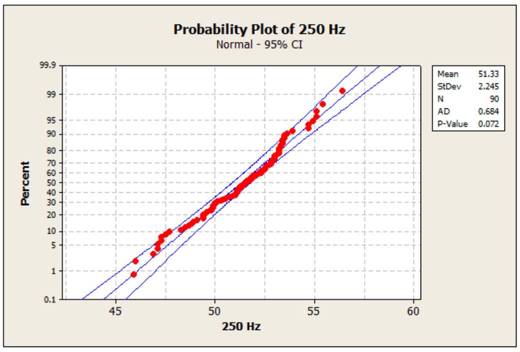
Figure 8-4: Normal Probability Plots - Timber Floor Tests

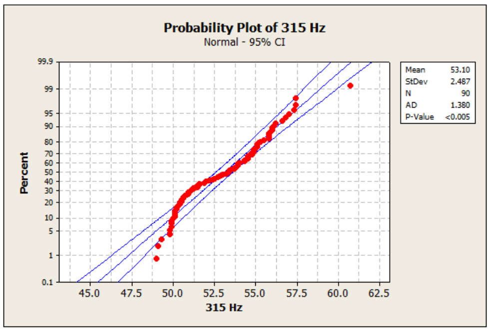
Visual inspection indicates the timber floor data appear to be approximately normal apart from the single figure values and frequency data above 1.25kHz which is likely to be due to the effects of amongst other things background noise increasing the variability at higher frequencies and skewing the data. The probability plots for the concrete floor tests are detailed below:

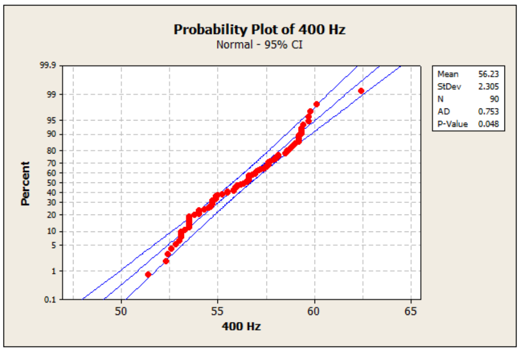


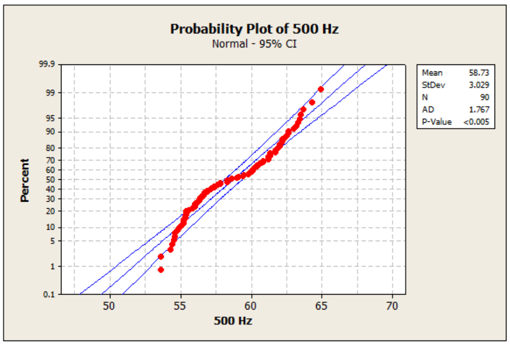

Probability Plot of 160 Hz
Normal - 95% CI
Mean 43.02
StDev 1.924
N 90
AD 0.772
P-Value 0.043
Percent
160 Hz

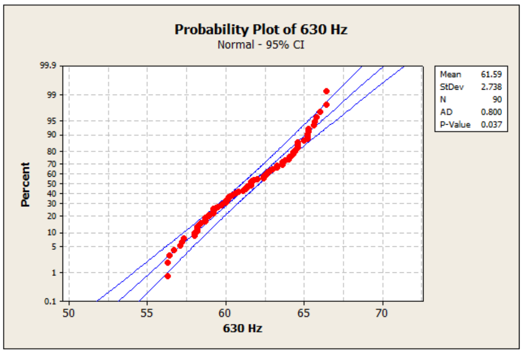

Probability Plot of 200 Hz
Normal - 95% CI
Mean 46.80
StDev 2.826
N 90
AD 0.968
P-Value 0.014
Percent
200 Hz

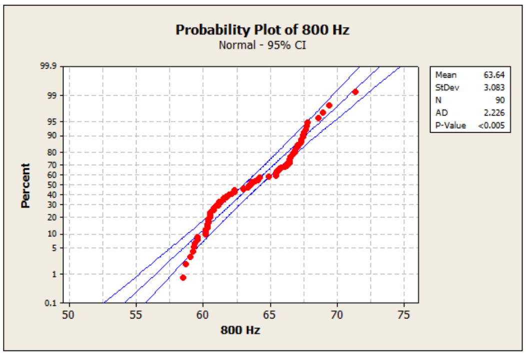

Probability Plot of 250 Hz
Normal - 95% CI
Mean 51.33
StDev 2.245
N 90
AD 0.684
P-Value 0.072
Percent
250 Hz


Probability Plot of 315 Hz
Normal - 95% CI
Mean 53.10
StDev 2.487
N 90
AD 1.380
P-Value <0.005
Percent
315 Hz


Probability Plot of 400 Hz
Normal - 95% CI
Mean 56.23
StDev 2.305
N 90
AD 0.753
P-Value 0.048
Percent
400 Hz


Probability Plot of 500 Hz
Normal - 95% CI
Mean 58.73
StDev 3.029
N 90
AD 1.767
P-Value <0.005
Percent
500 Hz


Probability Plot of 630 Hz
Normal - 95% CI
Mean 61.59
StDev 2.738
N 90
AD 0.800
P-Value 0.037
Percent
630 Hz


Probability Plot of 800 Hz
Normal - 95% CI
Mean 63.64
StDev 3.083
N 90
AD 2.226
P-Value <0.005
Percent
800 Hz

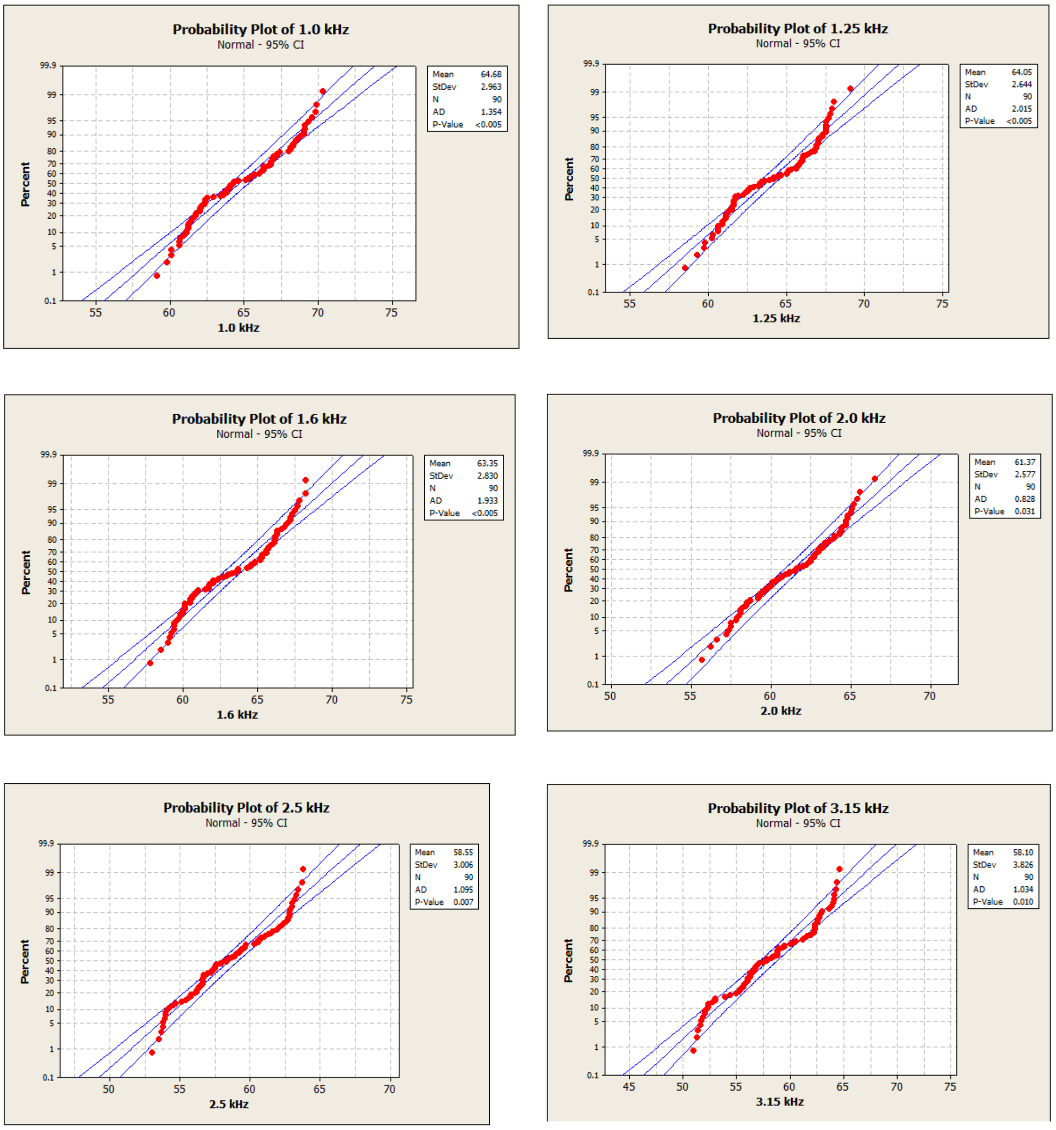


Figure 8-5: Normal Probability Plots - Concrete Floor Tests

Visual inspection indicates the concrete floor data appear to be within the confidence limits but the "s" shape of the data indicates at best it can be classed as "approximately normal".

In addition to visual inspection, there are quantitative tests for the normality of the data. One of these tests is called the Anderson Darling test. It has its own unique test statistic and a corresponding p-value. When the p-value for a quantitative test for normality is relatively large (p> 0.05) then we can accept the null hypothesis, H_o: x is normally distributed. When the p-value is relatively small (p≤ 0.01) then we must reject H_o and conclude the distribution is not normal; for the intermediate values of p (0.01<p<0.05) the test may be inconclusive.

The p-vales for the Anderson Darling test are located in the top right box on every graph.

In this study we will use the p-value provided a quantitative indicator of normality including where the data deviates from normality, and where ANOVA results need to be treated with caution. The p-values for the timber & concrete GRR are summarised in Table 8-1.

Table 8-1: D_{nT} Data – (90 test sample): σ and p-values for timber & concrete floor tests.

Floor	**Timber Test Data**		**Concrete Test Data**	
Hz	σ	**p-value**	σ	**p-value**
$D_{nT,w}$	0.99	0.005	1.97	<0.005
$D_{nT,w} + C_{tr}$	2.286	<0.005	1.54	<0.005
100 Hz	3.305	0.036	2.663	0.252
125 Hz	2.018	0.012	2.700	0.077
160 Hz	1.249	0.067	1.924	0.043
200 Hz	1.691	0.377	2.826	0.014
250 Hz	1.471	0.02	2.245	0.072
315 Hz	1.084	0.753	2.487	<0.005
400 Hz	0.8836	0.008	2.305	0.048
500 Hz	0.8378	0.633	3.029	<0.005
630 Hz	0.9848	0.288	2.738	0.037
800 Hz	1.232	0.014	3.083	<0.005
1.0 kHz	2.008	0.005	2.963	<0.005
1.25 kHz	2.206	0.055	2.644	<0.005
1.6 kHz	2.489	0.005	2.830	<0.005
2.0 kHz	2.084	0.005	2.577	0.031
2.5 kHz	2.446	0.005	3.006	0.007
3.15 kHz	4.116	0.005	3.826	0.01

Yellow boxes ($p<0.01$) reject H_o and conclude "x" is not normal. Green are inconclusive ($0.01<p<0.05$), white $p>0.05$ accept H_o and conclude distribution is normal.

Based on p-values, which indicate that the data may be non-normal caution is required when viewing the third octave band data in the range 315Hz – 3150Hz, for both floors.

There are six bands in the timber data and three in the concrete data where it can be concluded at the 95% confidence level that the data is normal. Further reference will be made to the p-values and the normality of the data if there is any notable difference in the ANOVA results across the frequency bands. The $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} data falls into the non-normal category, which might be due to low sample numbers. Normally a minimum sample size to determine normality would be 20 – 25. In the GRR design of experiment the test sample is only 6 floors which may be one reason the distribution appears to be non-normal. It is noted that the 6 floors are sampled repeated times by multiple operators. The repeated sampling of the same floors adds to the overall information on the measurement error for operator, instrument and part but does not improve the information about the probability distribution shape of the sample because the same 6 floors are tested. In addition the single figure values are an amalgam of the full frequency range results where some bands have skewed data due to the influence of background noise. The frequencies where background noise is within the 10dB range where corrections are made to the measured levels will affect the individual frequency distribution shape and may impact on the test for normality for the single figure values. It has been suggested by Osma [123] that in cases where the test sample is known to be non-normal, the average and range method ($\bar{X}$ and R) of assessing the gauge performance, is more appropriate than ANOVA. Although the $\bar{X}$ and R method breaks down the overall variation into three categories; part to part, repeatability and reproducibility it does not allow for an assessment of operator by part interaction which is an important part of this study.

There may be non-normality for both the timber floor and concrete floor DOE because of the constrained sample where the room size and volumes were as far as possible identical, in order to determine the part to part variability due to the construction of the floor. This sampling constraint artificially narrows the range and is against the requirement for a comprehensive GRR assessment. This is customarily included in the GRR study in order to detect bias, or the non linearity of the gauge [124] where a gauge may provide less variable results on a standard unit near the centre of the range rather than one at the extreme ends [68].

The range of performance was extended by incorporating the linearity test data. The same test for normality for the extended test sample is also carried out. Any data deviating away from normality is therefore identified.

As ANOVA is relatively robust in dealing with data that is non-normal, it was decided to proceed with the processing of the results for all GRR data whilst noting the graphical and numerical indicators for normality tests on the data. The ANOVA results from all data sets then were compared.

8.4 ANOVA

The results of the analysis of variance assessment using the previously described balanced two factor crossed random model with interaction (see Chapter 5: ANOVA) allow the components of variance to be determined for each GRR experiment. The data sets for the single figure sound insulation values and the third octave bands (100 – 3150Hz) are summarised from the statistical results[122, 125]. The analysis software ANOVA table printouts are given in Chapter 8 Informative Appendix:

8.4.1 Timber (lightweight) floor

The first objective of the ANOVA assessment is to determine if the measured data are revealing significant effects, which can be used to determine the components that are contributing to the variability, over and above the error we would normally expect (represented by repeatability or measurement error of the instrumentation). ANOVA generates the p-value which is usually compared with an alpha (α) of 0.05, based on a 95% confidence interval.

The tests for the main effects of the two factors "operator" and "part" is shown in Table 8-2: where $p<0.05$ there is evidence that these factors are statistically significant at the 95% level.

Table 8-2: Tests for main Effects of operators and parts: p-values for the Timber GRR

Two way ANOVA with interaction			
Timber GRR: p-values	Factor	Factor	Interaction α>0.25
	Parts	Operator	P*O
$D_{nT,w}$	0.000	0.000	0.973
$D_{nT,w} + C_{tr}$	0.018	0.000	0.519
100Hz	0.082	0.000	0.393
125Hz	0.885	0.000	0.054
160Hz	0.000	0.004	0.98
200Hz	0.000	0.004	0.585
250Hz	0.000	0.000	0.975
315Hz	0.000	0.000	0.953
400Hz	0.000	0.000	0.342
500Hz	0.000	0.000	0.655
630Hz	0.000	0.000	0.356
800Hz	0.000	0.005	0.794
1000Hz	0.000	0.004	0.557
1250Hz	0.000	0.000	0.893
1600Hz	0.000	0.000	0.28
2000Hz	0.000	0.000	0.305
2500Hz	0.000	0.000	0.005
3150Hz	0.000	0.000	0.048

The p-value follows the standard hypothesis test rationale:

$$H_o: \mu = \mu_0$$

$$H_A: \mu \neq \mu_0$$

H_o : Cells marked in yellow indicate (p>α) where α= 0.05 for most GRR experiments; p>0.05 means that we either have to accept H_o or reserve judgement.

Where the p-value for parts is greater than 0.05, the part variation is indistinguishable from repeatability variation. The converse is that if p<0.05 the gauge or, in the present case the test kit or instrumentation, can distinguish at least one part different from the rest. If p>0.05 for the operator, it indicates the reproducibility is indistinguishable from repeatability variation and it could be pooled with the error term. Only when p<0.05 can it be distinguished from repeatability.

The p-values from the timber GRR indicate that there is substantial evidence that both main effects are statistically significant at the 95% level and that their effects can be distinguished from the standard error for measurement.

There does not appear to be any evidence that the interaction term is significant. Most of the p-values are above the 0.25 term for alpha where the statistical software, using AIAG test protocols[102], pools the interaction term in with the error term thus becoming part of repeatability. See Figure 8-1 for how this works and is reported in the statistical printout.

With the knowledge that the main effects are generating statistically significant results we can start looking at the single figure sound insulation values for the timber floor tests. The $D_{nT,w}$ results are detailed in Table 8-3:

NB: for consistency the figures are reported to 3 decimal places to reflect the statistical software output. Rounding will be applied when commenting on measurement uncertainty and for data comparison as 1 decimal place is the practical measurement resolution of the instrumentation.

8.4.1.1 $D_{nT,w}$

Table 8-3: Timber Lightweight Floor - Major Components of Variance ($D_{nT,w}$)

dB (σ^2)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$	0.810	0.317	0.493	0.493	0.000	0.316	1.126
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$	0.900	0.560	0.700	0.700	0.000	0.560	1.060

where:

σ_r^2 = Repeatability (instrument Variance)

σ_R^2 = Reproducibility (Operator Variance)

σ_p^2 = Part to part variance

$\sigma_{p.o}^2$ = Operator by part variance

σ_{GRR}^2 = total gauge variance = $\sigma_r^2 + \sigma_R^2 + \sigma_{p.o}^2$ see section 8.7.2 of [126]

(NB: in timber case $\sigma_{p.o}^2 = 0$ for single figure values)

The ANOVA results show the instrumentation (represented by "repeatability"σ_r) is responsible for 0.56dB of the total standard deviation of the results; the operator (represented by σ_o) is responsible for 0.70dB of the total standard deviation. The standard deviation of the operator in this case is equivalent to the Reproducibility σ_R as the interaction term ($\sigma_{p.o}$) is zero.

The part (timber floor - represented by σ_p) is responsible for 0.56dB of the total standard deviation.

The instrumentation and the part are contributing similar amounts of uncertainty (0.56dB) to the total uncertainty relating to $D_{nT,w,}$ but the biggest contributor is the operator in this case. Note that the p-values for both the operator and part were (p=0.000) indicating that these values are highly significant.

For the $D_{nT,w}$ + C_{tr} single figure value the results are influenced by the low frequency spectrum adaptation term. See Table 8-4:

8.4.1.2 $D_{nT,w}$ + C_{tr}

Table 8-4: Timber Lightweight Floor - Major Components of Variance ($D_{nT,w}$ +C_{tr})

dB (σ^2)	σ_{GRR2}	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$ + C_{tr}	5.870	1.440	4.430	4.430	0.000	0.240	6.110
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$ + C_{tr}	2.420	1.200	2.110	2.110	0.000	0.490	2.470

The part to part component (represented by σ_p) for $D_{nT,w}$ + C_{tr} has a standard deviation of 0.49dB. This is similar to the $D_{nT,w}$ single figure value for the part to part variability (0.56dB).

The ANOVA results show the instrumentation (representing "repeatability" σ_r) is responsible for 1.20dB of the total standard deviation of the results; the operator (represented by σ_o) is responsible for 2.11dB of the total standard deviation. Both these results show at least twice the variability of the $D_{nT,w}$ single figure value for the instrument and the operator which follows the basic analysis results for timber floors where the relatively poor low frequency performance increases the standard deviation when the spectrum adaptation term is applied.

The standard deviation of the operator in this case is equivalent to the Reproducibility σ_R as the interaction term ($\boldsymbol{\sigma_{p.o}}$) is classed as not statistically significant and set to zero.

The part to part variability (represented by $\boldsymbol{\sigma_p}$) is responsible for 0.49dB of the total standard deviation.

The operator is still the major contributor to the overall measured variability with a standard deviation of 2.11dB.

It is possible to cardinally rank the importance of the components depending on size. For both single figure values on lightweight timber floors the operator is the factor that contributes most to the uncertainty followed by the instrument and then the part.

Table 8-5: measurement variability due to defined factors - ordered by magnitude

Measurand		$D_{nT,w}$	$D_{nT,w} + C_{tr}$
Order	Factor	dB	dB
1	Operator σ_o	0.7dB	2.1dB
2	Instrument σ_r	0.6dB	1.2dB
3	Part σ_p	0.6dB	0.5dB

Ideally the part measured would contribute the most and the operator and instrument contributions would be low. In GRR studies in the motor industry the AIAG[102] have a broad set of compliance criteria for the measurement system based on the percentage contributions of operator, instrumentation and part to the total variability, these are shown in Table 8-6:

Table 8-6: AIAG Measurement System Analysis Criteria (as % of total variance)

%GRR	Criteria
<10%	Measurement system is acceptable
10% - 30%	Measurement system may be acceptable
>30%	Measurement system needs improvement

Note: %GRR = $\%\sigma_r^2 + \%\sigma_R^2$

The percentage contributions to overall variability are shown in Table 8-7 for the timber floor case.

Table 8-7: Showing components of variance as percentage of total variability - Timber Floor GRR

Measurand	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$	71.96%	28.15%	43.81%	43.81%	0.00%	28.04%	100.00%
$D_{nT,w} + C_{tr}$	96.11%	23.59%	72.52%	72.52%	0.00%	3.89%	100.00%

Using the timber floor results and applying the AIAG criteria the conclusion would be that the measurement system needs improvement and the operators being the largest contributor and significantly above the 30% upper threshold need retraining! The contribution of the instrumentation in both cases gives a percentage contribution between 20-30%, enough in itself to create doubt about the usefulness of the measurement system.

Where the GRR is based on a constrained sample the part to part variability is likely to be small because the construction of the floor and room size have been selected to be similar. Normally under GRR DOE a sample with the full measurement range of the instrumentation would be chosen[68]. The relatively large variability in percentage terms of the instrumentation and the operator is therefore understandable. The percentage variability based on the component divided by the total variability is therefore less important in this context though the variabilities, as measured in decibels are significant because they show that the overall variability (σ_p) due to the part to part variability of the construction (workmanship) is 0.56dB $D_{nT,w}$ and 0.49dB $D_{nT,w}$ + C_{tr} which is relatively small compared with the total variability (σ_{Total}) of 1.06dB $D_{nT,w}$ and 2.47dB $D_{nT,w}$ + C_{tr}. This shows that the sound insulation performance of the floors is consistent when based on a single figure value.

8.4.2 Frequency Data - Results

ANOVA is performed across the frequency range for the D_{nT} values. This gives a much more detailed view of where the major regions of variability in the data lie. The frequency result tables are summarised below in Table 8-8 & Table 8-9 and the variances plotted in Figure 8-6.

Table 8-8: Timber Lightweight Floor (variance) - Major Components of Variance Frequency Data (D_{nT})

D_{nT} (var)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
100Hz	12.244	3.742	8.502	8.502	0.000	0.362	12.607
125Hz	4.612	2.031	2.582	2.091	0.491	0.000	4.612
160Hz	1.013	0.921	0.091	0.091	0.000	0.670	1.683
200Hz	1.907	1.551	0.356	0.356	0.000	1.212	3.119
250Hz	1.580	0.787	0.793	0.793	0.000	0.871	2.451
315Hz	0.992	0.432	0.560	0.560	0.000	0.343	1.335
400Hz	0.450	0.313	0.136	0.136	0.000	0.424	0.873
500Hz	0.494	0.317	0.177	0.177	0.000	0.287	0.781
630Hz	0.571	0.342	0.230	0.230	0.000	0.525	1.096
800Hz	0.418	0.359	0.059	0.059	0.000	1.317	1.735
1000Hz	1.057	0.855	0.201	0.201	0.000	3.578	4.634
1250Hz	2.224	0.770	1.454	1.454	0.000	3.466	5.690
1600Hz	6.424	1.584	4.840	4.840	0.000	0.828	7.252
2000Hz	4.016	1.329	2.687	2.687	0.000	0.996	5.012
2500Hz	4.111	1.079	3.033	2.539	0.494	2.810	6.921
3150Hz	9.750	2.493	7.257	6.626	0.631	10.054	19.804

Table 8-9: Timber Lightweight Floor (s.d.) - Major Components of Variance Frequency Data (D_{nT})

D_{nT} (s.d.)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
100Hz	3.50	1.93	2.92	2.92	0.00	0.60	3.55
125Hz	2.15	1.43	1.61	1.45	0.70	0.00	2.15
160Hz	1.01	0.96	0.30	0.30	0.00	0.82	1.30
200Hz	1.38	1.25	0.60	0.60	0.00	1.10	1.77
250Hz	1.26	0.89	0.89	0.89	0.00	0.93	1.57
315Hz	1.00	0.66	0.75	0.75	0.00	0.59	1.16
400Hz	0.67	0.56	0.37	0.37	0.00	0.65	0.93
500Hz	0.70	0.56	0.42	0.42	0.00	0.54	0.88
630Hz	0.76	0.58	0.48	0.48	0.00	0.72	1.05
800Hz	0.65	0.60	0.24	0.24	0.00	1.15	1.32
1000Hz	1.03	0.92	0.45	0.45	0.00	1.89	2.15
1250Hz	1.49	0.88	1.21	1.21	0.00	1.86	2.39
1600Hz	2.53	1.26	2.20	2.20	0.00	0.91	2.69
2000Hz	2.00	1.15	1.64	1.64	0.00	1.00	2.24
2500Hz	2.03	1.04	1.74	1.59	0.70	1.68	2.63
3150Hz	3.12	1.58	2.69	2.57	0.79	3.17	4.45

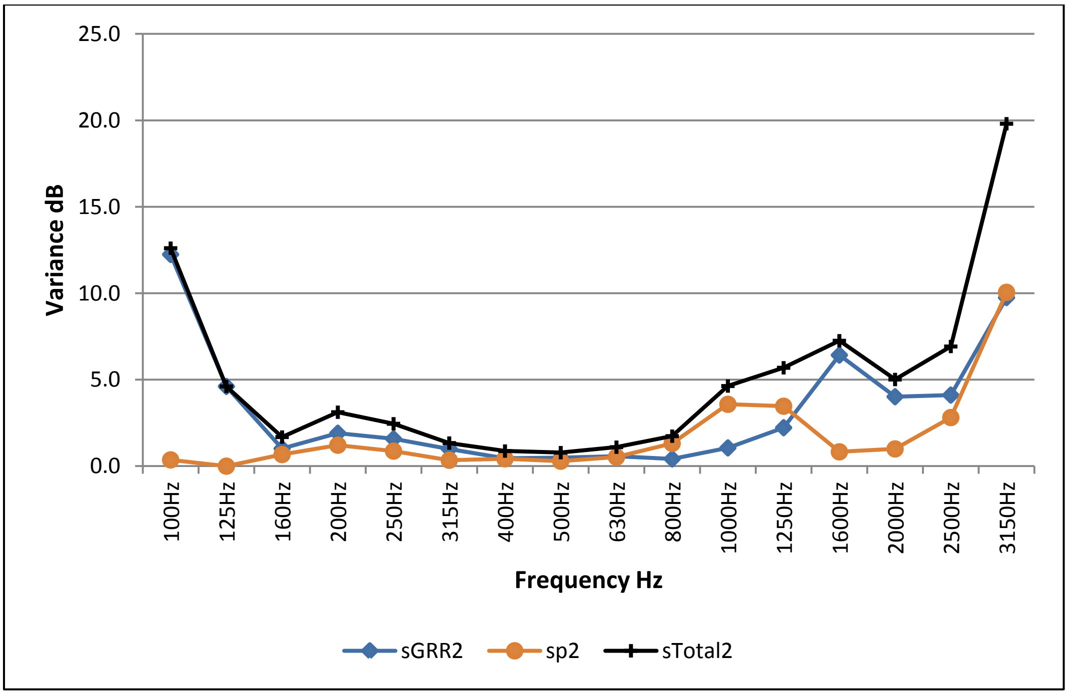


Figure 8-6: Timber Lightweight Floor - Components of Variance - σ_{GRR}^2, σ_p^2, σ_{Total}^2

In Figure 8-6 the total variance is characterised by a "U" shaped curve showing greater variance at the low and high ends of the frequency range. There is a clear indication that the gauge is the primary component responsible for the high variance at the low frequency bands 100Hz & 125Hz. The contribution from the part is small in comparison. From 160Hz to 315Hz the contribution to total variance from the gauge is slightly higher, between 400 -630Hz the variance of both is virtually equivalent and relatively small. From 800Hz to 1250Hz the part contributes more to the total variance than the gauge, at 1600Hz the part contribution drops and the gauge contribution rises significantly to dominate the contribution and also for the 2000Hz and 2500Hz bands. At 3150Hz the part and the gauge variances rise significantly and both contribute a similar amount to the total.

The important conclusions to be drawn from this graphical representation of the components of variance are at the 100Hz - 125Hz and 3150Hz bands. The part contributes very little to the overall variance at low frequency and it is noted at these frequencies, the p-values were not significant see Table 8-10.

Table 8-10: Section of table showing p-values for parts at low frequencies 100-125Hz

	Parts
$D_{nT,w}$	0.000
$D_{nT,w} + C_{tr}$	0.018
100Hz	0.082
125Hz	0.885

This means the part variability was indistinguishable from the measurement error (repeatability),

At 100 – 125Hz it is the gauge (combining both r+R) that is responsible for the high level of variance not the construction of the floor. The variance of the gauge also reflects the expected variance at the low frequency part of the spectrum where there is a non-diffuse field.

At high frequency where there is a significant rise in variability in the measured data both the gauge and the part contribute equal amounts. As previously discussed in Table 7-6 there may be a combination of factors contributing to the relatively high variance levels. The part variance may be due to the high performance of the timber floor at high frequency. This, combined with background noise, means variance increases after 2000Hz. From the gauges perspective this could be interpreted to mean that the test kit could not output sufficient sound power or the background noise is increasingly influential, resulting in increased variance.

The gauge contribution can be broken down further by looking at the repeatability, reproducibility, operator and interaction terms individually. These are plotted in Figure 8-7.

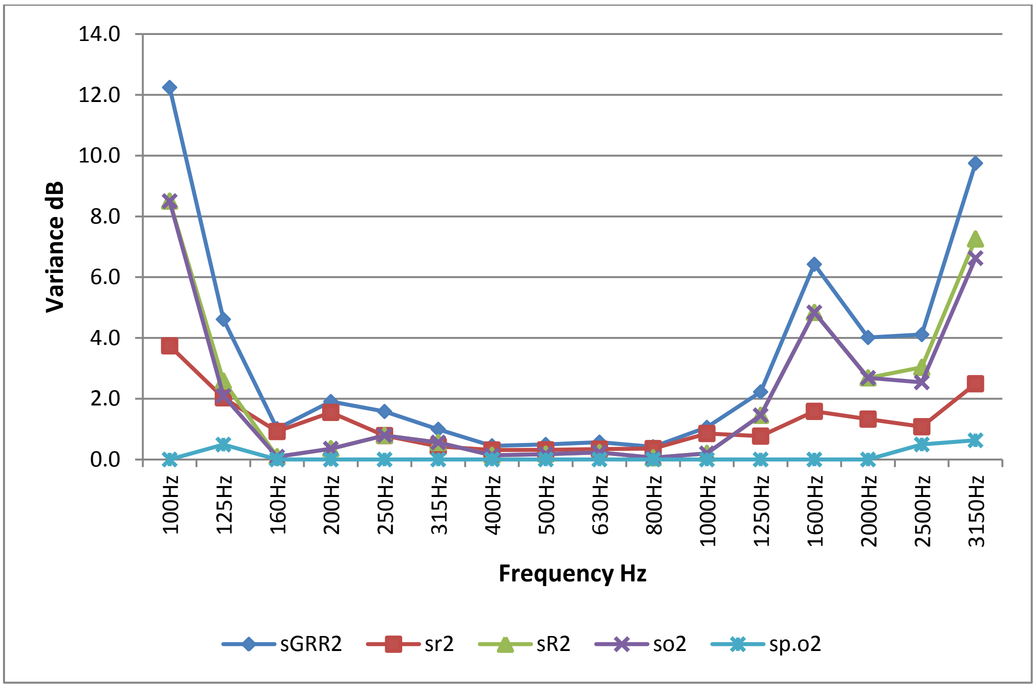


Figure 8-7: Lightweight Floor - Components of Variance - Timber GRR. σ_{GRR}^2, σ_r^2, σ_R^2, σ_o^2, $\sigma_{p.o}^2$

Both repeatability and reproducibility components follow a “U” shaped profile although the repeatability contribution to the Gauge overall variance is generally lower than the reproducibility contribution.

The reproducibility contribution is dominated by the operator component and apart from a couple of minor interaction effects the reproducibility curve duplicates the operator curve.

Repeatability is dominant at 160Hz – 200Hz, between 250Hz-800Hz the r & R contributions are similar and both r & R contributions rise after 1000-1250Hz but R is dominant. The increase in variance after 1250Hz for both r & R is likely to be due to the influence of background noise.

8.4.3 Concrete (Heavyweight) Floor

The tests for the main effects of the operator and part is shown in Table 8-11: for p<0.05 there is evidence that these factors are statistically significant at the 95% level.

Table 8-11: Tests for main Effects of operators and parts: p-values for the Timber GRR

Two way ANOVA with interaction			
Concrete GRR: p-values	Factor	Factor	Interaction α>0.25
	Parts	Operator	P*O
$D_{nT,w}$	0.000	0.001	0.066
$D_{nT,w} + C_{tr}$	0.001	0.002	0.003
100Hz	0.000	0.000	0.063
125Hz	0.049	0.004	0.014
160Hz	0.248	0.158	0.000
200Hz	0.000	0.016	0.000
250Hz	0.000	0.354	0.002
315Hz	0.000	0.075	0.198
400Hz	0.000	0.011	0.053
500Hz	0.000	0.042	0.058
630Hz	0.000	0.007	0.056
800Hz	0.000	0.007	0.338
1000Hz	0.000	0.000	0.005
1250Hz	0.000	0.000	0.005
1600Hz	0.000	0.015	0.039
2000Hz	0.000	0.001	0.644
2500Hz	0.000	0.058	0.065
3150Hz	0.000	0.038	0.183

H_o : Cells marked in yellow indicate ($p>\alpha$) where α= 0.05 for most GRR experiments; p>0.05 means that we either have to accept H_o or reserve judgement.

The operator has one p>0.05 at 160Hz. The part has 4 frequencies where p>0.05 at 160Hz, 250Hz, 315Hz and 2500Hz. At these frequencies the part and operator variance is indistinguishable from repeatability. For the other frequencies, the p-values from the concrete GRR indicate that there is substantial evidence that both main effects are statistically significant at the 95% level and that their effects can be distinguished from the standard error for measurement.

Only 2 of the p-values at 800Hz and 2000Hz for the part by operator interaction term are above α =0.25, indicating the interaction term is significant, and retained at all other frequencies.

The p-values indicate that there are some significant effects which can be investigated. The $D_{nT,w}$ results are detailed in Table 8-12 :

8.4.3.1 $D_{nT,w}$

Table 8-12: Concrete Heavyweight Floor - Major Components of Variance ($D_{nT,w}$)

dB (σ^2)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$	0.883	0.489	0.394	0.286	0.109	3.379	4.263
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$	0.940	0.700	0.630	0.530	0.330	1.840	2.060

NB: reported to 3 decimal places in line with statistical software output but rounded, where appropriate for discussion below.

The ANOVA results show the instrumentation (representing "repeatability"$\boldsymbol{\sigma_r}$) is responsible for 0.7dB of the total standard deviation of the results, the operator (represented by $\boldsymbol{\sigma_o}$) is responsible for 0.53dB of the total standard deviation. The standard deviation of the operator in this case is not equivalent to the Reproducibility $\boldsymbol{\sigma_R}$ as the interaction term is significant and must be accounted for separately. The interaction ($\boldsymbol{\sigma_{p.o}}$) contribution is 0.33dB.

The part (Concrete floor - represented by σ_p) is responsible for 1.84dB of the total standard deviation which is the largest component.

The instrumentation and the operator are contributing similar amounts of uncertainty (0.53 dB & 0.7dB respectively) to the total uncertainty relating to $D_{nT,w}$ but the biggest contributor is the part because the room sizes differed.

8.4.3.2 $D_{nT,w} + C_{tr}$

For the $D_{nT,w}$ + C_{tr} single figure value the results are affected by the low frequency spectrum adaptation term. See Table 8-13:

Table 8-13: Concrete Heavyweight Floor - Major Components of Variance ($D_{nT,w}$ + Ctr)

dB (σ^2)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$ + C_{tr}	1.791	0.800	0.991	0.573	0.417	0.829	2.620
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$ + C_{tr}	1.340	0.890	1.000	0.760	0.650	0.910	1.620

The concrete $D_{nT,w}$ + C_{tr} results show that the components of variability are relatively low within 1dB standard deviation. The instrumentation (representing "repeatability" σ_r) has a standard deviation of 0.89dB; the operator (represented by σ_o) has a standard deviation of 0.76dB. The standard deviation of the operator in this case is different from the reproducibility σ_R as the interaction term ($\sigma_{p.o}$) is 0.65dB.

The part to part component (represented by σ_p) for $D_{nT,w}$ + C_{tr} has a standard deviation of 0.91dB. This is significantly lower than the $D_{nT,w}$ single figure value for the part to part variability (1.84dB) which follows the basic analysis results for concrete floors.

The part is the major contributor to the overall measured variability with a standard deviation of 0.91dB. The variability of the instrumentation is very similar in this case, 0.89dB, the operator has the lowest standard deviation of 0.76dB.

The components are ranked in order of size. This is shown in Table 8-14 for both D_{nTw} and $D_{nT,w}$ + C_{tr} on heavyweight concrete floors.

Table 8-14: Concrete Floor - measurement variability due to defined factors - ordered by magnitude

Measurand		$D_{nT,w}$	$D_{nT,w}$ + Ctr
Order	Factor	dB	dB
1	Part σ_p	1.8dB	0.9dB
2	Instrument σ_r	0.7dB	0.9dB
3	Operator σ_o	0.5dB	0.8dB

In the concrete floor GRR, the part has the highest standard deviation, then the instrument and the operator, though the data tell us that the parts influence is proportionally greater in the $D_{nT,w}$ case (1.84dB) compared with repeatability and operator components than in the $D_{nT,w}$ + Ctr case where they are almost equal in size. The order is the reverse of the timber floor GRR. It is noted that the concrete floor GRR had non-identical test rooms which varied in size and volume (unlike the timber floor GRR where they were similar), the small changes in size and volume of the source and receiver rooms is likely to be the cause of this difference. The operator and instrument have very similar standard deviations for $D_{nT,w}$ + C_{tr} which, when rounded to 1 decimal place to reflect the instrument measurement resolution, are 0.8dB and 0.9dB respectively.

Ideally, in AIAG measurement system analysis, the part measured would contribute the most and be dominant and the operator and instrument contributions would be low. The AIAG [102] criteria can be compared with the percentage contributions of operator, instrumentation and part to the total variability, these are shown in Table 8-15:

Table 8-15: Showing components of variance as percentage of total variability - Concrete Floor GRR

Measurand	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$	20.72%	11.47%	9.25%	6.70%	2.55%	79.28%	100.00%
$D_{nT,w}$ + Ctr	68.35%	30.53%	37.81%	21.88%	15.93%	31.65%	100.00%

The percentage contribution of the gauge is lower than in the timber GRR but the conclusion is the same that the measurement system needs improvement in particular for $D_{nT,w} + C_{tr}$. This is likely to be influenced by the constrained sample chosen so the results in Table 8-15 so an AIAG comparison should be treated as indicative only. The most important results to take from this analysis are the standard deviations and variance of the individual sources of variability in dB, and the comparison of these values with the lightweight timber floor data.

8.4.4 Frequency Data

The results between 100Hz – 3150Hz are summarised in Table 8-16 & Table 8-17 and the variance plotted in Figure 8-8.

Table 8-16: Concrete heavyweight floor (variance) - Major components frequency data (D_{nT})

D_{nT} (var)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
100Hz	5.760	2.916	2.844	2.184	0.660	2.092	7.852
125Hz	6.942	3.666	3.276	1.930	1.346	0.886	7.828
160Hz	3.631	1.989	1.642	0.287	1.354	0.184	3.815
200Hz	3.436	1.367	2.069	0.835	1.235	5.621	9.058
250Hz	2.789	1.756	1.034	0.044	0.989	2.705	5.495
315Hz	1.682	1.378	0.303	0.153	0.151	5.381	7.062
400Hz	1.781	1.142	0.639	0.362	0.277	4.283	6.064
500Hz	1.844	1.293	0.551	0.248	0.303	8.763	10.607
630Hz	2.489	1.547	0.942	0.572	0.370	6.082	8.570
800Hz	2.910	2.350	0.560	0.560	0.000	7.952	10.862
1000Hz	2.176	0.896	1.280	0.858	0.422	8.041	10.217
1250Hz	1.997	0.829	1.168	0.781	0.387	6.110	8.107
1600Hz	1.258	0.796	0.462	0.244	0.218	8.075	9.333
2000Hz	1.093	0.826	0.267	0.267	0.000	6.643	7.736
2500Hz	0.962	0.695	0.267	0.112	0.155	9.610	10.571
3150Hz	0.772	0.604	0.168	0.097	0.071	16.479	17.251

Table 8-17: Concrete heavyweight floor (s.d.) - Major components frequency data (D_{nT})

D_{nT} (s.d.)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
100Hz	2.40	1.71	1.69	1.48	0.81	1.45	2.80
125Hz	2.63	1.91	1.81	1.39	1.16	0.94	2.80
160Hz	1.91	1.41	1.28	0.54	1.16	0.43	1.95
200Hz	1.85	1.17	1.44	0.91	1.11	2.37	3.01
250Hz	1.67	1.32	1.02	0.21	0.99	1.64	2.34
315Hz	1.30	1.17	0.55	0.39	0.39	2.32	2.66
400Hz	1.33	1.07	0.80	0.60	0.53	2.07	2.46
500Hz	1.36	1.14	0.74	0.50	0.55	2.96	3.26
630Hz	1.58	1.24	0.97	0.76	0.61	2.47	2.93
800Hz	1.71	1.53	0.75	0.75	0.00	2.82	3.30
1000Hz	1.48	0.95	1.13	0.93	0.65	2.84	3.20
1250Hz	1.41	0.91	1.08	0.88	0.62	2.47	2.85
1600Hz	1.12	0.89	0.68	0.49	0.47	2.84	3.05
2000Hz	1.05	0.91	0.52	0.52	0.00	2.58	2.78
2500Hz	0.98	0.83	0.52	0.33	0.39	3.10	3.25
3150Hz	0.88	0.78	0.41	0.31	0.27	4.06	4.15

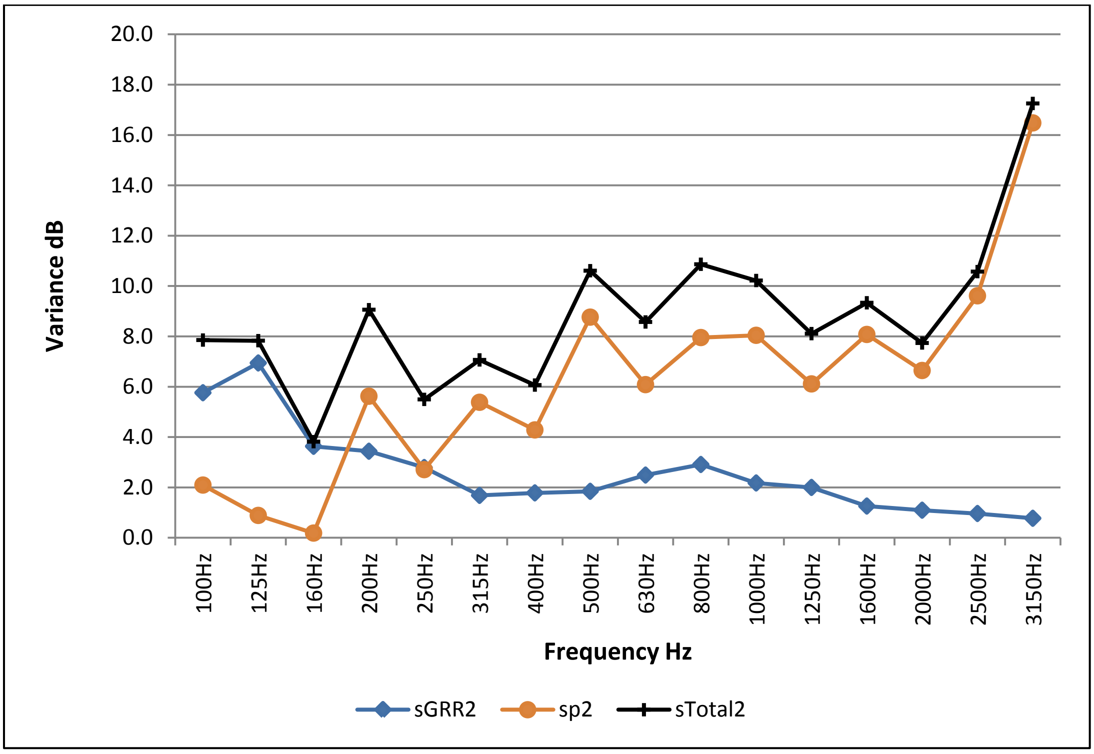


Figure 8-8: Concrete Heavyweight Floor - Components of Variance - σ_{GRR}^2, σ_p^2, σ_{Total}^2

In Figure 8-8 the total variance is characterised by a gradually rising line which rises steeply after 2000Hz. The gauge is the primary component at low frequency 100Hz –

160Hz as was the case for the timber floors but above 160Hz the part is more influential and dominant at 315Hz and above.

The p-value for the part at 160Hz was not significant meaning the part variability was indistinguishable from the measurement error (repeatability).

The gauge contribution can be broken down further by looking at the repeatability, reproducibility, operator and interaction terms individually. These are plotted in Figure 8-9.

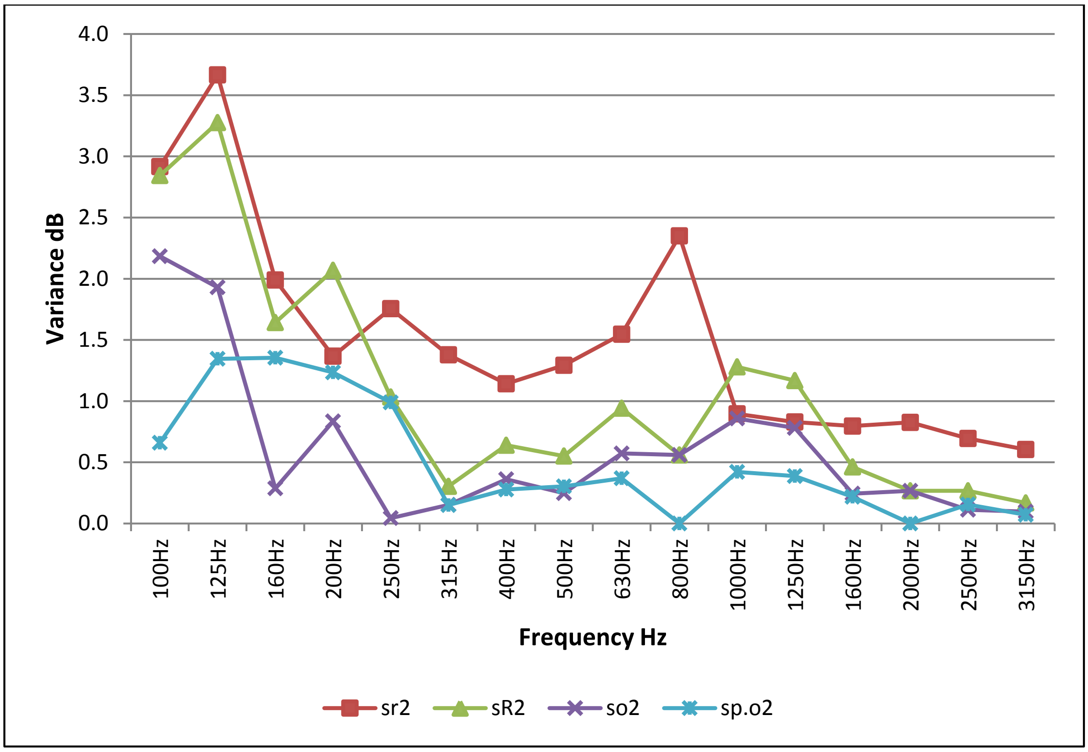


Figure 8-9: Concrete Heavyweight Floor - Components of Variance - σ_{GRR}^2, σ_r^2, σ_R^2, σ_o^2, $\sigma_{p.o}^2$

Both repeatability and reproducibility components follow a falling trend as frequency increases.

The repeatability component of variance is the largest value across the frequency range; the operator contribution is similar in only the 1000 – 1250Hz bands. It is noted that there is a significant interaction term in this data which has a higher variance than the operator term between 160Hz – 315Hz bands and, when it is combined with the operator term to form reproducibility, it exceeds the repeatability term in the 200Hz, 1000Hz and 1250Hz bands.

The p-value for the operator at 160Hz, 250Hz-315Hz and 2500Hz was not significant meaning the operator variability was indistinguishable from the measurement error (repeatability).

The repeatability contribution to the Gauge overall variance is generally higher than the reproducibility contribution. The instrumentation influence is dominant above the operator.

8.4.5 Timber GRR v Concrete GRR

The individual variance terms for the timber and concrete floor GRR studies are assessed individually starting with part to part variance (σ_p^2):

8.4.5.1 Part to Part Variance:

A graphical representation of the part to part variance for the timber and concrete data is shown in Figure 8-10.

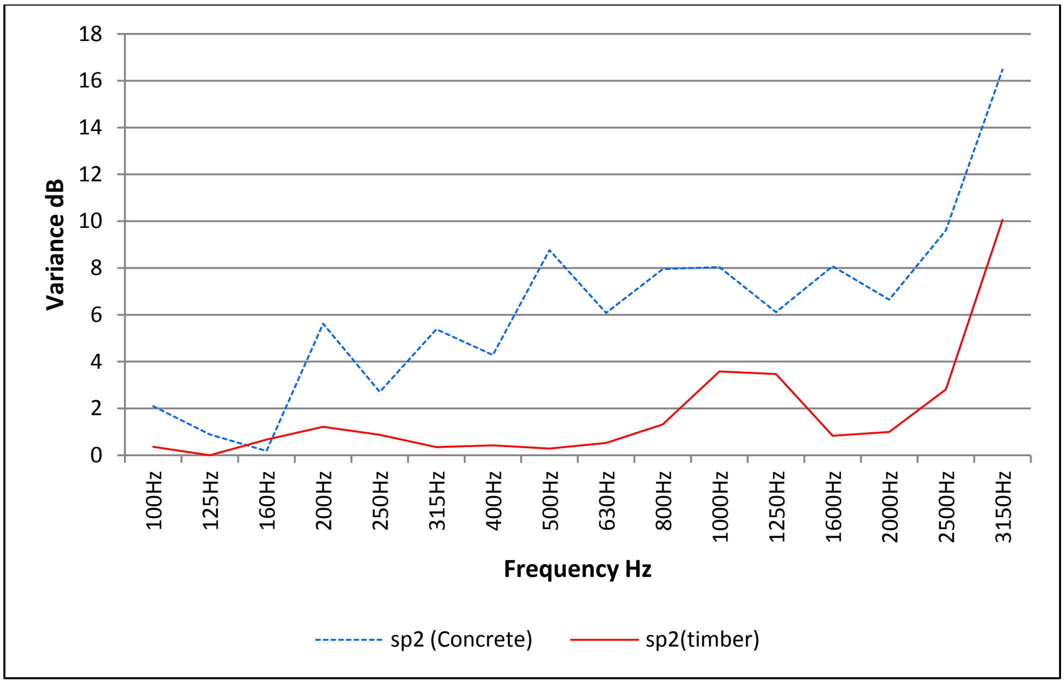


Figure 8-10: Part to Part Variance components - Timber v Concrete GRR

Apart from the 160Hz band, the part to part variance (in dB) in the timber GRR is significantly smaller across the frequency range than the values for the concrete GRR. Both part to part terms rise significantly after 2000Hz indicating that there is significant contribution to variability at high frequency in both floor types. Excluding 100Hz and 160Hz, the concrete floor variance is between 2 and 8.5dB higher than that for timberthe reason being the difference in test room sizes in the concrete test sample. To demonstrate how this variability occurs the concrete GRR data for 3150hz is used as an example.

The part variability at 3150Hz (between sample floors) is represented by the variance of sample mean $S_{\bar{y}}^2$. A high value for this term means that the parts themselves are different. The variance of sample mean for 3150Hz is 5.9dB which is relatively large. The error variance S_E^2 (repeatability) at this frequency band is low by comparison (0.7dB). The variance of sample mean is dominant. We conclude from this the floors were performing differently at 3150Hz i.e. it was not caused by measurement error. This data can be compared to the 160Hz data where the measurement error (or error variance) is 3.4dB and the variance of the sample mean is 0.21dB. In this situation we would conclude it is difficult to distinguish the part variability from the natural error in the measurement. These data are shown in Table 8-18.

Table 8-18: Heavyweight floor test sample for 3150Hz & 160Hz band showing individual repeat measurements made on each sample by all 5 operators

3150Hz			Treatment (k)						
Engineer	1	2	3	4	5	6			
SM	62.7	62.3	56	55.8	52.4	58.8			
SM	60	64.6	56.8	57.8	51	58.8			
SM	61.6	64.1	56.7	56	53.9	59.3			
SC	62.5	60.5	56.4	56.3	53	58.9			
SC	61.9	62.9	56.9	55.6	51.4	58.5			
SC	60.4	62.6	57	56.8	52	58.8			
SP	62.7	64.3	57.1	56.7	52.3	58.9			
SP	62.3	63.6	56.3	55.6	51.3	59.5			
SP	62.3	64	55.2	56.6	51.7	58.8			
DJ	61.2	62.3	54.4	55.3	51.9	58.2			
DJ	62.2	62.6	55	57.8	51.6	57.6			
DJ	62.1	61.3	55.7	56.3	52	58.3			
BW	63	64.1	55.4	56	52.8	59.5			
BW	62.2	63.9	57.5	55.9	53	58.8			
BW	61.5	64.4	56.2	56.3	52.4	60.2	Error Variance	Grand Mean	Var Sample Mean
							S_E^2	y bar bar	S_y^2
mean	61.9	63.2	56.2	56.3	52.2	58.9		58.1	
Var	0.7	1.5	0.8	0.5	0.6	0.4	0.7		
Variance of sample mean =	14.48225	25.65985309	3.7163272	3.1723568	35.05956	0.575912			5.90
F Ratio =	47.34608								

6 floor sample

Error (r) variance is rel. small

Variance or sample means significantly different between floors

160Hz			Treatment (k)						
Engineer	1	2	3	4	5	6			
SM	44.4	42	45.8	45.7	44.6	43.7			
SM	44.2	41.6	44	46.1	43.5	44.7			
SM	42.8	39	43.7	43.9	43.1	42.6			
SC	39.5	41.7	44.7	39.5	39.4	42.7			
SC	37.5	44.6	44.9	44.3	42.5	39.1			
SC	39.3	41.6	44.1	43	41.8	40.6			
SP	43.4	44.8	42	42.6	43.5	45.9			
SP	42.6	45.4	43.2	40.4	43.1	46.2			
SP	43.8	43.8	45.1	41.1	43.9	43.6			
DJ	40.1	43.7	45.8	42.6	42.7	44.7			
DJ	40.6	45.2	42.7	42.6	43.3	46			
DJ	41.6	42.5	43.9	43.9	43.9	43.5			
BW	41.6	43.2	42.5	43.1	42.3	42.6			
BW	41	41.1	45	42.5	44.5	45.7			
BW	46.6	38.8	43	40.7	43.2	44.5	Error Variance	Grand Mean	Var Sample Mean
							S_E^2	y bar bar	S_y^2
mean	41.9	42.6	44.0	42.8	43.0	43.7		43.0	
Var	5.6	4.2	1.4	3.4	1.6	4.1	3.4		
Variance of sample mean =	1.180844	0.1764	1.0133778	0.0484	5.05E-29	0.5184			0.21
F Ratio =	0.370623								

Error (r) variance is rel. Large due to non-diffuse field

Variance or sample means very similar between floors

In summary; for 3150Hz data the variance between the sample means is relatively large in comparison to the repeatability S_E^2. It is therefore relatively easy to distinguish. The part to part term is the major influence in the total variance for this floor at this frequency.

For the 160Hz band the repeatability S_E^2 is relatively large due to the natural variability of sampling sound pressure level in a non diffuse field. The variance of sample means S_y^2 is relatively low, all the floors perform in a similar way at this frequency.

We can say how important the contribution of this component of variance is by representing the variance as a percentage of the total variance.

