STUDIES IN MODAL DENSITY – ITS EFFECT AT LOW FREQUENCIES

M Wankling University of Huddersfield, Huddersfield, UK
B Fazenda University of Huddersfield, Huddersfield, UK

1 INTRODUCTION

The ability to objectively measure the reproduction quality of a small room at low frequencies has long been desired. Over many years, there have been attempts to produce recommendations, metrics, and criteria by which to define a particular room. These have often concentrated on some aspect of the modal distribution, such as spacing or density. Other attempts have focused upon the deviation from a desired frequency response.

Whilst the subjective validity of objective measures such as these has often been questioned, the notion that a transitional region between a modal and diffuse sound fields exists, dependant on the room volume and reverberation time continues to permeate much thinking. The calculation of this transitional frequency relies on the calculation of a desired modal density. In the case of the most well known definition, the Schroeder Frequency[1], the transitional frequency is that point where the density becomes sufficient that three modes lie within one bandwidth.

Although this idea may well be useful in some instances, such as defining points for the use of statistical sound field analysis, recent thought has cast some doubt over its relevance as a subjective frequency above which we may ignore modal issues[2]. This paper highlights a number of studies along with a new listening test, which help us to better understand the role of modal density upon subjective perception of modal soundfields.

2 OPTIMAL MODAL DENSITY

As frequency increases, modes occur closer and closer together, therefore increasing the modal density. Furthermore, as room volume increases, one also expects an increase in density at a fixed frequency. It can be observed that, if the aspect ratio of the room remains constant, as volume increases, the modal frequency response retains the same shape, but is 'squashed' into a narrower frequency band (Figure 1).

It is often stated that as a larger number of modes become concentrated in a given frequency range, the overall magnitude frequency response becomes flatter and thus is commonly associated with better quality reproduction. The following experimental work tests the subjective relevance of this assumption by searching for an optimal density.

2.1 Tests Omitting Mode-Shapes

The optimal modal density could be defined as that point where no further increase is necessary in order for a room to be perceived in the same way as an ideally smooth case. In an attempt to ascertain this, a number of subjective tests were deployed. Low frequency room responses were modelled using the modal decomposition Greens Function:

$$P_{\omega}(r) = j\omega\rho Q c^2 \sum_n \frac{P_n(r)P_n(r_0)}{X_n(\omega^2-\omega^2{}_n-2j\delta_n\omega_n)} \text{ (Equation 1).}$$

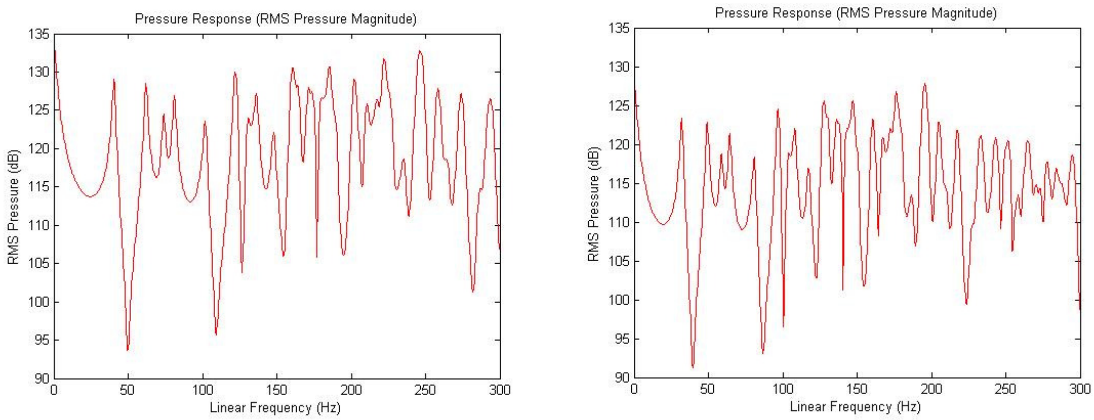


Figure 1 – a) 50m^3 room response b) 100m^3 room response

The room volume may then be increased, in turn increasing the density. As volume increases, subjects were asked if they could detect a difference between the variable room and a reference case which has a smooth response. This then identifies the detection threshold where the volume of the variable room produces an output which is perceptually the same as that of the reference. The density at any given frequency can then be extrapolated using an expression describing typical mode spacing in rectangular rooms[3].

During pilot testing, it became clear that such a threshold was achieved only if the mode-shapes ($P_n(r)$ and $P_n(r_0)$ - the coupling of source and receiver positions in Equation 1) were omitted from the model. Although unrealistic in rooms, this represents the best case scenario where all modes add constructively, resulting in a consistent smoothing of frequency response as the density increases. It is also noted that this represents the conditions assumed for room ratio metrics as suggested by Louden[4], Bonello[5], Bolt[6] etc.

