CINEMA SOUND: A NEW LOOK AT OLD CONCEPTS

Philip Newell — Acoustics Consultant, Moaña, Spain
Keith Holland — ISVR, University of Southampton, UK
Soledad Torres-Guijarro — Laboratorio Oficial de Metroloxía de Galicia, Ourense, Spain
Sergio Castro — Reflexion Arts, Vigo, Spain
Eliana Valdigem — Acoustics Engineer, Trofa, Portugal

1 THE BACKGROUND

Over the past 40 years or so it has been customary in most cinemas around the world to align the loudspeakers in accordance with the procedures set out by Dolby, amongst others, for the implementation of the X-curve, as shown in Figure 1. With the imminent advent of digital cinema, many of the old restrictions which have until now limited the bandwidths (for analogue soundtracks) and the bit-rates (for digital soundtracks) will no longer exist. With this in mind, a considerable number of people within the film industry have been asking for this opportunity to be taken advantage of in order to review the whole concept of cinema sound standards.

The X-curve, so called because of its principally experimental origins, was arrived at somewhat empirically. It was based on listening tests making comparisons between loudspeakers with a flat response, positioned in the close field, and conventional cinema loudspeakers in the far field in conventional cinema acoustics [1]. The results which the application of the X-curve yielded were said to give a good compatibility between the two sources, as a flat loudspeaker response at distances of ten or twenty metres was deemed to sound excessively bright, due to psychoacoustic phenomena. However, there can be some flexibility in the application of the X-curve, for example, cinema studio consultants may use it perhaps more as a guide, and make the final assessments in each room by the use of critical test films and trained ears, but in commercial cinemas, where no such expertise is usually available, the application can be rather more rigid.

The initial tests, leading to the adoption of the X-curve, took place in the early 1970s, coincident with the advent of commercially viable one-third-octave equalisers and their corresponding 'real time' analysers. Almost throughout the entire world of professional sound, these were seen as a radical new means of achieving uniform responses in rooms. Despite being severely flawed, the application of this concept spread like wild-fire, largely due to a great lack of understanding of both acoustics and psychoacoustics by the majority of its practitioners. In fact, in no conventional system can a 'correction' to the amplitude of the frequency response compensate for the responses of rooms, but the response curves of the real-time analysers were beguiling.

By the late 1970s, due to the frequent non-correlation of what the analysers were displaying and what the ears of the recording engineers were hearing, the use of such equalisation on music-recording-studio monitors began to be called into question. By the early 1980s, in the music industry, the whole concept of third-octave equalisation was often being flatly rejected. Nevertheless, in the live sound, discotheque and cinema worlds, the use of third-octave equalisation continued to be almost de rigueur.

There is however one great difference between the application of equalisation to cinema sound and its use in theatres and discotheques. In the latter cases there is always somebody in control who can adjust levels or equalisation as they, or the audience, may wish, but in cinemas, if the intelligibility is lacking for the all-important dialogue, then nothing can be done about it. The sound is fixed at the time of making the final mix, and in a complex soundtrack the individual levels and equalisations can be critical. For this reason it has long been necessary to try to standardise both the mixing and the public exhibition responses as closely as possible. From the confines of the dubbing theatres (mixing studios) it needs to be known what the responses in the reproduction environment will be. However, this standardisation of responses is rarely satisfactorily achieved.

Over the years, the cinema industry has tried to find ways of calibrating and standardising the rooms, but always with an eye on keeping the costs within commercially acceptable levels. They have moved from single microphone measurements to multiple, multiplexed arrays, and on to other means of achieving better spacial averaging for the measurements. However, spacial averaging is largely a statistical exercise. The fact that the average response is within limits does not in itself mean that the response is satisfactory over the whole area in question, or even a part of it. What is more, the fact remains that the application of equalisation to the loudspeakers in an attempt to correct a far-field response in a large room over a wide listening area is a non-starter. The whole current concept of cinema equalisation is considered by many to be grossly flawed.

There are several disturbing reasons for rejecting the current state of affairs. For example, it is well-known that the excess-phase (non-minimum-phase) part of a room response gives rise to the fact that third-octave equalisation cannot correct simultaneously the steady-state and transient components of a signal. The improvement of either one may lead to the degrading of the other. Any applied 'correction' to the steady state response, as is the current norm, will inevitably linearly distort the direct sound from the loudspeakers, and may significantly damage the transient responses. This is also a great irony because, despite the very important contribution which the direct sound makes to the naturalness of the sound, and which can also have a great bearing on the intelligibility of dialogue, the steady-state response correction (which can at best only correct the response in certain, spacially-dependent areas of the room) will distort the very important direct response at *all* places in the room. This seems to be a problem requiring serious further consideration.

As the first pass of the ear of the direct sound is important, it would seem self-evident that its integrity should be preserved as far as possible. Yet, despite this fact, the current alignment procedures lead to the distortion of both the amplitude *and* phase responses of the direct sounds, and this process is being perpetuated in both dubbing theatres and commercial cinemas the world over. The following sections of this paper will look at the extent to which this distortion is being currently applied, and will also look at ways to improve the situation.

2 ACTUAL RESPONSES OF REAL ROOMS

There is currently much debate taking place within the cinema industry about the relevance of the X-curve to modern cinemas, and especially regarding the compatibility of soundtracks with domestic format releases. Meta-data and automatic source detection are also on the agendas. However, these matters are outside the scope of this paper. Nevertheless, no matter whatever response curve be finally chosen as a standard, the fact remains that this response should be delivered to the cinema-going public as faithfully as possible. The findings in the following sections of this paper indicate that this is far from occurring to the desired degree.

In order to assess the current state of affairs, the responses of 20 Dolby Digital and Dolby Surround rooms, both studios and commercial cinemas, were measured at the standard calibration distance of two-thirds of the distance from the screen to the rear wall, and also at a distance of 2 metres from the central loudspeakers. The two-thirds distance plots, measured in one-third octave bands, are shown in Figure 2. Although these were single microphone measurements, it can be seen that the great majority of them at least approximately conform to the X-curve, signifying that some attempt has been made to standardise these rooms. However, Figure 3 shows the responses when measured at 2m from the screens. From these it can be seen that they are not in fact as similar as one may expect from inspection of the two-thirds distance plots. Figure 4 shows the higher-resolution, truncated responses of the 2m plots (essentially the direct sound from the loudspeakers, before the arrival of any room reflexions, truncated at 10ms from the first arrival). It can clearly be seen that no two responses are alike, nor barely even similar.

Given what was said in Section 1 about the need for the direct sound to be accurate if a natural sound were to be expected, the state of affairs depicted in Figures 3 and 4 would suggest that if the direct responses are as different from each other as they are shown to be, then there can be little hope of sonic compatibility from room to room. It is way beyond the realms of reason to expect that the transfer functions via the highly complex acoustic responses of the rooms could somehow transform these very dissimilar two-metre responses into one, uniform (or even reasonably similar) response at the two-thirds distance position (from the screen to the back wall) in each room. Essentially, if the direct signals are not similar to begin with, then no third-octave room equalisation is going to be able to correct the situation deep into a room.

3 LIMITATIONS OF THE MEASUREMENTS

As stated in Section 1, it is recommended that the measurements for the in-room calibrations are made by some sort of spacial averaging, but the measurements made for the purposes of this paper were made by single microphones. It is probable that the spacially averaged responses at the two-thirds distances would be closer to the X-curve if such averaging had been carried out, but the confirmation of the accuracy of the calibration was not the objective of this study. The prime aim was to measure the degree of linear distortion being applied to the direct signals in an attempt to standardise the far-field responses.

Some modal activity is clearly visible in the plots, but almost all of the measured 'Listening Position' responses show an acceptable tendency towards the X-curve, and as they were all done in Dolby calibrated rooms it is not unreasonable to expect that all the calibrations were within normally encountered tolerances. What is more, the authors of the paper are aware that in many commercial cinema calibration processes, a single microphone is, in fact, still commonly used. The principle purpose of Figure 2 is to confirm that no rooms were wildly out of calibration, although in two cases the calibration can be seen to be very dubious.

It was decided to take the close-field, third-octave measurements at two metres as a means of capturing predominantly the direct sound. Some advisers had suggested gating the responses and measuring from a greater distance, such as three or four metres, to ensure that there were fewer possibilities of suffering geometrical errors due to the different distances to the drivers, but in no cases were the loudspeakers measured for this paper inordinately large. It was also considered that despite the inevitable influences of the rooms, even on the two-metre measurements, if the differences that were suspected to exist were as large as they have been shown to be, then the steady-state, two-metre measurements were still valid for the purposes of this paper. The axis of measurement in each case was the result of an experienced compromise judgement, well within the coverage of the principal axis of the mid-range drivers, and equidistant from any multiple woofers (if applicable). The authors do not consider that any significant variations between the responses shown in Figures 2 and 3 were due to the proximity of the measuring microphones to the loudspeakers.

None the less, the gated responses have been incorporated into this paper, to help to give a more precise view of any upper bass and mid frequency anomalies, but the gating has rendered the frequencies below around 100 Hz unable to be measured.

4 SHORT COMMENTS ON THE MEASUREMENTS FOR EACH ROOM

Figure 5 shows the 2m truncated, 2m third-octave, and two-thirds-distance third-octave plots, one above the other, for each public cinema room. Each set will now be briefly discussed in turn. Cine 1 to 9 are the public cinemas; Dubb 1 to 11 are dubbing theatres (sound-stages / mixing-theatres). Figure 6 shows the equivalent plots for the dubbing theatres.

