



POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH
COMMITTEE ON ACOUSTICS

ARCHIVES of ACOUSTICS

QUARTERLY

Vol. 48, No. 2, 2023

WARSAW



ARCHIVES of ACOUSTICS

QUARTERLY, Vol. 48, No. 2, 2023

Research Papers

- W. Huang, J. Peng, T. Xie, *Study on Chinese speech intelligibility under different low-frequency characteristics of reverberation time using a hybrid method* 151
- A.R. Al-Obaidi, *Experimental diagnostic of cavitation flow in the centrifugal pump under various impeller speeds based on acoustic analysis method* 159
- Z. Hashemi, A. Fahim, M.R. Monazzam, *Investigation of the effect of non-flat surfaces on the performance of perforated acoustic absorber* 171
- J. Nurzyński, *The acoustic effect of windows installed in a wood frame façade* 183
- Q. Lu, M. Wang, M. Li, *Infrasound signal classification based on ICA and SVM* 191
- P. Nieradka, A. Dobrucki, *Study on the effectiveness of Monte Carlo filtering when correcting negative SEA loss factors* 201
- A. Sakalauskas, R. Jurkonis, A. Lukoševičius, *Three-dimensional freehand ultrasound strain elastography based on the assessment of endogenous motion: Phantom study* 219
- X. Yuan, C. Hao, Q. Jiang, *Effect of rotation on the piezoelectric wave impedance characteristics* 231
- Y. Yang, J. Fan, B. Wang, *Research on scattering feature extraction of underwater moving cluster targets based on the highlight model* 235

Review Papers

- M. Bujacz, B. Szttyler, N. Wileńska, K. Czajkowska, P. Strumiłło, *Review of methodologies in recent research of human echolocation* 249
- D.I. Popescu, *A study of the Romanian framework and the challenges in implementing the noise mapping legislation* 273

Technical Note

- R. Makarewicz, M. Buszkiewicz, *Corrected long-term time-average sound level of amplitude-modulated wind turbine noise* 281

Editorial Board

Editor-in-Chief: NOWICKI Andrzej (Institute of Fundamental Technological Research PAS, Poland)

Deputy Editor-in-Chief: WÓJCIK Janusz (Institute of Fundamental Technological Research PAS, Poland)

Associate Editors

General linear acoustics and physical acoustics

RDZANEK Wojciech P. (University of Rzeszów, Poland)

SNAKOWSKA Anna (AGH University of Science and Technology, Poland)

Architectural acoustics

KAMISIŃSKI Tadeusz (AGH University of Science and Technology, Poland)

MEISSNER Mirosław (Institute of Fundamental Technological Research PAS, Poland)

Musical acoustics and psychological acoustics

MISKIEWICZ Andrzej (The Fryderyk Chopin University of Music, Poland)

PREIS Anna (Adam Mickiewicz University, Poland)

Underwater acoustics and nonlinear acoustics

GRELOWSKA Grażyna (Gdańsk University of Technology, Poland)

MARSZAŁ Jacek (Gdańsk University of Technology, Poland)

Speech, computational acoustics, and signal processing

KOCIŃSKI Jędrzej (Adam Mickiewicz University, Poland)

Ultrasonics, transducers, and instrumentation

OPIELIŃSKI Krzysztof (Wrocław University of Science and Technology, Poland)

WÓJCIK Janusz (Institute of Fundamental Technological Research PAS, Poland)

Sonochemistry

DZIDA Marzena (University of Silesia in Katowice, Poland)

Electroacoustics

ŻERA Jan (Warsaw University of Technology, Poland)

Vibroacoustics, noise control and environmental acoustics

ADAMCZYK Jan Andrzej (Central Institute for Labor Protection – National Research Institute, Poland)

KLEKOT Grzegorz (Warsaw University of Technology, Poland)

KOMPAŁA Janusz (Central Mining Institute, Poland)

LENIOWSKA Lucyna (University of Rzeszów, Poland)

PIECHOWICZ Janusz (AGH University of Science and Technology, Poland)

PLEBAN Dariusz (Central Institute for Labor Protection – National Research Institute, Poland)

Journal Managing Editor: JEZIEWSKA Eliza (Institute of Fundamental Technological Research PAS, Poland)

Advisory Editorial Board

Chairman: WESOŁOWSKI Zbigniew (Polish Academy of Science, Poland)

BATKO Wojciech (AGH University of Science and Technology, Poland)

BLAUERT Jens (Ruhr University, Germany)

BRADLEY David (The Pennsylvania State University, USA)

CROCKER Malcolm J. (Auburn University, USA)

DOBRUCKI Andrzej (Wrocław University of Science and Technology, Poland)

HANSEN Colin (University of Adelaide, Australia)

HESS Wolfgang (University of Bonn, Germany)

KOZACZKA Eugeniusz (Gdańsk University of Technology, Poland)

LEIGHTON Tim G. (University of Southampton, UK)

LEWIN Peter A. (Drexel University, USA)

MAFFEI Luigi (Second University of Naples SUN, Italy)

PUSTELNY Tadeusz (Silesian University of Technology, Poland)

SEREBRYANY Andrey (P.P. Shirshov Institute of Oceanology, Russia)

SUNDBERG Johan (Royal Institute of Technology, Sweden)

ŚLIWIŃSKI Antoni (University of Gdańsk, Poland)

TITTMANN Bernhard R. (The Pennsylvania State University, USA)

TORTOLI Piero (University of Florence, Italy)

VORLÄNDER Michael (Institute of Technical Acoustics, RWTH Aachen, Germany)

Polish Academy of Sciences
Institute of Fundamental Technological Research PAS
Committee on Acoustics

Editorial Board Office

Pawińskiego 5B, 02-106 Warsaw, Poland

phone (48) 22 826 12 81 ext. 206

e-mail: akustyka@ippt.pan.pl <https://acoustics.ippt.pan.pl>

Indexed and abstracted in Science Citation Index Expanded (also known as SciSearch®)
and Journal Citation Reports/Science Edition

Abstracted in Acoustics Abstracts, in Applied Mechanics Reviews
and indexed in Journal of the Acoustical Society of America

Recognised by the International Institute of Acoustics and Vibration, I IAV

Edition co-sponsored by the Ministry of Science and Higher Education

PUBLISHED IN POLAND

Typesetting in L^AT_EX: JEZIEWSKA Katarzyna (Institute of Fundamental Technological Research PAS, Poland)

Research Paper

Study on Chinese Speech Intelligibility Under Different Low-Frequency Characteristics of Reverberation Time Using a Hybrid Method

Wuqiong HUANG^{(1),(2)*}, Jianxin PENG^{(1)*}, Tinghui XIE⁽³⁾

⁽¹⁾ *School of Physics and Optoelectronics, South China University of Technology
Guangzhou, China*

⁽²⁾ *State Key Laboratory of Subtropical Building Science, South China University of Technology
Guangzhou, China*

⁽³⁾ *School of Architecture and Art, Shijiazhuang Tiedao University
Shijiazhuang, China*

*Corresponding Author e-mail: hwq2020@scut.edu.cn (Wuqiong Huang); phjxpeng@163.com (Jianxin Peng)

(received March 15, 2022; accepted November 19, 2022)

Reverberation time (RT) is an important indicator of room acoustics, however, most studies focus on the mid-high frequency RT, and less on the low-frequency RT. In this paper, a hybrid approach based on geometric and wave methods was proposed to build a more accurate and wide frequency-band room acoustic impulse response. This hybrid method utilized the finite-difference time-domain (FDTD) method modeling at low frequencies and the Odeon simulation at mid-high frequencies, which was investigated in a university classroom. The influence of the low-frequency RT on speech intelligibility was explored. For the low-frequency part, different impedance boundary conditions were employed and the effectiveness of the hybrid method has also been verified. From the results of objective acoustical parameters and subjective listening experiments, the smaller the low-frequency RT was, the higher the Chinese speech intelligibility score was. The syllables, consonants, vowels, and the syllable order also had significant effects on the intelligibility score.

Keywords: low frequency; speech intelligibility; classroom; finite-difference time-domain method.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Concerning the reverberation time (RT) in a room, most studies only pay attention to the mid-high frequency RT, less care about the low-frequency part of RT in a room. RT is the primary index of the acoustical design for all kinds of halls, but the requirement of RT in the low-frequency range is still controversial. Around the 1960s, Beranek defined the RT ratio of low frequency (125–250 Hz) and medium frequency (500–1000 Hz) as the bass ratio (BR), and put forward to the ideal value of 1.1–1.5 for BR (BERANEK, 1962), namely the bass-rise reverberation characteristic. This bass-rise characteristic has been seen as desirable or at least tolerated in auditoria, especially in the USA (BARRON, 2010). However, a flatten RT curve has been more favorable in Europe recently, even BERANEK (2010)

questioned himself after numerous and elaborate investigations. After measuring many performing venues with good sound quality, FUCHS and STEINKE (2015) found that these buildings had a relatively flat frequency curve of RTs, hence they suggested that the BR close to 1 was more favorable to the low-frequency sound. In addition, ADELMAN-LARSEN (2015) emphasized the necessity to improve clarity by controlling the low-frequency RT in his analysis of large-scale venues.

The low-frequency sound in an auditorium can increase the feeling of warmth in the hall (BERANEK 1996), however intelligibility is more important than warmth in a speech hall, such as a classroom. In China, most of the Chinese acoustic standards are still based on the characteristic of the bass-rise RT (GB/T, 2005; GB, 2010; JGJ/T, 2012). Moreover, there are few sound standards for frequency below 500 Hz. As we

know, Chinese is a tonal language, which is different from the intonation language of western countries. For Chinese, vowels are longer than consonants, as the low-frequency vowels are easy to mask the mid-high frequency consonants, which ultimately affect the Chinese speech intelligibility in a room. Some studies have demonstrated acoustic problems by applying low-frequency sound-absorbing structures and sound-absorbing materials into actual buildings. By adding low-frequency sound absorption structures in the classroom, ZHA and LYU (2020) reduced the low-frequency noise and flattened the low-frequency RT characteristics, in which way they obtained the satisfactory sound environment for a classroom. PENG *et al.* (2020) compared the objective parameters and subjective speech perception in two classrooms with similar RTs, but one classroom had a better speech perception than the other. They suggested that this discrepancy might be due to the difference in a low-frequency RT or a background noise level. In the follow-up study, XU *et al.* (2021) used the Odeon software to calculate RTs of these two classrooms and then carried out the Chinese speech intelligibility listening test by a headphone, which eliminated the influence of background noise and confirmed that reducing the low-frequency RTs was helpful to improve the Chinese speech intelligibility in the classroom.

The room acoustic simulation is an important part of the architectural design process, which is convenient and cost-saving. However, popular methods of predicting room acoustic characteristics are based on a geometric acoustic model of ray-like sound propagation, considering that these kinds of software are suitable for small wavelengths in comparison to the dimensions of the enclosure and internal objects. At larger wavelengths, the ray-like assumption no longer holds and the phenomena such as diffraction caused by a low-frequency acoustic wave cannot be ignored (SOUTHERN *et al.*, 2013). Solving the low-frequency sound problem the method based on the acoustic wave theory should be used. Moreover, the finite-difference time-domain (FDTD) method can simulate the frequency-dependent boundaries with desired sound absorption characteristics through the digital impedance filter (KOWALCZYK, VAN WALSTIJN, 2008). As the FDTD method has become more mature, this method is showing its superiority in both simulation accuracy and calculation speed (BOTTELDOOREN, 1995; OXNARD, 2018). Even so, the FDTD method is very memory-intensive, especially when modeling large volumes or a wide frequency bandwidth such as the range of hearing. Thus, this work used a combined method of FDTD and a geometric method to get a synthetic wide frequency-band room impulse response (RIR), which includes the 63 Hz octave frequency band.

This study aims to investigate the effect of different RTs in low frequencies on speech intelligibility. It is

better to control RTs in the mid-high frequencies by modeling methods and change only the RTs in low frequencies. Through the combined method four kinds of low-frequency reverberation characteristics were established, and the differences in Chinese speech intelligibility in a classroom before and after the improvement were compared.

2. Experimental method

2.1. Room

In universities, most classrooms are large with a rectangular shape. In the study, a large classroom with volume of $15.82 \times 8.22 \times 4.9 \text{ m}^3$ was selected in the university, which has two windows and two metal doors on the left side, four windows on the right side and a blackboard on the front wall. All the walls and the ceiling are of plastered brick, the floor is covered with ceramic tiles, and the seats and desks are composed of multi-plywood. The RIRs of six receiving positions in the classroom were recorded by using the B&K 4189 microphone, and using the B&K 4296 dodecahedral loudspeaker as an omnidirectional sound source. The sound source (S) with a height of 1.4 m from the ground and all the receiving positions (R1–R6) with the same height of 1.2 m were arranged as shown in Fig. 1. During the measurement, the doors and windows were closed, and the RIRs were measured by using the swept-frequency method in the unoccupied classroom. Whereafter, RIRs were calculated by Dirac4.1 software, and objective acoustic parameters such as RT were obtained. After the calculation, the measured average RT of the six receiving positions was 3 s at 63 Hz, rising to 3.7 s at 125 Hz, and then decreasing gradually. In the Odeon model, the sound absorption coefficients were basically set according to the material of each surface in the actual classroom, while slightly adjusted to make the average RT of each receiving position close to the measured average RT. The sound absorption coefficients of all surfaces in the Odeon model are listed in Table 1.

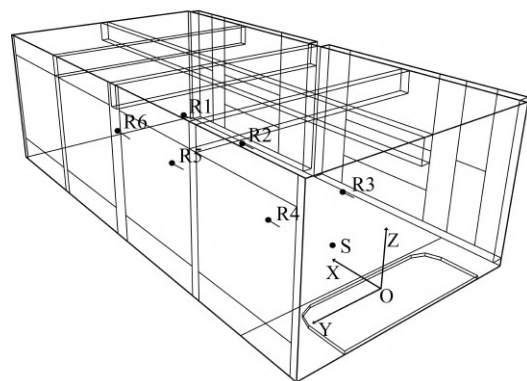


Fig. 1. Location of source and receiving points in the Odeon model.

Table 1. Sound absorption coefficients of the classrooms in the Odeon models.

Frequency [Hz]	500	1000	2000	4000
Floor	0.03	0.04	0.04	0.05
Glass windows	0.03	0.03	0.02	0.02
Seats	0.32	0.29	0.35	0.32
Front wall	0.03	0.04	0.04	0.07
Door	0.01	0.01	0.02	0.02
Ceiling	0.02	0.02	0.02	0.02
The walls of Model A&B	0.02	0.02	0.02	0.03
The walls of Model C&D	0.56	0.56	0.56	0.56

2.2. FDTD acoustic model

The formulation of the FDTD scheme used in this study is the numerical solution of the wave equation that governs sound wave propagation in an ideal isotropic medium:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 p, \quad (1)$$

where p denotes sound pressure and c denotes the sound speed which is set to 344 m/s, t is time in seconds and $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the 3D Laplacian operator. FDTD schemes for numerical simulation of the wave equation are derived by approximating time and space derivatives with finite difference operators according to KOWALCZYK and VAN WALSTIJN (2008). Assuming an equal distance between grid points in all directions, the 3D discretized wave equations take the form of Eq. (2):

$$\begin{aligned} p_{l,m,i}^{n+1} = & \kappa^2 (p_{l+1,m,i}^n + p_{l-1,m,i}^n + p_{l,m+1,i}^n \\ & + p_{l,m-1,i}^n + p_{l,m,i+1}^n + p_{l,m,i-1}^n) \\ & + 2(1 - 3\kappa^2) p_{l,m,i}^n - p_{l,m,i}^{n-1}, \end{aligned} \quad (2)$$

where $\kappa = cT/X$ denotes the Courant number, T is the time step, X is the grid spacing, l , m , and i denote the spatial indexes in x , y , and z directions, and n is the time index. To ensure numerical stability in simulations, the stability condition should be satisfied, that is $\kappa \leq \frac{1}{\sqrt{3}}$. Besides, the grid spacing should not be too long, generally less than one tenth of the wavelength. Combined with the stability condition, in this study the grid spacing was set to 0.06 m and the time step was set to 100.7 μ s, and the derivative of a Gaussian function was chosen as the excitation source. Under these conditions, the spectral characteristic of the excitation source is non-flat in the whole frequency band, which will distort the listening material. To eliminate the non-flat effects of the excitation source, the RIRs calculated by the FDTD method were corrected by an inverse-filtering technique (SAKAMOTO *et al.*, 2008).

In general, the reflected wave has a phase and an amplitude that differ from those of the incident wave, and such changes diverge with frequency. Assuming that a digital filter also has a frequency response in amplitude and phase, so the frequency-dependent boundary can be incorporated in a FDTD model with a digital filter. Since the infinite impulse response filter (IIR filter) and the specific impedance of the boundary have a similar form, this study expressed the boundary in terms of IIR filter. KOWALCZYK and VAN WALSTIJN (2008) presented the FDTD formulation of the digital impedance filter (DIF) in a rectilinear grid, and the update formula for a boundary node could be expressed as:

$$\begin{aligned} p_{l,m,i}^{n+1} = & \left[2\kappa^2 (p_{l-1,m,i}^n + p_{l,m-1,i}^n + p_{l,m,i-1}^n) \right. \\ & + \kappa^2 \left(\frac{g_x^n}{b_x} + \frac{g_y^n}{b_y} + \frac{g_z^n}{b_z} \right) + 2(1 - 3\kappa^2) p_{l,m,i}^n \\ & + \left(\frac{\kappa a_x}{b_x} + \frac{\kappa a_y}{b_y} + \frac{\kappa a_z}{b_z} - 1 \right) p_{l,m,i}^{n-1} \left. \right] \\ & \times \left(1 + \frac{\kappa a_x}{b_x} + \frac{\kappa a_y}{b_y} + \frac{\kappa a_z}{b_z} \right)^{-1}, \end{aligned} \quad (3)$$

where a , b , and g have been clarified in (KOWALCZYK, VAN WALSTIJN, 2008) and are not explained here. In this FDTD model, the DIFs combined Bessel high-pass filters and Bessel low-pass filters were used to obtain the sound absorption coefficient of each interface. From the parameters a and b of the filter, the corresponding specific acoustic impedance (ξ) can be obtained. According to the relationship among the specific acoustic impedance, the reflection coefficient and the absorption coefficient (α), the absorption coefficient can be expressed as:

$$\alpha = \frac{4 \operatorname{Re}(\xi)}{|\xi|^2 + 2 \operatorname{Re}(\xi) + 1}. \quad (4)$$

2.3. Verification

The RIRs obtained by the FDTD method were processed by low-pass filter with upper limit frequency of 355 Hz (355 Hz is corresponding to the upper cut-off frequency of 250 Hz octave band), while the RIRs obtained by Odeon software were processed by the high-pass filter with the lower cut-off frequency of 355 Hz. Then, these two kinds of simulation results were combined to obtain synthetic RIRs over the entire audible spectrum.

To verify the accuracy of the DIF boundary, a classroom model named Model A, was established, whose absorption coefficients of all boundaries were adjusted so that the average RTs of the six receiving positions were close to the measurements of the classroom.

Figure 2 exhibits that the simulated average RTs of six receiving positions are basically consistent with the measured results, which is within plus and minus 5% of the measurement. The comparison result has verified the effectiveness of this combined simulation method.

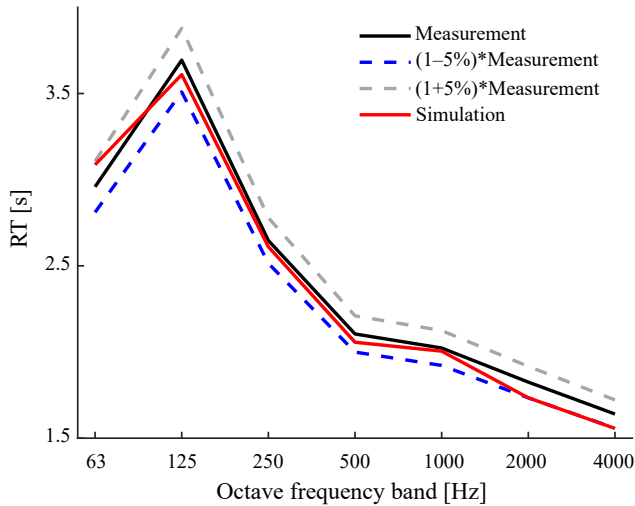


Fig. 2. Comparison of RTs between the measurement and simulation.

Based on Model A, Model B with relatively flat RT at low frequencies was established. According to the specification (GB, 2010), the RT in the classroom larger than 300 m³ should be lower than 0.8 s. Therefore, both of Model C and Model D with an average RT of 0.8 s at mid-high frequency were established. In addition, Model C and Model D have a rising and a flat RT at low frequencies, respectively. The RTs of these four models are shown in Fig. 3 and the sound absorption coefficients of these FDTD models are listed in Table 2. The absorption coefficients of FDTD models were obtained through Eq. (4). Most of the sound absorption processing were for plastered brick walls and

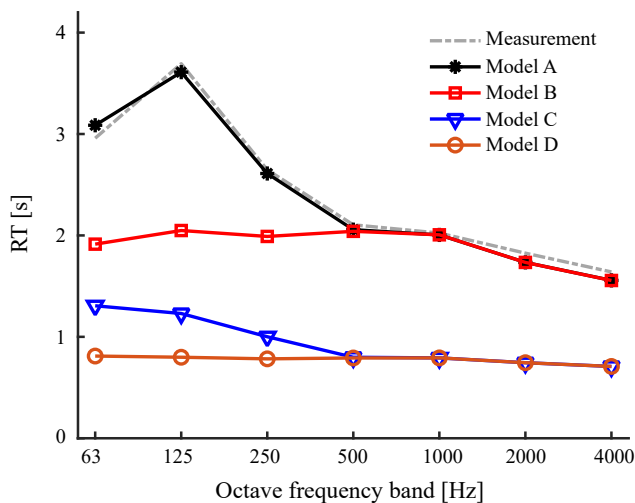


Fig. 3. RTs of measurement and four models.

Table 2. Sound absorption coefficients of the classrooms in the FDTD models. The rests models stand for models with the same absorption coefficient.

Boundary	Model	Frequency [Hz]		
		63	125	250
Floor	A	0.07	0.1	0.15
	D	0.14	0.13	0.14
	Rests	0.1	0.12	0.15
Glass windows	A	0.11	0.07	0.15
	Rests	0.12	0.06	0.2
Seats	D	0.28	0.28	0.34
	Rests	0.24	0.25	0.31
Front wall	A	0.1	0.13	0.2
	Rests	0.12	0.15	0.2
Wall	A	0.26	0.09	0.2
	B	0.25	0.25	0.25
	C	0.51	0.51	0.51
	D	0.88	0.81	0.73
Ceiling	A	0.1	0.07	0.15
	B	0.33	0.33	0.33
	C	0.4	0.4	0.4
	D	0.79	0.72	0.69

ceiling, while other boundaries were just slightly adjusted. From Fig. 3, Model A and Model B have the same RT at 500–4000 Hz, while Model C and Model D have the same RT at 500–4000 Hz. Model A and Model C have a rising RT at 63–500 Hz, while Model B and Model D have a flat RT at 63–500 Hz. It can be seen that this hybrid method can deal with the RT quantitatively, which is beneficial to analyze the effects of low-frequency RT on speech intelligibility.

2.4. Speech intelligibility test

As the simulation results of Model A were consistent with the measured results, this study only conducted listening tests on the simulated models. In a quiet room where the background noise level was less than 30 dBA, subjective evaluation of the speech intelligibility was conducted using recordings of each receiving position. These signals had been processed based on the Mandarin Chinese speech intelligibility test word list as specified by GB 15508-1995 (GB, 1995). Each receiving position used two different lists, which were recorded by a man and a woman reading, respectively. Each list has 75 syllables, which are randomly divided into 25 rows of three syllables with no coherent meaning. The lists keep the balance of difficulty and phonemic characteristic, where each consonant, vowel, or tone appears with the same frequency in each list. Each row is preceded by a carrier phrase, for example “the tenth row is ā, ér, jìng”, where the carrier phrase

gives a hint of row number and “ā, ér, jìng” stands for the three syllables. The words were recorded in an anechoic chamber, and spoken by a male or a female speaker at the rate of about 4 syllables per second. There is a pause of about 5 seconds between each row for the listener to write down the syllables. The testing word lists were convolved with simulated binaural RIRs by Cooledit Pro software.

Fourteen graduate students aged between 22–30 participated in this speech intelligibility test, who were trained and familiar with Chinese phonetic alphabet. All participants are native speakers of Mandarin Chinese and had absolute thresholds of less than 15 dB HL at octave frequencies between 125 Hz and 8000 Hz. For each test condition, up to 4 participants could participate in the test at the same time using the HP-S4 headphone amplifier, and each participant wore the same type of the Sennheiser HD580 headphone at a speech sound pressure level of 60–65 dBA. The listening material was played through Cool Edit Pro software and controlled by the tester. Finally, the results were scored by testers against the correct answer. As the tonal error rate was very low and some subjects accidentally mis-marked tones, the tonal results will not be discussed in the study. Only the consonant and the vowel in each syllable are correct, the syllable is counted as a correct syllable. The correct rate of syllables in each list is calculated by percentage, then the average score of all participants is the Chinese speech intelligibility score of each receiving position. The speech intelligibility score of consonants depends on the correct rate of consonants, no matter whether its vowel part is correct or not. Similarly, the speech intelligibility score of vowels does not consider whether the consonants are right or not.