To test for a detection threshold, the PEST (Parameter Estimation by Sequential Testing) method was used[7,8]. Subjects were required to identify the difference between a smooth reference case (modelled with a room of 100000m^3) and a varying volume determined by PEST rules. To ensure that the subject could not simply claim to hear a difference, an ABX procedure was employed. At each volume, three comparisons were made. If the samples were correctly identified three times consecutively, the volume was increased. However, a single incorrect answer represented a failure to detect a difference and therefore the volume was decreased. The requirement of three consecutive correct answers reduces the probability of the subject guessing to 12.5%, and while this is not as low as the typical statistical threshold (<5%), it was considered sufficient given the association with the PEST methodology, which would reduce the volume at the next comparison unless six consecutive guesses were made - a probability of just 1.6%.

Pure tones (0.4 second decaying sinusoids) at 63Hz, 125Hz and 250Hz were convolved with the modelled room response to produce test stimuli. Sample replay levels were weighted and presented according to the 90dB equal loudness contour. Eight subjects were tested, in a quiet studio space, isolated from the test machine.

2.1.1 Results

Figure 2 shows results for the mean value and standard deviation for room volumes where no detectable difference existed between the two cases compared.

In practice the results provide the preferred volume for a particular frequency. In order to extract the density, the modal bandwidth for the corresponding frequency has to be obtained from the damping conditions in the model (δ). Modal density can then be calculated as the number of eigenfrequencies within a modal bandwidth at a given volume. This can be achieved using Bolt's equation[3] as follows:

$$\frac{\delta N}{BW_{modal}} = \frac{4\pi F^2 V}{c^3} \quad \text{(Equation 2)}$$

where F is frequency, V is room volume. This density is indicated in Table 1.

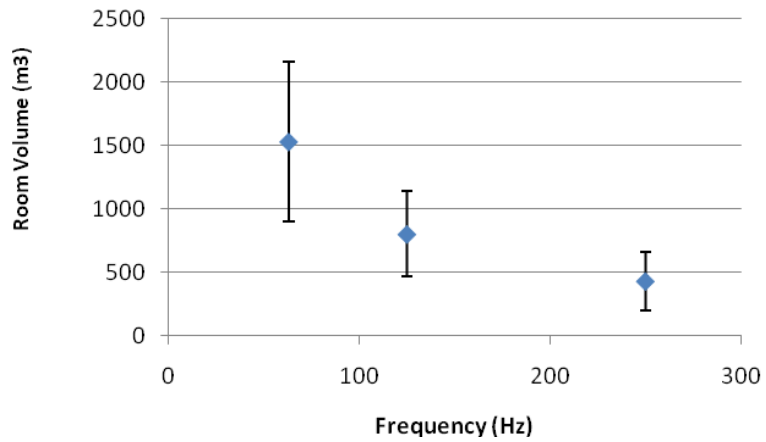


Figure 2 - Mean threshold volume for the detection of difference over three test frequencies

Frequency (Hz)	***63***	***125***	***250***
Modal Bandwidth as prescribed in the model - (2.2/RT)	2.17	2.63	3.75
Subjective Volume Threshold	1529	803	433
Subjective Modal Density (Eq. 4)	4.1	10.3	31.6

Table 1: Modal density according to bandwidth from model damping conditions and subjective volume threshold

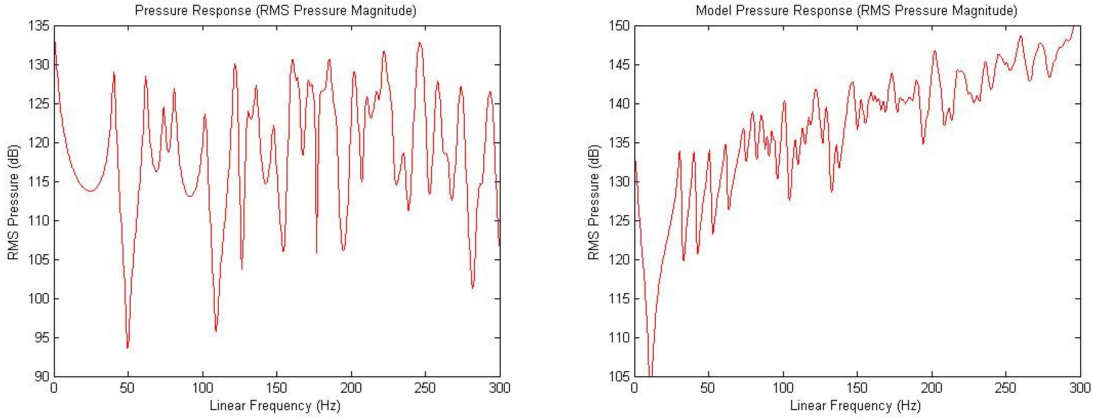
As can be seen, the required modal density appears to rise steeply with frequency, from an initial 4.1 modes per bandwidth, to around 31 at 250Hz. Consequently, no definition of a generic optimal modal density across frequency is possible from these results.