Cine1 The two-thirds-distance (Listening Position) plot shows a reasonable approximation to the X-curve, except for a boost around 200Hz The lower two bands are quite high.

The 2m plot shows a large peak centred on 160Hz, which could be an indication of trying to correct an 'uncorrectable' floor-reflexion dip around 125Hz. The relative levels of the two lowest bands are around 20dB lower that at the two-thirds-distance position.

The truncated plot is rather erratic, not only showing the boost at 160Hz but also a ragged mid-range response which is almost certainly a result of a large distance between the screen and the hard-fronted loudspeakers.

Cine2 The two-thirds-distance plot shows peaks of 3 to 5dB around 90, 250, 500, 800 and 1600Hz. These are mostly also apparent in the 2m plots. All of the three plots for this room are 'lumpy', and would not suggest that a very natural sound would be evident anywhere in the room. A relative bass rise exists at the two-thirds-distance position, compared with the 2m plot, but it is not as exaggerated as in Cine1.

Cine3 In this room, the responses at both 2m and the two-thirds distance are somewhat more similar to each other than in either Cine1 or Cine2. Even the low bass levels are proportional, but there is evidence of a boost of around 5dB in the 63 to 100Hz bands, again perhaps to try to deal with a floor dip.

Cine4 A peak is evident in both third-octave plots in the 500Hz band, and slightly lower in the truncated response. The response below the 63Hz band collapses. All the plots are very ragged.

Cine5 The truncated response shows 12dB of variation, and is very uneven in general. Again, this is a room with a large gap between the loudspeakers and the screen. At 2m, a peak is evident in the 500Hz band, perhaps intentionally introduced to support a general trough between 250Hz and 1kHz, but the peak in the truncated response does not correspond. If the dips are reflexion induced, then no amount of boost will correct them.

Cine6 There is little evidence of the application of the X-curve to this room, and the response characteristic between the 63 and 500Hz bands is very uneven, both at 2m and at the two-thirds distance, with some commonality between the two. This seems to suggest that the equalisation was bypassed. It does happen! The sharp fall above the crossover frequency of 500Hz is an indication that the relative balance of the highs and lows has also not been adjusted.

Cine7 In this room, unlike most of the others, the lower two frequency bands are proportionately lower at the two-thirds-distance position with respect to the 2m measurement. From the 800Hz band upwards, the far-field plot is commendably flat. Both the third-octave and the truncated 2m responses show evidence of a significant dip at 500Hz.

Cine8 These responses are all well out of specification, although there is evidence of the application of the X-curve over most of the response. There is also evidence of a gain reduction above 500Hz. When questioned about this, the projectionist said that there had been many complaints from the public about the aggressive sound, so he had turned down the level of the horns. This is, unfortunately, not a totally isolated incident.

Cine9 The plots for this room are all so far out of specification that there is little useful to be said about them.

None of the above rooms, all public cinemas, could really be considered to meeting the desired standards of reproduction.

Dubb1 The representation of the X-curve at the listening position is very respectable, although this seems to have been achieved at the expense of the flatness of the sub-1kHz response of the 2m response, which exhibits a response of about +4dB between 63Hz and 1kHz.

Dubb2 For a relatively large room, around 750m^3, the X-curve at the listening position is almost exemplary, and the 2m response is also very smooth. The rise in the 100Hz band in the 2m third-octave response is not evident in the truncated response, suggesting that this may be a localised reflexion in the region of the microphone, and not a part of the direct response from the loudspeaker.

Dubb3 Other than a peak in the 125Hz band at 2m, the responses in this room are commendable; generally +2dB at 2m and +3dB at the listening position. The peak at 125Hz seems to be an attempted correction for a floor dip, but the correction is perhaps subjectively worse than the problem.

Dubb4 The 2m plots are quite smooth, both in the truncated and third-octave form. The dip centred around 1.6kHz is evident in all the plots, but no reason for this is apparent. There is a broad response dip centred around 125Hz at the listening position, probably a reflexion from the floor, with some indication of an attempt to correct it in the 2m third-octave plots.

Dubb5 At 600Hz there is a boost at 2m which corresponds with a dip at the listening position. This seems to be evidence of an attempted correction, as the boost is also evident in the truncated response. The 125 to 250Hz bands do not correspond well at the two distances of measurement.

Dubb6 Although the listening position response is very good, the 2m response is quite poor, suggesting that it has been heavily equalised in order to try to achieve the response goal at the listening position. The truncated response is well in agreement with the 2m third-octave response, which suggests that the latter is not suffering from localised acoustic disturbances.

Dubb7 This room exhibits some characteristics which are the opposite of Dubb6, in that the 2m response is generally much better than the listening position response. The 400Hz range can be seen to be lacking on all the plots.

Dubb8 All the responses are respectable, but there is evidence of dips around 800Hz in the 2m plots which do not seem to be necessary from an equalisation point of view at the listening position.

Dubb9 There is no evidence of the application of the X-curve in any of these measurements. However, other than some sign of a resonance in the 250 and 315Hz bands at the listening position, the response is, overall, about +3dB. No truncated response was available for this room.

Dubb10 Both the 2m and the listening position responses for this room are respectable. The 80Hz dip at the listening position is probably a reflexion-induced dip. Again, no truncated response was available for this room.

Dubb11 Except for a dip in the 400 and 500Hz bands, the listening position response is excellent. The 2m plots show some evidence of a peak around 250Hz, but it is much more evident on the third-octave plot than on the truncated plot, suggesting that it could be a localised reflexion in the vicinity of the microphone. Nevertheless, the 2m plots are generally smooth.

5 SOME CONSIDERATIONS

The poor overall similarity of the 2m and two-thirds-distance responses implies that the frequency balances of the transient and more continuous sounds will not be the same. Where the truncated (gated) responses and the third-octave responses at 2m do not correspond with each other, the implication is that there is some localised effect resulting from reflexions or resonances in the vicinity of the microphone. The resonances and reflexions due to the cavities between the screens and the loudspeakers are considered to be part of the loudspeaker responses.

6 EQUALISATION-INDUCED DRIVER DISTORTION POTENTIAL

In some of the two metre measurements there is evidence of considerable equalisation boosts at low frequencies. As many of the loudspeakers in the rooms (especially the public cinema rooms) are already marginal in terms of their output capability, the applied boosts may be seriously risking distortion and driver damage.

7 DISCUSSION OF RESULTS

Figure 7 shows the overlaid plots of the dubbing theatres at both the two metre and the listening positions. If anything, overall, the listening position (far-field) plots are slightly more closely grouped than those at the 2m distance. This is not entirely unexpected, as the listening positions were the reference positions for the room calibrations, but the situation is the reverse of what would be desirable for accurate monitoring. The transient responses, which are considered by many professionals to relate closely to the naturalness, and hence standardisation of a soundtrack, must evidently be wildly different over the spread of the rooms measured, due to the fact that the 2m measurement are so different. Such differences in the direct sounds from room to room are not a good sign of compatibility.

Figure 8 shows the overlaid plots for the public cinemas. In these cases, the 2m (close-field) plots are clearly more tightly grouped that the far-field plots, which is only to be expected in rooms with less acoustic control. However, the degree of variability at both distances is alarming in terms of acoustic standardisation. Many of the responses are way outside the tolerance limits, shown dashed on some plots.

8 CONCLUSIONS

What seems to be evident from this study is that the standard X-curve for the equalisation of cinema soundtracks is far from being met in many public cinemas. Even in the more tightly controlled dubbing theatres (mixing studios) there is significant evidence of out of tolerance responses at low frequencies. The most disquieting part of the findings is that the far-field responses in the rooms seem to be no less accurate than the close-field responses. Normally, one would always expect a degradation in the smoothness of a pressure amplitude (frequency) response as one becomes more distant from the source in anything other than an anechoic chamber (free field). However, it seems that in order to try to do the impossible - flattening a room response via third-octave equalisation - the direct responses from the loudspeakers in the measured rooms have been linearly distorted beyond reason.

There is a certain logic in the belief that if the sounds do not come out of the loudspeakers with similar responses, there is little hope of them arriving at the listeners' ears with similar responses. Figures 5 and 6 are very informative in that nowadays, loudspeaker technology has arrive at a level where, if the need to provide the X-curve roll-offs at the frequency extremes were simply to be applied by means of parametric equalisation in the close field, then the responses at the 2m positions could be expected to be greatly more similar to each other than the travesties shown in the previously mentioned figures. These days, the loudspeakers, themselves, can be expected to be respectably flat. Whether the X-curve itself is a travesty, is perhaps outside of the scope of this paper, but it seems to be clear that there is ample justification for a thorough reassessment of the current cinema alignment procedures.

9 FURTHER EXPERIMENTATION

In the near future, work will begin on STI analysis of the captured responses, and auralisations will also be performed. From these it will be easier to assess the true degree of perceived sonic differences from room to room, and any effects which they may have on the intelligibility of the dialogue in soundtracks.

10 ACKNOWLEDGEMENTS

Thanks are due to Matt Desborough (Dolby UK), Branko Neskov (P&B, Lisbon, Portugal), David Sans (Soundub, Barcelona, Spain), Christian Beusch (Tonstudio Beusch, Zurich, Switzerland), and also to the NFTS in the UK, for their assistance in gathering many of the room responses and for sharing their experiences in discussions relating to the subject of the paper.

Special thanks are also due to Floyd Toole, for his long correspondence on the subject, his helpful comments on the analysis of the data, and his help with the preparation of some of the artwork.