3. Results

Figure 4 shows the Chinese speech intelligibility scores at each receiving position in four models. It can be obviously seen from the results that the Chinese speech intelligibility scores of Model B at each receiving position are higher than the scores of Model A, and the scores of Model D at each receiving position are higher than Model C, which indicates the intelligibility score of flat RT at a low frequency is higher than that of rising RT at a low frequency. Besides, the standard deviation of the Model A and Model B at each receiving position is basically larger than Model C and Model D. The scores of Model C and Model D are much higher than that of Model A and Model B, which indicates that the RT characteristic in the original classroom is insufficient. From the repeated measurement analysis of variance, the model ($F(3, 234) = 400.927$) and receiving position ($F(5, 78) = 49.377$) have significant effects on the Chinese speech intelligibility scores ($p < 0.001$). As there is a significant interaction be-

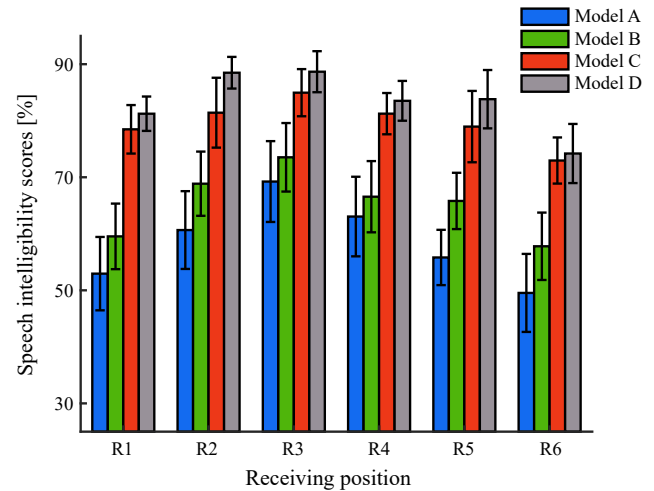


Fig. 4. Chinese speech intelligibility scores in four models.

tween the model and the listening location ($p < 0.001$), it needs paired comparison of the simple effect.

The results of the paired comparison indicates that the speech intelligibility scores of Model C and Model D are significantly higher than that of Model A and Model B ($p < 0.05$). For receiving positions R1, R2, R5, and R6, there is a significant difference between Model A and Model B ($p < 0.05$), whereas there is no significant difference between Model A and Model B for receiving positions R3 and R4 ($p > 0.05$). Meanwhile, between Model C and Model D, there is no significant difference for receiving positions R1, R3, R4, and R6, while there is a significant difference for receiving positions R2 and R5. Therefore, for most locations, flattening the low-frequency characteristics in the classroom with $RT_{(500-1000\text{ Hz})}$ of 2 s can significantly improve the speech intelligibility. However, flattening the low-frequency characteristics in the classroom with $RT_{(500-1000\text{ Hz})}$ of 0.8 s can improve the speech intelligibility but not significantly. Namely, it is more important to improve the entire-frequency RT than improve the low-frequency RT only.

4. Discussion

To further explore the effect of low-frequency RT on speech intelligibility, the following is analyzed in terms of syllables. Mandarin Chinese speech sounds range from very low (about 100–125 Hz) to very high frequencies (above 10 kHz or 12 kHz for some sounds). A Chinese syllable must contain an initial consonant and a vowel, or only contain a vowel. The vowels are low in frequency and high in sound energy, while the initial consonants are much higher in frequency and lower in sound energy (WU, 1964). GELFAND (1998) wrote in his book that low frequencies tended to be effective maskers over a very wide range of frequencies, while higher frequencies were not good maskers of

low frequencies. When the low-frequency RT is much longer than the mid-frequency RT, low-frequency reverberant sounds can be emphasized by room modes and then mask speech sounds in a classroom (WU *et al.*, 2014). For this speech intelligibility test, three syllables are in a row, and each row is preceded by a carrier phrase. Therefore, due to the difference in RTs, syllables in different positions and carrier phrase will affect the test results. The speech intelligibility scores of syllable, the consonant and the vowel (SCV) for tested syllable orders in turn are shown in Fig. 5.

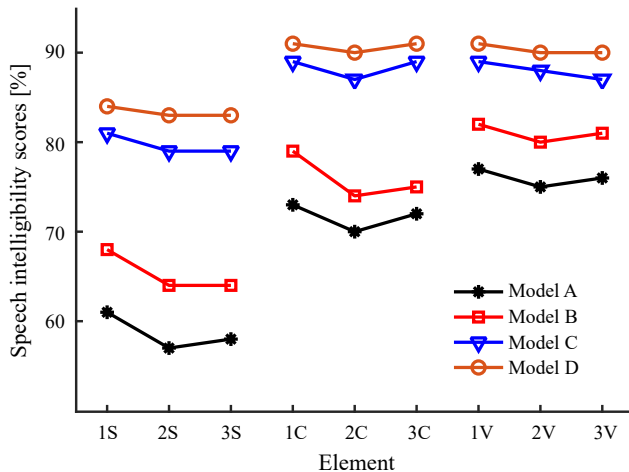


Fig. 5. Speech intelligibility scores of SCV. In the figure, 1S, 2S, and 3S denote syllable in the first, second, and third order, respectively; 1C, 2C, and 3C denote consonant in the first, second, and third order, respectively; 1V, 2V, and 3V denote vowel in the first, second, and third order, respectively.

The analyses of variance for average speech intelligibility scores of six receiving positions show that the model ($F(3, 468) = 614.176$), SCV ($F(2, 468) = 419.248$) and the orders ($F(2, 468) = 14.972$) have significant effects on the test ($p < 0.001$). Figure 5 shows that the scores of Model B are better than Model A, and Model D is better than Model C, which emphasized that the lower the low-frequency RT, the higher the speech intelligibility score. Due to the carrier phrase before the first syllables is fixed and slightly stopped, the first syllables will be less affected by the preceding syllables. Hence, the scores of syllables, consonants and vowels in first orders are significantly higher than the second and third orders ($p < 0.001$). The scores of vowels and consonants are significantly higher than syllables ($p < 0.001$), meanwhile the scores of vowels are significantly higher than consonants ($p < 0.001$).

There is a significant interaction between the model and SCV ($p < 0.001$), and the other interactions are not significant ($p > 0.05$). For the same element of SCV, there is a significant difference between each two models ($p < 0.05$), however, there is not significant difference between each two elements of SCV. For Mod-

els A and B, there is a significant difference between each two elements of SCV ($p < 0.001$). In Model A and Model B, the average score of consonants is lower than the average score of vowels in the same model (the consonant of Model A = 72%, the vowel of Model A = 76%, the consonant of Model B = 76%, the vowel of Model B = 81%). As both the average RTs of Model A and Model B at 500–1000 Hz are 2 s, while the reading speed is about 4 syllables per second, the RT is greater than the time interval between two syllables, which will bring serious interference. Since both the low frequencies and higher frequencies have a masking effect on the high frequencies, the front syllables of Model A and Model B have large masking effect on the higher frequency sound of the next syllables. Therefore, the scores of consonants are lower than vowels in Model A and Model B. For Model C and Model D, there is no significant difference between vowels and consonants ($p > 0.05$), but there is a significant difference between vowels and syllables ($p < 0.001$) or between consonants and syllables ($p < 0.001$). In Model C and Model D, the average scores of the initial consonant and vowel of the same model are basically the same (the consonant of Model C = 88%, the vowel of Model C = 88%, the consonant of Model D = 91%, the vowel of Model D = 91%). Owing to that the average RTs of Model C and Model D at 500–1000 Hz are only 0.8 s, the RT is too small to cause an obvious masking effect. Therefore, the scores of consonants are basically the same as that of vowels in Model C and Model D.

The above findings indicate that the vowel sounds have a certain masking effect on the consonant sounds, namely the low-frequency sound has a certain masking effect on the mid-high frequency sound, which confirmed the well-known “upward spread of masking” (OXENHAM, PLACK, 1998). The result of this study showed that too high reverberation time in the low frequency band significantly deteriorates the intelligibility of speech. To improve the speech intelligibility of a classroom, the average RT in the entire frequencies should be smaller to diminish the masking effect and the low-frequency RT is better not to be rising.

5. Conclusions

In this study, four models based on a large classroom were established by a hybrid method. The synthetic RIRs of the classroom models were obtained through the FDTD method modeling at low frequencies and the Odeon simulation at higher frequencies. The Chinese speech intelligibility listening test was conducted after verification. The average scores of each receiving position show that the scores of Model B are better than Model A, and Model D is better than Model C. The scores of syllables, consonants and vowels in first orders are significantly higher than the second and third orders. The scores of vowels and conso-

nants are significantly higher than syllables. The scores of consonants are lower than vowels in Model A and Model B, while the scores of consonants are basically the same as that of vowels in Model C and Model D. The above results indicate that RT in the entire frequencies should be smaller and the low-frequency RT is better to be flat to obtain better speech intelligibility.

References

1. ADELMAN-LARSEN N.W. (2015), Possible acoustic design goals in very large venues hosting live music concerts, *Auditorium Acoustics*, **37**(3): 308–316.
2. BARRON M. (2010), *Auditorium Acoustics and Architectural Design*, 2nd ed., Spon Press, London and New York.
3. BERANEK L.L. (1962), *Music, Acoustics and Architecture*, Wiley, New York.
4. BERANEK L.L. (1996), *Concert and Opera Halls: How They Sound*, Acoustical Society of America, New York.
5. BERANEK L.L. (2010), Listening to the acoustics in concert halls, [in:] *Proceedings of the International Symposium on Room Acoustics (ISRA 2010)*, Melbourne.
6. BOTTELDOOREN D. (1995), Finite-difference time-domain simulation of low-frequency room acoustic problems, *The Journal of the Acoustical Society of America*, **98**(6): 3302–3308, doi: 10.1121/1.413817.
7. FUCHS H.V., STEINKE G. (2015), Requirements for low-frequency reverberation in spaces for music: Part 2: Auditoria for performances and recordings, *Psychomusicology: Music, Mind, and Brain*, **25**(3): 282–293, doi: 10.1037/pmu0000089.
8. GB (1995), GB 15508-1995, Acoustics – Speech articulation testing method [in Chinese], *Standard of PR China*.
9. GB (2010), GB 50118-2010, Code for design of sound insulation of civil buildings [in Chinese], *Standard of PR China*.
10. GB/T (2005), GB/T 50356-2005, Code for architectural acoustical design of theater, cinema and multi-use auditorium [in Chinese], *Standard of PR China*.
11. GELFAND S.A. (1998), *Hearing: An Introduction to Psychological and Physiological Acoustics*, Marcel Dekker, New York.
12. JGJ/T (2012), JGJ/T 131-2012, Specification for acoustical design and measurement of gymnasium and stadium [in Chinese], *Standard of PR China*.
13. KOWALCZYK K., VAN WALSTIJN M. (2008), Modeling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM room acoustics simulations, *Journal of the Audio Engineering Society*, **56**(7/8): 569–584, <https://www.aes.org/e-lib/browse.cfm?elib=14401>.
14. OXENHAM A.J., PLACK C.J. (1998), Suppression and the upward spread of masking, *The Journal of the Acoustical Society of America*, **104**(6): 3500–3510, doi: 10.1121/1.423933.
15. OXNARD S. (2018), Investigating the stability of frequency-dependent locally reacting surface boundary conditions in numerical acoustic models, *The Journal of the Acoustical Society of America*, **143**(4): EL266–EL270, doi: 10.1121/1.5030917.
16. PENG J., LAU S.K., ZHAO Y. (2020), Comparative study of acoustical indices and speech perception of students in two primary school classrooms with an acoustical treatment, *Applied Acoustics*, **164**: 107297, doi: 10.1016/j.apacoust.2020.107297.
17. SAKAMOTO S., NAGATOMO H., USHIYAMA A., TACHIBANA H. (2008), Calculation of impulse responses and acoustic parameters in a hall by the finite-difference time-domain method, *Acoustical Science and Technology*, **29**(4): 256–265, doi: 10.1250/ast.29.256.
18. SOUTHERN A., SILTANEN S., MURPHY D.T., SAVIOJA L. (2013), Room impulse response synthesis and validation using a hybrid acoustic model, *IEEE Transactions on Audio, Speech, and Language Processing*, **21**(9): 1940–1952, doi: 10.1109/TASL.2013.2263139.
19. WU S., PENG J., BI Z. (2014), Chinese speech intelligibility in low frequency reverberation and noise in a simulated classroom, *Acta Acustica united with Acustica*, **100**(6): 1067–1072, doi: 10.3813/AAA.918786.
20. WU Z. (1964), The spectrographic analysis of the vowels and consonants in standard colloquial Chinese [in Chinese], *Acta Acustica*, **1**(1): 33–40, doi: 10.15949/j.cnki.0371-0025.1964.01.006.
21. XU S., PENG J., XIAO Y., HUANG W. (2021), The effect of low frequency reverberation on Chinese speech intelligibility in two classrooms, *Applied Acoustics*, **182**: 108241, doi: 10.1016/j.apacoust.2021.108241.
22. ZHA X., LYU H. (2020), Analysis and improvement of classroom acoustic environment [in Chinese], *Technical Acoustics*, **39**(4): 461–467, doi: 10.16300/j.cnki.1000-3630.2020.04.014.

Research Paper

Experimental Diagnostic of Cavitation Flow in the Centrifugal Pump Under Various Impeller Speeds Based on Acoustic Analysis Method

Ahmed Ramadhan AL-OBAIDI *Faculty of Engineering, Department of Mechanical Engineering, Mustansiriyah University*
Baghdad, Iraq; e-mail: ahmedram@uomustansiriyah.edu.iq

(received July 15, 2021; accepted December 8, 2022)

Condition monitoring in a centrifugal pump is a significant field of study in industry. The acoustic method offers a robust approach to detect cavitations in different pumps. As a result, an acoustic-based technique is used in this experiment to predict cavitation. By using an acoustic technique, detailed information on outcomes can be obtained for cavitation detection under a variety of conditions. In addition, various features are used in this work to analyze signals in the time domain using the acoustic technique. A signal in the frequency domain is also investigated using the fast Fourier method. This method has shown to be an effective tool for predicting future events. In addition, this experimental investigation attempts to establish a good correlation between noise characteristics and cavitation detection in a pump by using an acoustic approach. Likewise, it aims to find a good method for estimating cavitation levels in a pump based on comparing and evaluating different systems.

Keywords: experimental analysis; impeller speed; cavitation; acoustic signals.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Various pump industries have become increasingly interested in the phenomenon of cavitation. Several types of hydraulic machinery may be affected by cavitation. This includes all types of pumps and turbines for water (NELIK, 1999; AL-OBAIDI, 2018; 2019a; 2019b; SPRAKER, 1965). A variety of factors contribute to this problem (GIRDHAR, MONIZ, 2005; ARNOLD, STEWART, 1999; GAUTAM, 2012). As a consequence, all systems become noisier (LOBANOFF, ROSS, 2013; LIU, 2014; ANDERSON *et al.*, 2014). To detect cavitation in a pump, experimental acoustic signals are analyzed, and a variety of frequency ranges are studied (KARASSIK, MCGUIRE, 1998; BEEBE, 2004; AL-FAYEZ *et al.*, 2005). The performance of pumps has been investigated by many researchers using a variety of monitoring techniques. For example, noise was used to investigate cavitation as an indicator of a pump's performance (CHUDINA, 2003). Cavitation occurs in the inlet pump because of several undesired influences such as performance drop, pitting, erosion, and pump damage. The study analyzes acoustic signatures at var-

ious frequencies under different conditions. A higher frequency range is found to enhance cavitation at high mass flow.

Both vibrations and noise were used to monitor cavitation (ČERNETIČ *et al.*, 2008). Pumps were used with a semi-open impeller and a closed impeller, both of which have six blades. Each pump was studied separately for cavitation. The results indicated that every pump has a variety of noise and vibration spectra with many different types of frequencies. In addition, the results showed that the alteration levels for noise and vibration were around 10 to 15 dB with and without cavitation. The vibration method was used to investigate faults in the pump (ALBRAIK *et al.*, 2012). The flow rate was 30 m³/h, the head was 55 m, and the impeller speed was 2900 rpm for a closed impeller pump. Gradually closing the discharge valve and keeping the suction valve open showed that the net positive suction head available (NPSHA) was related to the net positive suction head required (NPSHR). In their study, NPSHR increased with the mass flow. The results also revealed that the vibration level increased as the mass flow grew.

FAROKHZAD *et al.* (2013) conducted a study on relationship between vibration and fault in a pump at various conditions. Two different configurations were found to be faulty: a faulty seal and a broken impeller. The authors found that significant alteration in vibration tendencies occurred with fault. ČUDINA (2003) investigated cavitation using the sound signal and investigated the sound level at a variety of discharge conditions. The results showed that the sound level increased as cavitation became fully developed. The noise in the pump increased or decreased depending on the mass flow and pump speed. Also, instability can happen in the pump due to cavitation, causing pump deterioration performance, pitting in material, and erosion. ČUDINA and PREZELJ (2009) examined cavitation by studying frequency of discrete audible spectra. The results found that the discrete frequency spectrum of audible noise was linked with cavitation in the pump. Moreover, the spectra frequency level peak augmented as the cavitation inception grew, and the maximum value occurred during the cavitation process.

Based on the above literature, it can be noticed that there is a lack of details connected to using various frequency range analyzes to detect cavitation phenomenon. It is essential that cavitations are investigated when the pump performs at high range conditions, for instance, at various mass flow rates and pump speeds. Using condition monitoring type acoustic approach has different advantages. It is easily recorded, evading safety dangers and the need for different ranges of temperature acoustic devices with their connected complications of mounting. The most significant disadvantage of the acoustic approach is that the sur-

rounding noise tends to contaminate the microphone (RAMROOP *et al.*, 2001; SAKTHIVEL *et al.*, 2010; JONES *et al.*, 2006; GRIST, 1998; KAMIEL, 2015). In this experimental work, the methodology used contains a variety of steps, as shown in Fig. 1.

2. Centrifugal pump experimental test rig

Figure 2 illustrates a flow loop system with a pump and sensors; clear PVC (polyvinyl chloride) pipes were used. In addition, the tank was made of plastic. A pump experimental test rig construction consisted of different significant stages. Designing a flow-loop pumping system was the first step. In the second stage, all requirements, such as installing the entire piping system, were conducted. A pressure transducer microphone, water flowmeter, and other equipment were also identified as part of the loop installation sensors.

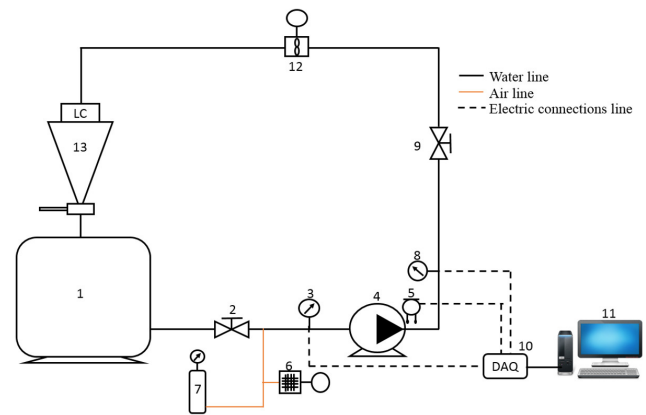


Fig. 2. Experimental setup components: 1 – water tank; 2 – suction valve; 3 – suction pressure transducer; 4 – centrifugal pump; 5 – microphone; 6 – air flowmeter; 7 – air supply; 8 – discharge pressure transducer; 9 – discharge valve; 10 – data acquisition system; 11 – PC; 12 – water flowmeter; 13 – hopper.

2.1. Time domain conventional statistical analysis

Peak value analysis

It is a significant statistical parameter to calculate the peak signal.

RMS value

The given equation defines the RMS value:

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}, \quad (1)$$

where N and x_i represent the total elements' number and the element signal.

Value of peak-to-peak

This is the third statistical parameter used in this work.

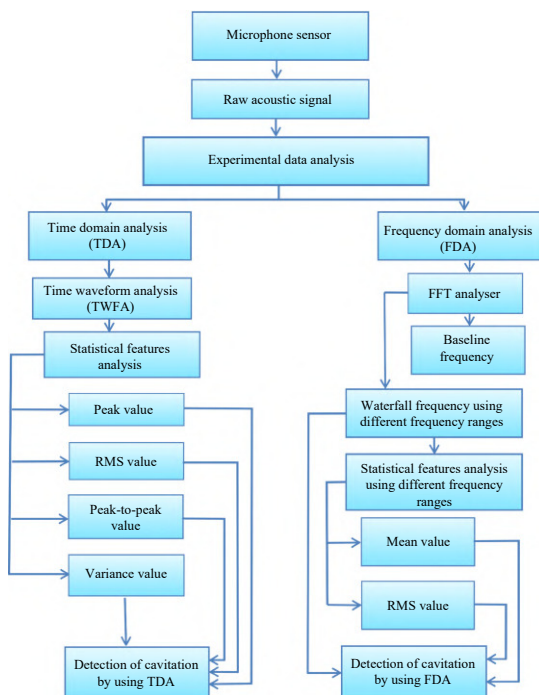


Fig. 1. Flow chart analysis of acoustic data processing.

Variance value

Variance value can be calculated using Eq. (2):

$$\text{Variance} = \frac{\sum (x_i - \bar{x})^2}{N - 1}, \quad (2)$$

where \bar{x} , x_i , and N are the total of elements, the set of elements, and the elements' mean values, respectively.

Frequency conventional analysis

The fast Fourier transform (FFT) is broadly employed to convert the different types of signals from the time analysis into the frequency-domain analysis. The outcome of the FFT offers various amplitudes of signals, and the FFT can be written as follows:

$$F(\omega) = \int_{-\infty}^{\infty} X(t) e^{j2\pi*ft} dt. \quad (3)$$

The inverse function of the FFT is defined as:

$$X(t) = \int_{-\infty}^{\infty} F(\omega) e^{j2\pi*ft} d\omega, \quad (4)$$

$$j = \sqrt{-1}, \quad (5)$$

$$\omega = 2\pi f, \quad (6)$$

where

$$e^{j2\pi*ft} = \cos 2\pi ft + \sin 2\pi ft. \quad (7)$$

$X(t)$ and $F(\omega)$ are the time signals and frequency domain, respectively.

Mean value analysis

The mean value is calculated using the following equation:

$$\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i. \quad (8)$$

The x , x_i , and N , are the total of elements, the set of elements, and the elements' mean values.

Root mean square analysis

From the frequency domain, it can be calculated by the root mean square (RMS) (ANDERSON *et al.*, 2014). The value of the H-NPSHR curve in the pump was provided by its manufacturer and the value of NPSHA, calculated based on varying conditions, is as follows:

$$\text{NPSHA} = \frac{P_{\text{atm}} + P_s}{\rho g} + \frac{V^2}{2g} - \frac{P_v}{\rho g} - H_i, \quad (9)$$

where P_{atm} – atmospheric air pressure [Pa], P_i – inlet pressure [Pa], V – water velocity in pipes [m/s], P_v – the pressure of water vapor [Pa].

3. Instantaneous outlet pressure (hydrostatic pressure) analysis

Figure 3 represents the signals of the outlet pressure (hydrostatic pressure) for various mass flows ranging from 152 to 378 l/min with an impeller speed of 2755 rpm in the centrifugal pump. The outlet pressure signals are changed when the mass flow changes, and the pressure magnitude decreases as the mass flow increases. This resulted from both mechanical and hydraulic losses as well as cavitation conditions. In the next part of this experiment, additional investigations will be necessary to detect cavitation.

Figure 4 describes the tendency of minimum, RMS, mean, and peak features for the amplitude of signals of outlet pressure. It is obvious that the features with the growing flow in the pump all follow a downward continuous tendency because of the same reasons stated before. The tendency for entire features rapidly reduces as the pump works at the mass flow greater than 350 l/min due to the high interaction flow in the impeller and volute parts as well as the cavitation occurrence.

Figure 5 illustrates the pump head measurements at various mass flows. According to experiments, the pump head decreases as the flow increases. Pressure alterations occur when there is interaction flow in the impeller due to the distribution pressure being nonuniform in the volute as well as when there is cavitation.

When it comes to determining normal and cavitation conditions, the NPSH is an influential parameter in the performance of the pump. Figure 6 describes the NPSHA and the NPSHR. The pump's flow in the pump can be altered by gradually throttling the valve at discharge suction and keeping the valve at the suction part fully open.

As can be seen, there are various regions of cavitation. In the first one, at low flow, no cavitation occurs in this area. At mass flow greater than 350 l/min, the cavitation starts to occur and an intersection between both curves has already formed, as shown in this figure. Furthermore, cavitation occurs more frequently at flow rates greater than 300 l/min. Also, it is observed that cavitation occurs between flow rates of 300 and greater than 350 l/min.