Based on our example above, the part to part variance at 3150Hz for the concrete floors is 96% of the total variance, at 160Hz it is 5%. This is illustrated in Figure 8-11:

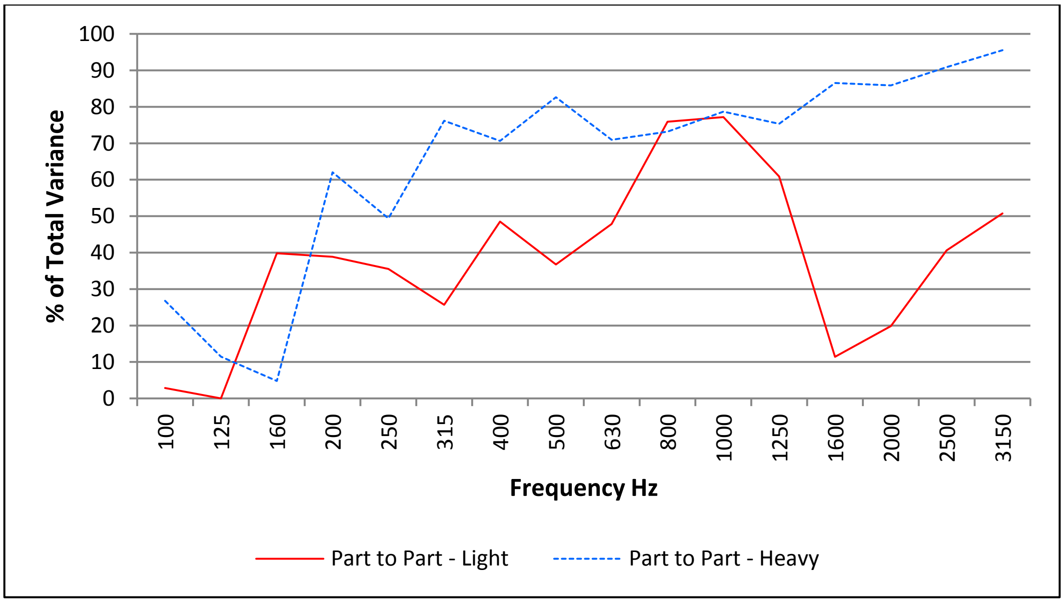
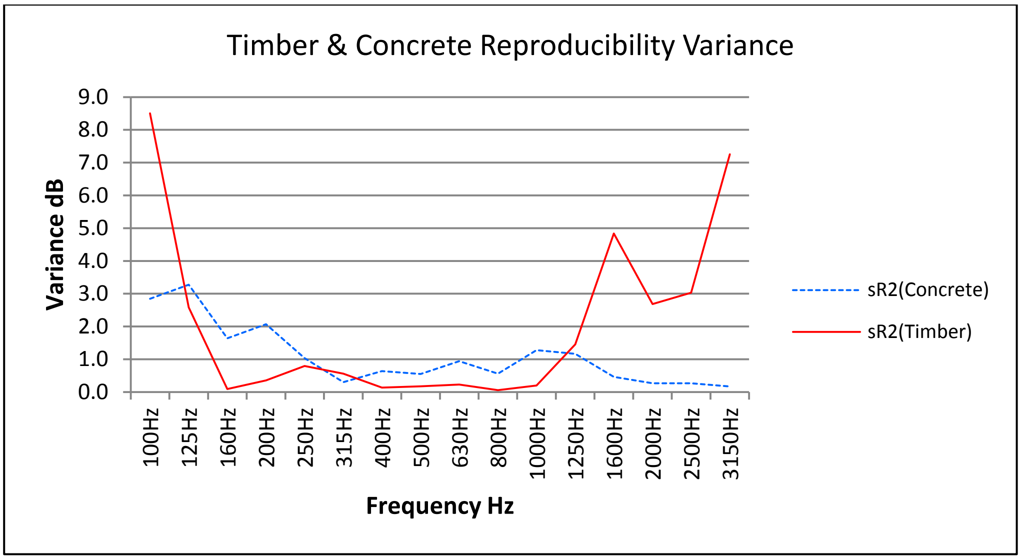


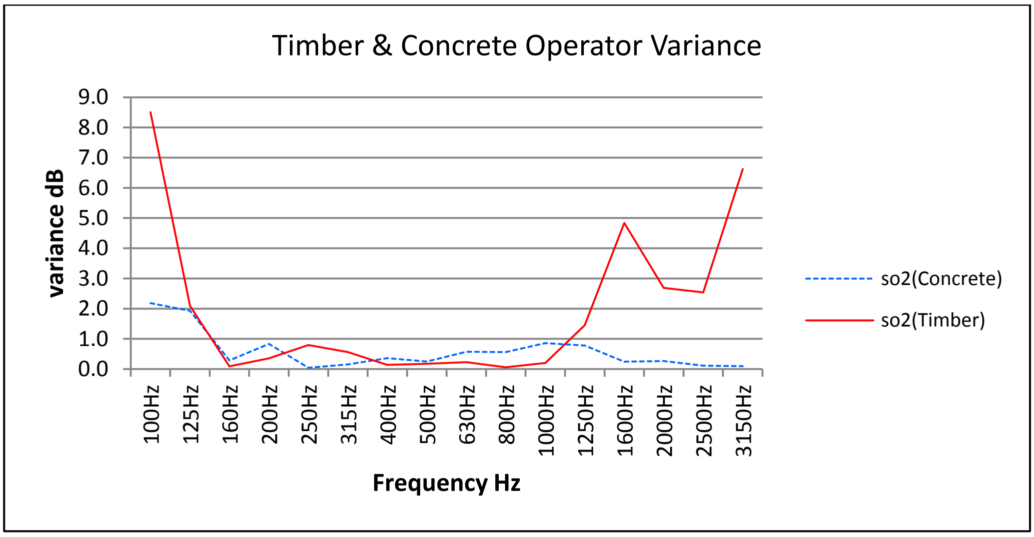
Figure 8-11: Timber & Concrete Part to Part Variance as a Percentage of Total Variance

Below 160Hz the part to part variance of the concrete floor contributed 27% or less of the total variance measured. Above 160Hz the part contributed more than 50% of the total variance measured, rising to 96% at 3150Hz. The timber GRR part to part variance contribution varies across the frequency range. This may be due to the constrained test sample rooms minimising the part variability. With room size fixed there is more scope for other external influences, such as the variability of the rooms' sound field and site background noise together with the limitations of the test kit to dominate and contribute a greater proportion of the total variance. The timber floor part to part variance was less than 4% at 100-125Hz, 40% of the total variance at 160Hz rising to 77% 1000Hz. It falls to 11% at1600Hz before rising again to 51% at 3150Hz.

8.4.5.2 Reproducibility (R)

The reproducibility component of variance σ_R^2 which equates to the operator variance σ_O^2 in the reduced model and may, in some cases, feature interaction between operator and part σ_{Op}^2 is plotted for the timber and concrete data in Figure 8-12:





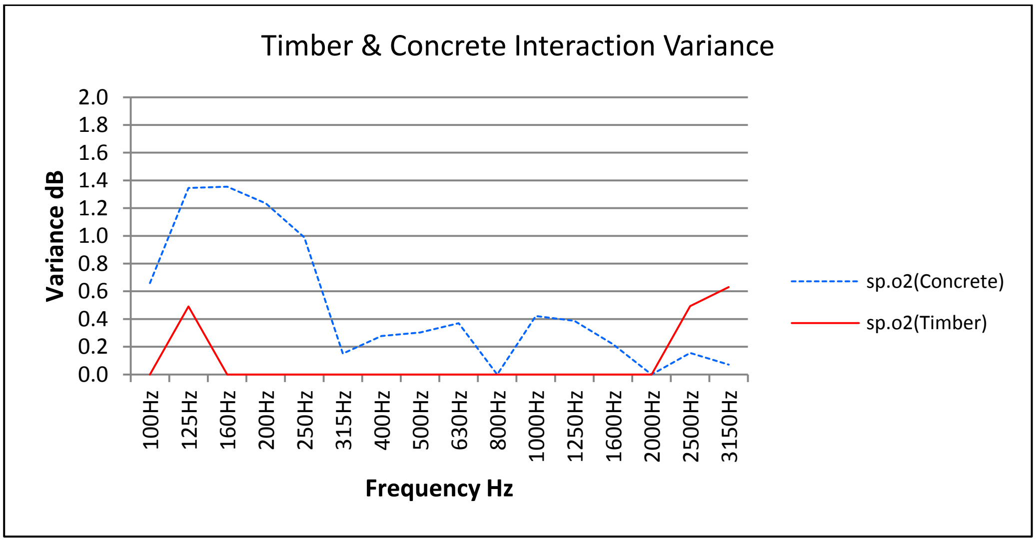


*Figure 8-12: Reproducibility, Operator and Operator*Part, Variance contributions*

The timber floor GRR has an Operator variance σ_O^2 of 8.5dB at 100Hz, falling to 2.1dB at 125Hz and is less than 1dB between 160Hz – 1000Hz. Above 1250Hz it varies between 1.4 – 6.6dB. The concrete floor GRR operator variance peaks at 100Hz (2.2dB) falling to 1.9dB at 125Hz and is less than 1dB between 160-3150Hz. There is an interaction variance σ_{Op}^2 , contribution of 0.5 – 0.6dB in the 125Hz, 2500Hz and 3150Hz bands in the timber GRR. Interaction affects all but the 800Hz and 2000Hz bands in the concrete GRR and generally has a falling trend from 1.4dB at 160Hz to 0.1dB at 3150Hz. The concrete floor GRR interaction terms have more influence on the reproducibility variance term σ_R^2 at low frequency between 125 – 250Hz where the range is 1.6 – 3.2dB approximately 1.2 – 1.4dB higher than the "Operator" variance term. The interaction term has an influence less than 1dB across the other frequency bands. It is notable, that the interaction between each of the factors, operator and part was not nil. This contravenes the assumption required in GUM[6] that the contributions from input variables are independent.

The reproducibility variance σ_R^2 is detailed as a percentage of total variance in Table 8-19:

Table 8-19: Reproducibility - Proportion of Total Variance (%)

%	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
Timber	67%	56%	5%	11%	32%	42%	16%	23%	21%	3%	4%	26%	67%	54%	44%	37%
Concrete	36%	42%	43%	23%	19%	4%	11%	5%	11%	5%	13%	14%	5%	3%	3%	1%

This data is plotted in Figure 8-13:

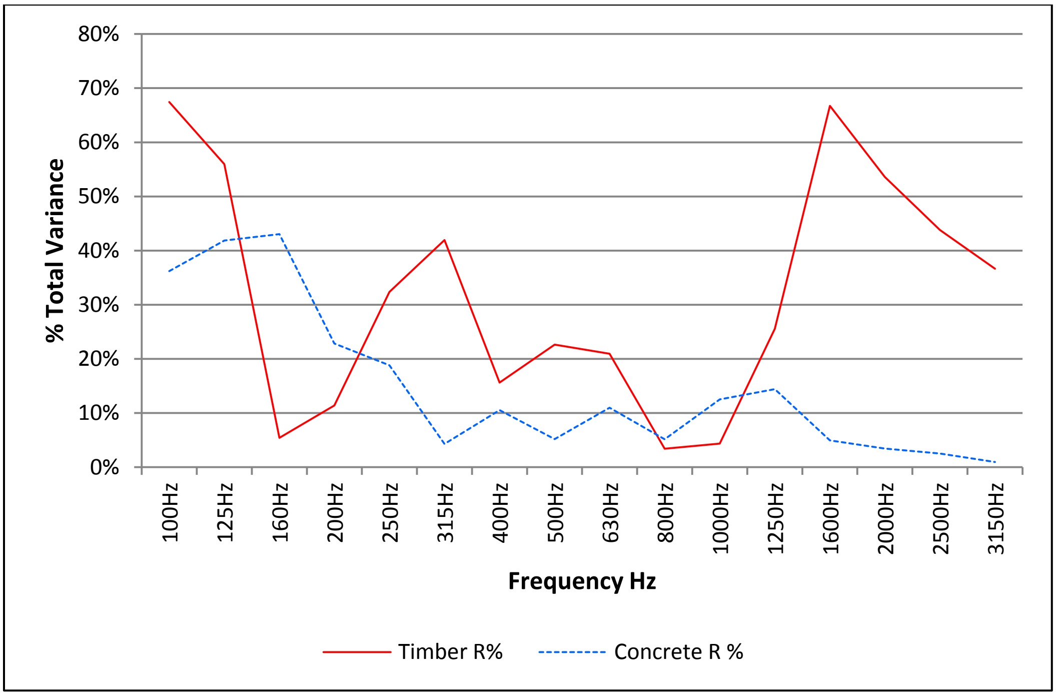


Figure 8-13: Timber & Concrete Reproducibility as a percentage of total variance

The concrete floor reproducibility contributions are influential on the total variance at low frequency where the room has a non diffuse field and is likely to cause variability. Reproducibility variance is 36-43% between 100 – 160Hz. It then trends downwards falling to 4% at 315Hz and varies between 5-14% between 400-1250Hz before falling from 14% at 1250Hz to 1% at 3150Hz. The relatively low percentage contribution in the concrete GRR after 250Hz, all under 20% and generally <10% indicates that the measurement system "may be acceptable" or "is acceptable" under AIAG criteria [102] and is able to distinguish individual parts (at least greater than one).

The timber floor reproducibility contributions to the total variance do not follow a consistent pattern and vary between 3% - 67% of the total variance depending on frequency. Reproducibility is influential at low frequency being 67 % & 56% at 100Hz and 125 Hz respectively but falls to 4% at 160Hz. The mid range frequencies 200Hz – 800Hz peak at 42% at 315Hz before falling to 3% at 800Hz. The higher frequencies exhibit a low contribution from reproducibility of only 4% at 1000Hz but rise to peak at 67% at 1600Hz before falling back to 37% at 3150Hz.

The operator effects drive the reproducibility for both GRR studies with the interaction term having more influence on the Concrete GRR Reproducibility Variance term at low frequency between 125 – 250Hz where the range is 1.6 – 3.2dB approximately 1.2 – 1.4dB higher than the "Operator" variance term. Interaction has only a relatively minor influence in certain frequencies in the timber GRR.

When reviewing the contribution of the reproducibility variance in percentage terms it is noted that the operator variance, measured in decibels is less than 1dB for the frequency bands between 160-1000Hz for timber and 160 -3150Hz for concrete. Because the operator variance is relatively small across this frequency range, any movement, in this case a fraction of a decibel will mean a significant and possibly misleading increase in percentage terms.

8.4.5.3 Repeatability (r)

The repeatability component of variance σ_E^2 is plotted for the timber and concrete GRR in Figure 8-14:

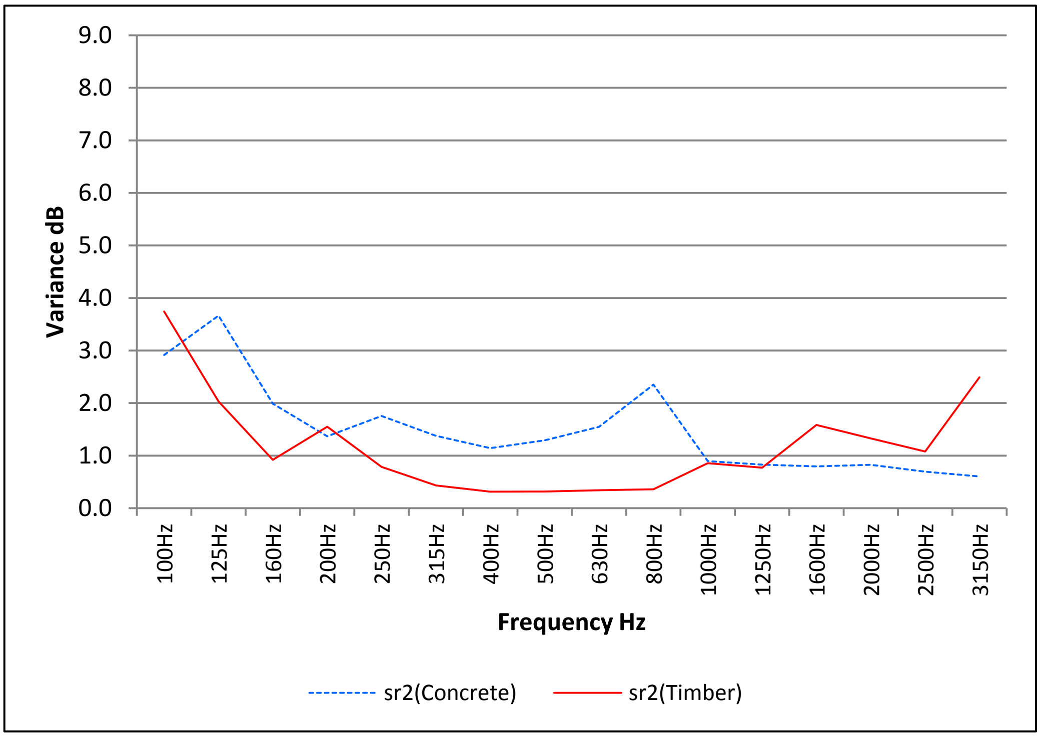


Figure 8-14: Timber & Concrete Repeatability Variance Contributions

Both the timber and concrete floor GRR have similar repeatability variance components across the frequency range. This is to be expected as the repeatability is meant to represent the instrumentation measurement error.

For the timber floor GRR the level falls from 3.7dB to 1.6db between 100-200Hz, it is then less than 1dB between 250 – 1250Hz inclusive rising to 2.5dB at 3150Hz.

For the concrete GRR the level rises from 2.9dB at 100Hz to 3.7dB at 125Hz before falling to 1.1dB at 400Hz. The level rises again peaking at 2.4dB at 800Hz before falling below 1dB from 1000 – 3150Hz.

The Reproducibility variance is detailed as a percentage of total variance in Table 8-20:

Table 8-20: Reproducibility - Proportion of Total Variance (%)

%	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Timber	30%	44%	55%	50%	32%	32%	36%	41%	31%	21%	18%	14%	22%	27%	16%	13%
Concrete	37%	47%	52%	15%	32%	20%	19%	12%	18%	22%	9%	10%	9%	11%	7%	4%

This data is plotted in Figure 8-15:

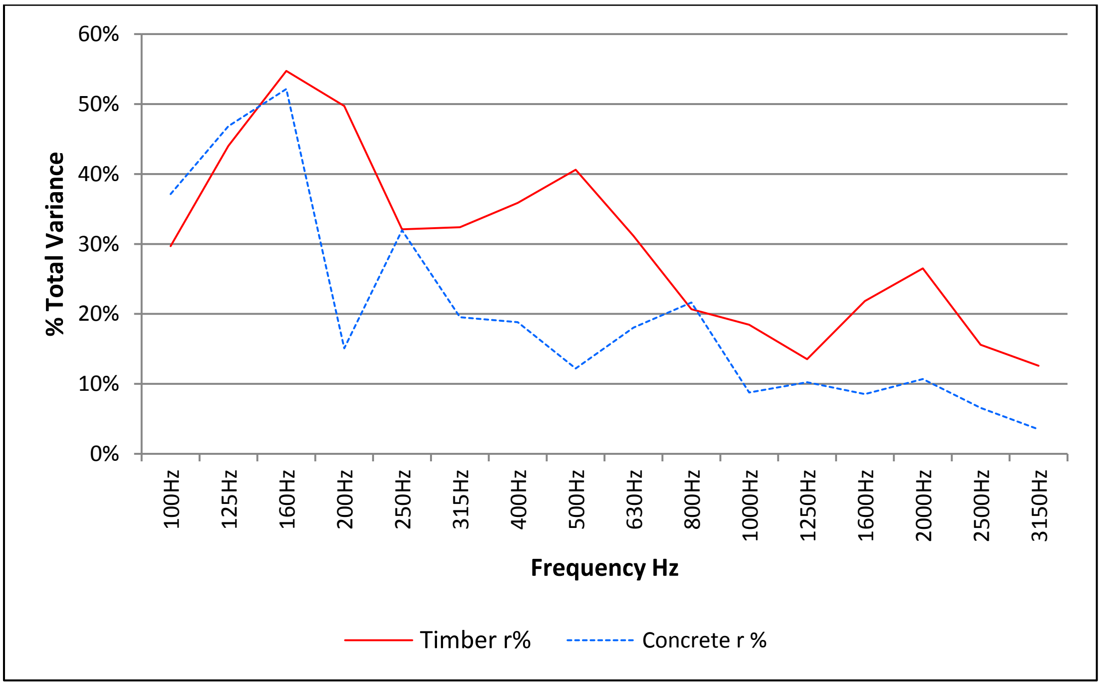


Figure 8-15: Timber & Concrete Repeatability as a percentage of total variance

Both the concrete and timber GRR repeatability variances are most influential at low frequency 100Hz-160Hz, rising from 30% - 55% in the case of timber and 37% - 52% in the case of concrete before falling to 13% and 4% at 3150Hz respectively.

The increased proportion of total variance at low frequency is expected as the non-diffuse field in the small rooms will affect both the operator and the instrument variance at the low frequency end of the spectrum. It is noted that the relatively low level in decibel terms and falling influence of the instrumentation as frequency increases tends to suggest that the repeatability component is apparently not significantly affected by background noise effects on either test site.

8.4.5.4 Gauge Variability

The gauge component of variance σ^2_{GRR} is the repeatability and reproducibility components combined. It also encompasses any interaction which may be present between the part and the operator factors.

The gauge variance is plotted for the timber and concrete GRR in Figure 8-16:

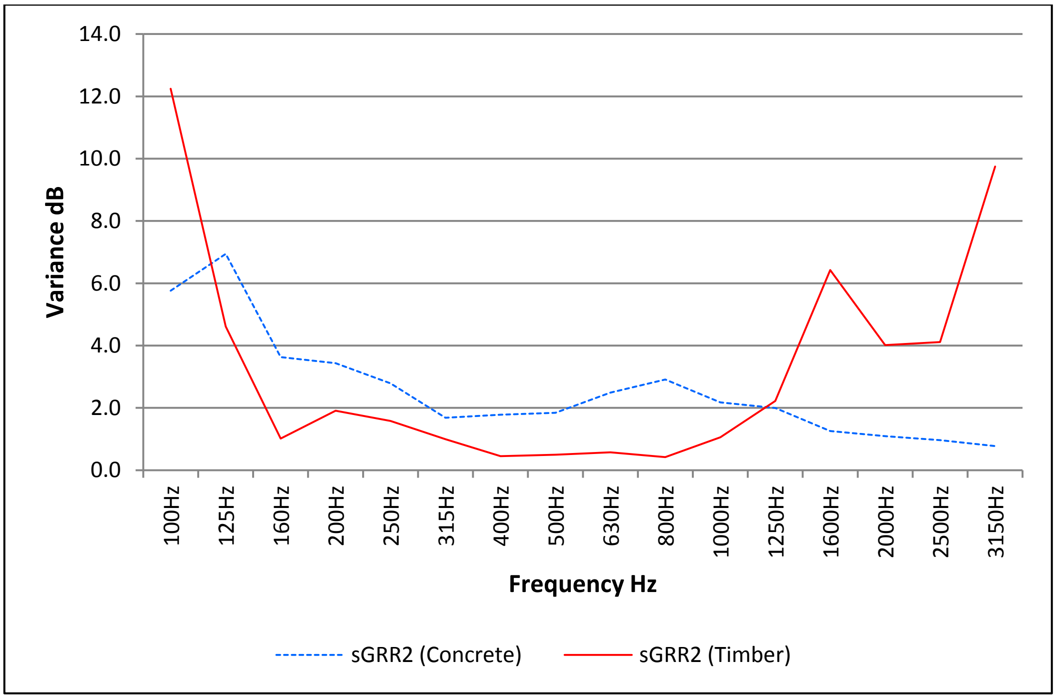


Figure 8-16: Timber & Concrete Gauge Variance Contributions

It is noted that the same test kit and instrumentation was used for both GRR experiments but not necessarily by the same operators.

The timber gauge variance is higher than the concrete gauge variance at 100Hz and 1250Hz – 3150Hz. Visual inspection of the shape of the gauge variance curve indicates the reproducibility variance component makes a substantial contribution and dictates the shape of the curve. The timber gauge variance is affected by the non diffuse field in the test rooms at low frequency and by the influence of relatively high background noise and/or a combination of inadequate noise output from the test kit and high performing floors. Between the range 160 – 1000Hz the timber Gauge variance is 0.4 and 1.9dB which is between 1 – 2dB lower than the concrete gauge variance.

It is noted that the operator variances for timber and concrete GRR are very similar across 125 – 1250Hz, the difference in the gauge variance across this range is therefore due to a combination of the concrete GRR repeatability variance which is generally greater than that of the timber GRR between 125 – 1250Hz and; the interaction between the operator and part which has a magnitude of 1 – 1.4dB between 125Hz – 250Hz and between 0 – 0.4 dB between 250Hz – 1250Hz.

In general terms, the concrete floor σ_{GRR}^2 the reproducibility variance is most influential at 200Hz, 100Hz and 1250Hz, repeatability is dominant at all other frequencies. See Figure 8-17.

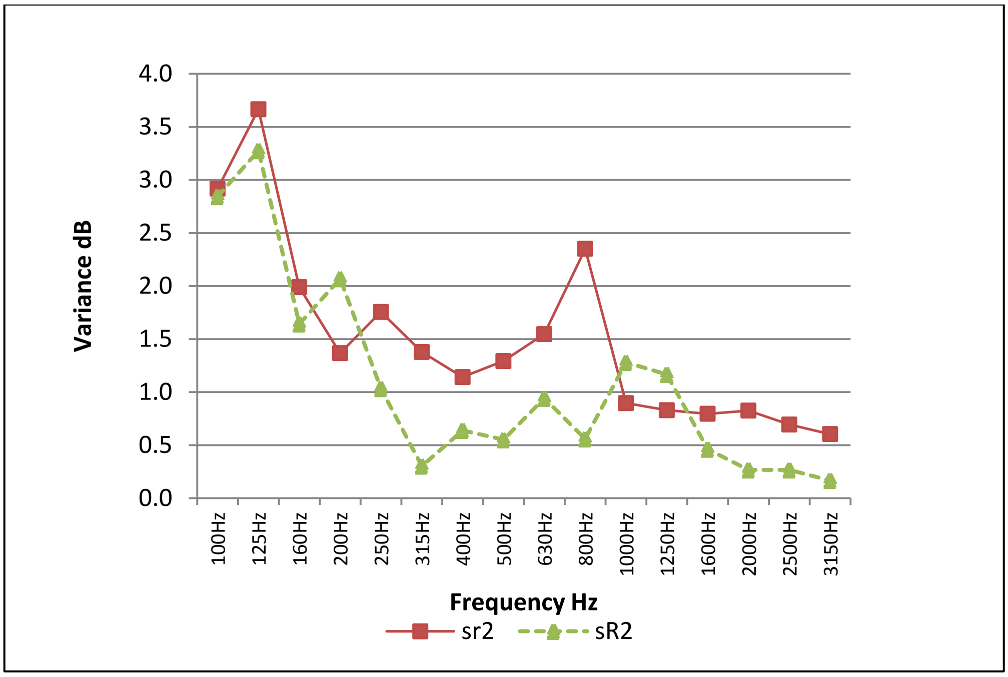


Figure 8-17: Concrete GRR R & r variance

For the timber floor σ^2_{GRR} the reproducibility variance is most influential at 100Hz, 125Hz and above 1250Hz, repeatability is dominant at 160Hz, 200Hz and 1000Hz and is similar for all other frequencies. See Figure 8-18.

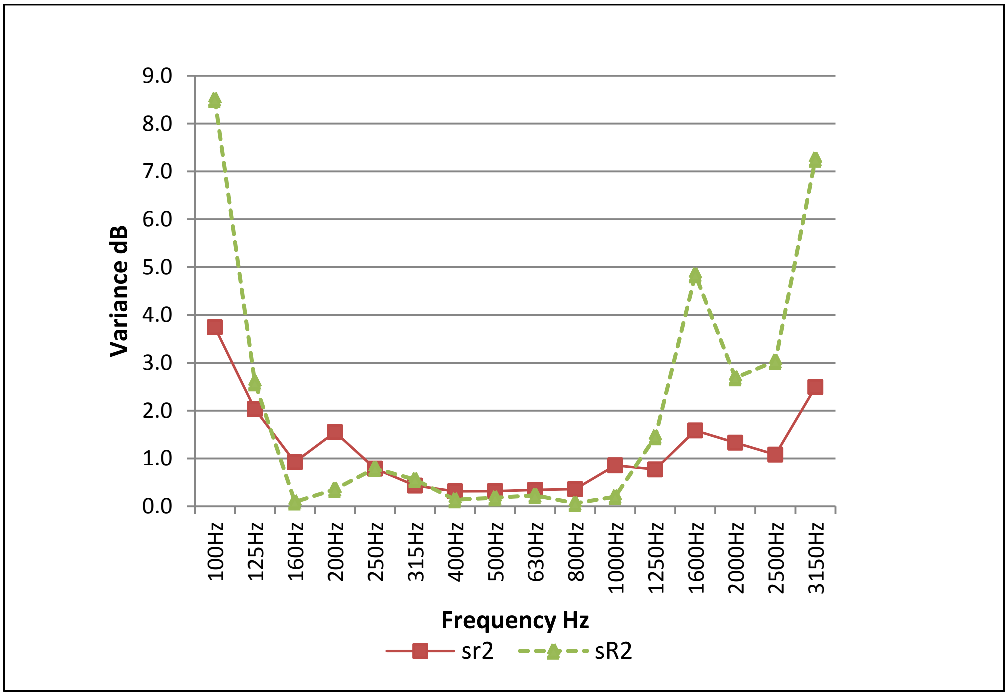


Figure 8-18: Timber GRR R & r variance

The percentage of the total variance the gauge contributes is calculated. This is detailed in Table 8-21 and is plotted in Figure 8-19.

Table 8-21: Gauge - Proportion of Total Variance (%)

%	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Timber	97%	100%	60%	61%	64%	74%	51%	63%	52%	24%	23%	39%	89%	80%	59%	49%
Concrete	73%	89%	95%	38%	51%	24%	29%	17%	29%	27%	21%	25%	13%	14%	9%	4%

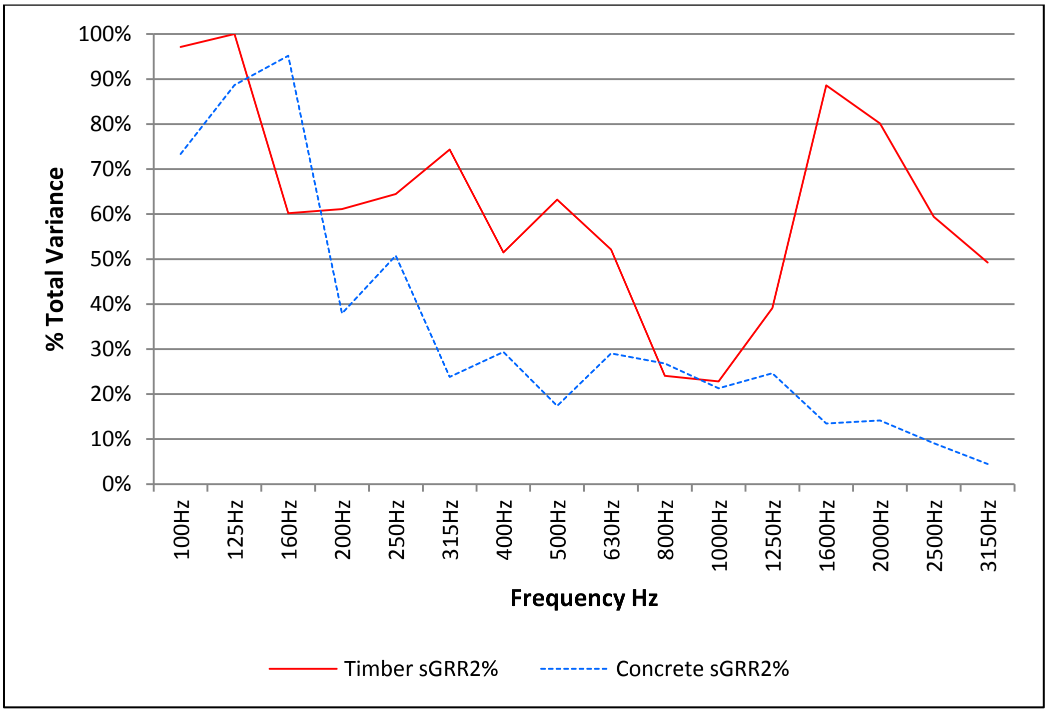


Figure 8-19: Timber & Concrete Gauge variance as a percentage of total variance

The influence of the gauge is greatest at the low frequency where the gauge for both timber and concrete GRR was the major component of variance. Its influence generally falls for the concrete GRR as frequency increases. For the timber GRR the level trends downward to 1000Hz, after 1000Hz it increases again. This is caused by the difficulty in measuring an accurate figure due to the influence of background noise.

Under AIAG criteria the only frequencies where the percentage contribution from the measurement system would be deemed "acceptable" would be in the 2500Hz and 3150Hz third octave bands for the Concrete GRR.

For the timber GRR the 30% AIAG criteria is only achieved in the 800Hz and 1000Hz bands where the measurement system "maybe acceptable". It should be noted however that this is based on a percentage of total variability in the sample and as the samples in both GRRs are constrained this artificially inflates the contribution from the other factors. The full extent of the measurement system is not tested.

In order to allow a proper assessment of the sound insulation measurement system under the AIAG criteria four additional test elements were added to the sample. The additional test elements were chosen to extend the total range and all data from the concrete and timber GRR was combined to create one large GRR test sample. The results from this are discussed in the Linear Test analysis.

8.4.6 Linearity Test Data

The reasons for including a "linearity test" in the analysis of the measurement system were discussed in paragraph 5.2.2.1. The interest in this particular experiment is in the measurement system itself and the measurement uncertainty offered by the the operators and the sound insulation test kit. The variability of the part is expected to be high as the test sample intentionally includes parts which artificially represent test elements at the extremes of the normal site measurement range.

The linearity test sample was included in order to test the gauge test capabilities for low performing and high performing test specimens. This is to ensure that the gauge under test, in our case the sound insulation measurement system, is observed over its full measurement range and not just a restricted sample based on a generic construction. This is in contrast to the initial GRR DOE which was manipulated to select carefully chosen floor constructions and room sizes in order to tell us something about the variability of the floor construction itself. The linear test is an examination of the measurement system not the construction being measured. Therefore the additional test elements are chosen only for their potential to produce an airborne sound insulation test result higher or lower than that sampled in the timber and concrete GRR studies. They do not represent any particular construction type and are a device to ensure the sound insulation range is extended to provide additional information about the capabilities of the measurement system. Internal plasterboard partions within dwellings were chosen to offer low performing test elements and two situations were used where sound insulation was tested between a line of three internal rooms within a residential dwellings using the middle room as the “test element” and the end room as the source and receiver rooms. The test room pairs were not restricted to small rooms (i.e. bedrooms circa 25m^3). Rooms classed as living rooms and dining rooms were used which had larger volumes. Test samples were chosen to represent sound insulation performance at the practical upper and lower range of sound insulation measurement in the field and a summary is detailed in the table below:

Table 8-22: Range of performance for the linear test sample only (4 test elements measured by all 5 operators)

Linear Test Sample	$D_{nT,w}$	$D_{nT,w}$ + Ctr
Mean	50.2	44.3
Min	34	27
Max	62	56

The four additional test elements were tested on the same site as the concrete GRR survey. They were tested a total of three times each by all 5 operators and the data was added to the samples from the timber and concrete floor experiments giving overall total GRR data test sample of 240. This data sample follows closely the recommended DOE advice for testing linearity and contains more data points giving greater degrees of freedom which improves the confidence in the results from the model. Visual inspection of the histogram shows that the addition of this data at the extreme ends of the performance range meant that the overall distribution shape is now significantly skewed and has a three peaked distribution, see Figure 8-20:

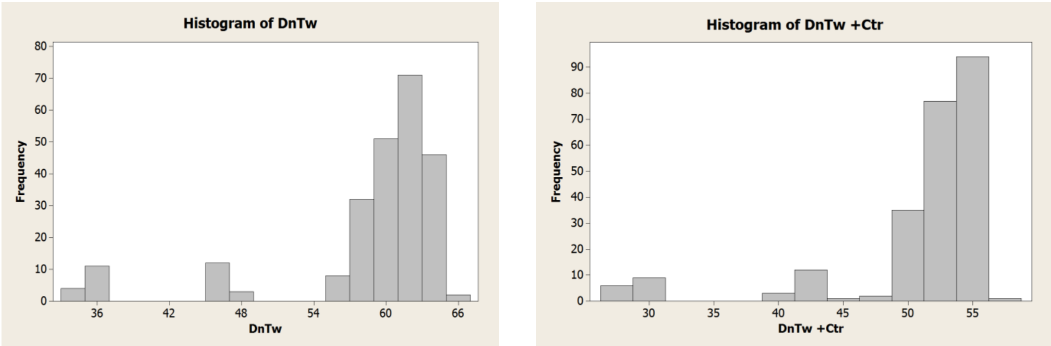


Figure 8-20: Histograms of $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} show a non-normal distribution shape for the Linear (All) Test Sample

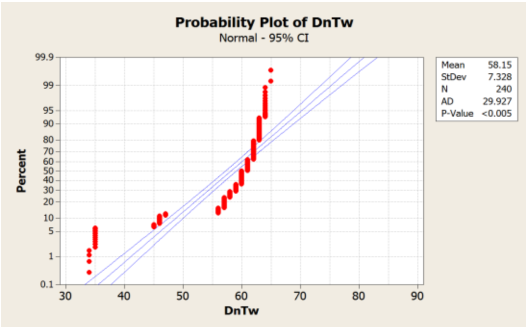
In addition there is clear indication from the p-values in Table 8-23 that the combined data sample of 240 surface tests deviates from normality for the linear GRR experiment. This is mainly because the three samples taken are from different construction populations and were selected in order to obtain suitable performing surfaces at the extreme ends of the normal measurement range in field testing test. The elements had to be chosen based on their likely sound insulation performance rather than their construction and they had to be readily available on site e.g. internal stud partitions between rooms.

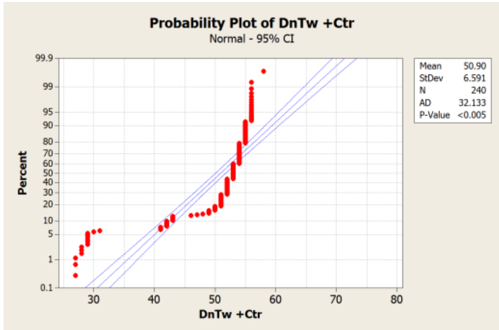
Table 8-23: D_{nT} Data – (240 test sample): σ and p-values for Linear (All) floor tests.

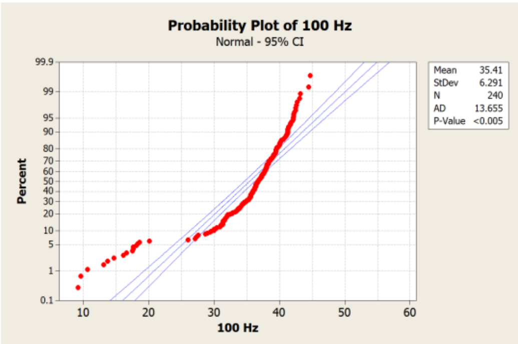
Floor	**Linear Data**	
Hz	σ	**p-value**
$D_{nT,w}$	7.328	<0.005
$D_{nT,w}$ +Ctr	6.591	<0.005
100 Hz	6.291	<0.005
125 Hz	7.316	<0.005
160 Hz	6.828	<0.005
200 Hz	6.597	<0.005
250 Hz	7.569	<0.005
315 Hz	8.528	<0.005
400 Hz	8.473	<0.005
500 Hz	7.728	<0.005
630 Hz	6.925	<0.005
800 Hz	6.488	<0.005
1.0 kHz	6.547	<0.005
1.25 kHz	7.285	<0.005
1.6 kHz	7.947	<0.005
2.0 kHz	8.247	<0.005
2.5 kHz	8.949	<0.005
3.15 kHz	11.12	<0.005

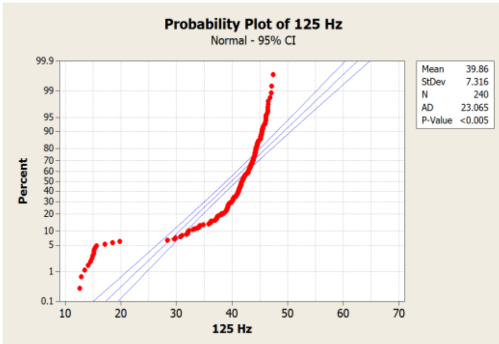
Yellow boxes (p<0.01) reject H_o and conclude "x" is not normal. All the boxes for the linear test data are yellow.

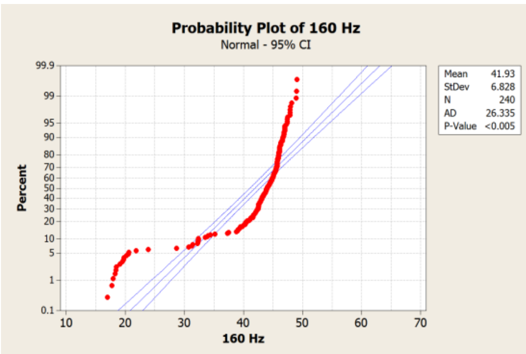
This is also reflected in the normal probability plots which show severely skewed data well outside the normal 95% confidence limits indicating non-normality in the test sample. See Figure 8-21 & Figure 8-22:

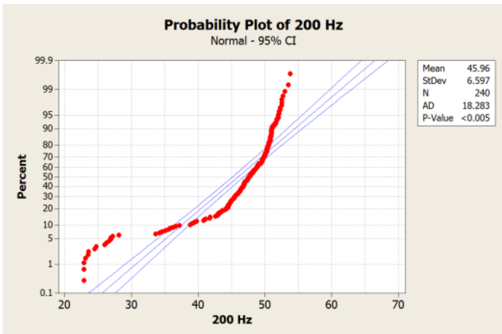


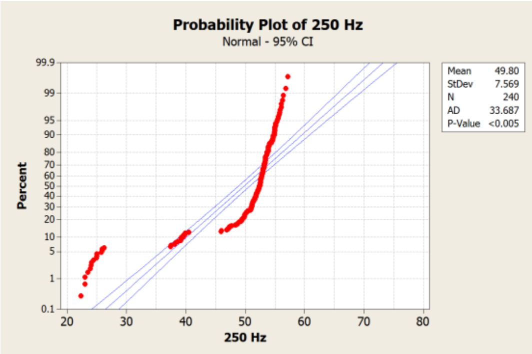


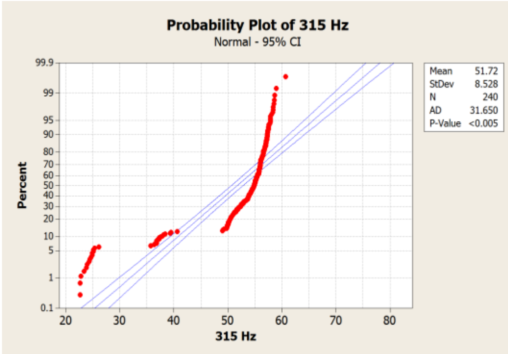


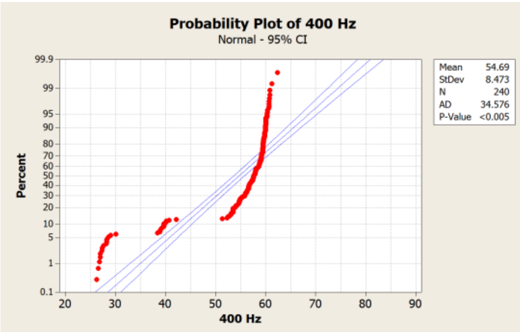


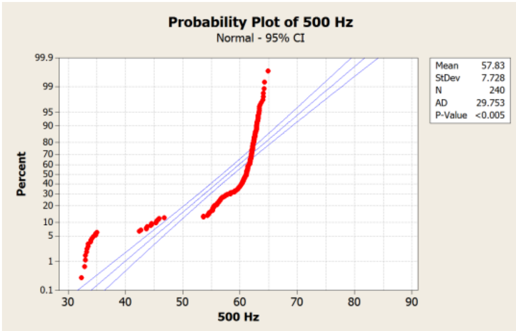


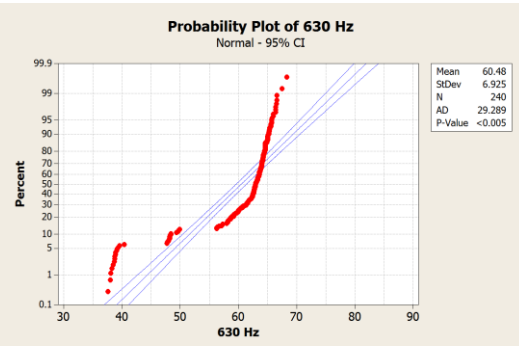


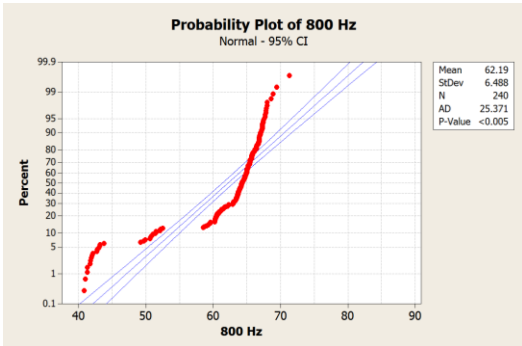

Probability Plot of 250 Hz
Normal - 95% CI
Mean 49.80
StDev 7.569
N 240
AD 33.687
P-Value <0.005
Percent
250 Hz

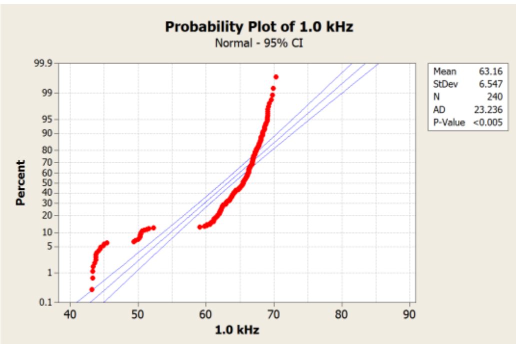

Probability Plot of 315 Hz
Normal - 95% CI
Mean 51.72
StDev 8.528
N 240
AD 31.650
P-Value <0.005
Percent
315 Hz

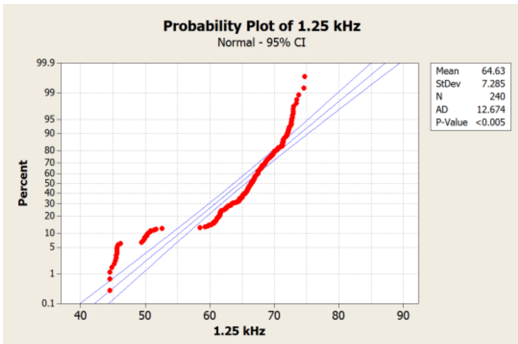

Probability Plot of 400 Hz
Normal - 95% CI
Mean 54.69
StDev 8.473
N 240
AD 34.576
P-Value <0.005
Percent
400 Hz


Probability Plot of 500 Hz
Normal - 95% CI
Mean 57.83
StDev 7.728
N 240
AD 29.753
P-Value <0.005
Percent
500 Hz


Probability Plot of 630 Hz
Normal - 95% CI
Mean 60.48
StDev 6.925
N 240
AD 29.289
P-Value <0.005
Percent
630 Hz


Probability Plot of 800 Hz
Normal - 95% CI
Mean 62.19
StDev 6.488
N 240
AD 25.371
P-Value <0.005
Percent
800 Hz


Probability Plot of 1.0 kHz
Normal - 95% CI
Mean 63.16
StDev 6.547
N 240
AD 23.236
P-Value <0.005
Percent
1.0 kHz


Probability Plot of 1.25 kHz
Normal - 95% CI
Mean 64.63
StDev 7.285
N 240
AD 12.674
P-Value <0.005
Percent
1.25 kHz

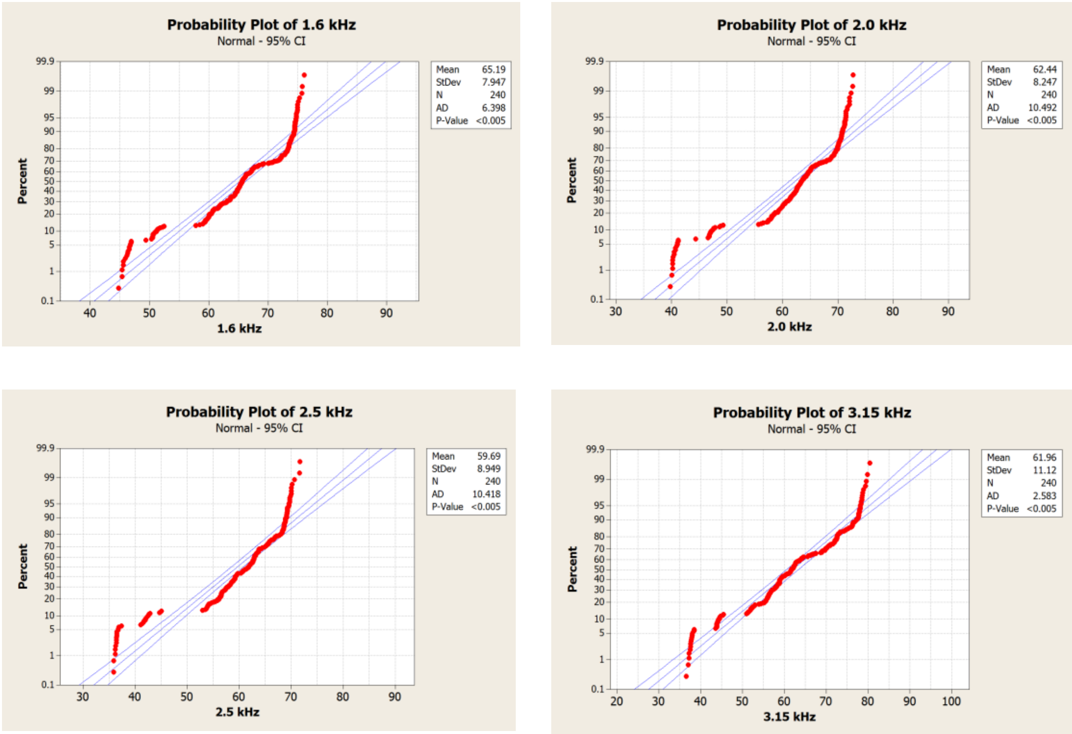
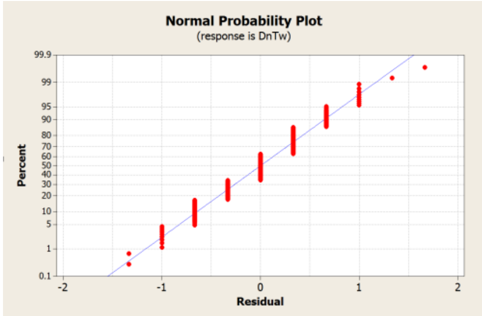
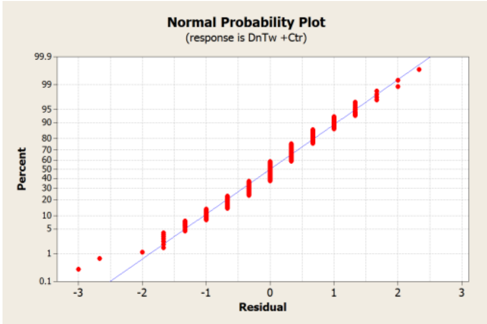
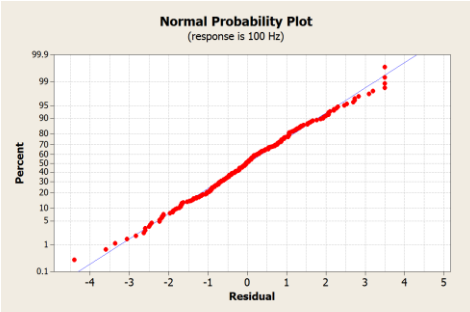
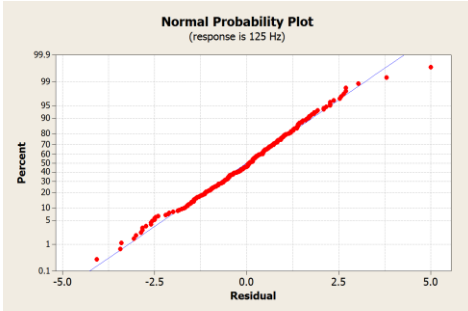


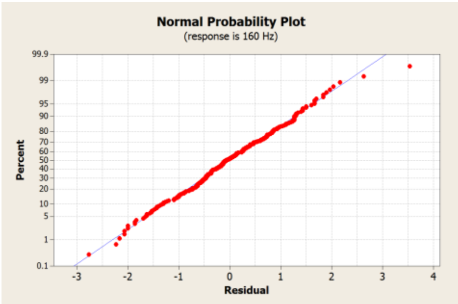
Figure 8-21: Residual Normal Probability Plots – Linear (All) Floor Tests showing non- normal distribution

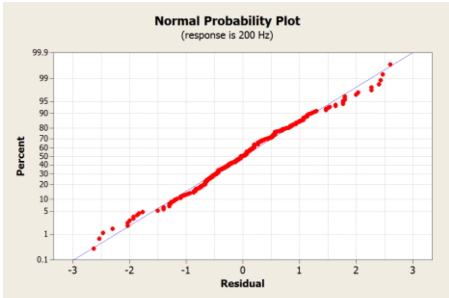

Normal Probability Plot
(response is DnTw)
Percent
Residual

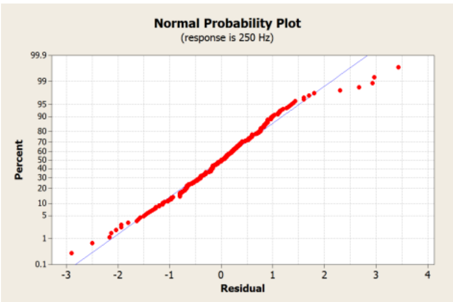

Normal Probability Plot
(response is DnTw +Ctr)
Percent
Residual

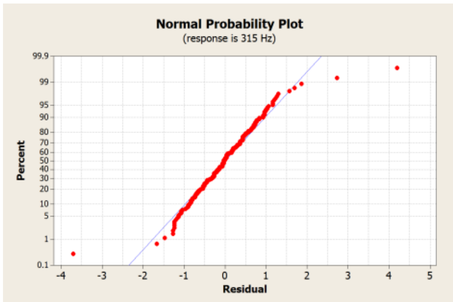

Normal Probability Plot
(response is 100 Hz)
Percent
Residual

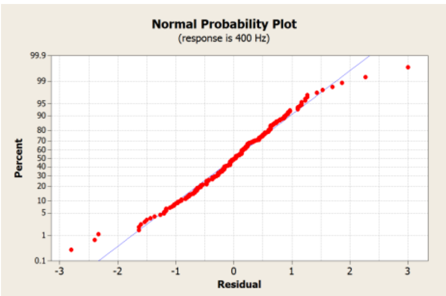

Normal Probability Plot
(response is 125 Hz)
Percent
Residual

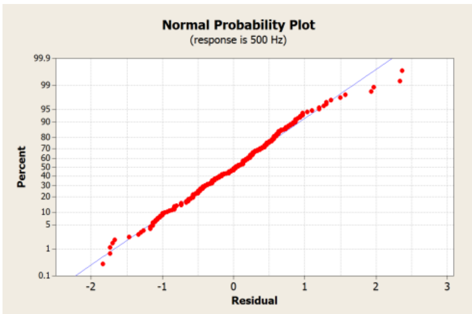

Normal Probability Plot
(response is 160 Hz)
Percent
Residual


Normal Probability Plot
(response is 200 Hz)
Percent
Residual


Normal Probability Plot
(response is 250 Hz)
Percent
Residual


Normal Probability Plot
(response is 315 Hz)
Percent
Residual


Normal Probability Plot
(response is 400 Hz)
Percent
Residual


Normal Probability Plot
(response is 500 Hz)
Percent
Residual

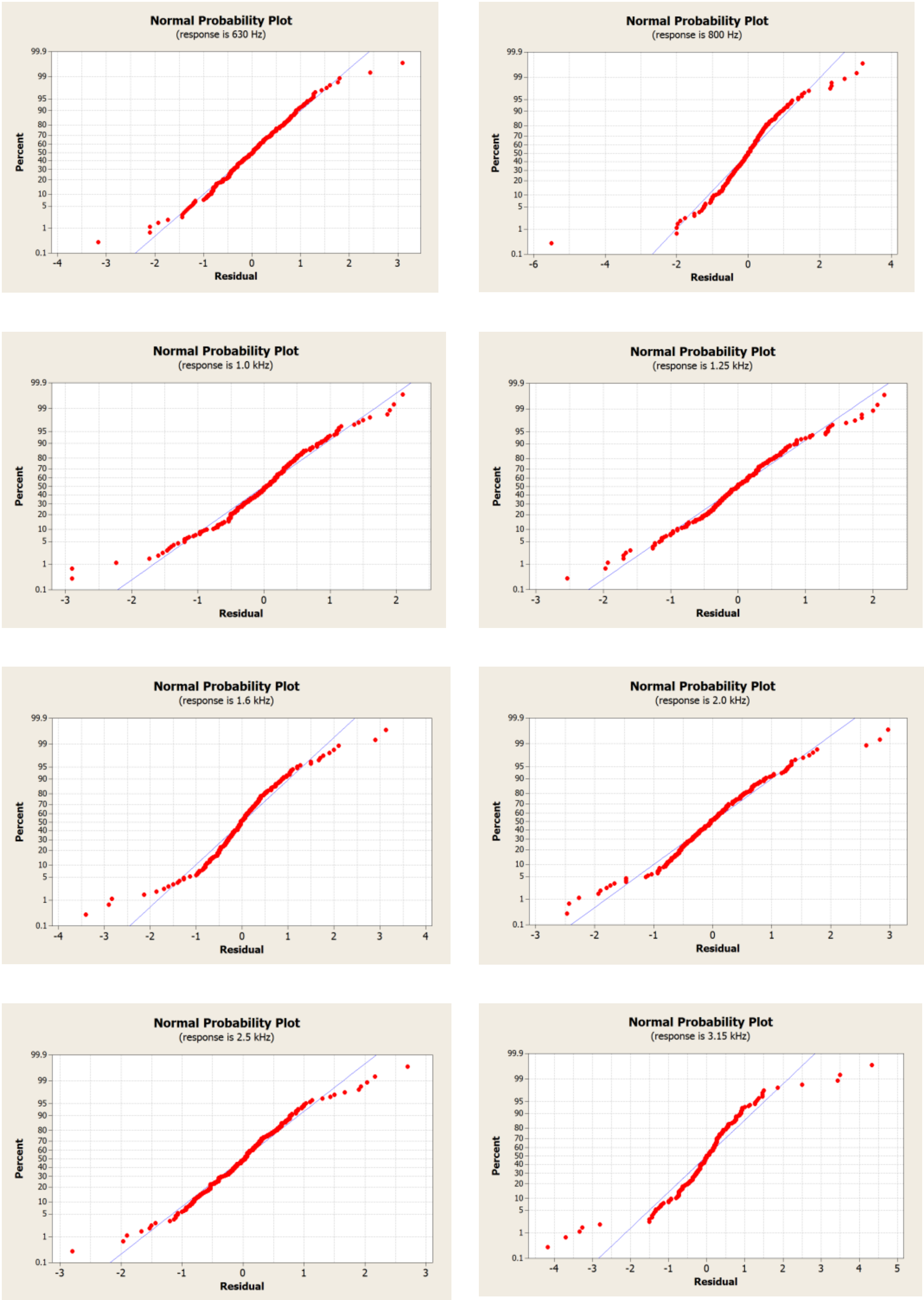


Figure 8-22: Residual Normal Probability Plots – Linear (All) Floor Tests

The normal probability plots of the residuals confirm the p-value table conclusion from Table 8-23 show that the data is not normal for the majority of the frequency range.

The repeatability and reproducibility test results for timber, concrete and the linear test are listed in Table 8-24:

Table 8-24: Repeatability & Reproducibility in dB: Timber / Concrete / Linear Data standard deviations

Floor	Timber		Concrete		Linear (All)	
Hz/dB	σ_{rTIM}	σ_{RTIM}	σ_{rCON}	σ_{RCON}	σ_{rLIN}	σ_{RLIN}
$D_{nT,w}$	1.6	2.0	2.0	1.8	1.7	1.7
$D_{nT,w}$ +Ctr	3.4	5.9	2.5	2.8	2.8	4.1
100 Hz	5.4	8.2	4.8	4.7	4.8	6.5
125 Hz	4.0	4.5	5.4	5.1	4.7	4.5
160 Hz	2.7	0.8	3.9	3.6	3.4	3.1
200 Hz	3.5	1.7	3.3	4.0	3.3	4.1
250 Hz	2.5	2.5	3.7	2.8	3.1	2.6
315 Hz	1.8	2.1	3.3	1.5	2.6	1.9
400 Hz	1.6	1.0	3.0	2.2	2.6	1.7
500 Hz	1.6	1.2	3.2	2.1	2.5	1.8
630 Hz	1.6	1.3	3.5	2.7	2.7	2.3
800 Hz	1.7	0.7	4.3	2.1	3.0	1.9
1.0 kHz	2.6	1.3	2.7	3.2	2.5	2.3
1.25 kHz	2.5	3.4	2.6	3.0	2.5	2.8
1.6 kHz	3.5	6.2	2.5	1.9	2.7	4.1
2.0 kHz	3.2	4.6	2.5	1.4	2.7	3.2
2.5 kHz	2.9	4.9	2.3	1.4	2.4	3.3
3.15 kHz	4.4	7.5	2.2	1.1	3.1	4.7

The $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} results are detailed in Table 8-25 & Table 8-26:

8.4.6.1 $D_{nT,w}$

Table 8-25: Linear GRR - Major Components of Variance ($D_{nT,w}$)

dB (σ^2)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
D_n $D_{nT,w\ Tw}$	0.7	0.4	0.4	0.2	0.2	56.3	57.0
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$	0.9	0.6	0.6	0.4	0.4	7.5	7.6

For the linear GRR the total variance is dominated by the part to part variance across the test sample as the size and shape of the room and construction of the separating element are now no longer fixed. For $D_{nT,w}$, the results show the instrumentation (representing repeatability σ_E^2) is responsible for 0.4dB of the total variance of the results. This term was 0.3dB for timber and 0.5dB for concrete, indicating the combined expanded sample size aggregates the pooled repeatability variance between samples. The reproducibility variance is the same as the repeatability term at 0.4dB; this is a combination of 0.2dB variance from the operator (represented by σ_O^2) and 0.2dB from the operator by part interaction term σ_{po}^2. The reproducibility and interaction variance was 0.4dB for timber with no interaction contribution and 0.4dB for the concrete GRR with 0.11dB interaction (0.29dB due to the operator).

