Proceedings of the Institute of Acoustics

IMPROVING SPEECH INTELLIGIBILITY IN UNDERGROUND STATIONS

BM Shield South Bank University, School of Engineering Systems and Design, London
LN Yang South Bank University, School of Engineering Systems and Design, London

1. INTRODUCTION

Underground stations are typically long enclosures, for which classical room acoustics is inappropriate [1-3]. A ray tracing computer model has been developed specifically for the prediction of the sound field in long enclosures [4,5]. The aim of the model is to enable the acoustics of long enclosures, particularly underground stations, to be studied. In particular the model has been designed to predict the acoustic parameters associated with speech intelligibility, including sound propagation, clarity index, and speech transmission index. This paper briefly describes the development and validation of the model. It shows how the model has been used to study the sound field in long enclosures with multiple sources, of which underground stations are particular examples, and discusses the effects predicted by the model of various factors upon speech intelligibility.

2. THE COMPUTER MODEL

The computer model that has been developed for the prediction of the sound field in underground stations is a ray tracing model. It takes account of the directivity of sources, diffraction around the ends of train running tunnels, and uses a statistical method to compensate for the loss of late reflections. The model has been developed to predict the sound field in long enclosures of both rectangular and circular cross section, as many underground stations have curved surfaces. The model, uniquely, uses the exact analytical description of a curved surface in the geometrical description of the space being modelled, as it has been found that approximating a curved surface, even with many planes, leads to focussing effects being underestimated [6].

The model predicts sound propagation, and room acoustics parameters relating to speech intelligibility. These are derived from the impulse response and include reverberation time, early decay time, deutlichkeit or definition (D50), and clarity index (C50).

The model also predicts the speech transmission index (STI) from calculation of modulation transfer functions. The method of calculating STI was based upon the ray tracing method of Rietschote et al [7], which has been modified to include air absorption and background noise, to make it suitable for predicting STI in a real long enclosure such as a station.

3. VALIDATION OF THE MODEL WITH A SINGLE SOURCE

The model has been validated using data from scale model measurements and measurements made in a real station [8,9]. In each case a single source was used.

The scale model was a 1:16 model of an underground station with curved surfaces. Measurements and predictions were made of sound pressure levels and early decay times at receiver points along the length of the platform.

As long enclosures are non-diffuse spaces, the sound pressure level decreases and the early decay time increases with increasing source-receiver distance when there is a single source. Comparison of measured and predicted values showed that the computer model correctly modelled both sound propagation and early decay times along the length the station, across the frequency range.

Measurements made in the scale model with the ends both completely open and partially open, as would be the case in a real station, showed that the station ends have more effect on early decay time than on sound pressure level. The effects of the ends are more noticeable at greater distances from the sound source. This suggested that treatment of the station ends could be a method of improving speech intelligibility, as adding absorbent to the ends would reduce reverberation time. The effect on the speech transmission index of adding absorption to the platform ends is discussed further below.

The ability of the model to accurately predict parameters relating to speech intelligibility was investigated using measurements made in a real station of rectangular cross section. A single loudspeaker source was placed approximately one third of the way along the length of the platform. The following parameters were measured and predicted: sound propagation, early decay time, speech transmission index, clarity index and definition. The measurements showed that early decay time increased with increasing source receiver distance, whereas the values of all the other parameters decreased along the platform.

It was found that, in general, all parameters were accurately predicted. Predictions of the speech transmission index were particularly accurate, both at individual frequencies and overall. The error in the predicted value of overall STI was less than 0.03 in the near field and 0.01 in the far field of the source.

4. THE SOUND FIELD WITH MULTIPLE SOURCES

Validation of the model showed that it accurately predicts the sound field in long enclosures of varying cross sectional shape with a single source. It was then used to investigate the sound field in a long enclosure with multiple sources, such as an underground station with speakers arranged along the length of the platform.

Measurements made previously by Barnett [10] in a real station, and by Orlowski [11] in a scale model station have shown that, although the sound field in a long enclosure is non-uniform when excited by a single source, with multiple sources the sound field resembles a diffuse field. These results have been confirmed by calculations by Kang [12]. The model described here was used to predict sound propagation and speech transmission index in a long semi-cylindrical enclosure, with five sources and ten sources. The predictions agreed with the findings of the previous workers. In both cases sound pressure levels were almost constant along the length of the spaces, as were the values of the STI.