11 REFERENCES

[1] Allen, Ioan; 'The X-Curve: Its Origins and History', SMPTE Journal, Vol. 115, Nos 7 & 8, pp. 264-275 (2006)

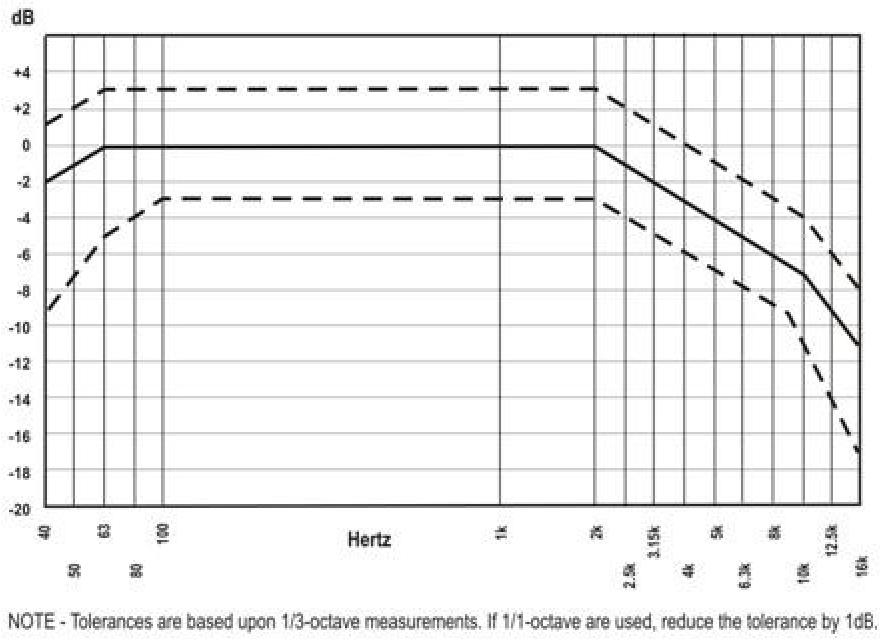

Figure 1 The X-Curve

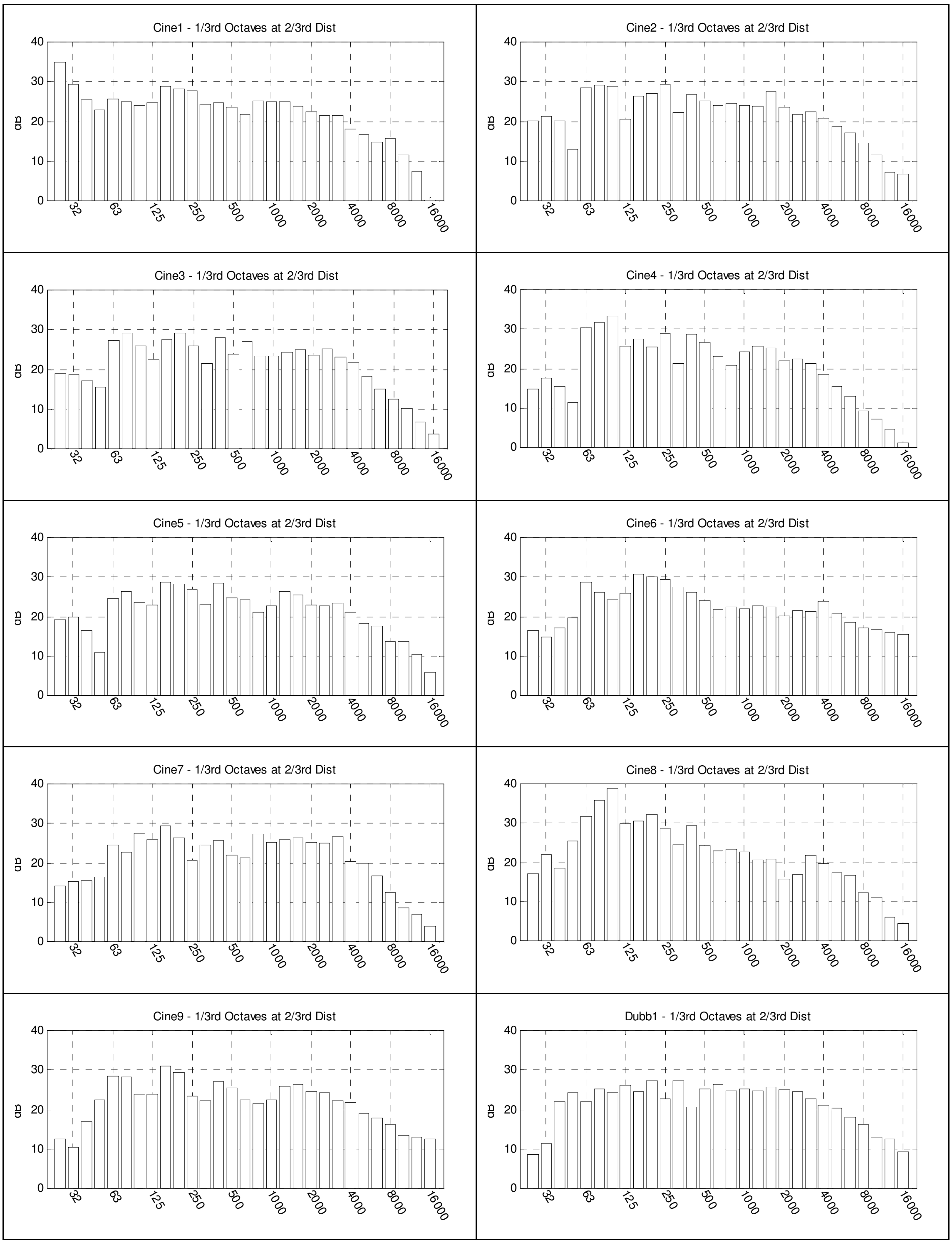

Figure 2a Third-Octave Responses at 2/3rd Distance from Screen to Back-wall: Room#1 to #10

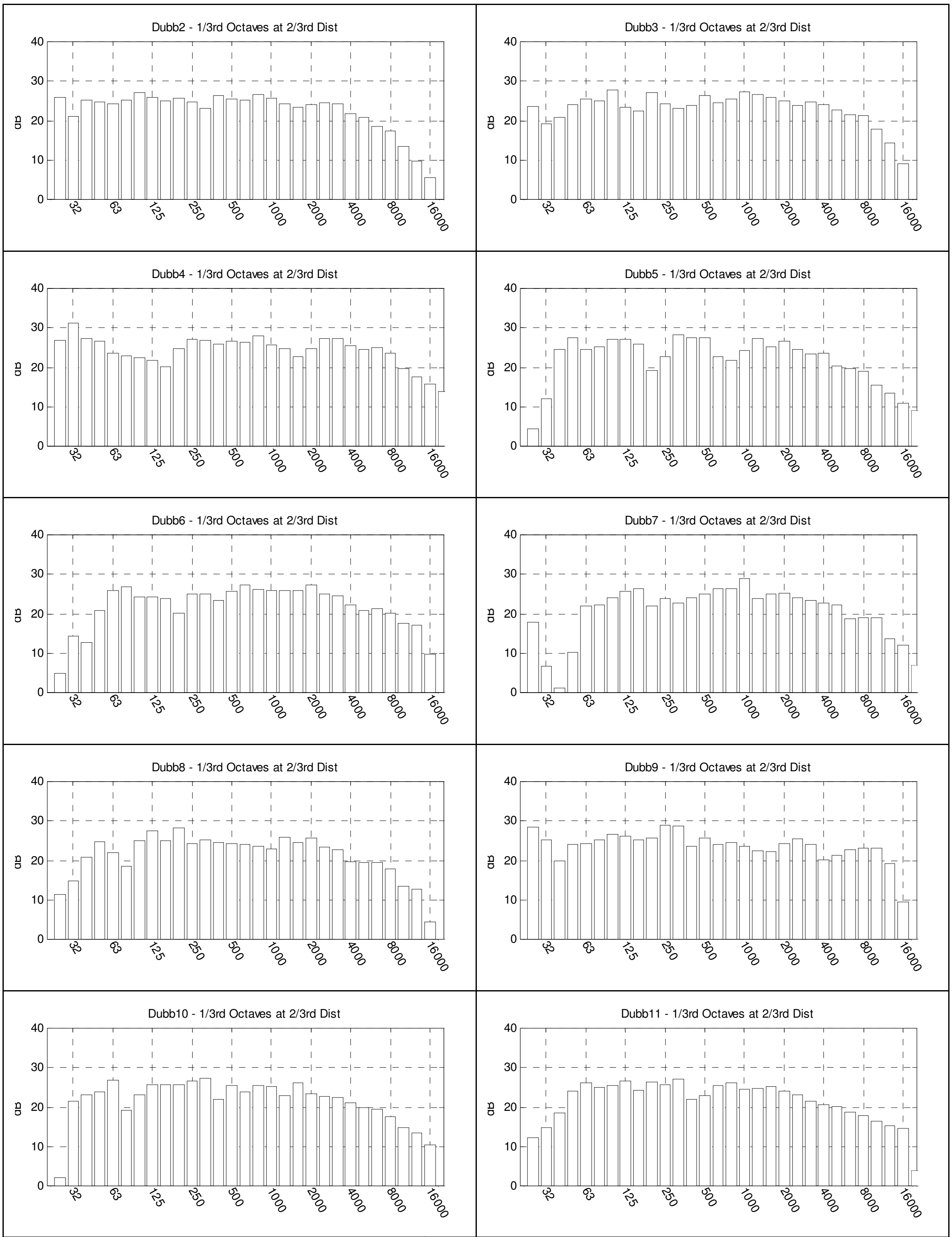

Figure 2b Third-Octave Responses at 2/3rd Distance from Screen to Back-wall: Room#11 to #20