Whilst of some interest, this test does not represent realistic scenarios in terms of either room model or indeed stimulus. To further understand perception of modal density, these issues must be addressed.

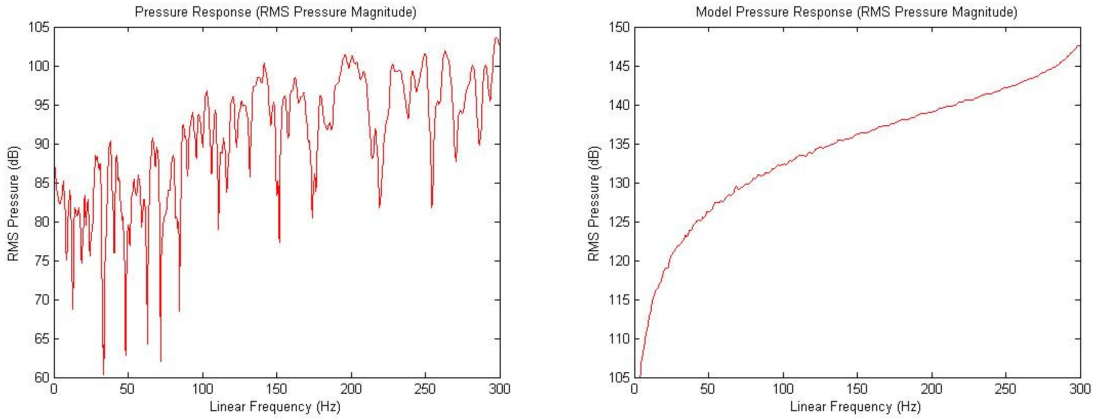
2.2 Including Mode-Shapes

When including the mode shapes in the model, it is not possible to directly test the optimal density using the PEST methodology. Figure 3 compares two room volumes with and without the shape functions. The difference when increasing room volume is clear. There is no systematic smoothing of the response. In fact, the characteristic peaks and dips do not disappear. This results in subjects being able to detect differences between a reference and a variable volume in all cases, and so the PEST can never converge.

It was therefore decided to run an ABX test consisting of ten paired comparisons. Here, not only realistic room models were employed, but also realistic (musical) stimuli. Sample A was a reference room. Two reference volumes were tested - 500m^3 and 10000m^3. Sample B varied in volume approaching the reference. Sample X was unknown and the subject was asked to identify it as either A or B.



a) with mode shapes, b) without mode shapes (room volume 50m3)



c) with mode shapes, d) without mode shapes (room volume 10000m3)

Figure 3 – Comparison of two rooms, with and without shape functions included

Each of the ten ABX tests was fixed at 10 trials. The same eight subjects were tested. Results are presented in Table 2 and Figure 4. In addition to the actual volume of the target room, the volume is indicated as a percentage to enable comparison between the two cases tested.

The same trends are evident for both room sets. Regardless of general volume, if the compared rooms are very different, detection is a simple task. This task remains relatively simple until the differences in volume are below 10%. At this point, it is noted that the two frequency responses become very similar and detection is no longer possible.

A chi-square test was carried out on the data to determine the significance of each result. Values for *p* indicate the success of detection in each case. Values below 0.05 report a significant detection whilst above this value no detection is validated. Therefore, the statistical results show the same trend for both room sets – large and small. It becomes increasingly difficult to detect a difference as the volume approaches that of the reference room. Above around 90%, the subjects are not able to tell the difference significantly.

Small Room	Reference Volume	500	500	500	500	500
	Test Room Volume	100	250	400	450	490
	% of reference	20%	50%	80%	90%	98%
	Mean correct identifications	9.22	8.56	8.33	8.11	6.56
	p	0.0000	0.0011	0.0042	0.0057	0.1512

Large Room Volume	Reference Volume	10000	10000	10000	10000	10000
	Test Room Volume	1000	5000	9000	9500	9990
	% of reference	10%	50%	90%	95%	99%
	Mean correct identifications	9.11	8.56	7.67	5.89	5.89
	p	0.0001	0.0008	0.0244	0.1342	0.9212

Table 2: Results and Chi-Square analysis showing the mean correct identifications and significance of each test - $p<0.05$ indicates the subjects could significantly identify different rooms. Percentages refer to the percentage volume of the test room (sample A) compared to the reference (sample B).