The STI calculation method was further validated for the case of multiple sources by comparing predictions of speech transmission index with measurements by Barnett [8] of a PA system in a real underground station. In this case nineteen loudspeakers were arranged along the length of the platform, and predictions were made at five receiver positions. There was good agreement between measured and predicted values.

5. PREDICTIONS OF SPEECH INTELLIGIBILITY IN DIFFERENT CONDITIONS

In order to investigate the effects on speech intelligibility of certain factors, the speech transmission index has been predicted under various conditions. The predictions were made in a hypothetical space based upon a typical underground station, the scale model of which was used in the earlier validation described above.

The space was assumed to be a long partially cylindrical enclosure with diameter 5.7 m and length 128 m, as shown in Figure 1.

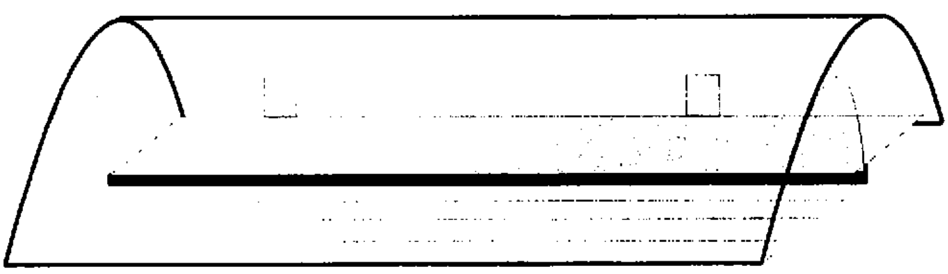

Figure 1. Geometry of hypothetical long enclosure

5.1 Effects of background noise

In order to investigate the effect of background noise on speech intelligibility, predictions of STI were made in varying background noise levels with one, two and five sources. The locations of the sources at 20 m intervals, and receivers at 10 m intervals along the length of the space are shown in Figure 2. Each source was assumed to have the same sound power level of 100 dB.

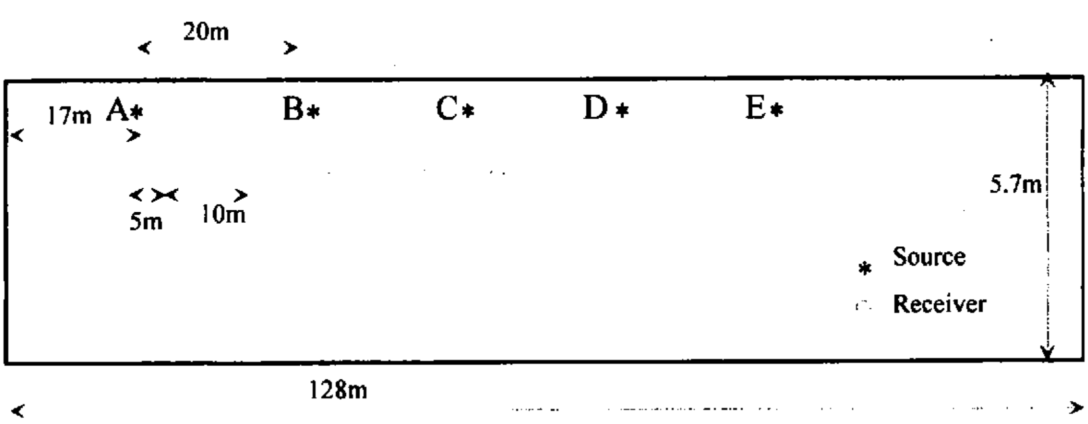

Figure 2. Positions of sources and receivers.

Predictions were made with the single source A, with two sources A and B, and with all five sources as shown in Figure 2. Source A was located 17 m from the left end of the enclosure, and the first receiver position 5 m to the right of source A.

For each combination of sources predictions were made in four differing background noise conditions: with no noise, and with signal to noise ratios of 50 dB, 20 dB and 10 dB (that is, in background noise levels of 0 dB, 39 dB, 69 dB and 79 dB, given a source sound power level of 100 dB).