4. Acoustic analysis signal in the time domain

In this experimental measurement study, cavitation is investigated at a variety of conditions using various types of features in the time analysis domain. Figure 7 illustrates the acoustic signals in time analysis waves. It is noted that the flow is lower than the flow of 300 l/min. As can be seen, there is no alteration in acoustic amplitude characteristics. The important change in signals occurs at a rate greater than 350 l/min, due to the interaction flow in the volute

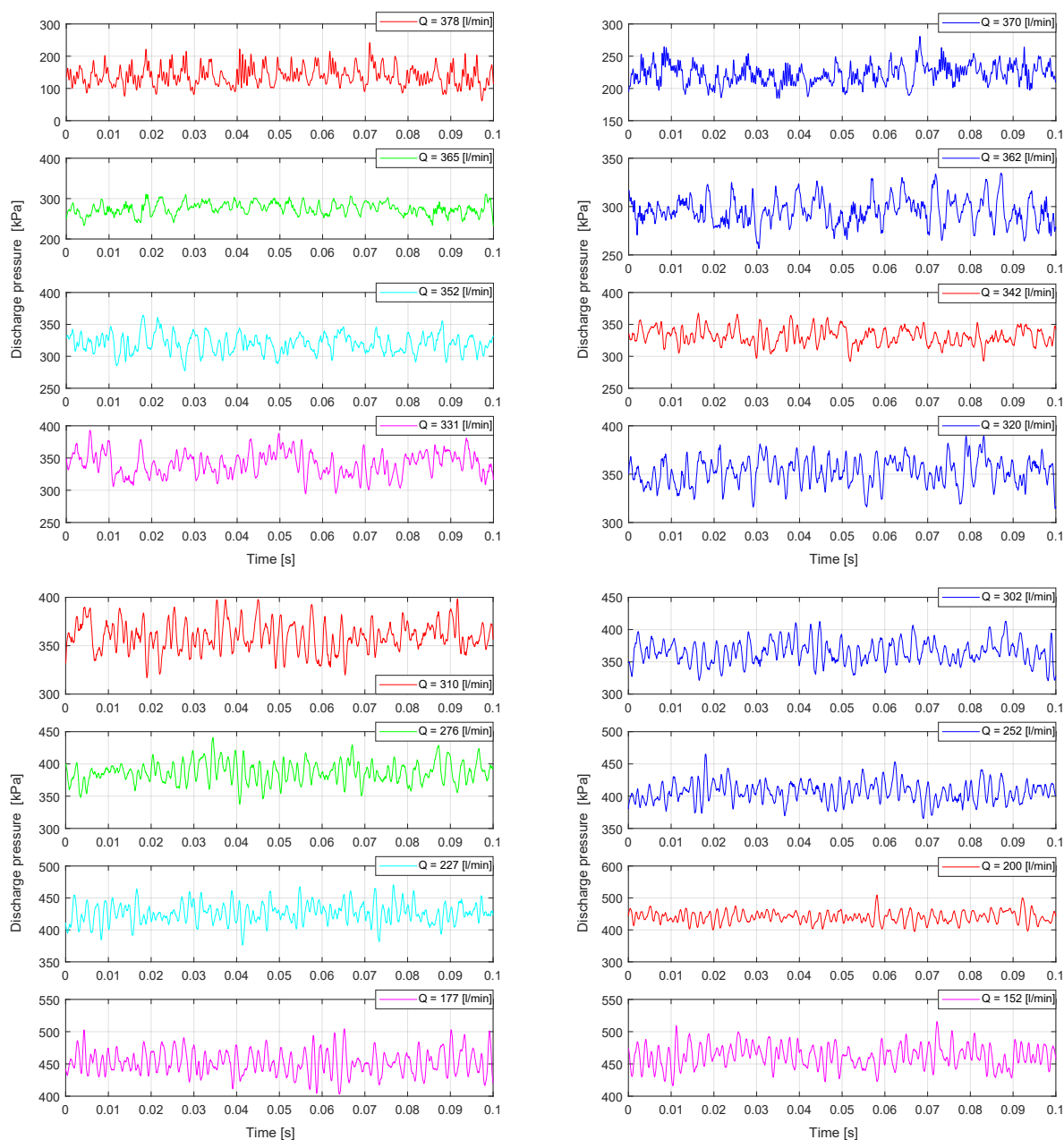


Fig. 3. Various signals of outlet pressure in time analysis.

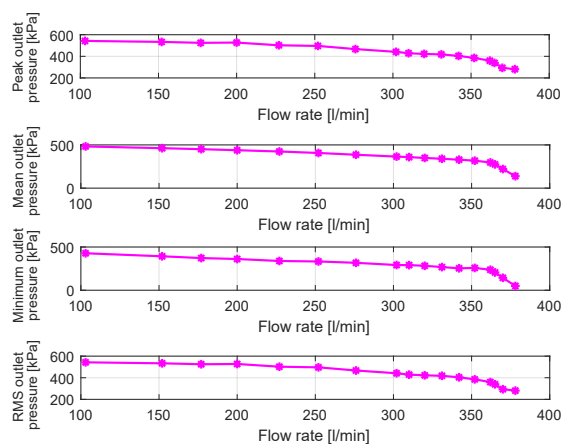


Fig. 4. Tendency of the peak, mean, minimum, and RMS features for the signals of outlet pressure.

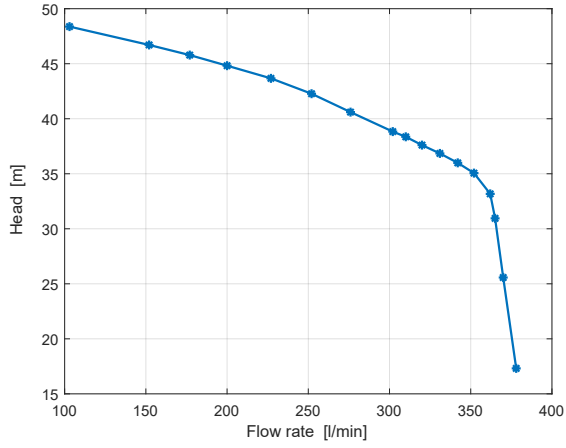


Fig. 5. Pump head at different conditions.

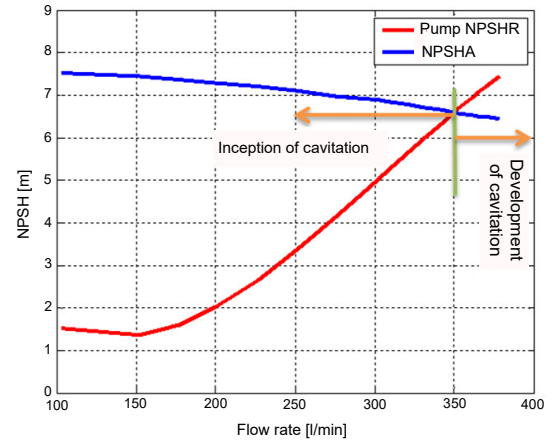


Fig. 6. Characteristics of NPSH with and without cavitation.

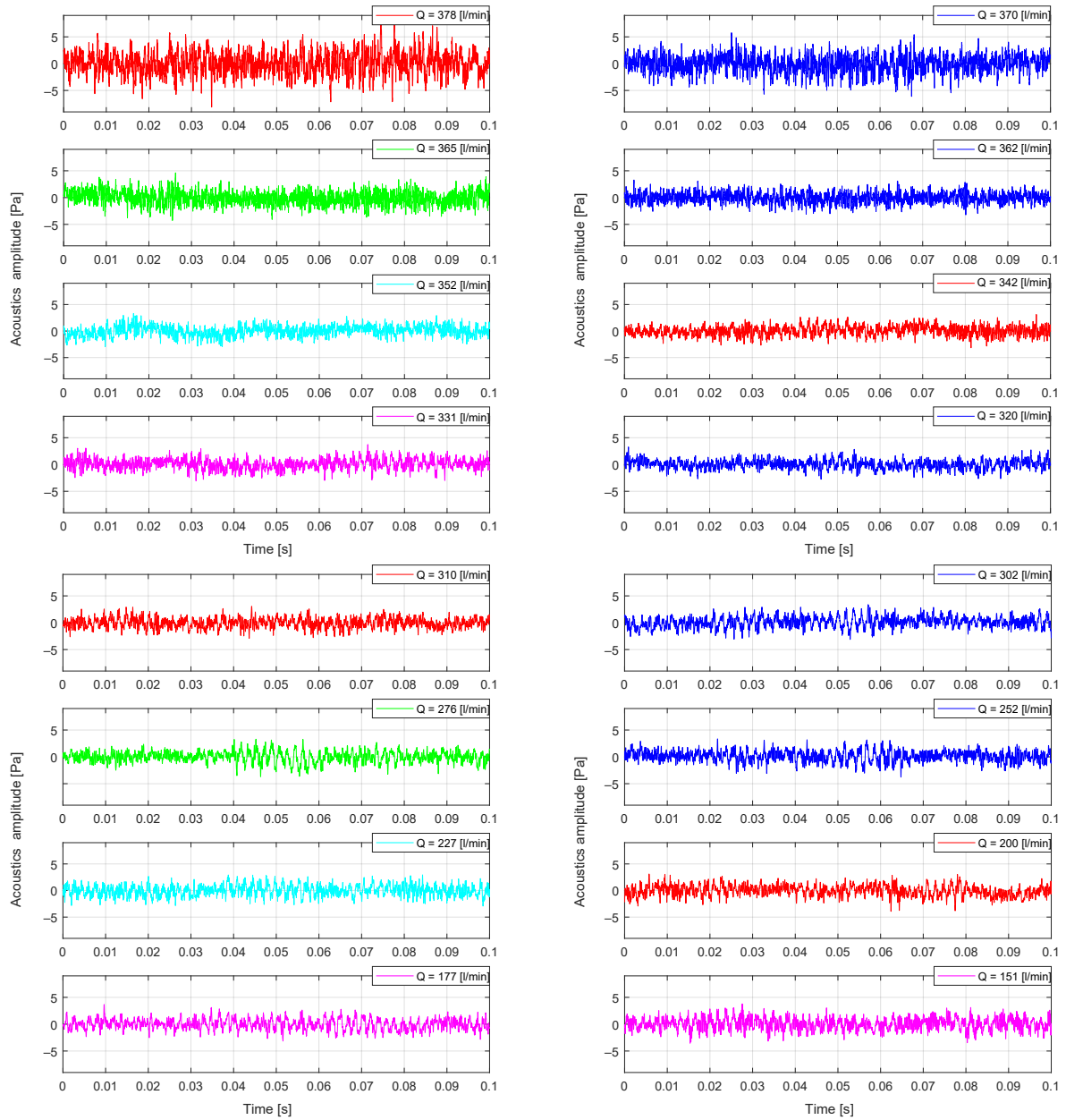


Fig. 7. Acoustic signature in time analysis domain.

and impeller and the cavitation resulting in an NPSHR value higher than NPSHA.

Further analysis conducted to find the relationship between the acoustic signature and the occurrence of cavitation is presented in the next part.

5. Acoustic signals analysis using features

Figure 8a shows the peak and RMS values of microphone signal to investigate cavitation. Because of the inner recirculation flow, there is increased hydraulic noise when the pump runs with a lower flow than the designation flow. Likewise, as the pump functions at a flow greater than 300 l/min, the noise level increases because of additional hydraulic noise, flow turbulence, and the cavitation process. However, when flow increases above 350 l/min, the signature of the acoustic

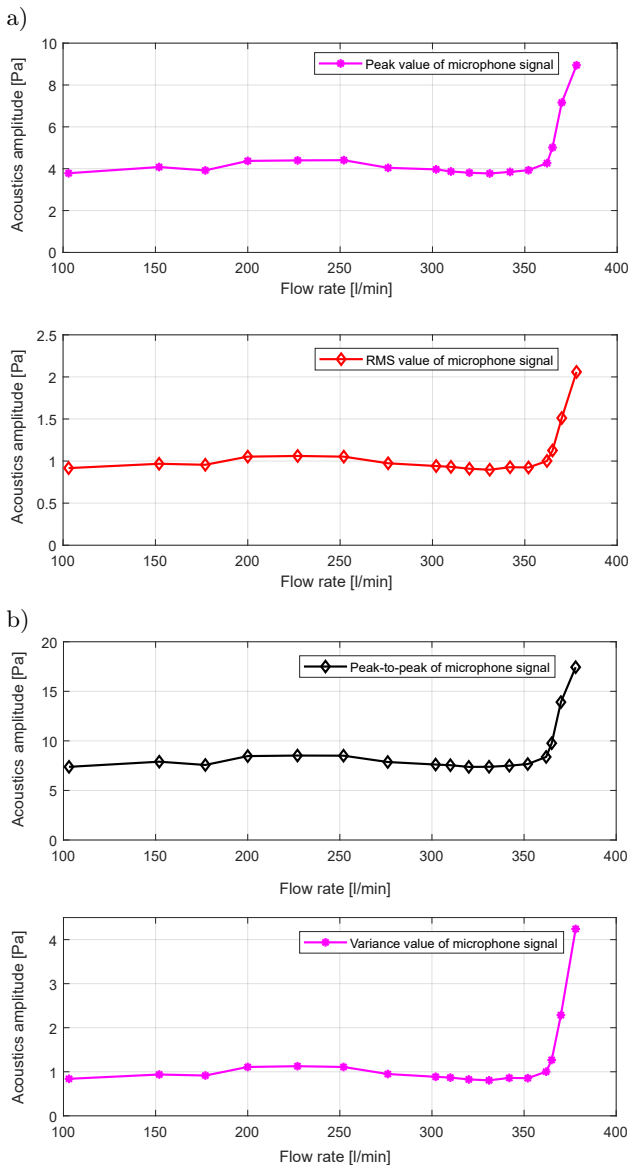


Fig. 8. Tendencies of the acoustics signature using peak, RMS, peak-to-peak, and variance signature features.

signal rapidly increases. The bubbles collapse, consequently altering noise that leads to the NPSHR being higher than the NPSHA due to the cavitation process.

Figure 8b shows an acoustic signature analysis for additional investigation of the acoustic signature using various statistical values of peak-to-peak feature and variance feature. Figure 8b shows that the above features tend to have a similar tendency to those in Fig. 8a. Since the flow inside the pump changes at different range conditions, the noise level increases. According to this observation, the cavitation phenomenon in the pump is highly affected by the mass flow. Furthermore, using the above feature can provide a good indicator of cavitation.

In this experiment, the purpose is to examine the cavitation process in a pump under various conditions. The 3D figure below explains the connection between various parameters. This figure compares acoustic signals in the frequency analysis. It is demonstrated how the acoustic signals vary in the pump. Figure 9a describes these signals in the low-frequency range of 0 Hz–1 kHz. Based on the results, the acoustic signal did not change when the pump operated below 350 l/min. However, the significant alteration in the

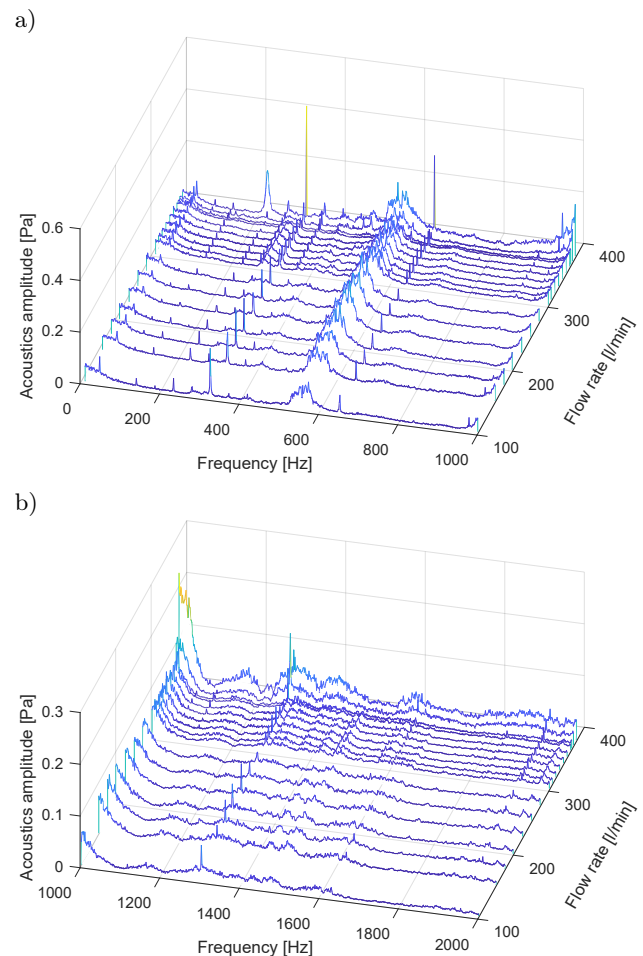


Fig. 9. Amplitude acoustic analysis at the range of frequency: a) 0 Hz–1 kHz; b) 1–2 kHz.

acoustic level was observed for the pump operating with 350 l/min and greater. Due to the interaction flow between the volute and impeller as well as cavitation, in general, the noise level increases. The dominant frequencies in the pump are the blade passing frequency of 229.5 Hz, the rotation impeller frequency of 45.9 Hz, and harmonics.

6. Acoustic signals in the frequency analysis domain

Figure 9b shows the signals in the frequency range from 1 kHz to 2 kHz. Based on this frequency range, the results revealed a slight alteration in amplitude signals as the pump operated at less than 350 l/min; nevertheless, a quick increase in amplitude signals is indicated under the flow greater than 350 l/min. Due to the occurrence of cavitation, the NPSHA was smaller than the NPSHR. Also, it was observed that the blade passing frequency (BPF) and rotational frequency (RF) dominated in this frequency range.

Figure 10 shows two ranges of frequencies: a) 2 to 10 kHz and b) 10 to 15 kHz. The flow rate of the pump

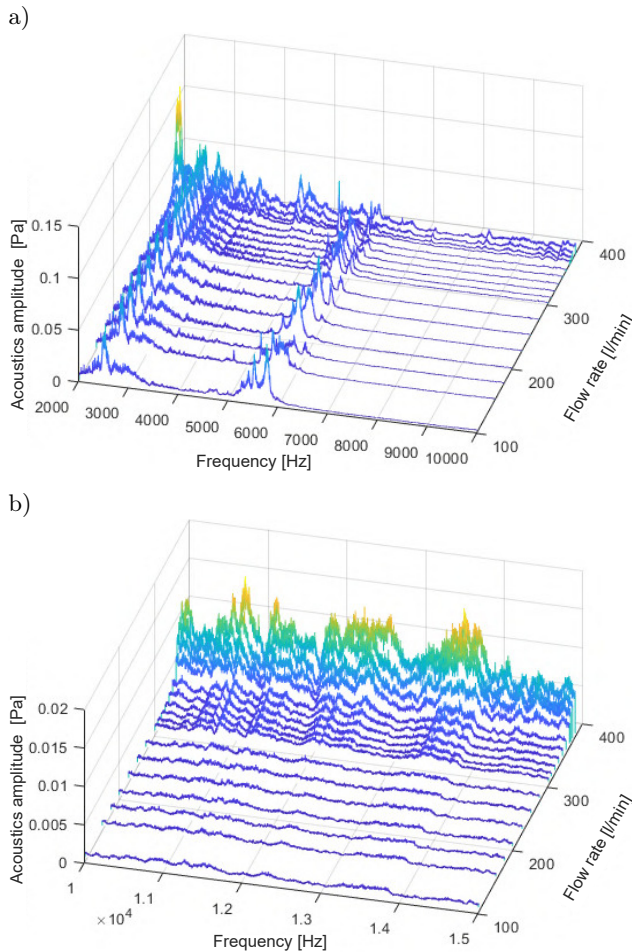


Fig. 10. Amplitude acoustic analysis at the frequency ranges: a) 2–10 kHz; b) 10–15 kHz.

is less than 350 l/min. No alteration in the acoustic signal levels is noted. A significant rise is detected when the flow exceeds 350 l/min. Cavitation growth occurs at high mass flow rates. As a result of cavitation, the noise level increases under high flow. The frequency range from 10 to 15 kHz was also effective in diagnosing cavitation.

6.1. Acoustic signals in the frequency analysis using the mean feature

Figure 11 illustrates acoustic amplitude signals mean value at the frequency ranges of 0 Hz–1 kHz, 1–2 kHz, 2–10 kHz, and 10–15 kHz. It is revealed that using the mean feature, there is no important alteration as the pump works under smaller than the design flow. Nevertheless, it is noted that the quick rise in acoustic level takes place under the flow greater than 350 l/min because of the cavitation occurrence. Also, the NPSHR is greater than the NPSHA in this case. The result indicates that using the mean value in the frequency domain offers a good approach to analyze cavitation.

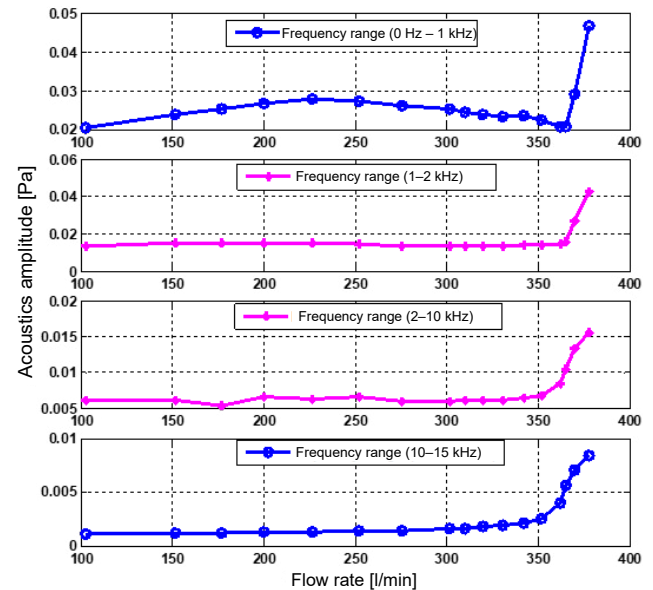


Fig. 11. Mean feature analysis in the frequency range between 0 Hz and 15 kHz.

6.2. Using the RMS feature in the frequency analysis domain

Figure 12 describes the acoustic signal analysis using the RMS feature for frequency ranges from 0 Hz to 15 kHz. All curves in this figure have the same tendency as the acoustic mean feature in Fig. 11. The amplitudes of acoustic signals rapidly increase at a flow higher than 350 l/min for all ranges of frequencies. Also, the results show that the differences in ampli-

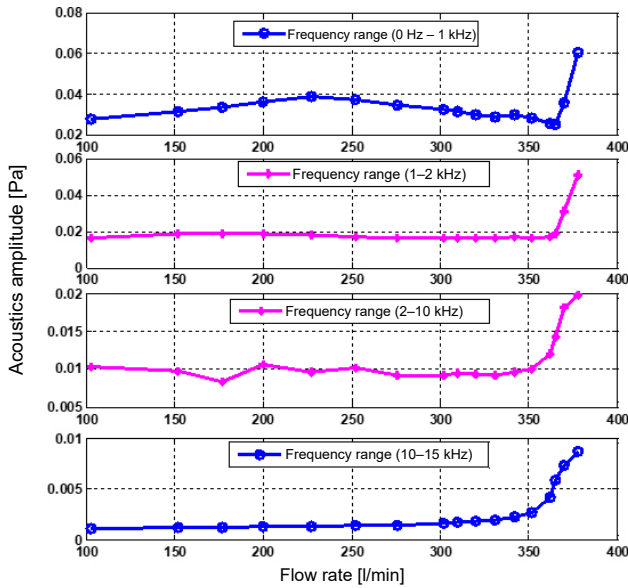


Fig. 12. RMS feature at the frequency range between 0 Hz and 15 kHz.

tudes of acoustic in frequency analysis at a variety of mass flows and frequency ranges agreed with cavitation characteristics in Fig. 6.

7. Influence of impeller speed on the cavitation occurrence

Figure 13 illustrates maximum statistical feature values. It is obvious that these values under an impeller speed of 2755 rpm are significantly higher than in other three remaining impeller speeds of 2610, 2320, and 2030 rpm. The peak value of the impeller speed of 2755 rpm is higher by about 13.4%, 15.5%, and 35.59% compared to 2610, 2320, and 2030 rpm, respectively. Similarly, the RMS value is higher by around 11.8%, 29.95%, and 52.6%, the peak-to-peak value by about 13.1%, 13.5%, and 33.3%, and the variance value by around 16.7%, 42.2%, and 75.89%, as summarized in Tables 1 and 2.

Table 1. Summary of the maximum statistical feature results for the acoustic amplitude at different impeller speeds.

Impeller speed [rpm]	Peak [Pa]	RMS [Pa]	Peak-to-peak [Pa]	Variance [Pa]
2755	0.045	0.043	0.016	0.0085
2610	0.039	0.036	0.016	0.0084
2320	0.035	0.027	0.011	0.0063
2030	0.023	0.016	0.0065	0.0034

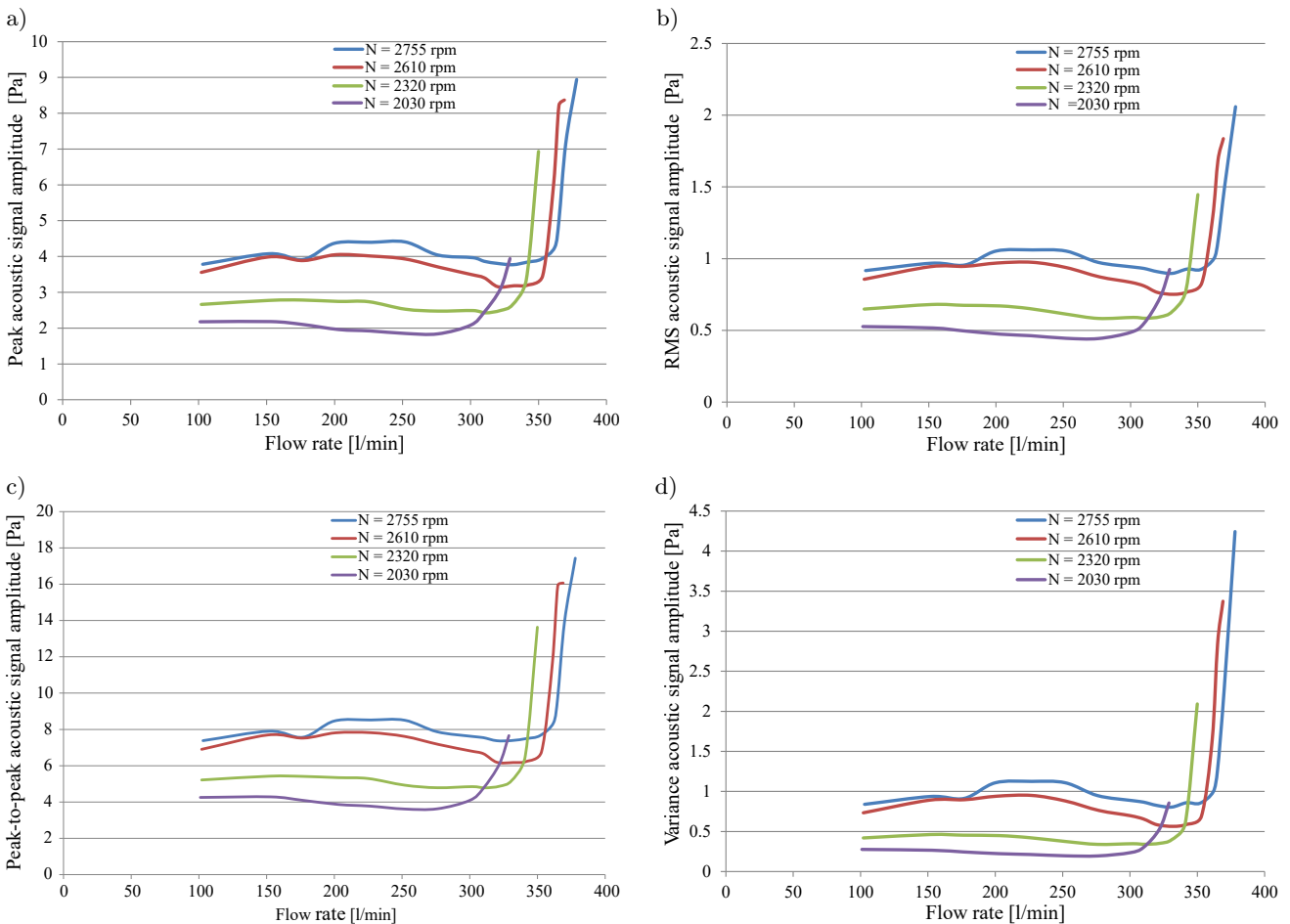


Fig. 13. Comparison of the results of various features at various impeller speeds.