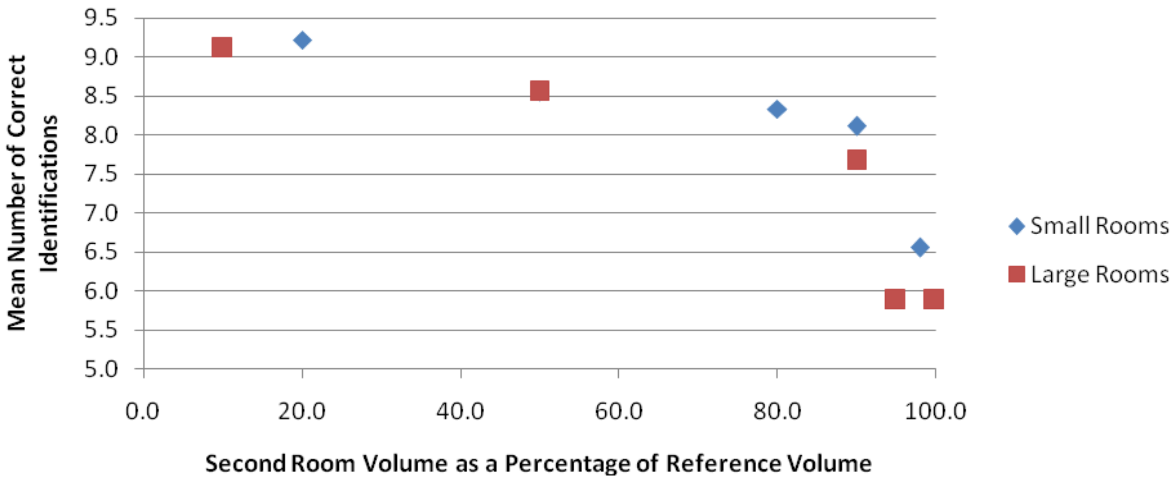


Figure 4: Correct answers in the identification of two room volumes

The interesting outcome is that even in the larger rooms, where modal density is significantly higher; there is no significant reduction of audibility of modal effects. If, as the Schroder Frequency theory implies, the sound field becomes more diffuse, then these results do not suggest that our perception follow those of diffuse conditions.

3 CONTRASTING MODAL DENSITIES

These studies show that, although a threshold can be observed where mode shapes are omitted, in realistic cases, it is not sufficient to rely on the modal density to indicate where a transition region lies in terms of a density which is perceived as the same as that of a smooth response. The question now remains, is there a perceptual improvement to low frequency reproduction *quality* as the density increases? A listening test has been designed which compares auralisations of a number of density scenarios. Auralisations are made in a similar manner to previous successful work[9,10].

Initially, real impulse response measurements were taken in a variety of generally rectangular rooms of varying volumes. The auralisations of these were compared with modelled rooms, in order to verify the credibility of modelled rooms in the listening test.

3.1 Real and Modelled Rooms

Five rooms were measured, using the WinMLS software. Impulse responses were measured up to 300Hz. Rooms were chosen to be as close to rectangular as possible, and volumes varied (approx.) $19m^3$ to $2500m^3$. The rooms were auralised in the same manner as modelled situations, with the real measurement low frequency transfer function replacing the modelled.

Whilst it was not possible to generate the exact transfer function of each room using the decomposition model, the characteristic sound of the real rooms may be produced when the decay time is carefully selected

3.2 Modal Density and Decay Time

Samples were initially produced with both varying low frequency decay times and volumes. However, it is immediately apparent that regardless of room volume, a longer decay time is always rated worse than a shorter time. Ratings therefore become based solely upon decay time and not a difference which may be caused by differing modal density. It is for this reason that comparisons of real rooms are not possible and why modelled rooms are so useful for subjective testing. A listening test was thus devised where modelled rooms were used and the decay kept constant. Only the density is varied, in an attempt to determine if an increase results in a subjective preference, and if so, can a linear trend can be shown?

4 LISTENING TEST

A listening test has been performed which compares pairs of samples, where the low frequency has been modelled, and auralised, with five differing room volumes. The decay times were kept constant, modelled using a typical reverberation time curve for small rooms (see Figure 5). Therefore the test replicates the case where decay time is constant, and yet modal density varies considerably.