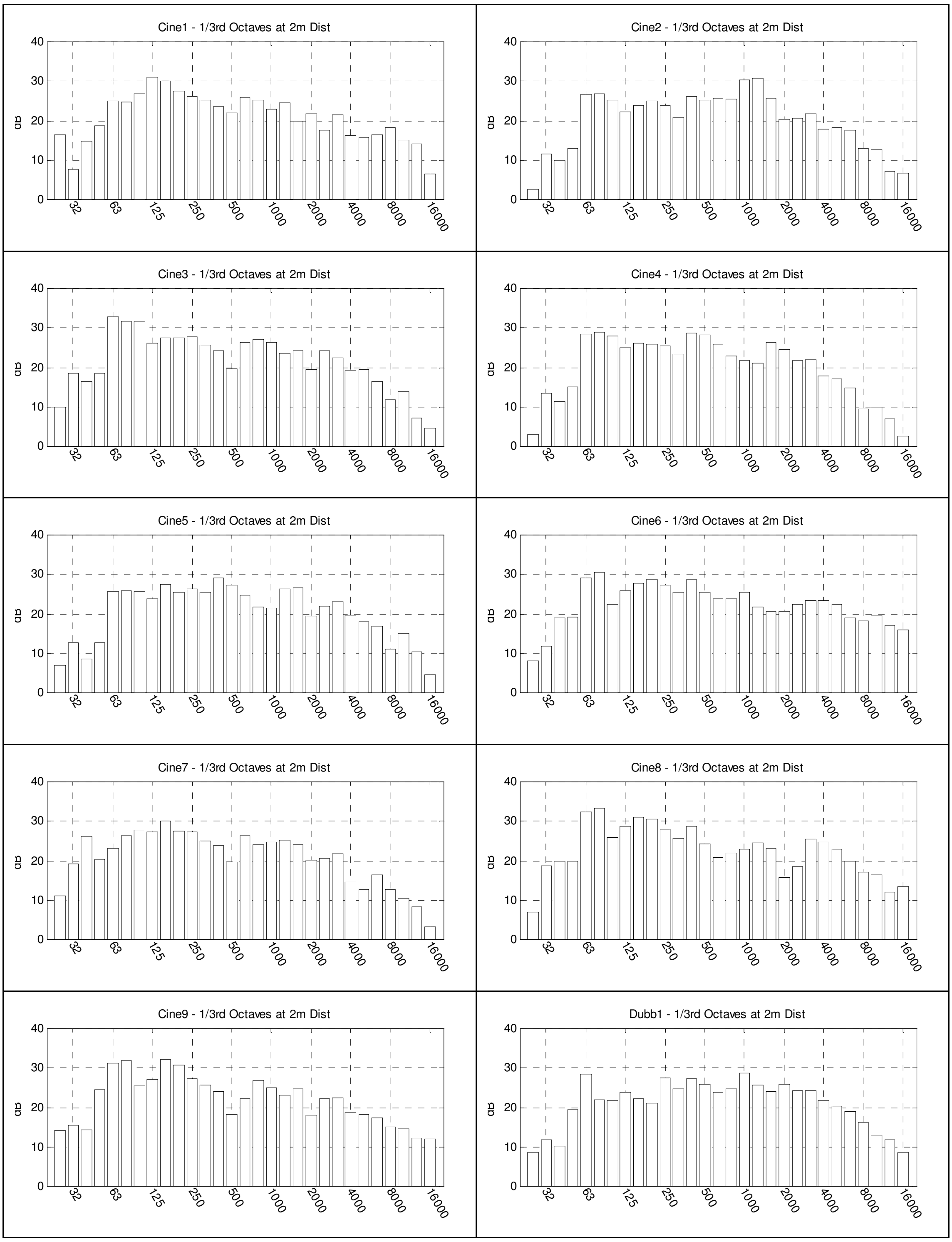

Figure 3a Third-Octave Responses at 2m Distance from Screen: Room#1 to #10

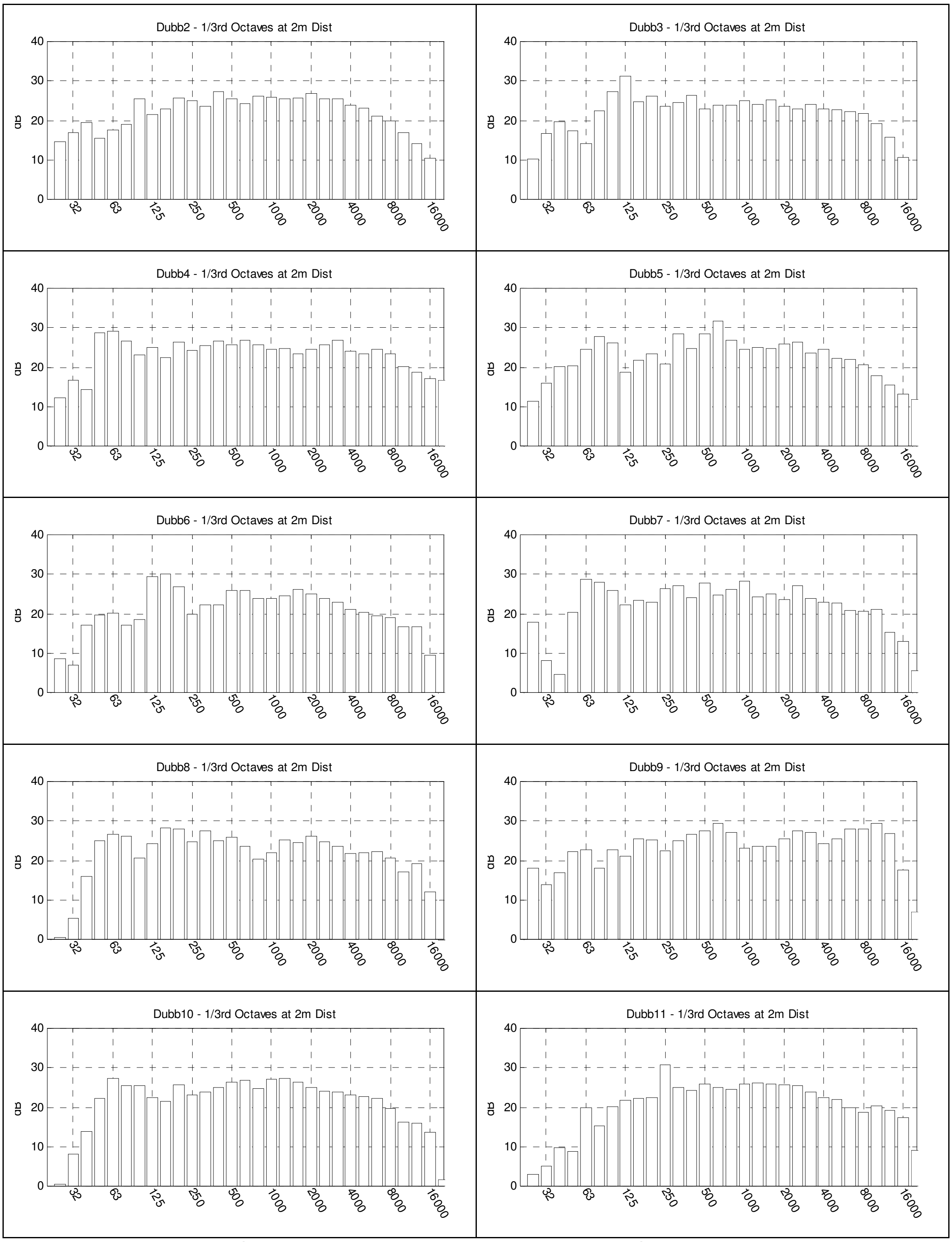

Figure 3b Third-Octave Responses at 2m Distance from Screen: Room#11 to #20