The predicted values of STI with no background noise and with signal to noise ratios of 20 dB and 10 dB are shown in Figures 3 to 5. It was found that for each source arrangement the STI values were very similar when there was no background noise and with a signal to noise ratio of 50 dB. The results also showed that with no background noise, the larger the number of sources the lower the STI values. This is due to interference between sources when there is no background noise.

It was also shown that, the poorer the signal to noise ratio, the greater the improvement in STI that can be achieved by having multiple sources, particularly in the far field of the single source. However the lower the signal to noise ratio the poorer the speech intelligibility is for any number of sources, as would be expected.

With a single source the STI decreases with increasing distance from the source in all noise conditions. However, with five sources at equal distances the STI values are almost constant along the length of the enclosure in all noise conditions. As stated above when there is background noise the use of a multiple source system improves the STI values. This is because the effect on the STI of the increase in reverberant sound is compensated for by the increase in the number of sources. The results show that, in high levels of background noise, the larger the number of sources, the greater the improvement in STI that can be achieved.

It can also be seen that, for a five source system, the changes in STI between different noise conditions are negligible when the signal to noise ratio is greater than 20 dB. However, with one or two sources the STI increases with increasing signal to noise ratio.

In underground stations there will always be background noise due to the passengers and train noise. The effect of this noise on the STI will be significant if the signal to noise ratio is poor. When the background level is high, a multiple source system will improve the speech intelligibility in the space. Although in low background noise conditions a single source gives higher STI values than a multiple source system in the near field of the source, a multiple source system generally has the following advantages in a long enclosure such as an underground station. It gives uniformity of the sound field and improves speech intelligibility along the whole space when the background noise level is high.

5.2 Effects of increasing the number of sources

To further investigate the effects on the STI of the number of sources predictions were made with ten sources at 10 m intervals, and the results compared with those obtained previously for five sources. Predictions were made at the same five receiver points as before, in three noise conditions: in no background noise, and with signal to noise ratios of 20 dB and 10 dB (that is, in noise levels of 69 dB and 79 dB).

The results showed that with no background noise and with a signal to noise ratio of 20 dB there was very little difference in STI values with five and ten sources. In all cases the STI values were approximately constant along the length of the space. In the higher level of background noise, when the signal to noise ration was 10 dB, the STI values decreased slightly and there was a small difference between the STI with five and ten sources, ten sources giving very slightly higher values. Thus increasing the number of sources has little effect on speech intelligibility when there is no or a low level of background noise, and gives a marginal improvement in higher levels of background noise

The presence of passengers on the station will affect speech intelligibility. To study this effect predictions were made of the STI with ten sources in varying background noise levels and with different absorption coefficients of the platform floor. The noise level conditions were as before, that is, no ambient noise, and with signal to noise ratios of 20 dB and 10 dB; and the values of the floor absorption coefficients were 0.05 and 0.02.

It was found that increasing the floor absorption coefficient marginally improved the STI values (by up to approximately 0.05) in all three noise conditions. However, with increasing background noise level, the improvement in the STI caused by increasing the floor absorption gets smaller; that is higher levels of background noise cancel the improvement in STI which is given by the higher floor absorption. Applying this to the case when passengers are on the platform, as roughly approximated by the floor absorption, the STI will only improve (over the situation when the station is empty) provided the passengers are very quiet and there is no noise of an approaching train.

5.3 Effects of increasing the absorption of platform ends

The ends of passenger platforms are usually constructed of the same material as the side walls, which is acoustically hard. These hard surfaces are not good for speech intelligibility since strong reflections from the platform ends will increase reverberation time and decrease speech intelligibility. They will significantly affect the sound field especially towards the ends of the platform, even though the surface area of the end is small compared with the whole surface area of the station. Theoretically, the speech intelligibility should be improved if absorbent is added to the ends of the platforms to reduce reflections. This has been investigated by using the computer model to calculate the speech intelligibility with and without absorbent treatment on the platform ends.

Predictions of STI were made with ten sources along the platform, without background noise and with a signal to noise ratio of 10 dB.

For the predictions it was assumed that the absorbent on the end of the platforms had absorption coefficient 1, which means that no sound would be reflected from the ends, as in the case of totally open ends.