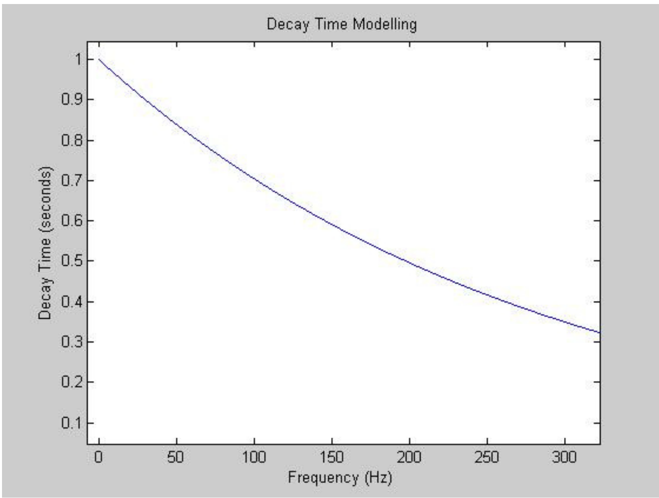
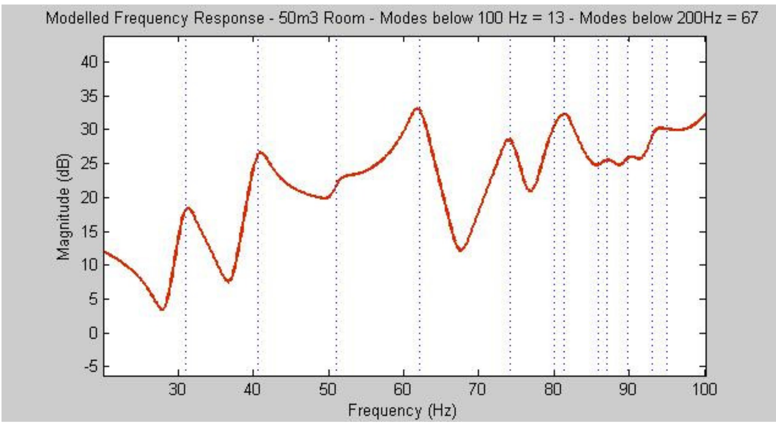


Figure 5 - Decay times were kept constant in the model for each room volume, but varied across frequency in a typical manner of a small well damped room.

Rooms were modelled at 50, 100, 250, 500 and 1000m^3. A point source was, in each case, assumed to lie in a tri-corner. The room ratio was kept constant (2.58 : 1.97 : 1) and the receiver located at a distance varying with room size in the x and y planes but at a constant height of 1.3m.

Example frequency responses can be seen in Figure 6. The vertical lines represent the modal frequencies. From here it is evident that the 1000m^3 room has a much higher density. In the first 100Hz, the 50m^3 room contains 13 modal frequencies, while the 1000m^3 room has 154.



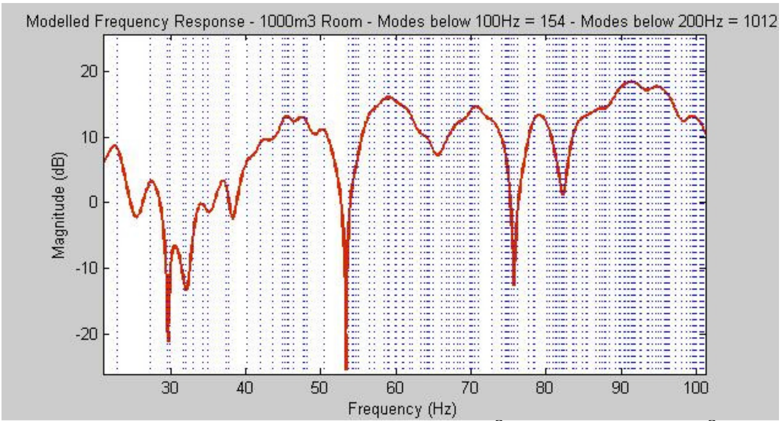


Figure 6 – Frequency responses of room 50m^3 (top) and 1000m^3 (bottom)

As in a typical room, the decay is frequency dependant. Therefore, the frequency range in which we are listening to modal effects may be of importance. We are able to alter this when auralising samples, by changing the crossover frequency between the modelled low frequency and the original music sample. The test was therefore conducted with the crossover at two frequencies, 100Hz and 200Hz.

4.1 Test Procedure

Pair-wise comparisons were chosen over direct preference scaling methods. It was noted that the differences between samples were small, and may differ due to more than one factor. A task requiring subjects to rate all samples on a single scale is difficult under these circumstances. Therefore, pair-wise comparison was chosen, with a simple worse/same/better question asked. Each sample was rated against each other including reversals of Sample A and B, eg. *Sample A = 50m^3, Sample B = 100m^3* and also *Sample A = 100m^3, Sample B = 50m^3*. Listeners were instructed to audition Samples A and B, and then make a decision based upon the overall quality of low frequency reproduction. They were encouraged not to spend a great deal of time on each comparison – if they couldn't detect a noticeable preference, it should be assumed the quality was the same.

The interface was created in the Matlab environment as shown in Figure 7. Before undertaking the test, subjects were given a short training phase where they were played the music sample that would be used in the test, with a variety of differing modal artefacts. This was designed for them to become accustomed with the sample and the likely degradation effects.