Table 2. Summary of the minimum statistical feature results for the acoustic amplitude at different impeller speeds.

Impeller speed [rpm]	Peak [Pa]	RMS [Pa]	Peak-to-peak [Pa]	Variance [Pa]
2755	3.79	0.92	7.39	0.85
2610	3.56	0.86	6.92	0.74
2320	2.67	0.65	5.22	0.43
2030	2.18	0.53	4.26	0.28

Figure 14 describes the signal investigation in the frequency analysis using the mean feature at various mass flows and impeller speeds for the frequency range from 0 Hz to 15 kHz.

The results indicate that the mean feature for the pump speed of 2755 rpm is higher than for $N = 2610$, $N = 2320$, and $N = 2030$ rpm. Similarly, it is noted that the impeller speed is related to the increase of acoustic signal amplitude. The mean values increase as the pump speed increases, as listed in Table 3. A comparison of results for the minimum mean acoustic value is presented in Table 4.

Table 3. Comparison results for the maximum mean acoustic values.

Impeller speed [rpm]	Mean value 0 Hz – 1 kHz [Pa]	Mean value 1–2 kHz [Pa]	Mean value 2–10 kHz [Pa]	Mean value 10–15 kHz [Pa]
2755	0.047	0.043	0.016	0.0085
2610	0.039	0.036	0.0158	0.0084
2320	0.035	0.027	0.011	0.0063
2030	0.023	0.016	0.0065	0.0034

Table 4. Comparison results for the minimum mean amplitude features.

Impeller speed [rpm]	Mean value 0 Hz – 1 kHz [Pa]	Mean value 1–2 kHz [Pa]	Mean value 2–10 kHz [Pa]	Mean value 10–15 kHz [Pa]
2755	0.021	0.014	0.0062	0.0012
2610	0.0189	0.012	0.0058	0.0011
2320	0.016	0.0088	0.0043	0.00105
2030	0.013	0.0063	0.0041	0.00103

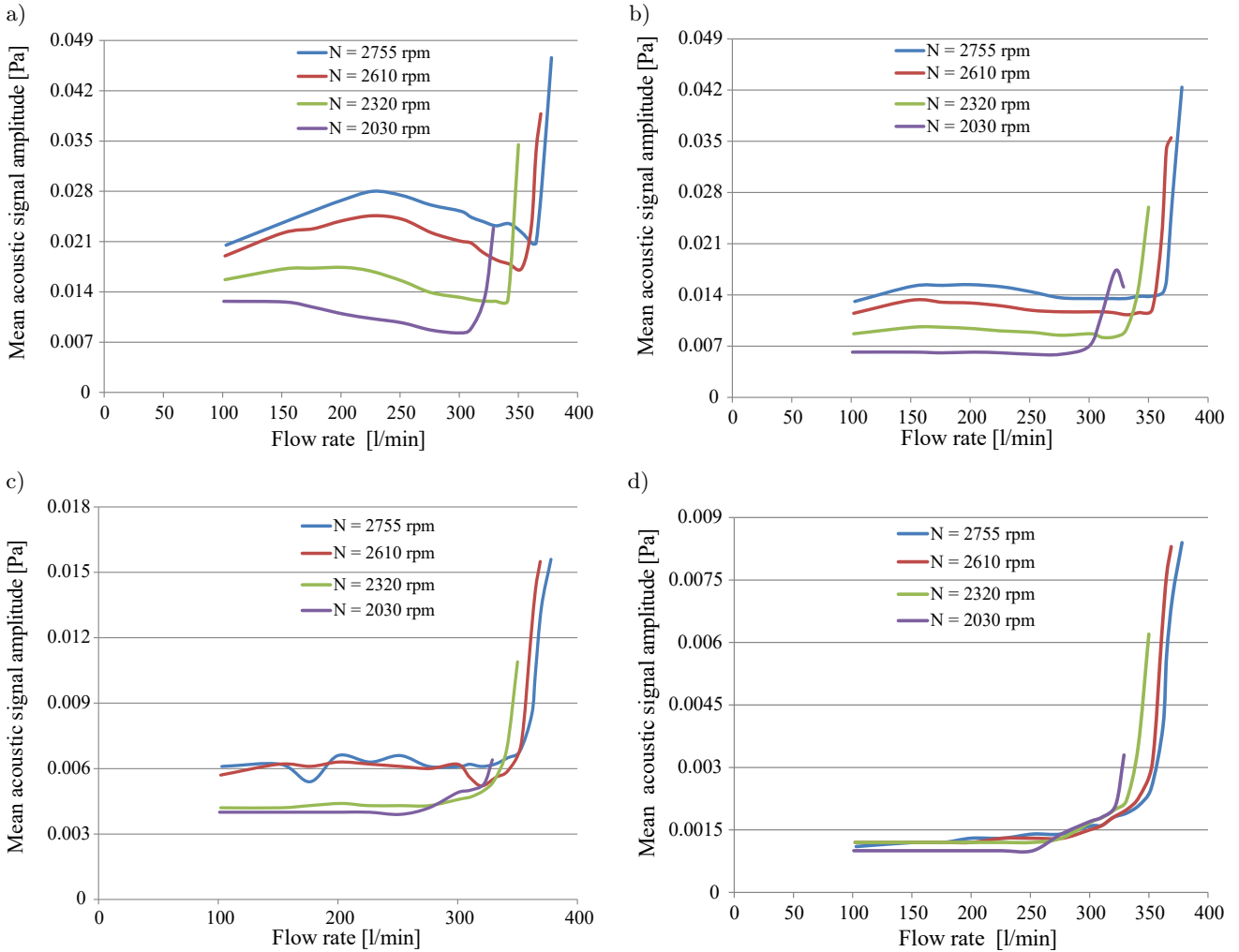


Fig. 14. Results comparison for mean feature at the frequency range between 0 Hz and 15 kHz.

Figure 15 illustrates the RMS analysis at the range frequency from 0 Hz to 15 kHz. The results reveal that the RMS value for the pump speed of 2755 rpm is significantly greater than for the cases with $N = 2610$, $N = 2320$, and $N = 2030$ rpm. It is also observed that the maximum and minimum values of the RMS features increase as the impeller speed increases, as listed in Tables 5 and 6.

Table 5. Comparison of maximum value results of the RMS feature.

Impeller speed [rpm]	Maximum value 0 Hz – 1 kHz [Pa]	Maximum value 1–2 kHz [Pa]	Maximum value 2–10 kHz [Pa]	Maximum value 10–15 kHz [Pa]
2755	0.061	0.052	0.019	0.009
2610	0.048	0.043	0.018	0.008
2320	0.047	0.031	0.014	0.007
2030	0.033	0.023	0.009	0.004

Table 6. Comparison of maximum value results of the RMS feature.

Impeller speed [rpm]	Minimum value 0 Hz – 1 kHz [Pa]	Minimum value 1–2 kHz [Pa]	Minimum value 2–10 kHz [Pa]	Minimum value 10–15 kHz [Pa]
2755	0.026	0.017	0.009	0.0014
2610	0.022	0.014	0.0085	0.0013
2320	0.017	0.008	0.0065	0.00125
2030	0.011	0.006	0.0055	0.00101

8. Conclusions

The following outcomes are derived in this investigation studying the influence of different mass flows and impeller speeds using acoustic signals on detecting cavitation occurrence. Time-domain analysis of the acoustic signal's amplitude in time by various features can provide a good indication of cavitation. The results

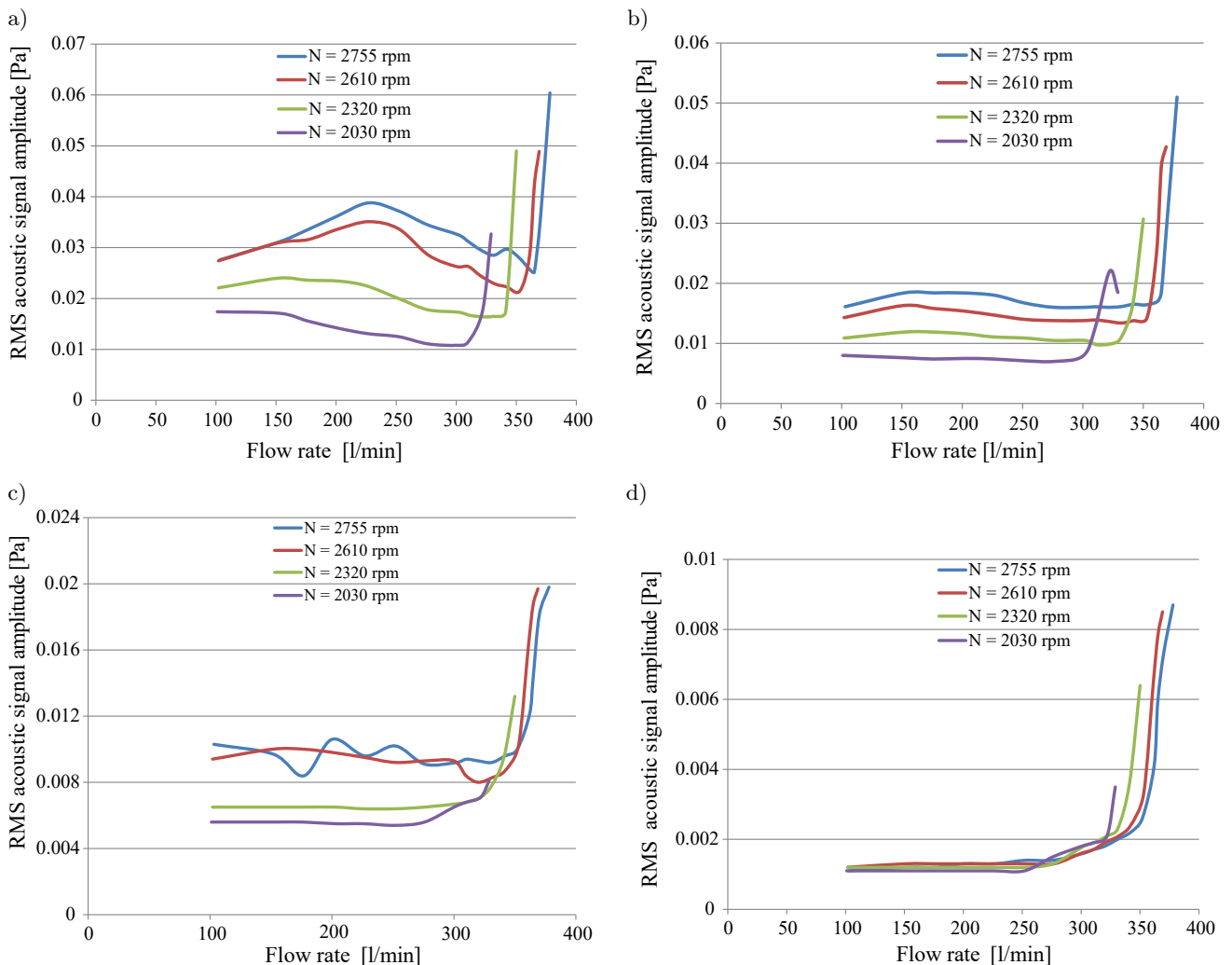


Fig. 15. RMS features comparison at the frequency range from 0 Hz to 15 kHz.

found that the value of the acoustic signal amplitude increased or decreased directly with the flow range in the pump. Acoustic signals in the time domain showed that the level of acoustic signal increased as the pump worked at the occurrence of cavitation. Also, studying the acoustic amplitude in frequency analysis is a satisfactory approach to study changes in cavitation. The dominant frequencies in the pump are the blade passing frequency, the rotation frequency, and its harmonics. Various frequency ranges offer good indications of determining cavitation by examining the amplitude of the acoustic signal in the frequency analysis. The results showed that mean and RMS features in the frequency analysis effectively detect cavities. The study of the acoustic signal amplitude under a variety of experimental tests displayed a good agreement with cavitation characteristics demonstrated in the NPSH plot. Furthermore, the result comparison in both time and frequency analyzes revealed that both mean and RMS features under an impeller speed of 2755 rpm are significantly greater than in other experimental cases. According to this study's experimental results, the acoustic method effectively detects the various levels of cavitation.

Acknowledgments

The author would like to thank Mustansiriyah University, Baghdad, Iraq, for its support.

References

1. ALBRAIK A., ALTHOBIANI F., GU F., BALL A. (2012), Diagnosis of centrifugal pump faults using vibration methods, [in:] *Journal of Physics: Conference Series, 25th International Congress on Condition Monitoring and Diagnostic Engineering (COMADEM 2012)*, **364**: 012139, doi: 10.1088/1742-6596/364/1/012139.
2. ALFAYEZ L., MBA D., DYSON G. (2005), The application of acoustic emission for detecting incipient cavitation and the best efficiency point of a 60kW centrifugal pump: Case study, *NDT & E International*, **38**(5): 354–358, doi: 10.1016/j.ndteint.2004.10.002.
3. AL-OBAIDI A. (2018), *Experimental and Numerical Investigations on the Cavitation Phenomenon in a Centrifugal Pump*, Ph.D. Thesis, University of Huddersfield, UK.
4. AL-OBAIDI A.R. (2019a), Experimental investigation of the effect of suction valve opening on the performance and detection of cavitation in the centrifugal pump based on acoustic analysis technique, *Archives of Acoustics*, **44**(1): 59–69, doi: 10.24425/aoa.2019.126352.
5. AL-OBAIDI A.R. (2019b), Monitoring the performance of centrifugal pump under single-phase and cavitation condition: A CFD analysis of the number of impeller blades, *Journal of Applied Fluid Mechanics*, **12**(2): 445–459, doi: 10.29252/JAFM.12.02.29303.
6. ANDERSON T. et al. (2014), *Introduction to Small Water Systems: A Course for Level 1 Operators*, Alaska, USA.
7. ARNOLD K., STEWART M. (1999), *Surface Production Operations, Design of Gas-Handling Systems and Facilities*, Vol. 2, Gulf Professional Publishing, Netherlands, doi: 10.1016/B978-0-88415-822-6.X5000-4.
8. BEEBE R.S. (2004), *Predictive Maintenance of Pumps Using Condition Monitoring*, Elsevier, Netherlands.
9. ČERNETIĆ J., PREZELJ J., ČUDINA M. (2008), Use of noise and vibration signal for detection and monitoring of cavitation in kinetic pumps, *The Journal of the Acoustical Society of America*, **123**(5): 3316, doi: 10.1121/1.2933777.
10. CHUDINA M. (2003), Noise as an indicator of cavitation in a centrifugal pump, *Acoustical Physics*, **49**(4): 463–474, doi: 10.1134/1.1591303.
11. ČUDINA M. (2003), Detection of cavitation phenomenon in a centrifugal pump using audible sound, *Mechanical Systems and Signal Processing*, **17**(6): 1335–1347, doi: 10.1006/mssp.2002.1514.
12. ČUDINA M., PREZELJ J. (2009), Detection of cavitation in operation of kinetic pumps. Use of discrete frequency tone in audible spectra, *Applied Acoustics*, **70**(4): 540–546, doi: 10.1016/j.apacoust.2008.07.005.
13. FAROKHZAD S., BAKHTYARI N., AHMADI H. (2013), Vibration signals analysis and condition monitoring of centrifugal pump, *Technical Journal of Engineering and Applied Sciences*, **4**: 1081–1085.
14. GAUTAM N. (2012), Construction working and advantages of centrifugal pump, *India Study Channel*, <https://www.indiastudychannel.com/resources/149329-Construction-Working-And-Advantages-Of-Centrifugal-Pump.aspx>.
15. GIRDHAR P., MONIZ O. (2005), *Practical Centrifugal Pumps, Design, Operation and Maintenance*, Elsevier, Netherlands, doi: 10.1016/B978-0-7506-6273-4.X5000-4.
16. GRIST E. (1998), *Cavitation and the Centrifugal Pump: A Guide for Pump Users*, Taylor & Francis, USA.
17. JONES G.M., SANKS R.L., BOSSERMAN B.E., TCHOBANOGLOUS G. [Eds.] (2006), *Pumping Station Design*, Gulf Professional Publishing, USA.
18. KAMIEL B.P (2015), *Vibration-based Multi-fault Diagnosis for Centrifugal Pumps*, Ph.D. Thesis, Curtin University.
19. KARASSIK I.J., MCGUIRE T. (1998), *Centrifugal Pumps*, 2nd ed., Springer, USA.
20. LIU G. (2014), *Effects of Geometrical Parameters on Performance of Miniature Centrifugal Pump*, Ph.D. Thesis, Nanyang Technological University, Singapore.

21. LOBANOFF V.S., ROSS R.R. (2013), *Centrifugal Pumps: Design and Application*, Elsevier, Netherlands.
22. NELIK L. (1999), *Centrifugal & Rotary Pumps: Fundamentals with Applications*, Taylor & Francis, USA.
23. RAMROOP G., LIU K., GU F., PAYNE B.S., BALL A.D. (2001), Airborne Acoustic Condition Monitoring of a Gearbox System, [in:] *2001 5th Annual Maintenance and Reliability Conference*.
24. SAKTHIVEL N., SUGUMARAN V., BABUDEVASENAPATI S. (2010), Vibration based fault diagnosis of monoblock centrifugal pump using decision tree, *Expert Systems with Applications*, **37**(6): 4040–4049, doi: 10.1016/j.eswa.2009.10.002.
25. SPRAKER W.A. (1965), The effects of fluid properties on cavitation in centrifugal pumps, *Journal of Engineering for Gas Turbines and Power*, **87**(3): 309–318, doi: 10.1115/1.3678264.

Research Paper

Investigation of the Effect of Non-Flat Surfaces on the Performance of Perforated Acoustic Absorber

Zahra HASHEMI⁽¹⁾ , Ali FAHIM^{(2)*} , Mohammad Reza MONAZZAM⁽³⁾ ⁽¹⁾ *Behbahan Faculty of Medical Sciences*
Behbahan, Iran⁽²⁾ *School of Engineering Science, College of Engineering, University of Tehran*
Tehran, Iran⁽³⁾ *School of Public Health, Tehran University of Medical Sciences*
Tehran, Iran

*Corresponding Author e-mail: a.fahim@ut.ac.ir

(received April 10, 2022; accepted December 12, 2022)

In order to investigate the effect of the surface shape on the performance of perforated panels, three non-flat shapes were considered for perforated panel with their absorption performance compared with the usual shape of the (flat) perforated panel. In order to simulate the absorption coefficient of a non-flat perforated panel, the finite element method was implemented by the COMSOL 5.3a software in the frequency domain. Numerical simulation results revealed that all the shapes defined in this paper improve the absorption coefficient at the mid and high frequencies. A and B shapes had a higher performance at frequencies above 800 Hz compared to the flat shape. Also, shape C had a relative superiority at all frequencies (1–2000 Hz) compared to the reference shape; this superiority is completely clear at frequencies above 800 Hz. The maximum absorption coefficient occurred within the 400–750 Hz range. After determining the best shape in terms of absorption coefficient (shape C), a perforated panel of 10 m² using fiberglass fibers and desired structural properties was built, and then it was also subjected to a statistical absorption coefficient test in the reverberation chamber according to the standard. The results of the statistical absorption coefficient measurement showed that the highest absorption coefficient was 0.77 at the frequency of 160 Hz. Also, to compare the experimental and numerical results, these conditions were implemented in a numerical environment and the statistical absorption coefficient was calculated according to the existing relationships. A comparison of the numerical and laboratory results revealed acceptable agreement for these two methods in most frequency spectra, where the numerical method was able to predict this quantity with good accuracy.

Keywords: perforated acoustic absorber; surface shape; statistical absorption coefficient; reverberation chamber; finite element method.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

When a single-frequency plane wave collides with a rigid wall, a static wave with a pressure amplitude twice the pressure amplitude of the incident wave is generated due to boundary conditions on the wall surface. In other words, all the acoustic energy of the incident wave is reflected by the wall and remains in the medium. The use of absorbent materials can reduce the wave energy reflected from the wall. Perfo-

rated panels are one of the most common resonant absorbers used for sound control. This type of absorbers is widely used due to their adjustable mechanical properties and ease of processing. The unique physical properties of the perforated panels have led to their application in complex mechanical systems such as magnetic resonance scanners (MRI) (LI, MECHEFSKE, 2010), cooling systems (ALLAM, ÅBOM, 2014), turbofan motor (JING *et al.*, 2008), as well as many buildings (YU *et al.*, 2016; 2017) and mufflers (FUCHS, ZHA,

1997; 2006). The perforated panel is a plate consisting of a number of orifices with a specified diameter and spacing, which are positioned on a rigid wall at a certain distance. The orifices on the plate are interconnected and parallel similar to many Helmholtz resonators. When the frequency of the incident wave is close to the natural frequency, the air column formed in the orifices is strongly vibrated and collides with the back wall. Acoustic energy is converted to thermal energy due to inertia and adhesion effects, whereby the noise decreases.

A good absorber has an air-like resistance (for example, a resistance close to 1 and a resistance-to-reactance ratio greater than 1) and a reactance far lower than that of air. This reactance is achieved by giving space to the back of the perforated panel. In such a structure, the characteristics of the orifices and the air gap of the back of the plate affect the resonance frequency. With the proper selection of structural parameters, appropriate absorption can be achieved within the specific frequency range without the need for porous materials. The absence of common porous materials allows for a “clean” system that is more suitable for hospitals, food, and pharmaceutical industries, and micro-electronics (LI, MECHEFSKE, 2010). Although perforated panels have been considered an alternative to porous and fibrous absorbers, they have lower performance than porous materials in terms of both the absorption rate and the absorption range. Recent studies to enhance the efficiency of perforated absorbers have indicated the importance of this class of absorbers and their role in reducing noise in specific frequency ranges.

WANG *et al.* (2014) achieved maximum absorption at different frequencies by utilizing the space behind the perforated panel and creating chambers with different depths. These frequencies were proportional to the depth of each sub-chamber. LEE and KWON (2004) increased the absorption coefficient significantly using multiple perforated panels. The use of absorbers on the back of the perforated panel is another way through which HASHEMI *et al.* (2019) improved the absorption range and absorption rate by placing foams of varying current resistance and changing their layout. Some studies have also used the resonance of the micro-perforated structure itself, especially at low frequencies (LEE *et al.*, 2005; CHANG *et al.*, 2010).

Few studies have looked at the apparent shape of adsorbent materials. For example, CHEN *et al.* (2000) examined the effect of the shape of porous absorbers behind the micro-perforated plate. The authors examined simple, semicircular, concave, and triangular shapes and concluded that the form of porous absorbers definitely affects the absorption coefficient at some frequencies. EASWARAN and MUNJAL (1993) studied the sound reflection coefficient from the foam edges using the Galerkin finite element method. BOLTON and GREEN (1986) investigated the rates of

adsorption coefficient and sound transmission loss in a panel composite structure by the finite element method. The results of both studies revealed that the porous materials with an edged shape improve the rates of absorption coefficient and the transmission loss in some frequency bands. WANG *et al.* (2019) conducted a study investigating the absorption properties of a corrugated perforated plate. For this purpose, a three-dimensional finite element model was used to estimate the absorption properties of the corrugated perforated plate. The results showed that the replacement of sinusoidal micro-perforated panel (MPP) changes the state of connection in the air-mass system; as a result, its absorption performance is different from that of flat MPP.

WANG and LIU (2020) showed in their study on corrugated micro-perforated panel absorber (MPPA) that when the wavelength of the sound wave is short, the absorption performance of corrugated MPP is better than that of the flat type, and at long wavelengths, there is no difference in the absorption rate of the two absorbers. Another finding of this study is better absorption performance at dip points (non-resonant frequencies) compared to the flat perforated absorption type, which can be used to control the resonance or reduce the accidental broadband noise in large spaces and buildings.