The predicted values of STI with and without absorbent treatment with no background noise are shown in Figure 6. The results when the signal to noise ratio is 10 dB are similar. In both cases the introduction of absorption slightly increases the STI values along the length of the space, by approximately 0.04 with no noise and 0.03 with signal to noise ratio of 10 dB.

6. CONCLUSIONS

In existing underground stations, where the construction and materials cannot easily be altered, and safety and construction standards must be maintained, the improvement of speech intelligibility is difficult. Usually the surface of underground stations is acoustically hard, causing strong reflections from the surface which decrease speech intelligibility. In this case, the best way to improve speech intelligibility is to partly refurbish the station including changing the electro-acoustics of the public address system. For new stations, it is possible at the design stage to ensure that the highest possible levels of speech intelligibility will be achieved.

The above results show that varying the number of sources can improve speech intelligibility in an underground station with a multiple source system. In a quiet environment, decreasing the number of sources can increase the speech intelligibility since it reduces the reverberation time. In a high background noise environment, increasing the number of the sources can improve speech intelligibility by increasing the signal to noise ratio. Changing the absorption coefficient of the floor of the platform can also give a slight improvement. The introduction of absorbent on the ends of the platform will also give some improvement in speech intelligibility by reducing the strong reflections from the ends

It has been seen that in a quiet environment, the STI for a single source system is better than that for a multiple source system. This implies that reducing the number of the sources can increase the speech intelligibility. During the off peak time when there are fewer passengers on the platform, the environment is likely to be quiet, so fewer sources will be needed.

The PA system should be adaptable to meet different conditions. For example, when the background noise is high the PA system could use all of the loudspeakers; in a quiet environment, only half or fewer of the loudspeakers need to be used, for example, every other loudspeaker or every third loudspeaker could be used.

The improvement of speech intelligibility in underground stations is a complicated issue related to the construction of the space, the background noise levels, the number of sources and their directivity. The use of the computer model described here provides an efficient method of investigating all these factors simultaneously.

7. REFERENCES

1. G Westerberg. A new underground subway station acoustical treatments. Proc. Inter-noise 86, 265-268, 1986.

2. P W,Barnett. Acoustics of underground platforms (Part 2). Proc. Institute of Acoustics, **16**, 433-443, 1994.

3. J Kang The unsuitability of the classic theory in long enclosures. Architectural Science Review, **39**, 89-94, 1996.

4.L N Yang. Computer modelling of speech intelligibility in underground stations. PhD Thesis, South Bank University, 1997.

5. L N Yang and B M Shield. Development of a ray tracing computer model for the prediction of the sound field in long enclosures. J. Sound and Vibration (accepted for publication 1999).

6.B M Shield and L N Yang. Two methods of modelling a curved surface in a ray tracing model. J. Building Acoustics (in press 1999).

7.H F van Rietschote, T Houtgast and J M Steeneken. Predicting speech intelligibility in rooms from the modulation transfer function IV: A ray tracing computer model, Acustica, **49,** 245-252, 1981.

8.L N Yang, T J Cox and B M Shield. Modelling of sound propagation in underground stations. Proc. Inter-noise 96, 3173-3175, 1996.

9.B M Shield and L N Yang. Modelling of acoustic parameters and speech intelligibility in long enclosures. Proc. 16th International Congress on Acoustics, 2157-2158, 1998.

10. P W Barnett. Speech intelligibility measurements taken at Old Street and St John's Wood underground stations. Research report 534, 1994.

11. R J Orlowski. Underground station scale modelling for speech intelligibility prediction. Proc. Institute of Acoustics, **16**, 167-172, 1994.

12. J Kang. Acoustics in long enclosures with multiple sources. Journal of the Acoustical Society of America, **99**, 985-989, 1996.