8.4.6.2 $D_{nT,w} + C_{tr}$

Table 8-26: Linear GRR - Major Components of Variance ($D_{nT,w}$ + Ctr)

dB (σ^2)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
$D_{nT,w}$ + Ctr	3.1	1.0	2.2	0.9	1.3	43.0	46.1
dB (σ)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
$D_{nT,w}$ + Ctr	1.8	1.0	1.5	0.9	1.1	6.6	6.8

As in the timber and concrete GRR the introduction of the spectrum adaptation term increases the variance due to the added low frequency variability in the correction term.

For $D_{nT,w} + C_{tr}$, the results for the Linear GRR show the instrumentation (representing "repeatability" σ_E^2) is responsible for 1.0dB of the total variance of the results, this components value was 1.4dB for timber and 0.8dB for concrete, indicating the value aggregates across the larger test sample. The reproducibility variance (σ_R^2) is 2.2dB; this is a combination of 0.9dB from the operator (represented by σ_O^2) and 1.3dB from the operator by part interaction term σ_{po}^2. The reproducibility variances were 4.4dB for timber with no interaction and 1.0dB for the concrete GRR with 0.4dB interaction and 0.6dB due to the operator. Again the larger sample aggregates the components of variance terms.

The uncertainty increases for the single figure values when the spectrum adaptation term is included with the reproducibility term increasing the most. This is due to the increase in the interaction term as it is dominant over the operator term. The frequency analysis allows the contribution of this component of variance to be examined in more detail.

8.4.7 Frequency Data - Results

In order to investigate the contribution of each of the components at each frequency, the ANOVA was carried out over the linear GRR for each of the D_{nT} values, between 100Hz – 3150Hz. This gives a detailed view, in absolute terms, of where the major regions of variability lie. The results are summarised in Table 8-27 and Table 8-28 graphically represented in Figure 8-23 .

Table 8-27: Linear GRR (variance) - Major Components of Variance Frequency Data (D_{nT})

D_{nT} (var)	σ_{GRR}^2	σ_r^2	σ_R^2	σ_o^2	$\sigma_{p.o}^2$	σ_p^2	σ_{Total}^2
100Hz	8.3	2.9	5.4	1.1	4.4	33.5	41.8
125Hz	5.4	2.8	2.5	1.0	1.5	51.4	56.7
160Hz	2.7	1.5	1.2	0.3	1.0	46.7	49.4
200Hz	3.5	1.4	2.1	1.1	1.0	42.7	46.3
250Hz	2.1	1.2	0.8	0.0	0.8	58.7	60.7
315Hz	1.3	0.9	0.5	0.2	0.2	75.9	77.2
400Hz	1.2	0.9	0.3	0.0	0.3	75.0	76.2
500Hz	1.2	0.8	0.4	0.2	0.3	62.2	63.4
630Hz	1.6	0.9	0.6	0.2	0.4	49.3	50.9
800Hz	1.6	1.1	0.5	0.1	0.4	43.0	44.6
1000Hz	1.4	0.8	0.7	0.2	0.5	44.0	45.5
1250Hz	1.8	0.8	1.0	0.2	0.8	54.5	56.3
1600Hz	3.0	0.9	2.1	0.4	1.7	64.0	67.0
2000Hz	2.2	0.9	1.3	0.2	1.1	70.0	72.2
2500Hz	2.1	0.7	1.3	0.2	1.2	82.9	85.0
3150Hz	4.1	1.3	2.8	0.7	2.1	127.1	131.2

Table 8-28: Linear GRR (s.d.) - Major Components of Variance Frequency Data (D_{nT})

D_{nT} (s.d.)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
100Hz	2.9	1.7	2.3	1.0	2.1	5.8	6.5
125Hz	2.3	1.7	1.6	1.0	1.2	7.2	7.5
160Hz	1.6	1.2	1.1	0.5	1.0	6.8	7.0
200Hz	1.9	1.2	1.5	1.1	1.0	6.5	6.8
250Hz	1.4	1.1	0.9	0.0	0.9	7.7	7.8
315Hz	1.2	0.9	0.7	0.5	0.5	8.7	8.8
400Hz	1.1	0.9	0.6	0.2	0.5	8.7	8.7
500Hz	1.1	0.9	0.7	0.4	0.5	7.9	8.0
630Hz	1.2	1.0	0.8	0.5	0.7	7.0	7.1
800Hz	1.3	1.1	0.7	0.3	0.6	6.6	6.7
1000Hz	1.2	0.9	0.8	0.4	0.7	6.6	6.7
1250Hz	1.3	0.9	1.0	0.5	0.9	7.4	7.5
1600Hz	1.7	1.0	1.4	0.6	1.3	8.0	8.2
2000Hz	1.5	0.9	1.1	0.5	1.0	8.4	8.5
2500Hz	1.4	0.9	1.2	0.4	1.1	9.1	9.2
3150Hz	2.0	1.1	1.7	0.9	1.4	11.3	11.5

The total variance is plotted with the gauge and part to part variance in Figure 8-23.

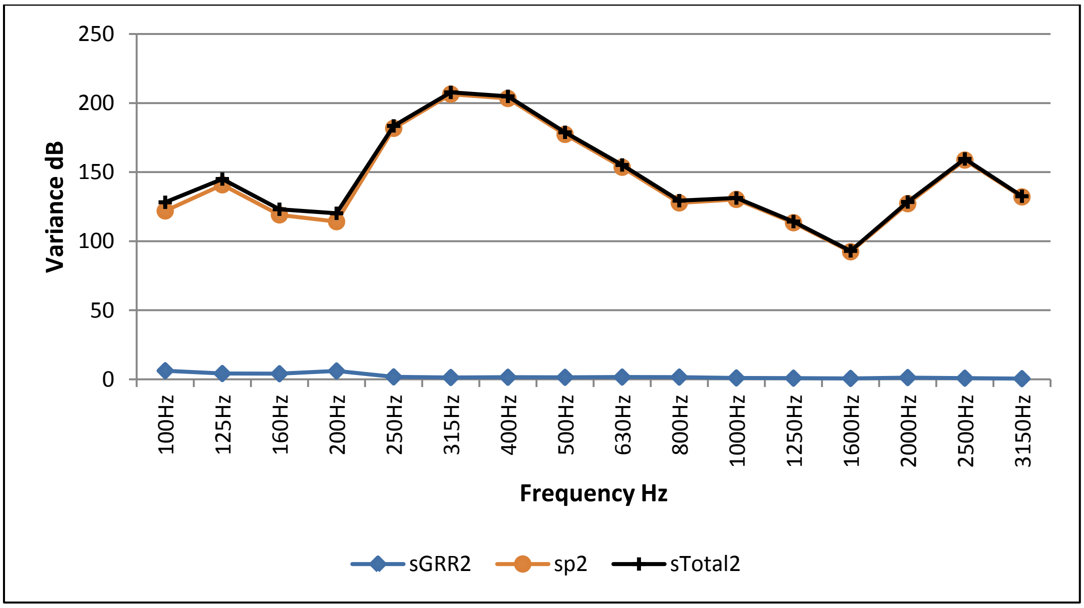


Figure 8-23: Linear GRR – Components of Variance - σ_{GRR}^2, σ_p^2, σ_{Total}^2

In the linear GRR the part to part variance is dominant across the full frequency range. It is only when the part to part contribution to the total variance is taken out that the measurement system contribution can be evaluated against a reasonable scale. This is shown in Figure 8-24.

The repeatability variance for the Linear GRR frequency data is higher at low frequency peaking at 2.9 dB at 100Hz and generally falling below, or just above, 1dB from 315Hz – 3150Hz. The repeatability is the dominant measurement system or “Gauge” uncertainty term between 250Hz – 1000Hz.

The reproducibility variance is highest at the low and high frequencies and a minimum in the middle range. It is higher than the repeatability variance at 100Hz, 200Hz and 1250 – 3150Hz. Interaction is the dominant component in reproducibility apart from 200Hz where the operator contributes 1.1dB (interaction is 1.0dB).

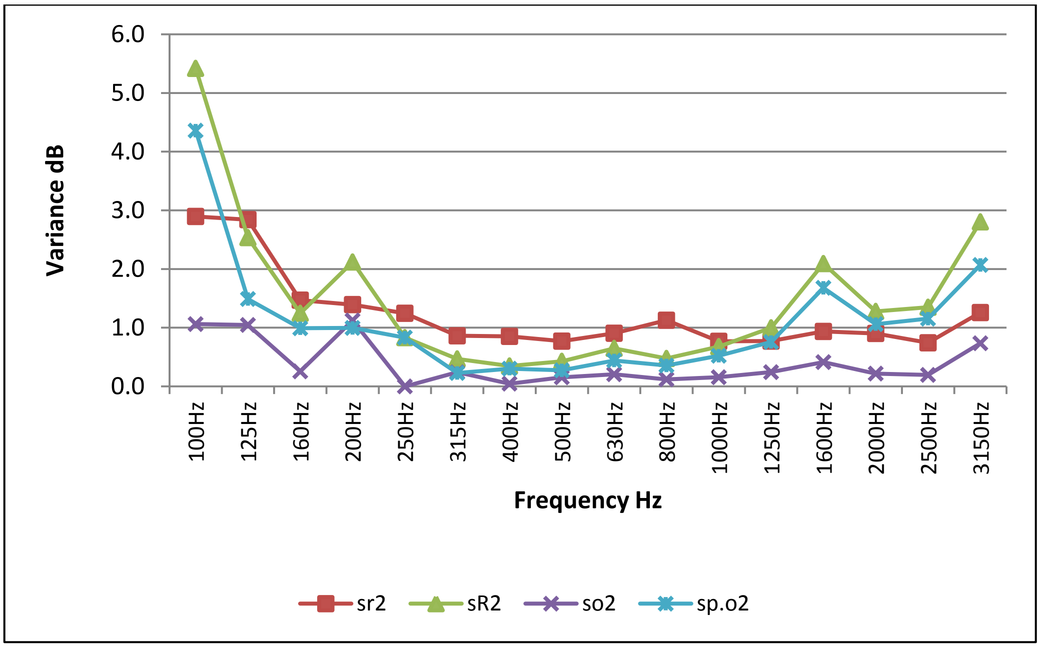


Figure 8-24: Linear GRR: Graphical representation of variance due to operator, interaction and repeatability & reproducibility.

The linear GRR individual repeatability, reproducibility and operator terms can be plotted on a graph with the timber and concrete GRR results to illustrate the effect of a combined GRR with extended range on the components of variance derived using ANOVA. The contribution of the measurement system repeatability and reproducibility is detailed in Figure 8-26 & Figure 8-27. The reproducibility is further sub-divided to show the operator contribution and the operator by part interaction in Figure 8-28 & Figure 8-29.

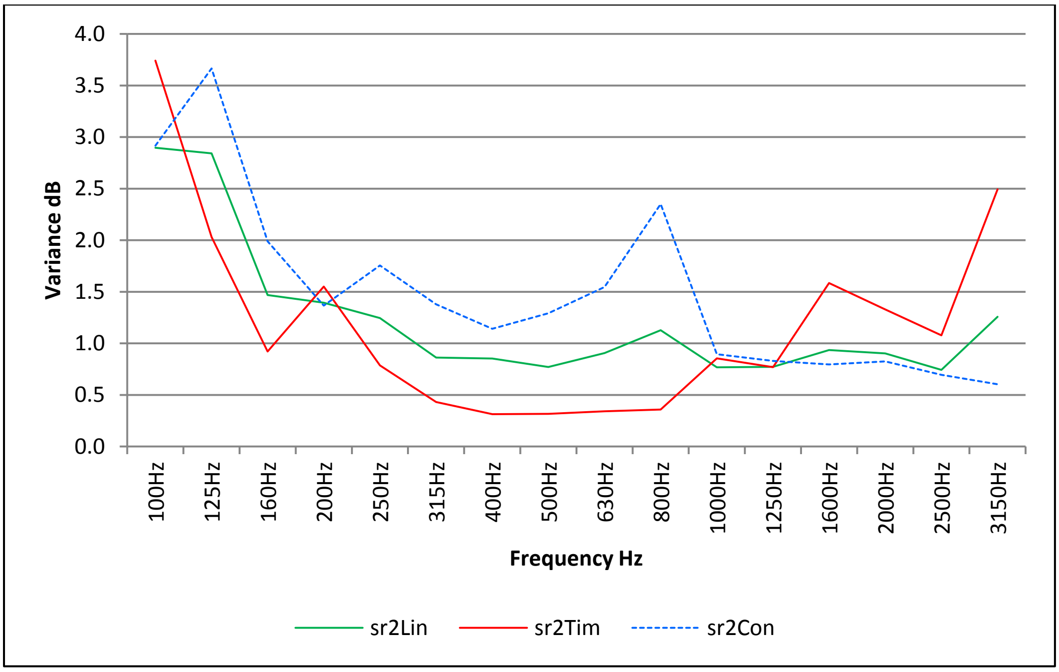


Figure 8-25: Repeatability Variance σ^2_r: Timber/Concrete/Linear GRR

The repeatability and reproducibility variance levels across the frequency range show how measurement system analysis is dependent on the construction being measured and to some extent the conditions on site.

The repeatability variance component associated with the instrumentation was generally lower for the timber GRR than it was for the concrete GRR. The low frequency range for both timber and concrete GRR <250Hz is influenced by room effects, i.e. by a non diffuse field.

The timber GRR was affected by background noise at high frequency as represented in the repeatability component of variance from 1250-3150Hz. For the concrete GRR there is some background noise effect which occurs in the mid range frequencies. This is the reason the GRR shows higher repeatability variance than the timber GRR, between 125 – 800Hz. For 1000Hz – 1250Hz repeatability is similar for both timber and concrete floors. The timber GRR is affected more and has higher repeatability in the range 1600Hz – 3150Hz.

The pooled repeatability, which incorporates the data from the timber and concrete GRR, plus four new test elements, displays a trend between the two larger GRR studies.

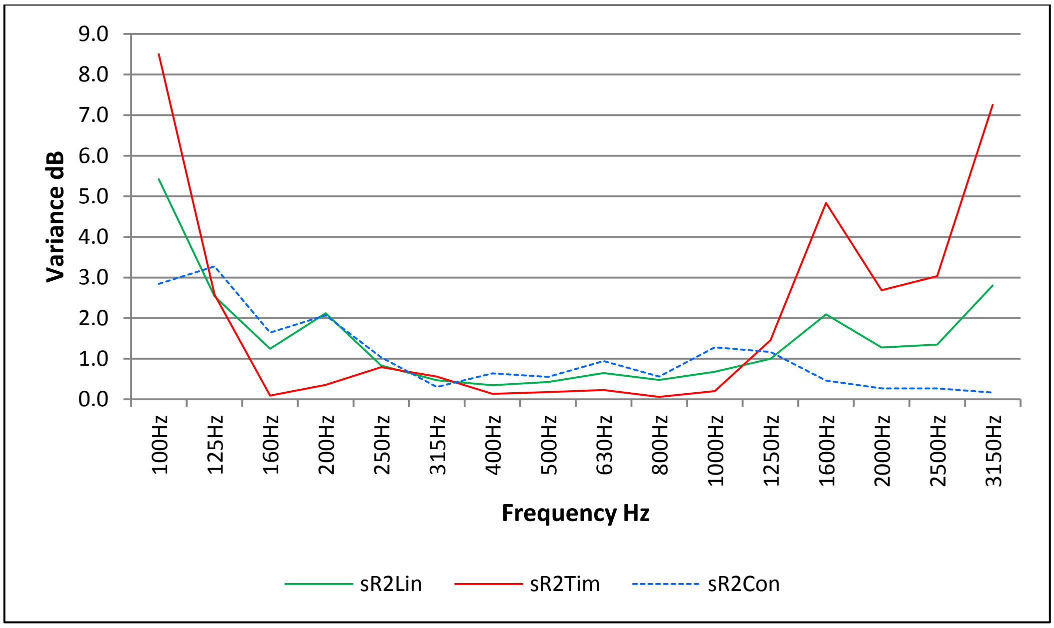


Figure 8-26: Reproducibility Variance σ^2_R: Timber/Concrete/Linear GRR

The reproducibility variances are also affected, below 250Hz, by the low modal density in the room and non diffuse field and in the timber GRR by the background noise correction term at the higher frequency bands 1250Hz – 3150Hz. The reproducibility derived from the combined data in the linear GRR ranges in between the variances of the two larger GRR studies though this does not necessarily mean that it will always take a middle route in all components of variance. The reproducibility component can be sub-divided into two further components for the operator and the operator by part interaction. These help describe where the variability associated with the reproducibility originates and also informs where the independence of these factors are compromised. They are detailed in Figure 8-27 & Figure 8-28.

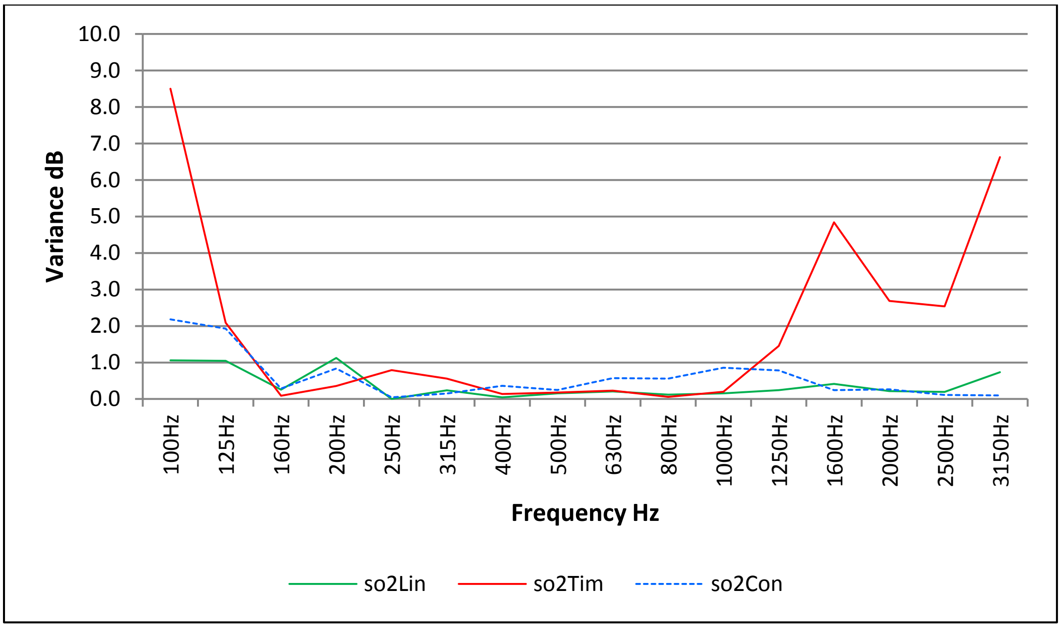


Figure 8-27: Operator Variance σ^2_o: Timber/Concrete/Linear GRR

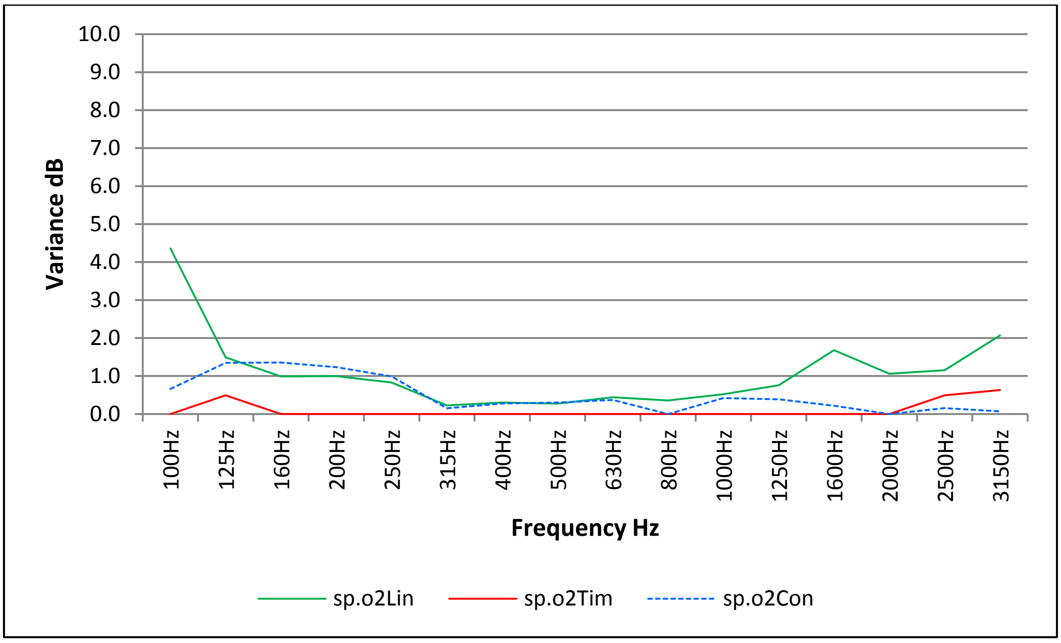


Figure 8-28: Interaction Variance $\sigma^2_{p.o}$: Timber/Concrete/Linear GRR

The timber reproducibility variance is calculated from a reduced model (without interaction) apart from 125Hz, 2500Hz & 3150Hz. It is therefore represented by the operator variance for the majority of the frequency range.

For the concrete GRR, interaction between part and operator is significant across most frequencies and it is the dominant factor in the reproducibility variance at the 125Hz – 250Hz bands.

The linear GRR operator variance is lower than both the timber and concrete GRRs at 100Hz, see Figure 8-27. This is enhanced by an interaction term that is higher than the timber and concrete GRRs, the result is that the reproducibility variance for the Linear GRR 100Hz band is between the timber and concrete GRR values. A similar situation also occurs at the 1000Hz – 3150Hz frequency bands, where there is significant interaction identified between the part and the operator for the Linear GRR (0.5 – 2.1dB) though the operator variance at these frequencies is relatively low (0.2 – 0.74dB). The contribution of the interaction and operator components is detailed in Figure 8-29.

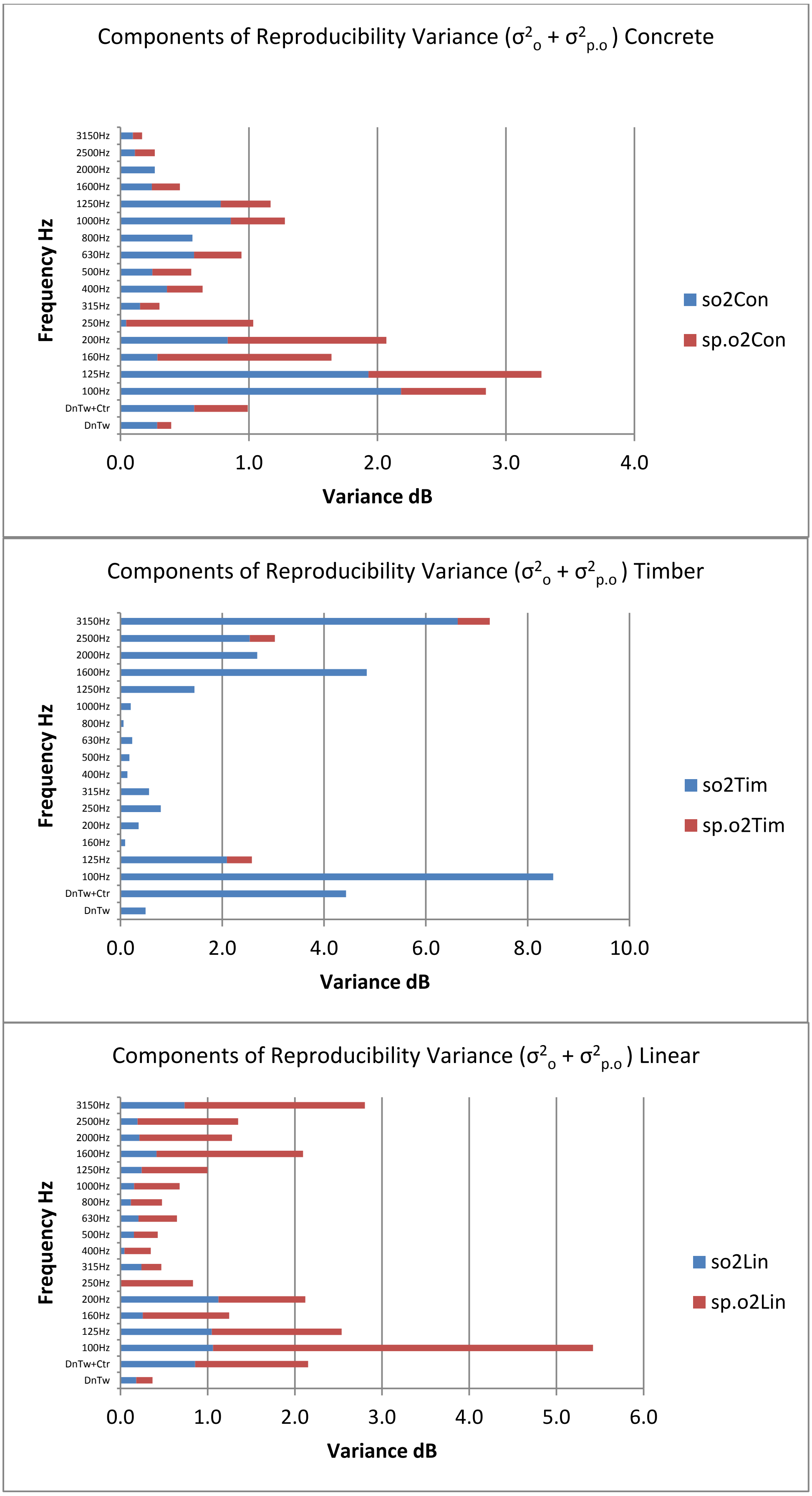


Figure 8-29: Reproducibility Variance by components

In the GRR for timber there is no significant interaction, a reduced model is fitted to the data and it is concluded that the factors contributing to the total measurement uncertainty are independent. For concrete the interaction contribution is significant for almost all frequencies and the full model is fitted which incorporates an interaction variance term. The combined study shows the interaction term for the linear GRR has a significant influence on the reproducibility variance, with peaks of >4dB at 100Hz, and is >0.5dB from 125 – 250Hz and from 1000-3150Hz.

8.5 Discussion

The data for the timber and concrete GRR together with the linearity test sample are representative of airborne sound insulation tests carried out in the field.

8.5.1 Repeatability

For repeatability, the data show that the concrete and timber GRR studies produce similar results, which are expected, as the same instrumentation was used in both studies. The inclusion of the additional tests to the combined study produced a “pooled” variance term for repeatability, which is detailed in Figure 8-25. Repeatability levels on this graph fall within a range of 3 – 4 dB over the frequency range, with peaks where background noise was influential.

8.5.2 Reproducibility

The results shown in Figure 8-26 for reproducibility are significantly different below 250Hz and above 1250Hz, but are of similar magnitude within that range. Visual inspection confirms that the curve is similar to the operator variance for both timber and concrete floors.

The operator variance incorporates the variability of the measurement method, the effects of the room on the sound pressure level and reverberation time frequency response, and site conditions (background noise). The magnitude of the operator’s effect varies across the frequency range and also between GRR studies as different influences dominate the field test environment.

We expect the operator variance term to take into account the predictable high levels of variance at low frequency which are expected to fall with increasing frequency. This is observed in both the timber and concrete studies though the timber GRR has a relatively high variance term at 100Hz, the variance due to the operator is virtually identical for both studies between 125Hz – 1000Hz and is generally <1dB.

Other factors clearly influence the higher frequency performance as the operator variance increases significantly in the timber GRR after 1250Hz. This is related to a combination of: high background noise, equipment limitations and the performance of the test element being measured. They combine to give a relatively high variance for the Timber GRR operator term >6dB at 3150Hz. The concrete GRR variance remains <1dB across the frequency range 160Hz-3150Hz.

Site influences such as background noise, that are out of the control of the operator but are responsible for the magnitude of the operator variance, can mask other effects such as the operator by part interaction. This is particularly evident in the concrete data, where it is clear these factors are not independent i.e. the operators were offering some additional effect related to how they carried out the test procedure and dependent on the room in which they measured. The interaction influence in the GRR increased when four additional tests were carried out and both the timber and concrete data sets were combined to create a larger sample with 5 operators measuring 16 parts, three times each.

8.5.3 Interaction

The presence of interaction in the measurement process is not considered in the British Standards BS5725[5, 7, 24, 25], UKAS Guidance M3003 [76] or GUM [6], although its identification and quantification is a requirement of the EA Guidelines on the expression of uncertainty in quantitative testing [127].

The interaction between operator and part was one of the reasons ANOVA was attractive, compared with alternative approaches. Even so, the level of interaction and the fact that it appeared to be so significant in one study and not the other was unexpected. The DOE selected floor tests in preference to wall tests because floors offered more than 2 choices where a loudspeaker can be placed. This gives more potential for randomisation in the measurement process, though it appears that the choice of position in this study may be affected by other influences.

The cause of the interaction is uncertain but is obviously driven by the operators and the choices they make during the testing process. Operators will naturally try to work as efficiently as possible during the survey; economising on the effort expended and co-ordinating their actions to take the minimum time between tests. As the equipment is heavy and unwieldy, this is likely to include the choice of test kit placement in the test room. The test rooms in the concrete GRR were non-identical and individual rooms introduced factors: geometry, power sockets, windows, etc., which are likely to have caused some operators to constrain their test method. For example, if operators choose different corners for the speakers, this might result in an operator by part interaction. A simple reason for this to occur could be due to the location of plug sockets and the length of the loudspeaker power leads (which are not identical for all test kits). They lift the loudspeaker and tripod from one side of the room to another and so, probably unwittingly, use the same two corners of a particular room, constrained by the cable length which is in reach of a conveniently sited socket. In addition, it is possible that the operators who are in the rooms for all measurements, may prefer to point the sound level meter in one direction, perhaps facing a window where they can see outside. As the test regime is highly regimented and repetitive it may also be possible that after a few tests the operators 'fix' on a certain arrangement for the microphones and the loudspeaker, it could even be related to the operators being left or right handed, or which way a door opened into a room presenting two corners to the operator as they carried equipment into the test space.

The presence of interaction is significant for two reasons. Firstly, because this has implications for DOE for inter-laboratory studies and round robin tests. Operators should be aware that this may occur and inform other participants of techniques designed to avoid it. They should also, as a minimum precaution, incorporate a check for interaction into their data analysis procedure. Secondly, interaction between key factors has other implications on the wider subject of uncertainty and the techniques to calculate its magnitude. It is noted that interaction has not been explicitly mentioned in recent papers using simulation models to determine uncertainty; see Goydke et al [59] and Wittstock [61]. Monte Carlo simulations, used to calculate the measurement uncertainty based on third octave band values, and which assumes independent input variables, will lack a component of variance, which in certain circumstances, will be influential at all frequencies. The results of these studies will therefore not be able to account for the components' contribution correctly. They need to incorporate an interaction term into the model to reflect the field test situation and this term may in turn be dependent on the type of room or construction they test.

8.5.4 Part to part

Apart from quantifying the repeatability and reproducibility, the three GRR data sets have provided insights into other factors and their contributions to uncertainty. The part being measured has a significant contribution even if building elements are nominally identical. This has repercussions when comparing results with the guideline reproducibility values in the international standard [8].

Apart from instrumentation variability, which is not reliant on the test sample being measured, a true comparison of reproducibility with the international standards "R" (Table A.2) [8] must incorporate the variance of the part. This is because the variance of the part is implicit in the values for R in the standard. In inter-laboratory tests the part to part variance results from the reconstruction of the test sample at each location e.g. the re-fixturing of the glazing test specimen in the laboratory wall opening or the re-building of the walls in each of the laboratories using standard blocks or plasterboard. This ensures the part to part variability is included in the data sample.

For the timber floor, the room shape, size and volume were identical in each of the six test pairs. This allowed an assessment of the part variability due to the floor construction. It is shown that the floor construction contributes to the total variance measured. For timber floors the part to part variance was calculated as 0.24dB for $D_{nT,w} + C_{tr}$, with one standard deviation of 0.5dB. The part to part contribution to total variance across the frequency range 100Hz – 3150Hz varied significantly with frequency and was between 0 – 10dB.

For the GRR of the concrete floor, the sample was of small non-identical rooms. The GRR part to part variance was 0.8dB, $D_{nT,w} + C_{tr}$. As the room varied in volume by 4 – 5 m3 the variance is not due to construction alone and a direct comparison of the timber and concrete floor part to part variance is not possible. However, since the room dimensional differences were small, it might be assumed that 0.8dB represents an upper limit of part to part variance for this type of concrete floor. The third octave band sound insulation performance produced variances of between 0.2 – 16.5dB, dependent on frequency.

With respect to the parts contribution to total variability, the concrete floor GRR results show the part to part variance is the dominant component i.e. it is >50% of the total variance between 315Hz – 3150Hz. It is expected that if the parts measured are not identical this will increase the measured variance and with it the proportional contribution to the total variance. This is one of the reasons the timber GRR data is more informative, because it allows the total variance to be partitioned into component parts and details the frequency region/s where the variance caused by the replication of the construction of the part is influential.

It also aligns the test scenario with the one conducted in the international standard to determine guideline values for repeatability and reproducibility. For example the highest part to part variance for the timber GRR was in the 3150Hz band (10.1dB), but this relatively large variance, the highest in any frequency band was only 51% of the total variance at this frequency. The contribution of the part to total variance at 800Hz and 1000Hz was 76% and 77%, respectively. The part is the most influential component of variance in this region but the variance measured in decibels was relatively low in comparison, 1.3dB and 3.5dB respectively. If the parts or test specimens are the same (in this case same construction and room size) and the part is seen to be the dominant element in the total variance measured then, improvements in instrumentation or measurement technique will not result in a significant reduction in the measurement uncertainty.

As an example of the importance of quantifying the part to part variance contribution consider the following, based on the timber 100Hz data:

Part to part variance = 10.1dB; repeatability variance = 2.5dB; reproducibility variance = 7.3dB. A new measurement technique is developed to reduce the reproducibility variance by 50%. Assuming no operator by part interaction, what will be the reduction in total standard deviation associated with the measurement process? See Table 8-29.

Table 8-29: Improving measurement technique - Impact of a 50% reduction in reproducibility variance

Var	σ^2_p	σ^2_r	σ^2_R	Standard Dev dB
dB	10.1	2.5	7.3	4.5
dB	10.1	2.5	3.7	4.0
				0.4

The improvement is 0.4dB.

If the impact of the part to part variance is not known e.g. all the testing is carried out on one part, a single test specimen, part to part variance would not be included. The only knowledge of uncertainty would be restricted to r and R. The predicted reduction in the total standard deviation, based on a 50% reduction in Reproducibility would be: see Table 8-30:

Table 8-30: Improving measurement technique – Ignoring Part to part variance - Impact of a 50% reduction in reproducibility variance

Var	σ^2_p	σ^2_r	σ^2_R	Standard Dev dB
dB	0	2.5	7.3	3.1
dB	0	2.5	3.7	2.5
				0.7

A 0.7dB reduction in standard deviation is more than would be obtained (0.4dB) when measuring a typical timber floor sample in the field.

For the single figure value the case is less extreme. For $D_{nT,w}$ $+C_{tr}$ the timber GRR part to part variance is 0.24dB; repeatability variance = 1.4dB; reproducibility variance = 4.4dB. This shows is that it is important to understand variance due to the part for the intended test sample, if developing a method of improving the reproducibility variance. In addition, if considering an inter-laboratory test experiment with a view to comparing the results with the guideline values for reproducibility, one should be aware of the part to part variance contribution and the impacts it has on the final results.

8.6 Conclusion

This chapter has detailed the results of the GRR study. The most popular types of timber and concrete floors have been tested in the field with an additional sample to extend the performance range and align it with the DOE guidance on conducting a GRR study [68]. The data is considered as representative of a typical sound insulation test in the field. In all cases, reasonable steps were taken to emphasize where the data may be adversely affected by external influences e.g. background noise.

The normality of the data also has been considered and, non-normal data have been identified. A reason for this may be that the time consuming measurement process has restricted sample sizes, leading to data which may appear non-normal even though the underlying population is known to be Gaussian. In all cases, ANOVA is required because it is relatively robust and insensitive to non-normal data. It also is useful because it highlights the presence of interaction between factors. Interaction between operator and part is present especially in the concrete floor GRR and when the additional test elements were added and the data combined. The significance of discovering interaction means that independence of input variables cannot be assumed. In addition, it also suggests that modelling or simulation techniques, to determine uncertainty, should be used with caution. Ignoring interaction will result in an inaccurate estimate of the reproducibility contribution.

The results of the GRR, discussed above, will be compared with the current standards for r & R and the new proposed standard for field testing sound insulation [9] in the next chapter.

9 Current Guideline Values – Standard Uncertainties

9.1 Introduction

In Chapter 8, the GRR data was analysed and the repeatability and reproducibility determined, together with more detailed information on the contribution from the operator and the part. In this chapter the results are compared with the new draft standard for measurement uncertainty in building acoustics [9]. The new definitions of uncertainty introduced by the draft standard are discussed and improvements suggested.

9.2 Repeatability & reproducibility – guideline values

The guideline values for r & R for measuring airborne sound insulation in the laboratory and field are detailed in ISO 140-2: 1991. This document is currently under revision and will be replaced by ISO 12999 (Working Draft) 2012, which uses the same methods from BS5725 Parts 1 & 2 to calculate r & R. In a new approach, ISO12999 presents the standard uncertainty values for three inter-laboratory tests and advises on the methodology that should be used for inter-laboratory experiments.

The information on uncertainty guideline values from the new proposed standard is reviewed in this chapter as the GRR data have previously been compared to the ISO140-2 reference values in Chapter 7.

9.3 ISO 12999

The working draft of ISO 12999 – Part 1 Sound Insulation references GUM and follows current conventions in defining the “Standard Uncertainty” relating to testing sound insulation in laboratories. This is a different descriptor to that used in ISO140-2 as it is a standard deviation, not a variance term, so a direct comparison with the ISO standard curves cannot be made without an appropriate correction.

To illustrate the changes in the curves for reproducibility and repeatability for ISO140-2 and ISO12999 the square root is taken of the ISO140-2 variance terms for R & r and are shown, alongside the standard uncertainty for R & r from ISO 12999, in Figure 9-1 and Figure 9-2:

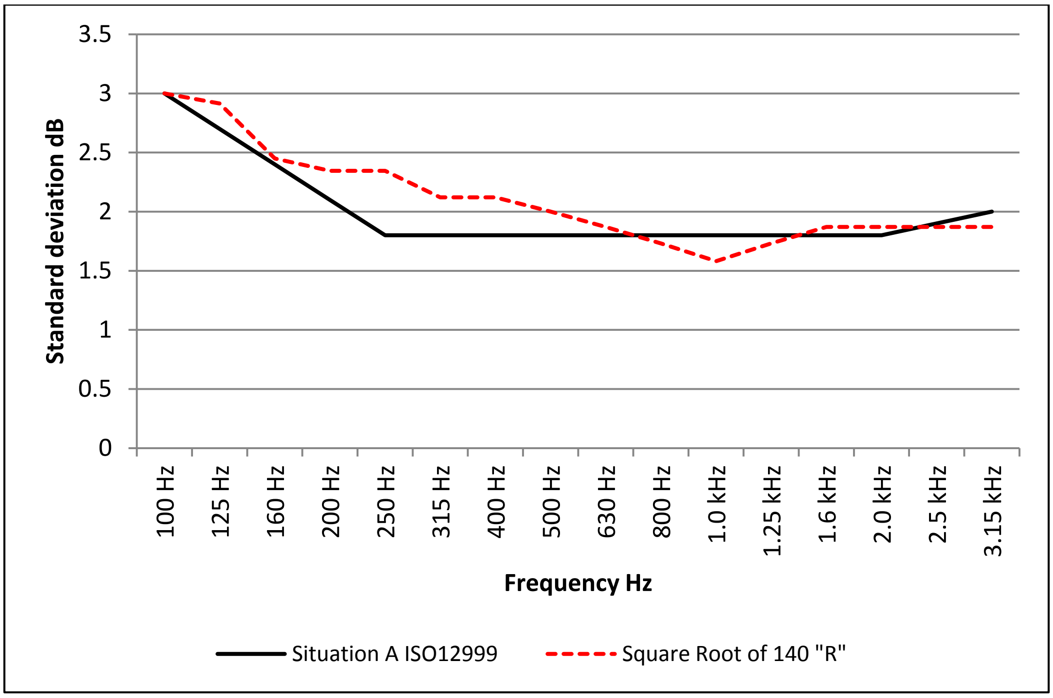


Figure 9-1: New ISO 12999 Reproducibility compared with square root of variance terms from ISO140-2

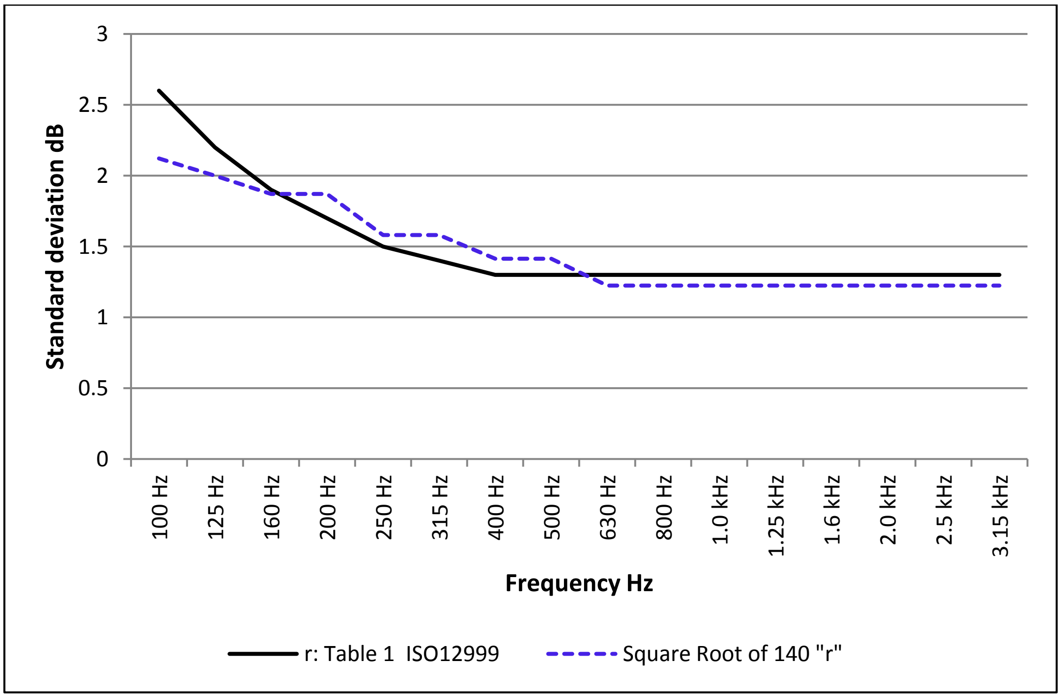


Figure 9-2: New ISO 12999 Repeatability compared with square root of variance terms from ISO140-2

The ISO12999 draft values for reproducibility are lower between 125 – 630Hz, higher at 1000Hz and similar above 1250Hz. For repeatability the values are higher at 100 – 125Hz, lower between 200 – 500Hz and broadly similar above 630Hz.

ISO 12999 defines also three measurement situations for inter laboratory studies. These are detailed in Table 2 in ISO 12999 and are reproduced in Table 9-1 :

Table 9-1: Inter Laboratory measurement situations - ISO12999: 2012

Situation	Description
A	Situation A is that a building element is to be characterized by measurements. In this case, the measurand is defined by the relevant part of ISO 140 including all additional requirements e.g. for the measurement equipment and especially for the test facilities. Therefore, all measurement results that may be obtained in another test facility or building also comply with this definition. The standard uncertainty thus is the standard deviation of reproducibility as determined by inter-laboratory measurements.
B	Situation B is described by the case that 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. The standard uncertainty obtained for this situation is called in-situ standard deviation.
C	Situation C handles the case that the measurement is simply repeated in the same test facility by the same operator using the same equipment. The standard uncertainty is the standard deviation of repeatability as determined by inter-laboratory measurements.

In order to provide a meaningful comparison, the appropriate guideline values need to be selected. Situation A and B describe different forms of reproducibility with “A” providing a better match for the GRR, as it allows for part to part variability as well as operator variability in the reproducibility component. Situation C describes repeatability but the values for Situation “C” only represent within laboratory repeatability and not the total repeatability across all laboratories. A closer match to the GRR design is detailed in Table 1 of ISO12999, see

Table 9-2: : Reproduction of Table 1 from ISO/DIS 12999-1 illustrating maximum standard deviation of repeatability

Table 1 — Maximum standard deviation of repeatability

frequency Hz	maximum standard deviation of repeatability dB
50	4,0
63	3,5
80	3,0
100	2,6
125	2,2
160	1,9
200	1,7
250	1,5
315	1,4
400	1,3
500	1,3
630	1,3
800	1,3
1 000	1,3
1 250	1,3
1 600	1,3
2 000	1,3
2 500	1,3
3 150	1,3
4 000	1,3
5 000	1,3

Table 1 from ISO/DIS 12999-1 is based on the "total average" $\bar{\bar{y}}$, where Laboratory x carries out n_x repeated measurements. These represent the maximum repeatability situation and reflect the pooled GRR repeatability data for all operators. Both forms of repeatability data are detailed in Table 9-3:

Table 9-3: Repeatability & Reproducibility values for laboratory tests (airborne sound insulation) ISO 12999.

dB	**100 Hz**	**125 Hz**	**160 Hz**	**200 Hz**	**250 Hz**	**315 Hz**	**400 Hz**	**500 Hz**	**630 Hz**	**800 Hz**	**1.0 kHz**	**1.25 kHz**	**1.6 kHz**	**2.0 kHz**	**2.5 kHz**	**3.15 kHz**
r: Situation A	3.0	2.7	2.4	2.1	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.9	2.0
r: Situation C	1.4	1.2	1	0.9	0.8	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
r: ISO12999 Table 1	2.6	2.2	1.9	1.7	1.5	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3

The GRR data are shown, with the ISO12999 values overlaid, in Figure 9-3 & Figure 9-4:

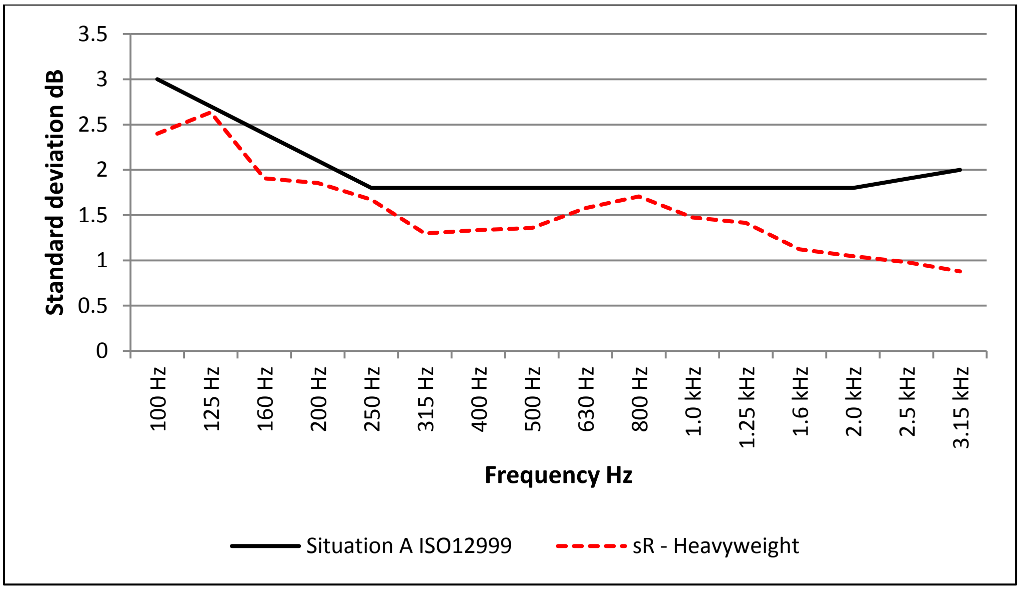


Figure 9-3: Reproducibility Comparison: Concrete Floor values for R & Situation A ISO12999

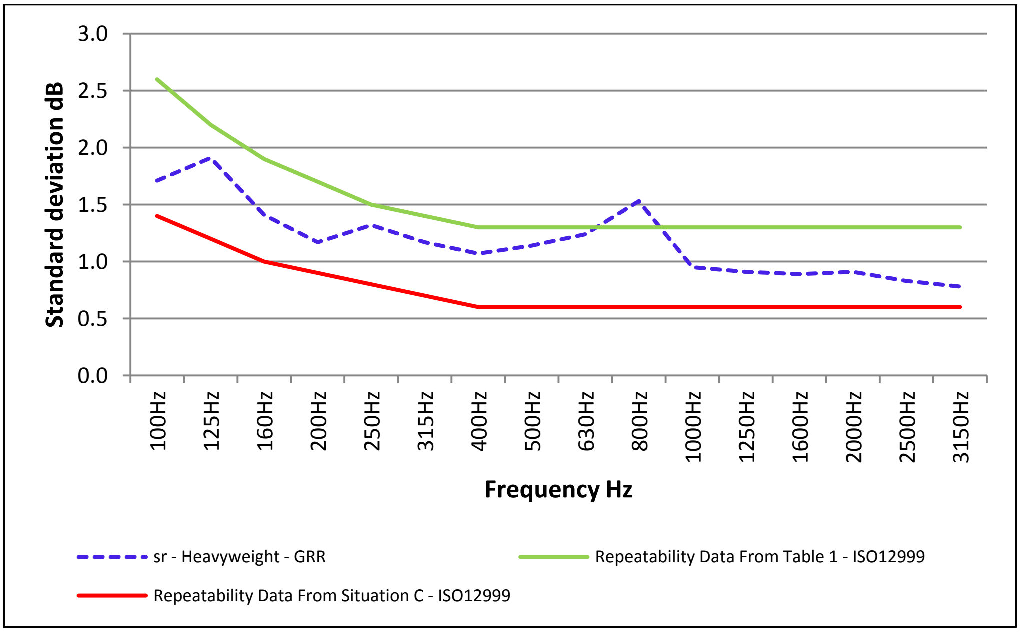


Figure 9-4: Repeatability Comparison: Concrete Floor values for r & Situation C & Max ISO 12999

The concrete floor gives relatively good agreement with the ISO 12999 values of reproducibility. All calculated values are below the draft ISO12999 curve for test situation "A". The repeatability maximum guideline values published in ISO12999 are only exceeded at 800Hz. The repeatability represented by situation "C" are lower than the measured repeatability across the full frequency range.

The timber floor data are superimposed on the ISO12999 guideline values in Figure 9-5 & Figure 9-6:

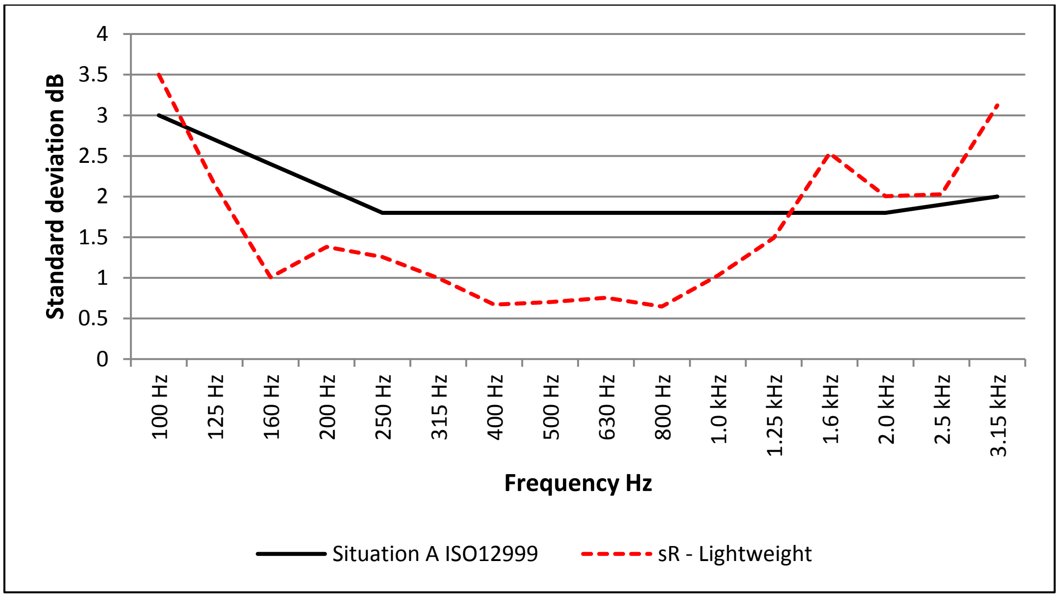


Figure 9-5: Reproducibility Comparison: Concrete Floor values for R & Situation A ISO12999

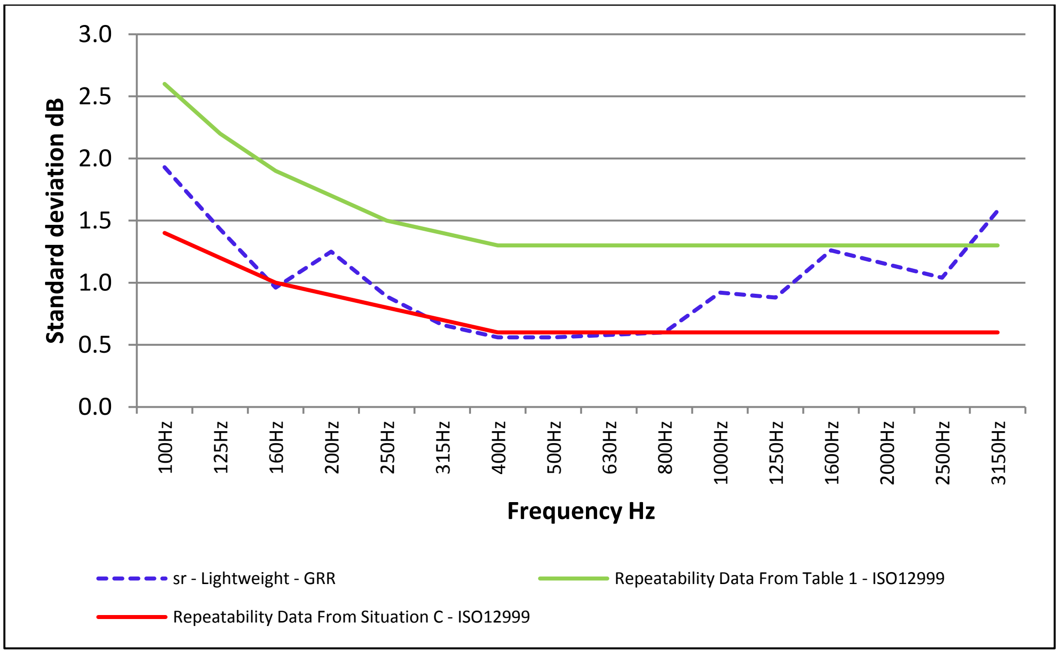


Figure 9-6: Repeatability Comparison: Timber Floor values for r & Situation C & Max ISO 12999

The timber floor gives relatively good agreement with the ISO 12999 values for reproducibility apart from the 100Hz band and above 1600Hz, where site background noise affected the measurements.

For repeatability the GRR data is below the maximum value curve apart from the 3150Hz band where measurements again are affected by site background noise. In the timber floor GRR the repeatability levels in some frequency bands are relatively low and in several third octave bands, “r” values are similar to those represented by situation “C” at 160Hz and 315 – 800Hz. Though measured standard uncertainties are higher than those described by any of the situations, the maximum standard should not necessarily be seen as incorrect.

ISO12999 recommends noting levels that are above the curves. It only advises considering a higher result as invalid if an error occurred e.g. if after the measurement the sound level meter was calibrated outside the tolerance range.

The GRR data generally gives good agreement with the standard uncertainties detailed in the draft ISO12999, Since the correct measurement and calibration procedures were followed by all operators, there is no reason to eliminate data which is above the guideline value curve and increases in the standard deviation of results can be explained by the test conditions that prevailed on site.

9.3.1 Single Figure Values

The ISO 12999 draft also presents standard uncertainties for single number values and Situation A, B and C are used to define the experimental test scenario. The reproducibility is described by situation A and repeatability is described by situation C. There is no maximum standard for repeatability detailed for the single figure values.

9.3.2 Reproducibility

Table 9-4 takes the single figure $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} reproducibility values for test situation “A” from ISO 12999-1 and compares them with the reproducibility values calculated for the timber and concrete GRR.

Table 9-4: Standard Uncertainties – Reproducibility - Single Number Values – ISO12999, Timber & Concrete Floors

Single number value: Reproducibility	Situation A dB	Timber dB	Concrete dB
$D_{nT,w}$	1.2	0.7	0.6
$D_{nT,w} + C_{tr}$	1.4	2.1	1.0

There is no straightforward mathematical relationship between the third octave band sound pressure levels measured and the single figure descriptor, and it does not necessarily follow that a data set that has lower standard uncertainties in all frequencies leads to a single figure descriptor which has a lower standard uncertainty. Conversely, third octave band standard uncertainties that are higher than the ISO12999 values do not necessarily result in a single figure descriptor with a higher standard uncertainty. As we have seen in previous chapters, two highly variable factors e.g $D_{nT,w}$ & C_{tr}, when added together can result in a descriptor with lower variability.

These situations are reflected in the GRR results. The standard uncertainties for the third octave band reproducibility data in concrete floor are all below the ISO12999 curve for situation A as seen in Figure 9-3. This is replicated in each of the single figure descriptors, which have standard uncertainties of 0.6dB and 1.0dB for $D_{nT,w}$ and $D_{nT,w}$ + C_{tr}, respectively, compared to the Situation “A” levels of 1.2dB and 1.4dB respectively. For the timber case, the reproducibility data shown in Figure 9-5 showed several frequency bands with higher levels of standard uncertainty, but the single figure value for $D_{nT,w}$ (0.7dB) is lower, compared to Situation A although the level for $D_{nT,w}$ +C_{tr} (2.1dB) is higher.

9.3.3 Repeatability

If the single figure descriptors are calculated from Situation “C” it is clear that this represents the lowest levels of variability and the test data are expected to be higher.

Table 9-5: Standard Uncertainties – Repeatability - Single Number Values – ISO12999, Timber & Concrete Floors

Single number value: Repeatability	Situation C dB	Timber dB	Concrete dB
$D_{nT,w}$	0.4	0.6	0.7
$D_{nT,w} + C_{tr}$	0.5	1.2	0.9

The repeatability data for concrete floors generally fall below the maximum standard repeatability but are greater than the Situation "C" descriptors for the single laboratory test. The timber floor repeatability data have individual third octave bands that match Situation C values but others are significantly above. The single number values calculated from this data are above the Situation C levels. This was expected as the draft ISO 12999 is inconsistent in that it provides no maximum standard single number values as it does for the third octave band values, only situation A, B and C are supplied. See Table 9-6.