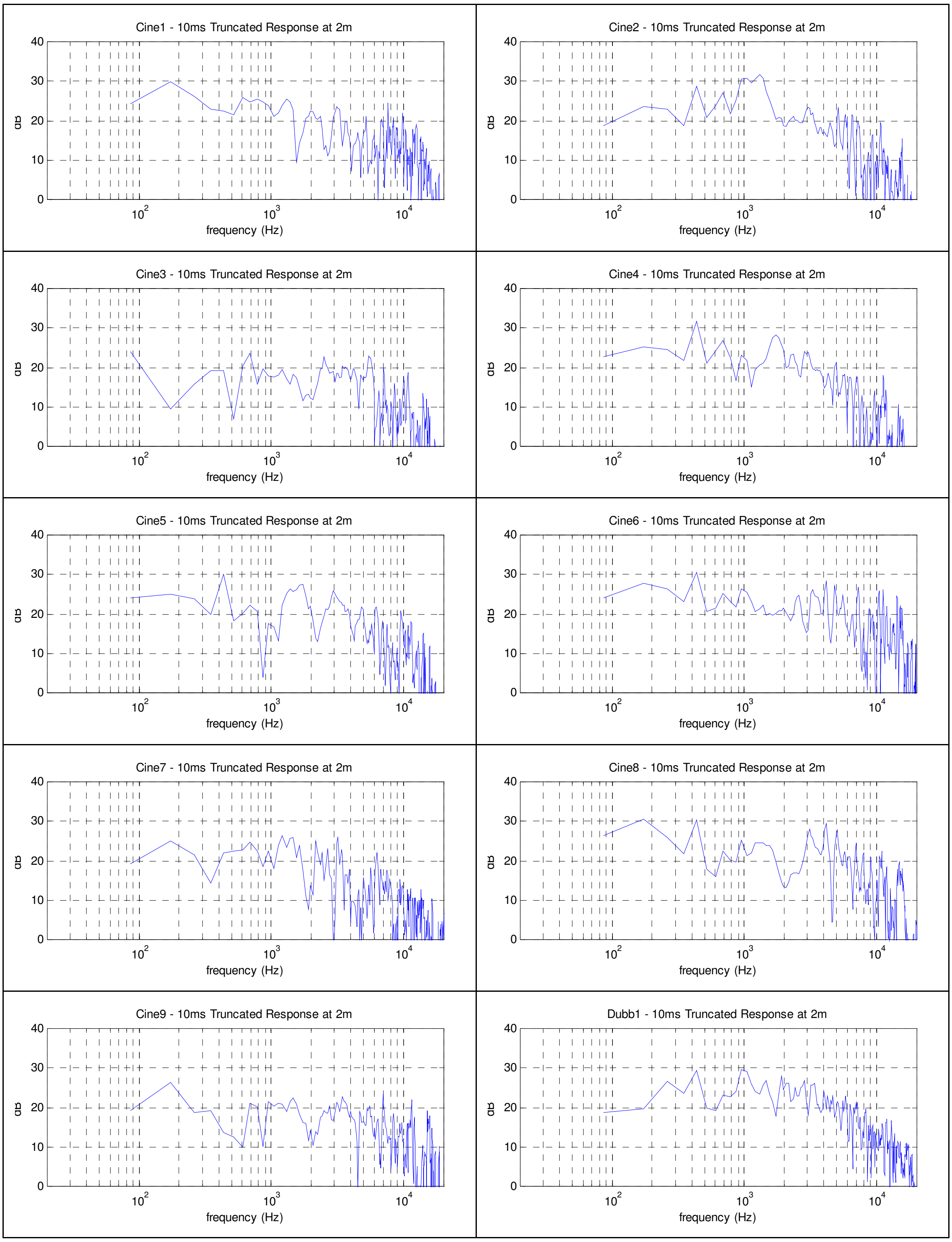

Figure 4a Narrowband 10ms Truncated Responses at 2m Distance from Screen: Room#1 to Room#10

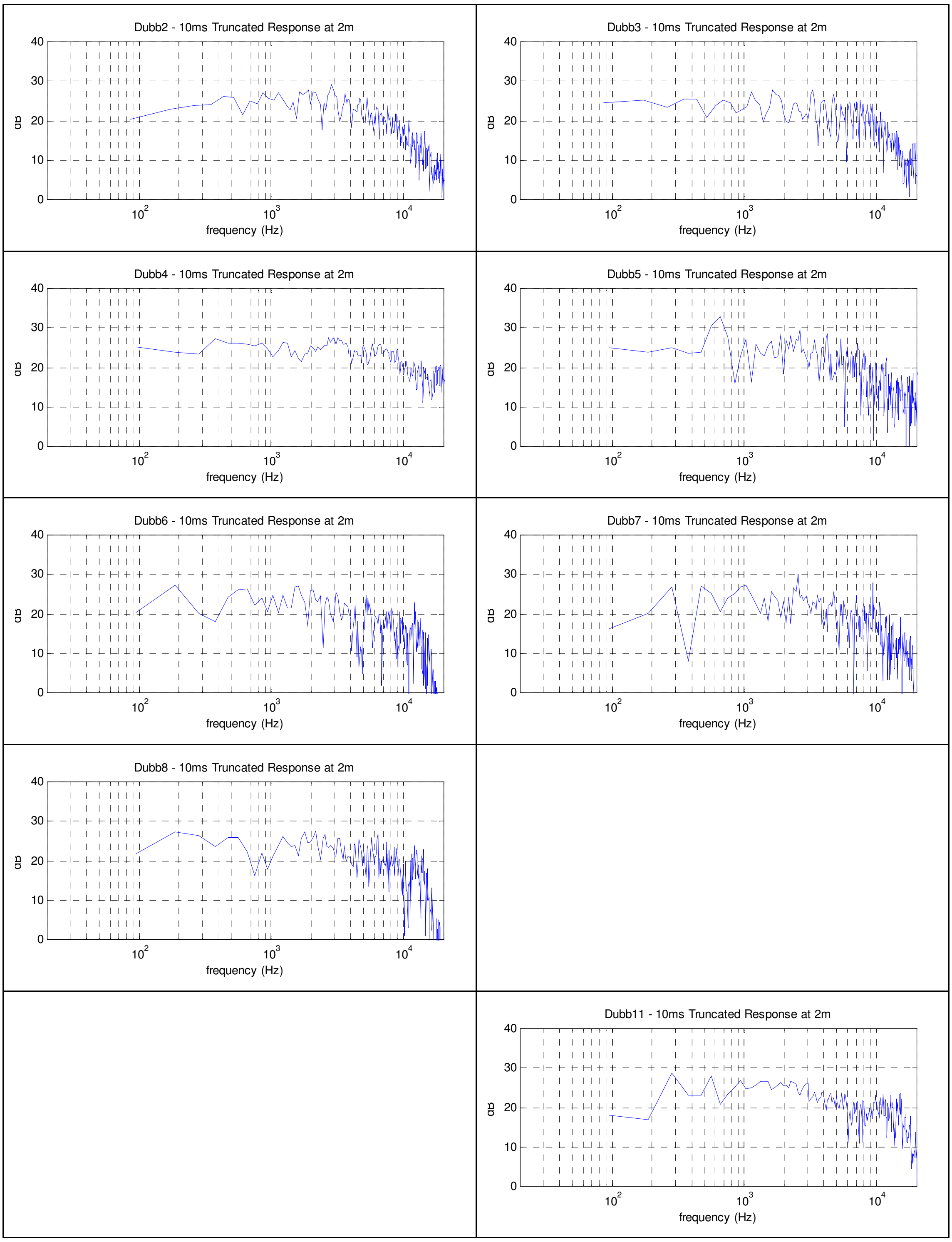

Figure 4b Narrowband 10ms Truncated Responses at 2m Distance from Screen: Room#11 to Room#20

