Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.jzus.zju.edu.cn; www.springerlink.com
E-mail: jzus@zju.edu.cn

Sound quality evaluation of high-speed train interior noise by adaptive Moore loudness algorithm*

Le LUO, Xu ZHENG†‡, Zhi-yong HAO, Wen-qiang DAI, Wen-ying YANG

(*Department of Energy Engineering, Zhejiang University, Hangzhou 310027, China*)

†E-mail: zhengxu@zju.edu.cn

Received Apr. 3, 2016; Revision accepted Oct. 8, 2016; Crosschecked Aug. 15, 2017

Abstract: An online experiment to acquire the interior noise of a China Railways High-speed (CRH) train showed that it was mainly composed of middle-low frequency components and could not be described properly by linear or A-weighted sound pressure level (SPL). Thus, the appropriate way to evaluate the high-speed train interior noise is to use sound quality parameters, and the most important is loudness. To overcome the disadvantages of the existing loudness algorithms, a novel signal-adaptive Moore loudness algorithm (AMLA) based on the equivalent rectangular bandwidth (ERB) spectrum was introduced. The validation reveals that AMLA can obtain higher accuracy and efficiency, and the simulated dark red noise conforms best to the high-speed train interior noise by loudness and auditory assessment. The main loudness component of the interior noise is below 27.6 ERB rate (erbr), and the sound quality of the interior noise is relatively stable between 300–350 km/h. The specific loudness components among 12–15 erbr stay invariable throughout the acceleration or deceleration process while components among 20–27 erbr are evidently speed related. The unusual random noise is effectively identified, which indicates that AMLA is an appropriate method for sound quality assessment of the high-speed train under both steady and transient conditions.

Key words: High-speed train; Sound quality evaluation; Equivalent rectangular bandwidth (ERB) spectrum; Adaptive Moore loudness algorithm (AMLA); Unusual random noise

http://dx.doi.org/10.1631/jzus.A1600287 **CLC number:** TB532

1 Introduction

By early 2015, the Chinese high-speed railway had exceeded 16000 km in route length, and this accounts for more than 60% of the worldwide total length (Sone, 2015; Tan *et al*., 2016). The noise and vibration problems of high-speed trains have been receiving more and more attention (Jin, 2014). On the one hand, the pass-by noise is mainly composed of the wheel-rail noise and aerodynamic noise (Mellet *et al*., 2006; Deng *et al*., 2014; Noh *et al*., 2014), and has a big impact on the life quality of the residents living along the railway (Gu, 2006). On the other, since many high-speed trains in China start at dawn and arrive at dusk, the interior noise throughout the long journey will greatly influence the riding comfort of the passenger, and has become one of the biggest complaints about the high-speed train. Therefore, it is important to find a proper way to evaluate the interior noise in agreement with the subjective feelings of the passengers, particularly for development of the next generation high-speed train (Soeta and Shimokura, 2013).

The traditional way to evaluate a sound is to use A-weighted sound pressure level (SPL), with linear SPL reduced based on the 40 phon equal-loudness curve. Although A-weighted SPL is supposed to accord with the subjective feelings of human ears, the evaluation will be inappropriate when the noise is mainly composed of middle-low frequency components, or the overall value of linear SPL exceeds

‡ Corresponding author

* Project supported by the Fundamental Research Funds for the Central Universities (No. 2016QNA4012), China

ORCID: Le LUO, http://orcid.org/0000-0002-9049-6171; Xu ZHENG, http://orcid.org/0000-0001-9000-6593

60 dB (Hellman and Zwicker, 1987). Thus, most research work has been focused on the psychoacoustic parameters of sound quality, and loudness is the most important parameter to evaluate the intensity of a sound. This has become the definitive model for calculating other parameters such as sharpness and roughness (Cook and Ali, 2012).

The primary loudness model was introduced by Fletcher and Munson (1933). Then the Stevens (1956) and Zwicker (1956) models were proposed and developed, leading to the standard ISO 532:1975 (ISO, 1975). Zhang *et al*. (2012) used the Zwicker loudness combined with noise rating (NR) and room criterion (RC) parameters to evaluate the noise in the tourist cabin of a China Railways High-speed (CRH) train. Park *et al*. (2015) researched the unsteady sudden variation of the interior noise when a Korea Train Express (KTX) passed another train running in the opposite direction or passed through a tunnel, based on the Zwicker loudness and wavelet transform. However, the frequency band of the Stevens or Zwicker model was divided based on the 1/3 octave principle, which may lead to inaccurate assessment. To overcome this deficiency, Moore *et al*. (1997) proposed a new loudness model which achieved a certain progress. First, equivalent rectangular bandwidth (ERB) was introduced to divide the audible frequency range into 372 nonlinear bands, which was more detailed than the original critical bandwidth named Bark. Then the cochlea excitation was calculated by analytic equations, which could extend to continuous signal processing. The Moore loudness model has been widely adopted and became the American national standard ANSI S3.4-2005 (ANSI, 2005).

However, the accuracy, efficiency, and application range of the traditional spectrum processing methods in the Moore loudness model are varied along with different input signals, which may affect the sound quality evaluation directly. In this paper, an adaptive Moore loudness algorithm (AMLA) is proposed, and the evaluation accuracy is validated by simulated signals. Then AMLA is applied to distinguish the different auditory feelings between the interior noise of a CRH train and four kinds of simulated colored noises. Finally, the sound qualities inside the CRH carriage under steady and transient conditions are analyzed by AMLA. The main procedure is shown in Fig. 1.

2 Development of AMLA

2.1 Computation procedure of Moore loudness

According to standard ANSI S3.4-2005, the computation procedure of the Moore loudness shown in Fig. 1 can be divided into four steps as follows:

1. The human ear is mainly composed of the outer ear, middle ear, and cochlea. When people hear a sound, the sound energy is magnified when transmitted from the outer ear to the middle ear, and attenuated from the middle ear to the cochlea. To demonstrate the specific acoustical effect of the human ear, the transfer function is determined by cubic spline interpolation. As shown in Fig. 2, the amplifying effect of the outer ear mainly works in bands above 1 kHz, and the amplitude fluctuation in a free field is much larger than that in a diffuse field. The attenuation value of the middle ear is rather large in the low frequency band, and varies in fluctuations as the frequency exceeds 1 kHz.

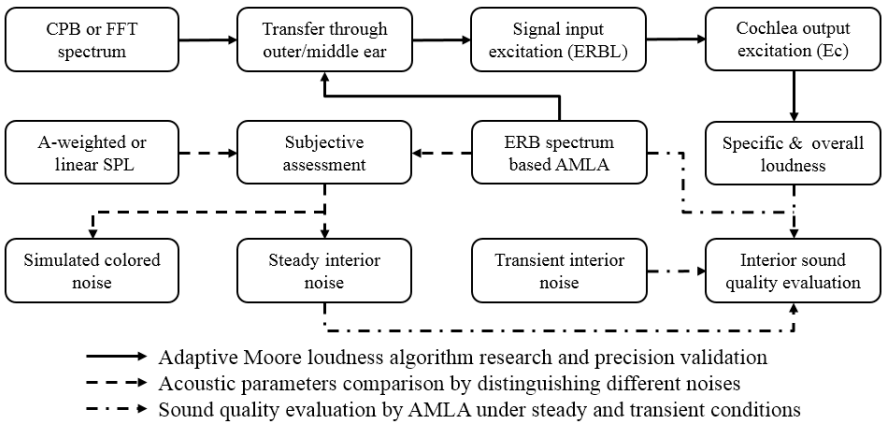


Fig. 1 Sound quality evaluation procedure of the high-speed train interior noise
CPB means constant percentage bandwidth, and FFT means fast Fourier transformation

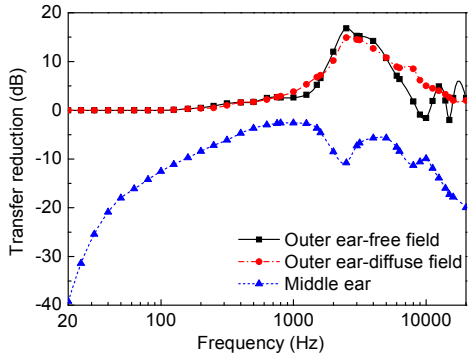


Fig. 2 Transfer function of the outer ear and middle ear (ANSI, 2005)

2. After the transmission through the outer ear and middle ear, the sound reaches the cochlea and generates the input excitation. Unlike the critical bandwidth proposed by Zwicker, Moore adopts ERB to describe the nonlinear characteristics of the human auditory system (Mao *et al*., 2013). Accordingly, signal center ERB is expressed as

$$\mathrm{ERB_o}=24.673\times(0.004368f_o+1), \tag{1}$$

where f_o is defined as the signal center frequency.