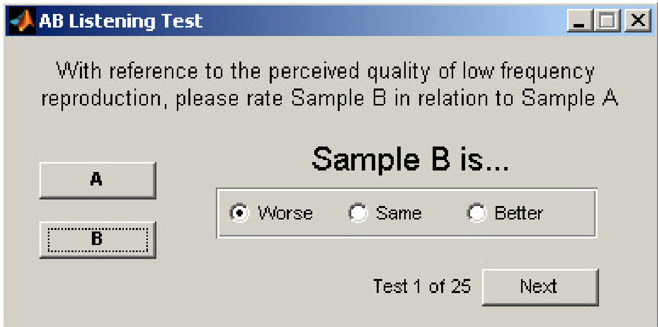


Figure 7 – Test GUI

Seven subjects were tested. Of these, all but one subject had prior listening test experience, and all had experience mixing music in a number of environments. Furthermore, five subjects had been through a listening panel screening test including audiometry and an introduction to critical listening comparisons. One subject reported tiredness before taking the test, although their results were similar to the others.

4.2 Results

The results for each subject's 25 comparisons were placed in a matrix, using -1, 0 and 1 for the ratings 'worse', 'same' and 'better' respectively. The matrix can then be analysed for both judgement errors and to determine the quality of each sample in relation to the others. This technique has been used successfully in the work of Huang et al.[11] in rating the annoyance of noise samples. An example matrix for subject 1 is shown in Table 4.

As can be seen from the table, the subject correctly identified each identical pair as the same (deep shaded cells scored 0). The scores in each row relate to the preference of B against the column A. For example, the highlighted score shows that Sample 4 was rated as better than Sample 2 (score of 1). When this pair were rated the other way around, we can see that Sample 2 was rated worse than Sample 4. If this had not been the case, there would be an inconsistency in preference.

	1	*2*	*3*	*4*	*5*	R_{row}
1	0	0	0	-1	-1	-2
2	-1	0	0	-1	0	-2
3	0	1	0	-1	-1	-2
4	1	*1*	1	0	0	3
5	0	1	1	-1	0	1
R_{col}	0	3	2	-4	-2	

Table 4 – Example matrix of subject 1

In order to analyse the results, the first step is to determine the level of judgement errors for each subject. Two types of error are evaluated here, self comparison errors (sc) and comparisons between two different samples (comp).

Error type	*Error recorded when:*
sc	$R_{ij} \neq 0$ (when i=j)
comp	$R_{ij} \neq -R_{ji}$

Table 5 – Misjudgement errors

For each matrix there are a total of five self comparisons (shaded dark grey in Table 4). There are also ten possible sources of error between comparisons of two samples. A total of 15 errors are therefore possible. The number of errors was calculated for each subject as a percentage of this total (Table 6).

	Subject						
	1	2	3	4	5	6	7
100Hz	33	53	13	87	33	60	60
200Hz	40	53	67	33	27	80	60

Table 6 - Percentage of judgement errors made by each subject

As can be seen, the error rate is in most cases, very high. The average error across subjects and the two test crossover frequencies was 49.9%. This reveals that consistency in judging preference between the various modal densities is very low. In practical terms, this can be related to the difficulty of the task and demonstrates that perceived quality is not directly related to the tested factor, i.e. room volume/density. Even though there are changes in perceived quality, these are not directly related to a controlled factor such as density. The mean error is also similar for both 100Hz and 200Hz crossover frequencies.

It is interesting to note however, that subjects were often rating using the 'better' and 'worse' options. It is not sufficient to say that all samples must be of the same quality – if this were so, there would have been significantly more 'same' ratings. One explanation for this is that the samples *were* similar, but subjects felt that they *should* make a choice, and this could be a source of bias in the results. Another possible explanation is that subjects were unsure what constituted a preference. Indeed, a number commented that there were different parts of the musical sample which appeared to be affected differently. For example, when the kick drum sounded acceptable, the bass guitar was degraded, and vice-versa.

These high error rates imply that it is not particularly worthwhile to analyse individual subject's results and that there is no consistent preference between these samples. However, as a result of many preference ratings being recorded, the matrices are now further analysed in order to note any interesting features.

An overall preference score may be extracted by using Equation 3. Note that the value of R_{col} reveals the opposite preference result to R_{row}, hence the minus sign.

$$\bar{R} = \frac{R_{row} + (-R_{col})}{2} \quad \text{(Equation 3)}$$

Tables 7 and 8 shows the average score determined by Equation 3 for each subject and each room volume:

	Subject							
Volume	1	2	3	4	5	6	7	Mean
$50m^3$	-1	-1	0	-0.5	-2	-2	0.5	-0.86
$100m^3$	-2.5	-3	-4	-2	1	-1.5	-1	-1.86
$250m^3$	-1.5	2	-1	0	-4	1.5	-2	-0.71
$500m^3$	3.5	1	2	0.5	3.5	0	2.5	1.86
$1000m^3$	1.5	1	3	2	1.5	2	0	1.57