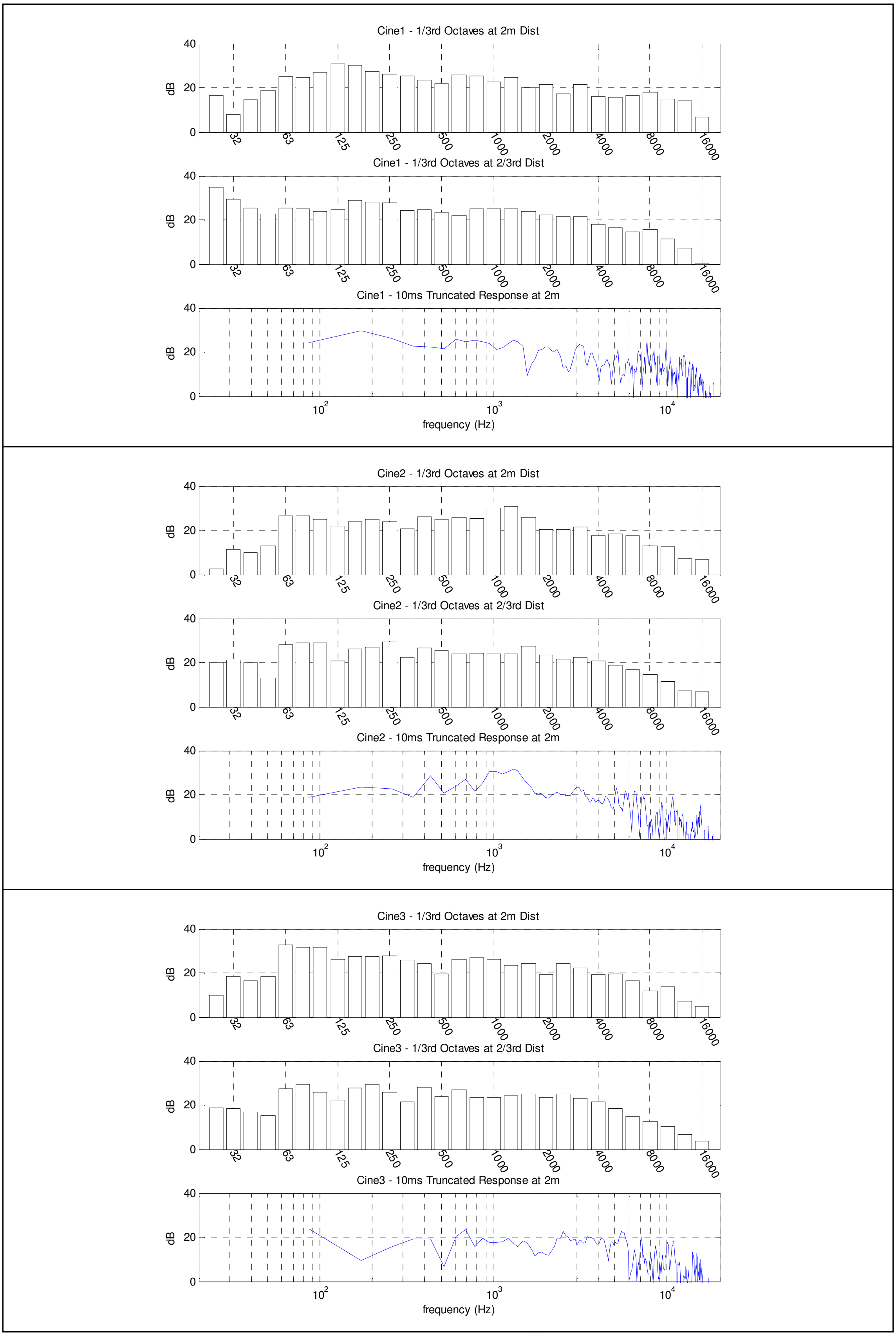

Figure 5a Responses by Room of Public Cinema Rooms #1 to #3

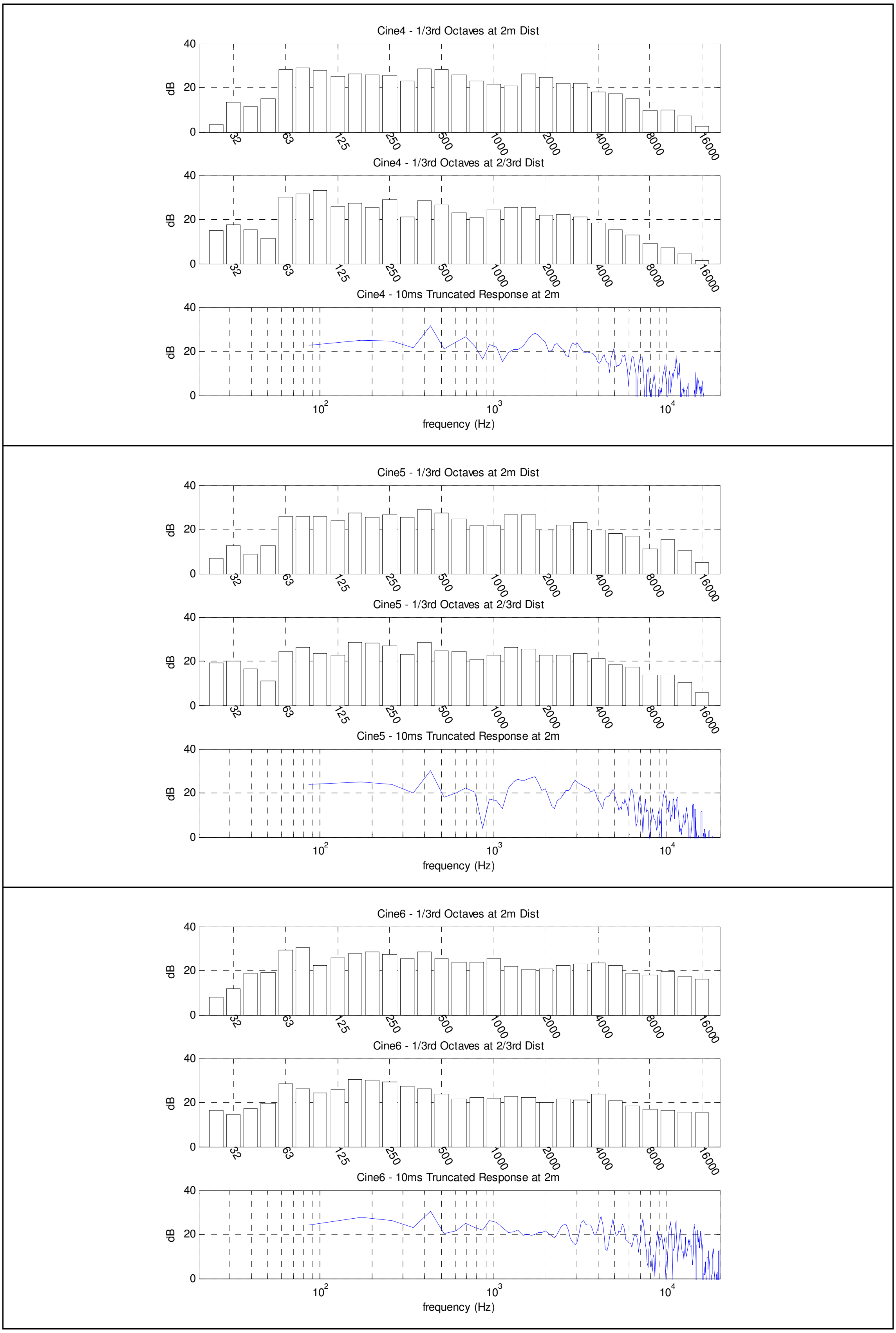

Figure 5b Responses by Room of Public Cinema Rooms #4 to #6

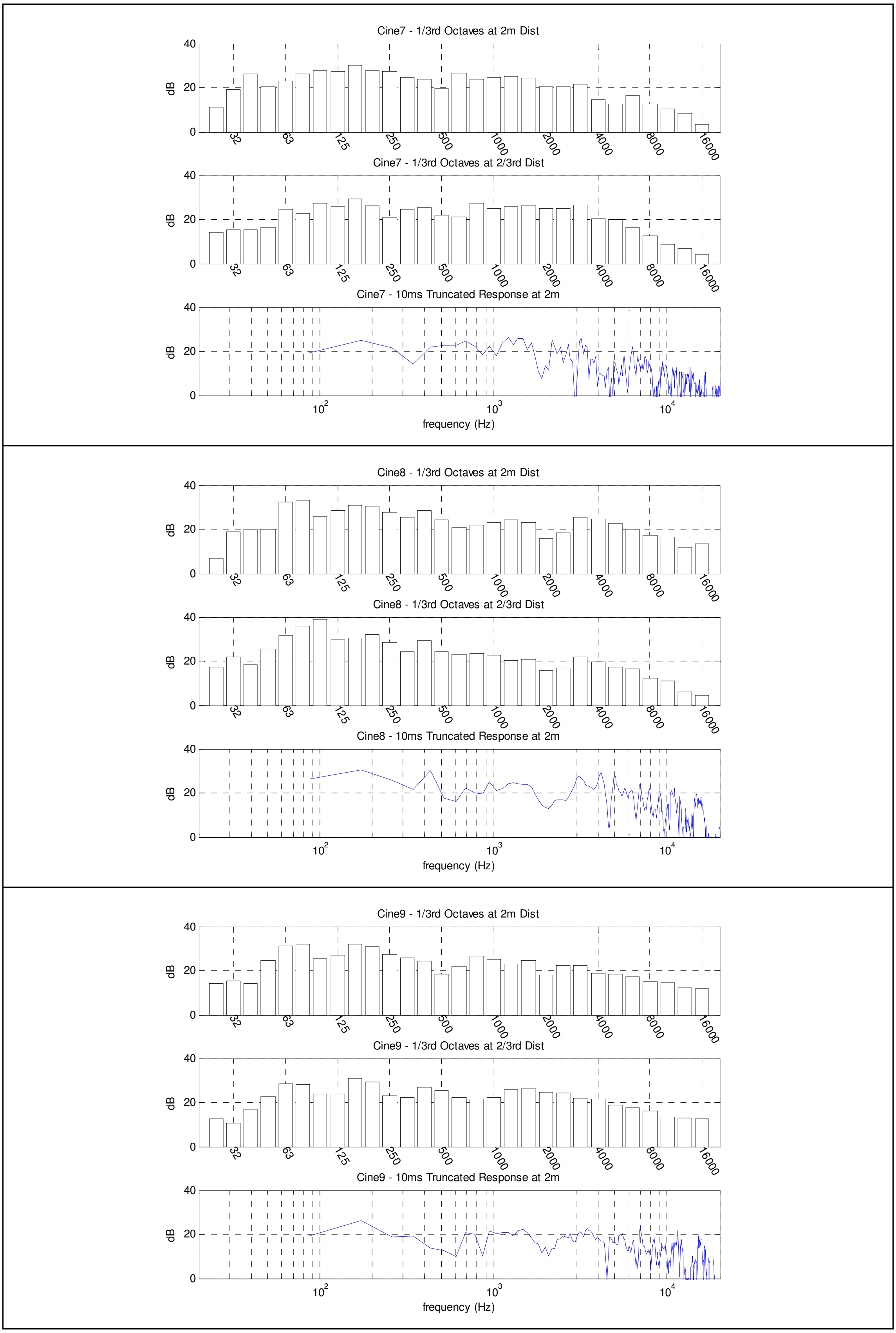

Figure 5c Responses by Room of Public Cinema Rooms #7 to #9