According to studies, the number of researches examining the surface shape of absorbers, especially perforated absorbers, is sparse. Thus, the main aim of this study is to investigate the effect of the surface shape of perforated absorbers on their absorption performance. Note that only three shapes are considered in this study and the results cannot be generalized for other shapes. The investigations in this study are conducted with two numerical finite element and experimental methods, which are described in more detail later. Regarding the article structure, theoretical and numerical adjustment sections are provided. The results of the numerical and experimental methods and the validation of the numerical method are presented in the following sections, and finally, the discussion and conclusions are presented.

2. Theoretical framework

The classical approach for such a system involves calculating the impedance of an orifice and then calculating the total impedance of the perforated panel according to the perforation percentage. The total impedance of the system depends on the perforation percentage, the diameter of the orifices, and the thickness of the panel. Impedance is a complex quantity with two real and imaginary terms:

$$Z_M = X + j\omega M. \quad (1)$$

The real part of the acoustic impedance X represents the energy propagation and viscous losses of the sound wave propagated through the orifices, and is known as resistance. The imaginary part of $j\omega M$ is called reactance and refers to the mass of air moving through the orifices. Obviously, the impedance depends on both the real and the imaginary terms, as well as other factors used depending on the model. Also, the shape of the orifices is assumed in a way that the heat-dissipated energy is insignificant compared to the viscous dissipated energy. There is no interaction between the orifices. If the orifices are too close to each other, such an assumption is not suitable for calculating the impedance and can be modified using the Fok function. MAA (1998) provided a well-known equation for determining the acoustic impedance of perforated panels:

$$Z_{M(MPP)} = \frac{\sqrt{2}\eta k}{\phi d} + \frac{j\omega\rho_0}{\phi} \cdot \left\{ \frac{0.85d}{\psi(\xi)} + t \left[1 - \frac{2}{k\sqrt{-j}} \frac{j_1(s\sqrt{-j})}{j_0(k\sqrt{-j})} \right]^{-1} \right\}, \quad (2)$$

where ρ_0 represents the density of air, ω is the angular frequency, and $j = \sqrt{-1}$ denotes imaginary numbers.

Here $k = d\sqrt{\frac{\omega\rho_0}{4\eta}}$, d denotes orifice diameter, t is the panel thickness, j_0 and j_1 indicate the Bessel functions of first class and order 0 and 1, η represents the viscosity coefficient, ϕ is the porosity, and s is the perforation ratio in percentage, ψ is the Fok function obtained by the following equation:

$$\psi(\xi) = (1 - 1.40925\xi + 0.33818\xi^3 + 0.06793\xi^5 - 0.02287\xi^6 + 0.03015\xi^7 - 0.01641\xi^8 + \dots)^{-1}, \quad (3)$$

where $\xi = 0.88d/b$, and b is the distance between the orifices. As mentioned, the perforated panels should have rigid backing wall with a certain distance to become a resonance absorber. The impedance of the perforated surface (Z_s) and the backspace is obtained from the following relation:

$$Z_s = Z_{M(MPP)} - j\rho_0 c_0 \cot\left(\frac{\omega D}{c_0}\right), \quad (4)$$

where D represents the distance from the back of the perforated panel to the solid wall, $Z_{M(MPP)}$ denotes the perforated panel impedance, and c_0 is the speed of sound in air. The absorption coefficient of the perforated absorber panel for the normal state is the ratio of the scattered wave pressure to the absorbed wave pressure obtained from the following relationship:

$$\alpha = 1 - \left| \frac{Z_s - \rho_0 c_0}{Z_s + \rho_0 c_0} \right|^2. \quad (5)$$

The oblique absorption coefficient can be obtained from the following equation:

$$\alpha_{\theta,\beta} = \frac{\iint_{S_{\text{Inlet}}} \text{Re} \left[(j\varphi) \left[\frac{\partial \varphi}{\partial z} \right]^* \right] dx dy}{K_0 (\cos \theta) L_x L_y}, \quad (6)$$

where L_x and L_y indicate the length and width of the plate, respectively, φ is the velocity potential, and $k_0 = \frac{\omega}{c_0}$ is the wave number in free space.

Finally, the absorption coefficient in the diffusion field can be written as follows:

$$\alpha_S = \frac{1}{2\pi} \int_0^{2\pi} \left(\int_0^{\pi/2} \alpha_{\theta,\beta} \sin 2\theta d\theta \right) d\beta, \quad (7)$$

where θ and β are the angles of elevation and azimuth, respectively.

3. Methods

3.1. Numeric adjustment

In order to simulate the absorption coefficient of a non-flat perforated panel, the finite element method was used by COMSOL 5.3a software in the frequency domain. The computational range included the space behind the perforated panel, the perforated panel itself, and the virtual channel. This software solves the wave equation by solving the Helmholtz equation in the frequency domain. Given the non-flat surface of the perforated panel, it is not possible to claim that the normal absorption coefficient is obtained if the angle of a sound wave is 0. Thus, in this study, the angle of the sound wave between two modes, 0 (minimum) and 90 (maximum), was considered 45 degrees. The modeling was done in a 3D environment based on drawing two cubes separated by a parabolic plane. The shape of the plate was determined using the specified equations. The upper cube corresponded to the virtual channel, while the lower cube corresponded to the volume behind the perforated panel. The acoustic field inside the virtual channel was connected to the space behind the plate through the perforated panel orifices. The length and width of the virtual channel and the back channel of the plate were 100 cm and 100 cm. Depending on the shape of the plates, the depth of the micro-perforated panel back was defined with the effective height and obtained by integrating the volume behind the perforated panel. The effective depth behind the perforated panel was assumed to be 10 cm. Rigid boundary conditions were considered for back-channel walls and periodic conditions at virtual channel boundaries. By applying the perforated boundary condition to the perforated panel, the acoustic impedance of the perforated panel relative to air was obtained according

to the Maa formula. The Dirichlet-to-Neumann boundary conditions (KELLER, GIVOLI, 1989), applied to the virtual channel inlet, would allow the sound wave to pass through this boundary without reflection.

To investigate the effect of the surface shape of the perforated panel, three designs: A, B, and C, were selected for the perforated panel surface (Fig. 1). Factors to consider in this study were the simplicity of the designs in terms of drawing in software and the construction plus use phase as well as aesthetics and decoration issues. Simulations were performed within the frequency range of 1 to 2000 Hz. Structural properties such as orifice diameter, perforation percentage, and plate thickness in all shapes were constant at 1 mm, 0.016%, and 1 mm, respectively (Table 1). The tetrahedral mesh type and mesh sizes were chosen uniformly. In addition to the shapes mentioned, a flat shape was also used as a reference for comparison with the selected designs. The simulations of the absorption coefficient were also compared with the flat shape results. Finally, according to the simulation results, the best

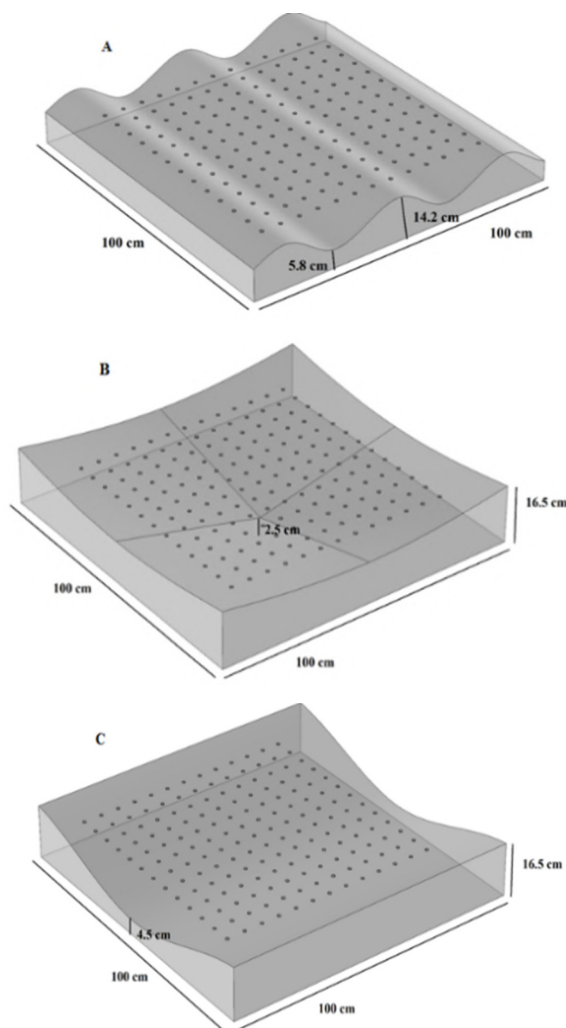


Fig. 1. Defined shapes for the perforated panel and their dimensions.

Table 1. Structural properties of the micro-perforated panel (numerical simulation).

Orifice diameter [mm]	Panel thickness [mm]	Effective height [mm]	Perforation ratio [%]
1	1	100	0.016

design in terms of absorption coefficient was chosen to construct at the required dimensions to determine the absorption coefficient in the reverberation chamber.

3.2. Measurement in the reverberation chamber

According to the results of the numerical solution, the best shape was selected in terms of the absorption coefficient. After making the desired piece (shape C – Fig. 1), the statistical absorption coefficient in the reverberation chamber was measured. The volume of the reverberation chamber was 100 m³ and the roof area was 31.18 m². The diffusers installed in the chamber also helped disperse the sound wave and convert it into a completely diffuse place. Measurements were made in accordance with ISO 354: 2003 standard (ISO 354, 2003). According to this standard, the media dimensions to measure the statistical absorption coefficient must be at least 9 to 12 m². In this study, a perforated panel of 10 m² was built by putting 40 pieces of 50 × 50 cm together. The area of the surfaces and base of the panels was 11.92 m². Considering the economic constraints and problems, the molding method was used to fabricate the desired piece. To make the pieces lighter and easier to shape, 3 mm thick fiberglass and 3 mm orifices with a 0.5% perforation ratio (Table 2) perpendicular to the non-flat surface were used (Fig. 2). Note that the main focus of this study was the form factor study; thus, the selection of other



Fig. 2. Non-flat perforated panel (shape C) of 50 × 50 cm dimension.

Table 2. Structural properties of the micro-perforated panel (experimental simulation).

Orifice diameter [mm]	Panel thickness [mm]	Effective height [mm]	Perforation ratio [%]
3	3	100	0.5

parameters such as orifice diameter, perforation percentage, and thickness was not among the priorities of this study. There were also many limitations during the construction phase and the most applicable for these parameters was considered.

Different arrangements were proposed for this shape of the panel. Nevertheless, since the periodic boundary conditions were chosen for the virtual channel boundaries in this study (Fig. 3), suggesting that the sound wave conditions in the next panels can be repeated as before, the regular arrangement was considered at this stage. According to the standard, the sample arrangement should not be parallel to the walls of the room and should be half a meter from each wall. The sound was transmitted by a 12-dimensional loudspeaker in two locations, at frequencies of 100 to 5000 Hz in 1.3 octave steps in different directions, with the microphone manually positioned at five preset locations, and the experiments were repeated three times in each case. Finally, the average statistical absorption coefficients were obtained.



Fig. 3. Arrangement of the perforated panel in the reverberation chamber. The layout of the audio microphone and speaker.

For comparing the experimental method with the numerical methods, according to Eq. (7), the absorption coefficient at angles ($\theta = 0-90^\circ$, $\beta = 0-180^\circ$) was calculated in the software separately; by averaging their values, the statistical absorption coefficient was obtained at 1.3 octave frequencies.

4. Results

4.1. Numerical results for selecting the best shape

The results of the finite element simulation were presented for the designs considered for the perforated panel compared to the flat shape (reference) and compared to each other. As mentioned earlier, the structural properties of the surface shape in all three shapes as well as the reference shape were very similar and the effective depth of all structures is 10 cm.

Figure 4 displays the simulation results of the absorption coefficient in the two flat (normal) perforated panels in A and reference shapes. The performance of both shapes at frequencies below 800 Hz was similar, but with increasing frequency, the perforated panel with a sinusoidal shape A performed better. At 240 Hz, there was a small peak in both diagrams, which could be due to the resonance of the chamber behind the perforated panel. Considering the similarity of properties such as orifice diameter, perforation percentage, and plate thickness in both samples, subsequent peaks occurred at both identical absorbers and at frequencies of 450 and 750, which is higher in absorber A with a negligible difference. Most of the absorption was observed within the frequency range of 450–800 Hz, with the performance of both absorbers declining above the frequency of 800 Hz and the absorption coefficient diminishing in the reference plate with a steeper slope.

As can be seen in Fig. 5, the perforated absorber B, which is actually an inverted pyramid scheme similar to absorber A, performs better at frequencies above 800 Hz compared to the flat shape. Meanwhile, it also functions better at frequencies below 800 Hz. For example, at the frequency of 400 Hz, it has an absorption coefficient of 0.82, while the absorption coefficient of the reference is 0.6. The next point to notice in this diagram is the higher resonance peak in the back chamber space at 240 Hz. Given the variable height of the back chamber as well as the higher height at

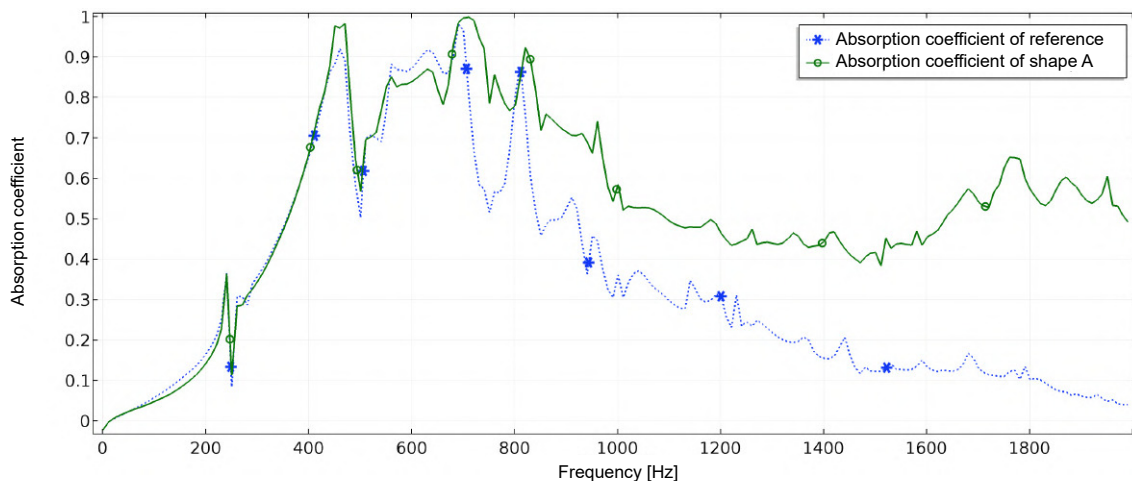


Fig. 4. Comparison of absorption coefficient in the two shapes of A and reference shape.

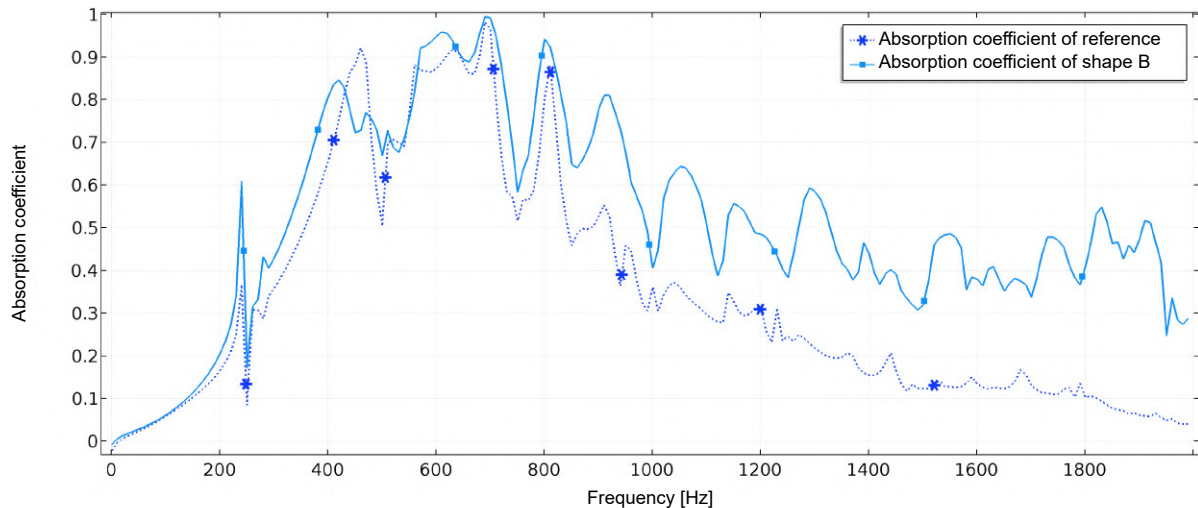


Fig. 5. Comparison of absorption coefficient in the two shapes of B and reference shape.

some points in absorber B, such a peak seems reasonable.

The results of shape C in Fig. 6 show relative superiority at all frequencies (1–2000 Hz) compared to the reference shape; this superiority is completely clear at frequencies above 800 Hz. The maximum absorption coefficient has occurred within the 400–750 Hz range.

Figure 7 well illustrates that at mid and high frequencies, the shapes defined in this study improve the performance of these absorbers, where the increased absorption rate by up to 3-fold is observed at some frequencies. Meanwhile, by designing and choosing the appropriate design and shape, the absorption pattern can be changed at lower frequencies, and a better absorption rate is achieved. The most important finding in these diagrams was the change in the absorption pattern in absorbers of unusual shape relative to the flat shape. On the other hand, further analysis

showed that the changes in the perforated panel improved the absorption coefficient and wider absorption bandwidth. In terms of the absorption pattern, the absorption diagrams of these three designs differ due to reasons discussed below.

4.2. Results of statistical absorption coefficient in the experimental method (reverberation chamber)

Figure 8 shows the statistical absorption coefficient by the experimental method in the reverberation chamber. As can be seen, the highest absorption coefficient was observed at low frequencies (less than 250 Hz) and with a peak value of 0.77 at 160 Hz. Also, at 400 Hz a small peak was observed. What is significant in this graph is the acceptable absorption coefficient at low frequencies, which can be very important due to the low porosity of the absorbers in these areas.

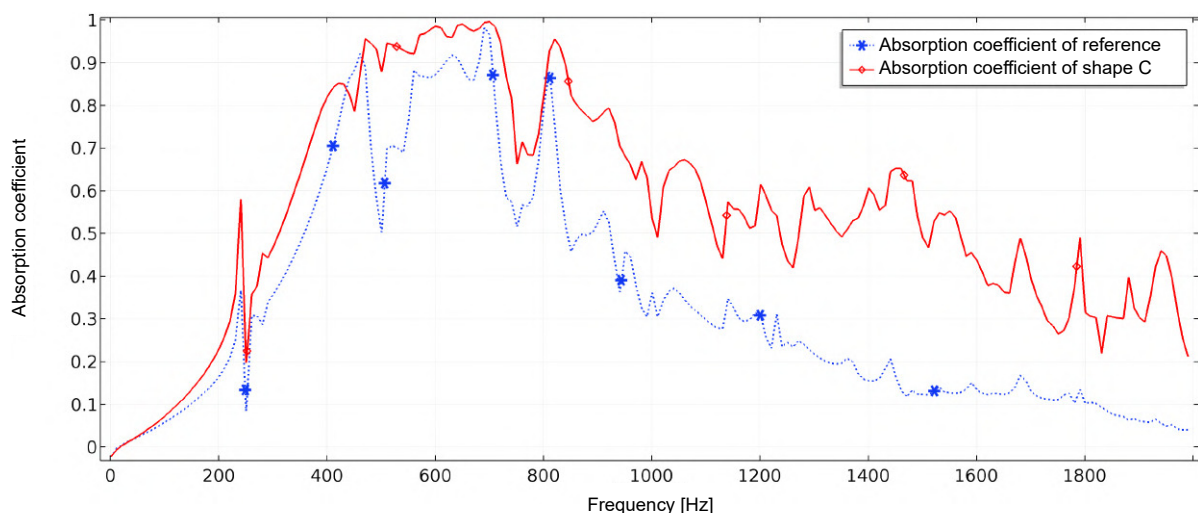


Fig. 6. Comparison of absorption coefficient in the two shapes of C and reference shape.

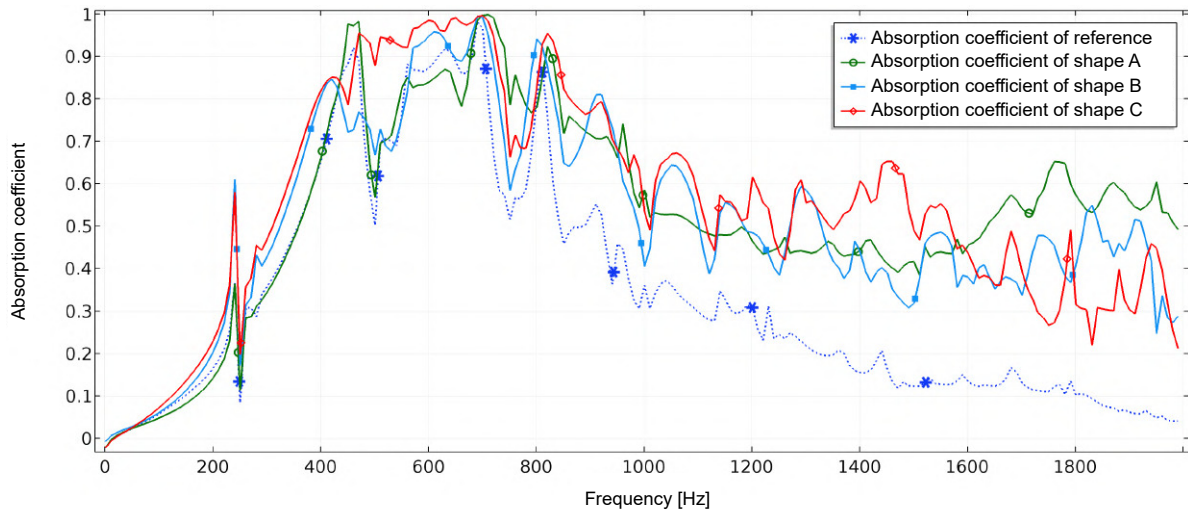


Fig. 7. Comparison of absorption coefficient in all studied shapes.

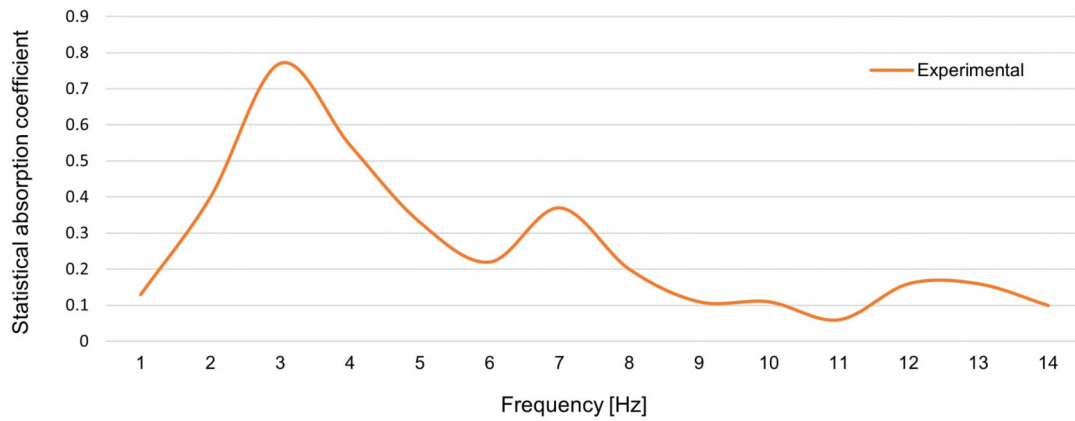


Fig. 8. Statistical absorption coefficient by the experimental method (shape C).

4.3. Comparison of the results of statistical absorption coefficient in FEM and experimental method

A comparison of the statistical absorption coefficient with experimental and numerical methods

(Fig. 9) indicated the acceptable agreement of both methods in estimating this quantity. If the mismatch of the results at the frequency of 400 Hz is ignored, the absorption pattern is the same in both methods, and the numerical method simulated the results with acceptable accuracy. At some points, the experimental

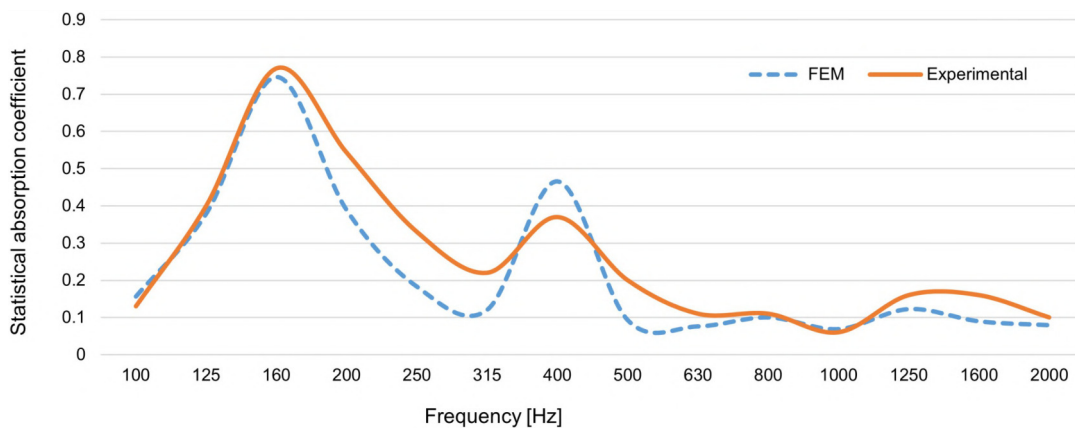


Fig. 9. Comparison of the statistical absorption coefficient with experimental and numerical methods.

absorption coefficient was higher than the value estimated by the numerical method, and the reasons are elaborated further.