The input excitation level of the transmitted signal called ERBL is calculated by summing up the power spectrum in an ERB. The related equations are as follows:

$$\mathrm{ERBL_o}=10\times\log\left(\frac{\sum_{f_o-\mathrm{ERB_o}/2}^{f_o+\mathrm{ERB_o}/2}W_aP_a^2}{P_0^2}\right), \tag{2}$$

$$W_a=(1+p_ag_a)\exp(-p_ag_a), \tag{3}$$

where P_a is the effective sound pressure of the signal, P_0 denotes the reference sound pressure 2×10^{-5} Pa, W_a is the weighting function with $g_a=|f_a-f_o|/f_o$, $p_a=4f_a/\mathrm{ERB_o}$, and f_a represents the frequency of the signal in the corresponding $\mathrm{ERB_o}$ range.

3. The output excitation of the cochlea E_c is calculated in this step. Moore considers that the loudness model should be divided into 372 cochlear filters from 1.8 ERB rate (erbr) to 38.9 erbr with an interval of 0.1 erbr. With the erbr of each cochlear filter ($\mathrm{erbr_c}$) known, the corresponding center frequency is defined as the cochlea center frequency f_c:

$$f_c=\frac{10^{\mathrm{erbr_c}/21.366}-1}{0.004368}. \tag{4}$$

Consequently, f_c represents the human hearing range between 50–15 000 Hz. With f_c known, the corresponding cochlea bandwidth $\mathrm{ERB_c}$ can be obtained by Eq. (1).

The weighting function W_b that determines the shape of the cochlea filter is basically the same as that in Eq. (3), but with different values for each parameter:

$$p_b=\begin{cases}4f_c/\mathrm{ERB_c}-0.35\times[4f_c/(\mathrm{ERB_c}\times30)]\\ \quad\times(\mathrm{ERBL_b}-51), & f_b\le f_c,\\ 4f_c/\mathrm{ERB_c}, & f_b>f_c,\end{cases} \tag{5}$$

$$g_b=\begin{cases}|f_b-f_c|/f_c, & |f_b-f_c|/f_c\le1,\\ 1, & 1<|f_b-f_c|/f_c\le4,\\ 4, & |f_b-f_c|/f_c>4,\end{cases} \tag{6}$$

where f_b is the signal frequency in the corresponding $\mathrm{ERB_c}$ range. $\mathrm{ERBL_b}$ represents the corresponding ERBL calculated in step 2.

E_c is defined as the sum of the sound energy in one $\mathrm{ERB_c}$:

$$E_c=\frac{\sum_{f_c-\mathrm{ERB_c}/2}^{f_c+\mathrm{ERB_c}/2}W_bP_b^2}{P_0^2}\cdot E_0, \tag{7}$$

where E_0 is the output excitation value of a 1 kHz sinusoid with a SPL of 0 dB, and P_b is the signal sound pressure in the corresponding $\mathrm{ERB_c}$ range.

4. After acquiring the output excitations of all the cochlea filters, the specific loudness N'_c with the unit of sone/erbr can be calculated as

$$N'_c=\begin{cases}C\left(\frac{2E_c}{E_c+E_{\mathrm{thrq}}}\right)^{1.5}\left[\left(GE_c+2E_{\mathrm{thrq}}\right)^{\alpha}-\left(2E_{\mathrm{thrq}}\right)^{\alpha}\right], & E_c\le E_{\mathrm{thrq}},\\ C\left[\left(GE_c+2E_{\mathrm{thrq}}\right)^{\alpha}-\left(2E_{\mathrm{thrq}}\right)^{\alpha}\right], & E_{\mathrm{thrq}}<E_c<10^{10},\\ C\left(\frac{E_c}{1.0707}\right)^{0.2}, & E_c\ge10^{10}.\end{cases} \tag{8}$$

According to the ANSI standard, C=0.046 871. E_{thrq} is the absolute threshold connected to f_c. G represents the low-level gain of the cochlea amplifier with G=E_{thrq}(500 Hz)/E_{thrq}(f_c), and α is a coefficient related to G. As shown in Fig. 3, the three parameters are variable when f_c<500 Hz but remain invariant when f_c≥500 Hz, with E_{thrq}=2.3067, G=1, and α=0.2.

Finally, the overall loudness with the unit of sone is obtained by summing up the specific loudness from all the cochlea filters.

$$N=\sum_{c=1}^{372}N_c'\times 2/10, \tag{9}$$

where multiplying by 2 means binaural presentation and dividing by 10 represents 0.1 erbr intervals.

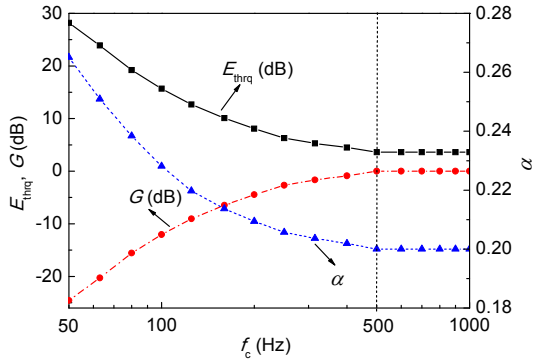


Fig. 3 E_{thrq}, G, and α curves as a function of f_c (ANSI, 2005)

2.2 Disadvantages of the existing algorithms

As for the frequency-domain signals with the amplitude already known, the Moore loudness can be calculated directly by the procedure mentioned above. However, the sounds in practical engineering are complex time-domain signals with various unknown components. Thus, the crucial factor of the computational accuracy is the input spectrum of the time-domain signal.

2.2.1 Constant percentage bandwidth spectrum

According to the discrete transform algorithm given in the ANSI standard, first the 1/3 octave bands spectrum of the signal is obtained by temporal filtering. Then the input spectrum is computed based on the width of each frequency band, namely n_k. When the bandwidth is less than 30 Hz, the components are spaced at 1 Hz intervals, with the amplitudes set equal to the SPL of the corresponding 1/3 octave band (SPL_k). When the bandwidth exceeds 30 Hz, the components are spaced at 10 Hz intervals, with the amplitudes set 10 dB higher than the SPL of the corresponding band. The acquired spectrum is named the constant percentage bandwidth (CPB) spectrum in this paper, and the equations are expressed as

$$\text{SPL(CPB)}=\begin{cases}\text{SPL}_k-10\times\log(n_k), & n_k\le 30,\\ \text{SPL}_k-10\times\log(n_k/10)-10, & n_k>30.\end{cases} \tag{10}$$

As shown in Fig. 4a, the CPB spectrum of a test signal has similar characteristics to the FFT spectrum of the colored noise, which represents the same SPL value in the same 1/3 octave band. As a result, the amplitude peaks of the original signal are blurred or even distorted. The more complex the signal is, the greater the calculation error may become. Thus, the Moore loudness algorithm based on CPB spectrum may be inaccurate while processing practical engineering signals.