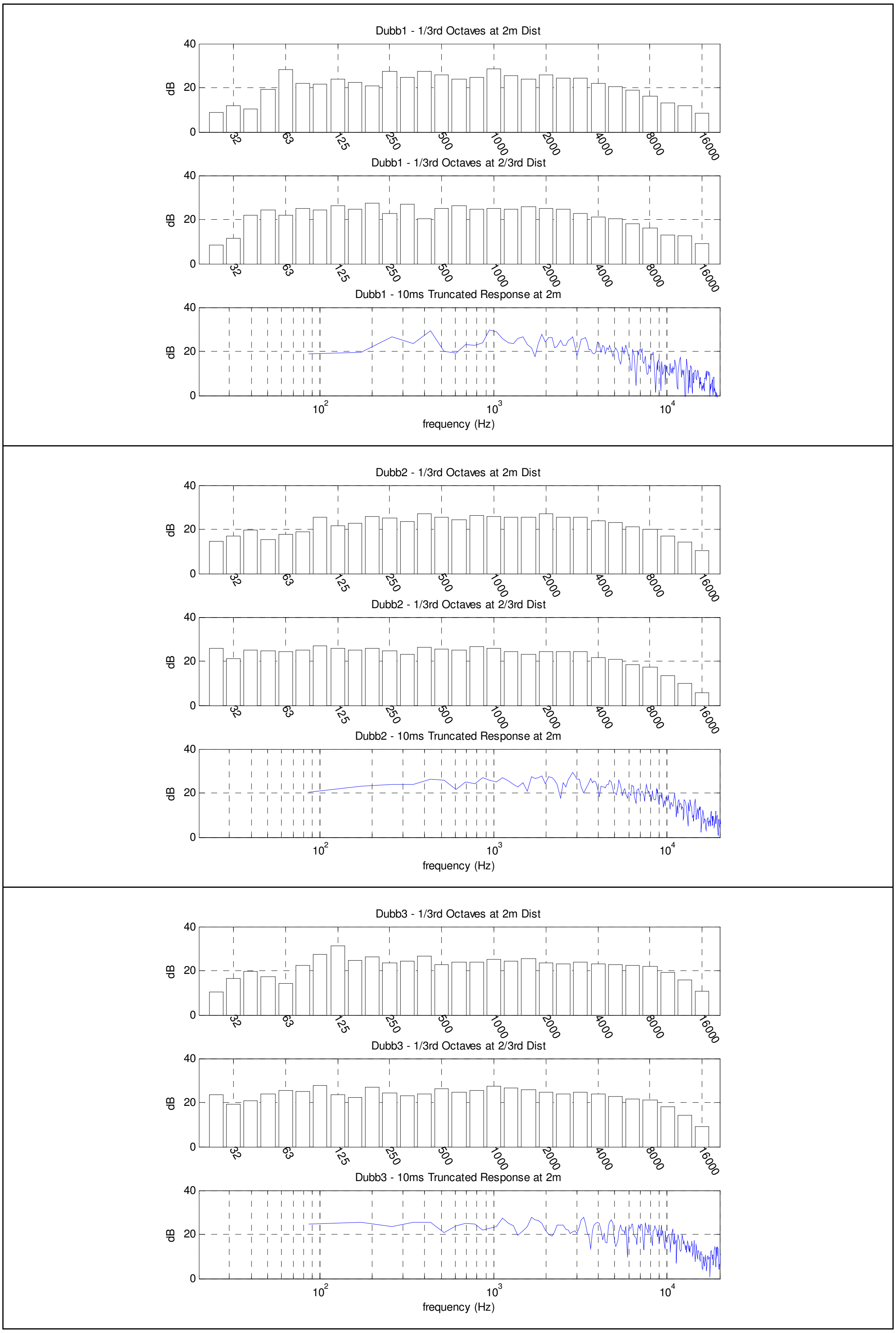

Figure 6a Responses by Room of Dubbing Theatres #1 to #3

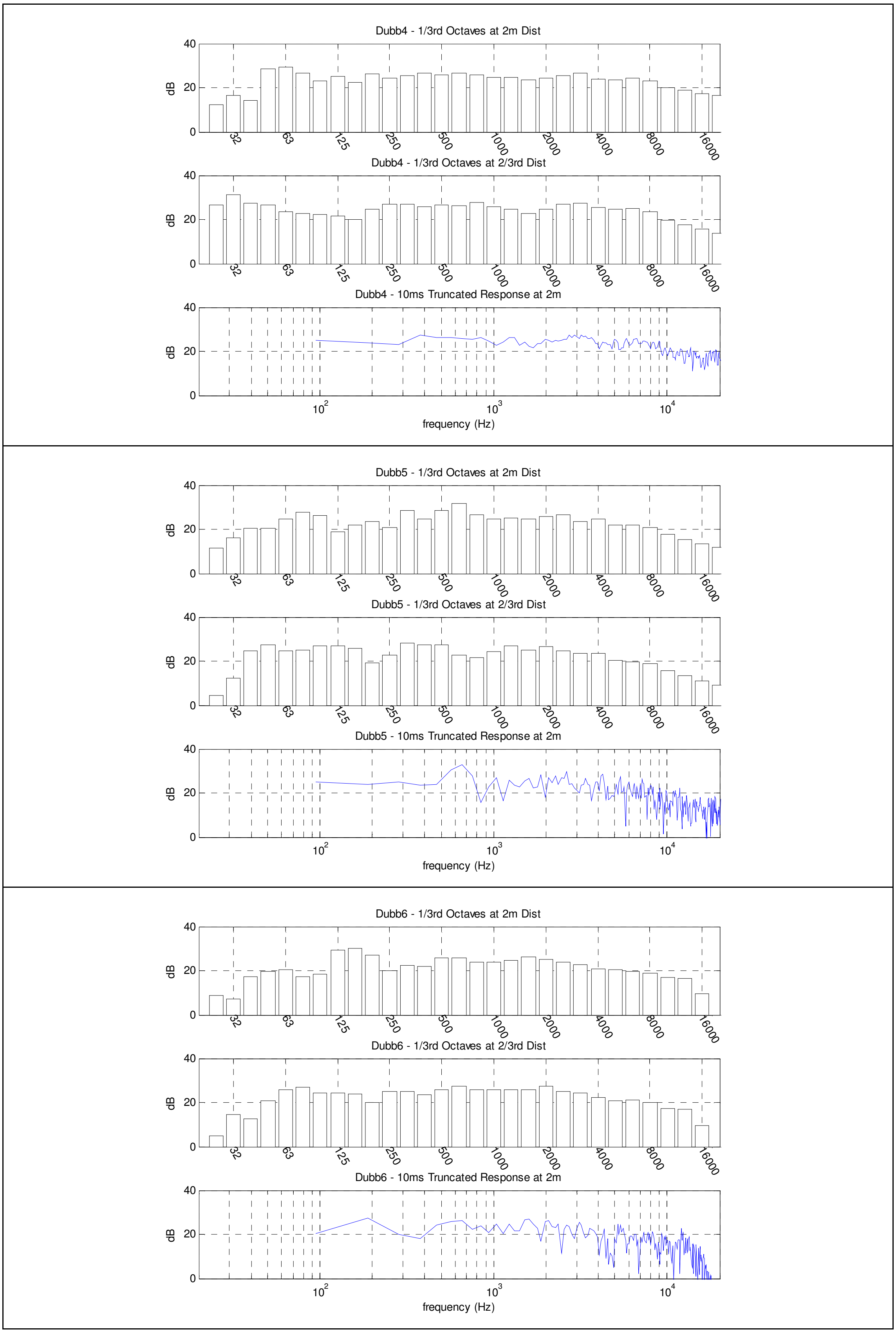

Figure 6b Responses by Room of Dubbing Theatres #4 to #6

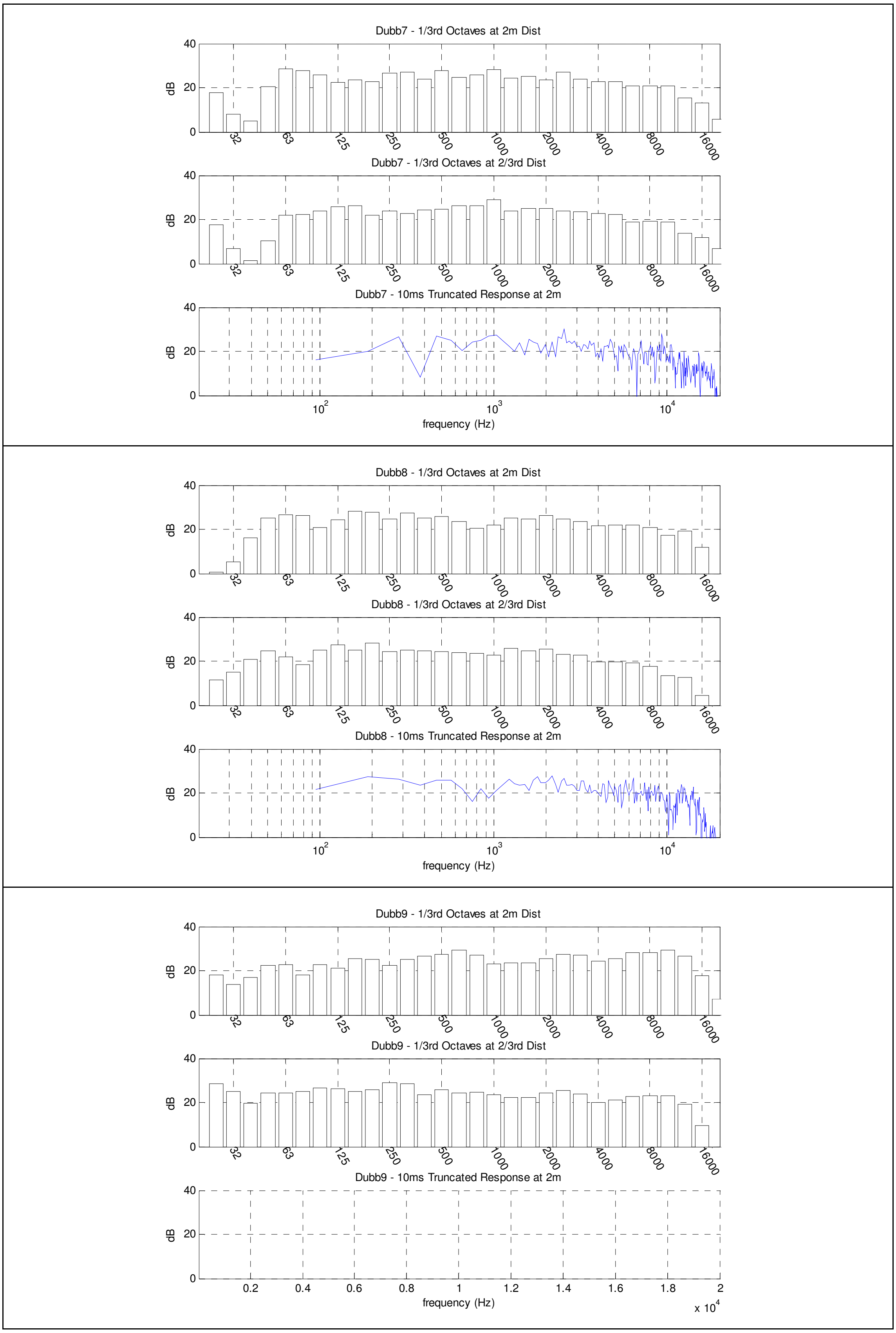

Figure 6c Responses by Room of Dubbing Theatres #7 to #9