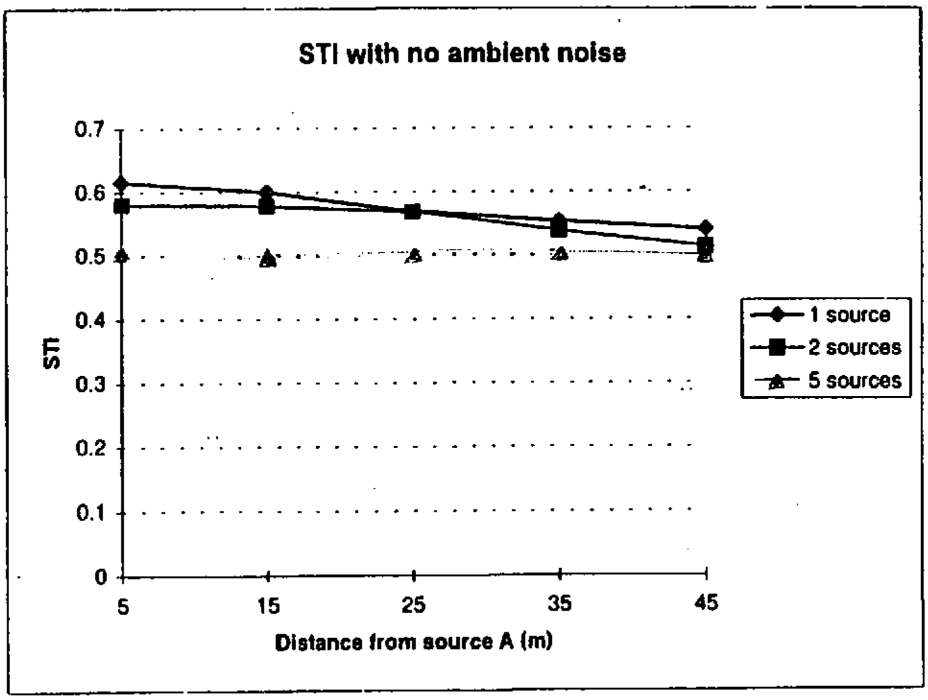

Figure 3. Predicted STI with no background noise.

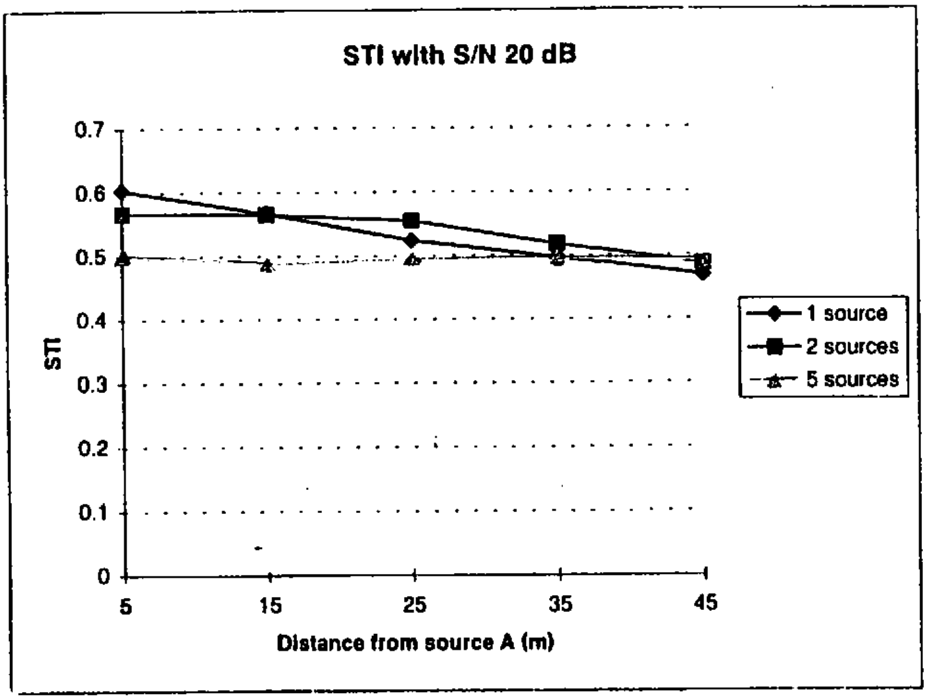

Figure 4. Predicted STI with 20 dB signal to noise ratio.