2.2.2 FFT spectrum

The FFT spectrum of the test signal can express the peak amplitude objectively as shown in Fig. 4b. However, with fixed frequency resolution, high calculation efficiency and accuracy are hard to achieve at the same time according to Heisenberg's uncertainty principle (Ding and Pei, 2013). Generally speaking, the higher the frequency resolution of the FFT spectrum is, the more accurate the computation of Moore loudness can be. However, the subsequent calculation efficiency may drop dramatically at the same time. Many scholars have studied the Moore loudness algorithm based on the FFT spectrum (Jiao *et al.*, 2012). Glasberg and Moore (2002) proposed a more efficient algorithm by synthesizing the spectrum through multiple parallel FFT computations, so that the spectrum can obtain higher resolution in the low frequency band and lower resolution in the high frequency band. However, the FFT spectrum leakage will vary with different window widths, which may result in nonconservation of the total energy, and eventually this restricts the applicability of the algorithm.

2.3 AMLA based on ERB spectrum

A time-domain signal can be transformed into various spectra in the frequency domain. Since the CPB spectrum cannot represent the peak characteristics

precisely and the FFT spectrum is restrained by the uncertainty principle, a novel spectrum with better frequency distribution is needed for the routine procedure of Moore loudness calculation.

As stated in Section 2.1, the input ERBL and output E_c of the cochlea are both the sum of the sound energy in the corresponding ERB range. If the input spectrum of the sound is directly transformed into the ERB form, the whole procedure of Moore loudness calculation can be considerably simplified. In addition, the nonlinear characteristics of the sound in accord with the human auditory system can be retained to the maximum extent.

To achieve this, first the power spectral density (PSD) spectrum of the acoustic signal is processed with an interval of 1 Hz.

$$\mathrm{PSD_j}=\left|\sum_{m}^{M}S_m\,\mathrm{e}^{-\mathrm{i}2\pi f_\mathrm{j}}\right|^2\Big/\Delta f,\quad f_\mathrm{j}=1,2,\cdots,f_\mathrm{s}/2,\quad (11)$$

where S_m is the original signal, i is the imaginary number, f_j is the signal frequency, f_s is the sampling frequency, and M is the signal length with $M=f_\mathrm{s}\times T$. As the frequency interval Δf equals 1 Hz, the value of the time span T is 1 s.

Then the SPL of each frequency component is calculated by summing the energy of the PSD spectrum in the corresponding ERB range.

$$\mathrm{SPL(ERB)}=10\times\log\left(\frac{\sum_{f_\mathrm{c}-\mathrm{ERB_c}}^{f_\mathrm{c}+\mathrm{ERB_c}}\mathrm{PSD_b}\,\Delta f}{P_0^2}\right),\quad (12)$$

where the subscript b is the same with that in Eq. (5).

Considering the energy conservation of time series, the frequency value is calculated as

$$f(\mathrm{ERB})=\frac{\sum_{f_\mathrm{c}-\mathrm{ERB_c}}^{f_\mathrm{c}+\mathrm{ERB_c}}f_\mathrm{b}\,\mathrm{PSD_b}}{\sum_{f_\mathrm{c}-\mathrm{ERB_c}}^{f_\mathrm{c}+\mathrm{ERB_c}}\mathrm{PSD_b}}.\quad (13)$$

We define the obtained spectrum as the ERB spectrum. Once the original sound is generated, the computation procedure of the ERB spectrum can be highly standardized. In other words, the ERB spectrum is exclusive and signal adaptive. The novel algorithm based on the ERB spectrum is named the adaptive Moore loudness algorithm.

As shown in Fig. 4c, the ERB spectrum distribution is relatively dense at the amplitude peaks of the FFT spectrum and sparse at the valleys. When f_s equals 65 536 Hz, the number of spectrum lines are 1867, 4097, and 372 for the case of CPB, FFT (with an internal of 8 Hz), and ERB, respectively. As the subsequent calculation efficiency is inversely correlated with the number of spectrum lines, it can be inferred that AMLA is the most time-saving method for complex sound signals with long time spans.

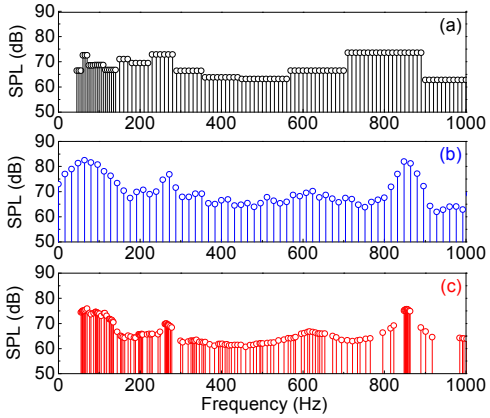


Fig. 4 CPB (a), FFT (b), and ERB (c) spectrum presentations of the time-domain signal

2.4 Validation of evaluation accuracy

To verify the precision of Moore loudness results acquired by different spectrum algorithms, three kinds of simulated noises were studied in this study, and the relative errors were calculated with the reference values given in the ANSI standard. Signal 1 was a 3 kHz pure tone with an SPL of 80 dB. Signal 2 was a complex tone containing components at 1, 1.6, and 2.4 kHz with an SPL of 60 dB for each component. Signal 3 was a broadband pink noise with a frequency range from 50 to 15 000 Hz and a spectrum level at 1 kHz of 0 dB SPL. All the noise signals were simulated with a sampling frequency of 65 536 Hz and time span of 1 s. The corresponding curves in time and frequency domain are shown in Fig. 5.

The overall loudness results are shown in Table 1. The relative errors of the algorithm based on CPB spectrum are 5.27%, 13.18%, and 2.09% for the case of the pure tone, complex tone, and pink noise,

respectively. Since the frequency spectra of CPB and the colored noise are quite similar, it can be inferred that with a high accuracy standard, the CPB spectrum-based algorithm is only suitable for colored noise, which will result in some limitations in application. As for the FFT spectrum based algorithm, the relative errors are all below 1.18% with the frequency resolution of 1 Hz, which indicates that the calculation precision is greatly improved. However, the calculation efficiency distinctly declines. As shown in Fig. 6, signal 4 is an interior noise sample of a high-speed train. When the frequency resolution drops with larger intervals, the loudness result varies from 32.6 sone to 37.6 sone, and the computing time decreases exponentially from 28.1 s to 0.8 s. To balance calculation precision and efficiency, a reasonable frequency resolution needs to be determined for each individual signal, which will lead to poor adaptiveness of the FFT spectrum-based algorithm.

By contrast, AMLA not only has the fastest computing speed (with 372 lines), but the relative errors are also the lowest (below 0.52%). Furthermore, for a certain time series the ERB spectrum is exclusive. The simulation results confirm that AMLA has higher precision and efficiency, and can be applied to various noise signal processing. Thus, AMLA is chosen to calculate the loudness of the high-speed train interior noise.