Table 7 - 100Hz Crossover Frequency

	Subject							
Volume	1	2	3	4	5	6	7	Mean
$50m^3$	-1	1.5	-2	-2.5	-3	-1	-2.5	-1.50
$100m^3$	-2.5	-4	-2	-2.5	2	-0.5	-1.5	-1.57
$250m^3$	-1.5	3	0	0	-2.5	0	-1.5	-0.36
$500m^3$	3	-0.5	2.5	2.5	3.5	1	3	2.14
$1000m^3$	2	0	1.5	2.5	0	0.5	2.5	1.29

Table 8 – 200Hz Crossover Frequency

Analysis using Equation 3 produces a ranking order, where the mean of the five rooms is zero. The greater the spread (maximum -5 and 5) the greater the perceived difference. Viewing Tables 7 and 8, we are able to see that the scores do appear to differ across the five rooms, although the spread is not particularly wide (±1.86 at 100Hz). This low spread suggests the samples are indeed perceived as similar. Interestingly, whilst the individual scores are not identical for the two test crossover frequencies, the rank order is preserved.

Whilst care should be taken when making claims using these scores, due to the high number of assessment errors observed, the results do provide some interesting insight and points of discussion. It is clear that rooms with larger volumes seem to be attracting higher scores, although the trend is not linear at all. This leads to the conclusion that particular interactions with room transfer function, not modal density, are responsible for perceived quality improvements. There does not appear to exist a difference between the two cross over frequencies tested which is worthy of discussion.

5 DISCUSSION

A number of listening tests have been conducted in an attempt to further understand the relationship between modal density and perceived audio quality at low frequencies. Each test seems to point towards a similar conclusion – that the modal density is not a defining factor when it comes to perception.

The first point to note is that the effect of increasing the modal density is not analogous to a flattening of the frequency response. There appears to be no point where a required density has been met in order that any further increase results in no change to the perceived sound. It is only when the frequency responses themselves are very similar that auralised samples becomes indistinguishable. If there is always a difference, it is likely to follow that some differences will be perceived as superior to others.

This brings us the question of preference. Although differing transfer functions are shown to sound *different*, this does not reveal whether the overall quality of one is preferable to another. It may be that whilst it is possible to distinguish absolute differences between two samples in an ABX type test, there is no real difference in perceived *quality*. The paired comparison listening test in Section 4 was therefore conducted.

Here we find that in terms of preference, results are inconsistent. The question must be asked, why is this so? The answer lies primarily in the similarity between samples, which made the task difficult to undertake. It is also possible that subjects adapted to the samples after a period of time and as they grew more familiar with the sample, could pick out elements which they preferred more easily. Whatever further analysis reveals, it must be treated with caution due to this inconsistency.

When analysing the results matrices using Equation 3, we do not see a linear trend of preference as modal density increases. However, the analysis does suggest that at both crossover frequencies, the best and worst volumes are 500m^3 and 100m^3 respectively. The individual frequency responses of the five rooms are shown in Figure 8. A visual analysis may give us insight as to why these rooms are rated as they are. The 100m^3 room contains an obvious dip with a centre frequency at around 86Hz which is likely to cause audible degradation. The 500m^3 room's response appears to fluctuate to a greater extent, but the smaller spacing between the peaks and dips might account for a better overall perception. On the other hand, there appears no great difference in characteristic between the rooms of 250 and 500m^3, and yet these were scored quite differently according to the analysis. This may of course be another indication that the analysis cannot be taken with great confidence.