5. Discussion

This study investigated the surface shape factor and its effect on the absorption performance of perforated panels. Initially, three simple shapes were defined as numerical tools in the software environment for the reasons mentioned, and with the same structural properties for these shapes, their absorption performance was compared to that of the flat shape. The results showed that the absorption performance increased at the mid and high frequencies, at least for the defined shapes. Also, at least in one of the shapes (C), there was a noticeable difference in the absorption coefficient at low frequencies, in addition to the high frequencies. In this study, the desired shape was constructed in 10 m^2 according to ISO 354 (2003) standard and its statistical absorption coefficient was measured in the reverberation chamber. Simultaneously, this quantity was simulated numerically.

Initially, these two issues should be explained:

- Why does the highest amount of absorption coefficient for all types of perforated absorbers occur at frequencies above 800 Hz?

With sound waves colliding with a hard and inflexible obstacle, the velocity of the particles will reach a maximum between a quarter and a third of the sound wavelength. If the absorber thickness is less than a quarter of the sound wavelength, its effect on absorption diminishes. This is the reason for acceptable absorption in the high-frequency spectrum using a thin absorber. As the shaping of the plates creates areas with variable heights, the maximum heights would lie within the range of $1/4$ of the wavelength or close to this range, and therefore, better absorption is observed at higher frequencies compared to lower frequencies. The absorber depth must be very high to cover the range of $1/4$ wavelength. In this regard, WANG *et al.* (2019) found that the absorption performance of a perforated wavy surface can be different from that of a flat perforated plate when the sound wavelength is less than the depth of air behind the perforated plate.

- Why did the absorption coefficient peak in Fig. 5 occur at 240 Hz for all types of perforated absorbers (three designs: A, B, and C)?

Different modes were tested to ensure no error in the numerical model, including the mesh size, mesh geometry, and different boundary conditions. In the present study, the rigid boundary condition was used for the chamber walls behind the perforated plate. This means that the chambers with an area of $100 \times 100\text{ cm}^2$ were separate from each other and might be subjected to the resonance phenomenon due to the dimensions at some

frequencies. In the next step, the periodic boundary condition was considered instead of the rigid condition, with the periodic boundary condition signifying that the entire absorber area was continuous and not separated by hard walls. As a result, the shape area would be different from the previous state, and consequently, the resonance frequency would be different. Thus, the peaks of 240 Hz were eliminated with this change. According to the above reasons, it can be concluded that the presence of peaks at a frequency of 240 Hz was not due to a computational error, as these peaks disappeared by changing the boundary conditions. Also, the peaks related to the frequency of 240 Hz could be related to the resonance of the volume of the chamber behind the perforated plate.

In resonance absorbers, the air gap actually plays the spring role, whose value is controlled by the depth measure. Sound absorption in perforated panels is influenced by the resonance of the volume of vibrating air in the orifice and backspace. The change in the geometric configuration of the backspace can effectively alter the mass-spring coupling between the chamber and the perforated panel (WANG *et al.*, 2010). Also, the results of the study conducted by LEE and LEE (2007) on a flexible micro-perforated panel show this can be done by adjusting the curvature of the panel and thereby bringing the resonance frequencies closer to each other, the overall value of absorption coefficient increased. In irregular chambers, the real part of impedance contains not only the resistance of the perforated panel itself but also the acoustic energy transmitted from the chamber.

Further, the acoustic energy transmitted from these chambers has a different shape relative to the regular shape of the back chamber, which in turn, is closely related to the distortion of the shape of the acoustic modes in non-uniform chambers. In response to these changes, the resistance is not uniformly distributed on the surface of the perforated panel, and it alters the pattern of absorption. The perforated panels with a non-uniform back chamber have also been considered with multiple absorption features. The local absorption properties in non-uniform chambers are attributed to the variable impedance conditions. In this regard, we can refer to a study by WANG *et al.* (2010) by creating a trapezoidal chamber on the back of the micro-perforated panel and comparing it to the usual (rectangular) shape of the back of the micro-perforated panel. The authors concluded that the shape of the back chamber was very effective in the performance of the perforated absorbers and it made a significant difference to the efficiency of these types of absorbers. This improves the absorption performance and widens the absorption range. A number of researchers partitioned the backspace of a micro-perforated panel in the form of a honeycomb. The results showed that the unusual shape of the backspace increases the coupling between

the acoustic modes in the back chamber thus enhancing the absorption band (SAKAGAMI *et al.* 2010; YANG, CHENG, 2016). These studies, in many ways, confirm the results of this study. Typically, in absorbers, several energy loss paths occur at the same time, such as reflection at the boundaries, deviation due to different speeds of sound propagation in separate materials, and air friction on the walls (KULHAVÝ *et al.*, 2018). When the soundwave hits a very flat surface on a plate, it will be reflected at the same angle. When the flat surface changes, i.e., it becomes shaped or angled, a part of the energy is propagated in a direction other than the reflection angle. This propagation and dispersion of the reflecting wave in different directions is called diffusion. Any non-flat surface can be considered a diffuser (SCHROEDER, 1975; 1979). The reflection mode or, in other words, the wave propagation in the back and front of the plate, depends entirely on the shape of the plate. One of the limitations of this study was a failure to deal with the contribution of each of these phenomena in changing the absorption pattern, yet definitely the complexity and interdependence of these factors have contributed to this.

In the remainder of this discussion, the phase difference caused by the discrepancy in the path length traveled by the waves entering the chamber behind the plate can also be investigated. The incident waves move a pressure wave to the bottom of the chamber, which is reflected upward after colliding with the bottom of the chamber. Considering the varying lengths of the traveled path in irregularly shaped chambers, phase difference occurs in these waves in the return path relative to each other. Whether these changes contribute to or attenuate the absorption is a complex phenomenon for which all factors should be examined. Nevertheless, it can be safely assumed that it changes the absorption pattern, which may also be associated with some of the changes in absorption curves. Also, the non-flatness of the perforated panel as the absorption surface means that the angle of the sound wave changes; obviously, the absorption coefficient depends on the angle of the radiation wave (MAA, 1998).

Changing the angle of the radiation wave can both support the absorption and reduce the absorption coefficient. Nevertheless, it is clear that when the surface is angled, the incident angle of the sound wave with the surface changes relative to its usual state. It may be argued that this would also contribute to the absorption in the shapes defined in this study.

The results of the statistical absorption coefficient in Fig. 8 show that the highest absorption rate was obtained at the low frequency of 160 Hz. Wave physics and the length of wavelength in this frequency region often render porous and fibrous absorbers ineffective in this area, and obtaining acceptable absorption coefficients in this region requires great weight constraints and thickness of these materials.

The most important advantage of resonance absorbers is the adjustable properties of these absorbers. Such a result is undoubtedly attributable to factors affecting the performance of perforated absorbers, including orifice diameter, perforation percentage, backplate depth, and panel geometry as well as configuration. The results showed that the shape of the surface, and thus, the variable height behind the perforated panel have been effective for the reasons mentioned above. Considering the limitations of this study, we were not able to test the flat specimen with the structural properties of the specimen in the reverberation chamber. On the other hand, in the numerical environment, the laboratory test conditions were simulated and compared with the results of the reverberation chamber. The results revealed that the agreement between the two methods is acceptable, especially at low frequencies. The absorption coefficient in the experimental method was higher in most frequency spectra. Several mechanisms are effective in increasing the absorption coefficient in the experimental method, such as absorption by the frame around the sample. In the numerical method, only the surface of the absorber is computed in the calculations of the absorption coefficient, but this is inevitable in the reverberation chamber and there is a little extra absorption surface compared to the numerical conditions.

In addition, the vibration of the perforated panel and the edge effect can be mentioned. The effect of the edges is indeed the phenomenon of diffraction or sound wave refraction on the sample edges. The edge diffraction causes excess amounts of acoustic energy to enter into the perforated chamber, i.e., the absorption panel area is greater than the calculated value. Nevertheless, in most parts of Fig. 8, there is a good agreement between the two methods despite the differences. For example, at the frequency of 400 Hz, the absorption rate in the numerical method is higher than in the experimental method. Of course, despite the high precision of the pieces, there may be differences in some properties that are unavoidable in experimental methods.

6. Conclusion

In this study, three shapes were defined for the perforated panel, and their absorption performance was compared to the usual shape of the perforated absorbers at frequencies of 1–2000 Hz based on the numerical finite element method. According to the results obtained in this study, it can be stated that:

- 1) Sound absorption in perforated panels is influenced by the resonance of the volume of vibrating air in the orifices and backspace. The change in the geometric configuration of the backspace can effectively change the vibro-acoustic coupling between the chamber and the perforated panel, which in turn, changes the absorption pattern.

- 2) Chambers with unusual shapes have multiple and local absorptions due to their geometry. The local absorption properties in non-uniform chambers are attributed to the variable impedance conditions.
- 3) The way the wave reflects in the back and front of the plate depends entirely on the plate shape. In this study, the defined shapes changed the reflection of the waves as part of the energy dissipation path. The plate shape also caused the phase difference of the waves in the chamber due to the difference in path travel, affecting the efficiency of the absorbers.
- 4) In this study, only three shapes were examined where the selection of shapes was not scientifically based, and only their simplicity, constructability and applicability were considered. Also, the contribution of each phenomenon to energy dissipation was not estimated, so another study is required to cover these limitations.
- 5) The results of the absorption coefficient showed that the major absorption in this type of absorber occurred at low frequencies (160 Hz), which can be very important due to the low porosity of the absorbers in these areas.
- 6) There was an acceptable agreement on the statistical absorption coefficient between the numerical and experimental methods. It can be stated that the use of numerical methods such as finite element to predict the acoustic properties of different media, both in terms of speed of work and economic issues, can be very helpful.

Acknowledgments

This study was financially supported by the Tehran University of Medical Sciences & Health Services (Grant No.: 36052). Also, the first author would like to Mr. Farhad Taleei, Mr. Moghadami, Mr. Broghany, Mr. Karkhane, and Mr. Emami for their support and cooperation during this study.

References

1. ALLAM S., ÅBOM M. (2014), Fan noise control using microperforated splitter silencers, *Journal of Vibration and Acoustics*, **136**(3): 031017, doi: 10.1115/1.4027245.
2. BOLTON J.S., GREEN E.R. (1986), Sound transmission through foam-lined double panel constructions, *The Journal of the Acoustical Society of America*, **79**(S1): S31–S31, doi: 10.1121/1.2023165.
3. CHANG D., LIU B., LI X. (2010), An electromechanical low frequency panel sound absorber, *The Journal of the Acoustical Society of America*, **128**(2): 639–645, doi: 10.1121/1.3459838.
4. CHEN W.-H., LEE F.-C., CHIANG D.-M. (2000), On the acoustic absorption of porous materials with different surface shapes and perforated plates, *Journal of Sound and Vibration*, **237**(2): 337–355, doi: 10.1006/jsvi.2000.3029.
5. EASWARAN V., MUNJAL M.L. (1993), Finite element analysis of wedges used in anechoic chambers, *Journal of Sound and Vibration*, **160**(2): 333–350, doi: 10.1006/jsvi.1993.1027.
6. FUCHS H.V., ZHA X. (1997), Acrylic-glass sound absorbers in the plenum of the Deutscher Bundestag, *Applied Acoustics*, **51**(2): 211–217, doi: 10.1016/S0003-682X(96)00064-3.
7. FUCHS H.V., ZHA X. (2006), Micro-perforated structures as sound absorbers – A review and outlook, *Acta Acustica united with Acustica*, **92**(1): 139–146.
8. HASHEMI Z., MONAZZAM M.R., FAHIM A. (2019), Estimation of sound absorption performance of complex perforated panel absorbers by numerical finite element method and examining the role of different layouts behind it, *Fluctuation and Noise Letters*, **18**(03): 1950013, doi: 10.1142/S0219477519500135.
9. ISO 354 (2003), Acoustics – Measurement of sound absorption in a reverberation room, *International Organization for Standardization*.
10. JING X., PENG S., SUN X. (2008), A straightforward method for wall impedance eduction in a flow duct, *The Journal of the Acoustical Society of America*, **124**(1): 227–234, doi: 10.1121/1.2932256.
11. KELLER J.B., GIVOLI D. (1989), Exact non-reflecting boundary conditions, *Journal of Computational Physics*, **82**(1): 172–192, doi: 10.1016/0021-9991(89)90041-7.
12. KULHAVÝ P., SAMKOVÁ A., PETRU M., PECHOCKIAKOVA M. (2018), Improvement of the acoustic attenuation of plaster composites by the addition of short-fibre reinforcement, *Advances in Materials Science and Engineering*, **2018**: 7356721, doi: 10.1155/2018/7356721.
13. LEE D.H., KWON Y.P. (2004), Estimation of the absorption performance of multiple layer perforated panel systems by transfer matrix method, *Journal of Sound and Vibration*, **278**(4–5): 847–860, doi: 10.1016/j.jsv.2003.10.017.
14. LEE Y.Y., LEE E.W.M. (2007), Widening the sound absorption bandwidths of flexible micro-perforated curved absorbers using structural and acoustic resonances, *International Journal of Mechanical Sciences*, **49**(8): 925–934, doi: 10.1016/j.ijmecsci.2007.01.008.
15. LEE Y.Y., LEE E.W.M., NG C.F. (2005), Sound absorption of a finite flexible micro-perforated panel backed by an air cavity, *Journal of Sound and Vibration*, **287**(1–2): 227–243, doi: 10.1016/j.jsv.2004.11.024.
16. LI G., MECHEFSKE C.K. (2010), A comprehensive experimental study of micro-perforated panel acoustic absorbers in MRI scanners, *Magnetic Resonance Materials in Physics, Biology and Medicine*, **23**(2): 177–185, doi: 10.1007/s10334-010-0216-9.

17. MAA D.Y. (1998), Potential of microperforated panel absorber, *The Journal of the Acoustical Society of America*, **104**(5): 2861–2866, doi: 10.1121/1.423870.
18. SAKAGAMI K., YAMASHITA I., YAIRI M., MORIMOTO M. (2010), Sound absorption characteristics of a honeycomb-backed microperforated panel absorber: Revised theory and experimental validation, *Noise Control Engineering Journal*, **58**(2): 157–162, doi: 10.3397/1.3294861.
19. SCHROEDER M.R. (1975), Diffuse sound reflection by maximum-length sequences, *The Journal of the Acoustical Society of America*, **57**(1): 149–150, doi: 10.1121/1.380425.
20. SCHROEDER M.R. (1979), Binaural dissimilarity and optimum ceilings for concert halls: More lateral sound diffusion, *The Journal of the Acoustical Society of America*, **65**(4): 958–963, doi: 10.1121/1.382601.
21. WANG C., CHENG L., PAN J., YU G. (2010), Sound absorption of a micro-perforated panel backed by an irregular-shaped cavity, *The Journal of the Acoustical Society of America*, **127**(1): 238–246, doi: 10.1121/1.3257590.
22. WANG C., HUANG L., ZHANG Y. (2014), Oblique incidence sound absorption of parallel arrangement of multiple micro-perforated panel absorbers in a periodic pattern, *Journal of Sound and Vibration*, **333**(25): 6828–6842, doi: 10.1016/j.jsv.2014.08.009.
23. WANG C., LIU X. (2020), Investigation of the acoustic properties of corrugated micro-perforated panel backed by a rigid wall, *Mechanical Systems and Signal Processing*, **140**: 106699, doi: 10.1016/j.ymssp.2020.106699.
24. WANG C., LIU X., LIXI H. (2019), On the sound absorption performance of corrugated micro-perforated panel absorbers, [in:] *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, **259**(7): 2336–2347.
25. YANG C., CHENG L. (2016), Sound absorption of microperforated panels inside compact acoustic enclosures, *Journal of Sound and Vibration*, **360**: 140–155, doi: 10.1016/j.jsv.2015.09.024.
26. YU X., CUI F.S., CHENG L. (2016), On the acoustic analysis and optimization of ducted ventilation systems using a sub-structuring approach, *The Journal of the Acoustical Society of America*, **139**: 279–289, doi: 10.1121/1.4939785.
27. YU X., LAU S.K., CHENG L., CUI F. (2017), A numerical investigation on the sound insulation of ventilation windows, *Applied Acoustics*, **117**(Part A): 113–121, doi: 10.1016/j.apacoust.2016.11.006.

Research Paper

The Acoustic Effect of Windows Installed in a Wood Frame Façade

Jacek NURZYŃSKI 

Building Research Institute

Warsaw, Poland; e-mail: j.nurzynski@itb.pl

(received September 15, 2022; accepted December 20, 2022)

The acoustic effect of windows installed in a prefabricated wood frame façade was considered. Windows inserted into a lightweight wall modify its structural scheme. The research aimed to investigate the possible interaction of the façade's main components and their actual contribution to the total sound insulation. The principal research question involved the prediction of the acoustic performance of the complete prefabricated panel from the performance of its basic elements, an opaque part and windows. As the frequency-dependent characteristics of the elements differ substantially, the use of single number values for prediction and accuracy was of particular interest. The study is based on laboratory measurements. Initially, two full-scale samples of an opaque wall and four windows were tested separately. Then, several variants of the façade consisting of various combinations of these elements were examined. The results of measurements were juxtaposed and compared with calculated values. The frequency-dependent experimental results were fairly consistent with calculations. The estimations based on single number quantities were also in good agreement with measurements. Thus, it may be concluded that the façade elements did not interact significantly, and the single number calculations give reliable results that can be used in practice.

Keywords: sound insulation; lightweight frame building; prefabricated façade; windows; protection against noise.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

The marked trend of lightweight, prefabricated residential buildings is evidently growing, so panelised and modular homes have become increasingly popular. The frame structure is beneficial in respect to construction, building physics, and energy conservation. Besides, these buildings are widely perceived as consistent with the sustainable development concept and providing good indoor environment quality (ŠUJANOVÁ *et al.*, 2019; LIEBL *et al.*, 2013). In this context, the acoustic performance of the façade is of primary importance particularly in the case of multi-storey dwellings situated in noisy areas. The review of the literature concerning acoustic properties of lightweight timber buildings, however, showed that the sound insulation of their envelope was rarely investigated and the major contribution was from analytical studies (CANIATO *et al.*, 2017). For prefabricated buildings, the façade panel, consisting of different components, becomes a final product placed on the market. The acoustic performance should then be determined for the entire panel,

especially with regard to the vision of an open building system and open manufacturing (NURZYŃSKI, 2007). Testing each possible combination of the opaque part and windows, however, seems impractical. Generally, the acoustic performance of a complex partition may be estimated from the performance of its elements. The formula for total sound insulation applies well to traditional massive buildings, but the structure of lightweight façade is very different. In acoustic terms, the components may interact making the estimations inadequate or imprecise.

The sound insulation of a window is usually determined empirically in the laboratory for a specific product of a specific type and dimensions, consisting of a strictly defined frame, the glazing, and sealing system. Extensive data based on laboratory measurements already exist for numerous products and may be used for the prediction of total sound insulation of a complete façade. The window, however, when tested in the laboratory is installed in a heavy massive partition so the mounting (edge) conditions differ considerably from the practical assem-

bly in a lightweight façade, which may influence the results (UTLEY, FLETCHER, 1969). Theoretical models, on the other hand, that may be used for calculating the sound insulation of windows, concentrate basically on the glazing (QUIRT, 1982; 1983). In fact, other components of the window, mainly the frame and sealing also play an important role in the entire system performance. As the components mutually interact (NURZYŃSKI, 2020), the results of simplified calculations usually differ significantly from the experimental data (TADEU, MATEUS, 2001).

The acoustic performance of the opaque wall, in turn, depends on its main structure (BRADLEY, BIRTA, 2001; DAVY *et al.*, 2019) but also largely on various details, dividers, and connections (QUIRT *et al.*, 1992; LJUNGGREN, ÅGREN, 2011). The installation of a window modifies the structural scheme of the wall as the studs, faces, and other elements are partly removed. Finding any experimental research work focused on the acoustic effect of these modifications or numerical simulations, however, is really difficult (CANIATO, 2020). Articles on the façade sound insulation mostly concern field measurements (KIM, KIM, 2007; BURATTI *et al.*, 2014), the reduction of external noise annoyance (RYU, SONG, 2019; AMUNDSEN *et al.*, 2011) and the low-frequency behaviour (KERÄNEN *et al.*, 2019; SCROSATI *et al.*, 2016).

The paper concentrates on the sound insulation of lightweight prefabricated façades and, in particular, the acoustic effect of windows installed in the opaque element. The basic research question concerned the possibility of predicting the acoustic performance of a complete panel based on the performance of basic components. As the frequency-dependent characteristics of an opaque element and a window differ substantially, the use of single number values for prediction and accuracy was of particular interest. Two samples of an opaque wall and four windows were tested separately in the laboratory. Then, complete façades composed of these elements were examined. Finally, the empirical sound insulation of the entire panels was juxtaposed and compared with the results of estimations.

Basic terms and symbols used in the article: R_w is the weighted sound reduction index [dB]; C is the spectrum adaptation term corresponding to spectrum no. 1 (ISO 717-1:2020, 2020) [dB]; C_{tr} is the spectrum adaptation term corresponding to spectrum no. 2 (ISO 717-1:2020, 2020) [dB]; f_c is the coincidence frequency [Hz].

2. Samples and materials

2.1. Opaque elements

The acoustic performance of the opaque part depends on its basic structure, details, and additional layers, i.e., external thermal insulation and internal tech-

nical cladding (DI BELLA *et al.*, 2014). Two samples of the wall, 4220×2760 mm, with the same basic structure and internal cladding but different thermal insulation were considered. The basic wall was supported with a frame constructed of wood studs, 180×60 mm, spaced at 600 mm, firmly secured on a perimeter framing. Faces, made of fire-resistant 12.5 mm plasterboards, were screwed on both sides to the studs. The plenum inside was filled with mineral wool. The technical cladding was made of 12.5 mm plasterboards supported with wooden battens (studs), 50×60 mm, fastened rigidly to the main frame. The plenum was filled with 50 mm of mineral wool (Fig. 1). Technical data on facing boards is presented in Table 1.

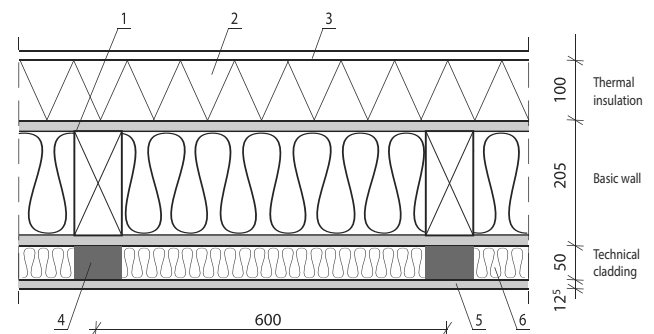


Fig. 1. Sample no. 1, façade with ETICS: 1) basic wall; 2) mineral wool 100 mm (lamella boards); 3) rendering; 4) studs; 5) boarding; 6) mineral wool 50 mm.

Table 1. Technical data on facing boards.

Board	Thickness [mm]	Density [kg/m ³]	Surface mass [kg/m ²]
Fire-resistant plasterboard	12.5	810	10.1
Fibre-cement board	8	1925	15.4

The first sample was equipped with an external thermal insulation composite system (ETICS), which in general reduces sound insulation in a certain frequency range due to the mass-spring-mass resonance (WEBER, 2003; SANTONI *et al.*, 2017). Several typical lightweight walls with ETICS made of mineral wool (MW) and expanded polystyrene (EPS), 100 mm and 120 mm thick, were initially examined to select the sample for further investigations. The results are presented in Fig. 2. The sound insulation characteristics were quite similar despite the different wall structures and various dynamic stiffness of the insulation. The values of the $R_w + C_{tr}$ index were within the range of 41–43 dB (NURZYŃSKI, 2022). The wall with ETICS consisted of 100 mm of mineral wool (lamella boards) and a thin rendering was finally selected as fairly representative of the wide range of lightweight façades of such a type (Fig. 1).

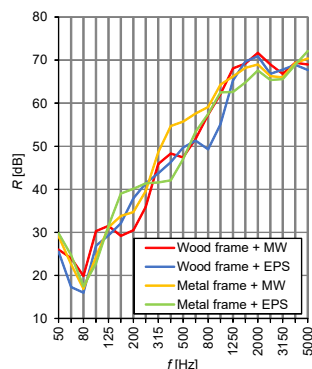


Fig. 2. Sound insulation of lightweight frame façades with ETICS, $R_w + C_{tr} = 41\text{--}43$ dB.

The second sample consisted of the same basic wall and had the same technical cladding but different thermal insulation. This was supported by a grid made of timber battens providing 30 mm ventilating cavity (Fig. 3). The cavity was opened by making two 30 mm slots in the external fibre-cement boarding along the bottom and upper edges of the test opening. Thermal insulation of such a type provides significantly better sound insulation and is recommended in the case of higher levels of outdoor noise.

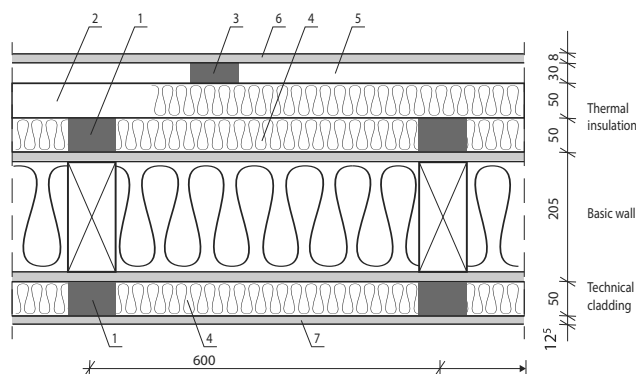


Fig. 3. Sample no. 2, façade with thermal insulation with a venting cavity: 1) studs; 2) horizontal battens; 3) vertical grid; 4) mineral wool 50 mm; 5) ventilating cavity 30 mm; 6) fibre-cement board 8 mm; 7) plasterboard 12.5 mm.