3 Noise assessment by various acoustic parameters

3.1 On-line experiment

The on-line experiment was conducted on a CRH train to get interior noise signals. In accordance with the Chinese national standard GB/T 12816:2006 (SAC, 2006), seven measuring points were set

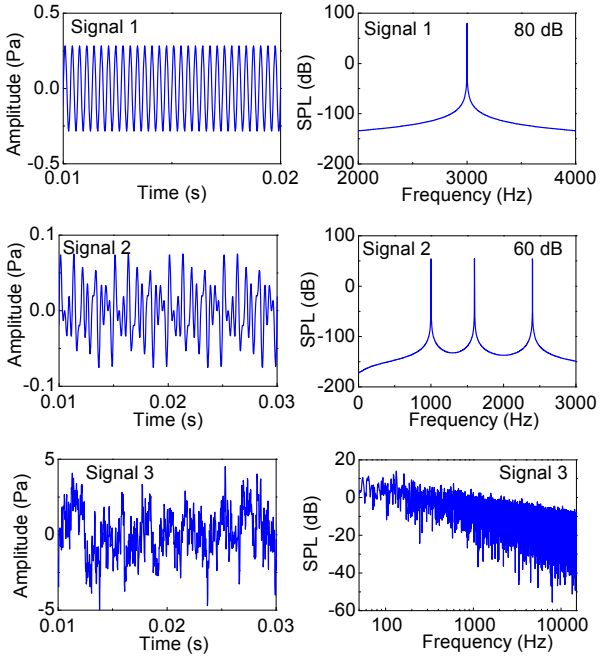


Fig. 5 Time and frequency curves of different simulated signals

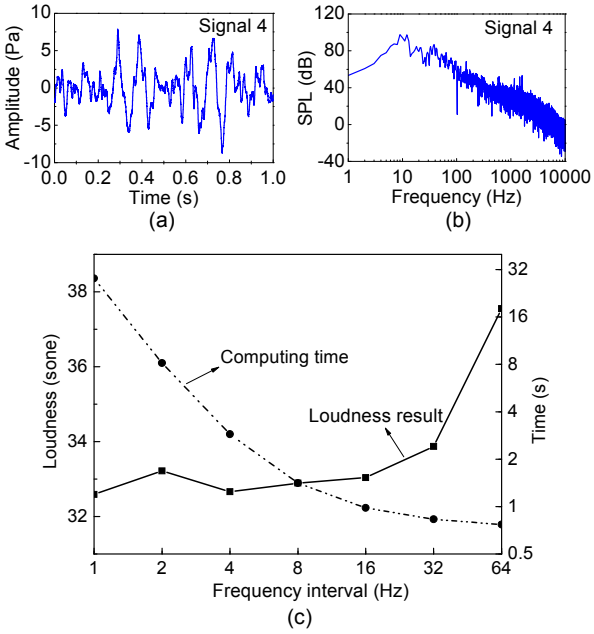


Fig. 6 Frequency resolution influence on the precision (a) and (b), and efficiency of the loudness calculation of an interior noise sample (c)

Table 1 Overall loudness results of different simulated signals

Signal	Reference	CPB spectrum	Relative error (%)	FFT spectrum	Relative error (%)	AMLA	Relative error (%)
Pure tone	27.50	28.95	5.27	27.62	0.44	27.62	0.44
Complex tone	12.75	14.43	13.18	12.90	1.18	12.77	0.16
Pink noise	3.83	3.91	2.09	3.86	0.78	3.85	0.52

inside the carriage. The spatial locations of different measuring points are shown in Fig. 7. The test equipment included a B&K 3050A-type front-end with 24 channels and several B&K 4190-type microphones. The on-line tests were conducted twice. In the first experiment, the interior acoustic environment was poor because the floor was bare, the seats were covered, and the glass doors between two carriages were open. The running speeds were between 250 km/h and 350 km/h. The interior acoustic environment was ameliorated for the second experiment. Carpet was laid on the floor, the seat covers were removed, and the glass doors remained closed. In addition, the test speeds were mostly between 350 km/h and 385 km/h. During data acquisition, the train was running on concrete track in an open field.

3.2 Acoustic characteristics discussion

Noise signals measured during the first and second on-line experiments are abbreviated as A series and B series, respectively. Since the acoustic environment at the central carriage is relatively stable during the whole experiment process, the measuring point P_3 is chosen to analyze the spectrum characteristics of the interior noise. Fig. 8 illustrates the SPL spectra between 20 Hz and 4 kHz at different running speeds. The interior noise signals all share a similar attenuation trend as the frequency rises. They are mainly composed of middle-low frequency components. For the case of the B series signal acquired at 350 km/h (B350-P_3 for short), the amplitude peak at 40 Hz is about 84 dB, and reduces to about 10 dB at 4 kHz. As for the A series signal acquired at the same running speed (A350-P_3), the SPL amplitude above 600 Hz is evidently larger than that of B350-P_3, and the spectrum peaks are much more dense.

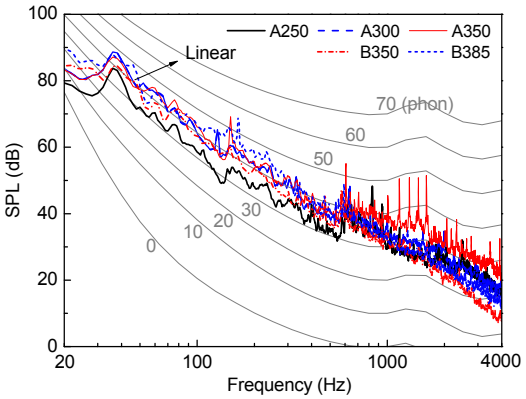
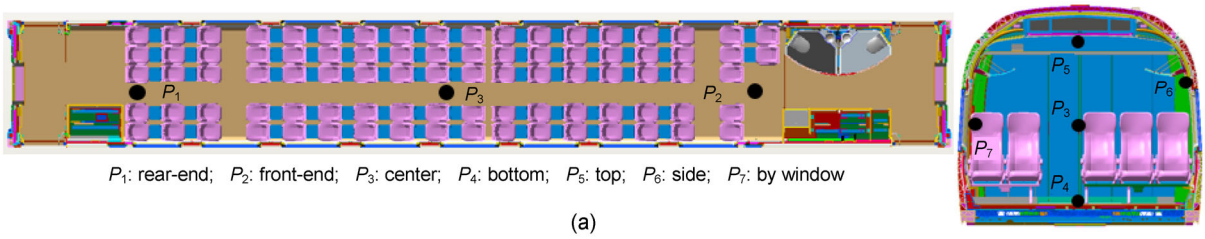
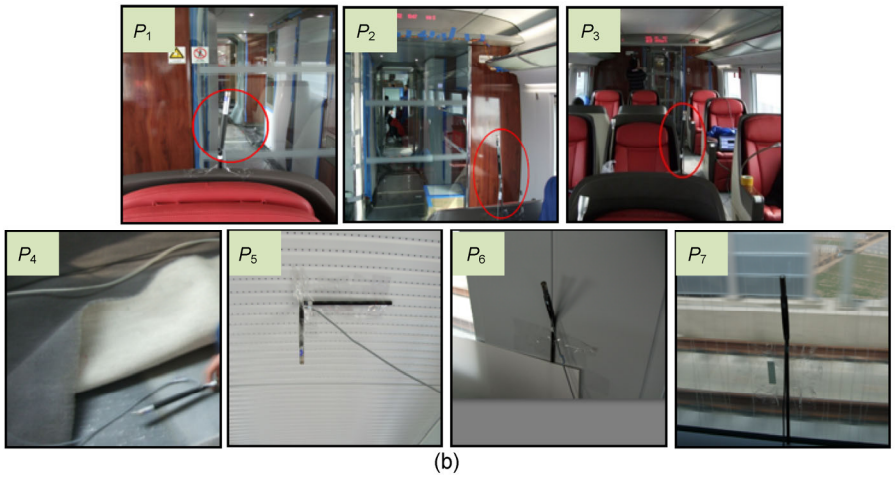


Fig. 8 Spectrum characteristics of the interior noise at different running speeds



P_1: rear-end; P_2: front-end; P_3: center; P_4: bottom; P_5: top; P_6: side; P_7: by window

(a)



(b)

Fig. 7 Schematic diagram (a) and photos (b) of the measuring point layout

In comparison with the equal-loudness contour (GB/T 4963:2007) (SAC, 2007), the whole SPL spectra are basically between the 30 phon and 40 phon equal-loudness curves, except for that of A350 between 1 kHz and 2 kHz, where the spectrum peaks nearly reach the 50 phon equal-loudness curve. Since the subtle difference in spectrum distribution of the interior noise may result in great auditory variations perceived by passengers, and the SPL spectrum can hardly represent the nonlinear characteristics of the interior noise, loudness calculated by AMLA should be applied to evaluate the interior sound quality of the high-speed train.