Table 9-6 Standard uncertainties for single-number values according to ISO 717-1 (Table 3 - replicated from Draft ISO/DIS 12999-1: 2012):

	situation A dB	situation B dB	situation C dB
R_w, R'_w, D_{nw}, $D_{nT,w}$	1,2	0,8	0,4
(R_w, R'_w, D_{nw}, $D_{nT,w}$) + C	1,3	0,9	0,4
(R_w, R'_w, D_{nw}, $D_{nT,w}$) + C_{tr}	1,4	1,0	0,5

9.4 Conclusion

This chapter has considered the current guideline values for uncertainty in ISO 140-2 and the new proposed standard ISO12999

ISO 12999 incorporates the references from the latest version of GUM and uses standard uncertainty as a descriptor for variability in measurement data. It also differentiates between reproducibility values and introduces the concept of an "in-situ" standard deviation in Situation B, where different laboratories measure the same element in one location. This helps reduce confusion where round robin tests have previously brought several laboratories to one location, to calculate the reproducibility from the resulting data and provide comparisons with "R" without taking into account the fact that some of the components of variance are missing e.g. Lang et al [36], Hall [38], Hoffmeyer [31].

One omission is noted in the draft: the single value sound insulation for repeatability, corresponding to the maximum standard third octave band data (see Table 1 instanced), inclusion of which would provide a comprehensive set of guideline values for the most common test situations. It is understood that this data will be provided once the investigations and current building acoustic research on measurement uncertainty has been completed.

In both current and proposed standards the GRR data is comparable with the relevant repeatability and reproducibility values. In addition, evaluation of the data against the new definitions of uncertainty in the latest draft of ISO12999 reinforces the view that the data is representative of what would be expected from a site test. Most of the higher uncertainty values are due to external influences, primarily the systematic error introduced by background noise. Any differences and it is concluded that obtaining values higher than the guidelines does not invalidate the measurement.

The next chapter looks at the GRR DOE and proposes an alternative quicker method of identifying the significance of the components of variance in the measurement.

10 Alternative method: testing significance of factors

10.1 Introduction

In Chapter 9 the GRR data was compared with the current and proposed guideline values for measurement uncertainty and the data shown to be consistent with expectations.

The comparison gives confidence to the statement that the ANOVA results are a true representation of the total uncertainty in the measurement process, including the contribution of the operator, part and instrumentation.

As a further check, a reduced experimental model was trialled using the concrete floor. The "Latin-square" uses a reduced factorial approach and serves to minimise the data sampling, whilst retaining statistical confidence in the results.

10.2 Latin-square

Latin-squares are an efficient method of blocking factors and are often used where an experimenter has one factor of interest and wants to control two (or more) sources of variation [70, 113, 128]. They can significantly reduce the number of runs required in an experiment but they may produce erroneous results if significant interaction between two or more variables is present. The statistical model for the Latin-Square is detailed in Montgomery [70, 113]:

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \begin{cases} i = 1, 2, \ldots., p \\ j = 1, 2, \ldots., p \\ k = 1, 2, \ldots., p \end{cases} \qquad \text{10-1}$$

The analysis of variance consists of partitioning the total sum of squares of the N=p^2 observations i.e N = 25 of p = 5, into components for rows, columns, treatments and error [70]:

$$SS_T = SS_{Rows} + SS_{Columns} + SS_{Treatments} + SS_E \qquad \text{10-2}$$

With degrees of freedom: $p^2 - 1 = p - 1 + p - 1 + p - 1 + (p - 2)(p - 1)$.

Assuming the distribution of the errors is normal and has a mean of zero and a variance σ^2 : this is represented by equation 10-3.

$$\epsilon_{ijk} \; is \; NID(0. \sigma^2) \qquad \textit{10-3}$$

The appropriate test statistic, for no differences in treatments means, is:

$$F_0 = \frac{MS_{Treatments}}{MS_E} \qquad \textit{10-4:}$$

Which is distributed as: $F_{p-1,(p-2)(p-1)}$, under the null hypothesis.

The test statistic gives a p-value which indicates the significance level. The p-value or calculated probability is the estimated probability of rejecting the null hypothesis (H_0) of a study question when that hypothesis is true. The null hypothesis in this case is an hypothesis of "no difference" e.g. no difference between operators. A sample output table is given in Figure 10-1 when the general linear model analysis of variance is used:

General Linear Model: 100 Hz versus Operator, Surface, Kit

Factor	Type	Levels	Values
Operator	random	5	1, 2, 3, 4, 5
Surface	random	5	1, 2, 3, 4, 5
Kit	fixed	5	1, 2, 3, 4, 5

Analysis of Variance for 100 Hz, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Operator	4	55.278	55.278	13.820	2.25	0.124
Surface	4	139.266	139.266	34.817	5.67	0.008
Kit	4	5.218	5.218	1.305	0.21	0.927
Error	12	73.699	73.699	6.142		
Total	24	273.462				

Figure 10-1: Minitab statistical output table: Latin-square analysis of 100Hz third octave band carried out on day 1.

10.2.1 GRR Latin-square design

The concept of the Latin-Square design is to provide a special sub set of a full factorial model that, if correctly designed, allows the significant factors to be identified. The square design comes from the need to match the number of factors with the sources of variation. For example, an experiment with three factors would be in our case: operator; test construction (Room); test kit; the Latin-square was carried out on the concrete floor site. It is noted that there was some interaction between parts and the operators, mainly in the low frequency bands 100Hz – 250Hz. (See Figure 10-2): the results in these frequencies should therefore be treated with a degree of caution.It should also be noted that in order for there to be a significant statistical test result there should be some variation in the measured result recorded. i.e. if all the floors perform identically no significance would be attributable to any of the factors under observation. Therefore the greater the variation in the measured results, the more potential there is to attach significance to one of the factors, assuming the reason for the variation was not just random error. Less variation in the data reduces the ability to determine significance through the random statistical experimental noise.

.

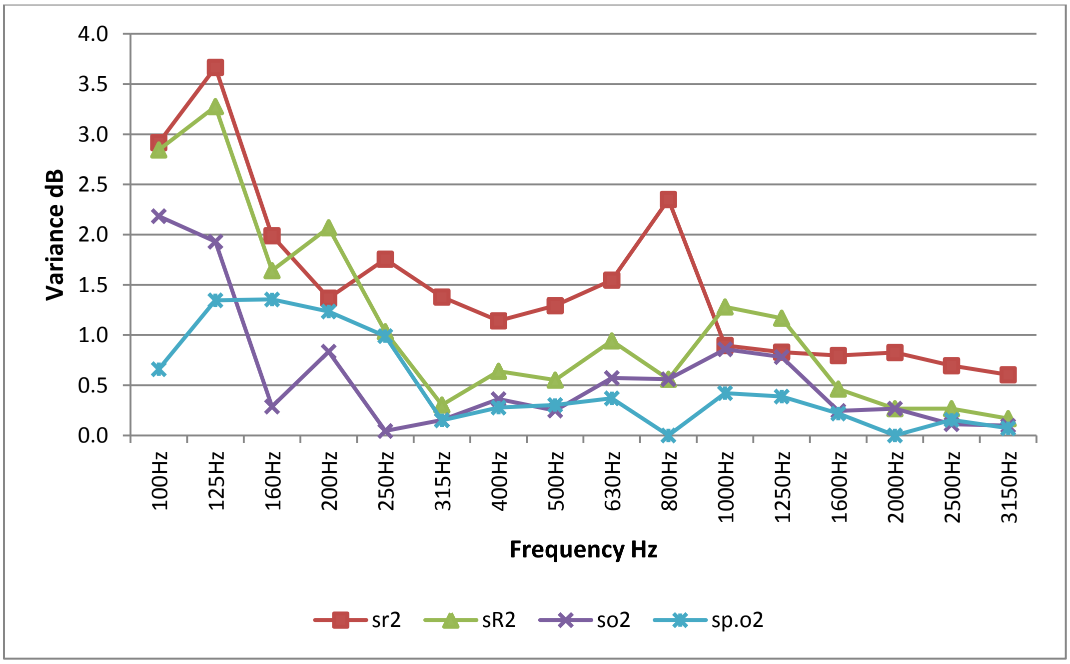


Figure 10-2: Concrete floor components of Variance - σ_{GRR}^2, σ_r^2, σ_R^2, σ_o^2, $\sigma_{p.o}^2$ - 100-3150Hz: From Ch8 showing interaction

The Latin-square test was conducted three times on site. Each of the operators tested every floor with one of the test kits (see matrix combinations used for each of the three test days in Figure 10-3). However, not all combinations of test kit, engineer and room were used. With 5 engineers, 5 rooms and 5 test kits a full "Factorial" design with one observation on each three way combination would require 125 observations. This is impractical in the case of sound insulation testing and statistically inefficient. The Latin-square method is an efficient method of showing how levels of one factor are assigned to combinations of levels of the other two factors. In this case the Latin-square requires 25 runs, which must meet certain conditions to make sure the experiment is balanced.

Day 1	*Operator*				
Surface	SP	SM	AA	BW	MT
A	5	2	1	3	4
B	3	5	4	1	2
C	4	3	2	5	1
D	1	4	5	2	3
E	2	1	3	4	5

Day 2	*Operator*				
Surface	SP	SM	AA	BW	MT
A	1	3	4	5	2
B	5	2	1	4	3
C	4	5	2	3	1
D	2	4	3	1	5
E	3	1	5	2	4

Day 3	*Operator*				
Surface	SP	SM	AA	BW	MT
A	3	5	1	4	2
B	2	3	5	1	4
C	4	1	2	5	3
D	1	4	3	2	5
E	5	2	4	3	1

Figure 10-3: Latin-square Matrices for each of the three daily test situations

10.2.2 Test data

In each of the 25 combinations a dB level is calculated and recorded. It could be the $D_{nT,w}$ value, the $D_{nT,w}$ + C_{tr} or the D_{nT} in each of the 16 third octave bands from 100Hz – 3150Hz, for each of the 16 observations. An example of the 25 sound insulation values for the $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} values is shown in Table 10-1.

Table 10-1: Results of Latin-square Test Day 1 - $D_{nT,w}$ and $D_{nT,w}$ +C_{tr}

Day Set	Kit	Operator	Surface	DnTw	DnTw +Ctr	Ctr	Operator
1	5	1	1	59	53	-6	SP
1	3	1	2	61	53	-8	SP
1	4	1	3	58	54	-4	SP
1	1	1	4	57	51	-6	SP
1	2	1	5	57	53	-4	SP
1	2	2	1	60	52	-8	SM
1	5	2	2	61	53	-8	SM
1	3	2	3	57	52	-5	SM
1	4	2	4	57	50	-7	SM
1	1	2	5	57	53	-4	SM
1	1	3	1	60	51	-9	AA
1	4	3	2	60	52	-8	AA
1	2	3	3	57	54	-3	AA
1	5	3	4	56	51	-5	AA
1	3	3	5	57	53	-4	AA
1	3	4	1	59	50	-9	BW
1	1	4	2	60	51	-9	BW
1	5	4	3	57	52	-5	BW
1	2	4	4	57	52	-5	BW
1	4	4	5	56	51	-5	BW
1	4	5	1	62	53	-9	MT
1	2	5	2	60	50	-10	MT
1	1	5	3	57	51	-6	MT
1	3	5	4	58	52	-6	MT
1	5	5	5	57	51	-6	MT

A spreadsheet summary of the p-value output from the statistical software for the D_{nTw} and $D_{nT,w}$ + C_{tr} values is shown in Table 10-2:

Table 10-2: p-values for $D_{nT,w}$ and $D_{nT,w}$ + C_{tr}

Latin Square Analysis - Summary Table			
Red Text < 0.05			
Latin Square 1		$D_{nT,w}$	$D_{nT,w}$ + C_{tr}
p-values			
Operator		0.306	0.308
Surface		0.000	0.499
Kit (response)		0.761	0.872
Latin Square 2		$D_{nT,w}$	$D_{nT,w}$ + C_{tr}
p-values			
Operator		0.215	0.566
Surface		0.000	0.151
Kit (response)		0.996	0.913
Latin Square 3		$D_{nT,w}$	$D_{nT,w}$ + C_{tr}
p-values			
Operator		0.582	0.147
Surface		0.001	0.169
Kit (response)		0.994	0.976

10.2.2.1 $D_{nT,w}$

From the observed test data it is concluded that the test kits are not significantly different (P = 0.761) and the operators are not significantly different (where the p-value is P = 0.306). This might indicate they are well trained and are therefore able to provide repeatable results

It is also concluded that the floors tested **are** significantly different at the 5% and 1 % level (where the p-value is P = 0.000).

10.2.2.2 $D_{nT,w} + C_{tr}$

In this case, it is concluded that the kits are not significantly different (P = 0.872) and the operators are not significantly different (P = 0.308).

The floors are not significantly different (P=0.151 to 0.499), which is unexpected, as the results for $D_{nT,w}$ indicated the floors tested were significantly different at a high confidence level for each of the three experiments conducted. The inclusion of the spectrum adaptation term reduces the variability between rooms from 56 – 62dB $D_{nT,w}$ to 50 – 54dB $D_{nT,w} + C_{tr}$ (see the results table in Table 10-1) and the ability for the latin square experiment to determine significance between factors falls. . This is not an isolated conclusion because the result is repeated for the Latin Square test carried out on each of the three days. Using the single figure value $D_{nT,w} + C_{tr}$ as the measurand (Sound Insulation performance indicator) makes it more difficult to attach significance to any of the factors through the statistical analysis because the results are more alike and the test is less able to apportion significance at a reasonable confidence level against the presence of experimental error.

The third octave band data was also analysed and the p-values for each of the three Latin-square experiments is given in Table 10-3, Table 10-4 & Table 10-5.

The significance of each factor in the determination of the D_{nT} frequency result is tested and compared against the 1% and 5% level.

Table 10-3: Latin-square Day 1

dB	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Operator	0.12	0.10	0.70	0.10	0.20	0.88	0.11	0.52	0.82	0.65	0.35	0.27	0.47	0.24	0.29	0.78
Floor	0.01	0.04	0.09	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kit (response)	0.93	0.53	0.61	0.72	0.19	0.57	0.94	0.57	0.18	0.10	0.77	0.26	0.78	0.30	0.16	0.83

Table 10-4: Latin-square Day 2

dB	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Operator	0.14	0.32	0.85	0.23	0.42	0.32	0.06	0.05	0.22	0.04	0.02	0.00	0.02	0.22	0.06	0.02
Floor	0.00	0.32	0.56	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kit (response)	0.27	0.94	0.45	0.68	0.67	0.48	0.48	0.89	0.90	0.28	0.59	0.66	0.64	0.61	0.53	0.74

Table 10-5: Latin-square Day 3.

dB	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Operator	0.03	0.26	0.25	0.47	0.73	0.14	0.19	0.52	0.17	0.32	0.02	0.40	0.33	0.82	0.37	0.43
Floor	0.00	0.07	0.15	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kit (response)	0.66	0.69	0.73	0.08	0.70	0.21	0.92	0.86	0.73	0.97	0.91	0.94	0.92	0.89	0.80	0.88

For the D_{nT} frequency results the floor tested is significant at the 1% level, across all Latin-squares at 200Hz and above. There are lower significance levels for the floor below 200Hz where it is less clear what factor has the greatest influence on the test result. Individual results from the tests show for the floor on Day 1, the 160Hz band was not significant and on days 2 and 3, the 125Hz and 160Hz bands did not show significance.

The operators on the experiment carried out on Day 2 showed some significance, mainly at the 5% level, in the mid to high frequency range. The test kits were not significant at any frequency on any of the test days, the measured result was therefore not reliant on which test kit was being used.

For the frequency data, the simple Latin-square analysis gives, with a high degree of confidence (p<1%), that the sound insulation values recorded are reliant on the floor being tested. The operator data appears to have some influence in selected frequency bands, but the test kit does not appear to feature as a significant factor at any frequency on any of the days. There are no significant differences that exist that depend on the test kits used.

The Latin square experiment confirms the result discovered previously in the GRR ANOVA in Chapter 8 that the part to part variability was confirmed to be the major component of variance in the concrete floor GRR, this is illustrated graphically ; see Figure 10-4.

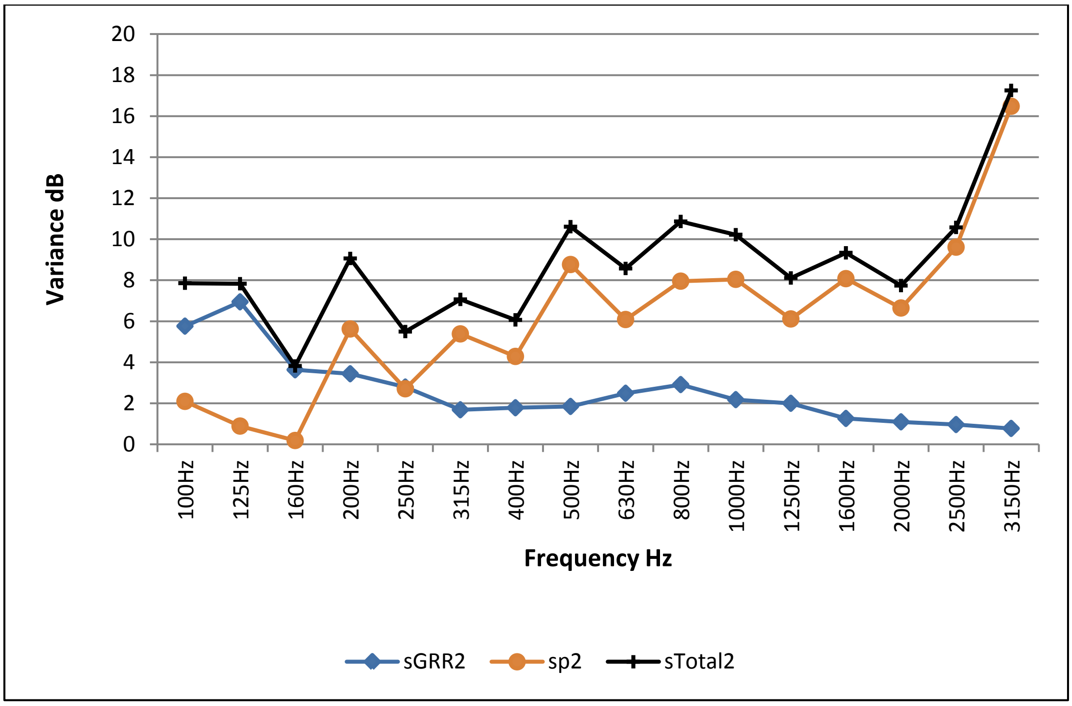


Figure 10-4: Concrete Heavyweight Floor - Components of Variance - σ_{GRR}^2, σ_p^2, σ_{Total}^2

The Latin-square test shows that this conclusion could have been obtained relatively quickly and with a high degree of confidence.

10.3 Conclusions

The Latin-square test is a robust statistical test of significance of the main factors likely to contribute to the final sound insulation value. It is relatively quick and easy to execute and allows the experimenter to determine what is likely to be the significant factor that determines the test result and therefore will be likely to be a major component of variance in any test situation, without having to carry out a full GRR.

In the case of the concrete floor, the results show that the floor element, is statistically significant (i.e. different) at the 1% significance level for the single figure descriptor $D_{nT,w}$ and for the majority of cases at the 1% level across the third octave band frequency range 100-3150Hz. It is noted that there is interaction in the concrete floor test data which is highest in the 125Hz and 160Hz bands but features at some appreciable level greater than 1dB, between 125-250Hz and therefore the test statistics in this lower frequency range should be treated with caution.

In the Latin square test experiment it is clear that the floor, or test element, which embodies the room shape, size, volume and the construction, is the key influence on the sound insulation result.

An unusual finding, revealed using this technique, was when the spectrum adaptation term C_{tr} is introduced .

The calculated results based on the $D_{nT,w} + C_{tr}$ performance of the floors showed that no individual factor; floor element, operator or kit was judged significant in determining the single figure result. This is likely to be due to the effect the spectrum adaptation term has on reducing the variability of the result for this particular test sample as previously discussed. Less variable data make it more difficult to obtain a statistically significant result using the Latin Square experiment analysis technique.

This chapter has shown that the Latin-square experiment confirms the findings of the concrete GRR that for the concrete floor test sample the part is the main factor affecting the test result and is likely to be the major component of variance. Because of the brevity of the testing process, it could be a useful statistical tool when used as a precursor to the GRR. The DOE could be fine tuned using the Latin-square, rather than from the GRR. If the opportunity had been presented to carry out the Latin Square on the timber floor there may have been a different result as the floors were chosen to be identical, including the rooms shape and size. In that case there may have been more significance attached to the operators or test kit rather than the floor being tested. The next chapter applies confidence intervals to the GRR results, in order to quantify the uncertainty associated with the point estimates of variance.

11 Confidence Intervals

11.1 Introduction

Chapter 10 illustrated a method whereby the significance of a factor can be easily determined and its importance checked using a small subset of the full factorial model. Previous chapters have also shown that the variability in the measurement process can be estimated quantitatively using statistical point estimates of the variances through ANOVA.

To be of practical use the size of the uncertainty or interval surrounding these estimates of variance is required. Confidence intervals are used to describe the uncertainty associated with the estimation process. This chapter discusses the options available and calculates the uncertainty for the 95% confidence coefficient. It is noted that the confidence limits calculated may not be of much practical use at all frequencies but the calculations are carried out to illustrate where the calculations do and do not work giving examples and explanations of the reasons why. It comments generally on the suitability and usefulness of the methods available to determine the confidence limits for variance components and identifies any shortcomings of the current procedures with reference to this research.

11.2 Confidence Intervals

So far, the thesis work has been focussed on determining the point estimates of the components of variance. The point estimate, in the absence of more or better data, is the best available estimate for a given parameter. Under most conditions, it is possible to supplement this figure with a statement about the uncertainty of the estimate. This is usually done by detailing a confidence interval [70, 113, 128] formed around the point estimate together with the degree of confidence. The degree of confidence is a probability expressed as a percentage e.g. 95%. In BS5725 no confidence interval for μ (overall mean of a test sample) is given, only repeatability (r) and reproducibility (R) limits are defined[66].

Current methods of determining the interval in which a point estimate (y) resides rely on applying simple multipliers to the standard uncertainties, based on the number of standard deviations from the mean. Using GUM and M3003 [76] this is simplified by the use of "coverage factors" which are the same as confidence intervals, as detailed in Table 11-1 below:

Table 11-1: Coverage Factors based on Coverage probability: M3003

Coverage probability p	Coverage factor k
90%	1.64
95%	1.96
95.45%	2.00
99%	2.58
99.73%	3.00

The table gives the coverage factor necessary for various levels of confidence, for a normal distribution. It would normally be applied to the combined standard uncertainty $u_c(y)$ to obtain the expanded uncertainty.

In GUM it is often necessary to find the upper and lower bound for the estimate (y) using the concept of expanded uncertainty. It is defined in GUM [6] as:

$$U_p = k.u_c(y) \qquad \textit{11-1:}$$

Where: U_p is the expanded uncertainty of output estimate (y) that defines an interval around the point estimate y: Y = y ± U_p , having a high specified level of confidence probability p.

k = a coverage factor

$u_c(y)$ = the combined standard uncertainty of (y). It depends on the uncertainties of the input variables $u(x_i)$:

As an example, the coverage factor for 95% and 99% are used to calculate the expanded uncertainty for the timber floor case, for $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} and are detailed in Table 11-2:

Table 11-2: Timber Floor Single Figure Standard Uncertainty (Total) & Expanded Uncertainty

dB	σ_{Total}	Coverage Probability 95%	Coverage Probability 99%	Expanded Uncertainty	
		Coverage Factor	Coverage Factor	U	U
		k	k	95%	99%
$D_{nT,w}$	1.1	1.96	2.58	2.1	2.7
$D_{nT,w}$ + C_{tr}	2.5	1.96	2.58	4.8	6.4

e.g. The timber standard uncertainty for $D_{nT,w}$ is 1.1dB multiplied by 1.96 gives an expanded uncertainty at the 95% confidence level of approximately 2.1dB

The coverage factor provides a symmetrical confidence interval around the sample mean, This simple approximation may not offer an accurate representation or fully describe the interval precision for components of variance determined by ANOVA. The GUM method has received criticism from Mahn [42] who points out that this may mask the uncertainty because very small sample sizes will have a large uncertainty in the sample standard deviation. Also the coverage factor approach does not identify the confidence intervals in the input values. This requires a more detailed approach.

11.3 Discussion

There are several methods for calculating the confidence intervals of variance components. Some simulation techniques by Leiva et al [129], the generalized confidence intervals (GCI) introduced by Weerahandi [130] and restricted maximum likelihood method (REML) originally proposed by Bartlett [131] may not, as pointed out by Borror et al [71], maintain the stated levels of confidence in the interval. They also require specialist software and computer simulation in order to model the intervals.

Two methods are selected to offer estimates of the variance components: the Satterthwaite approximation and the Modified Large Squares method.

Both these methods allow relatively straightforward calculations to be carried out on spreadsheets and are chosen to inform this thesis for this reason. NB: The Satterthwaite approximation in the context of confidence limits should not be confused with the popular method of calculating effective degrees of freedom which is widely referenced in the calculation of measurement uncertainty.

11.3.1 Satterthwaite Approximation

A method of constructing a confidence interval for variance components in a GRR is proposed by Montgomery et al [69]. It is based on the method originally developed by Satterthwaite [132, 133]. The Satterthwaite confidence interval, for the estimate of variance component $\hat{\sigma}^2_{repeatability}$, from Montgomery et al [69]is:

$$\frac{df_{E}.\,\hat{\sigma}^2_{repeatability}}{\chi^2_{\alpha/2,df_E}} \leq \hat{\sigma}^2_{repeatability} \leq \frac{df_{E}.\,\hat{\sigma}^2_{repeatability}}{\chi^2_{1-\alpha/2,df_E}} \qquad \textit{11-2:}$$

where df_E = error degrees of freedom in the ANOVA design.

$\hat{\sigma}^2_{repeatability}$ = MS_R

$\chi^2_{\alpha,k}$ = Chi-squared distribution with k degrees of freedom that has α% probability to the right.

NB: remember df_E increases from 60 to 80 in the reduced model without interaction (20 d.f. taken up by the operator by part interaction)

For example, with reference to the $D_{nT,w}$ ANOVA table for timber floors and calculating the 95% confidence interval:

Two-Way ANOVA Table Without Interaction					
Source	DF	SS	MS	F	P
Test Scenari	5	25.2556	5.05111	15.9369	0.000
Engineer	4	36.7778	9.19444	29.0096	0.000
Repeatability	80	25.3556	0.31694		
Total	89	87.3889			

Figure 11-1: ANOVA Table for Timber Floor GRR: $D_{nT,w}$

$$\hat{\sigma}^2_{repeatability} = 0.317$$

$$\frac{80\ (0.317)}{106.62} \leq \hat{\sigma}^2_{repeatability} \leq \frac{80\ (0.317)}{57.15} = 0.24dB \leq \hat{\sigma}^2_{repeatability} \leq 0.44dB$$

By subtracting the upper from the lower value and dividing by repeatability (0.44-0.24 = 0.2/0.317 = 0.63) the width of the confidence interval can also be expressed as percentage of $\hat{\sigma}^2_{repeatability}$ (63%).

Montgomery et al [69] emphasize that there may not be an exact confidence interval for $\hat{\sigma}^2_{reproducibility}$ and $\hat{\sigma}^2_{gauge}$ but approximate confidence intervals can easily be obtained.

For the full model (with interaction), the approximate 100(1 – α)% Satterthwaite confidence interval on $\hat{\sigma}^2_{reproducibility}$ from Montgomery et al [69] is:

$$\frac{\upsilon df_{E.}\,\hat{\sigma}^2_{reproducibility}}{\chi^2_{\alpha/2,\upsilon}} \leq \hat{\sigma}^2_{reproducibility} \leq \frac{\upsilon df_{E.}\,\hat{\sigma}^2_{reproducibility}}{\chi^2_{1-\alpha/2,\upsilon}} \qquad 11\text{-}3$$

where:

$$\upsilon = (\hat{\sigma}^2_{reproducibility})^2 \left\{ \frac{(\frac{1}{np})^2 MS_O^2}{o-1} + \frac{(p-1)/np)^2 MS_{Op}^2}{(o-1)(p-1)} + \frac{(1/n)^2 MS_R^2}{op(n-1)} \right\}^{-1} \qquad 11\text{-}4$$

The approximate 100(1 – α)% Satterthwaite confidence interval on $\hat{\sigma}^2_{gauge}$ [69] is:

$$\frac{udf_{E.}\,\hat{\sigma}^2_{gauge}}{\chi^2_{\alpha/2,u}} \leq \hat{\sigma}^2_{gauge} \leq \frac{udf_{E.}\,\hat{\sigma}^2_{gauge}}{\chi^2_{1-\alpha/2,u}} \qquad 11\text{-}5$$

where:

$$u = (\hat{\sigma}^2_{gauge})^2 \left\{ \frac{(\frac{1}{np})^2 MS_O^2}{o-1} + \frac{(p-1)/np)^2 MS_{Op}^2}{(o-1)(p-1)} + \frac{((n-1)/n)^2 MS_R^2}{op(n-1)} \right\}^{-1} \quad \text{11-6}$$

11.3.1.1 Interval width

The confidence intervals produced by the Satterthwaite approximation can be relatively wide, especially for reproducibility. This is due to the degrees of freedom in the model [69].

For $\sigma^2_{repeatability}$, the size of the confidence interval is related to the number of levels of the factors (parts, p and operators, o) and the amount of replication (n). From equation 11-2:

The confidence interval is determined by the ratios:

$$\frac{m}{\chi^2_{\alpha/2,m}} \; and \; \frac{m}{\chi^2_{1-\alpha/2,m}} \quad \text{11-7}$$

For the full model, m = *op(n-1).* Figure 11-2 gives an example using the concrete GRR data.

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Set	5	257.522	51.5044	63.1526	0.000
Engineer	4	23.822	5.9556	7.3025	0.001
Test Set * Engineer	20	16.311	0.8156	1.6682	0.066
Repeatability	60	29.333	0.4889		
Total	89	326.989			

Figure 11-2: ANOA Table for Concrete Floor GRR: $D_{nT,w}$

In the full model for concrete floors, $D_{nT,w}$, p = 6, o = 5 and n = 3 therefore m = 60 which, using equation 11-7 gives a confidence interval for the concrete floor $D_{nT,w}$ of:

$$0.35dB \leq \hat{\sigma}^2_{repeatability} \leq 0.72dB \qquad \textit{11-8}$$

For $\sigma^2_{repeatability}$ the precision of the confidence interval relies on the size of the product of *op(n-1). W*hen *n* is small, the biggest increase in m will be due to increase in *n*. Montgomery et al [69] have shown that there are diminishing returns for increases in m when m is large (>50). Therefore the improvement to the width of the confidence interval is only small for further increases in *n* when m is 60.

For $\sigma^2_{reproducibility}$ the impact on the precision of the confidence interval for *p*, *o* and *n* depends on v, see equation 11-4. The exact value for v is calculated for each third octave band, the lower v is, the wider the confidence interval. Montgomery et al [69] consider a special case that illustrates the limit of v as *p* tends to infinity as:

$$\lim_{p\to\infty} v = (o-1)\left(1+\frac{\sigma^2_{op}}{\sigma^2_o}\right) \qquad \textit{11-9}$$

Therefore when the ratio of σ^2_{op} to σ^2_o is near zero i.e. the interaction term is small in comparison with the variance of the operator (as it is in the case of the Timber GRR), the number of degrees of freedom in the confidence interval in equation 11-9 tends to (o – 1), in our case 4.

With o = 5 and m = 4, the approximate expected minimum width of the confidence interval for $\sigma^2_{reproducibility}$ can be calculated from the Chi-squared table, which is shown below:

$0.359\hat{\sigma}^2_{reproducibility}$ to $8.333\hat{\sigma}^2_{reproducibility}$

The confidence interval is almost 8 times (800%) the value of $\hat{\sigma}^2_{reproducibility}$. In order to reduce this, the number of operators would therefore have to be increased. Montgomery and Runger [69] also demonstrate that the same is true for replicates so that neither increasing the number of parts or replications will be effective in improving the estimate of $\sigma^2_{reproducibility}$ when σ^2_{op} is small in relation to σ^2_o.

11.3.1.2 Interpolation and confidence limits

The approximation of the interval comes in practice because υ and u will rarely be integer values and it is assumed there will have to be interpolation from the Chi-squared tables or rounding applied. In our approximations, there is no interpolation. Excel always looks at the whole number for υ and u when choosing a value e.g υ = 2.3, 2.4 & 2.9 will always return a chi-squared value relating to 2. Any instance where υ and u are less than unity are rounded upwards to allow an approximation to be obtained.

It is also the case that the degrees of freedom represented in the Chi squared tables by υ and u are low for the sound insulation test data. This leads to relatively wide confidence intervals based on this approximation. To illustrate this an example is calculated below using data from the concrete GRR.

Table 11-3: Section of Excel Calculation Sheet : confidence limit approximation for Reproducibility - Concrete Floor 500Hz band

Full Model (With Interaction) 95% Confidence Limits Concrete Floor (500Hz)

n	o	p			
3	5	6		a =0.95	
Mso	MSR	MSOP	o-1 df	Prob ((1-a)/2)	
6.665	1.293	2.202	1.70	0.025	0.975

$$\frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{1-\alpha/2,\nu}}$$

where

$$\nu = (\hat{\sigma}^2_{\text{reproducibility}})^2 \left\{ \frac{(1/np)^2 \text{MS}^2_O}{o-1} + \frac{[(p-1)/np]^2 \text{MS}^2_{OP}}{(o-1)(p-1)} + \frac{(1/n)^2 \text{MS}^2_R}{op(n-1)} \right\}^{-1} \quad (8)$$

Lower	5.023886	$\chi^2_{\alpha/2,\nu}$
Upper	0.000982	$\chi^2_{1-\alpha/2,u}$

	17.83	1/np2	MSO^2	Calc
Reprod =	0.31	0.003	44.422	0.034
	0.10			
v =	1.70	((p-1)/np)^2	MSOP^2	Calc
Integer v =	2	0.077	4.849	0.019
Op Var=	0.25	(1/n)^2	MSR^2	Calc
OPxP Var=	0.06	0.111	1.672	0.003

$$\frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{1-\alpha/2,\nu}}$$

Confidence Limits (dB)

v*repro/chi		repro		v*repro/chi
0.123	<	0.309	<	628.356

Using: equation 11-5 above to calculate υ a value of 1.7 is calculated which represents the degrees of freedom that must be used to look up the upper and lower chi squared values (0.000982 and 5.023886 respectively from the table). Note Microsoft Excel returns a value based on υ = 1 as it rounds down any non integer values i.e. there is no interpolation between integers. In order to approximate to the closest limit the decision was taken to round to the nearest integer value which in this case is 2. The confidence upper and lower limit can then be calculated using the value for reproducibility which is 0.309dB: e.g. the lower limit comes from:

(0.309 x 2) / 5.023886 = 0.123dB

The upper limit comes from the following calculation:

(0.309 x 2) / 0.000982 = 628.356dB!

NB: 3 decimal places for sound insulation measurement is not appropriate but has been retained in the example for demonstration purposes.

This example demonstrates how the low degrees of freedom affects the confidence limits for some of the reproducibility in this case. The results are affected in this case by the low number of operators in the DOE i.e. o = 5. Doubling the number of operators to 10 has a significant effect on reducing the size of the confidence interval see below:

Table 11-4: Section of Excel Calculation Sheet Satterthwaite confidence limit approximation for Reproducibility - Concrete Floor 500Hz band – Hypothetically increasing the number of operators to 10

Full Model (With Interaction) 95% Confidence Limits Concrete Floor (500Hz)

n	o	p			
3	10	6		a =0.95	
Mso	MSR	MSOP	o-1 df	Prob ((1-a)/2)	
6.665	1.293	2.202	3.08	0.025	0.975

$$\frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{\alpha/2,\,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{1-\alpha/2,\,\nu}}$$

where

$$\nu = (\hat{\sigma}^2_{\text{reproducibility}})^2\left[\frac{(1/np)^2 MS_O^2}{o-1} + \frac{[(p-1)/np]^2 MS_{OP}^2}{(o-1)(p-1)} + \frac{(1/n)^2 MS_R^2}{op(n-1)}\right]^{-1} \quad (8)$$

Lower	9.348404	$\chi^2_{\alpha/2,\,\nu}$
Upper	0.215795	$\chi^2_{1-\alpha/2,\,u}$

$$\frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{\alpha/2,\,\nu}} \leqslant \sigma^2_{\text{reproducibility}} \leqslant \frac{\nu\hat{\sigma}^2_{\text{reproducibility}}}{\chi^2_{1-\alpha/2,\,\nu}}$$

	39.85	1/np2	MSO^2	Calc
Reprod =	0.28	0.003	44.422	0.015
	0.08			
v =	3.08	((p-1)/np)^2	MSOP^2	Calc
Integer v =	3	0.077	4.849	0.008
Op Var=	0.25	(1/n)^2	MSR^2	Calc
OPxP Var=	0.03	0.111	1.672	0.002

Confidence Limits (dB)

v*repro/chi		repro		v*repro/chi
0.089	<	0.278	<	3.868

In this study the availability of resources restricted the number of operators to 5. In reality, because of the significant effort required in sound insulation testing and limitations to the allowable time on site it would not have been possible to extend the DOE to include more. Ten operators in this case would have required a minimum of 6 days on site.

11.3.1.3 Full and Reduced Model Fitting

The confidence intervals were calculated using the Satterthwaite approximation for $\sigma^2_{repeatability}$, $\sigma^2_{reproducibility}$ and σ^2_{gauge} for both concrete and timber ANOVAs. In each case the most appropriate model is selected, based on either a full model (with interaction) or the reduced model (without interaction), using the AIAG criteria [102], to remove interaction term from the model (α>0.25).

If $MS_{op} < MS_R$, this leads to a negative estimate of reproducibility. Then it is assumed $\hat{\sigma}^2_{reproducibility}$ = 0, unless the reduced model has been fitted and the interaction term is removed. Fitting the reduced model, where the interaction term is not significant, will lead to increase in the precision of the confidence interval, because the reduced model benefits from an increase in the degrees of freedom.

Based on these assumptions the calculated confidence intervals are detailed in Table 11-5 to Table 11-10:

11.3.1.4 Timber Floors:

Table 11-5: Repeatability Variance – Timber Floor GRR

Repeatability - No Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.24	1.08	2.81	#VALUE!	0.69	1.16	0.59	0.32	0.24	0.24	0.26	0.27	0.64	0.58	1.19	1.00	#VALUE!	#VALUE!
Estimate	0.32	1.44	3.74	-	0.92	1.55	0.79	0.43	0.31	0.32	0.34	0.36	0.86	0.77	1.58	1.33	-	-
Higher	0.44	2.02	5.24	#VALUE!	1.29	2.17	1.10	0.61	0.44	0.44	0.48	0.50	1.20	1.08	2.22	1.86	#VALUE!	#VALUE!
Repeatability - With Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.26	1.05	2.64	1.46	0.77	1.14	0.66	0.36	0.22	0.24	0.24	0.28	0.63	0.62	1.08	0.92	0.78	1.80
Estimate	0.37	1.46	3.67	2.03	1.07	1.59	0.91	0.49	0.30	0.33	0.33	0.39	0.87	0.85	1.51	1.27	1.08	2.49
Higher	0.54	2.16	5.44	3.01	1.59	2.36	1.35	0.73	0.45	0.49	0.49	0.57	1.29	1.27	2.23	1.89	1.60	3.69
Repeatability	α>0.25	α>0.25	α>0.25		α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.24	1.08	2.81	1.46	0.69	1.16	0.59	0.32	0.24	0.24	0.26	0.27	0.64	0.58	1.19	1.00	0.78	1.80
Estimate	0.32	1.44	3.74	2.03	0.92	1.55	0.79	0.43	0.31	0.32	0.34	0.36	0.86	0.77	1.58	1.33	1.08	2.49
Higher	0.44	2.02	5.24	3.01	1.29	2.17	1.10	0.61	0.44	0.44	0.48	0.50	1.20	1.08	2.22	1.86	1.60	3.69
Negative Error Value	0.08	0.36	0.93	0.57	0.23	0.39	0.20	0.11	0.08	0.08	0.09	0.09	0.21	0.19	0.40	0.33	0.30	0.70
Positive Error Value	0.13	0.58	1.50	0.98	0.37	0.62	0.31	0.17	0.13	0.13	0.14	0.14	0.34	0.31	0.63	0.53	0.52	1.20
% of Repeatability	65%	65%	65%	76%	65%	65%	65%	65%	65%	65%	65%	65%	65%	65%	65%	65%	76%	76%

Table 11-6: Reproducibility Variance – Timber Floor GRR

Reproducibility - No Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.16	1.50	2.82	#VALUE!	0.00	0.03	0.24	0.17	0.03	0.04	0.06	0.00	0.02	0.47	1.64	0.88	#VALUE!	#VALUE!
Estimate	0.49	4.43	8.50	#VALUE!	0.09	0.36	0.79	0.56	0.14	0.18	0.23	0.06	0.20	1.45	4.84	2.69	#VALUE!	#VALUE!
Higher	4.21	37.19	71.77	#VALUE!	1.14	3.59	6.88	4.81	1.26	1.59	2.04	0.64	2.02	12.33	40.63	22.74	#VALUE!	#VALUE!
Reproducibility - With Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.21	1.90	3.64	0.94	0.02	0.10	0.25	0.24	0.06	0.06	0.07	0.01	0.05	0.62	2.07	1.15	1.13	2.89
Estimate	0.49	4.43	8.51	2.19	0.08	0.35	0.78	0.55	0.14	0.18	0.23	0.06	0.20	1.45	4.85	2.69	2.64	6.75
Higher	9.05	82.11	157.73	40.57	77.54	13.90	10.85	10.27	8.14	2.44	3.21	57.29	7.89	26.80	89.85	49.91	48.89	125.16
Reproducibility	α>0.25	α>0.25	α>0.25		α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.16	1.50	2.82	0.94	0.00	0.03	0.24	0.17	0.03	0.04	0.06	0.00	0.02	0.47	1.64	0.88	1.13	2.89
Estimate	0.49	4.43	8.50	2.19	0.09	0.36	0.79	0.56	0.14	0.18	0.23	0.06	0.20	1.45	4.84	2.69	2.64	6.75
Higher	4.21	37.19	71.77	40.57	1.14	3.59	6.88	4.81	1.26	1.59	2.04	0.64	2.02	12.33	40.63	22.74	48.89	125.16
Negative Error Value	0.34	2.93	5.68	1.25	0.09	0.32	0.56	0.39	0.11	0.13	0.17	0.06	0.18	0.98	3.20	1.81	1.51	3.86
Positive Error Value	3.71	32.76	63.27	38.38	1.05	3.23	6.09	4.25	1.12	1.41	1.81	0.58	1.82	10.88	35.79	20.06	46.25	118.40
% of Reproducibility	821%	805%	811%	1811%	1248%	1000%	838%	827%	900%	876%	861%	1080%	994%	815%	806%	814%	1811%	1811%

Table 11-7: Gauge Variance – Timber Floor GRR

Gauge - No Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.42	2.74	5.99	#VALUE!	0.75	1.33	0.88	0.51	0.29	0.30	0.34	0.30	0.74	1.11	3.01	1.92	#VALUE!	#VALUE!
Estimate	0.81	5.87	12.24	#VALUE!	1.01	1.91	1.58	0.99	0.45	0.49	0.57	0.42	1.06	2.22	6.42	4.02	#VALUE!	#VALUE!
Higher	2.97	32.03	56.81	#VALUE!	1.46	2.98	4.33	2.93	0.85	0.98	1.25	0.63	1.67	8.90	35.13	15.43	#VALUE!	#VALUE!
Gauge - With Interaction																		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.41	2.74	6.01	2.58	0.73	1.34	0.84	0.50	0.29	0.31	0.34	0.30	0.73	1.09	3.03	1.93	2.00	4.86
Estimate	0.80	5.87	12.26	4.61	0.99	1.90	1.56	0.98	0.45	0.49	0.57	0.41	1.05	2.21	6.44	4.03	4.11	9.75
Higher	2.88	31.99	56.97	11.32	1.44	3.03	4.16	2.84	0.85	1.03	1.26	0.63	1.65	8.74	35.34	15.53	14.05	46.02
Gauge	α>0.25	α>0.25	α>0.25		α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25	α>0.25		
Is $MS_{OP}<MS_R$?	YES!	YES!	No	No	YES!	YES!	YES!	YES!	No	YES!	No	YES!	YES!	YES!	No	No	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.42	2.74	5.99	2.58	0.75	1.33	0.88	0.51	0.29	0.30	0.34	0.30	0.74	1.11	3.01	1.92	2.00	4.86
Estimate	0.81	5.87	12.24	4.61	1.01	1.91	1.58	0.99	0.45	0.49	0.57	0.42	1.06	2.22	6.42	4.02	4.11	9.75
Higher	2.97	32.03	56.81	11.32	1.46	2.98	4.33	2.93	0.85	0.98	1.25	0.63	1.67	8.90	35.13	15.43	14.05	46.02
Negative Error Value	0.39	3.13	6.25	2.03	0.26	0.58	0.70	0.48	0.16	0.19	0.23	0.12	0.32	1.12	3.42	2.10	2.12	4.89
Positive Error Value	2.16	26.15	44.56	6.71	0.45	1.07	2.75	1.94	0.40	0.48	0.68	0.21	0.61	6.68	28.71	11.42	9.94	36.27
% of Gauge	314%	499%	415%	190%	70%	87%	218%	244%	124%	137%	159%	78%	88%	351%	500%	337%	293%	422%

NB: Where cell returns "#VALUE!" reduced model not fitted, Confidence interval defaults to full model values.

11.3.1.5 Concrete Floors:

Table 11-8: Repeatability Variance – Concrete Floor GRR

Repeatability - No Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	1.76	#VALUE!	#VALUE!	#VALUE!	0.62	#VALUE!	#VALUE!
Estimate	-	-	-	-	-	-	-	-	-	-	-	2.35	-	-	-	0.83	-	-
Higher	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	3.29	#VALUE!	#VALUE!	#VALUE!	1.16	#VALUE!	#VALUE!
Repeatability - With Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.35	0.58	2.10	2.64	1.43	0.98	1.26	0.99	0.82	0.93	1.11	1.64	0.65	0.60	0.57	0.62	0.50	0.44
Estimate	0.49	0.80	2.92	3.67	1.99	1.37	1.76	1.38	1.14	1.29	1.55	2.27	0.90	0.83	0.80	0.86	0.70	0.60
Higher	0.72	1.19	4.32	5.43	2.95	2.03	2.60	2.04	1.69	1.92	2.29	3.37	1.33	1.23	1.18	1.27	1.03	0.90
Repeatability - Fitted Model												α>0.25						
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.35	0.58	2.10	2.64	1.43	0.98	1.26	0.99	0.82	0.93	1.11	1.76	0.65	0.60	0.57	0.62	0.50	0.44
Estimate	0.49	0.80	2.92	3.67	1.99	1.37	1.76	1.38	1.14	1.29	1.55	2.35	0.90	0.83	0.80	0.86	0.70	0.60
Higher	0.72	1.19	4.32	5.43	2.95	2.03	2.60	2.04	1.69	1.92	2.29	3.29	1.33	1.23	1.18	1.27	1.03	0.90
Negative Error Value	0.14	0.22	0.82	1.03	0.56	0.38	0.49	0.39	0.32	0.36	0.43	0.59	0.25	0.23	0.22	0.24	0.19	0.17
Positive Error Value	0.24	0.39	1.41	1.77	0.96	0.66	0.85	0.66	0.55	0.62	0.75	0.94	0.43	0.40	0.38	0.41	0.34	0.29
% of Repeatability	76%	76%	76%	76%	76%	76%	76%	76%	76%	76%	76%	65%	76%	76%	76%	76%	76%	76%

Table 11-9: Reproducibility Variance – Concrete Floor GRR

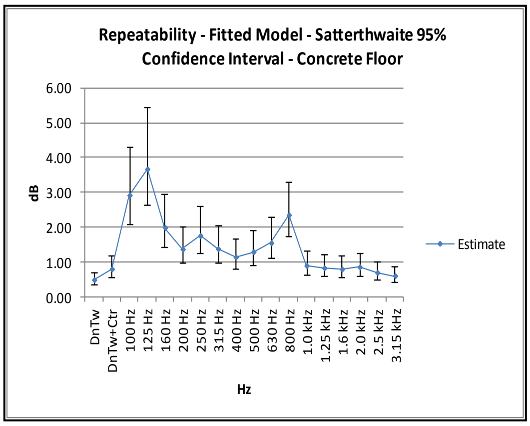
Reproducibility - No Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	0.07	#VALUE!	#VALUE!	#VALUE!	0.05	#VALUE!	#VALUE!
Estimate	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	0.56	#VALUE!	#VALUE!	#VALUE!	0.27	#VALUE!	#VALUE!
Higher	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	5.60	#VALUE!	#VALUE!	#VALUE!	2.55	#VALUE!	#VALUE!
Reproducibility - With Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.10	0.21	0.74	0.89	0.11	0.44	0.05	0.04	0.11	0.12	0.26	0.15	0.30	0.28	0.08	0.07	0.03	0.02
Estimate	0.31	0.66	2.32	2.20	0.56	1.08	0.24	0.18	0.42	0.31	0.65	0.57	0.94	0.86	0.29	0.26	0.14	0.11
Higher	4.27	9.13	32.20	130.30	568.44	64.07	246.55	186.16	16.48	628.36	38.29	22.42	13.11	11.93	11.37	10.42	145.61	113.02
Reproducibility - Fitted Model												α>0.25						
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.10	0.21	0.74	0.89	0.11	0.44	0.05	0.04	0.11	0.12	0.26	0.07	0.30	0.28	0.08	0.07	0.03	0.02
Estimate	0.31	0.66	2.32	2.20	0.56	1.08	0.24	0.18	0.42	0.31	0.65	0.56	0.94	0.86	0.29	0.26	0.14	0.11
Higher	4.27	9.13	32.20	130.30	568.44	64.07	246.55	186.16	16.48	628.36	38.29	5.60	13.11	11.93	11.37	10.42	145.61	113.02
Negative Error Value	0.21	0.45	1.57	1.30	0.45	0.64	0.19	0.15	0.30	0.19	0.38	0.49	0.64	0.58	0.21	0.19	0.11	0.09
Positive Error Value	3.97	8.47	29.88	128.10	567.88	62.99	246.31	185.98	16.07	628.05	37.65	5.04	12.16	11.07	11.08	10.15	145.47	112.90
% of Reproducibility	1358%	1358%	1358%	5884%	101806%	5884%	101806%	101806%	3923%	203612%	5884%	989%	1358%	1358%	3923%	3923%	101806%	101806%

Table 11-10: Gauge Variance – Concrete Floor GRR

Gauge - No Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	2.05	#VALUE!	#VALUE!	#VALUE!	0.73	#VALUE!	#VALUE!
Estimate	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	2.91	#VALUE!	#VALUE!	#VALUE!	1.09	#VALUE!	#VALUE!
Higher	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	4.68	#VALUE!	#VALUE!	#VALUE!	1.84	#VALUE!	#VALUE!
Gauge - With Interaction																		
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.55	1.08	3.50	4.38	2.55	2.21	2.02	1.24	1.21	1.30	1.66	2.05	1.30	1.15	0.86	0.72	0.70	0.56
Estimate	0.88	1.79	5.76	6.94	3.63	3.44	2.79	1.68	1.78	1.84	2.49	2.92	2.18	2.00	1.26	1.09	0.96	0.77
Higher	1.75	3.53	12.90	12.72	5.66	6.56	4.22	2.47	2.98	2.82	4.29	4.68	5.17	4.42	2.08	1.82	1.48	1.17
Gauge - Fitted Model												α>0.25						
Is MS_{OP}<MS_{R}?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	YES!	No	No
dB	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Lower	0.55	1.08	3.50	4.38	2.55	2.21	2.02	1.24	1.21	1.30	1.66	2.05	1.30	1.15	0.86	0.73	0.70	0.56
Estimate	0.88	1.79	5.76	6.94	3.63	3.44	2.79	1.68	1.78	1.84	2.49	2.91	2.18	2.00	1.26	1.09	0.96	0.77
Higher	1.75	3.53	12.90	12.72	5.66	6.56	4.22	2.47	2.98	2.82	4.29	4.68	5.17	4.42	2.08	1.84	1.48	1.17
Negative Error Value	0.33	0.71	2.26	2.56	1.09	1.22	0.76	0.45	0.57	0.54	0.83	0.86	0.88	0.84	0.40	0.36	0.27	0.21
Positive Error Value	0.87	1.74	7.14	5.77	2.03	3.12	1.43	0.79	1.19	0.98	1.81	1.77	3.00	2.42	0.82	0.75	0.52	0.40
% of Gauge	136%	137%	163%	120%	86%	126%	79%	73%	99%	83%	106%	90%	178%	164%	97%	102%	81%	79%

NB: Where cell returns "#VALUE!" reduced model not fitted, Confidence interval defaults to full model values.

The confidence intervals for $\sigma^2_{repeatability}$ $\sigma^2_{reproducibility}$ and σ^2_{gauge} are shown in figures Figure 11-3 to Figure 11-5 for both timber and concrete GRR.



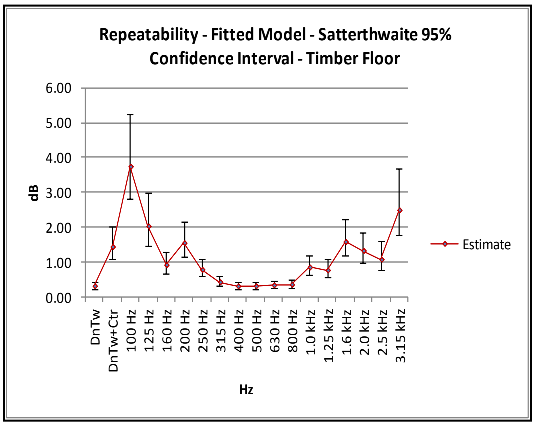
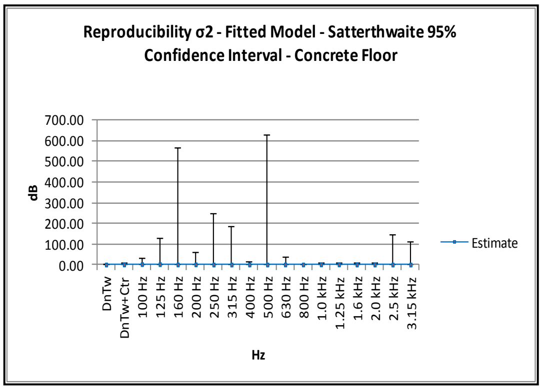


Figure 11-3: Satterthwaite Repeatability Confidence Limits Concrete & Timber GRR



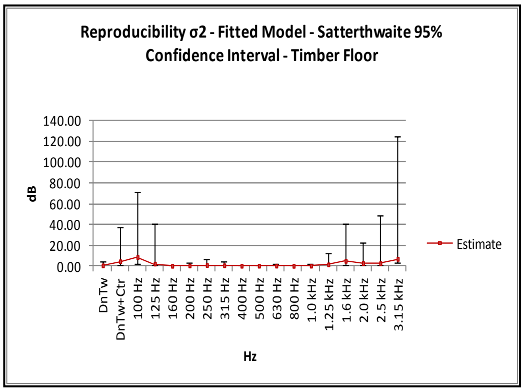
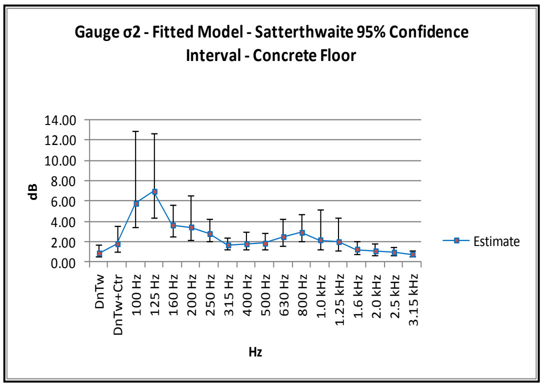


Figure 11-4: Satterthwaite Reproducibility Confidence Limits Concrete & Timber GRR



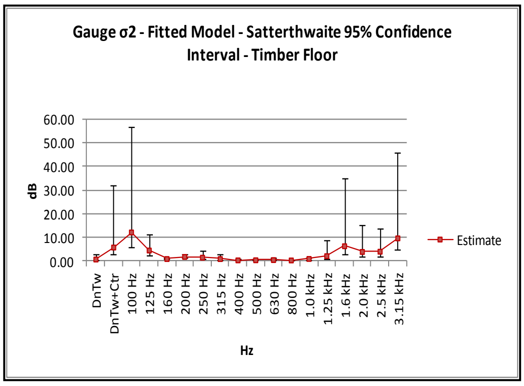


Figure 11-5: Satterthwaite Gauge Confidence Limits Concrete & Timber GRR

The confidence intervals are summarised in Table 11-11 and Table 11-12.

Table 11-11: Concrete GRR Confidence Interval Range - dB

dB/95%	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Concrete Repeatability Confidence Range	0.37	0.61	2.22	2.79	1.52	1.04	1.34	1.05	0.87	0.99	1.18	1.53	0.68	0.63	0.61	0.65	0.53	0.46
Concrete Reproducibility Confidence Range	4.17	8.92	31.45	129.40	568.33	63.63	246.50	186.13	16.37	628.23	38.03	5.54	12.80	11.65	11.29	10.35	145.58	112.99
Concrete Gauge Confidence Range	1.20	2.45	9.40	8.34	3.12	4.35	2.19	1.23	1.76	1.52	2.63	2.63	3.88	3.27	1.22	1.11	0.78	0.61

Table 11-12: Timber GRR Confidence Interval Range - dB

dB/95%	DnTw	DnTw+Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	1.25 kHz	1.6 kHz	2.0 kHz	2.5 kHz	3.15 kHz
Repeatability Confidence Range	0.21	0.94	2.43	1.55	0.60	1.01	0.51	0.28	0.20	0.21	0.22	0.23	0.56	0.50	1.03	0.86	0.82	1.90
Reproducibility Confidence Range	4.05	35.69	68.95	39.64	1.14	3.56	6.64	4.63	1.23	1.55	1.98	0.64	2.00	11.86	38.99	21.86	47.76	122.27
Gauge Confidence Range	2.55	29.28	50.81	8.74	0.71	1.65	3.45	2.42	0.56	0.68	0.91	0.33	0.93	7.80	32.12	13.52	12.06	41.16

For the timber GRR, interaction between operators and parts is not significant at $\alpha > 0.25$ for the majority of frequency data and the interaction term is removed. The reduced model is generally applied with the full model only applied to the 125Hz, 2500Hz and 3150Hz bands.

The Concrete GRR data exhibits significant operator by part interaction and the full model is generally applied to all but the 800Hz frequency band.

The Satterthwaite approximations to the confidence intervals for $\sigma^2_{repeatability}$ are generally larger for the concrete GRR data than for timber. They also show that for timber floor data the confidence intervals are relatively precise at 65% of $\hat{\sigma}^2_{repeatability}$. These are narrower intervals than the $\hat{\sigma}^2_{repeatability}$ for the concrete floor where the confidence interval is at 76% of $\hat{\sigma}^2_{repeatability}$. The precision of the confidence interval is determined by subtracting the lower limit of repeatability from the upper limit of repeatability and dividing the result by the estimate for repeatability. For example using data for $D_{nT,w} + C_{tr}$ for the timber floor GRR.

If the lower estimate of the confidence interval is 1.05dB the upper estimate is 2.16dB and the estimate for repeatability is 1.46dB then (all expressed as standard deviations):

(2.16 – 1.05)/1.46 = 0.76; expressed as a percentage it is, 76%

The improved precision in the timber model is due to the reduced model being fitted to the timber floor data in the absence of significant interaction (see Figure 11-6).