2.2. Windows

Four single-wing PVC windows, 1230×1480 mm, were tested individually in the laboratory. The windows were of the same system and manufacturer, had the same dimensions but different glazing. Single- and double-chamber insulated glass units (IGU) were used, respectively:

- Window A: double chamber IGU 4/16/4/16/4;
- Window B: double chamber IGU 10/12/4/12/6;
- Window C: double chamber IGU 44.1Si/12/4/12/44.1Si;
- Window D: single chamber IGU 66.2Si/24/86.2Si.

The glass units consisted of monolithic panes, 4, 6, and 10 mm thick, and laminated glass, 44.1Si, 66.2Si, 86.2Si, composed of two panes bonded with one or two layers of PVB ductile film (e.g., 66.2Si means two 6 mm panels bonded with two layers of laminate). The distance between panels in the unit was 12, 16, and 24 mm, respectively (see the denotation of samples).

2.3. Façades

The windows were installed successively in the opening made in the examined opaque element, and the installation was in accordance with field practice (Fig. 4). The openings were cut out in the elements after completing the first series of sound insulation measurements. The arrangement was representative for small rooms of about 12 m^2 of floor area (bedrooms).



Fig. 4. Sample of the façade installed in the laboratory.

3. Testing methods

The samples of the façade were tested in a facility consisting of two reverberant rooms of irregular shape so that the opposite surfaces were not parallel. The volume of the sending and receiving rooms was 100 m^3 and 93 m^3 , respectively. In order to suppress flanking transmission, the rooms were separated by an acoustic break. Additionally, a sound insulating lining was applied on the walls and ceiling in the receiving room. The samples were installed in accordance with ISO standard (ISO 10140-2:2021, 2021), and the external face of the wall was on the sending room side. The windows were tested in another facility consisting of two smaller reverberant rooms, 88 m^3 and 52 m^3 for the sending and receiving rooms, respectively. The rooms were separated by an acoustic break, and a sound insulating lining was applied in the receiving room. The windows were installed in a double, massive filler wall constructed of calcium silicate blocks $250 \text{ mm} + 200 \text{ mm}$, separated with vibration brake filled with mineral wool. The windows were fastened in the opening and sealed on the perimeter. The facilities, measurement procedures and the equipment complied with the requirements of respective ISO standard (ISO 10140-4:2021, 2021). A dual channel analyser and rotating microphones were used for the measurements. Average sound pressure levels in 1/3 octave bands were measured in the source and receiving rooms, and integrated over time and space.

4. Results of measurements and discussion

4.1. Windows

The windows were tested successively in the test opening provided in the heavy filler wall between both rooms of the laboratory. The values of the $R_w + C_{tr}$ index ranged from 31 dB to 45 dB with an average step of 5 dB (Table 2). This covers a comprehensive performance spectrum of commonly used windows equipped with typical insulated glass units (MISKINIS *et al.*, 2015). The frequency-dependent characteristics were very different due to the IGU structure, the type of panes, the mass per unit area and the distance between them (Fig. 5). The fundamental resonance of the double (triple) glazing system and the coincidence of single panels determined the shape of the sound insulation plots.

Table 2. Sound insulation of windows (single number values).

Window	R_w [dB]	$R_w + C$ [dB]	$R_w + C_{tr}$ [dB]
A (4/16/4/16/4)	36	35	31
B (10/12/4/12/6)	40	39	36
C (44.1Si/12/4/12/44.1Si)	45	43	39
D (66.2Si/24/86.2Si)	47	47	45

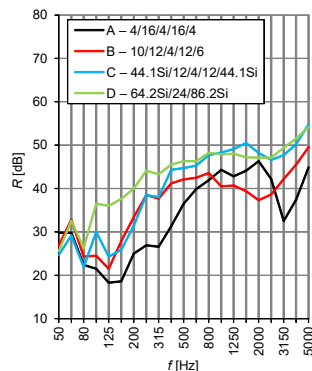


Fig. 5. Sound insulation of windows with different glazing, results of measurements.

4.2. Opaque elements

Two samples of the opaque element were examined. The measurements were taken in three different phases of their construction, i.e., for the basic wall without any insulating layers, with insulation but without external finishing (i.e., without the rendering in the sample no. 1 and the fibre-cement board in the sample no. 2), and finally for the complete façades. The sound insulation characteristics of the first sample are presented in Fig. 6. The insulating layers applied to the basic wall (without finishing) increased greatly the sound insulation at middle and high frequencies. Due to low frequency behaviour and whole structure resonance,

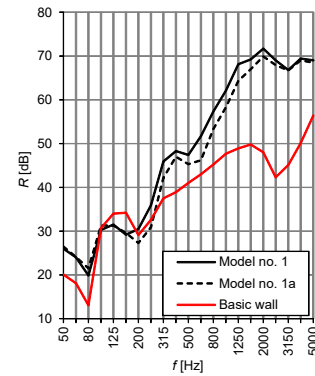


Fig. 6. Sample no. 1. Sound insulation measured in different phases of construction (model no. 1a without rendering).

however, the single number value of $R_w + C_{tr}$ did not improve at all. This equals 40 dB with and without insulation (Table 3). Implementation of the rendering, that is applying just the thin external finishing of ETICS, slightly moved the mass-spring-mass resonance towards low frequencies (Fig. 6) and, in effect, the single number quantity increased by 2 dB.

Table 3. Sound insulation of opaque elements (single number quantities).

Opaque element	R_w [dB]	$R_w + C$ [dB]	$R_w + C_{tr}$ [dB]
Basic wall	44	43	40
Sample no. 1a (without rendering)	47	44	40
Sample no. 1	50	47	42
Sample no. 2a (without external face)	53	51	46
Sample no. 2	60	58	54

The thermal insulation system with a venting cavity applied in the second sample was significantly more effective, particularly at low and middle frequencies (Fig. 7). The application of a bare insulation made of mineral wool was highly beneficial only at middle

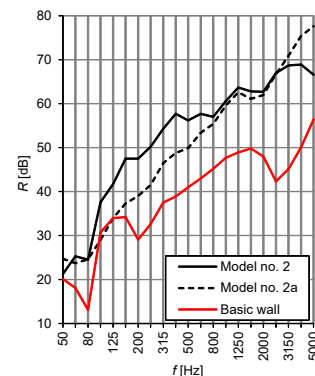


Fig. 7. Sample no. 2. Sound insulation measured in different phases of construction (model no. 2a without external faces).

and high frequencies, while the fibre-cement faces improved sound insulation considerably in the range of 100–630 Hz. Surprisingly, however, a pronounced lowering was observed in the range of 800–2500 Hz, where the sound insulation of the complete façade is quite the same as that of the façade without external finishing made of fibre-cement boards. This was probably caused by the ventilating slots made in the faces that opened the cavity along the upper and lower edges of the façade. The $R_w + C_{tr}$ index without external faces was 46 dB, whereas for the complete façade it was 54 dB (Table 3). The coincidence effect of a fibre-cement board, 8 mm thick, was observed in the high frequency range (theoretically calculated $f_c = 5200$ Hz), whereas low frequency behaviour was determined by the fundamental resonance of the basic wall (Fig. 7). The examined samples of opaque elements are fairly representative of a comprehensive range of lightweight external walls used in real life in prefabricated residential buildings.

4.3. Complete façades

The frequency-dependent characteristics of windows and opaque elements tested separately were very different (Figs. 5–7). This means that both components contribute to the total sound insulation of the entire façade in different ways. The measurement results for sample no. 1 examined in various wall–window configurations are shown in Fig. 8. Generally, the window determined the total sound insulation at middle and high frequencies. In the range beneath 800 Hz, the characteristics of sample 1B, 1C, 1D (with windows) were practically the same as for the stand-alone opaque part of the façade. However, the differences observed above 800 Hz did not significantly influence the single number quantities; the values of the $R_w + C_{tr}$ index gained 41–42 dB regardless of the window's presence and the type of glazing (Table 4). Thus, the final effect expressed in terms of $R_w + C_{tr}$ was approximately the same, which clearly indicates the need for optimisation in the façade designing process.

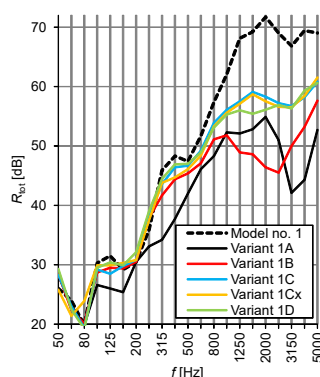


Fig. 8. Sample no. 1. Total sound insulation of complete façade with different windows.

Table 4. Total sound insulation of complete façades with different windows (single number quantities).

Sample no.	R_w [dB]	$R_w + C$ [dB]	$R_w + C_{tr}$ [dB]
1	50	47	42
1A	43	42	37
1B	46	44	41
1C	49	46	41
1D	49	47	42
2	60	58	54
2A	44	43	39
2B	48	47	44
2C	52	50	46
2D	52	51	48

The result obtained for the modified variant 1C without internal technical cladding (1Cx) is interesting as the removal of the cladding practically had no acoustic effect (Fig. 8). This confirms that such a supplementary layer, when rigidly fastened to the main structure of the lightweight frame wall, does not improve its sound insulation (NURZYŃSKI, 2022). The arrangement 1A had a window of lower sound insulation and, consequently, lowering in total sound insulation may be observed nearly across the entire frequency range (Fig. 8). The value of the $R_w + C_{tr}$ index dropped by 5 dB in comparison with the standalone opaque element.

The acoustic performance of the second sample was utterly determined by windows. Subsequent to their assembling, the total sound insulation decreased dramatically in almost the entire frequency range (Fig. 9). It is interesting, however, that the installation of windows brought about some slight improvements at low frequencies. The $R_w + C_{tr}$ index of samples 2B, 2C, and 2D dropped by 6–10 dB compared to the stand-alone opaque element. The differences between subsequent samples were relatively small, measuring just 2 dB (Table 4). The use of the window A seems impractical, as in this case the index dropped by 15 dB.

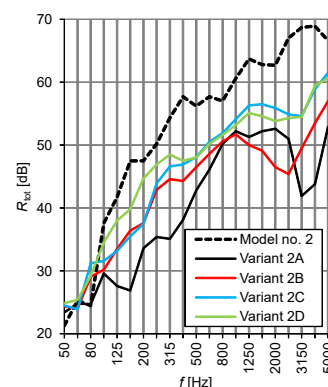


Fig. 9. Sample no. 2. Total sound insulation of complete façade with different windows.

The differences between single number quantities for respective variants of both samples, with a different opaque part but the same window, are presented in Table 5. The results confirm the need for façade optimisation, as the use of a highly insulated opaque element had rather limited effects compared to a wall equipped with common ETICS. The optimal wall–window configuration should be determined in reference to local requirements (RASMUSSEN, RINDEL, 2010; RASMUSSEN, 2010), while the development of a uniform acoustic categorisation for the whole prefabricated panels would be helpful for designers, manufacturers, and other stakeholders referring to existing acoustic classification schemes for residential buildings (NURZYŃSKI, 2007; CASINI *et al.*, 2016).

Table 5. Differences in single number values for respective variants of both samples.

Variant no.	ΔR_w [dB]	$\Delta(R_w + C)$ [dB]	$\Delta(R_w + C_{tr})$ [dB]
2A–1A	1	1	2
2B–1B	2	3	3
2C–1C	3	4	5
2D–1D	3	4	6

5. Estimations versus empirical results

Empirical sound insulation of the façade elements, opaque parts, and windows, was used for calculating total sound insulation of the complete façade:

$$R_{\text{tot}} = -10 \log \left(\sum_{i=1}^n \frac{S_i}{S} 10^{-0.1 R_i} \right), \quad (1)$$

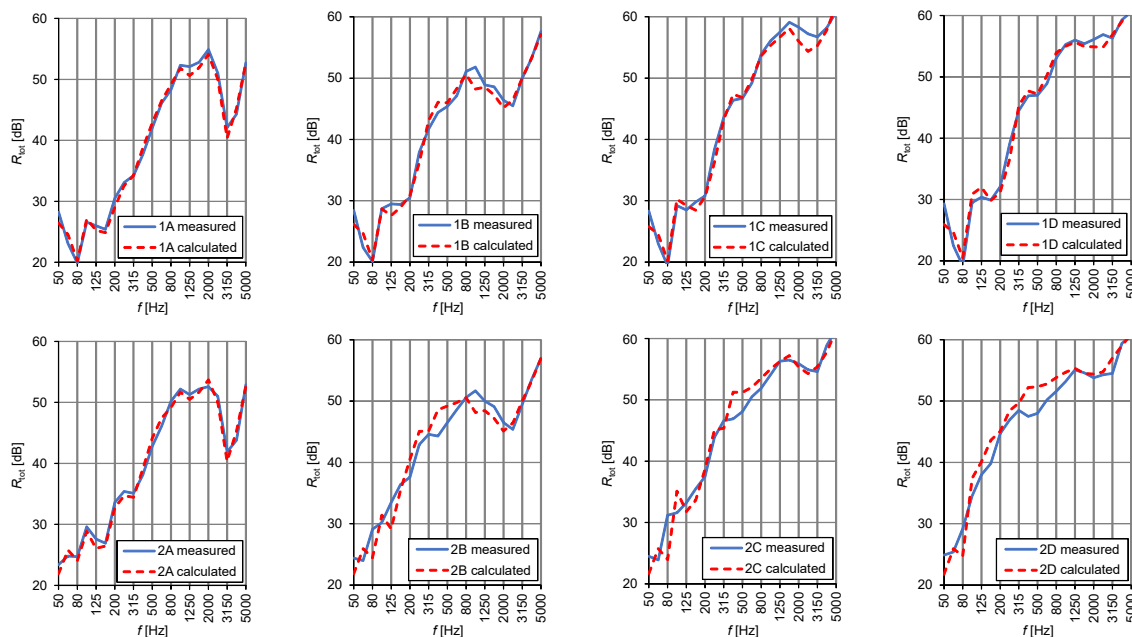


Fig. 10. Total sound insulation of complete façades with different windows: comparison of experimental and calculated values.

where R_{tot} is the total sound insulation of the complete façade [dB], S is the total area of the façade [m²], R_i is the sound insulation of part i of the façade [dB], S_i is the area of a part i of the façade [m²].

The results of calculations were fairly consistent with the experimental values for practically the entire frequency range (Fig. 10). This consistency is somewhat surprising, especially as the opaque elements and windows were tested in two distinctly different test facilities.

The estimations based directly on single number quantities were also in good agreement with the results of measurements. The calculated values rounded to an integer; in almost all cases, they were the same as measured. The sample 2D, however, formed an exception, as the calculated indices were higher by 2–3 dB. This was probably caused by window assembly failures that brought about tiny cracks locally decreasing the airtightness of the system. Some interaction between the wall and the window, however, can also be a reason for this discrepancy.

6. Conclusions

The acoustic effect of windows installed in lightweight frame façades was examined. The study investigated the acoustic interaction of façade components and aimed to verify whether the formula for the total sound insulation of a complex partition is applicable in this case. The frequency-dependent experimental results for the entire façade panels were fairly consistent with calculated values. The estimations based on single number quantities were also in good agreement with measurements. Thus, it may be concluded that

the façade elements did not interact significantly and that the single number calculations give reliable results that can be used in practice. This conclusion provide directions for the further work on the assessment of prefabricated frame buildings and may be useful while optimising their acoustic performance. The development of an acoustic categorisation and the marking scheme for prefabricated façade panels considered as final building products would be a helpful tool for manufacturers and designers.

The windows and opaque elements influenced total sound insulation in a different frequency range. The installation of windows B, C, and D in the opaque wall with ETICS (sample no. 1) reduced sound insulation considerably in the high frequency bands, mostly above 800 Hz. Due to low and medium frequency behaviour, however, the single number values of $R_w + C_{tr}$ were practically the same. The acoustic performance of the second sample was utterly determined by windows. Surprisingly, however, their installation brought about some improvements at low frequencies. The removal of the technical cladding in the sample no. 1 had practically no acoustic effect, which confirms that such a supplementary layer, when rigidly fastened to the main structure of a lightweight frame wall, does not improve sound insulation. These observations and conclusions may be useful for designers and engineers working on external wall structures. Future research should be focused on the acoustic effect of another technical elements installed in façade panels such as slot ventilators, shutter boxes, air transfer devices, electrical raceways, etc.

References

1. AMUNDSEN A.H., KLÆBOE R., AASVANG G.M. (2011), The Norwegian façade insulation study: The efficacy of façade insulation in reducing noise annoyance due to road traffic, *The Journal of the Acoustical Society of America*, **129**(3): 1381–1389, doi: 10.1121/1.3533740.
2. BRADLEY J.S., BIRTA J.A. (2001), On the sound insulation of wood stud exterior walls, *The Journal of the Acoustical Society of America*, **110**(6): 3086–3096, doi: 10.1121/1.1416200.
3. BURATTI C., BELLONI E., MORETTI E. (2014), Façade noise abatement prediction: New spectrum adaptation terms measured in field in different road and railway traffic conditions, *Applied Acoustics*, **76**: 238–248, doi: 10.1016/j.apacoust.2013.08.016.
4. CANIATO M. (2020), Sound insulation of complex façades: A complete study combining different numerical approaches, *Applied Acoustics*, **169**: 107484, doi: 10.1016/j.apacoust.2020.107484.
5. CANIATO M., BETTARELLO F., FERLUGA A., MAR-SICH L., SCHMID C., FAUSTI P. (2017), Acoustic of lightweight timber buildings: A review, *Renewable and Sustainable Energy Reviews*, **80**: 585–596, doi: 10.1016/j.rser.2017.05.110.
6. CASINI D., CELLAI G., FOGOLA J., SCAMONI F., SECCHI S. (2016), Correlation between facade sound insulation and urban noise: A contribution to the acoustic classification of existing buildings, *Building Acoustics*, **23**(3): 145–158, doi: 10.1177/1351010X16670173.
7. DAVY J.L., FARD M., DONG W., LOVERDE J. (2019), Empirical corrections for predicting the sound insulation of double leaf cavity stud building elements with stiffer studs, *The Journal of the Acoustical Society of America*, **145**(2): 703–713, doi: 10.1121/1.5089222.
8. DI BELLA A., GRANZOTTO N., PAVARIN C. (2014), Comparative analysis of thermal and acoustic performance of building elements, [in:] *Proceedings of Forum Acusticum. European Acoustics Association, EAA*, Kraków.
9. ISO 10140-2:2021 (2021), Acoustics – Laboratory measurement of sound insulation of building elements – Part 2: Measurement of airborne sound insulation.
10. ISO 10140-4:2021 (2021), Acoustics – Laboratory measurement of sound insulation of building elements – Part 4: Measurement procedures and requirements.
11. ISO 717-1:2020 (2020), Acoustics – Rating of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation.
12. KERÄNEN J., HAKALA J., HONGISTO V. (2019), The sound insulation of façades at frequencies 5–5000 Hz, *Building and Environment*, **156**: 12–20, doi: 10.1016/j.buildenv.2019.03.061.
13. KIM M.J., KIM H.G. (2007), Field measurements of façade sound insulation in residential buildings with balcony windows, *Building and Environment*, **42**(2): 1026–1035, doi: 10.1016/j.buildenv.2005.10.036.
14. LIEBL A., SPÄH M., BARTLOMÉ O., KITTEL M. (2013), Evaluation of acoustic quality in wooden buildings, [in:] *Proceedings of 42nd International Congress and Exposition. Noise Control Engineering, INTER-NOISE Noise Control for Quality of Life 2*, pp. 1492–1500, Innsbruck.
15. LJUNGGREN F., ÅGREN A. (2011), Potential solutions to improved sound performance of volume based lightweight multi-storey timber buildings, *Applied Acoustics*, **72**(4): 231–240, doi: 10.1016/j.apacoust.2010.11.007.
16. MISKINIS K., DIKAVICIUS V., BLIUDZIUS R., BANIONIS K. (2015), Comparison of sound insulation of windows with double glass units, *Applied Acoustics*, **92**: 42–46, doi: 10.1016/j.apacoust.2015.01.007.
17. NURZYŃSKI J. (2007), Evaluation of acoustic performance of multifamily buildings, open building manufacturing approach, [in:] *Proceedings of 36th International Congress and Exhibition on Noise Control Engineering, INTER-NOISE 2007*, pp. 1770–1779, Istanbul.
18. NURZYŃSKI J. (2020), Influence of sealing on the acoustic performance of PVC windows, [in:] *Research in Building Physics*, pp. 595–603, doi: 10.1201/9781003078852-82.

19. NURZYŃSKI J. (2022), Sound insulation of lightweight external frame walls and the acoustic effect of additional thermal insulation, *Applied Acoustics*, **190**: 108645, doi: 10.1016/j.apacoust.2022.108645.
20. QUIRT J.D. (1982), Sound transmission through windows I. Single and double glazing, *The Journal of the Acoustical Society of America*, **72**(3): 834–844, doi: 10.1121/1.388263.
21. QUIRT J.D. (1983), Sound transmission through windows II. Double and triple glazing, *The Journal of the Acoustical Society of America*, **74**(2): 534–542, doi: 10.1121/1.389819.
22. QUIRT J.D., WARNOCK A., HALLIWELL R., BIRTA J. (1992), Influence of stud type and spacing, screw spacing, and sound absorbing material on the sound transmission through a double panel wall specimen, *The Journal of the Acoustical Society of America*, **92**(4): 2470–2470, doi: 10.1121/1.404482.
23. RASMUSSEN B. (2010), Sound insulation between dwellings – Requirements in building regulations in Europe, *Applied Acoustics*, **71**(4): 373–385, doi: 10.1016/j.apacoust.2009.08.011.
24. RASMUSSEN B., RINDEL J.H. (2010) Sound insulation between dwellings – Descriptors applied in building regulations in Europe, *Applied Acoustics*, **71**(3): 171–180, doi: 10.1016/j.apacoust.2009.05.002.
25. RYU J., SONG H. (2019), Effect of building façade on indoor transportation noise annoyance in terms of frequency spectrum and expectation for sound insulation, *Applied Acoustics*, **152**: 21–30, doi: 10.1016/j.apacoust.2019.03.020.
26. SANTONI A., BONFIGLIO P., DAVY J.L., FAUSTI P., POMPOLI F., PAGNONCELLI L. (2017), Sound transmission loss of ETICS cladding systems considering the structure-borne transmission via the mechanical fixings: Numerical prediction model and experimental evaluation, *Applied Acoustics*, **122**: 88–97, doi: 10.1016/j.apacoust.2017.02.008.
27. SCROSATI C. *et al.* (2016), Uncertainty of facade sound insulation by a Round Robin Test. Evaluations of low-frequency procedure and single numbers, *Building and Environment*, **105**: 253–266, doi: 10.1016/j.buildenv.2016.06.003.
28. ŠUJANOVÁ P., RYCHTÁŘIKOVÁ M., MAYOR T.S., HYDER A. (2019), A healthy, energy-efficient and comfortable indoor environment, a review, *Energies*, **12**(8): 1414, doi: 10.3390/en12081414.
29. TADEU A.J.B., MATEUS D.M.R. (2001), Sound transmission through single, double and triple glazing. Experimental evaluation, *Applied Acoustics*, **62**(3): 307–325, doi: 10.1016/S0003-682X(00)00032-3.
30. UTLEY W.A., FLETCHER B.L. (1969), Influence of edge conditions on the sound insulation of windows, *Applied Acoustics*, **2**(2): 131–136, doi: 10.1016/0003-682X(69)90015-2.
31. WEBER L. (2003), Uniform acoustical characterization of External Thermal Insulation Systems, IBP report B-BA 6/2002, Stuttgart.

Research Paper

Infrasound Signal Classification Based on ICA and SVM

Quanbo LU[✉], Meng WANG, Mei LI*

*School of Information Engineering, China University of Geosciences
Beijing, China*

*Corresponding Author e-mail: maggieli@cugb.edu.cn

(received July 7, 2022; accepted December 7, 2022)

A diagnostic technique based on independent component analysis (ICA), fast Fourier transform (FFT), and support vector machine (SVM) is suggested for effectively extracting signal features in infrasound signal monitoring. Firstly, ICA is proposed to separate the source signals of mixed infrasound sources. Secondly, FFT is used to obtain the feature vectors of infrasound signals. Finally, SVM is used to classify the extracted feature vectors. The approach integrates the advantages of ICA in signal separation and FFT to extract the feature vectors. An experiment is conducted to verify the benefits of the proposed approach. The experiment results demonstrate that the classification accuracy is above 98.52% and the run time is only 2.1 seconds. Therefore, the proposed strategy is beneficial in enhancing geophysical monitoring performance.

Keywords: independent component analysis; fast Fourier transform; support vector machine; infrasound signal.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Infrasound (generally less than 20 Hz) is a low-frequency sound produced by natural and anthropogenic events. The frequency of infrasound signals typically is under 20 Hz (GI, BROWN, 2017; MCKEE *et al.*, 2018). Although it cannot be heard by the human auditory system, it widely exists in the world around us. Infrasound can be produced by natural events such as earthquakes, tsunamis, mudslides, tornadoes, and volcano eruptions (LIU *et al.*, 2021). Human induced events such as missile launches, ship navigation, and nuclear explosions can produce infrasound (ZHAO *et al.*, 2021). Infrasound is low frequency, long wavelength sound wave, accessible to diffraction, and not easily absorbed by the medium (MAYER *et al.*, 2020; CÁRDENAS-PEÑA *et al.*, 2013; CANNATA *et al.*, 2011). Therefore, infrasound can be employed in natural disaster monitoring.