3.3 Noise distinguishing

In order to verify the availability of the interior sound quality evaluation by AMLA, four colored noises, namely white, pink, brown, and dark red noise (Zhou *et al.*, 2012) with the same A-weighted SPL value of B350-P_3 were adopted for our noise distinguishing research. By listening to the sound playback, the corresponding auditory feelings of the colored noises range from "noisy" to "low and deep".

The specific loudness was calculated by AMLA, and the corresponding curves are shown in Fig. 9, while the overall loudness results are listed in Table 2. The specific loudness of the interior noise signal B350-P_3 drops as erbr increases. The highest peak is located around 2.8 erbr, whose corresponding frequency band is 64–97 Hz, while the highest peak of the objective linear SPL is located around 40 Hz as shown in Fig. 8. As for the deviation, on the one hand, the sound attenuation of the outer ear in the low frequency band is relatively large as shown in Fig. 2, which reduces the peak value at around 40 Hz. On the other, owing to the non-linear auditory characteristics of human ears, the ERB may overlap in frequencies, which means one frequency component may participate in multiple cochlea excitation calculations at different erbr. As a result, the highest subjective peak of specific loudness may shift from the objective peak of the linear spectrum.

In general, the specific loudness of the white noise and pink noise is mainly concentrated on the high erbr levels. As for the brown noise and dark red noise, the main specific loudness components gradually shift from high erbr levels to low erbr levels. Among all the colored signals, the specific loudness curve of the dark red noise is the most similar to that of B350-P_3, as well as the overall loudness value listed in Table 2. Through sound playback assessment, it has been verified that the auditory feelings of B350-P_3 and the simulated dark red noise are quite close to each other. The results confirm that AMLA can effectively substitute for the human ear in evaluating the interior sound quality of the high-speed train.

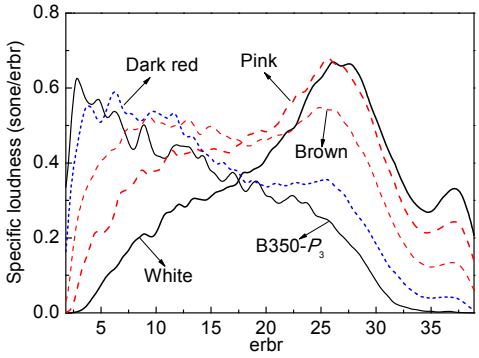


Fig. 9 Specific loudness curves of different noises

Table 2 Overall loudness results with the same A-weighted SPL value

Item	SPL (dBA)	Loudness (sone)
B350-P_3	65.7	21.2
White noise	65.7	25.4
Pink noise	65.7	29.2
Brown noise	65.7	28.8
Dark red noise	65.7	24.4

4 Interior sound quality evaluation by AMLA

In order to evaluate the sound quality inside the carriage comprehensively, loudness variations at different running speeds and spatial locations were analyzed in detail using AMLA. For practical running states, time-domain signals should be long enough to avoid the influence of random factors, such as track irregularity (Ning *et al.*, 2016), the switch between straight and curve paths (Matsumoto *et al.*, 2005), and passing through tunnels or by bridges (Zhang *et al.*, 2016). During the on-line experiment, all the steady conditions had lasted for 60 s, except for that at 385 km/h, which only lasted for 28 s. The final specific loudness curve and overall loudness value were linear-averaged through multiple computations.

4.1 Sound quality analysis at different running speeds

As stated in Section 3.2, noise signals at the central carriage P_3 are chosen to analyze loudness variations at different speeds. The time-erbr color map of the specific loudness is shown in Fig. 10, with the speed information given on the top right corner. For the case of signal A250, the outstanding specific loudness peaks locate at 12.3 erbr and 14.3 erbr, corresponding to 586–679 Hz and 783–898 Hz of the signal spectrum, respectively. As the speed rises to 300 km/h (A300 for short), the main components are below 8.7 erbr, which indicates that the interior noise below 387 Hz has clearly increased and may be easily noticed by passengers. As for the signal A350, the specific loudness in high erbr levels shows a significant growth, especially the peaks at 19.3 erbr and 25.9 erbr. Consequently, the middle-high frequency components around 1504–1702 Hz and 3301–3704 Hz account for a relatively high percentage of the overall loudness and may cause a resounding feeling for passengers.

As for the noise signal B350, the carriage space is more enclosed since the glass doors at both ends of the carriage are closed, and the air borne sound transfer paths from the adjacent carriages are cut off. In addition, with the seat covers removed, the middle-high frequency noise is clearly reduced because of the high absorption coefficient of the seat materials. The similar sound-absorption effect happens with carpet laid on the carriage floor. As a result, the specific loudness of B350 is mainly composed of low erbr components, and this differs greatly from that of A350. The result reveals that measures like using sound-absorbing materials and enhancing air tightness of the carriage are both effective in sound quality improvement. As the train speed reaches 385 km/h (B385), the specific loudness in high erbr levels changes little, while components under 8.7 erbr present an obvious increase. In particular, the amplitude peak at 4.9 erbr has exceeded 0.8 sone/erbr.

The average specific loudness curves at different speeds are shown in Fig. 11. All the specific loudness components above 27.6 erbr decrease sharply, which indicates that the main frequency band of the interior noise affecting sound quality assessment is approximately below 4 kHz. In contrast with A250, the specific loudness of A300 is greater under 10.1 erbr,

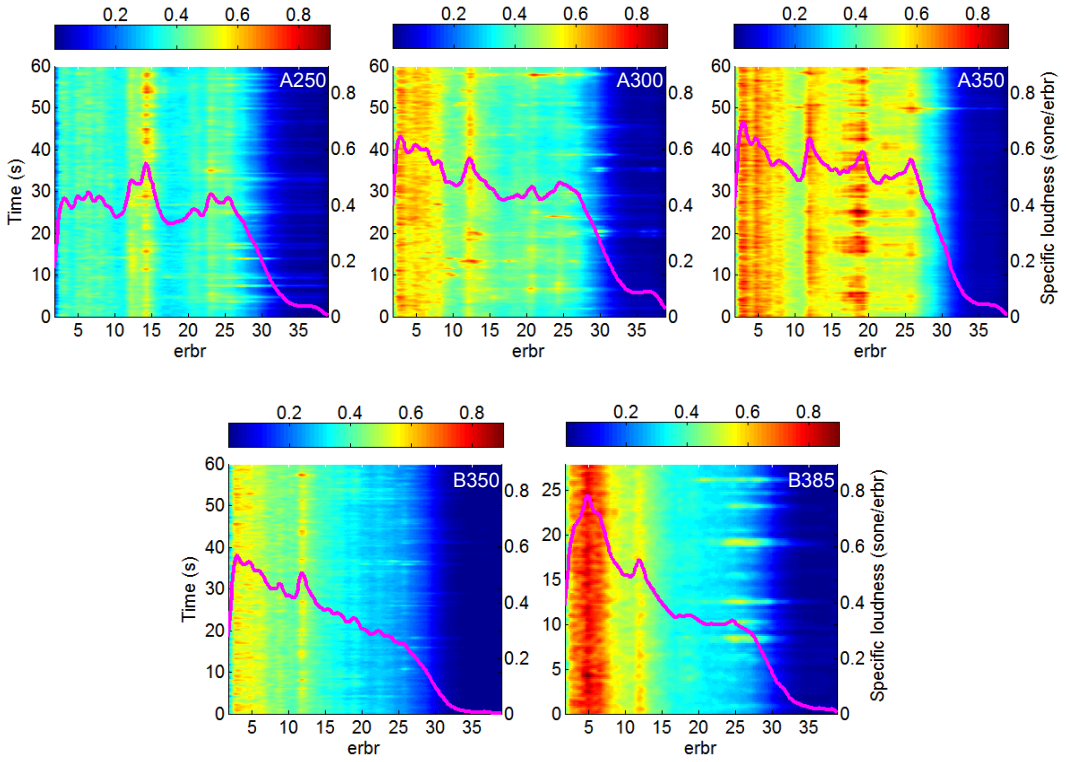


Fig. 10 Time-erbr color maps of the specific loudness at different running speeds

while A350 shows a higher amplitude between 10.8–27.6 erbr. However, the specific loudness of B350 decreases in the middle-high erbr level due to the modified acoustic environment, while that of B385 increases rapidly under 8.7 erbr.