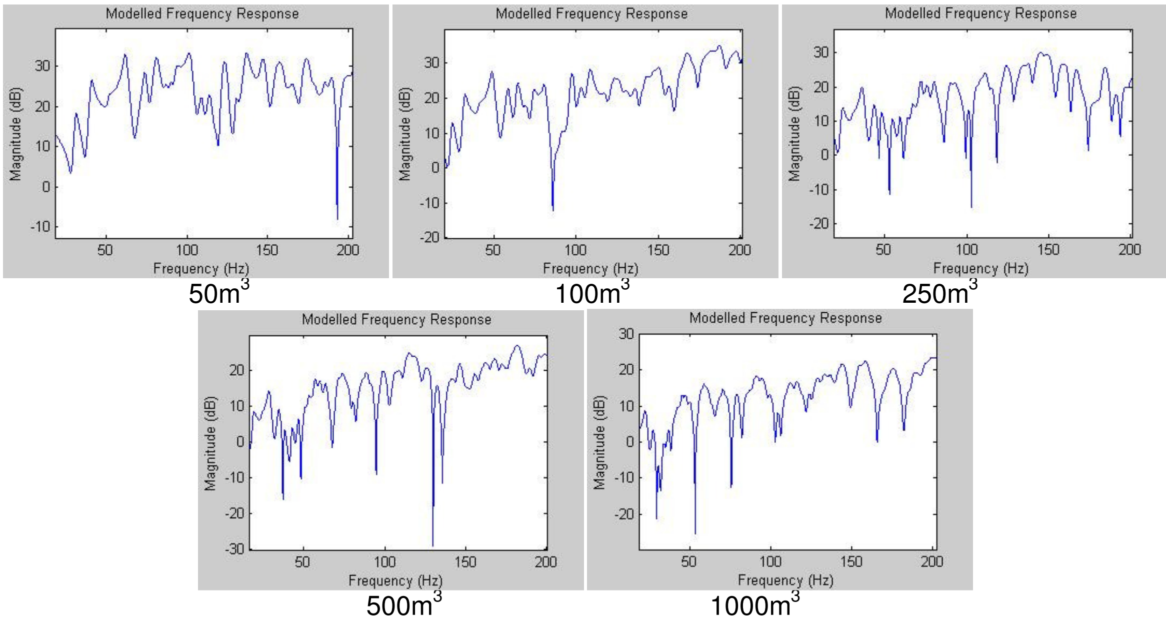


Figure 8 – Frequency Responses of the five rooms

The key issue then is the interaction of mode shapes. This interaction within a room causes a highly individual frequency response for that room and source and receiver position. The interaction continues to produce peaks and dips within the response, regardless of the density reached. This experiment shows that density itself should perhaps be disregarded as an indicator of quality. There are of course, factors other than modal density which help determine the perceived quality of a room, and in particular it should be remembered that this experiment maintains a fixed decay time across the rooms. Differing amounts of absorption, likely sources of reproduced audio and even different transient and steady state responses may all play their part in the differences within rooms of differing volumes. What is clear is that we cannot simply state that a room above a particular volume will only suffer modal problems up to a transition point based on that volume.

Further investigation is required to study our perception of the differing transfer function characteristics at low frequencies. Initial experiments also seemed to suggest a direct relationship between decay time and preference. When the decay time was modified, the preference was always that of a lower value. Finally, it is suggested that physiological factors such as the response of the ear may play a much more important role in any transition frequency and the perception of individual transfer functions at low frequency.

6 CONCLUSION

In these investigations into the perception of modal density at low frequencies, a number of results have been shown:

- Increasing the modal density does not provide a linear improvement approximating a reference case where realistic transfer functions are used.
- It is clear that an increased decay time in the low frequency region is a worse case than a shorter decay time, regardless of modal density.
- Where the decay time is kept constant, there is very little difference in perception of music within rooms of differing densities.
- Instead, the specific frequency response would appear of greatest importance when judging the quality of reproduction within a room.

These findings have a direct consequence on the validity of objective measures based upon a modal density in terms of their subjective performance. A greater understanding of the perception of individual transfer functions is required which can then be used as a basis for more generalised recommendations. This is the focus of further research.

7 REFERENCES

1. M. Schroeder, *Statistical parameters of the frequency response curves of large rooms*, J. Audio Eng. Soc., vol. 35 (5), 1987, pp. 299-305
2. B. Fazenda and M. Wankling, *Optimal Modal Spacing and Density for Critical Listening*, Proc. 125th AES Convention. San Francisco (2008)
3. R.H. Bolt, *Normal Modes of Vibration in Room Acoustics: Angular Distribution Theory*, J. Acoust. Soc. Am., vol. 11, 1939, pp. 74-79.
4. M. M. Louden, *Dimension-Ratios of Rectangular Rooms with Good Distribution of Eigentones*, Acustica, vol. 24 (5), 1971, pp. 101-104
5. O. J. Bonello, *A New Criterion for the Distribution of Normal Room Modes*, J. Audio Eng. Soc, vol. 19, pp. 597-606
6. R. H. Bolt, *Note on Normal Frequency Statistics for Rectangular Rooms*, J. Acoust. Soc. Am., vol. 18 (1), 1946, pp. 130-133
7. M.M. Taylor and C.D. Creelman, *PEST: Efficient Estimates on Probability Functions*, J. Acoust. Soc. Am., vol. 41, 1967, pp. 782-787.
8. M.M. Taylor, S.M. Forbes, and C.D. Creelman, *PEST reduces bias in forced choice psychophysics,* J. Acoust. Soc. Am., vol. 74, 1983, p.1367-1374
9. B. Fazenda, M.R. Avis, and W.J. Davies, *Perception of Modal Distribution Metrics in Critical Listening Spaces-Dependence on Room Aspect Ratios*, J. Audio Eng. Soc., vol. 53 (12), 2005, pp. 1128-1141.
10. M. Avis, B.M. Fazenda, and W.J. Davies, *Thresholds of detection for changes to the Q factor of low-frequency modes in listening environments*, J. Audio Eng. Soc., vol. 55 (7-8), 2007, pp. 611-622.
11. Y. Huang et al. *Pair-Wise Comparison Experiment on Subjective Annoyance Rating of Noise Samples With Different Frequency Spectrums but Same A-Weighted Level,* Applied Acoustics, vol. 69 (12), 2008, pp. 1205-1211