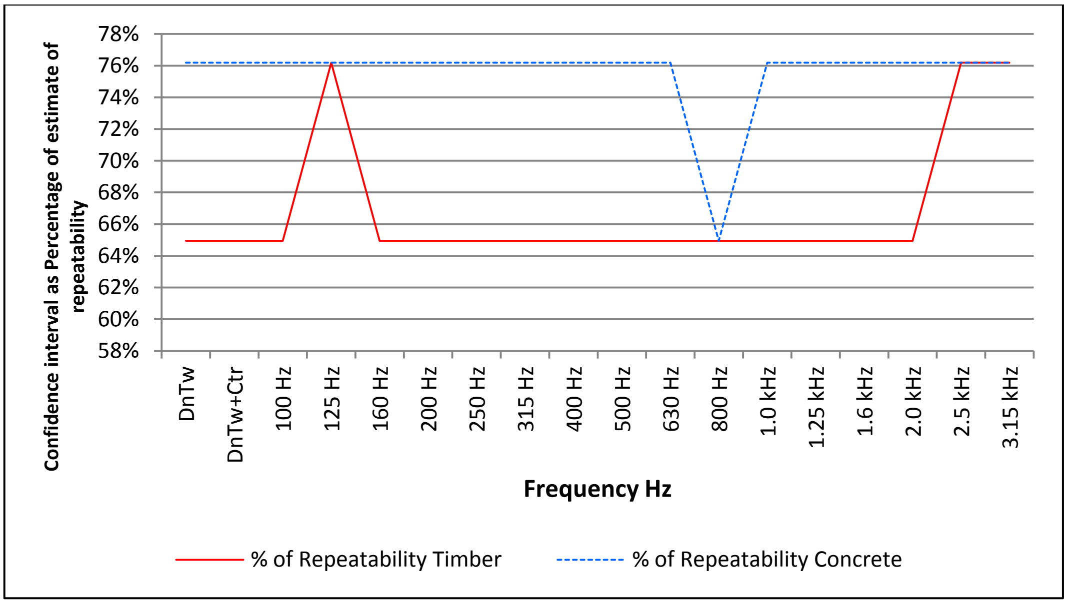


Figure 11-6: Satterthwaite Confidence Interval Precision - $\boldsymbol{\sigma^2_{repeatability}}$

With respect to the 95% confidence interval, for $\sigma^2_{repeatability}$ it is 0.37dB and 0.61dB for $D_{nT,w}$ and $D_{nT,w}$ + C_{tr}, respectively, for concrete GRR it is 0.21dB and 0.94dB for the timber GRR. It varies across the frequency range generally falling as frequency increases (see Figure 11-7).

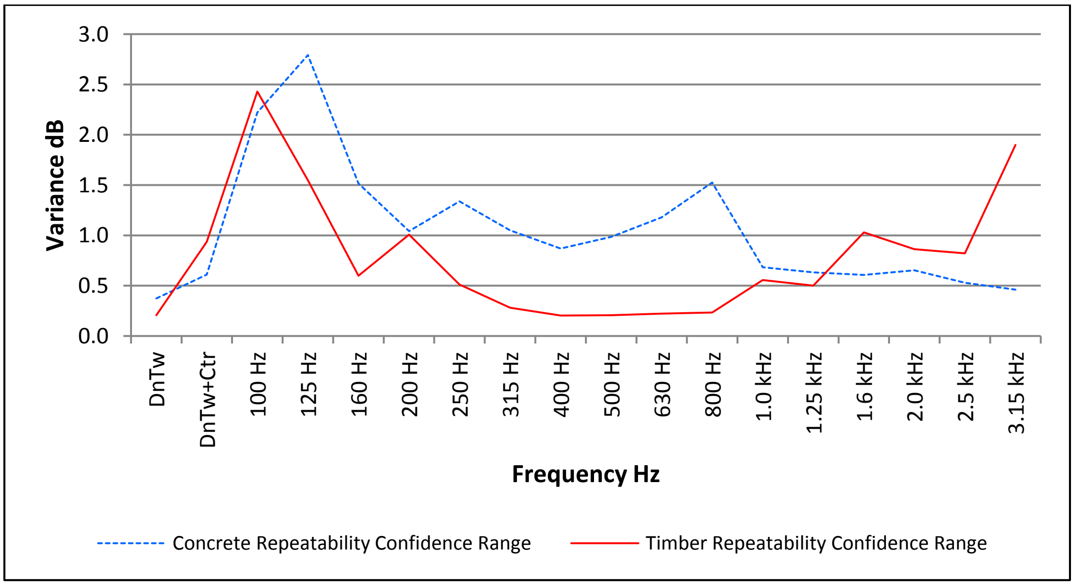


Figure 11-7: 95% Confidence Interval Size for variance repeatability component - Concrete & Timber GRR $\boldsymbol{\sigma^2_{repeatability}}$

For $\sigma^2_{reproducibility}$, the approximations give wide confidence limits and in some frequency bands, significant asymmetry for both timber and concrete GRR. The concrete confidence intervals are largest in the 160Hz and 500Hz bands, where the upper limit dictates the precision and the calculated levels are 101806% and 203612% of $\hat{\sigma}^2_{reproducibility}$. The calculated values for both 160Hz and 500Hz bands are 1.2367 and 1.697 respectively, both returning the chi squared value for 1 but with υ = 1.697 as multiplier, rounding up to 2. This is an example of how rounding can cause significant variability in the confidence interval precision. Its impact is especially acute at low values of υ.

The confidence interval approximations for $\sigma^2_{reproducibility}$ are relatively wide for both GRR studies. The minimum confidence interval for timber floors is 805% of $\hat{\sigma}^2_{reproducibility}$, and 992% for concrete floors (reduced model fitted). This is due to υ being small. The timber and concrete GRR data regularly result in υ less than (o – 1) and the low degrees result in reduced precision for this component of variance. See Figure 11-8.

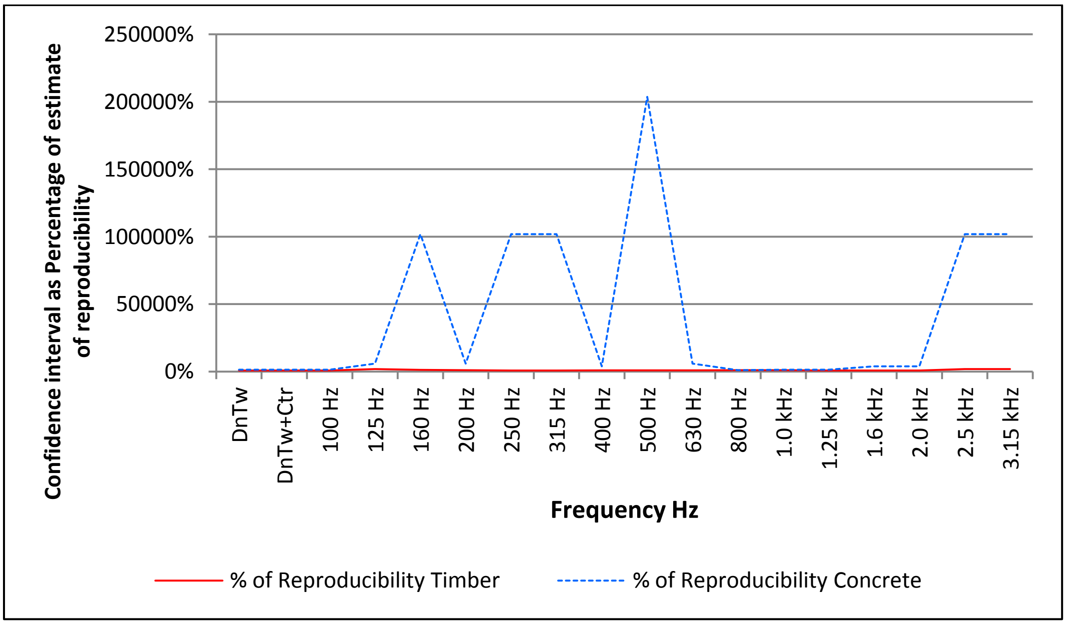


Figure 11-8: Satterthwaite Confidence Interval Precision - $\boldsymbol{\sigma^2_{reproducibility}}$

With respect to the 95% confidence interval for $\sigma^2_{reproducibility}$, it is 4.2dB and 8.9dB for $D_{nT,w}$ and $D_{nT,w}$ + *C*tr, respectively, for concrete GRR and 4.0dB and 35.7dB for timber.

The ranges for both timber and concrete data sets are both affected by low degrees of freedom. This is particularly true for the concrete data, across the third octave range, as in the majority of cases, the model includes interaction. The size of the confidence interval for timber is relatively small between 160Hz – 1000Hz, being below 2dB for 160Hz and 400 – 1000Hz and below 6dB for 200-315Hz. This is shown in Figure 11-9 and separately for the timber GRR in Figure 11-10.

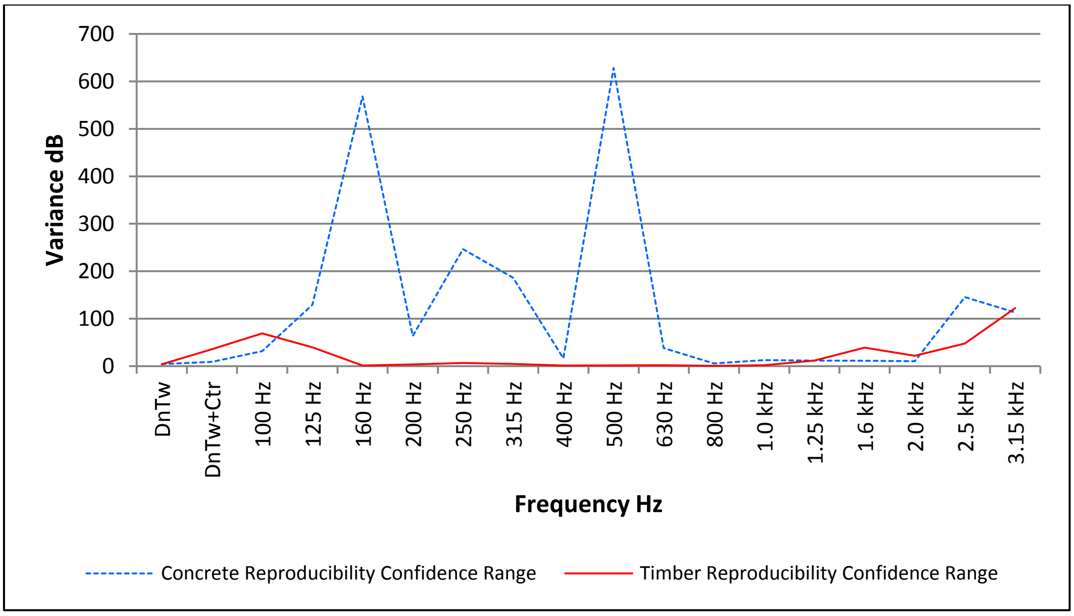


Figure 11-9: 95% Confidence Interval for variance reproducibility component - Concrete & Timber GRR $\boldsymbol{\sigma^2_{reproducibility}}$

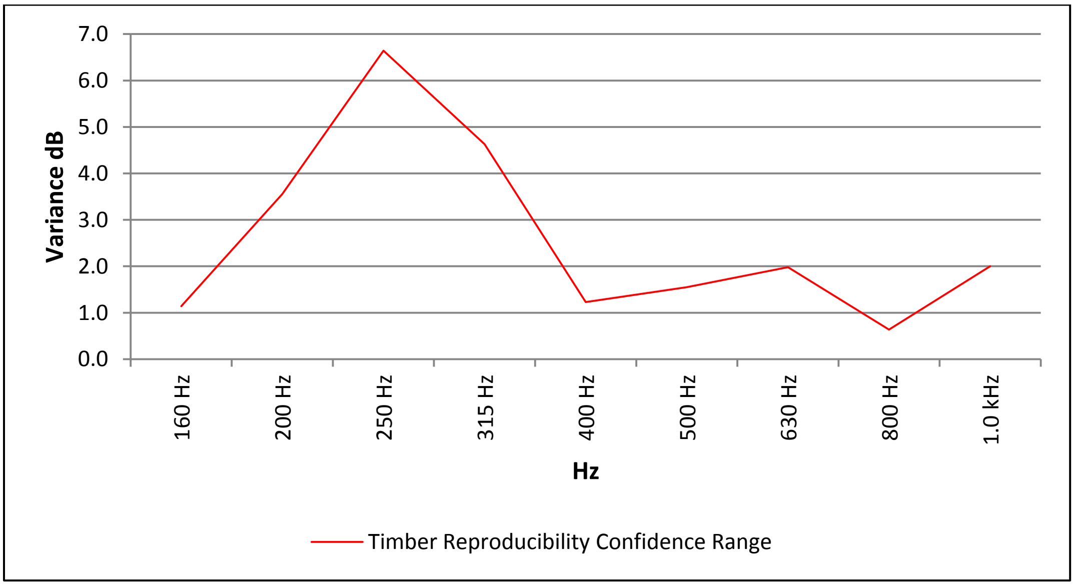


Figure 11-10: 95% Confidence Interval -Timber GRR – 160-1000Hz $\boldsymbol{\sigma^2_{reproducibility}}$

The Satterthwaite approximation for σ^2_{gauge} confidence intervals has greater precision than for $\sigma^2_{repeatability}$ or $\sigma^2_{reproducibility}$. This is due to u being relatively large. This is because the point estimate for $\hat{\sigma}^2_{gauge}$ is the summation of both $\hat{\sigma}^2_{repeatability}$ and $\hat{\sigma}^2_{reproducibility}$ and consequently is generally greater than unity. This results in the square term $\left(\hat{\sigma}^2_{gauge}\right)^2$, increasing in size.

The precision of the confidence intervals for σ^2_{gauge} concrete GRR data ranges from 73% to 178% and the timber GRR data range from 70% to 500%. The areas where the confidence intervals precision is lower (i.e. 500%) generally follow where the point estimate for $\hat{\sigma}^2_{gauge}$ is higher. This is illustrated in Figure 11-11.

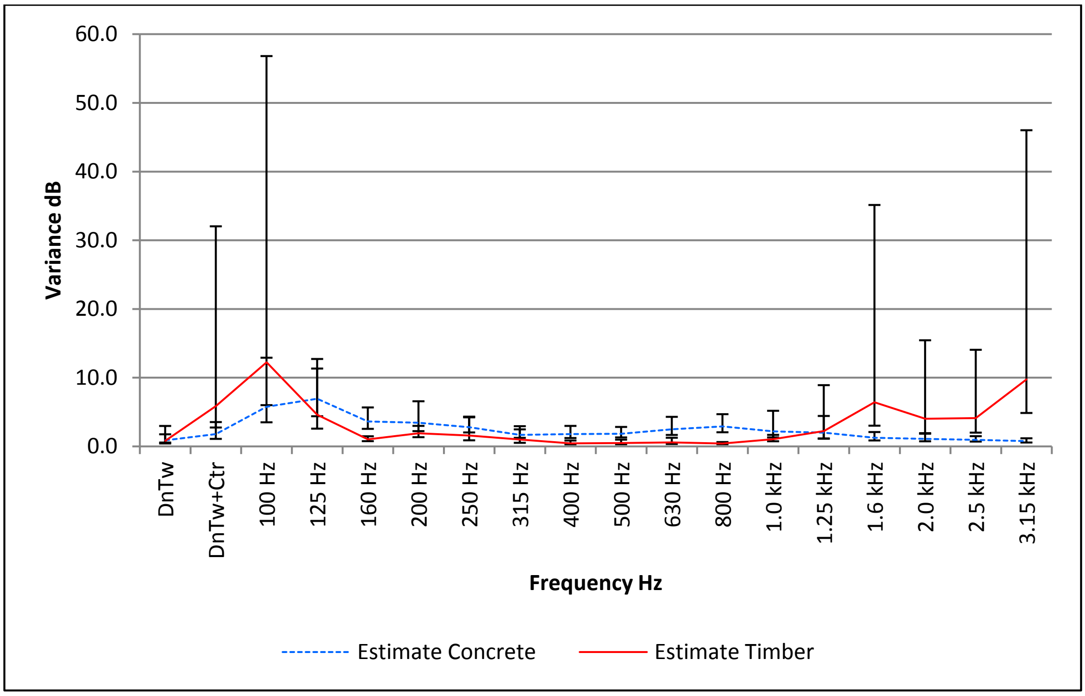


Figure 11-11: Point Estimates for Timber & Concrete GRR $\hat{\boldsymbol{\sigma}}^2_{gauge}$ *showing 95% confidence intervals*

With respect to the 95% confidence interval. For σ^2_{gauge} , it is 1.2dB and 2.5dB for $D_{nT,w}$ and $D_{nT,w}$ + C_{tr} respectively for concrete GRR and 2.6dB and 29.3dB for the timber GRR. The ranges for both timber and concrete data benefit from both the relatively high degrees of freedom as u and w are relatively high. The larger range for the timber GRR $D_{nT,w}$ + C_{tr} value is due to a large value for $\hat{\sigma}^2_{gauge}$ = 5.9dB. The ranges are shown in Figure 11-12.

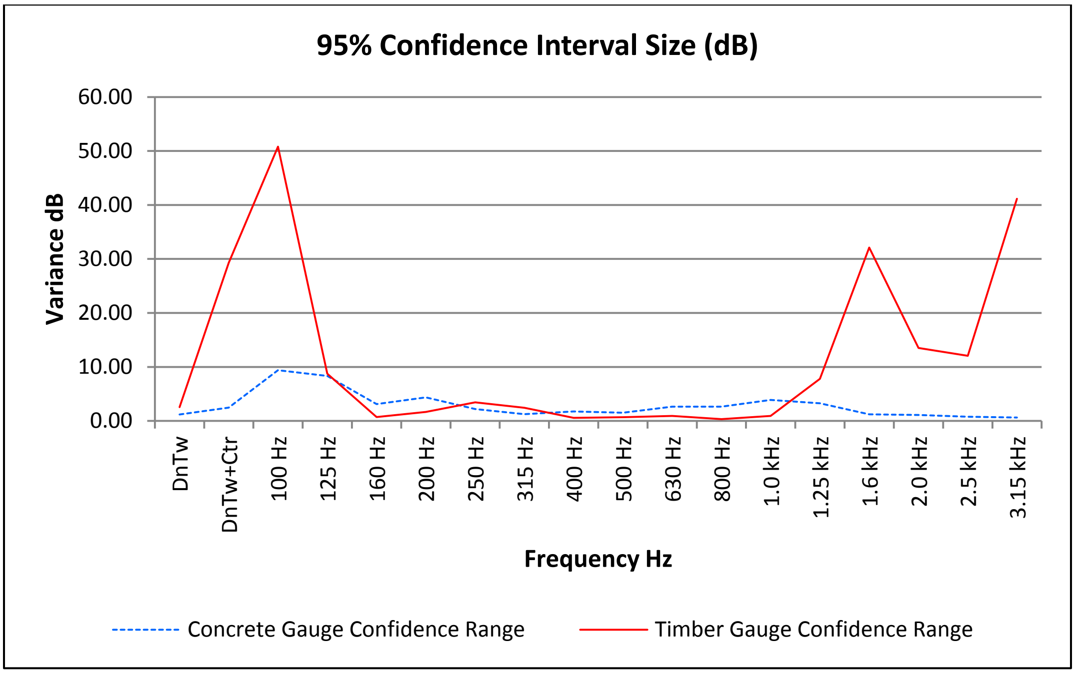


Figure 11-12: 95% Confidence Interval size for variance component "Gauge" - Concrete & Timber GRR σ^2_{gauge}

The Timber GRR has significantly higher point estimates of $\hat{\sigma}^2_{gauge}$ than concrete GRR at 1600 – 3150Hz. This is reflected in lower precision over this range; 293 - 500% as opposed to 79 – 102% for concrete; also greater intervals: 12.1 – 41.2dB as opposed to 0.6 – 1.2dB for concrete. See Figure 11-13.

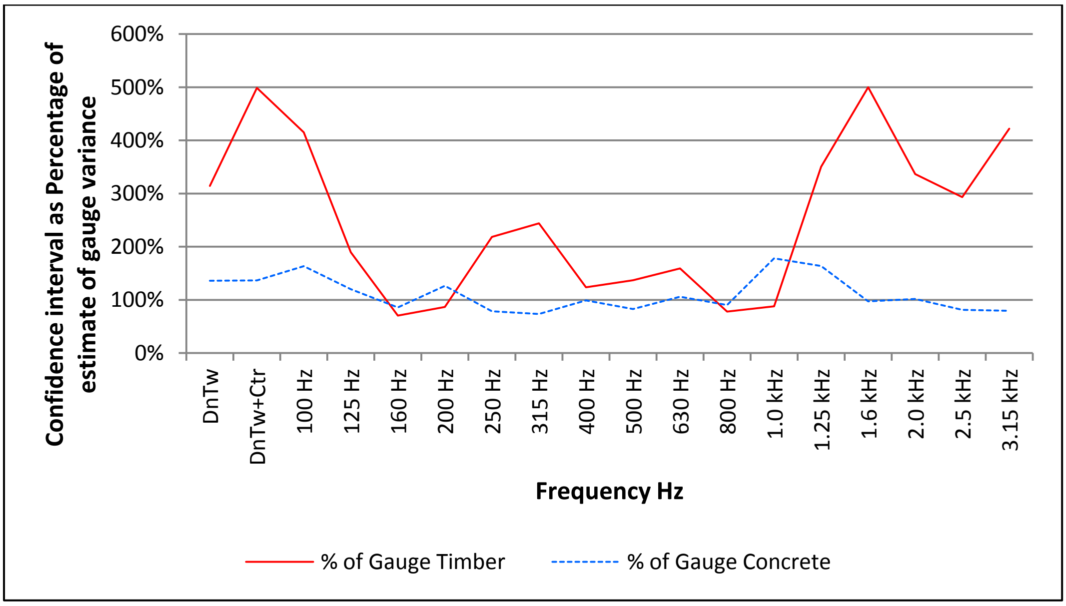


Figure 11-13: Satterthwaite Confidence Interval Precision (%) - σ^2_{gauge}

In summary, the Satterthwaite method of calculating confidence intervals for components of variance, determined by ANOVA, delivers relatively consistent confidence intervals for $\sigma^2_{repeatability}$ and σ^2_{gauge} with reasonable precision. For $\sigma^2_{reproducibility}$ the intervals vary, depending on the magnitude of the point estimate $\hat{\sigma}^2_{reproducibility}$ and the degrees of freedom, based on the number of operators (o – 1) for the reduced model, and v for the full model. In the GRR, the values for v are relatively low and less than (o – 1) in most cases. This reduces the precision of the confidence interval, which can be highly asymmetrical with the majority of the confidence interval range contained in the upper interval.

An additional factor relates to the $\sigma^2_{reproducibility}$ and σ^2_{gauge} confidence intervals, which are both approximations. It is unlikely v and u will be integer values and so the confidence interval will be affected by rounding and interpolation. Recent research into confidence intervals for ANOVA components has shown that there are alternative strategies that avoid the problem of interpolation [134-136].

These are useful, because for the majority of cases, exact confidence intervals for GRR components are not available and they have to be constructed using a different approach. One such approach is proposed by Borror et al (1997) [71] and is called the Modified Large Sample (MLS) method.

11.3.2 Modified Large Sample Method

The term "modified large sample" (MLS) was introduced by Graybill et al [137]. The method is based on the relationship between the expected mean squares obtained from the ANOVA method and the corresponding variance components. It can yield exact confidence intervals in some circumstances and very close to exact intervals in the majority of cases. Empirical evidence suggests that MLS intervals generally produce confidence coefficients at least as great as the stated level [71, 138] although it may be slightly wider than by other methods.

The resulting confidence intervals are functions of the expected mean squares. It is favoured in this study, compared with simulation [139, 140] when calculating confidence intervals. This is because once the elements are defined the calculation, although relatively cumbersome, can be carried out in a spreadsheet, unlike other methods.

Using a balanced two-factor crossed random model with interaction, see equation11-10:

$$Y_{ijk} = \mu_Y + P_i + O_j + (PO)_{ij} + E_{ijk} \qquad \textit{11-10}$$

Where: μ_Y is a constant and $P_i, O_j, (PO)_{ij}, E_{ijk}$ are jointly independent normal random variables with means of zero and variances $\sigma_P^2, \sigma_O^2, \sigma_{PO}^2, and\ \sigma_E^2$ respectively.

The ANOVA for the above model in equation 11-10 is described in Appendix 14.5.

The parameters to be considered are detailed in Table 11-13.

Table 11-13: Definition of Parameters in a gauge R&R study

Symbol	Definition
μ_Y	Mean of population of measurements
γ_P	Variance of the part
γ_O	Variance of the operator
γ_{PO}	Variance (Interaction) between part & operator
γ_M	Variance of the measurement system

The MLS constants used in confidence intervals for model in equation 11-10 are detailed in Table 14-11 in Appendix 14.6.

The confidence limits can be computed for these parameters for the timber and concrete floors.

11.4 Computed confidence intervals timber floor

Because the timber floor GRR has identical parts it provides confidence intervals which allow an estimation of the variability of the construction as a “part” as well as the other components of variance. Confidence intervals for each can be determined and are provided by Burdick et al [67]. Each can be calculated for the single figure values and third octave band frequency data.

The results are tabulated for the 95% confidence intervals. See Table 11-14 and Table 11-15:

Table 11-14: 95% confidence intervals - Timber Floor: Mean of population of measurements, variance of the part, and variance of the measurement system (gauge): $\boldsymbol{\mu_y, \gamma_P, \gamma_M}$

95%	μ_y			γ_P			γ_M		
dB	Estimate μ	Lower	Upper	Estimate γ_p	Lower	Upper	Estimate γ_m	Lower	Upper
DnTw	62.6	1.1	1.1	0.3	0.1	2.0	0.8	0.5	4.5
DnTw+Ctr	53.2	2.7	2.7	0.2	0.0	1.9	5.9	3.0	38.6
100 Hz	35.3	3.7	3.7	0.3	0.0	3.4	12.3	6.6	75.5
125 Hz	43.6	1.9	1.9	0.0	0.0	0.4	4.6	3.0	21.2
160 Hz	46.0	1.0	1.0	0.7	0.3	4.4	1.0	0.8	2.1
200 Hz	49.6	1.4	1.4	1.2	0.4	7.8	1.9	1.5	5.2
250 Hz	53.3	1.5	1.5	0.9	0.3	5.5	1.6	1.0	7.6
315 Hz	56.5	1.1	1.1	0.4	0.1	2.2	1.0	0.6	5.2
400 Hz	59.4	0.8	0.8	0.4	0.1	2.7	0.5	0.3	1.6
500 Hz	62.1	0.8	0.8	0.3	0.1	1.8	0.5	0.3	1.9
630 Hz	63.8	1.0	1.0	0.5	0.2	3.3	0.6	0.4	2.4
800 Hz	64.9	1.3	1.3	1.3	0.5	8.0	0.4	0.3	1.0
1.0 kHz	65.8	2.1	2.1	3.6	1.4	21.8	1.1	0.8	2.9
1.25 kHz	70.3	2.5	2.5	3.5	1.3	21.1	2.2	1.2	13.1
1.6 kHz	72.7	2.9	2.9	0.8	0.2	5.5	6.4	3.3	42.2
2.0 kHz	69.5	2.3	2.3	1.0	0.3	6.4	4.0	2.2	24.1
2.5 kHz	67.2	2.7	2.7	2.8	1.0	17.7	4.1	2.4	23.6
3.15 kHz	73.4	4.7	4.7	10.1	3.7	61.9	9.7	5.3	59.6

Table 11-15: 95% confidence intervals - Timber Floor $\boldsymbol{\sigma_O^2, \sigma_{po}^2, \sigma_E^2}$

95%	σ_O^2			σ_{PO}^2			σ_E^2		
dB	Estimate σ_O^2	Lower	Upper	Estimate σ_{PO}^2	Lower	Upper	Estimate σ_E^2	Lower	Upper
DnTw	0.50	0.2	4.2	-0.1	-0.1	0.0	0.4	0.3	0.5
DnTw+Ctr	4.43	1.5	37.2	0.0	-0.3	0.5	1.5	1.0	2.2
100 Hz	8.5	2.9	71.7	0.1	-0.7	1.6	3.7	2.6	5.4
125 Hz	2.1	0.6	18.7	0.5	-0.1	1.8	2.0	1.5	3.0
160 Hz	0.1	0.0	1.1	-0.2	-0.4	0.0	1.1	0.8	1.6
200 Hz	0.4	0.1	3.6	-0.1	-0.4	0.5	1.6	1.1	2.4
250 Hz	0.8	0.3	6.9	-0.2	-0.3	0.0	0.9	0.7	1.4
315 Hz	0.6	0.2	4.8	-0.1	-0.2	0.0	0.5	0.4	0.7
400 Hz	0.1	0.0	1.3	0.0	-0.1	0.1	0.3	0.2	0.4
500 Hz	0.2	0.1	1.6	0.0	-0.1	0.1	0.3	0.2	0.5
630 Hz	0.2	0.1	2.0	0.0	-0.1	0.1	0.3	0.2	0.5
800 Hz	0.1	0.0	0.6	0.0	-0.1	0.1	0.4	0.3	0.6
1.0 kHz	0.2	0.0	2.0	0.0	-0.2	0.3	0.9	0.6	1.3
1.25 kHz	1.5	0.5	12.3	-0.1	-0.3	0.1	0.9	0.6	1.3
1.6 kHz	4.8	1.7	40.6	0.1	-0.2	0.8	1.5	1.1	2.2
2.0 kHz	2.7	0.9	22.7	0.1	-0.2	0.6	1.3	0.9	1.9
2.5 kHz	2.5	0.8	22.0	0.5	0.1	1.4	1.1	0.8	1.6
3.15 kHz	6.6	2.2	56.5	0.6	-0.1	2.2	2.5	1.8	3.7

The lower limits are left as calculated in the table although any <0 will be assumed to be zero.

The confidence interval for the measurement of the mean of population of measurements μ_y is symmetrical and the estimates and their confidence limits are plotted in Figure 11-14 - Figure 11-19:

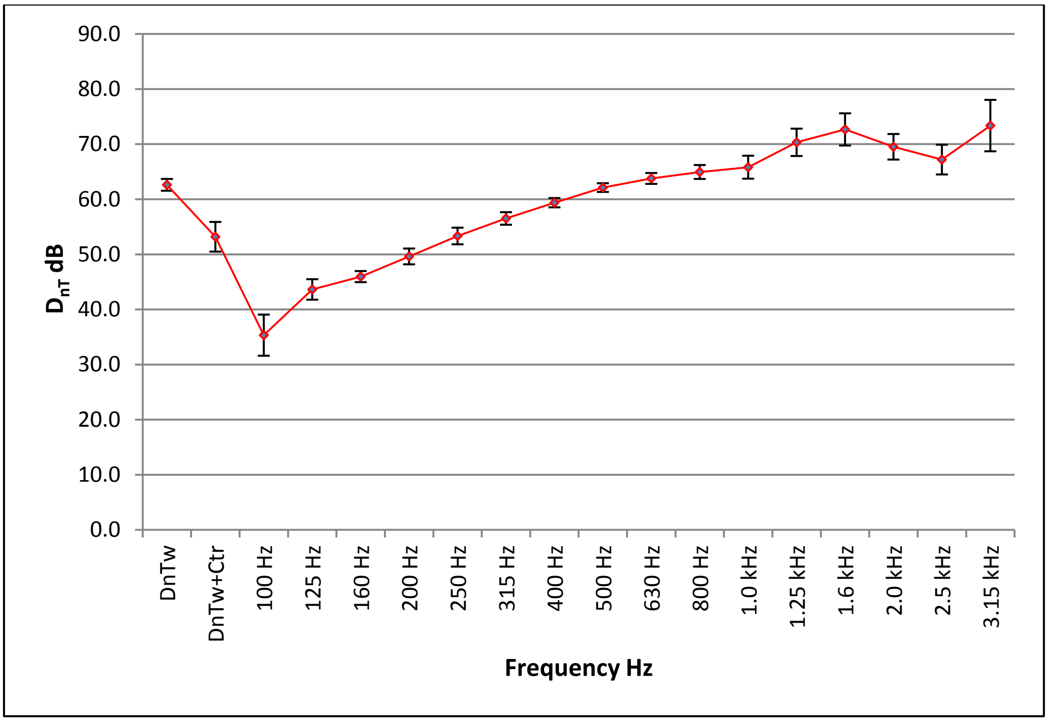


Figure 11-14: Mean of population measurements 95% confidence interval MLS – Timber Floor.

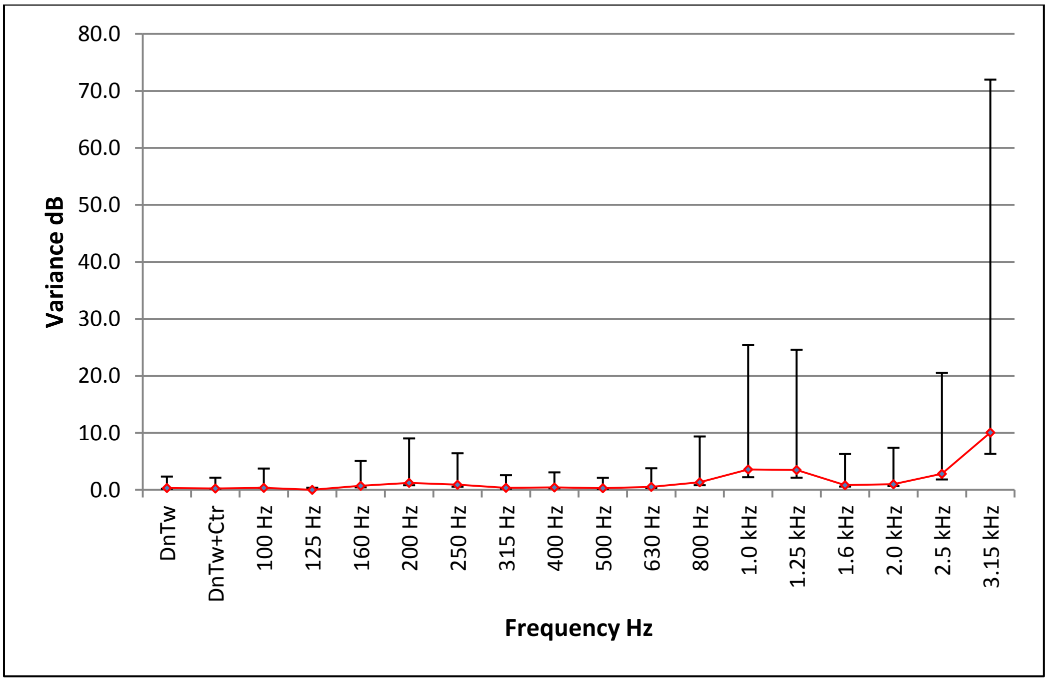


Figure 11-15: 95% Confidence interval for MLS variance of the part – Timber Floor

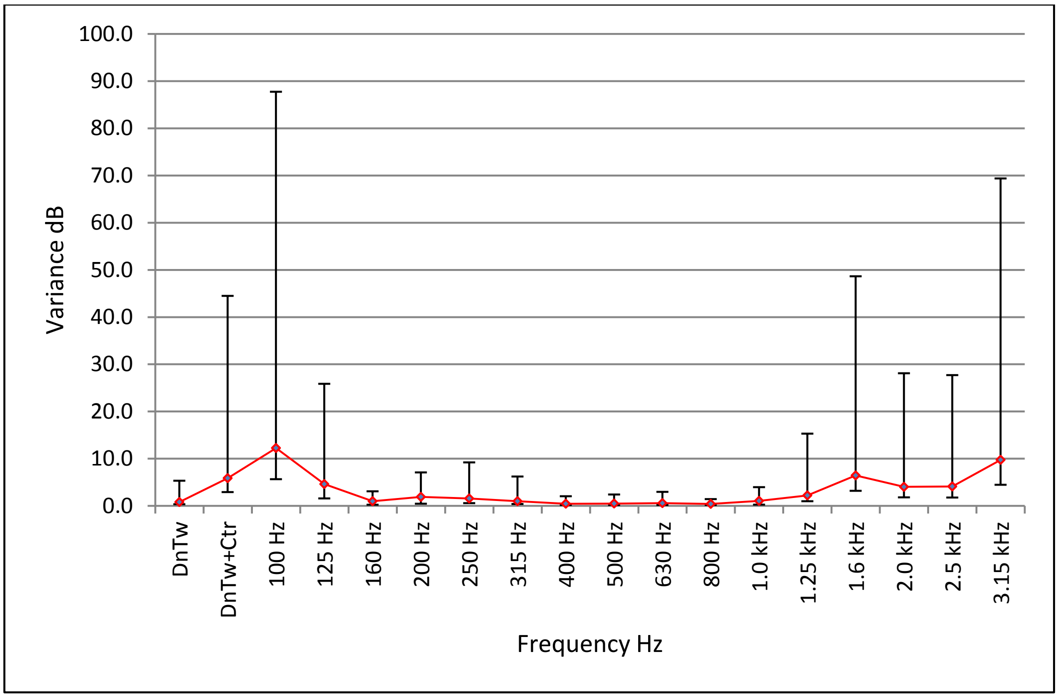


Figure 11-16: 95% Confidence interval for MLS variance of the measurement system – Timber Floor

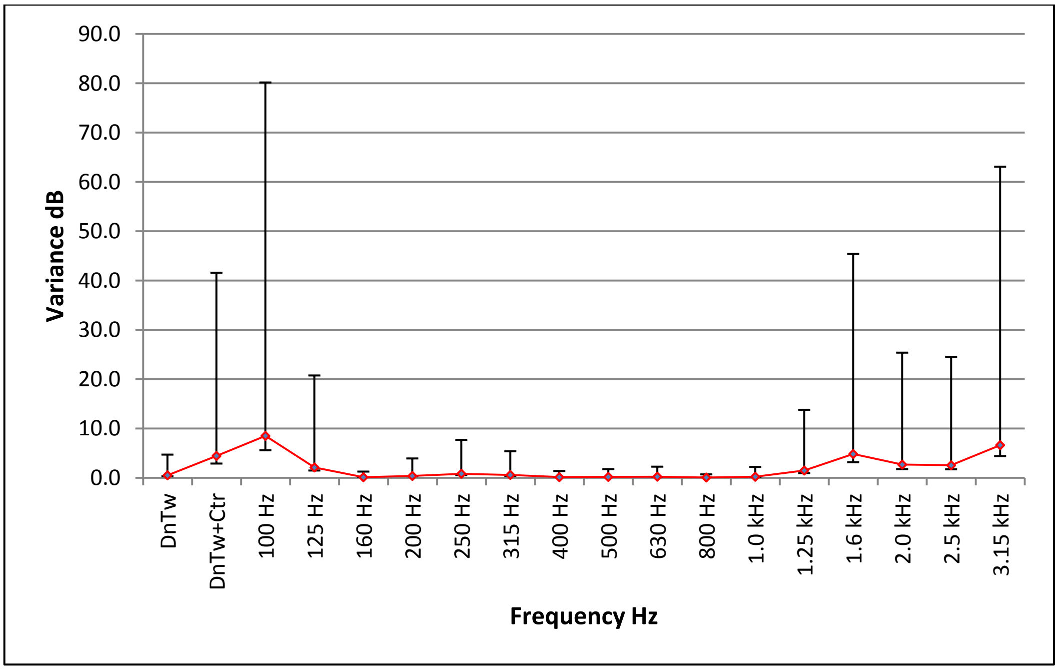


Figure 11-17: 95% Confidence interval for MLS variance of the operator – Timber Floor

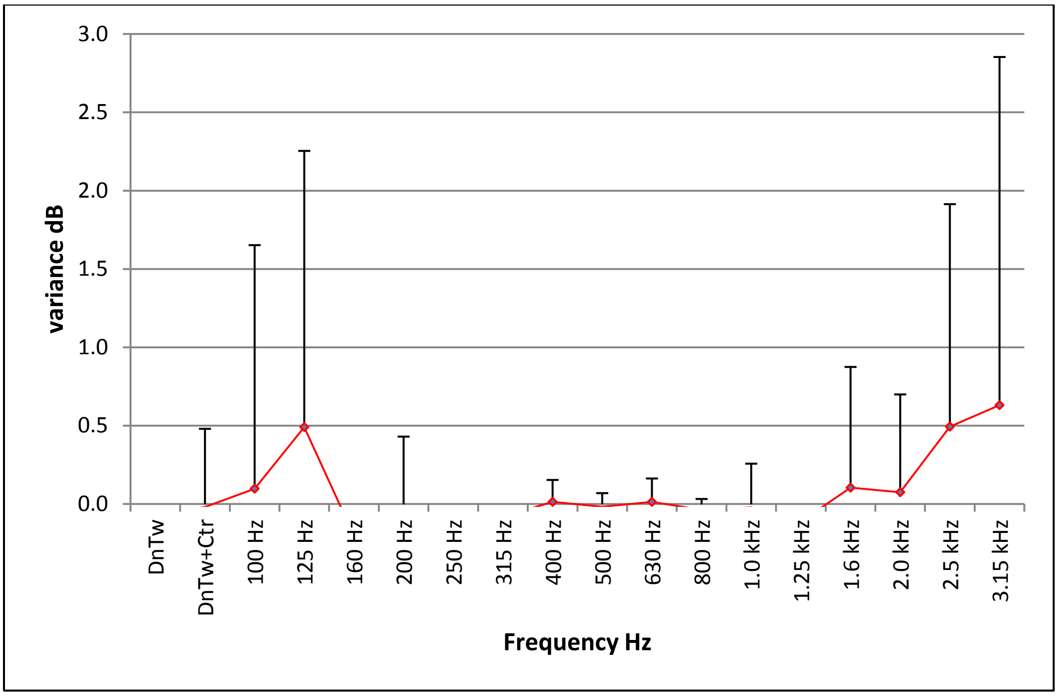


Figure 11-18: 95% Confidence interval for MLS variance of the part & operator interaction – Timber Floor

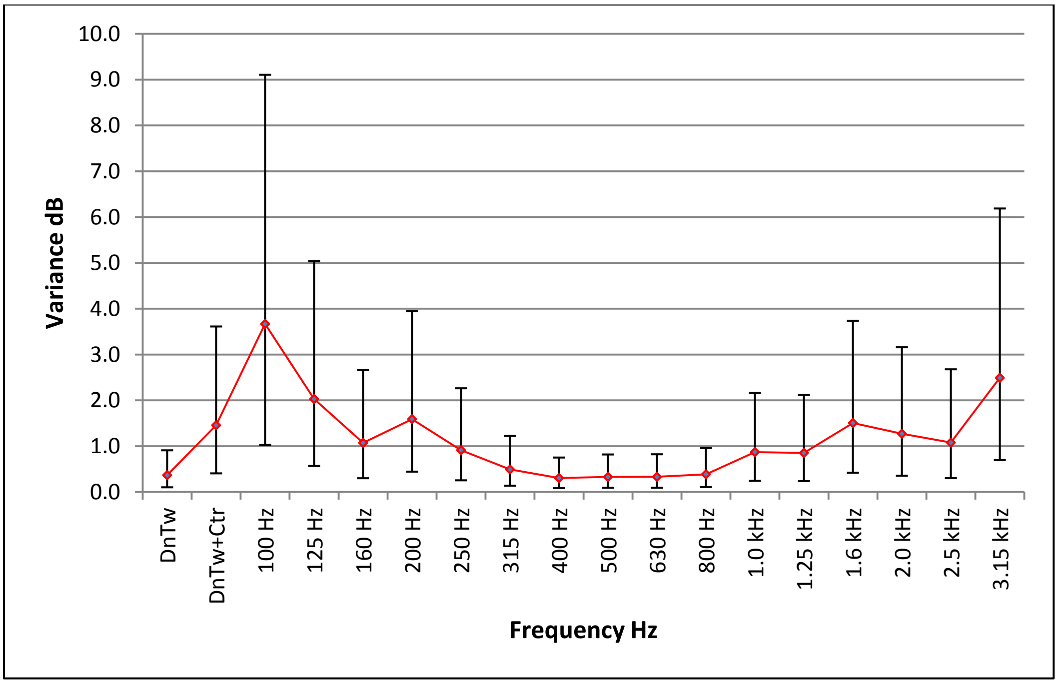


Figure 11-19: 95% Confidence interval for MLS variance of the repeatability – Timber Floor

11.5 Computed Confidence Intervals Concrete Floor

The concrete floor GRR has non-identical parts and therefore there is an element of variability in the part that dominates the total variability in the experiment. The confidence levels for the GRR are tabulated for the 95% confidence intervals in Table 11-16 and Table 11-17:

Table 11-16: 95% confidence intervals - Concrete Floor $\mu_y, \gamma_P, \gamma_M$

95%	μ_y			γ_P			γ_M		
dB	Estimate μ	Lower	Upper	Estimate γ_p	Lower	Upper	Estimate γ_m	Lower	Upper
DnTw	59.0	2.1	2.1	3.4	1.3	20.6	0.9	0.6	3.3
DnTw+Ctr	53.0	1.4	1.4	0.8	0.2	5.7	1.8	1.3	6.8
100 Hz	38.5	2.5	2.5	2.2	0.6	15.1	5.8	4.0	23.7
125 Hz	41.4	2.2	2.2	0.9	0.0	7.9	6.9	5.1	24.3
160 Hz	43.0	1.2	1.2	0.2	0.0	3.1	3.6	2.7	8.6
200 Hz	46.8	2.8	2.8	5.6	2.0	35.5	3.4	2.5	11.7
250 Hz	51.3	1.9	1.9	2.7	0.8	17.8	2.8	2.1	5.5
315 Hz	53.1	2.5	2.5	5.4	2.0	33.0	1.7	1.3	3.7
400 Hz	56.2	2.3	2.3	4.3	1.6	26.4	1.8	1.3	5.3
500 Hz	58.7	3.2	3.2	8.8	3.3	53.4	1.8	1.4	4.6
630 Hz	61.6	2.8	2.8	6.1	2.3	37.5	2.5	1.9	7.8
800 Hz	63.6	3.1	3.1	7.9	3.0	48.6	2.9	2.2	8.0
1.0 kHz	64.7	3.2	3.2	8.0	3.0	49.1	2.2	1.5	9.3
1.25 kHz	64.0	2.9	2.9	6.1	2.3	37.4	2.0	1.4	8.5
1.6 kHz	63.3	3.1	3.1	8.1	3.1	49.1	1.3	1.0	3.7
2.0 kHz	61.4	2.8	2.8	6.6	2.6	40.2	1.1	0.8	3.4
2.5 kHz	58.5	3.3	3.3	9.6	3.7	58.2	1.0	0.7	2.3
3.15 kHz	58.1	4.3	4.3	16.5	6.4	99.4	0.8	0.6	1.9

Table 11-17: 95% confidence intervals - Concrete Floor $\sigma_O^2, \sigma_{po}^2, \sigma_E^2$

95%	σ_O^2			σ_{PO}^2			σ_E^2		
dB	Estimate σ_O^2	Lower	Upper	Estimate σ_{PO}^2	Lower	Upper	Estimate σ_E^2	Lower	Upper
DnTw	0.29	0.1	2.7	0.1	0.0	0.4	0.5	0.4	0.7
DnTw+Ctr	0.57	0.1	5.6	0.4	0.1	1.2	0.8	0.6	1.2
100 Hz	2.2	0.6	20.0	0.7	-0.2	2.4	2.9	2.1	4.3
125 Hz	1.9	0.3	19.0	1.3	0.1	4.1	3.7	2.6	5.4
160 Hz	0.3	-0.3	4.8	1.4	0.5	3.5	2.0	1.4	2.9
200 Hz	0.8	0.1	8.9	1.2	0.5	3.1	1.4	1.0	2.0
250 Hz	0.0	-0.3	2.3	1.0	0.3	2.7	1.8	1.3	2.6
315 Hz	0.2	0.0	2.0	0.2	-0.2	0.8	1.4	1.0	2.0
400 Hz	0.4	0.0	3.8	0.3	-0.1	1.0	1.1	0.8	1.7
500 Hz	0.2	0.0	2.9	0.3	-0.1	1.1	1.3	0.9	1.9
630 Hz	0.6	0.1	5.8	0.4	-0.1	1.3	1.5	1.1	2.3
800 Hz	0.5	0.1	5.5	0.1	-0.4	1.1	2.3	1.6	3.4
1.0 kHz	0.9	0.2	8.0	0.4	0.1	1.2	0.9	0.6	1.3
1.25 kHz	0.8	0.2	7.2	0.4	0.1	1.1	0.8	0.6	1.2
1.6 kHz	0.2	0.0	2.6	0.2	0.0	0.7	0.8	0.6	1.2
2.0 kHz	0.3	0.1	2.5	0.0	-0.2	0.2	0.9	0.6	1.3
2.5 kHz	0.1	0.0	1.4	0.2	0.0	0.6	0.7	0.5	1.0
3.15 kHz	0.1	0.0	1.1	0.1	-0.1	0.4	0.6	0.4	0.9

Again, the lower limits are left as calculated in the table although any <0 would normally be assumed to be zero.

The estimates and their confidence limits are shown in Figure 11-20 – Figure 11-25:

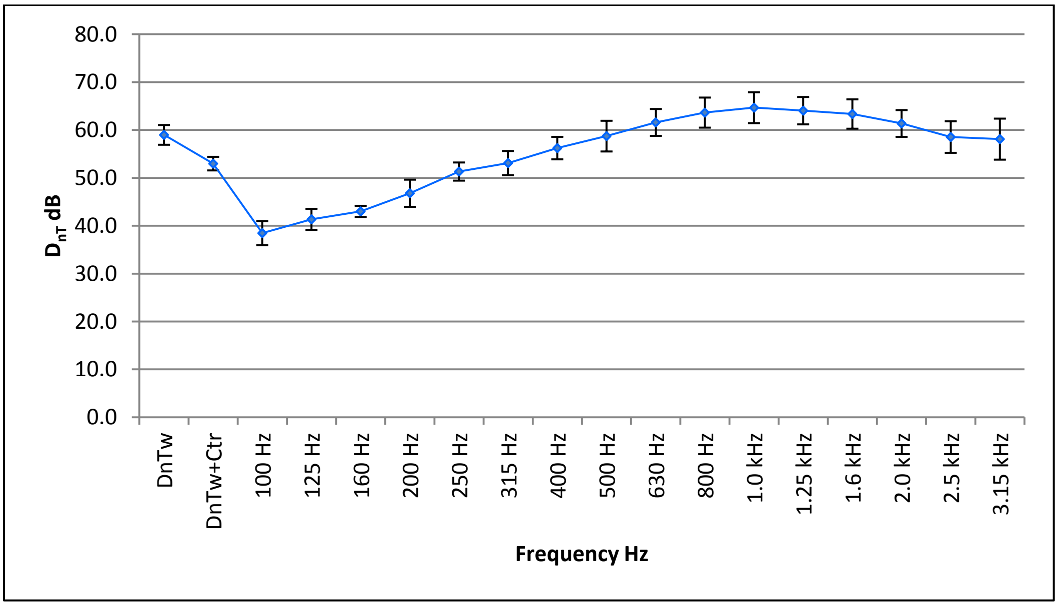


Figure 11-20: Mean of population measurements 95% confidence interval MLS – Concrete Floor

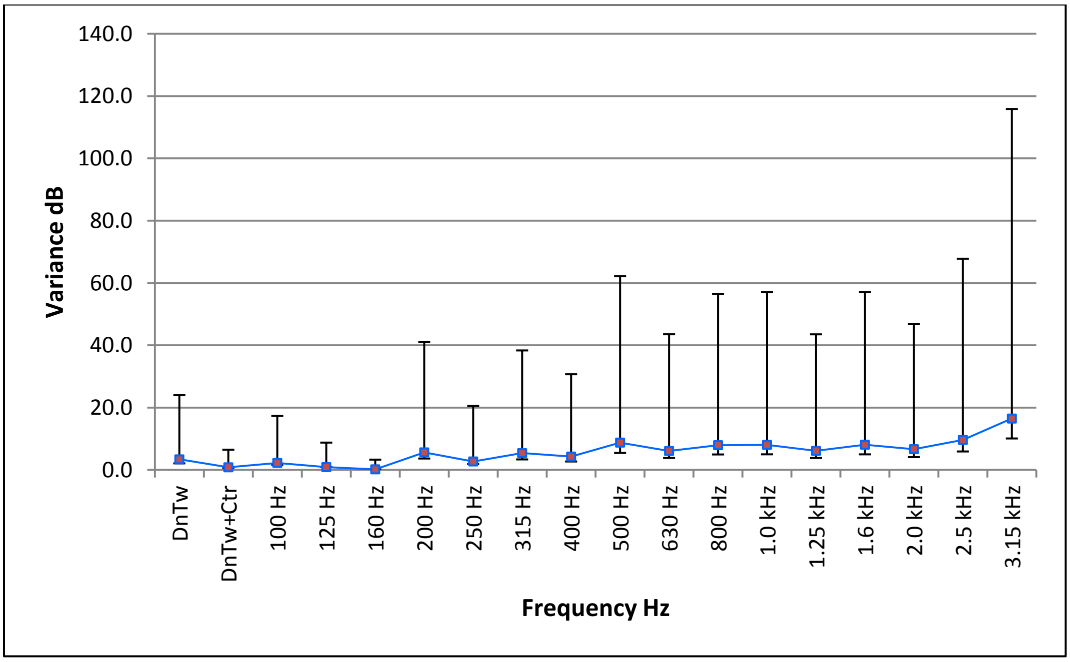


Figure 11-21: 95% Confidence interval for MLS variance of the part – Concrete Floor

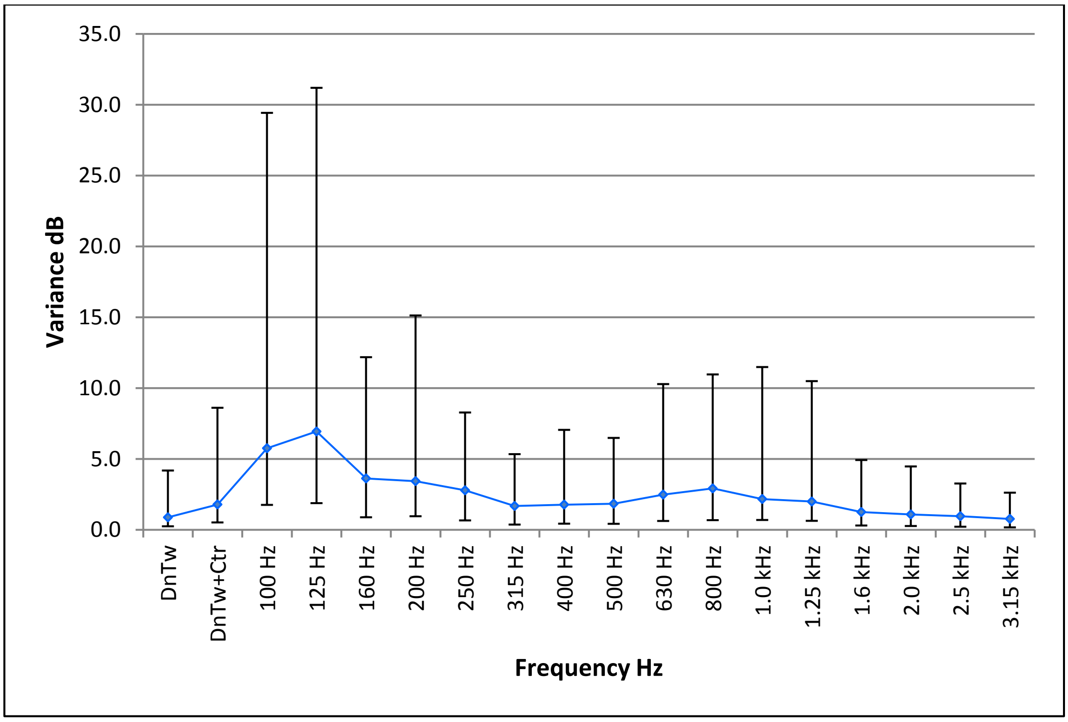


Figure 11-22: 95% Confidence interval for MLS variance of the measurement system – Concrete Floor

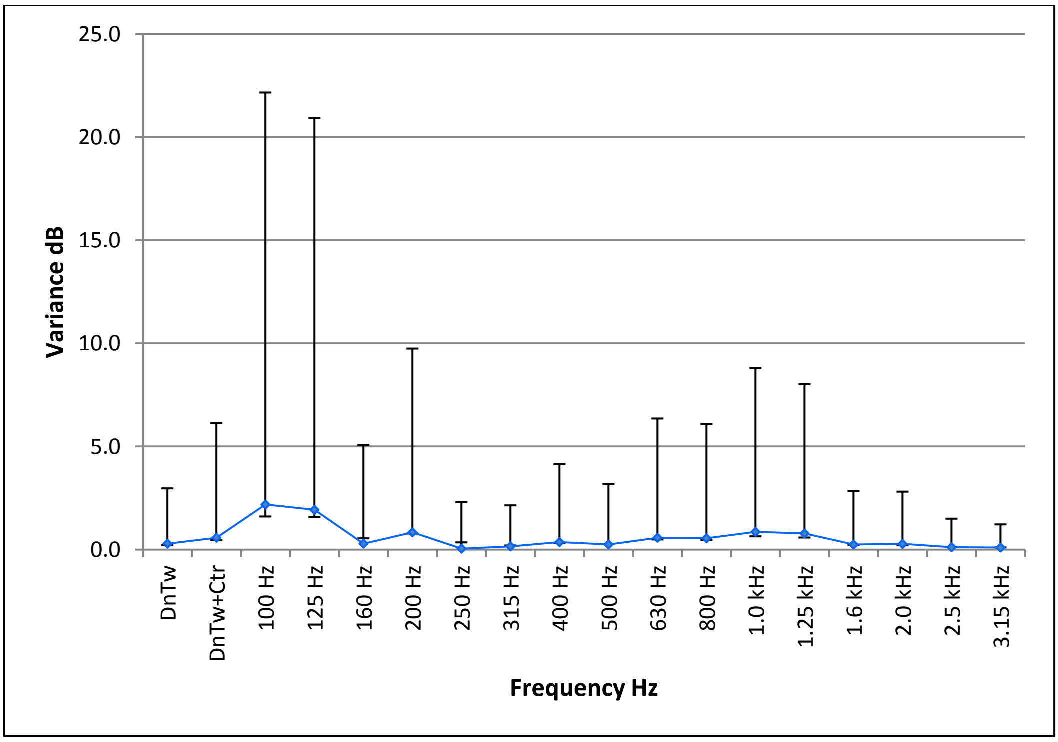


Figure 11-23: 95% Confidence interval for MLS variance of the operator – Concrete Floor

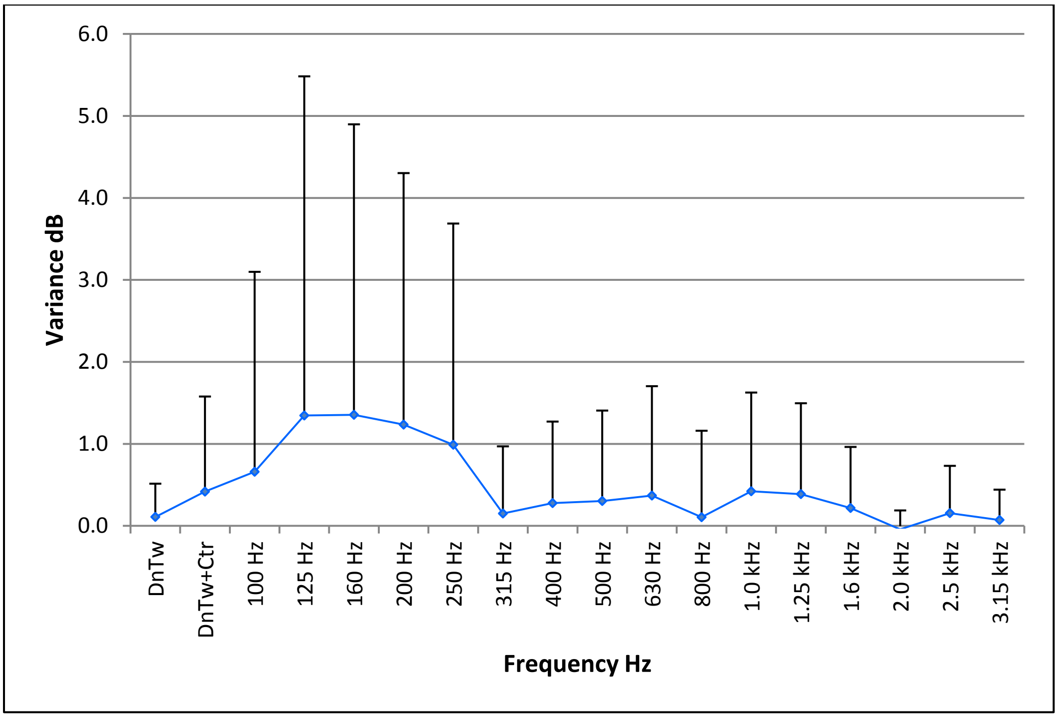


Figure 11-24: 95% Confidence interval for MLS variance of the part & operator interaction – Concrete Floor

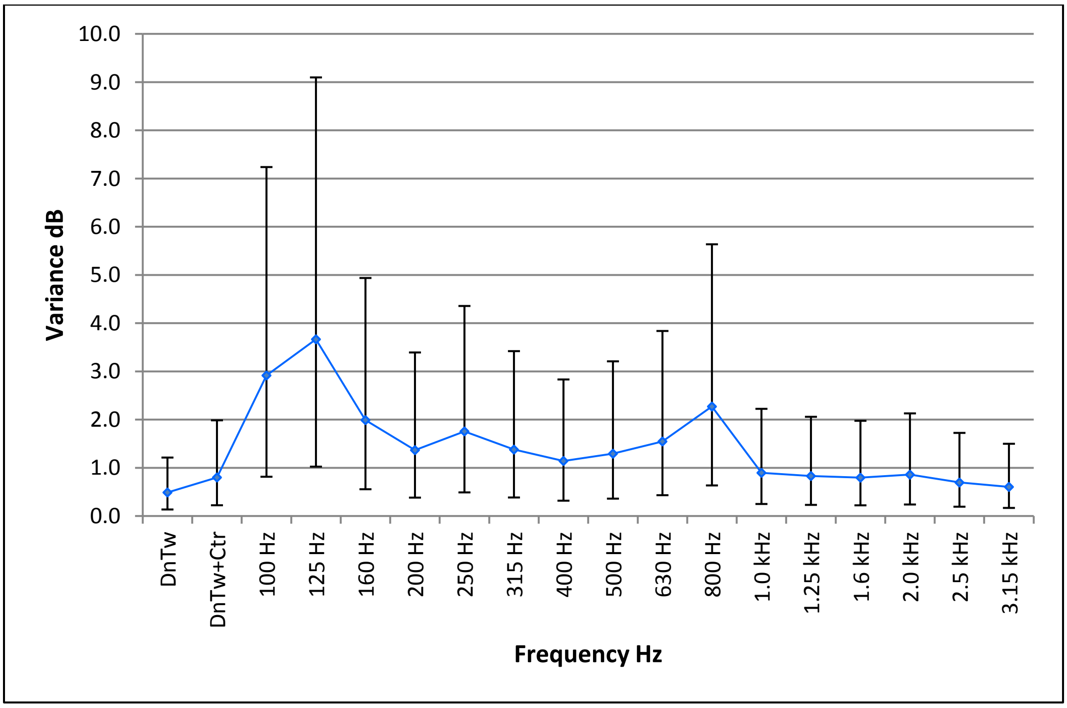


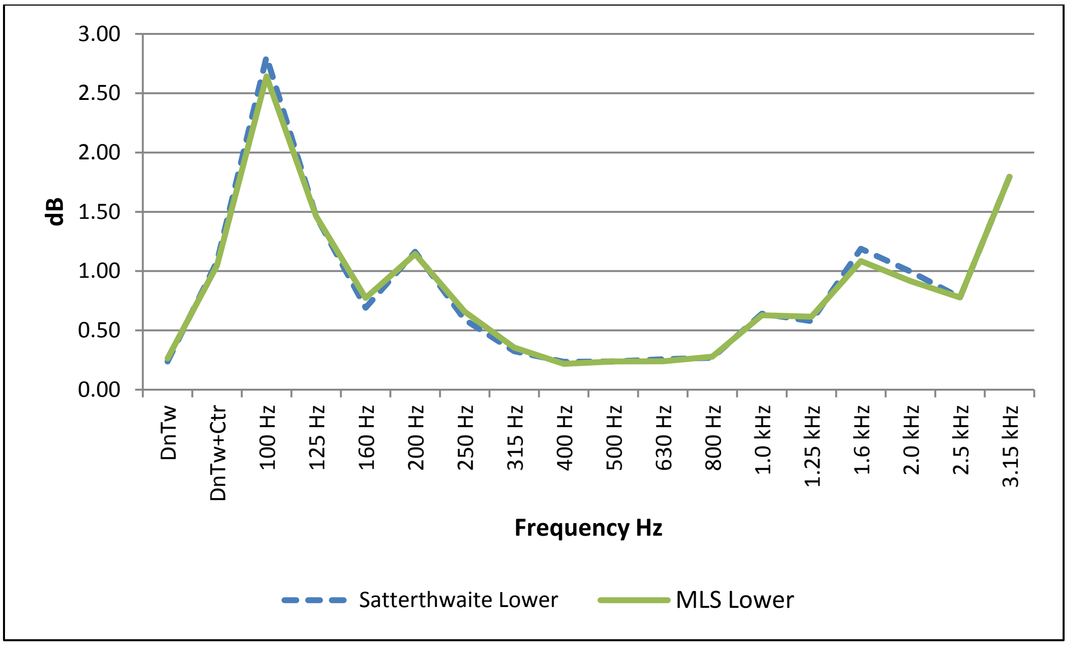
Figure 11-25: 95% Confidence interval for MLS variance of the repeatability – Concrete Floor

What is notable about the confidence intervals for $\gamma_{p,}$ σ_o^2 & σ_{po}^2 is their magnitude and their asymmetry. For the part variance component $\gamma_{p,}$ the timber floor shows much lower variance and smaller confidence intervals across the frequency range than the concrete floor. This characteristic of the timber GRR confidence interval is primarily due to the variability in room size for the concrete GRR.