The overall results at P_3 under different running speeds are shown in Fig. 12. The overall loudness values of A series are 23.2, 29.4, and 31.6 sone for the case of A250, A300, and A350, respectively, and the value of B350 is 21 sone, which is 2.2 sone lower in comparison with A250. This interesting phenomenon reveals that the sound quality of B350 is even better than that of A250 by loudness evaluation.

However, with linear SPL evaluation, the overall values of B350 and A300 are quite close. When using A-weighted SPL evaluation, the overall value of B350 is between A250 and A300. The indication is that the evaluation of loudness conforms best to the subjective assessment of the researchers. In addition, the overall loudness grows slowly between 300 km/h and 350 km/h, which indicates that the interior sound quality is relatively stable in this speed range.

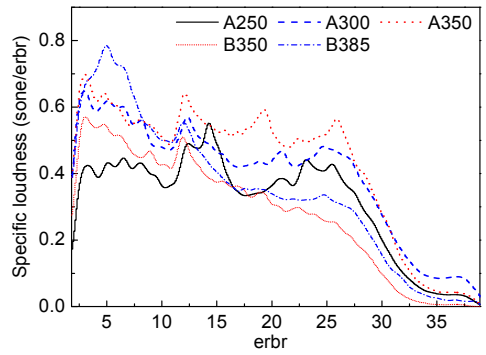


Fig. 11 Average specific loudness comparison at different running speeds

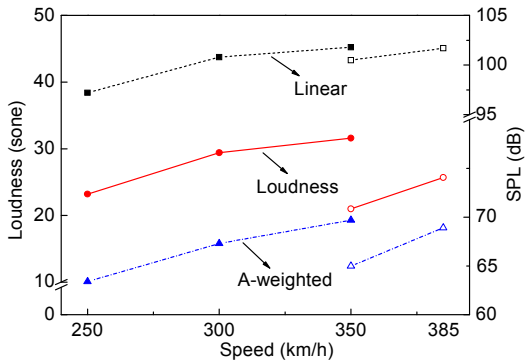


Fig. 12 Overall results comparison at different running speeds for A (■, ●, ▲), and B (□, ○, △) series

However, below 300 km/h or above 350 km/h the noise variation may be easily perceived.

4.2 Sound quality analysis at different spatial locations

Noise signals of B350 are chosen to analyze the sound quality characteristics at different measuring points. Firstly, the wheel-rail contact force around the bogie area, where P_1 and P_2 locate, is excited by the random rail roughness. On the one hand, the contact force stimulates the vibration of wheels and rails, which leads to the major radiated noise distribution at both ends of the carriage. On the other, it is transmitted into the secondary suspension force through the bogie suspension system, and the vibration of carriage body at both ends is excited (Zheng *et al.*, 2016). As a result, the interior noise at P_1 or P_2 is much louder than that of P_3 as shown in Fig. 13. Besides, P_2 is close to the pantograph system of the adjacent coach, so the specific loudness at 12.0 erbr is much higher than that of P_1. As for the measuring points P_4–P_7 at the central carriage, the specific loudness curves all share the similar tendency to decrease as erbr rises. Although the different noise sources have various spectrum characteristics, the loudness has changed little in erbr distribution after the sound attenuation and absorption of the car-body and interior materials.

All the average specific loudness curves at different locations are shown in Fig. 14. It is noticeable that the specific loudness of P_2 is the highest throughout the whole erbr range, and the max amplitude can reach up to 0.8 sone/erbr. The overall value comparison at different locations is shown in Fig. 15. The overall loudness of P_2 is 27.9 sone, which is 0.9 sone and 6.9 sone higher than that of P_1 and P_3, respectively. Among the three locations (P_1, P_2, and P_3), the sound quality at P_3 is the best. The overall loudness ranking of the measuring points surrounding the central carriage is P_5>P_6>P_7>P_4>P_3, which agrees with the subjective assessment of the researchers, and cannot be properly identified by either linear SPL or A-weighted SPL.

4.3 Sound quality variation under transient conditions

As stated above, the loudness calculated by AMLA is an effective way of evaluating the interior

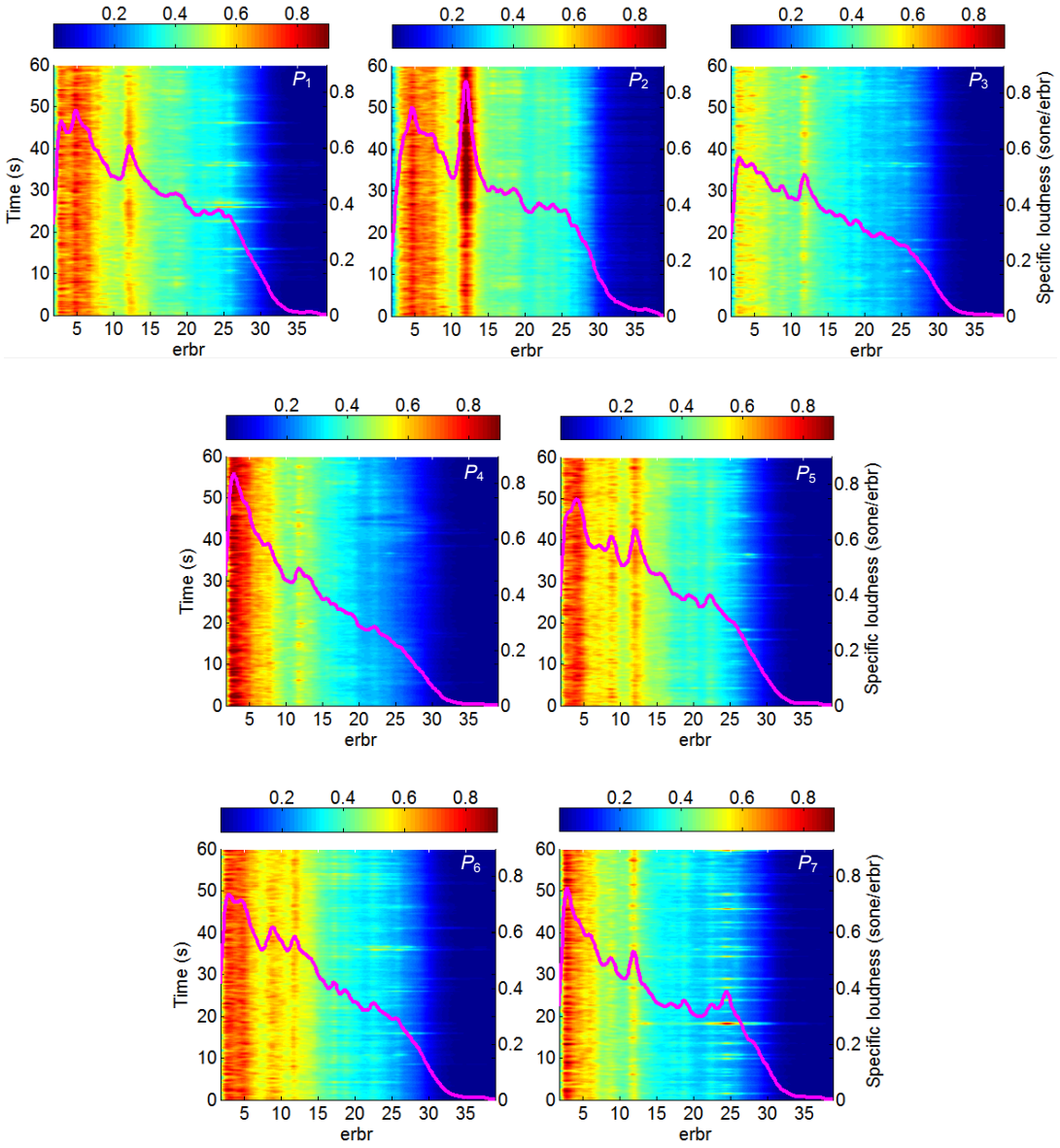


Fig. 13 Time-erbr color maps of the specific loudness at different spatial locations