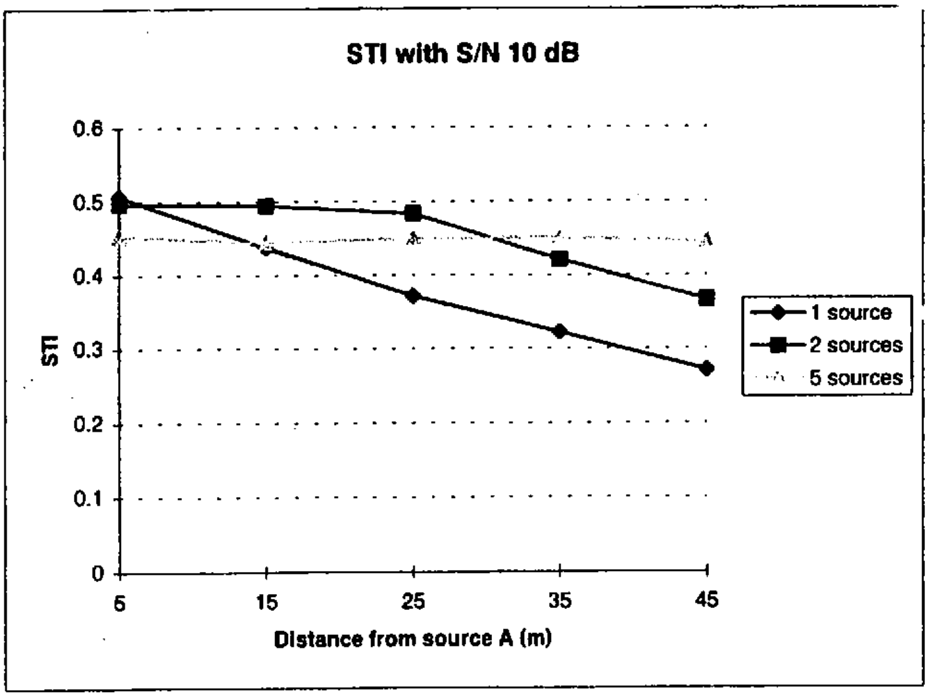

Figure 5. Predicted STI with 10 dB signal to noise ratio.

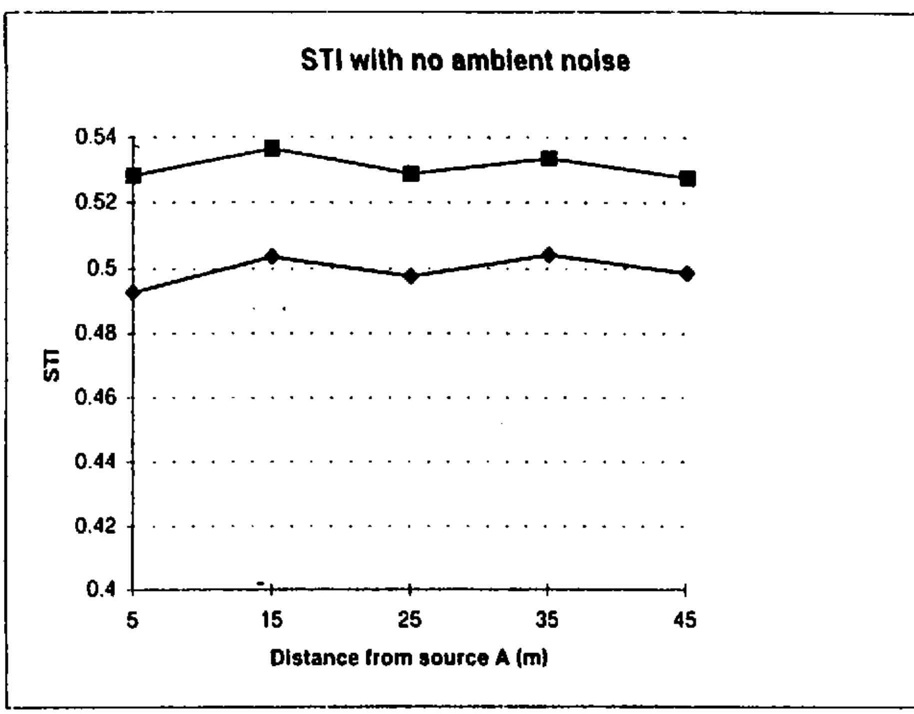

Figure 6. Predicted STI with no background noise, with (■) and without (♦) absorption on platform ends