The σ_o^2 has relatively large confidence intervals at the low frequency end of the spectrum e.g. 100Hz and 125Hz bands for both timber and concrete GRR. The timber floor also exhibits relatively large confidence intervals above 1250Hz where site background noise affected measurements. An additional consideration, at both 2500Hz and 3150Hz, is the reduction in degrees of freedom caused by fitting the full model. In these bands (v = 3).

The confidence intervals for the interaction term σ_{po}^2 are only relevant for the concrete GRR (apart from the 800Hz band in the timber GRR). The confidence interval is asymmetrical and there is virtually no lower limit. This is because for most frequency bands in the Concrete GRR, negative lower value are shown which defaults to 0. For the timber GRR, σ_{po}^2 is shown for information only as its estimate is negative for a number of frequency bands. It is also not significant where (α>0.25). As stated previously, where this occurs, the interaction term is removed from the model.

The MLS approximations can be assessed against the Satterthwaite confidence intervals for repeatability and the measurement system (Gauge). The upper and lower limits are shown for the timber and concrete case. In addition, for the timber GRR, the $\sigma_{reproducibility}^2$ for the Satterthwaite approximation is compared to the MLS confidence interval for σ_o^2 ; see Figure 11-26 – Figure 11-31.



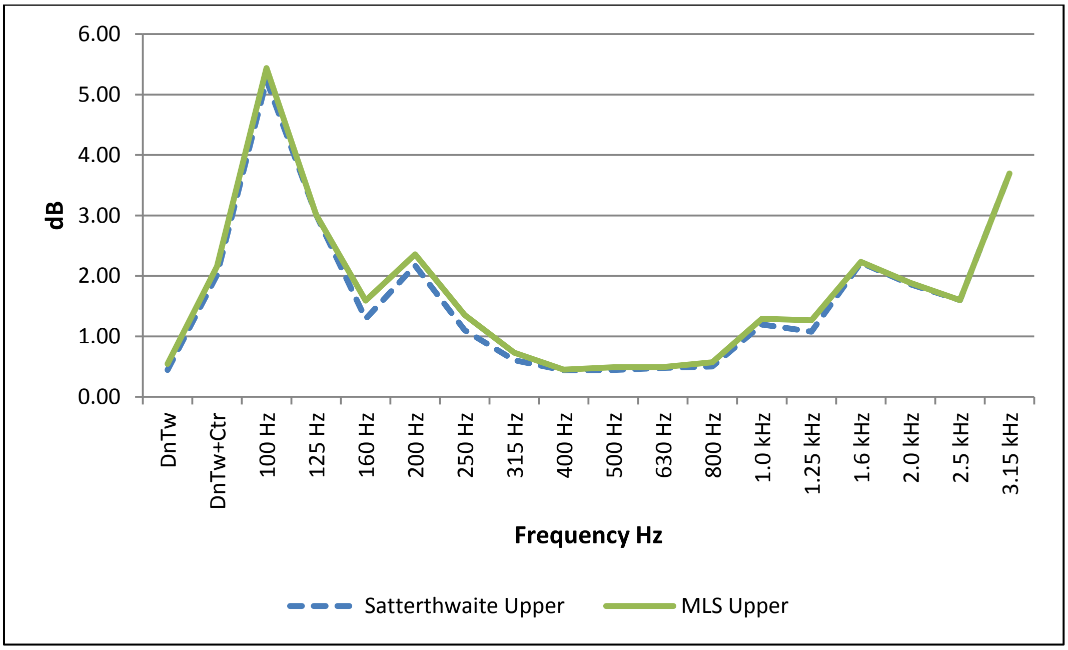
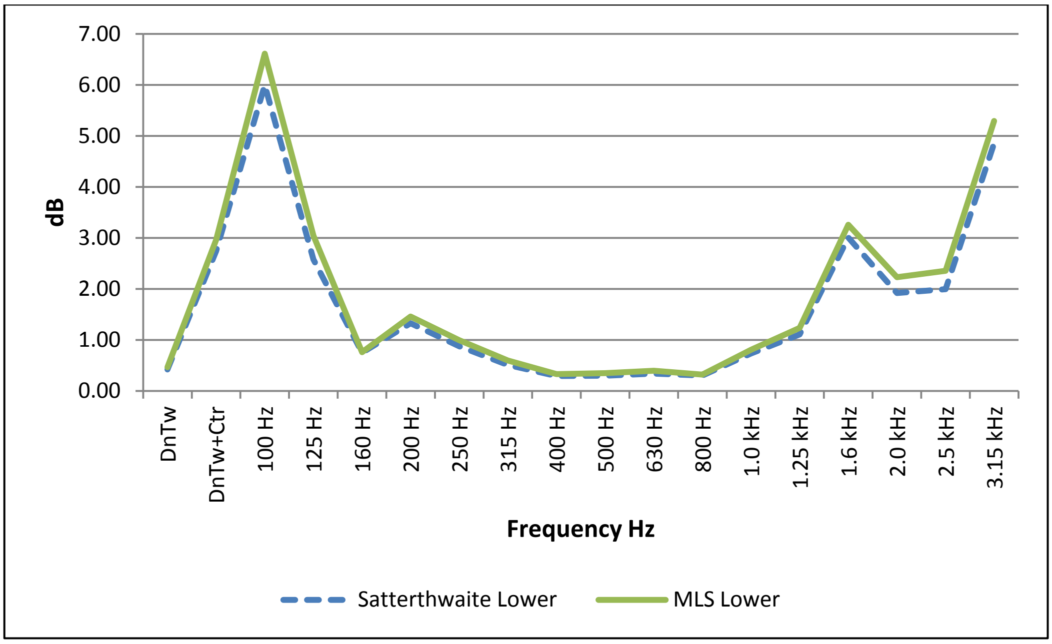


Figure 11-26: Timber GRR – Repeatability Confidence Limit approximations: Satterthwaite & MLS

The upper and lower confidence interval approximations by Satterthwaite and MLS methods produce relatively similar results for the timber repeatability data.



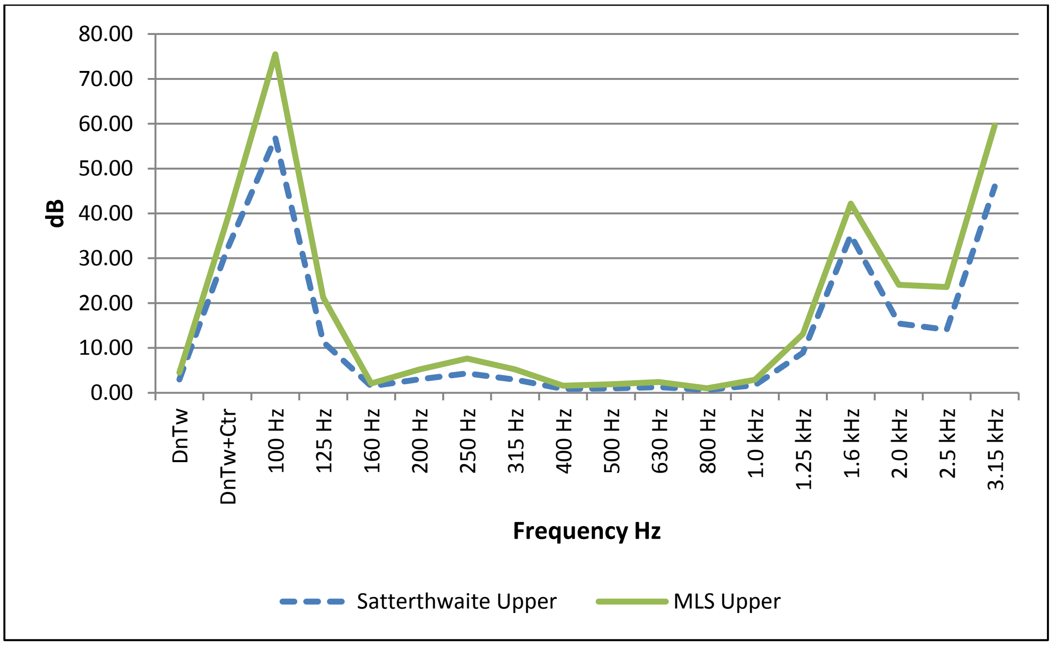
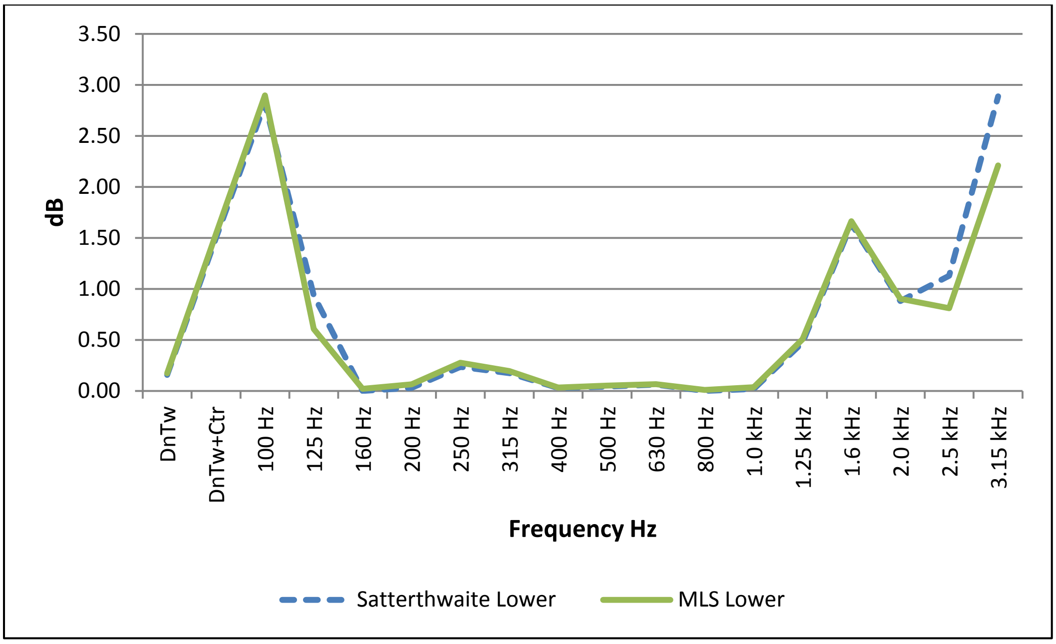


Figure 11-27: Timber GRR – Gauge Confidence Limit approximations: Satterthwaite & MLS

The confidence interval approximations were generally lower for the Satterthwaite method than for the MLS. The differences were more pronounced in the upper interval values, especially for the frequencies 100-125Hz and for 2000-3150Hz.



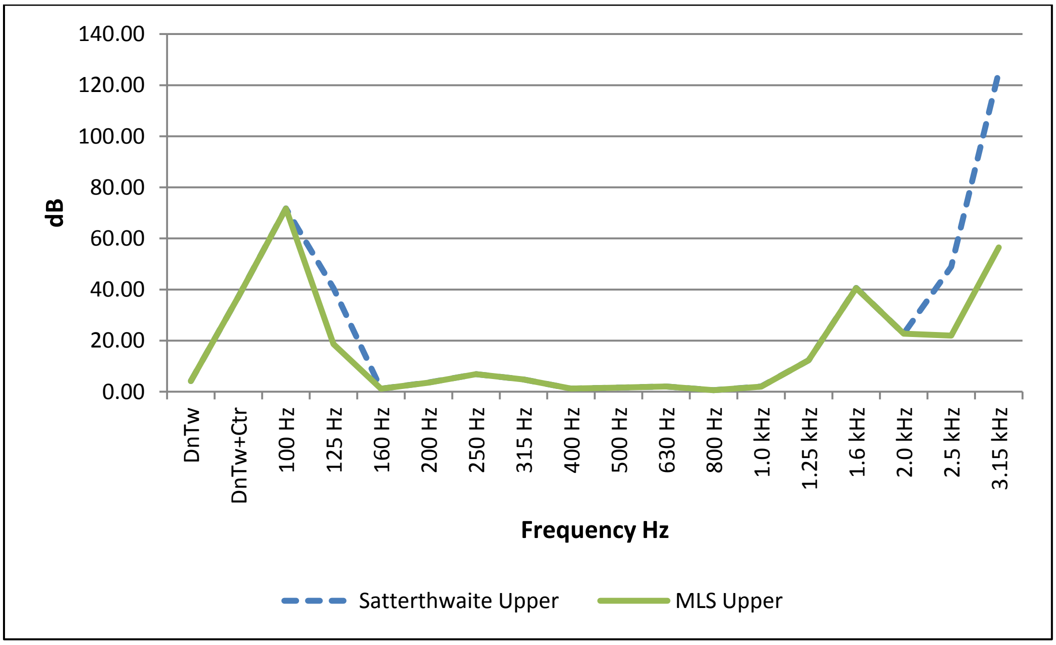


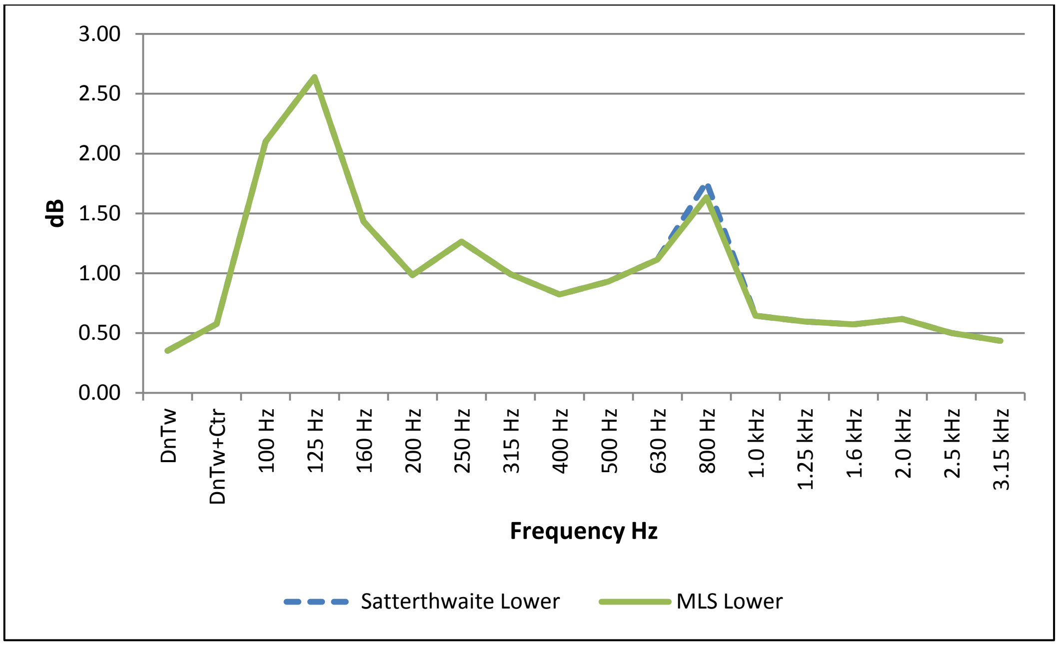
Figure 11-28: Timber GRR – Operator (Reproducibility) Confidence Limit approximations: Satterthwaite & MLS

For the timber GRR the reduced model was fitted over the majority of frequencies, so:

$$\hat{\sigma}^2_{reproducibility} = \hat{\sigma}^2_O \quad (11\text{-}11)$$

There is close agreement over the assessment range where this is the case.

The full model (with interaction) was fitted for the Satterthwaite approximation in the 125Hz, 2500Hz and 3150Hz bands. The Satterthwaite approximation gives higher levels than MLS for the upper and lower limits on the confidence interval.



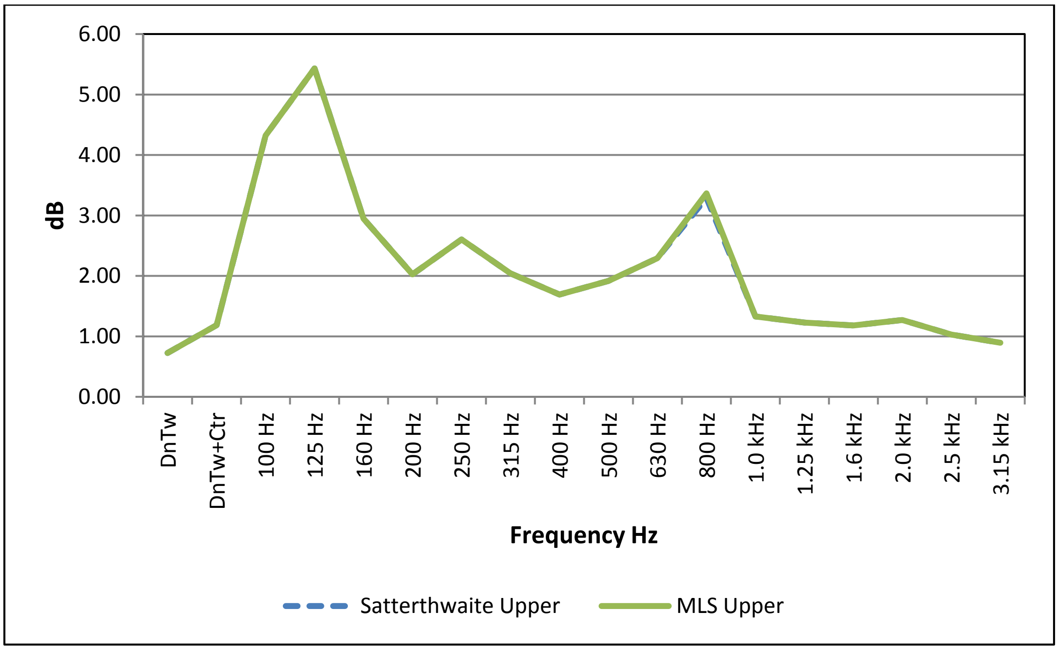


Figure 11-29: Concrete GRR – Repeatability Confidence Limit approximations: Satterthwaite & MLS

For the concrete GRR repeatability, the upper and lower confidence interval approximations by Satterthwaite and MLS give similar values.

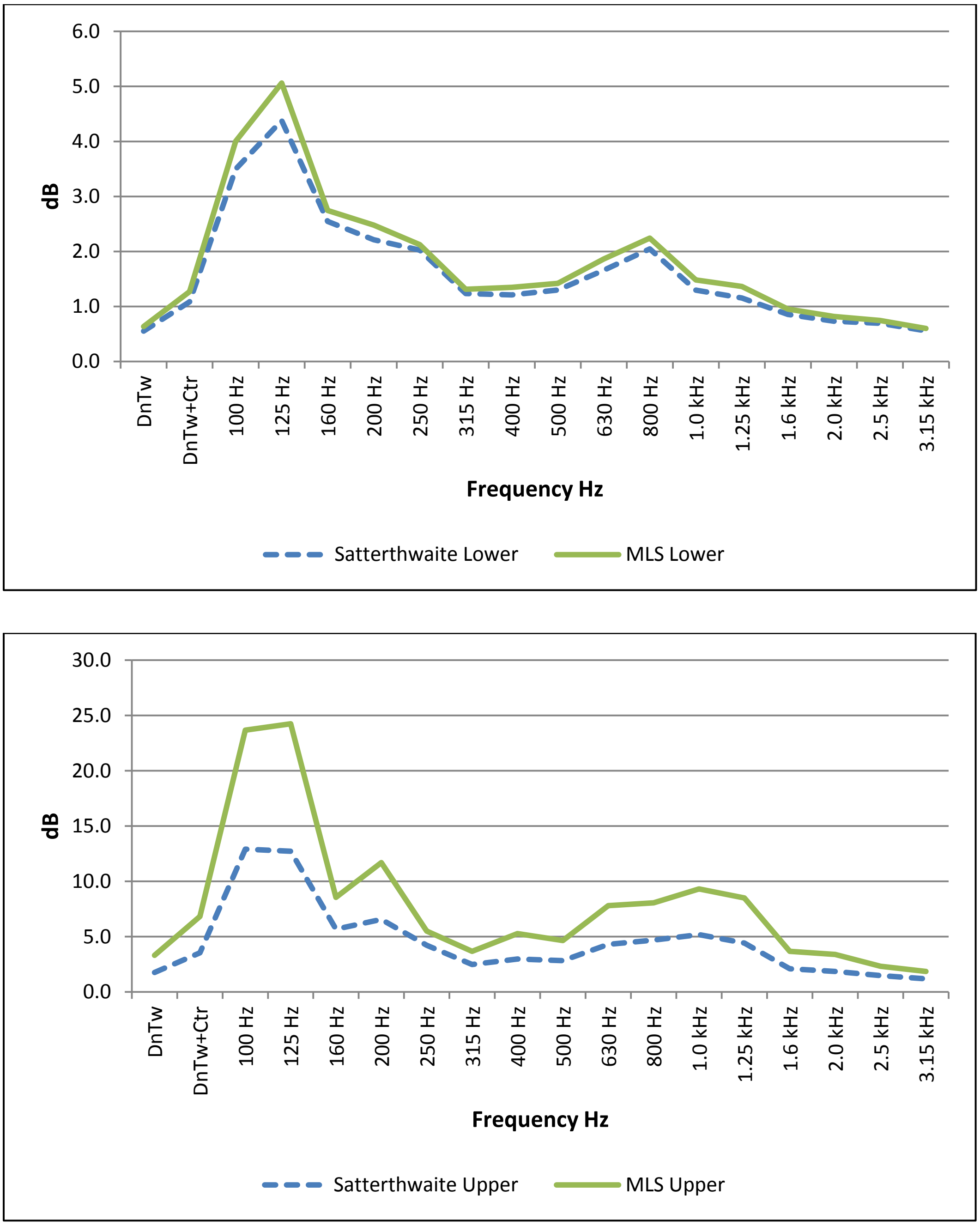
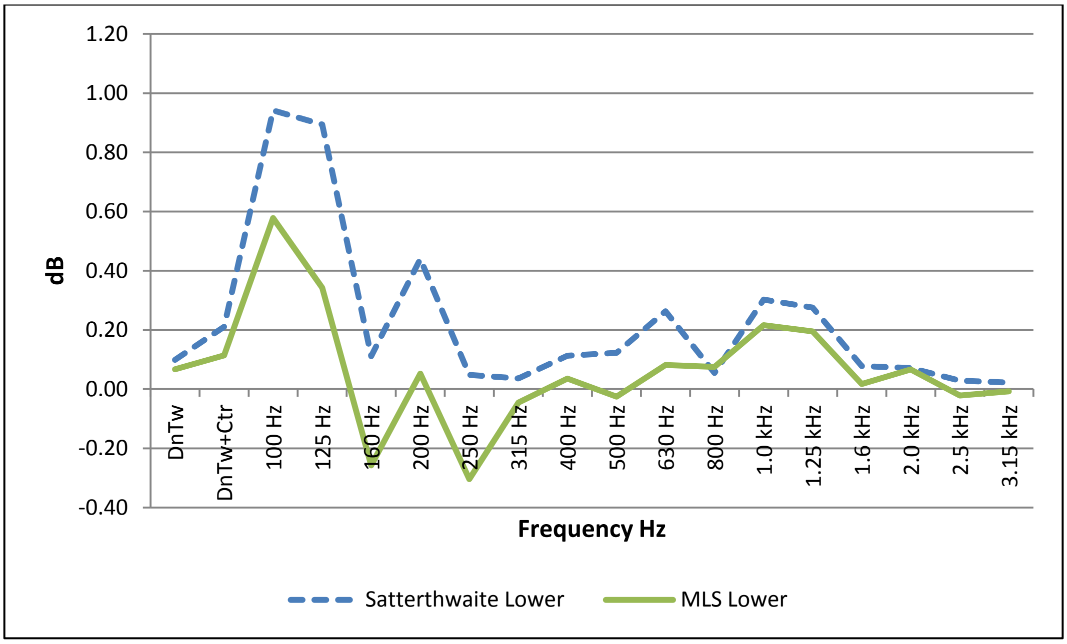


Figure 11-30: Concrete GRR – Gauge Confidence Limit approximations: Satterthwaite & MLS

The confidence interval approximations for $\hat{\sigma}^2_{gauge}$ were generally lower than for the Satterthwaite method than the MLS. The bottom graph in Figure 11-30 shows the differences were more pronounced in the upper interval values, especially for the lower frequencies 100-125Hz.



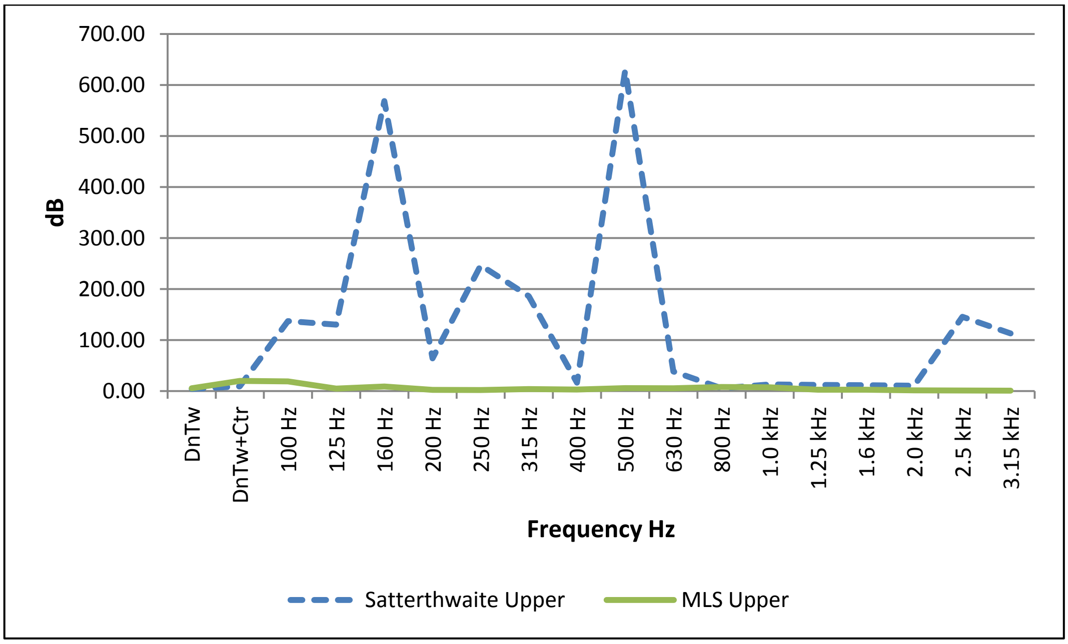


Figure 11-31: Concrete GRR – Operator (Reproducibility) Upper Confidence Limit approximations: Satterthwaite & MLS

For the concrete data, the full model (with interaction) was fitted in all but the 800Hz case. The MLS method does not have a documented calculation of confidence intervals for $\sigma^2_{reproducibility}$, when interaction is present in the model. Most of the research literature assume reduced models with no interaction. For the concrete data the Satterthwaite approximation was constrained by relatively low values for υ. In all frequency bands υ < (o-1) and in some cases the Chi-squared value was based on υ = 1.

As a result there is a lack of precision in the Satterthwaite confidence interval approximation for $\sigma^2_{reproducibility}$.

11.6 Conclusions

To understand the measurement process, both point and interval estimates are required. The use of ANOVA allows the components of variance to be both identified and quantified using standard formulae. However, it is noted that the confidence limits calculated may not be of any practical use at all frequencies for many of the variance components in this study due to their size and poor precision

Two calculation methods, developed for ANOVA, have been used to estimate the confidence intervals relating to the components of variance. Both the Satterthwaite approximation and MLS method are used and they result in similar upper and lower confidence limits for $\sigma^2_{repeatability}$, for both timber and concrete data

The Satterthwaite confidence interval values for σ^2_{gauge} are lower than those estimated by the MLS method and narrower, which is in agreement with the findings of Borror et al [71]. This does not automatically result in an improvement in the precision of the confidence interval, as the precision of the confidence interval is related to the estimated value of $\hat{\sigma}^2_{gauge}$. The point estimator for $\hat{\sigma}^2_{gauge}$ is generally lower for timber than for concrete floors, even though the precision is not as good.

The Satterthwaite approximations for $\sigma^2_{reproducibility}$ give a confidence interval which is, at least, 8 times the value of $\hat{\sigma}^2_{reproducibility}$. This is due to the Chi-squared value returned for low levels of (o-1) degrees of freedom for the reduced model or for low values of v for the full model. This adversely affects both calculation processes and results in confidence intervals that are large for both data sets. Both models are used in calculating the confidence intervals in the GRR and it has been shown that options for improving the confidence interval are limited. This means GRR studies, which typically have few operatives (because it substantially increases the size of the experiment) must therefore be content with relatively imprecise estimates of $\sigma^2_{reproducibility}$, consider restructuring the DOE, or find an alternative method to determine the confidence intervals.

The wide confidence limits, determined using Satterthwaite and MLS for this study have shown both the limitations of the DOE and the calculation processes themselves. The results obtained suggest that the software orientated computer simulation approach adopted by GCI, REML[93] and others [95] is likely to be more useful for future work if they can be used to provide narrower confidence intervals that are more precise, though this is beyond the scope of this study.

12 Uncertainty and Historical Data

12.1 Introduction

The previous chapters have identified the major components of variance in the measurement process and evaluated their contribution to the total uncertainty obtained. It has been noted that the part which is being measured, in this case the separating floor construction, has its own variability and this appears to be significantly different for timber and concrete floors.

To demonstrate how this research might be applied to existing or previously measured data which has not had the uncertainty calculated, the measurement uncertainty due to the part is derived from the sound insulation data published in 1958 by Parkin et al [13].

12.2 Parkin Revisited

In order to examine the contribution of the construction to the uncertainty seen in the measurement process the DOE focussed on selecting rooms of the same shape and size to minimise the dimensional influences on the variance of the part. A requirement therefore, is to obtain historical data which has several examples of airborne sound insulation tests being undertaken on the same floor construction in the same shape and size rooms. Parkin et al [13] carry out such a test for a simple concrete floor which has been previously detailed in Figure 1-1 in Chapter 1. There are 29 test results in total for the concrete floor construction and the total variance in the results can be calculated in each third octave band. The data is plotted in Figure *12-1*:

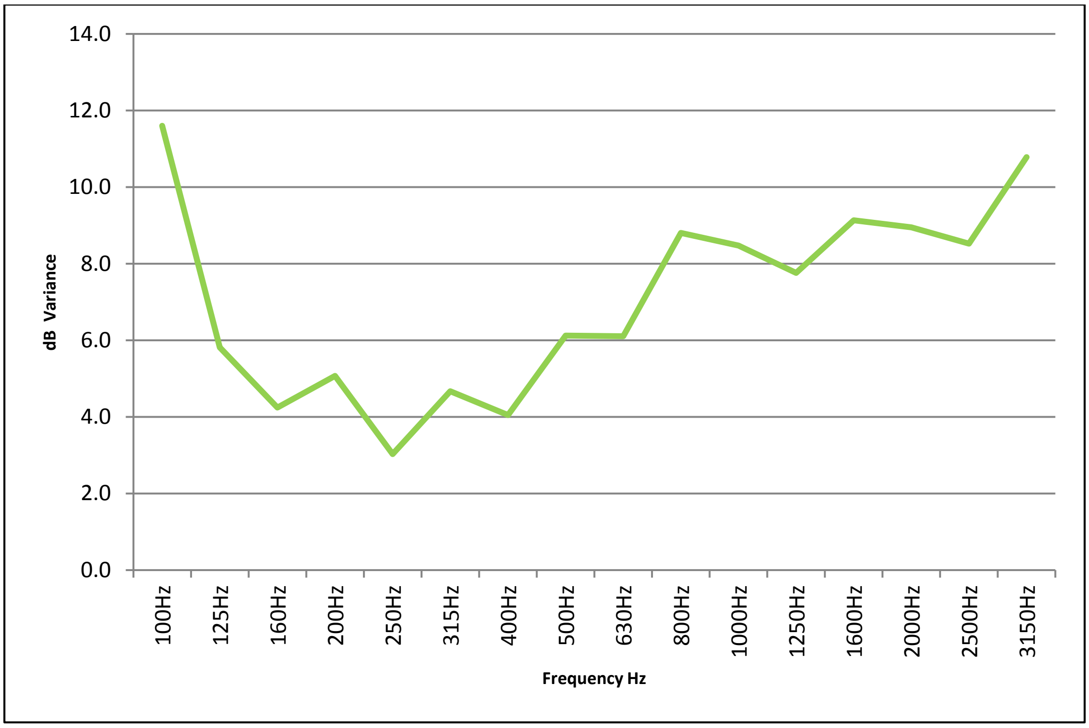


Figure 12-1: Total Variance - Simple Concrete Floor - Parkin et al (1960)

The corresponding third octave band variances for the concrete GRR are shown below in **Error! Reference source not found.** for ease of reference:

Table 12-1: Third Octave Band Standard Uncertainties – Concrete floor

D_{nT} (s.d.)	σ_{GRR}	σ_r	σ_R	σ_o	$\sigma_{p.o}$	σ_p	σ_{Total}
100Hz	2.4	1.7	1.7	1.5	0.8	1.5	2.8
125Hz	2.6	1.9	1.8	1.4	1.2	0.9	2.8
160Hz	1.9	1.4	1.3	0.5	1.2	0.4	2.0
200Hz	1.9	1.2	1.4	0.9	1.1	2.4	3.0
250Hz	1.7	1.3	1.0	0.2	1.0	1.6	2.3
315Hz	1.3	1.2	0.6	0.4	0.4	2.3	2.7
400Hz	1.3	1.1	0.8	0.6	0.5	2.1	2.5
500Hz	1.4	1.1	0.7	0.5	0.6	3.0	3.3
630Hz	1.6	1.2	1.0	0.8	0.6	2.5	2.9
800Hz	1.7	1.5	0.8	0.8	0.0	2.8	3.3
1000Hz	1.5	1.0	1.1	0.9	0.7	2.8	3.2
1250Hz	1.4	0.9	1.1	0.9	0.6	2.5	2.9
1600Hz	1.1	0.9	0.7	0.5	0.5	2.8	3.1
2000Hz	1.1	0.9	0.5	0.5	0.0	2.6	2.8
2500Hz	1.0	0.8	0.5	0.3	0.4	3.1	3.3
3150Hz	0.9	0.8	0.4	0.3	0.3	4.1	4.2

The average simple floor performance can be compared with the concrete floor used in the GRR: These data are shown in Figure *12-2*:

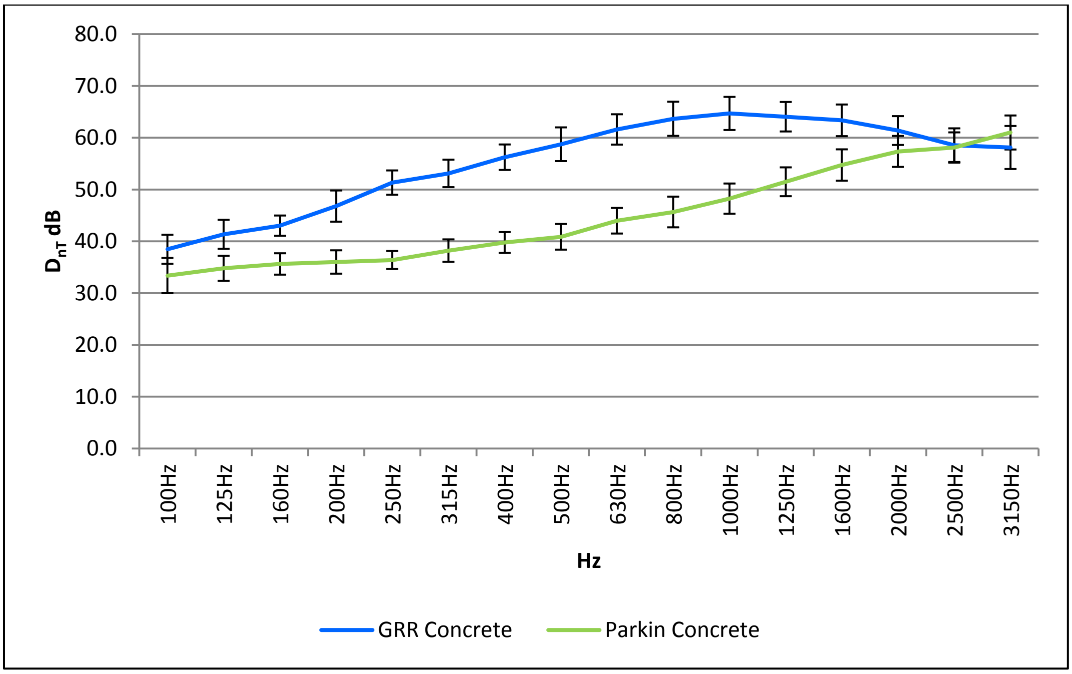


Figure 12-2*: GRR Floor & Parkin Concrete Floor mean D_{nT} values & standard deviations*

The total variance in both the Concrete and the Parkin floor is shown in Figure 12-3:

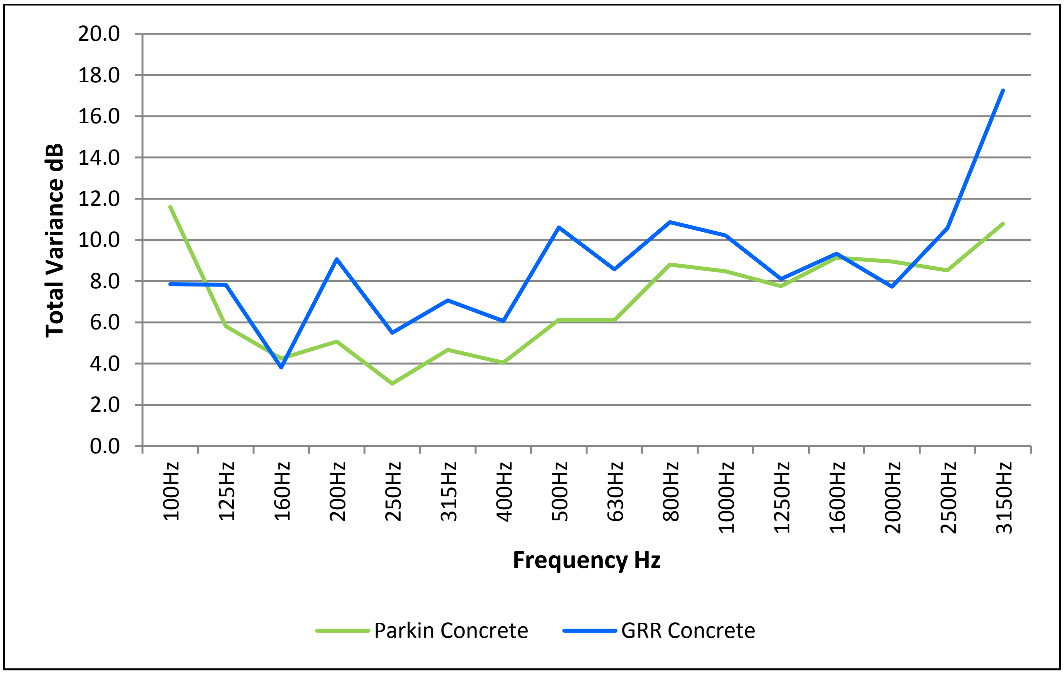


Figure 12-3: Comparison of the total variance for concrete GRR Floor and Parkin concrete floor

The total variance of the concrete GRR floor is greater than the variance of the Parkin floor between 200Hz – 1000Hz, though it appears to follow a similar signature. The differences at the low and high frequency ends may be due to the non-diffuse field and background noise contributions, respectively.

If it is assumed that the instrumentation variance (repeatability) is relatively stable and fixed and the contribution of the operator and part & operator interaction (reproducibility) is similar to the concrete floor GRR, it is possible, using the principle that uncertainty is summed in quadrature, to subtract the combined variance of the measurement system based on the concrete GRR from the total variance measured by Parkin et al for their simple concrete floor. What remains is an estimate for the variance of the "Part" measured by Parkin; see equation 12-1:

$$PPS\sigma_{total}^{2} - GRR\sigma_{gauge}^{2} = \widehat{PPS\sigma_{part}^{2}} \qquad \textit{12-1}$$

Where:

$PPS\sigma_{total}^{2}$ = Total variance measured by Parkin et al for the simple concrete floor;

$GRR\sigma_{gauge}^{2} =$ Variance of the measurement system calculated from concrete GRR;

$\widehat{PPS\sigma_{part}^{2}}$ = Estimated variance produced by the Parkin concrete floor.

It is noted that care must be exercised in comparing these results as several of the third octave bands in the concrete GRR were affected by background noise. It is further noted that the Parkin data set is likely to have also been affected in some way by background noise contribution on site but in principle the resulting figure is an approximation to the variance associated with the construction measured.

The part to part variance for the Parkin concrete flooris thus determined and is compared to thepart to part variances for the GRR concrete floor and timber floor in Figure 12-4:

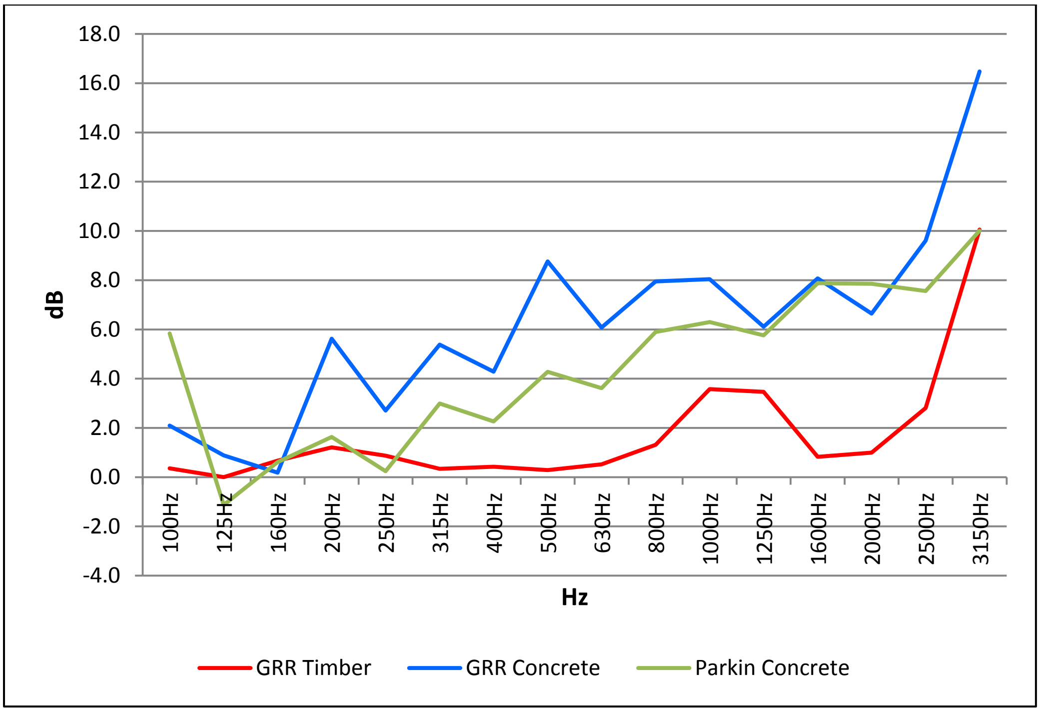


Figure 12-4: Timber & Concrete Part Variances from GRR - Parkin et al: Simple Concrete Floor

The Parkin floor variance generally increases with frequency. As Parkin's test was carried out on 29 rooms of similar size, it provides an appropriate comparison with the timber floor variance, since the contribution from the room has been minimised. The variance at 100Hz is significantly greater for the Parkin floor, which is likely to be due to the non diffuse field. It becomes negative at 125Hz. This is a feature of the simple subtraction process employed and is always likely to occur when the variance of the part is genuinely low and a zero value would normally be assigned. Results at 160Hz – 200Hz are similar to the timber floor but increase above 250Hz and are significantly different until 3150Hz. This comparison agrees with the findings of the previous comparison of the timber floor with a concrete floor sample by Craik et al [57]. In that case, the variability of the construction, or "workmanship" was seen to be higher in the simple concrete floor than the more complex timber floor construction [121].

The comparison with the concrete GRR floor must take into account the contribution likely to be due to the non-identical rooms, which as previously discussed may lead to an increase in the part variance in the GRR test floor. It is noted that the variance for both concrete floors has the same trend above 160Hz, where the variance due to the parts rise and fall in unison. This effect is not apparent in the lightweight timber floor. This correlation may suggest that there is a variance signature which is related to the part being measured.

12.3 Conclusion

This chapter has looked at a practical use for the measurement uncertainty data when it has been partitioned into its component parts. The Parkin data has been revisited and, as it conveniently lists a relatively large sample of measurement on a simple concrete floor, the part to part variability has been estimated using a simple subtraction of the effects of the gauge in the concrete GRR from the total variance measured by Parkin. Caution is required when considering this data as it has been previously shown that there is an influence associated with background noise in the concrete GRR data and there would be expected to be some contribution from background noise associated with the Parkin field tests. The exact effects of the background noise contribution are therefore unknown.

A comparison of Parkin's concrete floor with the timber and concrete samples collected in this study reveals that the variance of the part appears to follow the Concrete GRR curve pattern, if not exactly the magnitude, though this may be because the GRR example has non-identical room sizes which inflate the size of the part variance term. Notwithstanding the rooms' effect, the concrete data in this comparison suggest that the construction itself has a variability signature independent of the size of the rooms measured. This view is reinforced to some extent by the previous comparison of the timber floor with the concrete floor measured by Craik. The lightweight timber floor was seen to have lower part variability than the simple concrete floor and the variance curve shape is significantly different, although the data set we are observing is limited and in the case of the Parkin floor the part variability is inferred not calculated. In addition the comparisons rely on previous survey results by others, where the full data set has not been made available for assessment, and further work would be required in order to determine if this outcome is a component of the construction itself and not just a random effect.

13 Conclusion

13.1 Findings

The available literature highlights the problems with using GUM to calculate the components of measurement uncertainty, when testing airborne sound insulation in the field. The method described in BS5725 is also inadequate, being unable to allow any partitioning of the reproducibility variability.

Additional reviews of the available literature revealed a suitable approach developed primarily for the quality control process in manufacturing but specific to the evaluation of uncertainty in the measurement process. An advanced ANOVA design, coupled with a careful DOE, optimised the information gathered from field testing that was constrained by time.

Two experimental GRR were carried out: on a lightweight timber floor construction; on a heavyweight concrete floor. Preliminary analysis of both data sets was carried out, including on the source and receiver room loudspeaker position sound pressure levels. Each set was scrutinised for outliers and to ensure the data were representative and consistent with a field survey.

The uncertainty associated with single figure values $D_{nT,w}$ and $D_{nT,w} + C_{tr}$ for concrete floor GRR were most affected by the variability caused by the part which had a standard deviation (standard uncertainty) of 1.8dB for $D_{nT,w}$ and 0.9dB for $D_{nT,w} + C_{tr}$. This was expected as the room sizes used in the DOE were similar but not identical and the part to part variance was clearly dominant across the majority of the frequency range 100 – 3150Hz.. The instrument or measurement system was the next most dominant factor then the operator component, though both operator and instruments had standard uncertainties less than 1dB for the single figure values.

For the timber floor GRR the part was the least influential component of variance in the experiment. The was due to the room sizes being blocked intentionally in order to scrutinise the variability due to the construction. The dominant factor in the timber GRR was the operator and the standard uncertainty for this was three times greater for $D_{nT,w} + C_{tr}$ (2.1dB) than it was for $D_{nT,w}$ (0.7dB). The instrumentation also showed greater variability for the $D_{nT,w} + C_{tr}$ single figure value (1.2dB) as opposed to 0.6dB for $D_{nT,w}$.

The data for both floor constructions were analysed using the methods in BS5725 and a comparison made of the repeatability and reproducibility reference values in the International Standard.

The timber floor repeatability results were similar to the ISO 140-2 curve, up to 800Hz, above which the background noise caused the measured levels to increase and rise above the reference curve. A similar background noise effect was observed for reproducibility and calculated levels were above the reference curve at 1000Hz and above. The results for the concrete floor repeatability were higher than the ISO reference curve and peaked in the mid range at 800Hz due to the influence of background noise. The reproducibility levels were generally lower than the ISO 140-2 reference curve apart from the mid range of 630Hz-1000Hz and a peak at 800Hz due to background noise.

In general both surveys produced BS5725 results that were in line with expectations, based on the prevailing site conditions and exhibited increased levels of variability where the background noise influences were strongest.

Because the timber floor site provided a situation with similar room sizes a further comparison was made with the concrete floor results of Craik et al [57]. The same assessment procedure used by Craik et al was followed, which resulted in the calculation of the apparent sound reduction index between 50Hz – 5000Hz. Both surveys showed similar repeatability results, but in the reproducibility results, the timber floor was lower apart from above 2000Hz, when the background noise caused higher standard deviation in the timber floor. According to Craik et al the reproducibility was due to "workmanship" which suggests that either the workmanship on the timber floor was good leading to lower variability or the floor construction, although complex, does not have a significant bearing on the result.

The statistical analysis of the data was extended using ANOVA and followed a two way model with interaction. The ANOVA also covered an additional test sample which was introduced to extend the measurement range of the study. The additional test sample was added to the combined timber and concrete data set and a GRR assessment carried out.

The individual floor and combined data samples were tested and visually inspected for normality. Data that appeared to be non-normal was likely to be due to the relatively low sample numbers obtained, but the ANOVA was still used as it is relatively resistant to non normal data.

The presence of interaction was investigated and, when found to be significant, was included in the model. Where interaction was not significant, the reduced model was fitted and, in line with AIAG protocols, the variability due to the interaction term was added to the repeatability.

Interaction was observed in all GRR samples though it was especially evident in the concrete floor data.

The causes of interaction between the part and the operator are unclear but are likely driven by the operator and the choices they make. Some suggested possibilities centre on operators trying to work as efficiently as possible during the survey; economising on the effort expended and co-ordinating their actions to take the minimum time between tests. This is likely to include the choice of test kit placement in the test room due to room geometry, power socket locations, windows, etc which may affect the operator choice and constrain their test method. In addition, it is possible that the operators who are in the rooms for all measurements, may prefer to point the sound level meter in one direction and stand behind it whilst operating the meter. As the test regime is highly regimented and repetitive it may also be possible that after a the operators 'fix' on a certain arrangement for the microphones and the loudspeaker when visiting the same or similar rooms, it could even be related to the operators being left or right handed, or which way a door opened into a room presenting two corners to the operator as they carried equipment into the test space.

The significance of discovering interaction reinforces the conclusion that for the measurement of sound insulation in the field, GUM is not a suitable method for determining measurement uncertainty because the independence of input variables assumption is violated. It also suggests that the results of uncertainty studies that have used computer models based on GUM should be viewed with caution, unless some allowance has been made in the model for the interaction value.

The results of the ANOVA were compared with the draft International Standard for measurement uncertainty in Building Acoustics, ISO 12999. The new standard employs new terminology to describe uncertainty which uses standard deviations instead of variance for the reference values. They are known as "standard uncertainties" and follow the GUM nomenclature. New research underpins their calculation and they cannot be compared directly with the previous reference values in ISO 140 without a mathematical correction. The differences, although noticeable, are relatively small. Small changes in the experimental design may affect the uncertainty of measurement and in order to deal with this ISO12999 also proposes new measurement situations to describe how the uncertainty may be affected and provides reference values for each. It is therefore important if assessing compliance to select the correct one for any comparison. For this study data situation “A” from ISO 12999 describes the reproducibility test scenario which provides the most appropriate comparison although it is more likely the more reliable repeatability comparison is from the maximum standard deviation values for repeatability given in Table 1 of ISO12999 which better represent a pooled variance from many measurement systems. It is noted that the single number sound insulation values for repeatability corresponding to the maximum standard third octave band data from its Table 1 are missing. It is understood from the authors of the document that the research into measurement uncertainty and in particular the correlation effects associated with the frequency data and prediction of single figure values is still unfinished and as ISO 12999 is still in draft it is planned to update the information when it is available.

The GRR data is comparable with the currently published repeatability and reproducibility values given in ISO 12999 which reinforces the view that the data is representative of what would be expected from a test situation on site.

To supplement the GRR an additional test of significance, the Latin-square, was undertaken on the concrete floor site, to determine the dominant factor in the variability of sound insulation. As a reduced factorial experiment a combination of test kit, operator and floor are constructed in a matrix which allows the significance of the study, the Latin-square was repeated on each of the three days on site with different combinations of factors each time. The results show the test kits were not significant at any frequency; the operators were only significant in a few third octave bands and not for the single figure values.

For the frequency data, the simple Latin-square analysis gives, with a high degree of confidence ($p<1\%$), that the sound insulation values recorded are reliant on the floor being tested. The operators show some difference between each other and appear to have some influence in selected frequency bands, but the test kit does not appear to feature as a significant factor at any frequency on any of the days and we can conclude that no significant differences exist that depend on the test kits used.

The Latin square experiment confirms the result discovered previously in the GRR ANOVA in Chapter 8 that the part to part variability was confirmed to be the major component of variance in the concrete floor GRR

The floor measured, was universally significant on all three days across all but two or three low frequency bands. It was significant for $D_{nT,w}$ at the 1% level but was not significant for $D_{nT,w}$ + C_{tr} on any of the three days. Less variable data make it more difficult to obtain a statistically significant result using the Latin Square experiment analysis technique

The spectrum adaptation term reduces the variability between rooms from 56 – 62dB $D_{nT,w}$ to 50 – 54dB $D_{nT,w}$ + C_{tr} and as a result, the ability for the latin square experiment to determine significance between factors falls. This is not an isolated conclusion because the result is repeated for the Latin Square test carried out on each of the three days. Using the single figure value $D_{nT,w}$ + C_{tr} as the measurand (Sound Insulation performance indicator) the test is less able to apportion significance at a reasonable confidence level against the presence of experimental error.