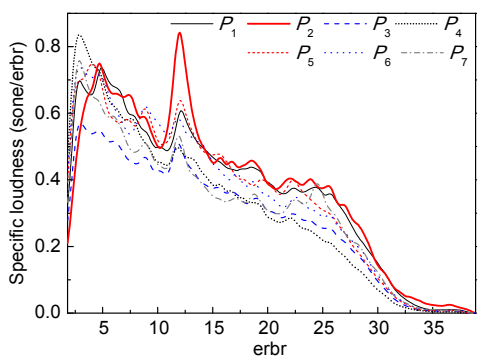


Fig. 14 Average specific loudness comparison at different spatial locations

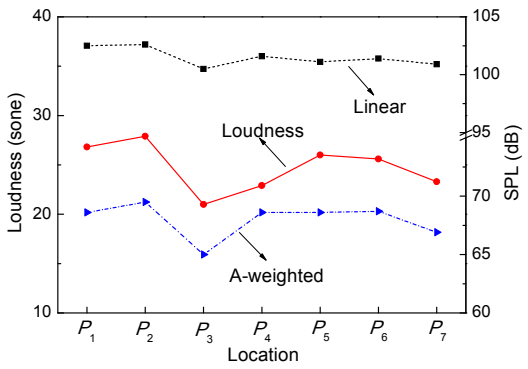


Fig. 15 Overall value comparison at different spatial locations

sound quality of the high-speed train under steady conditions. But in practice, passengers can easily perceive the noise variation when the train accelerates or decelerates. Therefore, a long time series under transient conditions namely A-P_3 is chosen to study the sudden sound quality variation of the high-speed train.

During data acquisition, the signal A-P_3 lasted for approximately 340 s and recorded four processes of the train running states, which were 0–40 s as the adjustment process, 40–190 s as the rapid acceleration process, 190–240 s as the gradual deceleration process, and 240–340 s as the gradual acceleration process. By listening to the sound playback, researchers noticed a sudden noise growth during 80–100 s (T_1 period for short), and a slow noise attenuation during 200–240 s (T_2 period for short). In addition, several unusual random noises like impact sounds, such as the one around 300 s (T_3), were also perceived. In order to evaluate the interior noise variation, AMLA was adopted together with the linear and A-weighted SPLs. Each 1 s time-domain sample was considered as a stationary signal, and the overall value was calculated with a time increment of 0.5 s.

The overall value variation in time-domain is shown in Fig. 16. The fluctuation of the linear SPL is so intense that the increasing trend during the T_1 period cannot be recognized, and the corresponding peaks representing the random noises also cannot be found. As for the A-weighted SPL curve, the overall value increase during the T_1 period is well computed with a narrow fluctuation range throughout the whole recording process. However, the trends during the T_2 period and the stabilized overall value after 240 s are evidently inconsistent with the actual running states. By contrast, the trend of the overall loudness curve is well matched with the different running states of the high-speed train, and the appearance of loudness peaks, such as the one at T_3, is also in accord with the auditory assessment of the researchers. The result confirms that Moore loudness is the most appropriate parameter for evaluating the sudden sound quality variation inside the carriage.

To obtain a further understanding of the sudden variation of the Moore loudness, the corresponding time-erbr color map is shown in Fig. 17. The specific loudness components among 12–15 erbr are relatively stable. They may be caused by some cooling device of the traction system underneath the carriage, which would launch and work continuously once the running speed reaches the critical value. That is why the overall loudness increases in an evident manner during the T_1 period in Fig. 16. As for the specific loudness above 20 erbr, the major components shift gradually from 20 erbr to 27 erbr along with the acceleration process, and vary in the opposite way when decelerating. In addition, an amplitude reduction throughout the whole erbr range appears during the T_2 period and leads to the overall loudness decreasing accordingly. Furthermore, as stated in Section 4.1, since the major specific loudness components below 8.7 erbr present an obvious rise after 240 s, it is inferred that the high-speed train is running close to 385 km/h. As for the random noise at T_3, it is a broad-erbr-range component, which may be caused by the instantaneous impact between the wheel and rail because of the track irregularity.

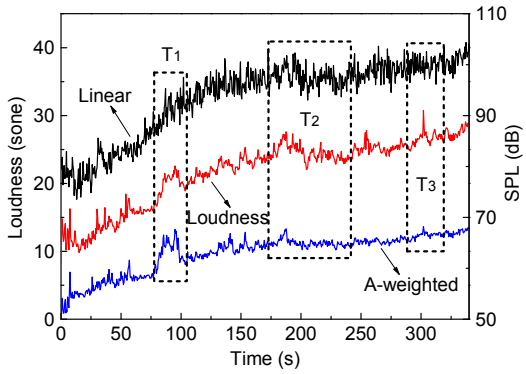


Fig. 16 Overall value variation as a function of time under transient conditions

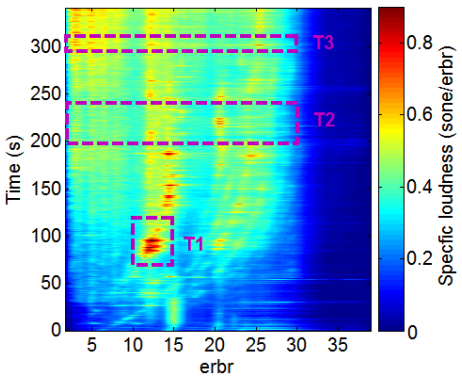


Fig. 17 Time-erbr color map of the specific loudness under transient conditions

To sum up, AMLA is also appropriate for the sound quality evaluation under transient running conditions. With the distribution color map of the specific loudness, an unusual noise precaution and low noise improvement can be focused on the frequency band corresponding to the certain erbr.

5 Conclusions

This paper proposes a signal-adaptive Moore loudness algorithm namely AMLA, and applies it to evaluate the interior sound quality of a CRH train. The innovative points and conclusions are summarized as follows:

1. The core idea of AMLA is the signal-adaptive ERB spectrum, which retains the nonlinear auditory characteristics of human ears. Compared with the existing algorithms based on CPB or FFT spectrum, AMLA can obtain higher accuracy and efficiency.

2. The interior noise of the high-speed train is mainly composed of middle-low frequency components, which has similar auditory feeling with the dark red noise. AMLA can effectively distinguish the interior noise from the colored noises with the same A-weighted SPL value.

3. The interior sound quality under different steady conditions is analyzed in depth comprehensively by AMLA, for the first time in the area of high-speed train research. Results show that the main loudness component of the interior noise is below 27.6 erbr, and the sound quality is relatively stable when the train runs at 300–350 km/h.

4. The specific loudness components among 12–15 erbr remain invariable throughout the acceleration or deceleration process, while components among 20–27 erbr are speed-related. The overall loudness curve and specific loudness color map can effectively identify the unusual random noises under transient conditions. This should be helpful for the next generation high-speed train design to provide low noise and high comfort.

References

ANSI (American National Standards Institute), 2005. Procedure for the Computation of Loudness of Steady Sounds, ANSI S3.4-2005. National Standards of America.