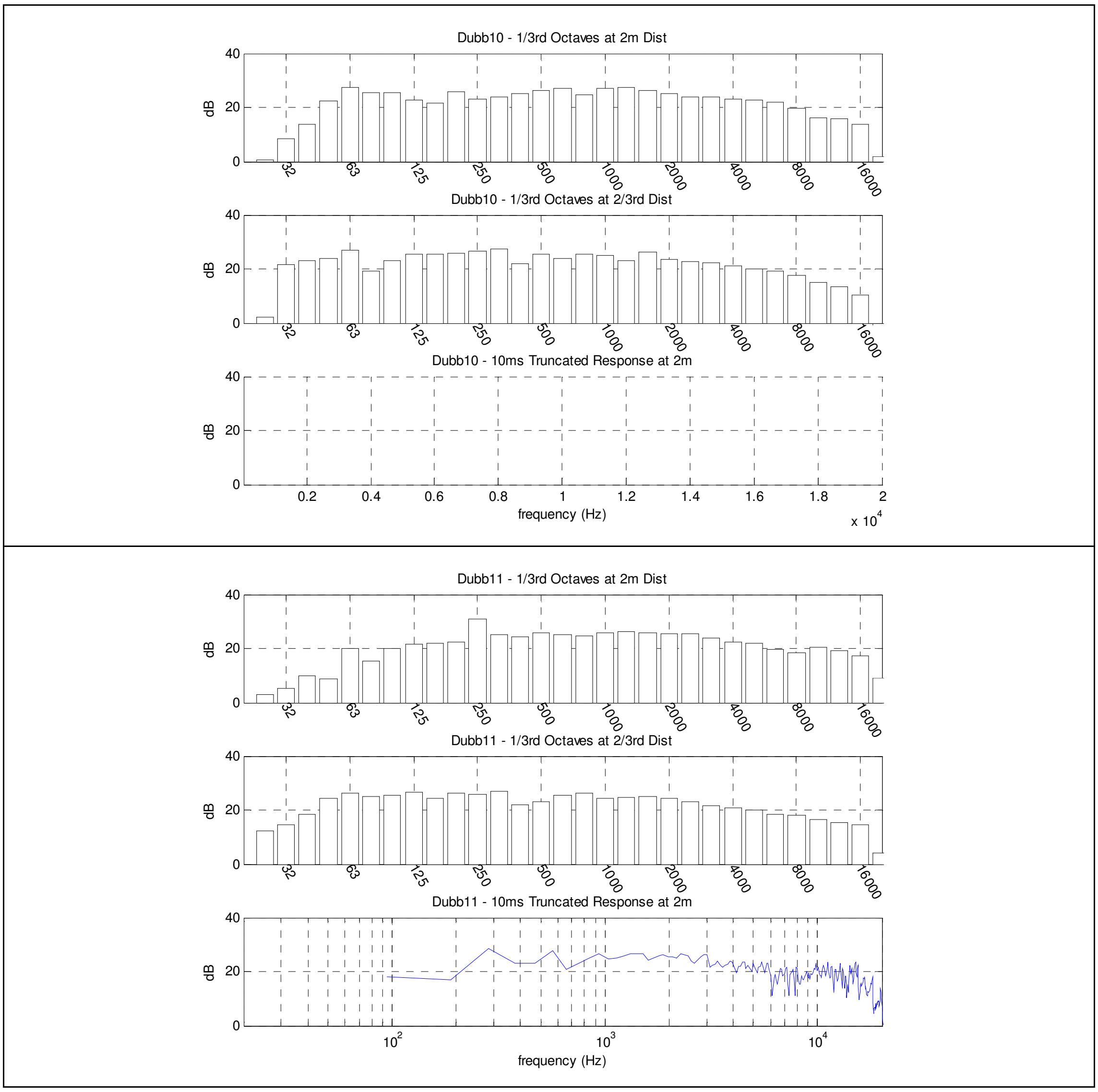

Figure 6d Responses by Room of Dubbing Theatre #10 to #11

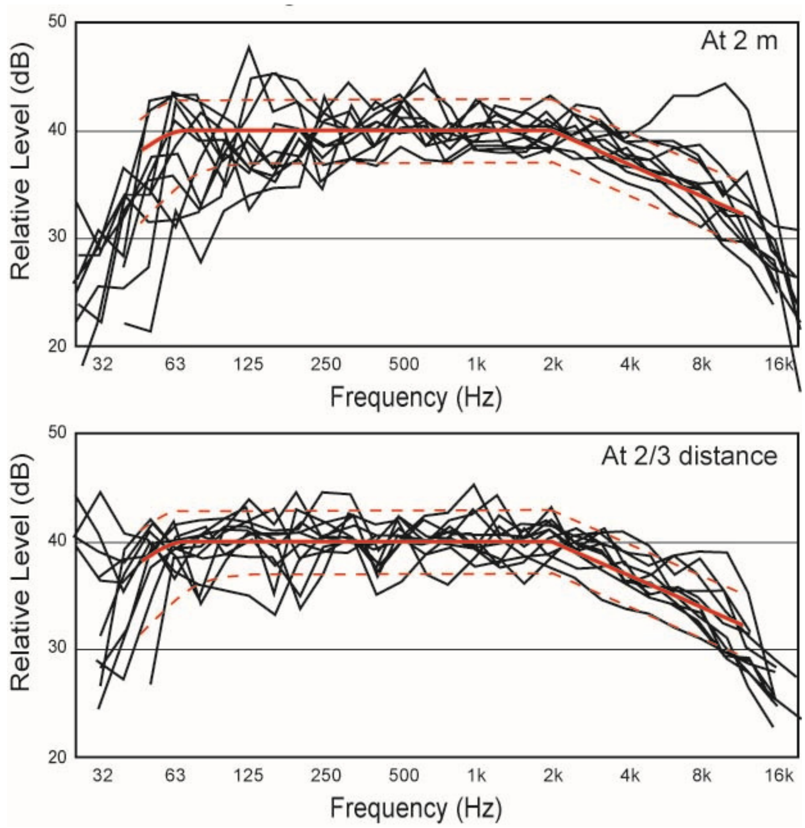

Figure 7 Overlaid Third-Octave Responses of Dubbing Theatres

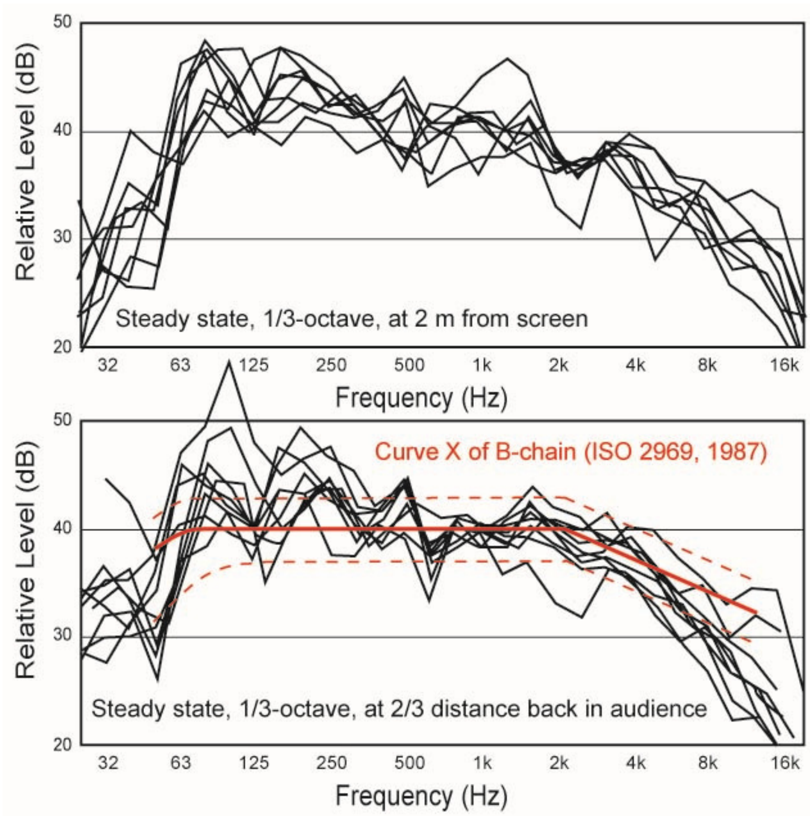

Figure 8 Overlaid Third-Octave Responses of Public Cinema Rooms