The Latin-square results consistently emphasized the importance of the variability of the part in the total variability in the concrete GRR sound insulation measurement process. If the opportunity had been presented to carry out the Latin Square on the timber floor there may have been a different result as the floors were chosen to be identical, including the rooms shape and size. In that case there may have been more significance attached to the operators or test kit rather than the floor being tested. As it is relatively fast, the Latin-square can be used as a method to identify the factors that influence test results and theDOE can be fine tuned using the Latin-square, rather than from the GRR.

In order to complete the work investigating measurement uncertainty the point estimates for the components of variance require confidence intervals.. Two methods, developed for ANOVA, were used to calculate the confidence intervals, using a spreadsheet design. The Satterthwaite and Modified Large Sample methods give similar results for the upper and lower limits of the confidence intervals. However based on the GRR DOE the resulting low degrees of freedom result in confidence intervals with relatively low levels of precision. In addition, due to the variance component requirement to be greater than zero, the confidence limits lower bounds can tend to zero. The upper limit is not bound by this constraint and is free to increase based on the statistical approximation.

With both floors' GRR data this leads to confidence limits for reproducibility, operator, gauge and part that are highly asymmetrical and given that sound insulation is measured on a logarithmic scale, the size of the interval, measured in decibels, renders them of little practical use. The options for improving these confidence limits are shown to be limited because they are largely dependent on increasing the operator numbers to improve confidence interval precision and reduce the range . This means GRR studies, which typically have few operatives either because large numbers of operators are not available, or because it substantially increases the size of the experiment must therefore be content with relatively imprecise estimates of for example, $\sigma^2_{reproducibility}$ or find an alternative method of approximation.

The wide confidence limits determined using Satterthwaite and MLS suggest that the software orientated computer simulation approach adopted by GCI, REML[93] and others [95] may be a more desirable method of calculation though this is beyond the scope of this study. Notwithstanding the shortcomings of the imprecise confidence limits produced for some of the variance components, the measurand and repeatability have confidence limits that are of practical use in sound insulation measurement and should accompany any uncertainty budget and be stated where uncertainty is defined.

13.2 Application of findings

The results of the study provide estimates for the major components of variance for two types of floor construction. The uncertainties include results for the single figure values $D_{nT,w}$ and $D_{nT,w} + C_{tr}$ and for the frequency range 100 – 3150Hz.One of the main aims of this research was to obtain quantitative estimates of the components of variance associated with airborne sound insulation testing in the field because the United Kingdoms Accreditation Service require an uncertainty evaluation for all measurements undertaken by certified laboratories. Based on the GRR results we are able to identify all significant components of uncertainty and make a reasonable estimation of their uncertainty contribution. If this information is recorded in the appropriate test certification it will ensure that reported results do not give a false impression of uncertainty for the testing of airborne sound insulation in the field.

13.3 Further work

The accumulation of reliable data in sound insulation measurement is time consuming and is usually constrained and influenced by the number of participants on hand, the availability of a suitable site and the impact of background noise.

It is noted that the findings of this study are limited because measurement uncertainty is only available for two types of separating floor construction and there will probably have to be more work investigating other constructions before the measurement uncertainty in airborne sound insulation can be reported with confidence.

It is beneficial to know that ANOVA and specifically the design of experiment associated with GRR provide a reliable means to separate out the components of variance efficiently. It is also advantageous to know the limitations of the GRR design and future work may incorporate modifications to the process to reduce the need for 3 repeat measurements and include more operators in the DOE to improve the number of degrees of freedom and also the precision and range of the confidence intervals of the variance components.

In addition it would be desirable to obtain a set of test data for a basic concrete floor construction in order to compare the results with the GRR concrete floor data and that of Parkin et al. Further samples of lightweight timber floors would also prove useful in accumulating a robust data set to promote comparison.

In addition the presence of interaction between the operator and part in the concrete GRR was unexpected and the reasons for it are unclear. Further work, to identify the reasons for this, would be desirable and would require a carefully designed experiment to investigate the influence of constraining the loudspeaker and microphone positions in a room.

As previously discussed it is also important to address the confidence limits on the variance components and in this respect further work could be on using the alternative approaches e.g. GCI proposed by Weerahandi [130] or REML see Bartlett [131] which require specialist software modelling tools.

There are also likely to be additional components of variance associated with the reproducibility component which can be isolated and their influence determined using advanced ANOVA. Three and four way ANOVAs are common in statistical analysis and although the influence of interaction terms reduces as the number of factors involved in its calculation increases, more could be made of the DOE to optimise a GRR study. The identification and quantification of the most influential components of variance e.g. types of part variation by construction, would mean that estimates of uncertainty for individual test situations could be obtained using simple summation in quadrature. For example; estimating the expected uncertainty in measurement if all participants used the same test kit in an experiment rather than use different ones, or if they measured floors of the same construction but not all participants measured the same floors.

This thesis research was on airborne sound insulation. A complementary study, using GRR for impact sound insulation, could also be carried out with the relevant extension to the study as mentioned above.

14 Appendices

14.1 Appendix 1 – GRR Data

During the survey electronic measurements were stored in the sound level meters memory and downloaded for each of the testers at the end of each day

For each floor surveyed the measurements made were recorded on a test sheet that was filled in on site, see Figure 6-4 and the unique measurement number noted down for each sound pressure level measurement and each reverberation time measurement.

This was transferred to an electronic record sheet in a spreadsheet like the one in Table 14-1:

Table 14-1: Electronic record sheet noting down the unique run numbers for the stored electronic record

Gage R&R Test Reproducibility Sheet						
Tester	WAW			Day (1 2 or 3)		3
Room Details - Room Combination No	2					
Source Room	27					
Receive Room	26					
Date	01/05/2009					
Time Test on Room Started	09:18					
BG Receive Room Run No	118					
Rev Times: 2 rev times in each of 3 mic positions						
Airborne Test						
	1	2	3	4	5	
Source 1 Room Run No	98	99	100	101	102	
Source 2 Room Run No	108	109	110	111	112	
	1	2	3	4	5	
Receive 1 Room Run No	103	104	105	106	107	
Receive 2 Room Run No	113	114	115	116	117	
	1	2	3	4	5	6
Rev Times	119	120	121	122	123	124

NB: In all 90 electronic record sheets were created for each of the timber and concrete GRR.

The run numbers were collated in one summary sheet which formed a coded resource to analyse the raw data. A small section of this sheet is shown below:

Table 14-2: Selected section of coded record sheet showing test room scenario (1-6) operator initials, day, background noise reading number and the 5 positional numbers for each loudspeaker measurement in the source room and in the receive room and all the rev time numbers.

					Source 1					Source 2					Receive 1					Receive 2					Reverberation Times					
Test Scenario	Engineer	Day	~Row	BG	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5	RT 1	RT 2	RT 3	RT 4	RT 5	RT 6
				2	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	7
1	SM	1	3	1	15	16	17	18	19	25	26	27	28	29	20	21	22	23	24	30	31	32	33	34	3	4	5	6	7	8
1	SM	2	34	1	14	15	16	17	18	25	26	27	28	29	19	20	21	23	24	30	31	32	33	34	2	3	4	5	6	7
1	SM	3	65	22	49	50	51	52	53	59	60	61	62	63	54	55	56	57	58	64	65	66	67	68	23	24	25	26	27	28
1	SC	1	96	1	14	15	16	17	18	24	25	26	27	28	19	20	21	22	23	29	30	31	32	33	2	3	4	5	6	7
1	SC	2	127	2	15	16	17	18	19	25	26	27	28	29	20	21	22	23	24	30	31	32	33	34	3	4	5	6	7	8
1	SC	3	158	1	27	28	29	30	31	38	39	40	41	42	32	33	34	35	37	43	44	45	46	47	2	3	4	5	6	7
1	SP	1	189	133	134	135	136	137	138	144	145	146	147	148	139	140	141	142	143	149	150	151	152	153	187	188	189	190	191	192
1	SP	2	220	289	263	264	265	266	267	273	274	275	276	277	268	269	270	271	272	278	279	280	281	282	318	319	320	321	322	323
1	SP	3	251	93	67	68	69	70	71	77	78	79	80	81	72	73	74	75	76	82	83	84	85	86	121	122	123	124	125	126
1	DJ	1	282	166	140	141	142	143	144	150	151	152	153	154	145	146	147	148	149	155	156	157	158	159	200	201	202	203	204	205
1	DJ	2	313	277	278	279	280	281	282	288	289	290	291	292	283	284	285	286	287	293	294	295	296	297	337	338	339	340	341	342
1	DJ	3	344	67	68	69	70	71	72	78	79	80	81	82	73	74	75	76	77	83	84	85	86	87	128	129	130	131	132	133
1	BW	2	375	655	591	592	593	594	595	601	602	603	604	605	596	597	598	599	600	606	607	608	609	610	631	632	633	634	635	636
1	BW	2	406	302	269	270	271	272	273	279	280	281	282	283	274	275	276	277	278	284	285	286	287	288	289	290	291	292	293	294
1	BW	3	437	97	77	78	79	80	81	87	88	89	90	91	82	83	84	85	86	92	93	94	95	96	125	126	127	128	129	130
2	SM	1	3	35	43	44	45	46	47	53	54	55	56	57	48	49	50	51	52	58	60	61	62	63	64	65	66	67	68	69
2	SM	2	34	40	35	36	37	38	39	46	47	48	49	50	41	42	43	44	45	51	52	53	54	55	62	63	64	65	66	67
2	SM	3	65	1	29	30	31	32	33	39	40	41	42	43	34	35	36	37	38	44	45	46	47	48	14	15	16	17	18	19
2	SC	1	96	34	41	42	43	44	45	52	53	54	55	56	46	47	48	49	50	57	58	59	60	61	62	63	64	65	66	67
2	SC	2	127	35	36	37	38	39	40	46	47	48	49	50	41	42	43	44	45	51	52	53	54	55	62	63	64	65	66	67
2	SC	3	158	14	48	49	50	51	52	58	59	60	61	62	53	54	55	56	57	63	64	65	66	67	15	16	17	18	19	20
2	SP	1	189	160	161	162	163	164	165	171	172	173	174	175	166	167	168	169	170	176	177	178	179	180	193	194	195	196	197	198
2	SP	2	220	290	291	292	293	294	295	301	302	303	304	305	296	297	298	299	300	306	307	308	309	310	324	325	326	327	328	329
2	SP	3	251	94	95	96	97	98	99	105	106	107	108	109	100	101	102	103	104	110	111	112	113	114	127	128	129	130	131	132
2	DJ	1	282	167	168	169	170	171	172	178	179	180	181	182	173	174	175	176	177	183	184	185	186	187	194	195	196	197	198	199
2	DJ	2	313	304	305	306	307	308	309	315	316	317	318	319	310	311	312	313	314	320	321	322	323	324	331	332	333	334	335	336
2	DJ	3	344	94	95	96	97	98	99	104	105	106	107	108	100	101	102	103	115	110	111	112	113	114	122	123	124	125	126	127
2	BW	2	375	656	611	612	613	614	615	621	622	623	624	625	616	617	618	619	620	626	627	628	629	630	637	638	639	640	641	642
2	BW	2	406	301	249	250	251	252	253	259	260	261	262	263	254	255	256	257	258	264	265	266	267	268	295	296	297	298	299	300
2	BW	3	437	118	98	99	100	101	102	108	109	110	111	112	103	104	105	106	107	113	114	115	116	117	119	120	121	122	123	124
3	SM	1	3	145	184	185	186	187	188	194	195	196	197	198	189	190	191	192	193	199	200	201	202	203	158	159	160	161	162	163
3	SM	2	34	169	176	177	178	179	180	186	187	188	189	190	181	182	183	184	185	191	192	193	194	195	170	171	172	173	174	175
3	SM	3	65	122	101	102	103	104	105	111	112	113	114	115	106	107	108	109	110	116	117	118	119	120	129	130	131	132	133	134
3	SC	1	96	134	141	142	143	144	145	151	152	153	154	155	146	147	148	149	150	156	157	158	159	160	135	136	137	138	139	140
3	SC	2	127	171	173	174	175	176	177	183	184	185	186	187	178	179	180	181	182	188	189	190	191	192	200	201	202	203	204	205
3	SC	3	158	68	82	83	84	85	86	92	93	94	95	96	87	88	89	90	91	97	98	99	100	101	69	70	71	72	73	74
3	SP	1	189	67	68	69	70	71	72	78	79	80	81	82	73	74	75	76	77	83	84	85	86	87	121	122	123	124	125	126
3	SP	2	220	1	2	3	4	5	6	12	13	14	15	16	7	8	9	10	11	17	18	19	20	21	54	55	56	57	58	59
3	SP	3	251	133	134	135	136	137	138	144	145	146	147	148	139	140	141	142	143	149	150	151	152	153	193	194	195	196	197	198
3	DJ	1	282	98	70	71	72	73	74	80	81	82	83	84	75	76	77	78	79	85	87	88	89	90	132	133	134	135	136	137
3	DJ	2	313	1	2	3	4	5	6	16	17	18	19	20	7	8	9	10	11	21	22	23	24	25	69	70	71	72	73	74
3	DJ	3	344	135	136	137	138	139	140	147	148	149	150	151	142	143	144	145	146	152	153	154	155	156	196	197	198	199	200	201
3	BW	2	375	490	448	449	450	451	452	458	459	460	461	462	453	454	455	456	457	463	464	465	466	467	231	232	233	234	235	236
3	BW	2	406	757	178	179	180	181	182	188	189	190	191	192	183	184	185	186	187	193	194	195	196	197	231	232	233	234	235	236
3	BW	3	437	183	163	164	165	166	167	173	174	175	176	177	168	169	170	171	172	178	179	180	181	182	184	185	186	187	188	189

The raw data was stored in a large spreadsheet 1156 columns wide and 93 rows deep. All third octave bands from 6.3Hz to 20KHz were stored for the sound pressure level measurements and from 50Hz – 10KHz for the rev times.

A small section of background noise measurements for 1 room test scenario is shown Table 14-3:

Table 14-3: 1 section of background measurement results for test scenario 1

Background																																			
6.3 Hz	8.0 Hz	10 Hz	12.5 Hz	16 Hz	20 Hz	25 Hz	31.5 Hz	40 Hz	50 Hz	63 Hz	80 Hz	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 kHz	.25 kH	1.6 kHz	2.0 kHz	2.5 kHz	.15 kH	4.0 kHz	5.0 kHz	6.3 kHz	8.0 kHz	0.0 kH	.2.5 kH	.6.0 kH	0.0 kH
7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
34.4	33.3	29.9	30.5	35.9	38.7	28.8	33.1	38.1	25.9	31.3	19.4	19.6	20.6	26.2	20.1	16.3	16.1	16	15.8	14	15.4	14.8	12.7	15.1	14	10.5	10.9	11.6	10.6	9.2	8.6	7.9	7.4	6.6	6.7
41.8	37.1	34.4	33.7	36.2	41.6	31.9	28.5	30.8	26.6	23.4	26.7	23.1	18.6	19.3	17	14.5	16.3	16.1	17.8	16.1	15.2	16	13.9	12.1	9.7	9.8	9.7	11.1	12.2	9.1	9.7	9.7	10.8	6.7	8.5
34.9	34.3	32.3	35.5	39.1	38.4	30.9	41.8	41.8	27.5	23.2	27.5	29.4	23.8	20.5	17.6	16.6	18.4	17.8	17.6	16.5	18.6	18.5	15	12.4	12	11.5	11.9	12.7	12.4	10	10	8.5	7.2	6.6	6
34.2	38.4	45.7	33.5	34.7	35.3	30.7	37.2	38	28.8	34.8	23.6	29.1	30.6	25.3	25.2	21.2	20.3	20.6	19.4	20.2	22.6	22.2	20.6	21.9	24.2	20.7	20.7	19.5	18.6	14.7	11.7	10.6	9	9	7.7
42.3	38.4	39.3	39.9	42.4	42.4	37	39.5	41	31.9	28.8	27.6	28	23.8	29.4	25.5	19.1	25.2	25	23.4	23.6	21.5	21.6	20	19.1	19.7	20.4	20.6	19.8	17.9	14.3	12	10.9	9.4	9.2	8.8
43.5	35.6	37.7	38.3	39	37.8	37.9	38.2	40.7	28.7	27	22.4	19.5	19.2	24.1	23.1	21.6	22.4	21.2	21.6	20.9	20.6	20.7	20.2	19.4	19	17.9	17.8	17	16.5	13.6	12.1	10.5	8.4	13	8.1
49.9	47.8	44.8	41.2	46	47.7	40.7	39.6	31.6	28.3	28	21.2	24.2	27.4	31.6	28.3	20.8	19.9	17.8	18.9	17.8	17.6	18	15.6	13	12.8	13.3	12.2	10.9	8.9	7.5	8	7.8	7.2	7	6.8
62.1	61.4	48.8	43	43.4	42.5	37.8	33.6	35.3	30.6	32	26.3	26.4	31.1	30.8	29.2	28.5	28	26.4	25.7	23.2	23.7	22.2	21.4	21.5	22.4	21.5	20.4	18.2	15.5	12.5	11.1	10.7	9.2	7.3	7.1
22.7	24.3	22	23.3	23.6	25.5	26.3	25.5	27.2	26.9	29.9	36.6	30.2	29.7	39.1	34.2	29.6	31	29.8	30.2	30.1	29.9	29.9	30.7	31.9	31.3	31.2	31.1	31	31.8	29.2	28.4	27.9	26.7	25.8	24.9
63.8	61.5	58.1	54.9	58.5	56.7	57.2	51.8	55.9	37.3	24.8	30.6	31.6	30.5	31.7	25.5	25.6	22.9	23.2	21	21.6	21.7	21.5	19.3	18	17.4	16.3	15.4	12	10	8.9	8.2	7.9	7.4	6.8	6.3
53.2	56.8	52.7	46	45.8	51.2	47.1	41.2	47.9	39.9	39.8	30	25.4	23.4	35.9	28.1	23.2	19.2	21.2	20.3	18.6	18.5	17.3	19.4	19.7	18.3	15.8	13.9	14.1	9.5	5.7	4.7	4.7	4.1	3.4	3.1
49	57.2	54.2	50.6	52.3	49.9	39.4	41.6	45.5	31.2	30.8	36.6	31.3	27.9	39.4	33.6	26	29.1	27.7	24.7	22.9	22.3	20.3	17.3	16.5	14.3	13.9	13.7	15.6	12.4	9.2	8.7	7.5	6.1	4.5	4.2
53.5	56.1	51.7	49.4	57.6	58.2	52.8	47	52.8	40.5	36.3	33.6	27.9	29.6	34.3	27.2	23.8	21.5	21	19.6	18.9	18.4	17.7	15.6	12.7	10.8	11	9.9	8.4	8.5	7	6.9	7.2	7.1	7.1	7.3
51.5	41.8	32.2	31	34.9	34.6	30.4	34.3	32.5	20.3	22.9	23.7	22.5	17.4	23.7	21.9	15	14.5	16.2	14.6	15.5	16.3	15.7	12.8	10.3	8.7	10.1	8.7	8.1	8.4	7.4	7.1	7.2	7.2	7.1	7.2
44.3	40.6	38.4	38.4	40.3	49.8	33.8	36.9	49.8	43.7	28.5	35	26.9	28	31.2	25.2	21.5	23	24	20.8	23.2	23.1	21.3	17.9	14.2	11.9	10.3	10.1	10	9.1	7.1	7	7.3	7.2	7.1	7.7

The results of the timber GRR tests are shown in Table 14-4 .

All this data is analysed in Minitab and the program output for the timber floor GRR is detailed in section 14.2.

Table 14-4: Timber GRR Results Single figure vales, C_{tr} and D_{nT} for all frequency bands.

			Results			DnT															
Test Scenario	Engineer	Day	DnTw	Ctr	DnTw +Ctr	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz	400 Hz	500 Hz	630 Hz	800 Hz	1.0 k	1.25 k	1.6 k	2.0 k	2.5 k	3.15 k
1	SM	1	64	-8	56	38.8	46.4	45.7	48.2	53.6	57.1	60.1	63.2	65.7	66.8	68	72.8	75	72.4	69.6	78.5
1	SM	2	62	-9	53	35	43.1	44.4	46.8	52.6	55.3	59.4	63.4	64.1	65.1	67.4	72.6	74.8	71.5	67.9	78
1	SM	3	63	-8	55	37.6	44.6	44.5	47.9	52.5	55.4	60	63.4	65	65.8	67.8	72.2	74.8	72.1	68.7	78.2
1	SC	1	64	-9	55	38.1	45.5	46	50	54.4	57.2	59.5	62.8	64.4	66.2	66.8	71.3	73.4	71.4	68.1	74.7
1	SC	2	63	-8	55	37.6	43.9	46.7	49.4	53.4	56.4	59.5	61.8	64.2	65.7	65.9	69.8	69	67.9	66.2	70.2
1	SC	3	63	-7	56	39.4	44.1	46.6	50	54.4	57.4	60.3	62.5	64.2	65	66.3	68.8	68.4	66	64.3	66.2
1	SP	1	64	-10	54	36.1	44.1	45.5	52.3	56.3	58.4	60.7	64.1	66.4	64.7	66.4	72.9	74.5	70.7	68.2	76.8
1	SP	2	64	-9	55	38.1	43.6	45.4	49.1	53.9	57.5	60.9	64.2	65.1	65.7	67.4	72.3	74.1	70.9	69.3	78
1	SP	3	64	-8	56	40.8	44.5	46.1	47.7	53.9	58.6	60.3	64	66.6	65.7	68.6	74.6	76.1	72.7	69.7	78.4
1	DJ	1	63	-10	53	34	45.1	46.1	52.5	55.5	56.6	58.9	62.9	65	64.1	66	71.9	74.4	71.5	69.1	77.8
1	DJ	2	62	-10	52	34	44.1	44.1	47.6	51.9	56.5	59.1	64.1	64.6	65.8	66.8	72.2	75.3	71.8	68.6	77.8
1	DJ	3	63	-10	53	34.2	42.9	44.3	50.1	53.7	55.3	58.9	63.2	64.4	65.3	67	72.5	75	71.3	68.5	78.9
1	BW	2	63	-11	52	33.6	44.3	46.8	47.9	52.3	54.9	59.6	62.9	65	65.8	67	72.6	74.5	71.2	69.1	78.2
1	BW	2	62	-11	51	31.6	42.2	45	48.7	52.6	55.9	58.9	62.9	64.6	64.6	65.2	71.6	74.6	70.8	68.6	78
1	BW	3	61	-12	49	29.4	40.1	44.8	47.9	52.5	57	59.6	62.4	64.3	65.5	66	71.7	73.5	71.4	68.8	77.2
2	SM	1	63	-9	54	36.4	47	46.6	50.1	51.1	55.5	57.9	61.5	64.6	65.4	65.5	72.1	73.2	69.3	65.2	73.3
2	SM	2	62	-9	53	34.6	41.3	44.8	51	52	54.8	58.1	61.8	64	63.4	63.7	71.3	72.8	68.4	64.4	72.6
2	SM	3	63	-9	54	37.1	44.6	45.6	51.8	51.1	54.1	58.2	62.1	64.3	64.4	66.3	71.9	73	70.2	65.2	74.1
2	SC	1	62	-8	54	35.3	45.3	46.9	50.4	52.6	56	58	61.4	63.6	64.4	63.5	68.5	69	66.9	63.8	70.4
2	SC	2	61	-8	53	35.3	41.6	45.7	50.4	51.1	54.7	57.4	62.2	62.8	62.9	63.1	67.7	69.3	66.6	63.3	69.3
2	SC	3	62	-8	54	36.7	42.2	45.6	50.3	51.7	56.3	57.9	61.8	62.1	63.1	63.2	69.4	70.6	67.5	64.5	70.5
2	SP	1	63	-9	54	35	46.5	47.3	50.2	55.6	57.9	59.8	61.9	64.2	64	64.8	71.2	73.2	67.3	63.5	72.6
2	SP	2	63	-8	55	37.4	46.4	46	49.7	52.7	56	58.2	61.9	64.1	63.9	64.4	71.3	73.3	68.9	65.1	72.9
2	SP	3	63	-7	56	39.1	46.1	45.8	51	53.4	56.3	59.6	63.2	64.9	63.7	65	70.6	73.4	68.8	63.7	72
2	DJ	1	62	-10	52	32.8	46	47.9	53.8	52.6	54.3	57.8	60.5	62.5	63.8	65	71.4	73.6	68.1	64	72.5
2	DJ	2	62	-10	52	33.3	44.9	45.7	52	52.5	55.1	57.9	61	63.6	63.6	64.3	71.3	73.6	69.8	65.8	73.2
2	DJ	3	63	-8	55	37.3	45.2	45.8	51.1	53.7	56.2	58.1	62.2	62.8	63.7	64.8	70.4	72.9	69.1	65	72.7
2	BW	2	62	-12	50	30.3	42	46.8	49.1	52.2	55.7	58.5	61.9	64.5	63.7	63.6	70	73.2	69.3	65.4	72.7
2	BW	2	61	-10	51	32.8	41.9	47.4	49.9	50.2	54.5	57.8	61.2	62.9	63.9	63.9	70.6	72.2	68.7	65.6	72.6
2	BW	3	60	-12	48	28.7	39.3	44.1	49.6	51.3	54.9	56.8	60	63.7	63	64.8	70.5	72.3	68	65.6	72.9
3	SM	1	63	-10	53	34.3	43.6	45.4	50	52	56.1	59.9	61.9	64.1	65	67.1	68.9	73.5	72.1	71.7	74.4
3	SM	2	62	-9	53	34.3	43.3	44.8	50.1	52.8	55.8	59.8	61.1	63.9	64.3	66.4	68.6	72.5	70.2	68.9	71.7
3	SM	3	62	-9	53	36.3	41.7	42.3	47.4	50.9	55.9	60	61.5	62.9	64.7	66.6	69.1	72.3	70.1	68.7	72.7
3	SC	1	62	-8	54	35.9	47.1	44.1	48.9	53.4	57.5	61.3	62.5	62.8	64.7	61.8	64.7	67.3	64.7	62.7	67.2
3	SC	2	62	-10	52	33.5	46.8	44.4	48.1	53.1	57.3	59.4	61.1	62.9	64.4	66.8	67.6	67.4	65.2	62.9	66.3
3	SC	3	62	-7	55	37.7	45.9	45.2	49.8	52.6	55.8	59.4	61.4	62.2	63.7	65.5	66.6	67.3	66	62.9	67.6
3	SP	1	63	-9	54	36.4	44.2	45.9	52.7	55.3	57.9	60.9	62.6	63.8	66.8	62	68.5	72.8	68.6	68.5	72.5
3	SP	2	64	-8	56	39.6	45.5	46.3	49.4	54.9	57.3	60.4	61.4	62.9	64.7	65.9	68.6	72.2	70	68.3	72.1
3	SP	3	63	-7	56	39.1	44.3	45.9	49.9	53	56.8	60	62.7	63.5	64.2	66.8	68.7	70.8	69.3	67.8	70.2
3	DJ	1	62	-11	51	32.5	43.4	45.7	50.5	52.7	57	60.6	62.2	63.8	65.7	63.5	69.6	73.7	71.1	71.6	75.3
3	DJ	2	62	-10	52	33.7	43.8	45.5	50.9	52.2	55.7	59.5	61.4	62.3	64.6	66.1	69	71.8	70	68.5	71.3
3	DJ	3	63	-10	53	36.4	42	42.6	49.6	52.6	57.2	60.7	62.1	63.3	65	66.1	67.6	72	70.7	68.6	71.8
3	BW	2	62	-13	49	30	39	44.8	51.8	51.7	57.1	60	61.2	63	63.9	66.2	68	72	70	68.7	72.1
3	BW	2	61	-12	49	29.6	39.9	43.4	49.6	51.8	56.5	60.1	61.9	62.7	64.3	65.6	68.3	71.9	70.6	69.7	72.6
3	BW	3	62	-12	50	31.4	40.5	45.3	49.4	51.6	55.9	59.1	62	62.6	64.2	66.1	67.8	72	70	69.3	72.1
4	SM	1	64	-11	53	34.6	46.3	49	51	53.9	56.6	59.6	62.2	64.2	65.9	68.2	72	73.6	69.6	70	77.7
4	SM	2	63	-9	54	35.8	45.3	47.4	49.7	52.9	56	59.8	62.3	64	65	67	71.5	73.8	70.4	69.5	77.8
4	SM	3	62	-8	54	37.3	41.5	45.1	46.9	52.3	55.5	59.4	61.7	63.6	64.8	67.2	71.4	74.5	70.9	68.7	78.5
4	SC	1	63	-10	53	34.9	46.1	49.1	49.5	53	57.2	57.6	62	63	64.7	66.3	69.6	70.2	66.3	65.8	71.6
4	SC	2	62	-7	55	38.3	44.2	47.9	47.3	53.2	56.6	58.1	61.2	63.9	65.2	66.4	68.8	67.6	65.1	63.3	66.8
4	SC	3	63	-8	55	38.9	43.8	47	47.9	52.8	56.7	59.5	62.1	62.8	64.6	66.3	70.2	71.4	69.6	66.7	72
4	SP	1	64	-9	55	37.3	45.9	47	49.6	53	57.9	59.1	63	64.6	65.5	67.7	71.4	74.4	70.4	70	79.4
4	SP	2	63	-8	55	38.1	42.1	47	48.6	54	57.2	59.6	63.1	64.5	65.5	67.5	71.3	73.9	71	68.9	78.6
4	SP	3	64	-9	55	37.1	46.1	48.9	52.1	54.4	58.5	60.8	62.6	64	64.9	67.9	72	74.7	71.5	69.2	78.7
4	DJ	1	63	-12	51	31.9	44.5	47.4	50	53.7	56	59.4	62.5	63.7	64	65.9	71.3	74.6	70.8	70	79.6
4	DJ	2	63	-12	51	31.3	43.6	46.7	50.3	53.3	57	59.5	62.4	64.5	65.2	66.9	71.4	74.9	72.8	70.6	80.4
4	DJ	3	63	-12	51	31.3	44.7	47	51	54.6	56.9	60.1	61.6	64.3	65.2	67.8	72.3	75	72.1	68.9	79.9
4	BW	2	61	-14	47	27.1	39.1	46.1	47.9	53.1	56.9	59.2	61.7	63.3	65.1	65.8	69.8	71.1	68.9	67.6	71
4	BW	2	62	-12	50	31.7	40.2	47.9	48.2	51.6	56.1	58	62.1	64.3	64.7	66.9	70.5	73	70.6	68.3	78.2
4	BW	3	62	-12	50	31	41.1	45.9	50.5	52.1	55.5	58.9	61.4	62.6	64	66.2	69.2	71.6	69.7	68.9	74.9
5	SM	1	62	-9	53	34.6	43.6	45.3	48.9	52.4	55.9	59	60.5	62.4	63.4	63	68.7	73.8	69.2	66.9	69
5	SM	2	62	-7	55	40.2	44.1	44.9	46.9	52.1	55.1	60.3	62.1	63.5	64.8	63.1	70	74.7	70.2	66.8	69
5	SM	3	62	-8	54	38.2	41.6	45.6	48.2	52.5	55.8	59.7	61.7	63.1	63.8	62.3	68	73.9	69.5	65.8	69
5	SC	1	62	-9	53	35.4	45.6	46.9	46.9	53.7	57.1	60.1	62.3	62.2	63.5	63.5	67.4	67	65.8	62.9	65.8
5	SC	2	61	-7	54	37	45.4	46	47.5	53.4	56.6	58.7	60.6	61.9	63.7	62.2	62.8	63.1	62.6	60.9	63.8
5	SC	3	62	-6	56	40.3	45.1	45.5	47.5	53.3	56.9	59.1	62.3	63.4	63.3	63	65.8	67.7	65.2	62.2	65.6
5	SP	1	64	-9	55	37.2	44.1	47.5	51.3	57.2	58.7	60.7	63	63.1	65.2	63.8	69.5	73.5	69.4	67.1	71
5	SP	2	62	-8	54	37.9	41.2	44.3	46.3	52.7	56.4	59.7	62.1	63.2	63.8	64.1	69.4	74.8	70	66.6	69.7
5	SP	3	63	-9	54	36.2	43.9	45.7	48.9	52.9	57.3	60.4	63.1	63.8	64.6	64.4	69.9	74.5	70.1	66.7	69.8
5	DJ	1	62	-8	54	36.2	46.1	45.9	48.8	53.9	55.2	59.1	62.7	62.3	63	62.7	68.4	73.6	70.1	67.2	70.6
5	DJ	2	62	-8	54	36.8	44.3	45.8	48.3	55	55.9	59.5	62.9	63.5	63.7	62.2	68.1	73.5	69.7	65.6	68.7
5	DJ	3	62	-8	54	37.8	42.9	45.6	48.5	53	56	59.3	61.1	62.2	63.2	62.1	68.1	74	69.4	64.9	69.2
5	BW	2	61	-10	51	32.1	43.8	45.7	46.4	50.2	55.1	59.8	62	63.2	63.4	63	69.9	73.8	68.8	66	70
5	BW	2	61	-15	46	26	40.2	46.1	46	52.7	56.3	59.5	62.7	63	65	62.9	68.1	73.4	68.6	66.2	69.8
5	BW	3	61	-12	49	30	39.5	45.1	46.4	52.5	55.4	59.4	61.8	62.6	63.5	62.5	66.5	68.5	65.3	63.8	65.7
6	SM	1	63	-10	53	34.2	44.2	46.4	50.8	55.1	56.1	58.6	61.6	64.6	66.9	68.7	72.6	74.3	70.9	69.1	76.4
6	SM	2	64	-9	55	38	45	45	50.3	54.9	56.5	59.5	62.4	64.9	67.7	68.6	74.7	75.7	71.7	70.2	77.7
6	SM	3	63	-9	54	36	42.9	46.3	51.1	53.3	55.7	59.1	60.9	64	66.7	67.6	72.8	74	70.5	69.2	76.2
6	SC	1	63	-8	55	37.3	43	46.1	50.6	55.8	57.6	59.5	61.8	63.6	66.1	67.2	69.8	70	67.1	65.6	69.7
6	SC	2	63	-8	55	38.3	41.1	47	48.9	55.7	57.8	59.4	61.2	63.9	66.4	66.7	68.4	67.1	65	62.7	66.6
6	SC	3	64	-10	54	36.4	44.3	46.9	50.7	54.7	57.5	59.2	61.4	64.6	66.8	69.1	72.1	72.9	70.3	68.2	73.3
6	SP	1	64	-8	56	39.3	43.8	46.1	52	55.9	58.3	59.8	62.6	64.9	66.8	68.1	73	74.5	70.4	69.1	75.9
6	SP	2	65	-7	58	43.1	45.1	48	50.3	56.1	58.4	59.8	62.8	65.6	67.8	68.9	73.8	75.8	71.3	69.6	77.5
6	SP	3	65	-9	56	40.1	43.6	46	50.6	56.8	59	60	62.9	65.6	67.3	69.2	73.4	74.6	70.8	68.4	76.3
6	DJ	1	64	-9	55	36.6	44.6	46.1	52.5	56.1	55.7	59.1	62	63.3	66.7	68.1	73.5	75.1	71.5	69.3	78.4
6	DJ	2	64	-13	51	31.3	45.7	48.2	52.5	55	58.1	60.1	62.4	65	67.3	69	72.9	74.1	71.3	69.9	76.4
6	DJ	3	63	-7	56	39.2	43	45.5	49.7	54.9	56.6	58.8	62.1	64	67.3	68.8	72.7	74.5	70.7	68.8	76.3
6	BW	2	63	-11	52	33.6	41.7	45.5	50.8	53.7	57.9	60	61.8	62.8	66.8	67.7	72.7	74.4	70.6	69.3	77.1
6	BW	2	62	-12	50	31	40.6	46.3	51	53.2	57.1	58.8	61.4	64.2	67.1	68.3	72.9	74.3	71.3	69.7	76.5
6	BW	3	62	-11	51	31.6	41.8	45.7	51.7	54.6	57.6	59.5	60.6	62.7	66	67.2	70.9	72.3	70	68.3	75.4

14.2 Measurement System – Sound Insulation Test Kit:

Both Gauge repeatability and reproducibility experiments were carried out using 5 sets of test equipment. The make and model numbers are detailed below. Each of the 5 operators used their own test kit for the duration of the experiment. Each operator is UKAS trained and accredited, has at least three years experience in testing sound insulation in the field and has been independently scrutinised and their abilities verified by UKAS on at least one occasion during the annual Laboratory UKAS surveillance visit.

It should be noted that all test kits are controlled under a UKAS Accredited Laboratory and consequently the sound level meters are calibrated on a 2 year rotation and their calibrators are calibrated annually.

Measurements were made using the following equipment:

- Norsonic 140 or Norsonic 118 Sound Level Meter & windshield
- Norsonic 1251 Calibrator
- JBL EON 10 G2 Loudspeaker & in built amplifier
- Rane ME30B or DBX 131 – Graphic equaliser
- Wireless Transmitter & Receiver – Sennheiser SK100 G2
- Loudspeaker tripod and sound level meter tripod, power leads (6 – 8ft) power supply extension reel

The equipment has traceable UKAS calibration.

In addition the sound level meter was calibrated immediately prior to and immediately after the survey was carried out and the calibrated level notedon the test sheets by the operator an example of which is also detailed below:

The calibration was as follows:	
Before	114dB
After	114dB

14.3 Sound Insulation Measurement Procedure

The test procedure followed by each of the operators is detailed below. This covers the basic British and International standard requirements.

1. *Prior to attending site, a test kit should be allocated for the survey and the equipment should be checked to ensure that it is fully functional and within calibration. The test kit should compromise of the following;*

Quantity	*Item*	*Type*
1	*Sound Level Meter*	*Norsonic 118/140*
1	*Acoustic Calibrator*	*Norsonic 1251*
1	*Loudspeaker*	*JBL eON 10*
1	*Graphic Equaliser*	*Rane ME30B or DBX 131*
1	*Wireless Kit*	*Sennheiser SK 100 G2*
-	*Cables*	*Power Leads, Audio Leads.*

2. All relevant site plans and client specific requirements should be identified in the client folder for the project. The "Sound Insulation Test Record Sheet" QF/18-3 should be printed for site use.

3. Upon arrival on site, personnel must not enter site without the required protective clothing. Safety Helmet, Fluorescent Jacket and Safety boots should be worn at all times when on a building site. Ear defenders should be worn during source room measurements, and reverberation time measurements.

4. Upon arrival, report to the site managers office. If construction details have not been obtained request the information from the site manager. Also ensure the correct plot numbers are recorded.

5. With the site manager, where possible, conduct a visual survey of the surfaces to be tested, ensuring that doors and windows are closed/closable. Request the site managers assistance in ensuring workmen on site are quiet for the duration of the test. Ensure that an adequate power supply is available. Point out any areas which may compromise the sound insulation performance of the surface to be tested to the site manager. If, in your opinion, the surface is not testable due to compromising or unfinished details, advise the site manager.

6. When points 1-5 have been completed, record the agreed surfaces to be tested, together with construction details and observations on the Sound Insulation Test Record Sheet for the project.

7. If the rooms are of an unequal size, the source room must be the larger of the two rooms

8. The room dimensions should be measured and the room volumes should be recorded on the Sound Insulation Test Record Sheet.

9. Unpack the tripod and speaker, and set up in the identified source room in preparation for the first of two (where practicable) source side readings. The speaker should not be mounted close to the separating element to be assessed. The speaker should not be closer than 0.5M to any room boundary. Where two speaker positions are used the speaker positions should not be closer than 1.4m apart

10. *If a graphic equaliser is required for the 6dB adjacent bands (see below), connect the wireless kit receiver to the graphic eq input. Then connect the graphic eq output to the loudspeaker. If a graphic eq is not required, connect the wireless kit receiver to the speaker input.*

11. Unpack the Norsonic 118/*140* Sound Level Meter and connect the Seihnheiser wireless transmitter unit to the Norsonic 118 output. Turn on the Norsonic 118 SLM and wait three minutes to boot and adapt to site conditions.

12. Check that the windows and doors of the source room are closed.

13. Seat the Norsonic Acoustic Calibrator firmly onto the microphone. Turn on the calibrator and calibrate the instrument by pressing the “CAL” key, adjust the sensitivity accordingly using the “plus/minus” keys, press “Enter” to set the calibration. Calibrate the sound level meter according to it’s UKAS requirements and record both the calibration level and the sensitivity level on the Sound Insulation Test Record Sheet.

14. Check that the meter is set to record L_{eq} 6 second readings and store in the meters memory. Ensure that the internal pink noise generator is set at an appropriate gain for the test conditions.

15. Leaving the amplifier and speaker off, go to the designated receive room, ensure doors and windows are closed and take five 6 second background sound pressure level measurements at varied axis throughout the room recording the reference numbers on the Sound Insulation Test Record Sheet. Care should be taken to ensure the measuring distance is not within 0.5M of the room boundaries and that each subsequent reading should not be within 0.7M of any other microphone position where practicable. It is important to ensure the cooperation of any workmen/clients on site. Every effort should be made to ensure that the background is representative of what would be normal conditions for the building.

16. Return to the source room, ensure doors and windows are closed, ensure hearing protection is worn. Turn on the amplifier and, ensuring the microphone is not within 0.5 meters of the room boundaries or within 1m from the source, take a reading of the source room level (typically set at between 100 – 110dB linear @ 1m from the source), press the “Pause cont” button on the NOR118/*140* SLM and repeat this (a minimum of) 4 times at different representative positions in the room (Each subsequent reading should not be within 0.7M of any other microphone position where practicable. The positions used should be on different axes.) Until the meter has been running for a total time of 30 seconds. The meter will (energy) Log average each of the 5 x 6 second runs internally. Make a visual check of the tabulated 1/3rd octave data on the meter screen to ensure that no third octave band is greater than 6dB to any other adjacent band. Particular attention should be paid to the frequencies 100Hz, 125Hz and 160Hz in this respect. If there is a greater difference than 6dB in any adjacent third octave band the **reading should not be stored –**

adjust the 1/3rd octave band settings on the Graphic Equaliser in the frequencies which do not comply until a suitable spectrum is achieved or the maximum limits for adjusting the frequency spectrum using the graphic equaliser are reached - repeat the readings again and check the data prior to saving. If the data meets the requirements above, store the run and record the run number on the Sound Insulation Test Sheet.

17. In the receive room, remove hearing protection, turn on the pink noise generator and listen to the noise levels, recording any subjective observations as to possible, leakage through the surface tested or flanking sound issues on the Sound Insulation Test Record Sheet.

18. Bearing in mind the prescriptive minimum distances from room boundaries and other microphone positions, take a minimum of five readings pressing the "Pause cont" button each time. After a minimum of 5 readings have been taken, store the energy averaged level in the NOR118/*140* memory, recording the run number on the Sound Insulation Test Record Sheet.

19. Return to the source room; select a further speaker position, a minimum distance of 1.4m from the original position, either increase (or decrease) the loudspeaker height using the speaker tripod vertical extension to ensure that the two loudspeaker positions are not taken in the same axis, and repeat steps 16 – 18. NB The source equalisation should not be changed for the second speaker position.

20. Remove the equipment from the source room and set up in the receive room.

21. Set up the meter to record Reverberation Time Data. Position the Loudspeaker in the receive room taking care to ensure that the speaker is more than 0.5m from any of the room boundaries. Ensure that the microphone positions use to measure the reverberation time are set up to be greater than 2m from the source, 1.5m from each other and 1m from any room boundary where practicable (as detailed in ISO 354:2003 para 7.1.2). Take a Reverberation Time (RT) measurement. Check the tabulated data to ensure that no question mark symbols "?" appear after data within the measurement range. If a question mark symbol "?" appears, do not immediately store the data, check against the revised manufacturers "Recommended lower limit" for the relevant 1/3rd octave band (The revised table is kept with the NOR118/*140* instrument as a cross reference). If the "?" error indicator is valid repeat the measurement and if necessary change position. Check the tabulated data to ensure that where practicable no single third octave band RT measurement is greater than 1.5 times any adjacent band. Save the reverberation time record and note the run number on the Sound Insulation Test Record Sheet.

22. Repeat the RT measurement step 21 for a further minimum 5 runs, making 6 readings in total. Measurements should be taken using at least one speaker position, with two readings at three selected microphone positions.

23. If in the opinion of the test engineer room geometry or site layout dictate that BE EN ISO 140 – 4 does not cover the test situation, please refer to BN EN ISO 140-14 for advice on sound insulation testing procedure.

24. After the sound insulation survey is complete repeat stage 13 above and record the calibration level and sensitivity on the Sound Insulation Test Record Sheet.

25. As soon as is practicable, download the data from the meter to the appropriate project folder. Do not delete the data from the meter until the raw data has been verified and the $D_{nT,w}$ + C_{tr} calculated.

14.4 Appendix 2: Minitab Output – Timber Floor GRR

Gage R&R for D_{nTw}

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.2556	5.05111	30.1060	0.000
Engineer	4	36.7778	9.19444	54.8013	0.000
Test Scenari * Engineer	20	3.3556	0.16778	0.4576	0.973
Repeatability	60	22.0000	0.36667		
Total	89	87.3889			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.2556	5.05111	15.9369	0.000
Engineer	4	36.7778	9.19444	29.0096	0.000
Repeatability	80	25.3556	0.31694		
Total	89	87.3889			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.81014	71.96
Repeatability	0.31694	28.15
Reproducibility	0.49319	43.81
Engineer	0.49319	43.81
Part-To-Part	0.31561	28.04
Total Variation	1.12575	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.90008	5.40046	84.83	15.33
Repeatability	0.56298	3.37787	53.06	9.59
Reproducibility	0.70228	4.21367	66.19	11.96
Engineer	0.70228	4.21367	66.19	11.96
Part-To-Part	0.56179	3.37076	52.95	9.57
Total Variation	1.06101	6.36608	100.00	18.07

Number of Distinct Categories = 1

Gage R&R for $D_{nT,w}$

Gage R&R Study - ANOVA Method

Gage R&R for D_{nTw} +C_{tr}

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.022	5.0044	3.5774	0.018
Engineer	4	324.822	81.2056	58.0500	0.000
Test Scenari * Engineer	20	27.978	1.3989	0.9611	0.519
Repeatability	60	87.333	1.4556		
Total	89	465.156			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	25.022	5.0044	3.4720	0.007
Engineer	4	324.822	81.2056	56.3384	0.000
Repeatability	80	115.311	1.4414		
Total	89	465.156			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	5.87273	96.11
Repeatability	1.44139	23.59
Reproducibility	4.43134	72.52
Engineer	4.43134	72.52
Part-To-Part	0.23754	3.89
Total Variation	6.11027	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2.42337	14.5402	98.04	88.90
Repeatability	1.20058	7.2035	48.57	44.04
Reproducibility	2.10508	12.6305	85.16	77.22
Engineer	2.10508	12.6305	85.16	77.22
Part-To-Part	0.48738	2.9243	19.72	17.88
Total Variation	2.47190	14.8314	100.00	90.68

Number of Distinct Categories = 1

Gage R&R for $D_{nT,w}$ +C_{tr}

Gage R&R Study - ANOVA Method

```
* NOTE * The average measurement is not greater than the lower spec
limit,
         indicating the measurements are very far from the target. No
         %Tolerance is calculated.


Gage R&R for 100 Hz

Gage name:       Earl Shilton
Date of study:   27/4/2009
Reported by:     WAW
Tolerance:
Misc:            Lightweight Floors
```

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	45.876	9.175	2.3160	0.082
Engineer	4	627.132	156.783	39.5761	0.000
Test Scenari * Engineer	20	79.231	3.962	1.0798	0.393
Repeatability	60	220.127	3.669		
Total	89	972.365			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	45.876	9.175	2.4519	0.040
Engineer	4	627.132	156.783	41.8985	0.000
Repeatability	80	299.358	3.742		
Total	89	972.365			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	12.2442	97.13
Repeatability	3.7420	29.68
Reproducibility	8.5023	67.44
Engineer	8.5023	67.44
Part-To-Part	0.3622	2.87
Total Variation	12.6065	100.00

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)
Total Gage R&R	3.49918	20.9951	98.55
Repeatability	1.93442	11.6065	54.48
Reproducibility	2.91587	17.4952	82.12
Engineer	2.91587	17.4952	82.12
Part-To-Part	0.60184	3.6110	16.95
Total Variation	3.55056	21.3033	100.00

Number of Distinct Categories = 1

Gage R&R for 100 Hz

Gage R&R Study - ANOVA Method

* NOTE * The average measurement is not greater than the lower spec limit,
indicating the measurements are very far from the target. No %Tolerance is calculated.

Gage R&R for 125 Hz

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	5.878	1.1757	0.3356	0.885
Engineer	4	164.547	41.1368	11.7429	0.000
Test Scenari * Engineer	20	70.062	3.5031	1.7249	0.054
Repeatability	60	121.853	2.0309		
Total	89	362.341			

Alpha to remove interaction term = 0.25

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	4.61239	100.00
Repeatability	2.03089	44.03
Reproducibility	2.58150	55.97
Engineer	2.09076	45.33
Engineer*Test Scenari	0.49074	10.64
Part-To-Part	0.00000	0.00
Total Variation	4.61239	100.00

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)
Total Gage R&R	2.14765	12.8859	100.00
Repeatability	1.42509	8.5506	66.36
Reproducibility	1.60670	9.6402	74.81
Engineer	1.44595	8.6757	67.33
Engineer*Test Scenari	0.70053	4.2032	32.62
Part-To-Part	0.00000	0.0000	0.00
Total Variation	2.14765	12.8859	100.00

Number of Distinct Categories = 1

Gage R&R for 125 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 160 Hz

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	54.872	10.9743	23.6187	0.000
Engineer	4	10.264	2.5661	5.5227	0.004
Test Scenari * Engineer	20	9.293	0.4646	0.4328	0.980
Repeatability	60	64.413	1.0736		
Total	89	138.842			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	54.872	10.9743	11.9114	0.000
Engineer	4	10.264	2.5661	2.7852	0.032
Repeatability	80	73.706	0.9213		
Total	89	138.842			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	1.01270	60.18
Repeatability	0.92133	54.75
Reproducibility	0.09138	5.43
Engineer	0.09138	5.43
Part-To-Part	0.67020	39.82
Total Variation	1.68290	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	1.00633	6.03799	77.57	315.94
Repeatability	0.95986	5.75915	73.99	301.35
Reproducibility	0.30229	1.81372	23.30	94.90
Engineer	0.30229	1.81372	23.30	94.90
Part-To-Part	0.81866	4.91194	63.11	257.02
Total Variation	1.29727	7.78361	100.00	407.28

Number of Distinct Categories = 1

Gage R&R for 160 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 200 Hz

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	98.680	19.7359	13.7530	0.000
Engineer	4	31.812	7.9529	5.5420	0.004
Test Scenari * Engineer	20	28.700	1.4350	0.9028	0.585
Repeatability	60	95.367	1.5894		
Total	89	254.558			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	98.680	19.7359	12.7260	0.000
Engineer	4	31.812	7.9529	5.1281	0.001
Repeatability	80	124.067	1.5508		
Total	89	254.558			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	1.90651	61.13
Repeatability	1.55084	49.72
Reproducibility	0.35567	11.40
Engineer	0.35567	11.40
Part-To-Part	1.21234	38.87
Total Variation	3.11885	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	1.38076	8.2846	78.18	89.75
Repeatability	1.24533	7.4720	70.52	80.94
Reproducibility	0.59638	3.5783	33.77	38.76
Engineer	0.59638	3.5783	33.77	38.76
Part-To-Part	1.10106	6.6064	62.35	71.57
Total Variation	1.76603	10.5962	100.00	114.79

Number of Distinct Categories = 1

Gage R&R for 200 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 250 Hz

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	69.281	13.8562	33.7563	0.000
Engineer	4	60.264	15.0659	36.7034	0.000
Test Scenari * Engineer	20	8.210	0.4105	0.4500	0.975
Repeatability	60	54.727	0.9121		
Total	89	192.481			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	69.281	13.8562	17.6130	0.000
Engineer	4	60.264	15.0659	19.1507	0.000
Repeatability	80	62.936	0.7867		
Total	89	192.481			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	1.57999	64.46
Repeatability	0.78670	32.09
Reproducibility	0.79329	32.36
Engineer	0.79329	32.36
Part-To-Part	0.87130	35.54
Total Variation	2.45129	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	1.25698	7.54187	80.28	45.31
Repeatability	0.88696	5.32178	56.65	31.97
Reproducibility	0.89067	5.34401	56.89	32.10
Engineer	0.89067	5.34401	56.89	32.10
Part-To-Part	0.93343	5.60061	59.62	33.64
Total Variation	1.56566	9.39396	100.00	56.43

Number of Distinct Categories = 1

Gage R&R for 250 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 315 Hz

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	27.869	5.5738	22.2605	0.000
Engineer	4	42.056	10.5141	41.9909	0.000
Test Scenari * Engineer	20	5.008	0.2504	0.5079	0.953
Repeatability	60	29.580	0.4930		
Total	89	104.513			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	27.869	5.5738	12.8919	0.000
Engineer	4	42.056	10.5141	24.3185	0.000
Repeatability	80	34.588	0.4323		
Total	89	104.513			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.99244	74.33
Repeatability	0.43235	32.38
Reproducibility	0.56009	41.95
Engineer	0.56009	41.95
Part-To-Part	0.34276	25.67
Total Variation	1.33520	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.99621	5.97728	86.21	25.97
Repeatability	0.65753	3.94519	56.90	17.14
Reproducibility	0.74839	4.49037	64.77	19.51
Engineer	0.74839	4.49037	64.77	19.51
Part-To-Part	0.58546	3.51275	50.67	15.26
Total Variation	1.15551	6.93306	100.00	30.12

Number of Distinct Categories = 1

Gage R&R for 315 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 400 Hz

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	33.3459	6.66918	19.4053	0.000
Engineer	4	11.0771	2.76928	8.0578	0.000
Test Scenari * Engineer	20	6.8736	0.34368	1.1338	0.342
Repeatability	60	18.1867	0.30311		
Total	89	69.4832			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	33.3459	6.66918	21.2901	0.000
Engineer	4	11.0771	2.76928	8.8404	0.000
Repeatability	80	25.0602	0.31325		
Total	89	69.4832			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.449699	51.49
Repeatability	0.313253	35.86
Reproducibility	0.136446	15.62
Engineer	0.136446	15.62
Part-To-Part	0.423728	48.51
Total Variation	0.873427	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.670596	4.02357	71.75	14.00
Repeatability	0.559690	3.35814	59.89	11.69
Reproducibility	0.369386	2.21631	39.52	7.71
Engineer	0.369386	2.21631	39.52	7.71
Part-To-Part	0.650944	3.90567	69.65	13.59
Total Variation	0.934573	5.60744	100.00	19.52

Number of Distinct Categories = 1

Gage R&R for 400 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 500 Hz

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	23.1219	4.62438	16.6351	0.000
Engineer	4	13.9882	3.49706	12.5798	0.000
Test Scenari * Engineer	20	5.5598	0.27799	0.8421	0.655
Repeatability	60	19.8067	0.33011		
Total	89	62.4766			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	23.1219	4.62438	14.5842	0.000
Engineer	4	13.9882	3.49706	11.0289	0.000
Repeatability	80	25.3664	0.31708		
Total	89	62.4766			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.493746	63.23
Repeatability	0.317081	40.60
Reproducibility	0.176665	22.62
Engineer	0.176665	22.62
Part-To-Part	0.287153	36.77
Total Variation	0.780899	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.702671	4.21602	79.52	12.32
Repeatability	0.563099	3.37859	63.72	9.87
Reproducibility	0.420316	2.52189	47.56	7.37
Engineer	0.420316	2.52189	47.56	7.37
Part-To-Part	0.535867	3.21520	60.64	9.39
Total Variation	0.883685	5.30211	100.00	15.49

Number of Distinct Categories = 1

Gage R&R for 500 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 630 Hz

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	41.0757	8.21513	22.1386	0.000
Engineer	4	17.9118	4.47794	12.0674	0.000
Test Scenari * Engineer	20	7.4216	0.37108	1.1188	0.356
Repeatability	60	19.9000	0.33167		
Total	89	86.3090			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	41.0757	8.21513	24.0547	0.000
Engineer	4	17.9118	4.47794	13.1118	0.000
Repeatability	80	27.3216	0.34152		
Total	89	86.3090			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.57132	52.12
Repeatability	0.34152	31.15
Reproducibility	0.22980	20.96
Engineer	0.22980	20.96
Part-To-Part	0.52491	47.88
Total Variation	1.09623	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.75586	4.53515	72.19	12.08
Repeatability	0.58440	3.50638	55.82	9.34
Reproducibility	0.47938	2.87626	45.79	7.66
Engineer	0.47938	2.87626	45.79	7.66
Part-To-Part	0.72451	4.34703	69.20	11.58
Total Variation	1.04701	6.28206	100.00	16.73

Number of Distinct Categories = 1

Gage R&R for 630 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 800 Hz

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	100.598	20.1196	72.7421	0.000
Engineer	4	5.683	1.4207	5.1366	0.005
Test Scenari * Engineer	20	5.532	0.2766	0.7163	0.794
Repeatability	60	23.167	0.3861		
Total	89	134.980			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	100.598	20.1196	56.0857	0.000
Engineer	4	5.683	1.4207	3.9604	0.006
Repeatability	80	28.698	0.3587		
Total	89	134.980			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.41773	24.07
Repeatability	0.35873	20.67
Reproducibility	0.05900	3.40
Engineer	0.05900	3.40
Part-To-Part	1.31739	75.93
Total Variation	1.73512	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.64632	3.87792	49.07	9.72
Repeatability	0.59894	3.59365	45.47	9.01
Reproducibility	0.24290	1.45739	18.44	3.65
Engineer	0.24290	1.45739	18.44	3.65
Part-To-Part	1.14778	6.88667	87.13	17.27
Total Variation	1.31724	7.90345	100.00	19.82

Number of Distinct Categories = 2

Gage R&R for 800 Hz

Gage R&R Study - ANOVA Method

Gage R&R for 1.0 k

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	272.596	54.5191	67.4927	0.000
Engineer	4	17.924	4.4811	5.5475	0.004
Test Scenari * Engineer	20	16.156	0.8078	0.9272	0.557
Repeatability	60	52.273	0.8712		
Total	89	358.949			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	272.596	54.5191	63.7381	0.000
Engineer	4	17.924	4.4811	5.2389	0.001
Repeatability	80	68.429	0.8554		
Total	89	358.949			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	1.05679	22.80
Repeatability	0.85536	18.46
Reproducibility	0.20143	4.35
Engineer	0.20143	4.35
Part-To-Part	3.57758	77.20
Total Variation	4.63437	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	1.02800	6.1680	47.75	14.82
Repeatability	0.92486	5.5491	42.96	13.33
Reproducibility	0.44881	2.6929	20.85	6.47
Engineer	0.44881	2.6929	20.85	6.47
Part-To-Part	1.89145	11.3487	87.86	27.27
Total Variation	2.15276	12.9166	100.00	31.03

Number of Distinct Categories = 2

Gage R&R for 1.0 k

Gage R&R Study - ANOVA Method

Gage R&R for 1.25 k

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	263.786	52.7573	101.970	0.000
Engineer	4	107.780	26.9451	52.080	0.000
Test Scenari * Engineer	20	10.348	0.5174	0.606	0.893
Repeatability	60	51.247	0.8541		
Total	89	433.161			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	263.786	52.7573	68.5224	0.000
Engineer	4	107.780	26.9451	34.9969	0.000
Repeatability	80	61.594	0.7699		
Total	89	433.161			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	2.22410	39.09
Repeatability	0.76993	13.53
Reproducibility	1.45418	25.56
Engineer	1.45418	25.56
Part-To-Part	3.46582	60.91
Total Variation	5.68993	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	1.49134	8.9481	62.52	17.67
Repeatability	0.87746	5.2647	36.79	10.40
Reproducibility	1.20589	7.2354	50.55	14.29
Engineer	1.20589	7.2354	50.55	14.29
Part-To-Part	1.86167	11.1700	78.05	22.05
Total Variation	2.38536	14.3121	100.00	28.26

Number of Distinct Categories = 1

Gage R&R for 1.25 k

Gage R&R Study - ANOVA Method

Gage R&R for 1.6 k

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	69.990	13.9980	7.6960	0.000
Engineer	4	354.792	88.6979	48.7652	0.000
Test Scenari * Engineer	20	36.378	1.8189	1.2076	0.280
Repeatability	60	90.373	1.5062		
Total	89	551.533			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	69.990	13.9980	8.8350	0.000
Engineer	4	354.792	88.6979	55.9825	0.000
Repeatability	80	126.751	1.5844		
Total	89	551.533			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	6.42403	88.59
Repeatability	1.58439	21.85
Reproducibility	4.83964	66.74
Engineer	4.83964	66.74
Part-To-Part	0.82758	11.41
Total Variation	7.25161	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2.53457	15.2074	94.12	27.48
Repeatability	1.25872	7.5523	46.74	13.65
Reproducibility	2.19992	13.1995	81.69	23.85
Engineer	2.19992	13.1995	81.69	23.85
Part-To-Part	0.90971	5.4583	33.78	9.86
Total Variation	2.69288	16.1573	100.00	29.20

Number of Distinct Categories = 1

Gage R&R for 1.6 k

Gage R&R Study - ANOVA Method

Gage R&R for 2.0 k

Gage name: Earl Shilton
Date of study: 27/4/2009
Reported by: WAW
Tolerance:
Misc: Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	81.346	16.2692	10.8617	0.000
Engineer	4	198.774	49.6934	33.1766	0.000
Test Scenari * Engineer	20	29.957	1.4978	1.1769	0.305
Repeatability	60	76.360	1.2727		
Total	89	386.437			

Alpha to remove interaction term = 0.25

Two-Way ANOVA Table Without Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	81.346	16.2692	12.2420	0.000
Engineer	4	198.774	49.6934	37.3927	0.000
Repeatability	80	106.317	1.3290		
Total	89	386.437			

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	4.01588	80.13
Repeatability	1.32896	26.52
Reproducibility	2.68692	53.61
Engineer	2.68692	53.61
Part-To-Part	0.99601	19.87
Total Variation	5.01189	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2.00397	12.0238	89.51	24.53
Repeatability	1.15281	6.9168	51.49	14.11
Reproducibility	1.63918	9.8351	73.22	20.06
Engineer	1.63918	9.8351	73.22	20.06
Part-To-Part	0.99801	5.9880	44.58	12.21
Total Variation	2.23873	13.4324	100.00	27.40

Number of Distinct Categories = 1

Gage R&R for 2.0 k

Gage R&R Study - ANOVA Method

Gage R&R for 2.5 k

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	223.550	44.7100	17.4683	0.000
Engineer	4	193.042	48.2606	18.8555	0.000
Test Scenari * Engineer	20	51.190	2.5595	2.3726	0.005
Repeatability	60	64.727	1.0788		
Total	89	532.509			

Alpha to remove interaction term = 0.25

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	4.11130	59.40
Repeatability	1.07878	15.59
Reproducibility	3.03252	43.81
Engineer	2.53895	36.68
Engineer*Test Scenari	0.49357	7.13
Part-To-Part	2.81004	40.60
Total Variation	6.92133	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	2.02763	12.1658	77.07	27.41
Repeatability	1.03864	6.2319	39.48	14.04
Reproducibility	1.74141	10.4485	66.19	23.54
Engineer	1.59341	9.5604	60.57	21.54
Engineer*Test Scenari	0.70255	4.2153	26.70	9.50
Part-To-Part	1.67632	10.0579	63.72	22.66
Total Variation	2.63084	15.7851	100.00	35.57

Number of Distinct Categories = 1

Gage R&R for 2.5 k

Gage R&R Study - ANOVA Method

Gage R&R for 3.15 k

Gage name:	Earl Shilton
Date of study:	27/4/2009
Reported by:	WAW
Tolerance:	
Misc:	Lightweight Floors

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Test Scenari	5	776.00	155.200	35.3917	0.000
Engineer	4	494.60	123.651	28.1973	0.000
Test Scenari * Engineer	20	87.70	4.385	1.7589	0.048
Repeatability	60	149.59	2.493		
Total	89	1507.89			

Alpha to remove interaction term = 0.25

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	9.7497	49.23
Repeatability	2.4931	12.59
Reproducibility	7.2566	36.64
Engineer	6.6259	33.46
Engineer*Test Scenari	0.6307	3.18
Part-To-Part	10.0543	50.77
Total Variation	19.8040	100.00

Lower process tolerance limit = 45

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	3.12245	18.7347	70.16	33.03
Repeatability	1.57896	9.4738	35.48	16.70
Reproducibility	2.69380	16.1628	60.53	28.49
Engineer	2.57407	15.4444	57.84	27.23
Engineer*Test Scenari	0.79416	4.7650	17.85	8.40
Part-To-Part	3.17085	19.0251	71.25	33.54
Total Variation	4.45016	26.7010	100.00	47.07

Number of Distinct Categories = 1

Gage R&R for 3.15 k

14.5 Appendix 3: balanced two-factor crossed random model with interaction

Table 14-5: ANOVA for model [2]

Source of variation	Degrees of freedom	Mean square	Expected Mean Square
Parts (P)	P – 1	S_P^2	$\theta_P = \sigma_E^2 + r\sigma_{PO}^2 + or\sigma_P^2$
Operators (O)	O – 1	S_O^2	$\theta_O = \sigma_E^2 + r\sigma_{PO}^2 + pr\sigma_O^2$
P x O	(p – 1)(o – 1)	S_{PO}^2	$\theta_{PO} = \sigma_E^2 + r\sigma_{PO}^2$
Replicates	po(r – 1)	S_E^2	$\theta_E = \sigma_E^2$

The Mean Squares and means for model 11-10 are detailed in Table 14-6:

Table 14-6: Mean squares and means for model [2]

Statistic	Definition
S_P^2	$\frac{or\Sigma_i\ (\bar{Y}_{i**} - \bar{Y}_{***})^2}{(p-1)}$
S_O^2	$\frac{pr\Sigma_j\ (\bar{Y}_{*j*} - \bar{Y}_{***})^2}{(o-1)}$
S_{PO}^2	$\frac{r\Sigma_i\ \Sigma_j\ (\bar{Y}_{ij*} - \bar{Y}_{i**} - \bar{Y}_{*j*} + \bar{Y}_{***})^2}{(p-1)(o-1)}$
S_E^2	$\frac{\Sigma_i\ \Sigma_j\ \Sigma_k\ (\bar{Y}_{ijk} - \bar{Y}_{ij*})^2}{po(r-1)}$
$\bar{Y}_{i**}$	$\frac{\Sigma_j\ \Sigma_k\ Y_{ijk}}{or}$
$\bar{Y}_{*j*}$	$\frac{\Sigma_i\ \Sigma_k\ Y_{ijk}}{pr}$
$\bar{Y}_{ij*}$	$\frac{\Sigma_k\ Y_{ijk}}{r}$
$\bar{Y}_{***}$	$\frac{\Sigma_i\ \Sigma_j\ \Sigma_k\ Y_{ijk}}{por}$

Distributional results for balanced two-factor crossed random model 11-10 are detailed in Table 14-7:

Table 14-7: Distributional results for model 11-10

Result	
1	$\bar{Y}_{***}$, S_P^2, S_O^2, S_{PO}^2 and S_E^2 are jointly independent
2	(p - 1)S_P^2 / θ_p is a chi squared random variable with p – 1 degrees of freedom
3	(o - 1)S_o^2 / θ_o is a chi squared random variable with o – 1 degrees of freedom
4	(p - 1)(o - 1)S_{po}^2 / θ_{po} is a chi squared random variable with (p – 1)(o – 1) degrees of freedom
5	po(r - 1)S_E^2 / θ_E is a chi squared random variable with po(r – 1) degrees of freedom
6	$\bar{Y}_{***}$ is a normal random variable with a mean μ_Y and variance $\frac{\theta_p + \theta_o - \theta_{po}}{por}$

The covariance structure for the balanced two-factor crossed random model in 11-10 are detailed in Table 14-8.