Cook, V.G.C., Ali, A., 2012. End-of-line inspection for annoying noises in automobiles: trends and perspectives. *Applied Acoustics*, **73**(3):265-275. http://dx.doi.org/10.1016/j.apacoust.2011.06.019

Deng, Y., Xiao, X., He, B., *et al.*, 2014. Analysis of external noise spectrum of high-speed railway. *Journal of Central South University*, **21**(12):4753-4761. http://dx.doi.org/10.1007/s11771-014-2485-3

Ding, J.J., Pei, S.C., 2013. Heisenberg's uncertainty principles for the 2-D nonseparable linear canonical transforms. *Signal Processing*, **93**(5):1027-1043. http://dx.doi.org/10.1016/j.sigpro.2012.11.023

Fletcher, H., Munson, W.A., 1933. Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America*, **5**(2):82-108. http://dx.doi.org/10.1121/1.1915637

Glasberg, B.R., Moore, B., 2002. A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, **50**(5):331-342.

Gu, X.A., 2006. Railway environmental noise control in China. *Journal of Sound and Vibration*, **293**(3-5):1078-1085. http://dx.doi.org/10.1016/j.jsv.2005.08.058

Hellman, R., Zwicker, E., 1987. Why can a decrease in dB(A) produce an increase in loudness? *Journal of the Acoustical Society of America*, **82**(5):1700-1705. http://dx.doi.org/10.1121/1.395162

ISO (International Organization for Standardization), 1975. Acoustic-method for Calculation Loudness Level, ISO 532:1975. ISO.

Jiao, Z.X., Liu, W., He, L.S., 2012. Three methods for calculating Moore's loudness. *China Measurement and Test*, **38**(1):5-8 (in Chinese).

Jin, X.S., 2014. Key problems faced in high-speed train operation. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, **15**(12):936-945. http://dx.doi.org/10.1631/jzus.A1400338

Mao, J., Hao, Z.Y., Zheng, K., *et al.*, 2013. Experimental validation of sound quality simulation and optimization of a four-cylinder diesel engine. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, **14**(5):341-352. http://dx.doi.org/10.1631/jzus.A1300055

Matsumoto, A., Sato, Y., Ohno, H., *et al*., 2005. Improvement of bogie curving performance by using friction modifier to rail/wheel interface: verification by full-scale rolling stand test. *Wear*, **258**(7-8):1201-1208. http://dx.doi.org/10.1016/j.wear.2004.03.063

Mellet, C., Létourneaux, F., Poisson, F., *et al*., 2006. High speed train noise emission: latest investigation of the aerodynamic/rolling noise contribution. *Journal of Sound and Vibration*, **293**(3-5):535-546. http://dx.doi.org/10.1016/j.jsv.2005.08.069

Moore, B., Glasberg, B.R., Baer, T., 1997. A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, **45**(4):224-240.

Ning, J., Lin, J., Zhang, B., 2016. Time–frequency processing of track irregularities in high-speed train. *Mechanical Systems and Signal Processing*, **66-67**:339-348.

http://dx.doi.org/10.1016/j.ymssp.2015.04.031

Noh, H., Choi, S., Hong, S., *et al*., 2014. Investigation of noise sources in high-speed trains. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **228**(3):307-322. http://dx.doi.org/10.1177/0954409712473095

Park, B., Jeon, J., Choi, S., *et al.*, 2015. Short-term noise annoyance assessment in passenger compartments of high-speed trains under sudden variation. *Applied Acoustics*, **97**:46-53. http://dx.doi.org/10.1016/j.apacoust.2015.04.007

SAC (Standardization Administration of the People's Republic of China), 2006. The Limiting Value and Measurement Method for the Interior Noise in the Railway Passenger Coach, GB/T 12816-2006. National Standards of the People's Republic of China (in Chinese).

SAC (Standardization Administration of the People's Republic of China), 2007. Acoustics–Normal Equal-loudness-level Contours, GB/T 4963-2007. National Standards of the People's Republic of China (in Chinese).

Sone, S., 2015. Comparison of the technologies of the Japanese Shinkansen and Chinese high-speed railways. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, **16**(10):769-780. http://dx.doi.org/10.1631/jzus.A1500220

Soeta, Y., Shimokura, R., 2013. Survey of interior noise characteristics in various types of trains. *Applied Acoustics*, **74**(10):1160-1166. http://dx.doi.org/10.1016/j.apacoust.2013.04.002

Stevens, S.S., 1956. Calculation of the loudness of complex noise. *Journal of the Acoustical Society of America*, **28**(5):807-832. http://dx.doi.org/10.1121/1.1908487

Tan, P., Ma, J.E., Zhou, J., *et al.*, 2016. Sustainability development strategy of China's high speed rail. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, **17**(12):923-932. http://dx.doi.org/10.1631/jzus.A1600747

Zhang, J., Xiao, X.B., Wang, D., *et al.*, 2012. Characteristics and evaluation of noises in the tourist cabin of a train running at more than 350 km/h. *Journal of the China Railway Society*, **34**(10):23-29 (in Chinese). http://dx.doi.org/10.3969/j.issn.1001-8360.2012.10.004

Zhang, X., Li, X., Hao, H., *et al.*, 2016. A case study of interior low-frequency noise from box-shaped bridge girders induced by running trains: its mechanism, prediction and countermeasures. *Journal of Sound and Vibration*, **367**:129-144. http://dx.doi.org/10.1016/j.jsv.2016.01.004

Zheng, X., Hao, Z.Y., Wang, X., *et al.*, 2016. A full-spectrum analysis of high-speed train interior noise under multi-physical-field coupling excitations. *Mechanical Systems and Signal Processing*, **75**:525-543. http://dx.doi.org/10.1016/j.ymssp.2015.12.010

Zhou, J., Liu, D., Li, X., *et al.*, 2012. Pink noise: effect on complexity synchronization of brain activity and sleep consolidation. *Journal of Theoretical Biology*, **306**:68-72. http://dx.doi.org/10.1016/j.jtbi.2012.04.006

Zwicker, E., 1956. On the loudness of continuous noises. *The Journal of the Acoustical Society of America*, **28**(4):764. http://dx.doi.org/10.1121/1.1905031

中文概要

题　目：基于自适应 Moore 响度算法研究高速列车车内声品质

目　的：高速列车的车内噪声以中低频为主，传统的线性和 A 计权声压级都无法客观描述人耳的听觉感受。本文旨在探索 Moore 响度应用于车内声品质分析的可行性。

创新点：1. 提出了一种自适应 Moore 响度算法（AMLA），该算法可有效提升计算的精度和效率；2. 采用 AMLA 分析了高速列车车内噪声在不同工况下的声品质特征。

方　法：1. 基于信号的等矩形带宽（ERB）谱，提出 AMLA 方法的理论；2. 参照 ANSI 标准中的仿真信号，评价 AMLA 的计算精度和效率；3. 采用 AMLA 辨别有色噪声信号与车内噪声样本，验证声品质分析的有效性；4. 结合在线搭载试验，运用 AMLA 分析稳态工况（不同行车速度和空间位置等）和瞬态工况（加速和减速等）下的车内声品质特征。

结　论：1. 相比传统方法，AMLA 方法的计算精度和效率相对较高，且适用范围更广；2. 高速列车的车内噪声与深红噪声信号具有相似的特征响度分布；3. 稳态工况下，车内噪声的响度成分集中在 27.6 erbr 以内，在 300~350 km/h 的速度区间内车内声品质较稳定，空间分布特征为“端部大、中间小”；4. 瞬态工况下，车内噪声在 20~27 erbr 内的响度成分与列车速度密切相关，而在 12~15 erbr 内的成分相对稳定。

关键词：高速列车；声品质评价；等矩形带宽谱；自适应 Moore 响度算法；车内异响