Table 14-8: Covariance structure for model in 11-10

Condition	Covariance (Y_{ijk}, $Y_{i'j'k'}$)
$i = i'$, $j = j'$, $k \neq k'$, (same part and same operator)	$\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2$
$i = i'$, $j \neq j'$ (same part with different operator)	σ_P^2
$i \neq i'$, $j = j'$ (same operator with different parts)	σ_O^2
$i \neq i'$, $j \neq j'$ (different parts and operators)	0

Variation in the measurement system is attributed to all sources except parts and is detailed in Table 14-9.

Table 14-9: GRR Parameters and point estimators for model in 11-10.

Gauge R&R Notation	Model Representation	Point Estimator
μ_Y	μ_Y	$\hat{\mu}_Y$
γ_P	σ_P^2	$\hat{\gamma}_P = \frac{S_P^2 - S_{PO}^2}{or}$
γ_O	σ_O^2	$\hat{\gamma}_O = \frac{S_O^2 - S_{PO}^2}{pr}$
γ_{PO}	σ_{PO}^2	$\hat{\gamma}_{PO} = \frac{S_{PO}^2 - S_E^2}{r}$
γ_M	$\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2$	$\hat{\gamma}_M = \frac{S_O^2 + (p-1)S_{PO}^2 + p(r-1)S_E^2}{pr}$
γ_R	$\frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}$	$\hat{\gamma}_R = \frac{\hat{\gamma}_P}{\hat{\gamma}_M}$

Where the definitions are detailed in Table 14-10.

Table 14-10: Definition of Parameters in a gauge R&R study

Symbol	Definition
μ_Y	Mean of population of measurements
γ_P	Variance of the part
γ_O	Variance of the operator
γ_{PO}	Variance (Interaction) between part & operator
γ_M	Variance of the measurement system

14.6 Appendix 4: MLS confidence limits

Table 14-11: Modified Large Sample constants.

Constant	Definition
G_1	$1 - F_{\alpha/2:\infty,p-1}$
G_2	$1 - F_{\alpha/2:\infty,o-1}$
G_3	$1 - F_{\alpha/2:\infty,(p-1)(o-1)}$
G_4	$1 - F_{\alpha/2:\infty,po(r-1)}$
H_1	$F_{1-\alpha/2:\infty,p-1} - 1$
H_2	$F_{1-\alpha/2:\infty,o-1} - 1$
H_3	$F_{1-\alpha/2:\infty,(p-1)(o-1)} - 1$
H_4	$F_{1-\alpha/2:\infty,po(r-1)} - 1$
F_1	$F_{1-\alpha/2:p-1,(p-1)(o-1)}$
F_2	$F_{\alpha/2:p-1,(p-1)(o-1)}$
F_3	$F_{1-\alpha/2:p-1,o-1}$
F_4	$F_{\alpha/2:p-1,o-1}$
F_5	$F_{1-\alpha/2:o-1,(p-1)(o-1)}$
F_6	$F_{\alpha/2:o-1,(p-1)(o-1)}$
F_7	$F_{1-\alpha/2:(p-1)(o-1),po(r-1)}$
F_8	$F_{\alpha/2:(p-1)(o-1),po(r-1)}$
G_{13}	$\frac{(F_1)^2 - G_1^2 F_1^2 - H_3^2}{F_1}$
G_{23}	$\frac{(F_5 - 1)^2 - G_2^2 F_5^2 - H_3^2}{F_5}$
G_{34}	$\frac{(F_7 - 1)^2 - G_3^2 F_7^2 - H_4^2}{F_7}$
H_{13}	$\frac{(1 - F_2)^2 - H_1^2 F_2^2 - G_3^2}{F_2}$
H_{23}	$\frac{(1 - F_6)^2 - H_2^2 F_6^2 - G_3^2}{F_6}$
H_{34}	$\frac{(1 - F_8)^2 - H_3^2 F_8^2 - G_4^2}{F_8}$

The following confidence intervals for μ_Y , γ_P , γ_O, γ_{PO}, and γ_M are described below.

14.6.1.1 Interval for μ_Y

The recommended definition [67] for the confidence interval for μ_Y is provided by Milliken & Johnson [141] for a 100(1 - α)%. The lower limit is:

$$L = \bar{Y}_{***} - C\sqrt{\frac{K}{por}} \quad \text{14-1:}$$

and the upper limit is:

$$U = \bar{Y}_{***} + C\sqrt{\frac{K}{por}} \qquad \text{14-2:}$$

where:

$$K = S_P^2 + S_O^2 - S_{PO}^2$$

and:

$$C = \frac{S_P^2\sqrt{F_{1-\alpha:1,p-1}} + S_O^2\sqrt{F_{1-\alpha:1,o-1}} - S_{PO}^2\sqrt{F_{1-\alpha:1,(p-1)(o-1)}}}{K}$$

NB: If K<0, then replace K with S_{PO}^2 and C with $\sqrt{F_{1-\alpha:1,(p-1)(o-1)}}$.

14.6.1.2 Interval for γ_P

The recommended definition for γ_P is provided by Ting et al [142]. The upper and lower limits for an approximate 100(1 - α)% confidence interval are [67]:

$$L = \hat{\gamma}_P - \frac{\sqrt{V_{LP}}}{op} \qquad \text{14-3:}$$

And the upper limit is:

$$U = \hat{\gamma}_P + \frac{\sqrt{V_{UP}}}{op} \qquad \text{14-4:}$$

Where:

$$V_{LP} = G_1^2 S_P^2 + H_3^2 S_{PO}^4 + G_{13}^2 S_P^2 S_{PO}^2$$

$$V_{UP} = H_1^2 S_P^4 + G_3^2 S_{PO}^4 + H_{13}^2 S_P^2 S_{PO}^2$$

In this case negative bounds are increased to zero.

14.6.1.3 Interval for γ_O

The variability of the operator is given by:

$$\gamma_O = \sigma_O^2 = \frac{[\theta_O - \theta_{PO}]}{pr} \qquad \text{14-5}$$

The recommended definition for Interval for γ_O is provided by Ting et al[142] and a direct substitution into the same formulae as γ_P eqn [9] &[10] the upper and lower limits for an approximate 100(1 - α)% confidence interval are:

The lower limit after Burdick et al[67] is:

$$L = \hat{\gamma}_O - \frac{\sqrt{V_{LO}}}{pr} \qquad \text{14-6:}$$

And the upper limit is:

$$U = \hat{\gamma}_O + \frac{\sqrt{V_{UO}}}{pr} \qquad \text{14-7:}$$

Where:

$$V_{LO} = G_2^2 S_O^2 + H_3^2 S_{PO}^4 + G_{23} S_O^2 S_{PO}^2$$

$$V_{UO} = H_2^2 S_O^4 + G_3^2 S_{PO}^4 + H_{23} S_O^2 S_{PO}^2$$

In this case negative bounds are increased to zero.

14.6.1.4 Interval for γ_{PO}

The upper and lower limits for an approximate 100(1 - α)% confidence interval for γ_{PO} are:

The lower limit after Burdick et al[67]is:

$$L = \hat{\gamma}_{PO} - \frac{\sqrt{V_{LPO}}}{r}$$ 14-8:

The upper limit is:

$$U = \hat{\gamma}_{PO} + \frac{\sqrt{V_{UPO}}}{r}$$ 14-9:

Where:

$$V_{LPO} = G_3^2 S_{PO}^4 + H_4^2 S_E^4 + G_{34} S_{PO}^2 S_E^2$$

$$V_{UPO} = H_3^2 S_{PO}^4 + G_4^2 S_E^4 + H_{34} S_{PO}^2 S_E^2$$

In this case negative bounds are increased to zero.

14.6.1.5 Interval for γ_M

The variability of the measurement process γ_M after Burdick et al[67]is:

$$\gamma_M = \sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2 = \frac{[\theta_O + (p-1)\theta_{PO} + p(r-1)\theta_E]}{pr}$$ 14-10:

The recommended definition for Interval for γ_M is provided by Graybill and Wang [137].

The upper and lower limits for an approximate 100(1 - α)% confidence interval are:

The lower limit after Burdick et al[67]is:

$$L = \hat{\gamma}_M - \frac{\sqrt{V_{LM}}}{pr}$$ 14-11:

The upper limit is:

$$U = \hat{\gamma}_M + \frac{\sqrt{V_{UP}}}{pr}$$ 14-12:

Where:

$$V_{LM} = G_2^2 S_O^4 + G_3^2 (p-1)^2 S_{PO}^4 + G_4^2 p^2 (r-1)^2 S_E^4$$

$$V_{UM} = H_2^2 S_O^4 + H_3^2 (p-1)^2 S_{PO}^4 + H_4^2 S_E^4$$

Once the individual constants have been computed the confidence intervals can be constructed in a spreadsheet [73] for each GRR.

14.7 Appendix 5: Concrete GRR Test Site Information

14.7.1 Meteorological Data:

Table 14-12: Meteorological conditions on concrete GRR test site

29/06/2010						
Room	**Temp AM (°C)**	**Temp PM (°C)**	**Humidity (%)**	**Humidity (%)**	**Pressure (Pa)**	**Pressure (Pa)**
1	26.1	26.4	63.3	58.8	1012.2	1011.2
2	26.1	26.7	64.3	58.2	1012.2	1011.2
3	24.9	24.0	62.3	60.0	1012.2	1011.2
4	25.4	24.8	66.3	63.2	1012.3	1011.2
5	27.8	26.8	60.1	57.7	1011.9	1011.0
6	27.2	26.7	60.0	59.3	1012.1	1010.7
7	25.9	25.8	65.9	59.1	1012.1	1011.0
8	24.9	.24.5	65.7	60.4	1011.8	1010.3
10	26.7	24.3	55.3	53.2	1011.6	1011.2
11	26.2	25.9	55.1	52.9	1011.7	1011.1
30/06/2010						
Room	**Temp AM (°C)**	**Temp PM (°C)**	**Humidity (%)**	**Humidity (%)**	**Pressure (Pa)**	**Pressure (Pa)**
1	24.2	23.3	67.3	69.0	1012.3	1010.8
2	22.4	22.8	67.1	68.9	1012.3	1010.6
5	24.0	23.9	67.5	71.4	1011.9	1010.3
6	24.5	24.6	66.8	68.3	1011.8	1010.4
10	24.7	24.8	60.4	65.2	1011.5	1010.0
3	24.5	24.0	63.9	61.7	1011.8	1010.5
4	24.7	24.2	60.8	62.7	1012.1	1010.5
7	24.9	.24.5	65.7	60.4	1011.8	1010.3
8	25.2	24.7	59.6	59.7	1011.8	1010.2
11	25.2	25.4	57.4	58.2	1011.5	1009.8

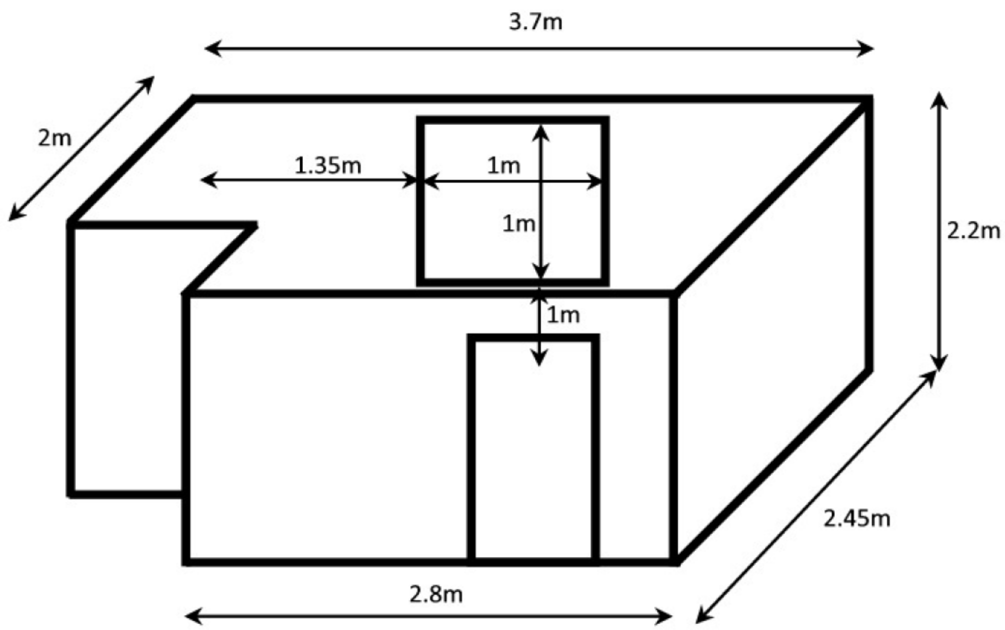
This shows that the temperature, humidity and the barometric pressure inside the test rooms is stable during the survey process and unlikely to be a significant cause of uncertainty.

14.7.2 Test Rooms – GRR: GRR Rooms - 6 Room Pairs in total

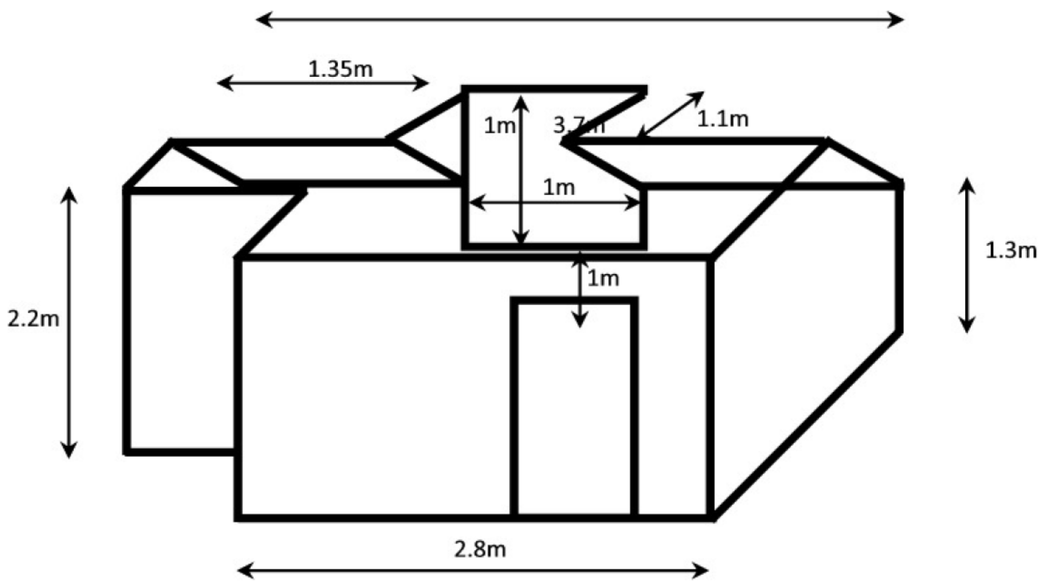
Flat No			Flat No		
Source room	Room Type	Vol m3	Receiver Room	Room Type	Vol m3
11	Bed	25.2	7	Bed	21.7
7	Bed	21.7	3	Bed	21.7
8	Bed	21.7	4	Bed	21.7
9	Bed	19.2	6	Bed	22.7
6	Bed	22.7	2	Bed	22.7
5	Bed	19.0	1	Bed	19.0

14.7.3 Room Dimensions

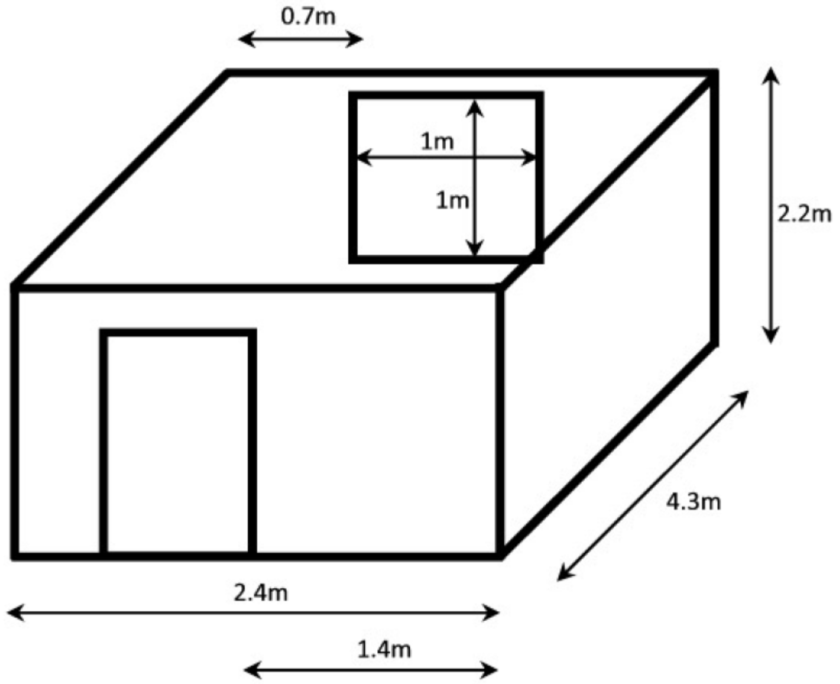
Rooms 5, 1



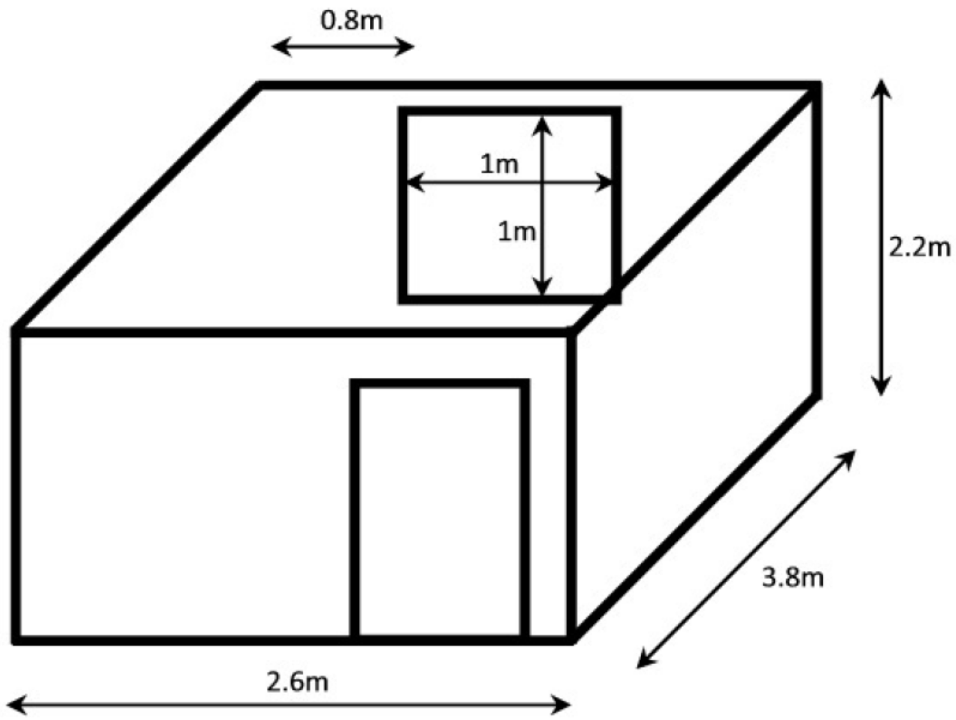
Room 9



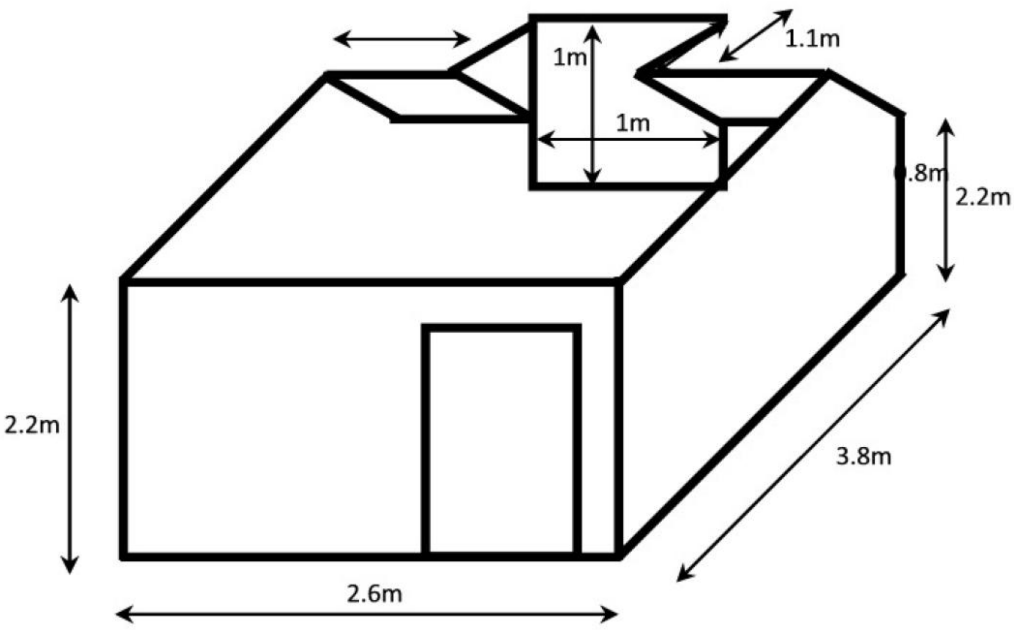
Rooms 10, 6 and 2



Rooms 7,3,11,8,4



Room 11



14.8 Appendix 6: List of Symbols and constants

Symbol	Description
$D_{nT,w}$	Weighted standardised level difference (dB)
C_{tr}	Spectrum adaptation term for traffic
D_{nT}	Standardised level difference (dB)
r	Repeatability – Within laboratory variance
R	Reproducibility – Between laboratory variance
δq	Total uncertainty (added in quadrature)
δx	Uncertainty of input x
δy	Uncertainty of input y
Y	The measurand (not measured directly
X_1, , X_2,……, X_N	N directly measured quantities
$u_c^2(y)$	Combined standard uncertainty
m	general mean (expectation)
B	laboratory component of bias under repeatability conditions;
e	random error occurring in every measurement under repeatability conditions.
V	is the receiving room volume in cubic m.
S	is area of separating element
A	is the equivalent absorption area in square m.
D	is level difference
R’	Apparent sound reduction index (dB)
ε^2	Normalised variance

S_T	Total area of the room surfaces (m^2)
d	Distance (m)
p	Sound pressure (Pa)
N	Mode count in a frequency band
B	Bandwidth (Hz)
T	Period (s), averaging time (s), reverberation time(s), temperature (°C)
f_S	Schroeder cut-off frequency (Hz)
R	Sound reduction index (dB)
B_{meas}	Measured static barometric pressure (hPa)
B_N	Normalised barometric pressure 1013hPa
T_{meas}	Measured Temperature (K)
T_N	Normalised temperature 293K
R_N	Normalised sound reduction index (dB)
R_{meas}	Measured sound reduction index (dB)
$L_{n,N}$	Normalized impact sound ressure level
$L_{n,meas}$	Measured impact sound pressure level
SS	Sum of squares
SS_E	Variation around the measurement instrument
SS_{O*P}	Variation around operators' and parts' mean
SS_P	Variation around parts' mean
SS_O	Variation around operators' mean
MS	Mean square: sum of squares divided by degrees of freedom. (NB: MS are always variances)

Symbol	Description
MS_E	Instrument variance
MS_{O*P}	Variance between parts and operators
MS_P	Part variance
MS_O	Operator variance
DF	Degrees of freedom
S_{xi}^2	Between laboratory variance
S_{pooled}^2	Within laboratory variance pooled.
$\widehat{\ }$	A sample estimate
$\overline{\ }$	Mean value, average
$\hat{\sigma}_L^2$	Between laboratory estimate (dB)
σ_r	Repeatability standard deviation (dB)
σ_R	Reproducibility standard deviation (dB)
μ_Y	Population mean
S_{total}^2	Total variance
S_E^2	Total variance within the sample
S_P^2	Total variance between samples
O_i	Random variable representing the effect of the operator
P_j	Random variable representing the effect of the part
$(OP)_{ij}$	Random variable representing the effect of the operator by part interaction
$R_{k(ij)}$	Random variable representing the replications on the measurement
σ_o^2	Variance of operator
σ_p^2	Variance of part
σ_{op}^2	Variance of the operator by part interaction
σ_R^2	Variance of the replication of measurement
$\overline{Y}_{i**}$	Dot notation: summation over a subscript, ith treatment level mean
$\overline{Y}_{*j*}$	Dot notation: summation over a subscript, jth treatment level mean
$\overline{Y}_{ij*}$	Dot notation: summation over a subscript, kth replicates mean
$\overline{Y}_{***}$	Dot notation: summation over all treatments and grand mean
γ_P	Variance of the process – part to part variance
γ_M	Variance of the measurement system
γ_R	Ratio of process variance to measurement variance
L'	Total length of all edges (m)
S'	Total surface area of the space (m^2)
C_o	Speed of sound m/s – assumed 340.3m/s
f	Band centre frequency (Hz)
N(f)	Statistical modal density
L	The adjusted signal level in (dB)
L_{sb}	The level of signal and background noise combined (dB)
L_b	The background noise level in (dB)
σ_T	Standard deviation of timber floor
σ_C	Standard deviation of concrete floor
S_R^2	Reproducibility variance – Between laboratory variance
S_r^2	Repeatability variance – Within laboratory variance
F	F Statistic – Used in test of variance: F is a ratio of sample variances which the expected value = 1 indicates no difference

Symbol	Description
P	p-value: experimental test statistic
VarComp	It is the variance component for each source
%Contribution (of VarComp)	The proportional contribution of a particular source to the total variation in the model
α	Confidence interval parameter e.g. α =0.05 for 95%
H_o	Null Hypothesis : $\mu = \mu_0$
H_A	Alternative hypothesis: $\mu \neq \mu_0$
σ_{GRR}^2	Component of variance associated with the gauge ($\sigma_r^2 + \sigma_R^2 + \sigma_{p.o}^2$)
σ_r^2	Component of variance associated with the repeatability
σ_R^2	Component of variance associated with the reproducibility
σ_o^2	Component of variance associated with the operator
$\sigma_{p.o}^2$	Component of variance associated with the operator by part interaction
σ_p^2	Component of variance associated with the part
σ_{Total}^2	Total variance
σ_{GRR}	Standard deviation associated with the gauge
σ_r	Standard deviation associated with the repeatability
σ_R	Standard deviation associated with the reproducibility
σ_o	Standard deviation associated with the operator
$\sigma_{p.o}$	Standard deviation associated with the operator by part interaction
σ_p	Standard deviation associated with the part
σ_{Total}	Total standard deviation
$\boldsymbol{y}$	Output estimate
$\boldsymbol{k}$	A coverage factor
$\boldsymbol{U_p}$	expanded uncertainty of output estimate (y)
$\hat{\boldsymbol{\sigma}}^2_{repeatability}$	Estimate of repeatability
$\hat{\boldsymbol{\sigma}}^2_{reproducibility}$	Estimate of reproducibility
$\hat{\boldsymbol{\sigma}}^2_{gauge}$	Estimate of gauge
$\boldsymbol{u}$	Constant for the confidence interval of the gauge
$\boldsymbol{v}$	Constant for the confidence interval for reproducibility.
$\chi^2_{\alpha/2,m}$	Chi-squared distribution upper limit
$\chi^2_{1-\alpha/2,m}$	Chi-squared distribution lower limit
$\boldsymbol{\theta_P}$	Expected mean square for part
$\boldsymbol{\theta_O}$	Expected mean square for operator
$\boldsymbol{\theta_{PO}}$	Expected mean square for part and operator
$\boldsymbol{\theta_E}$	Expected mean square error
$\boldsymbol{\mu_y}$	Mean of population of measurements
$\boldsymbol{\gamma_P}$	Variance of the part
$\boldsymbol{\gamma_M}$	Variance of the measurement system
$\boldsymbol{\gamma_O}$	Variance of the operator
$\boldsymbol{\gamma_{PO}}$	Variance (interaction) between the part and operator
$\boldsymbol{PPS\sigma^2_{total}}$	Total variance measured by Parkin et al for the simple concrete floor;
$\boldsymbol{GRR\sigma^2_{gauge}}$	Variance of the measurement system calculated from concrete GRR
$\widehat{\boldsymbol{PPS\sigma^2_{part}}}$	Estimated variance produced by the Parkin concrete floor

15 References

1. Knudsen, V.O., *MEASUREMENT AND CALCULATION OF SOUND-INSULATION* Journal of the Acoustical Society of America, 1930. **2**: p. 129-140.
2. HMSO, *Approved Document E: Resistance to the passage of sound*. 2003, HMSO.
3. *Technical Handbooks 2011 Domestic Noise*, in *Section 5 Noise*. 2011, May 2011: Scotland.
4. BSI, *BS EN ISO/IEC 17025 - General requirements for the competence of testing and calibration laboratories*. 2005, BSI.
5. *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.*, in *BS ISO 5725-2:1994* 1994.
6. (JCGM), J.C.f.G.i.M., *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*, in *JCGM 100: 2008 - GUM 1995 with minor corrections*. 2008.
7. *Accuracy (Trueness and precision) of measurement methods and results - Part 1 General principles and definitions*, in *BS ISO 5725-1:* . 1994.
8. *ISO 140-2: 1991 - Acoustics - Measurement of sound insulation in buildings and of building elements*, in *Part 2: Determination, verification and application of precision data.* 1991, BSI.
9. ISO, *ISO/CD 12999-1 Acoustics - Determination and application of measurement uncertainties in building acoustics - Part 1: Sound Insulation*. 2011.
10. Board, A.C.o.t.B.R., *Sound Insulation and Acoustics - Post War Building Studies No14*, D.o.S.I. Research, Editor. 1944.
11. Rasmussen, B. and J.H. Rindel, *Sound insulation between dwellings - Descriptors applied in building regulations in Europe.* Applied Acoustics, 2010. **71**(3): p. 171-180.
12. Gray, P.G., A. Cartwright, and P.H. Parkin, *Noise in three groups of flats with different floor insulations*, N.B. Studies, Editor. 1958, HMSO.
13. Parkin, P.H., H.J. Purkis, and W.E. SCHOLES, *Development of Grading System: Part II*, in *Field Measurements of Sound Insulation between Dwellings*. 1960, HMSO.
14. *Digest No88 - Part 1 - Sound Insulation of Dwellings*, B.R. Station, Editor. 1954, HMSO. p. 1-4.
15. *Digest No89 - Part 2 - Sound Insulation of Dwellings*, B.R. Station, Editor. 1954, HMSO.
16. Parkin, P.H. and E.F. Stacey, *Recent Research on Sound Insulation in Houses & Flats.* Journal of the Royal Institute of British Architects, 1954. **61**(9): p. 372-376.
17. *BS EN ISO 717-1: Acoustics-rating the sound insulation in buildings and of building elements. Part 1. Airborne sound insulation*, in *BS EN ISO 717-1*. 1996.
18. HMSO, *Approved Document E: Resistance to the passage of sound*, HMSO, Editor. 1985.
19. *Digest No88 - Part 1 - Sound Insulation of Dwellings (Revised)*, B.R. Station, Editor. 1964, HMSO.
20. HMSO, *Approved Document E: Resistance to the passage of sound*. 1992, HMSO.
21. *British Standard Recommendations for field and laboratory measurement of airborne & impact sound transmission in buildings*, in *BS 2750:*. 1956.
22. *BS EN ISO 140-4. Acoustics-measurement of sound insulation in buildings and of building elements. Part 4. Field measurements of airborne sound insulation between rooms*, in *BS EN ISO 140-4.* . 1998.

23. BSI, *ISO 354: 2003 Acoustics — Measurement of sound absorption in a reverberation room*. 2003.
24. *Accuracy (trueness and precision) of measurement methods and results - Part 3: Intermediate measures of the precision of a standard measurement method.*, in *BS ISO 5725-3:* . 1994.
25. *Accuracy (trueness and precision) of measurement methods and results - Part 4: Basic methods for the determination of the trueness of a standard measurement method.*, in *BS ISO 5725-4:* . 1994.
26. *Accuracy (trueness and precision) of measurement methods and results - Part 6 Use in practice of accuracy values*, in *BS ISO 5725-6:* . 1994.
27. A. Farina, et al., *Intercomparison of Laboratory Measurements of Airborne Sound Insulation of Repeatability and Reproducibility Values Partitions for the Determination of Repeatability and Reproducibility Values.* Building Acoustics, 1999 **6**(2): p. pp. 127-140(14).
28. A. Schmitz, A. Meier, and G. Raabe, *Inter-Laboratory Test of Sound Insulation Measurements on Heavy Walls: Part 1.* Building Acoustics, 1999 **6**(3-4): p. 159-169(11).
29. A. Schmitz, A. Meier, and G. Raabe, *Inter-Laboratory Test of Sound Insulation Measurements on Heavy Walls: Part 2.* Building Acoustics, 1999. **6**(3-4): p. 171-186(16).
30. Carvalho, O.P.A., *Reproducibility in inter-laboratory impact sound insulation measurements*, in *ICSV13 - Vienna*. 2006: Vienna.
31. D. Hoffmeyer, J Christensen, and H.S. Olesen *Nordic Intercomparison Programme in the Field of Acoustics - 3. Measurement: Field measurements of Airborne Sound insulation*. 1995.
32. Fausti, P., R. Pompoli, and R.S. Smith, *An Intercomparison of Laboratory Measurements of Airborne Sound Insulation of Lightweight Plasterboard Walls.* Building Acoustics, 1999. **6**(No2): p. 14.
33. Muellner, H., *Inter-laboratory test - measurement of airborne and impact sound insulation of lightweight floors, focussing on the extended frequency range below 100Hz*, in *Forum Acusticum 2011*. 2011: Allborg, Denmark. p. 2599-2604.
34. S.D.Kristensen and B. Rasmussen, *Repeatability and Reprodicibility of sound Insulation Measurements*. 1984 Danish Acoustical Institute
35. LCJ van Luxemburg and H. Martin, *Repeatability and Reproducibility of laboratory airborne sound insulation measurements: A Dutch Precision Experiment*. 1986, TPD - Ministry of Housing.
36. Lang, J., *A Round Robin on Sound Insulation in Buildings.* Applied Acoustics, 1997. **Vol. 52**(NO. 314): p. pp. 225-238.
37. Scamoni., F.S., C. Mussin, M. Galbusera, E. Bassanino, M. Zambon,G. Radaelle, S., *Repeatability and reproducibility of field measurements in buildings*, in *EuroNoise09*. 2009: Edinburgh.
38. Hall, R., *Sound insulation measurements in buildings*, in *Euronoise09*. 2009: Edinburgh.
39. Scrosati, C.S., F. Bassinino,B. Mussin,M. Zambon,G., *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 Engineering, 2013. **61**(3): p. 291-306.
40. Wieland Weise and V. Wittstock, *Using Round Robin Test Results for the Accreditation of Laboratories in the Field of Building Acoustics in Germany.* BUILDING ACOUSTICS, 2005 **12**(Number 3): p. Pages 189 - 206.
41. Cocchi, A. and G. Semprini, *Sound Insulation and Flanking Transmission*, in *Euronoise*. 2003: Naples. p. 1-11.

42. Mahn, J.P., John, *On the Uncertainty of the EN12354-1 Estimate of the Flanking Sound Reduction Index Due to the Uncertainty of the Input Data.* BUILDING ACOUSTICS, 2009. **16**.
43. SCHOLES, W.E., *A NOTE ON THE REPEATABILITY OF FIELD MEASUREMENTS OF AIRBORNE SOUND INSULATION?* Journal of Sound and Vibration, 1969. **10(l)**: p. 1-6.
44. Wittstock, V. and C. Bethke, *The Role of Static Pressure and Temperature in Building Acoustics.* BUILDING ACOUSTICS, 2003. **10**(No2): p. 159-176.
45. Kjaer, B., *Para 3.13.3 Effect of Humidity*, in *B&K Technical Documentation Microphone Handbook*. B&K.
46. Schroeder, M.R., *Effect of Frequency and Space Averain on the Transmission Responses of Multimode Media.* Journal of the Acoustical Society of America, 1969. **46**: p. 277-283.
47. Waterhouse, R., *Interference patterns in Reverberant Sound Fields.* Journal of the Acoustical Society of America, 1955. **27**(2): p. 247-258.
48. Waterhouse, R., *Noise measurement in reverberant rooms.* Journal of the Acoustical Society of America, 1973. **54**(4): p. 931-934.
49. Waterhouse, R. and D. Lubman, *Discrete versus Continuous Space Averaging in a Reverberant Sound Field.* Journal of the Acoustical Society of America, 1970. **48**(No1 Part 1): p. 1-5.
50. Waterhouse, R.V., *Sampling Statistics for an Acoustic Mode.* Journal of the Acoustical Society of America, 1970. **47**(No4 Pt1): p. 961-967.
51. Waterhouse, R.V., *Statistical Properties of Reverberant Sound Fields.* Journal of the Acoustical Society of America, 1968. **43**(No6): p. 1436-1444.
52. Lubman, D., *Fluctuations of Sound with Position in a Reverberant Room.* Journal of the Acoustical Society of America, 1968. **44**(No6): p. 1491 - 1502.
53. Lubman, D., *Spatial Averaging in a Diffuse Sound Field - Letters to the Editor.* Journal of the Acoustical Society of America, 1969. **46**(3 Pt1): p. 532-534.
54. Lubman, D., R. Waterhouse, and C.-S. Chien, *Effectiveness of continuous spatial averaging in a diffuse sound field.* Journal of the Acoustical Society of America, 1971. **53**(2): p. 650-659.
55. Craik, R.J.M., *On the Accuracy of Sound Pressure Level Measurements in Rooms.* Applied Acoustics, 1990. **29**: p. 25-33.
56. Sewell, E.C., *SOUND INSULATION OF STEPPED OR STAGGERED WALLS OF PLASTERED MASONRY.* Journal of Sound and Vibration, 1982. **84**(4): p. 463-480.
57. Craik, R.J.M. and J.A. Steel, *The Effect of Workmanship on Sound Transmission through Buildings - Part 1 - Airborne Sound* Applied Acoustics, 1989. **27**: p. 57-63.
58. Craik, R.J.M. and D. Evans, *The Effect of Workmanship on Sound Transmission through Buildings - Part 2 - Airborne Sound.* Applied Acoustics, 1989. **27**: p. 137-145.
59. Goydke H., Seibert B.R.L. , and S. W., *Considerations on the evaluation of uncertainty values of building acoustic single number quantities*, in *Euronoise 2003*. 2003: Naples.
60. Wittstock, V., *Uncertainties in building acoustics.* , in *Proceedings of Forum Acusticum*. 2005: Budapest.
61. Wittstock, V., *On the Uncertainty of Single-Number Quantities for Rating Airborne Sound Insulation.* ACTA ACUSTICA UNITED WITH ACUSTICA, 2007. **Vol. 93**: p. 375 - 386.
62. Lyn J, et al., *Empirical versus modelling approaches to the estimation of measurement uncertainty caused by primary sampling* The Analyst, 2007. **132**: p. 1231-1237.

63. Taibo, L. and H. Glasserman De Dayan, *ANALYSIS OF VARIABILITY IN LABORATORY AIRBORNE SOUND INSULATION DETERMINATIONS.* Journal of Sound and Vibration, 1983. **91**((3)): p. 319-329.
64. Davern W, A.D., P, *First report on Australasian comparison measurements of sound absorption coefficients*. 1980, Commonwealth Scientific and Industrial Research Organization - Division of Building Research.
65. Davern W.A., D.P., *Second report on Australaisian comparison measurements of sound absorption coefficients*. 1985, Commonwealth Scientific & Industrial Research Organisation - Division of Building Research.
66. Deldossi, L., Zappa, D., *ISO 5725 and GUM: comparison and comments.* Accreditation and Quality Assurance, 2009. **14**: p. 159-166.
67. Burdick, R.K., Borror,C.M, Montgomery,D.C., *Design & Analysis of Gauge R&R Studies.* 2005: SIAM.
68. Montgomery, D.C., Runger,G.C, *Gauge Capability and Designed Experiments. Part 1: Basic Methods.* Quality Engineering, 1993. **6**(1): p. 115-135.
69. Montgomery, D.C., Runger,G.C, *Gauge Capability Analysis and Designed Experiments. Part II: Experimental Design Models and Variance Component Estimation.* Quality Engineering, 1993. **6**(2): p. 289-305.
70. Montgomery, D.C., *Design and Analysis of Experiments*. 8th ed. 2013: John Wiley & Sons Inc.
71. Borror, C.M., Montgomery, D.C., Runger, G.C., *Confidence intervals for variance components from gauge capabilitiy studies.* Quality and Reliability Engineering International, 1997. **13**: p. 361-369.
72. Burdick, R.K., Borror,C.M, Montgomery,D.C., *A Review of Methods for Measurement Systems Capability Analysis.* Journal of Quality Technology, 2003. **35**(4): p. 342-354.
73. Burdick, R.K., Borror,C.M, Montgomery,D.C., *Computer Programs - Excel Spreadsheet for calculation of MLS confidence intervals*, in *Design and Analysis of Gauge R&R Studies*. 2005, SIAM.
74. Taylor, J.R., *An Introduction to Error Analysis - The study of uncertainties in physical measurement*. 2nd edition ed. 1997, USA: University Science Books.
75. Kirkup, L. and R.B. Frenkel, eds. *An introduction to uncertainty in measurement using the GUM*. 2006, Cambridge University Press
76. *M3003 The Expression of Uncertainty and Confidence in Measurement*. 2007, UKAS.
77. Walter Bich, M.G. Cox, and P.M. Harris, *Evolution of the 'Guide to the Expression of Uncertainty in Measurement'.* Metrologia (2006) **43**: p. S161-S166.
78. *Statistics - Vocabulary and symbols - Part 2: Applied Statistics*, in *BS ISO 3534-2:2006*. 2006.
79. Ellison, S.L.R. and V.J. Barwick, *Using validation data for ISO measurement uncertainty estimation.* The Analyst. **Vol. 123** p. (1387-1392).
80. BIPM, *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*. 2008.
81. Mahn, J.P., John, *On the probability density functions of the terms described by the EN12354 prediction method.* Building Acoustics, 2008. **15**(4): p. 262-287.
82. Waterhouse, R., *Statistical properties of reverberant sound fields.* Journal of the Acoustical Society of America, 1967. **43**(6): p. 1436-1444.
83. Bodlund.K, *STATISTICAL CHARACTERISTICS OF SOME STANDARD REVERBERANT SOUND FIELD MEASUREMENTS.* Journal of Sound and Vibration, 1976. **45**(4): p. 539-557.
84. 101:, J., *Guide to the expression of uncertainty in measurement*, in *Supplement 1 - Propagation of distributions using a monte carlo method.* 2008.

85. Mahn, J., *Prediction of Flanking Noise Transmission in Lightweight Building Constructions: A Theoretical and Experimental Evaluation of the Application of EN12354-1*. 2008, Canterbury.
86. Bessac F and B. B., *Sound power level measurement in reverberant room: estimating the uncertainties and their laboratory use.* Journal of Noise Control Engineering, 2007. **55**(1).
87. Chien, C.-S., R. Waterhouse, and D. Lubman, *Spherical averaging in a diffuse sound field.* Journal of the Acoustical Society of America, 1975. **57**(No4): p. 972-975.
88. Hopkins, C., *On the efficacy of spatial sampling using manual scanning paths to determine the spatial average sound pressure level in rooms.* J.Acoust. Soc Am., 2011. **129**(5): p. 3027-3034.
89. Fothergill, L.C. and R.S. Alphey, *The Effect on Sound Insulation of Small Holes Through Solid Masonry Walls.* Applied Acoustics, 1987. **21**: p. 247-251.
90. BSI, *BS EN ISO 717-1: 2013*, in *Acoustics — Rating of sound insulation in buildings and of building elements Part 1: Airborne sound insulation (ISO 717-1:2013)*. 2013.
91. Fothergill, L.C. and N. Hargreaves, *The Effect on Sound Insulation between Dwellings when Windows Are Close to a Separating Wall.* Applied Acoustics, 1992. **35**: p. 253-261.
92. Sewell, E.C. and J.E. Savage, *Effect of Associated Walls on the Sound Insulation of Concrete Party Floors.* Applied Acoustics, 1987. **20**: p. 297-315.
93. Deldossi, L., Zappa, D., *Confidence Intervals for Variance Components in Measurement System Capability Studies.* Communication In Statistics- Theory And Methods. **41**(16-17).
94. Schroeder, M.R., *The ''Schroeder frequency'' revisited.* Journal of the Acoustical Society of America, 1996. **99**(5): p. 3140-3141.
95. Leiva, R.A.G., F.A.,, *Confidence Intervals for Variance Components in the Balanced Two-Way Model with Interaction.* Communication In Statistics-Simulation and Computation, 1986. **15**: p. 301-322.
96. Hopkins, C. and P. Turner, *Field measurement of airborne sound insulation between rooms with non-diffuse sound fields at low frequencies.* Applied Acoustics 2005. **66**: p. 1339–1382.
97. Weise, W., *Measurement Uncertainties for Sound Field Levels in Rooms.* BUILDING ACOUSTICS, 2003 **10**(Number 4): p. Pages 281 - 287.
98. Torjussen, M., *Sound Insulation of Staggered Masonry Walls*. 2010, University of Southampton.
99. Persuy, G., *Dispersion of sound insulation lightweight plasterboard walls: the approach of the french gypsum industries association.* , in *Proceedings of the Symposium Managing Uncertainty in Noise Measurement and Prediction*. 2005: Le Mans, France.
100. Mandel, J., *Repeatability and Reproducibility.* Journal of Quality Technology, 1972. **4**(2).
101. Smith, R.S., Pompoli, R. Fausti, P., *An Investigation into the Reproducibility Values of the European Inter-Laboratory Test for Lightweight Walls.* Building Acoustics, 1999. **6**(3-4).
102. AIAG, *Automotive Industry Action Group: Measurement System Analysis*. 4th ed. 2010.
103. Wheeler, D., J., *EMP III: Evaluating the Measurement Process & Using Imperfect Data*. 2006: SPC Press.
104. AIAG, *Automobile Industry Action Group: Measurement System Analysis*. 3rd ed. 2002.

105. AIAG, *Automotive Industry Action Group: Measurement System Analysis*. 2nd ed. 1995.
106. AIAG, *Automotive Industry Action Group: Measurement System Analysis*. 1990.
107. Tsai, P., *Variable Gauge Repeatability and Reproducibility Study Using The Analysis of Variance Method.* Quality Engineering, 1988. **1**(1): p. 107-115.
108. Montgomery, D.C., *Design & Analysis of Experiments*. 5th ed. 2001: John Wiley.
109. Vardeman, S.B., VanValkenburg,E.S., *Two-Way Random-Effects Analyses and Gauge R&R Studies.* Technometrics, 1999. **41**(3): p. 202-211.
110. Brigham, A., B., *Gauge R&R Measurement System Analysis (MSA) Carl Zeiss*. 2005.
111. Srirattanasak, S., *How can MiniTab help in quality improvement - Hitachi Global Storage Technologies - Hyper Balance Failure Reduction Project*. 2007.
112. Searle, S., R., *Linear Models*. 1971: John Wiley & Sons.
113. Mason, R., L. Gunst,R,F. Hess,J,L., *Statistical Design and Analysis of Experiments: With Applications to Engineering and Science.* 2nd ed. 2003: Wiley.
114. *Robust Details Handbook*, in *Part E robust details*. 2012.
115. Gibbs BM, M.S., *Airborne sound level difference between dwellings at low frequencies.* BUILDING ACOUSTICS, 2004. **11**(1): p. 17.
116. Maluski S, G.B., *Application of a finite-element model to low-frequency sound insulation in dwellings.* Journal of the Acoustical Society of America, 2000. **104**(4): p. 10.
117. Whitfield, W., A., *Work Instruction - Airbrone Sound Insulation Test Procedure*, in *Quality Management System*. 2004, noise.co.uk ltd. p. 3.
118. Hopkins, C., *Sound Insulation*. 2007: Butterworth - Heinemann.
119. Lubman, D., *Precision of reverberant sound power measurements.* Journal of the Acoustical Society of America, 1974. **56**(2): p. 523-533.
120. Dah-You, M., *Frequency & Directional Distribution of Normal Modes of Vibration in a Rectangular Room.* Science Record, 1957. **1**(No6): p. 37-42.
121. Whitfield W.A and B.M. Gibbs. *VARIATION IN FIELD MEASUREMENT OF AIRBORNE SOUND INSULATION OF LIGHTWEIGHT TIMBER FLOORS*. in *Noise in the Built Environment,*. 2010. Ghent.
122. *MiniTab 16*. 2012. p. Statistical Analysis Software.
123. Osma, A., *An assessment of the robustness of gauge repeatability and reproducibility analysis in automotive components.* Journal of Automobile Engineering, 2011. **225**(Proc IMechE Part D).
124. Erdmann, T.P.D., R.J.M.M. Bisgaard, S., *Quality Quandries: A Gage R&R Study in a Hospital.* Quality Engineering, 2010. **22**: p. 46-53.
125. *Minitab*. 2010, Minitab Ltd.
126. Montgomery, D.C., *Statistical Quality Control: A Modern Introduction.* 6th ed. 2009: John Wiley & Sons.
127. *EA-4/16 - EA Guidelines on the expression of uncertainty in quantitative testing*, E.c.-o.f. Accreditation, Editor. 2003.
128. Mathews, P.G., *Design of Experiments with MINITAB*. 2005: ASQ.
129. Leiva, R.A., Graybill,F.A., *Confidence intervals for variance components in the balanced two-way model with interaction.* Communications in statistics: Simulation and Computation, 1986. **15**: p. 301-322.
130. Weerahandi, S., *Generalized confidence intervals.* Journal of the American Statistical Association, 1993. **88**: p. 899-905.
131. Bartlett, M.S., *Properties of Sufficiency and Statistical Tests.* Proceedings of the Royal Society : A: Mathematical, Physical and Engineering Sciences, 1937. **160**.
132. Satterthwaite, F.E., *Synthesis of variance.* Psychometrika, 1941. **6**: p. 309-316.

133. Satterthwaite, F.E., *An approximate distribution of estimates of variance components.* Biometrics Bulletin, 1946. **2**: p. 110-114.
134. Burdick, R.K., Graybill, F.A., *Confidence Intervals on Variance Components*. 1992: Marcel Dekker, New York.
135. Burdick, R.K., Larsen,G.A., *Confidence intervals on measures of variability in R&R studies.* Journal of Quality Technology, 1997. **29**: p. 261-273.
136. Burdick, R.K., Larsen, G.A., *Confidence intervals on parameters in a repeatability and reproducibility study'*, in *Third International Applied Statistics in Industry Conference*. 1995: Dallas.
137. Graybill, F.A., Wang,C.M., *Confidence intervals on non-negative linear combinations of variances.* Journal of the American Statistical Association, 1980. **75**: p. 869-873.
138. Cappelleri, J.C., Ting,N., *A modifed large-sample approach to approximate interval estimation for a particular intraclass correlation coefficient.* Statistics in Medicine, 2003. **22**: p. 1861-1877.
139. Chiang, A.K.L., *Improved confidence intervals for a ratio in an R&R study.* Communications in statistics: Simulation, 2002. **31**: p. 329-344.
140. Hamada, M., Weerahandi,S., *Measurement System Assessment Via Genaralized Inference.* Journal of Quality Technology, 2000. **32**(3): p. 241-253.
141. Milliken, G.A., Johnson, D.E., *Analysis of Messy Data: Designed Experiments*. Vol. 1. 1984: Lifetime Learning, Belmont CA.
142. Ting, N., Burdick, R.K., Graybill, F.A.,Jeyaratman, S.,Lu, T.-F.C., *Confidence intervals on linear combinations of variance components that are unrestricted in sign.* Journal of Statistical Computation and Simulation, 1990. **35**: p. 135-143.