Some scholars study infrasound signal classification algorithms and apply them to monitor infrasound signals. THÜRING *et al.* (2015) classified the infrasound data from the avalanche control site near Lavin in the Swiss Alps via SVM. The false detection rate was reduced from 65% to 10%, and the classification performance was significantly improved. Reliable help was

provided for establishing the automatic detection system of infrasound avalanche (IEZZI *et al.*, 2019). TSYBUL'SKAYA *et al.* (2012) classified atmospheric infrasonic signals based on the theory of testing statistical hypotheses. HAM *et al.* (2008) used radial basis function (RBF) network as the subnetworks of parallel neural network classifier bank to classify six different infrasound events. The classification accuracy reached more than 93%. Through data mining classification algorithms, the feature extraction can be conducted on the signal to achieve a better classification. LIU *et al.* (2014) used three types of feature extraction techniques (spectral entropy, discrete wavelet transformation (DWT), and Hilbert-Huang transform (HHT)) to extract the feature vector of four types of infrasound signals. The signal feature was extracted by back propagation neural network and SVM for classification. As a result, SVM has a greater classification accuracy (LI *et al.*, 2016). However, these methods do not separate the signal from the noise, which may limit their accuracy.

This research provides an approach for monitoring infrasound signal. The proposed technique first applies a blind source separation (BSS) method based on ICA to extract useful signals from mixed infrasound sig-

nals. Then fast Fourier transform (FFT) is carried out for feature extraction. Finally, SVM is utilized to classify the infrasound signals based on the retrieved features. The infrasound signal experiment is conducted to validate the superiority of our proposed technique. It provides a practical mechanism for real-time monitoring and analysis of infrasound signals (CHERNOGOR, SHEVELEV, 2018).

The remainder of this work is organized as follows. The methodologies and algorithms used in this paper are briefly discussed in Sec. 2. In Sec. 3, the performance of the proposed approach is compared in an experiment. Section 4 shows the experimental results through the analysis of different methods. Finally, conclusions are drawn in Sec. 5.

2. Methods

2.1. Source signal extraction based on independent component analysis

BSS is a well-known concept for separating mixed signals (SASTRY *et al.*, 2021). The word “blind” refers to the fact that source signals can be separated even if little information about them is available (MIKA, KLECZKOWSKI, 2011). One of the most widely-used examples of BSS is to separate voice signals of people speaking at the same time. This is called the cocktail party problem. This problem aims to detect or extract the sound with a single object even though different sounds in the environment are superimposed on another (MIKA, KLECZKOWSKI, 2011). Independent component analysis (ICA) is an analysis method of high-order statistics. It can separate a non-Gaussian characteristic component from mixed signals (QIAN *et al.*, 2019). Besides, ICA can remove the aliasing noise from signals without destroying the details. ICA has been applied in many fields and is one of the most powerful tools for data analysis.

As shown in Fig. 1, presenting a linear model of ICA, let x_1, x_2, \dots, x_n be an n -dimensional random observation mixed signal, which is a linear combination of s_1, s_2, \dots, s_m . Each observation $x_i(t)$ is a sample of the random variable. Let the mixed random variables and independent sources have zero mean. The model is defined in matrix form. Let $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$ be an n -dimensional random observation vector and $\mathbf{S} = (s_1, s_2, \dots, s_m)^T$ be an m -dimensional unknown source signal. Then the ICA can be expressed as:

$$\mathbf{X} = \mathbf{A}\mathbf{S} = \sum_{i=1}^m a_i s_i, \quad i = 1, 2, \dots, m. \quad (1)$$

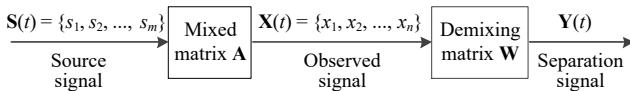


Fig. 1. Linear model of ICA.

Among them, s_i is the independent component, the mixed matrix $\mathbf{A} = [a_1, a_2, \dots, a_n]$ is a full rank matrix, and a_i is the base vector of the mixing matrix. In Eq. (1), we can see that each observation data x_i is obtained by the independent source s_i with different a_i linear weighting. The independent source s_i is an implicit variable. It cannot be directly measured. The mixed matrix \mathbf{A} is also an unknown matrix, so the only available information is the observed random vector \mathbf{X} . Without any restriction to estimate \mathbf{S} and \mathbf{A} from \mathbf{X} , the solution of Eq. (1) must be multiple solutions. However, according to the statistical characteristics of \mathbf{X} , ICA will give the unique solution of the equation under certain constraints. The unique solution can extract independent components.

The unknown source signals are independent, which is a fundamental basic assumption. As a result, the probability density function can be written as:

$$p(s) = \prod_{i=1}^m p_i(s_i), \quad (2)$$

where $p(s)$ represents the probability density function of the source signal, $p_i(s_i)$ signifies the independent component's probability density function (MIKA, KLECZKOWSKI, 2011), s is the source signal, s_i stands for the independent component, and m is chosen as dimension.

The source signals are independent in the ICA model (MIKA, KLECZKOWSKI, 2011). The independent signals will have a non-Gaussian distribution. Non-Gaussianity is measured by kurtosis. In order to simplify the model, it is assumed that the unknown mixing matrix \mathbf{A} is a square matrix $m = n$. The purpose of ICA is to find a transformation matrix. The linear transformation \mathbf{X} is used to obtain the n -dimensional output vector:

$$\mathbf{Y} = \mathbf{W}\mathbf{X} = \mathbf{W}\mathbf{A}\mathbf{S}, \quad (3)$$

where \mathbf{Y} is the approximate signal of \mathbf{S} , namely $\mathbf{Y} = \widehat{\mathbf{S}}$, the initial weight vector for the unmixing matrix \mathbf{W} is arbitrary. However, due to \mathbf{A} and \mathbf{S} being unknown, \mathbf{Y} has two uncertainties (the uncertainty of amplitude of separated signals and the uncertainty of order of separated signals). These uncertainties have no impact on the outcome.

2.2. Fast Fourier transform

COOLEY and TUKEY (1965) skillfully used the periodicity and symmetry of the \mathbf{W}_n factor to construct a fast algorithm for discrete Fourier transform (DFT), namely FFT. In the following decades, FFT was further developed. At present, the radix-2 and split-radix algorithms are commonly used.

When discussing the mathematical transformation of images, we consider images as functions with two

variables x and y . First, the Fourier transformation of a two-dimensional continuous function is introduced. Let $f(x, y)$ be a function of two independent variables x and y . It satisfies $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)| dx dy < 0$.

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\delta(ux+vy)} dx dy, \quad (4)$$

where $F(u, v)$ is the Fourier transform of $f(x, y)$ and $f(x, y)$ is the inverse Fourier transform:

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{-j2\delta(ux+vy)} du dv. \quad (5)$$

The amplitude spectrum, phase spectrum and energy spectrum of the Fourier transform are as follows:

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{1/2}, \quad (6)$$

$$\varphi(u, v) = \arctg[I(u, v)/R(u, v)], \quad (7)$$

$$E(u, v) = R^2(u, v) + I^2(u, v), \quad (8)$$

where $R(u, v)$ and $I(u, v)$ represent the real part and imaginary part of the Fourier transform, respectively.

2.3. Classification using SVM

CORTES and VAPNIK (1995) proposed SVM, a pattern recognition method developed based on statistical theory. In order to get the best generalization ability, SVM seeks the optimum balance between model complexity and learning ability based on limited sample information (AMARNATH, 2016). In the field of pattern recognition, SVM is mainly employed to handle the challenge of data classification. It demonstrates a number of distinct advantages in solving tiny samples, non-linear pattern recognition, and high-dimensional pattern recognition (AMARNATH, 2016). SVM may be used to solve a wide range of machine learning.

As shown in Fig. 2, the sample C1 is a positive sample, the sample C2 is a negative sample, and a linear function $g(x) = w^T x + b$ is required to separate C1 from C2. This is the case in two-dimensional space, and in three-dimensional space to separate C1 from C2 a face is required, and in the n -dimensional space an $n-1$ -dimensional hyper-plane is required to be

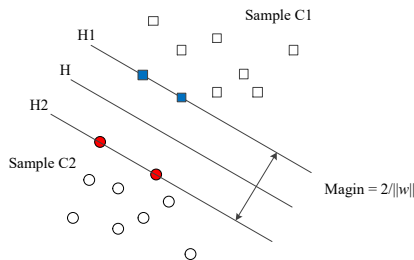


Fig. 2. SVM classification model.

separated. So, the separation of the hyperplane is expressed as:

$$g(x_i) = \langle w^T, x_i \rangle + b = 0. \quad (9)$$

The geometric interval between H1 and H, H2 and H is:

$$d = y_i \cdot (w^T x + b) \cdot \frac{1}{\|w\|}. \quad (10)$$

According to Eq. (10), it is necessary to find the nearest point (support vector) of the distance hyperplane in the sample, optimize w and b , and maximize the distance from the support vector to the hyperplane. It is a quadratic programming (QP):

$$\min \frac{1}{2} \|w\|^2 \quad (11)$$

subject to $y_i [(w^T x_i) + b] - 1 \geq 0 \quad (i = 1, 2, \dots, n).$

According to the Lagrange multiplier method, w can be expressed as:

$$w = \alpha_1 x_1 y_1 + \alpha_2 x_2 y_2 + \dots + \alpha_n x_n y_n = \sum_{i=1}^n (\alpha_i x_i y_i), \quad (12)$$

where α_i represents Lagrange multiplier, x_i represents sample points, y_i represents the category label of the i -th sample, and n is the number of samples. In the formula, only the sample point (support vector) belonging to H1 and H2 is not equal to zero, and these non-zero sample points only determine the classification function. Substituting Eq. (12) into Eq. (9) produces:

$$w^T x + b = \left(\sum_{i=1}^n \alpha_i y_i x_i \right)^T x + b = \sum_{i=1}^n \alpha_i y_i \langle x_i, x \rangle + b, \quad (13)$$

where $\langle x_i, x \rangle$ is the Kernel function $K(x_i, x)$ of SVM. The Kernel function can convert a sample from a low-dimensional space to a high-dimensional space, allowing it to be separated linearly (AMARNATH, 2016). At present, the choice of Kernel function mainly relies on experience. However, because the RBF is preferable in general, the RBF function is chosen as the Kernel function in this research (CHERNOGOR, SHEVELEV, 2018):

$$K(x, x_i) = \exp \left(-\frac{\|x - x_i\|^2}{\sigma^2} \right). \quad (14)$$

3. Experiment

3.1. Data set and tool

The data used in this paper originates from the International Monitoring System (IMS) with the help of the Comprehensive Nuclear-Test-Ban Treaty Beijing National Data Center (LIU *et al.*, 2014). This study divides infrasound incidents into three types. The data is gathered from sex separate infrasound sensor arrays located all around the world. This study uses 611 sets of data. The details of infrasound data collected from various regions are shown in Table 1. Earthquake,

Table 1. Infrasound data summary.

Event type	Data source (IMS Station Code)	Geographic coordinate	Number of signals	Total	Sampling frequency [Hz]
Earthquake	I14CL	(−33.65, −78.79)	74	203	20
	I30JP	(35.31, 140.31)	124		20
	I59US	(19.59, −155.89)	6		20
Tsunami	I10CA	(50.20, −96.03)	4	218	20
	I22FR	(−22.18, 166.85)	53		20
	I30JP	(35.31, 140.31)	113		20
	I52GB	(−7.38, 72.48)	66		20
Volcano	I30JP	(35.31, 140.31)	189	189	20

tsunami, and volcano eruption are the three types of infrasound events (Li *et al.*, 2016).

All 611 infrasound signal recordings have a sampling frequency of 20 Hz. Figure 3 depicts the infrasound stations' map.

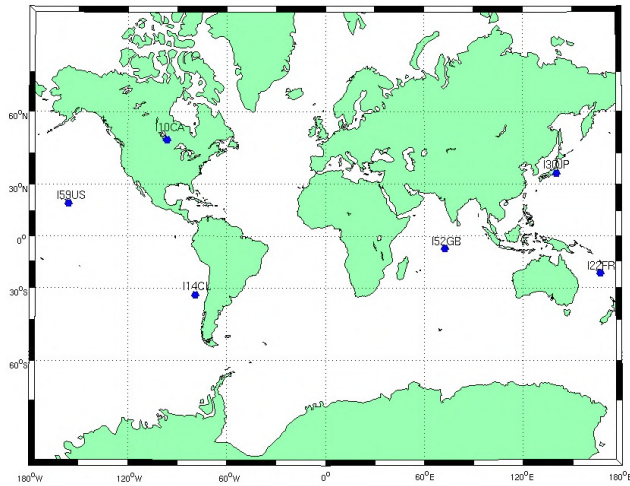


Fig. 3. Map of the infrasound stations.

3.2. Experiment setup

Figure 4 depicts the various categorization model frameworks, while Table 1 contains the data information. Figure 4a shows that the infrasonic signal is transformed by FFT. As a result, the feature extraction obtains three types of feature vectors. Each class is randomly divided into two groups: the training group and the testing group. The proportion of the training group and the testing group is around two to one. The SVM classification model is first trained by the training group. The SVM classification model is next tested by the testing group. Finally, the classification results and accuracy are given. In Fig. 4b ICA is added to remove aliasing noise from the signal without destroying the details of the signal, and then the separated signal is transformed by FFT. As a result, feature extraction obtains three types of feature vectors. Each class is randomly divided into two

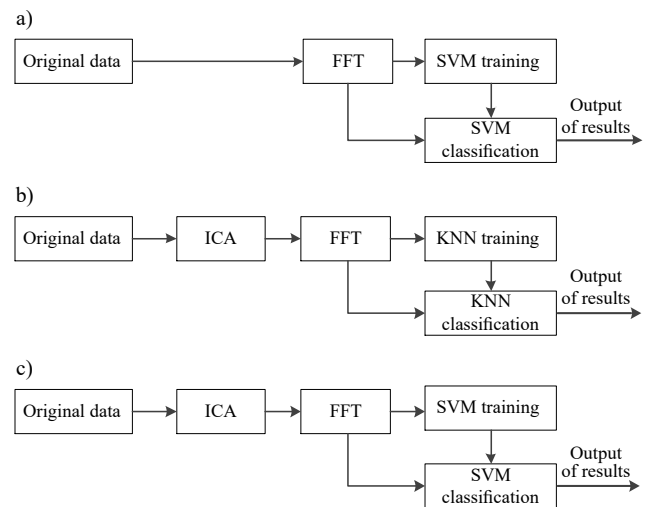


Fig. 4. Classification model framework.

groups: the training group and the testing group. Data from the training group outnumbers data from the test group by around two to one. The K-nearest neighbor (KNN) classification model is first trained by the training group. The KNN classification model is then tested by the testing group. Finally, the classification results and accuracy are given. In Fig. 4c ICA is added to remove aliasing noise in the signal without destroying the details of the signal, and then the separated signal is transformed by FFT. As a result, feature extraction obtains three types of feature vectors. Each class is randomly divided into two groups: the training group and the testing group. Data from the training group outnumbers data from the test group by around two to one. The SVM classification model is first trained by the training group. The SVM classification model is next tested by the testing group. Finally, the classification results and accuracy are given.

3.3. Data preprocessing

The original plot of three infrasonic events is shown in Fig. 5. In Fig. 5, we can see that it is difficult to sep-

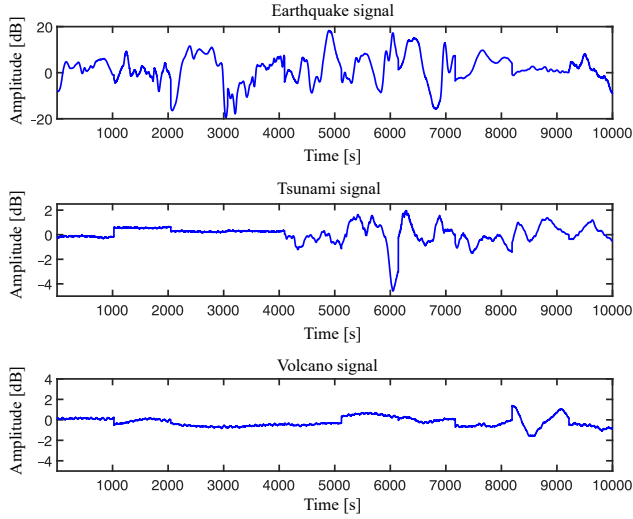


Fig. 5. Original plot of three infrasonic events.

arate different infrasound signals. Therefore, we should extract the feature vectors from the infrasound signal. The feature vectors of three infrasonic events extracted by FFT are presented in Fig. 6. Figure 7 shows the feature vectors of three infrasonic events based on ICA and FFT.

SVM is a supervised classification algorithm. SVM is trained by the training set to obtain the optimal parameters. The final classification result of SVM is obtained by the testing set. The training set data and testing set data of SVM are presented in Table 2. There are 611 infrasound signals, including earthquake, tsunami, and volcano eruption, collected from six infrasound stations. The sampling frequency is 20 Hz. Due to the different lengths of infrasound signals obtained at each infrasound station, all signals need to be truncated, and the data length used in all tests is 1024 points.

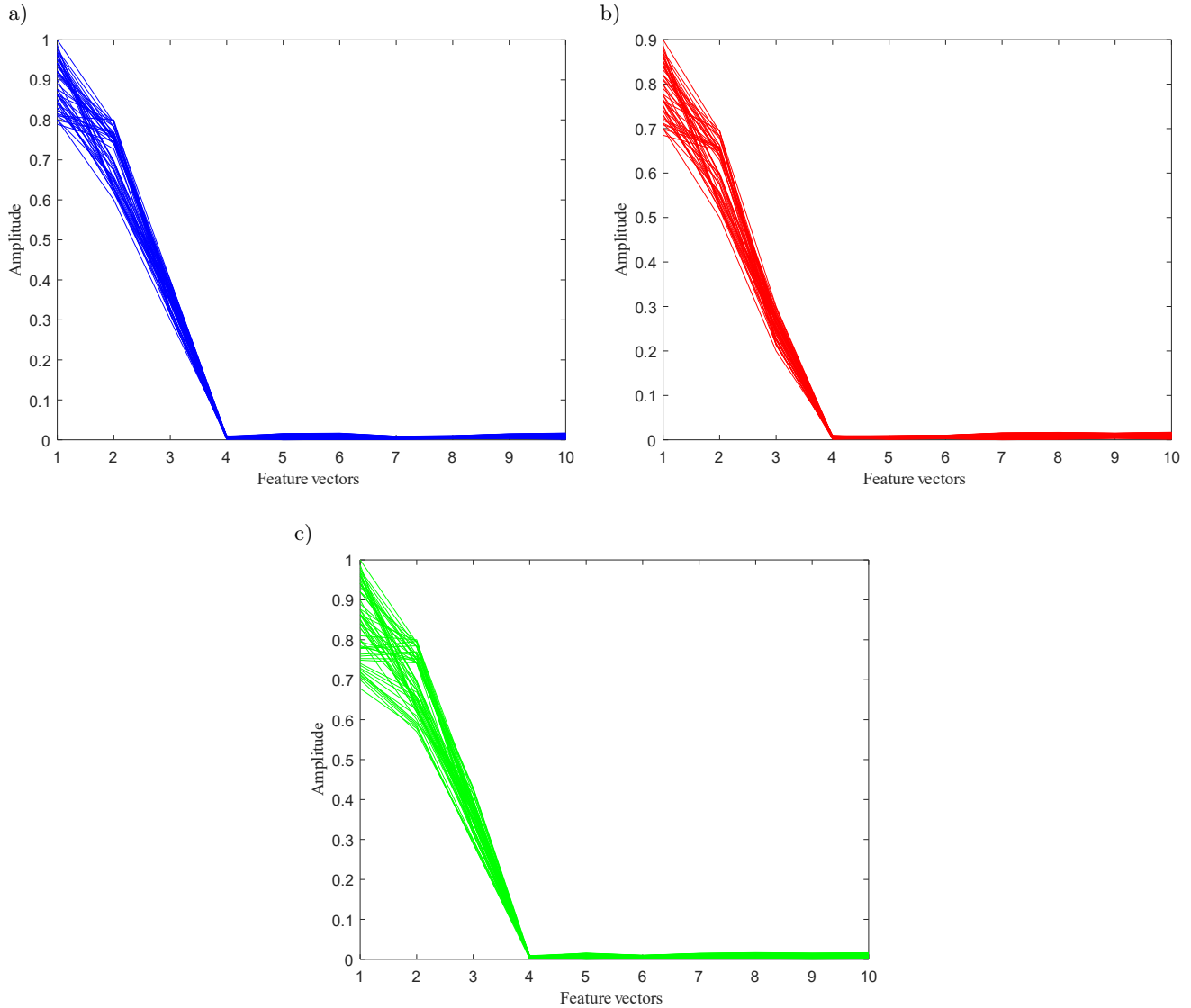


Fig. 6. Feature vectors extracted by FFT: a) feature vectors of earthquake; b) feature vectors of tsunami; c) feature vectors of volcano eruption.

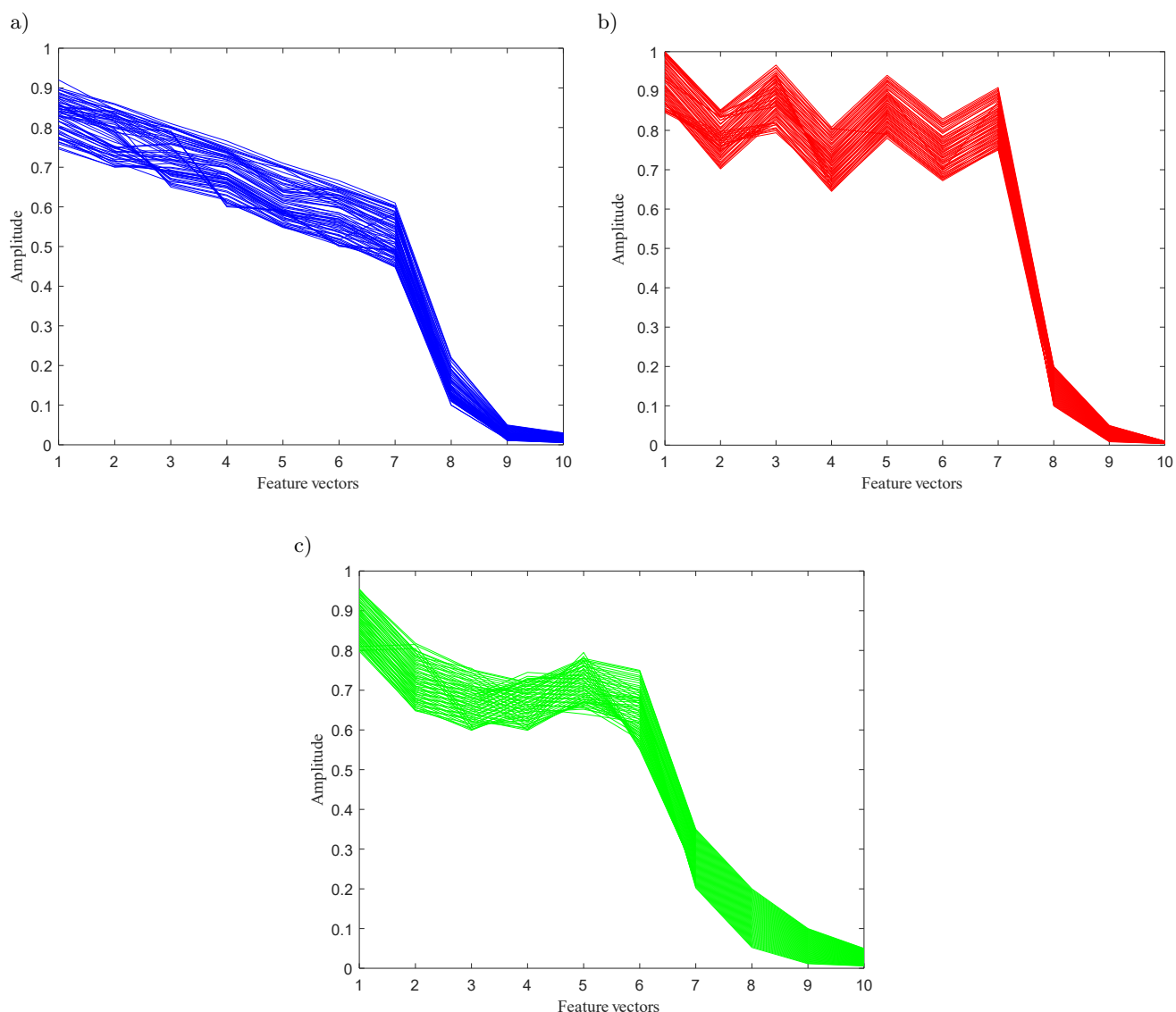


Fig. 7. Feature vectors extracted by ICA and FFT: a) feature vectors of an earthquake; b) feature vectors of a tsunami; c) feature vectors of a volcano eruption.

Table 2. Infrasound classes used for training and testing.

Event type	Class number	Number of vectors	Number of vectors used for training	Number of vectors used for testing
Earthquake	1	204	136	68
Tsunami	2	218	146	72
Volcano	3	189	126	63
Total	–	611	408	203

3.4. Experiment results and discussion

As shown in Fig. 6, the feature vectors of three infrasound events have some similarities. However, there are some differences between them and can be used for infrasound signal classification. Figure 7 shows that the distinction of the feature vectors among different

classes is obvious, and the amplitude of the eigenvalues is large. The results show that the proposed methods can extract feature vectors.

The final classification results are shown in Figs. 8 and 9. In these figures, the anticipated label is represented by the abscissa, whereas the true label is depicted by the ordinate.

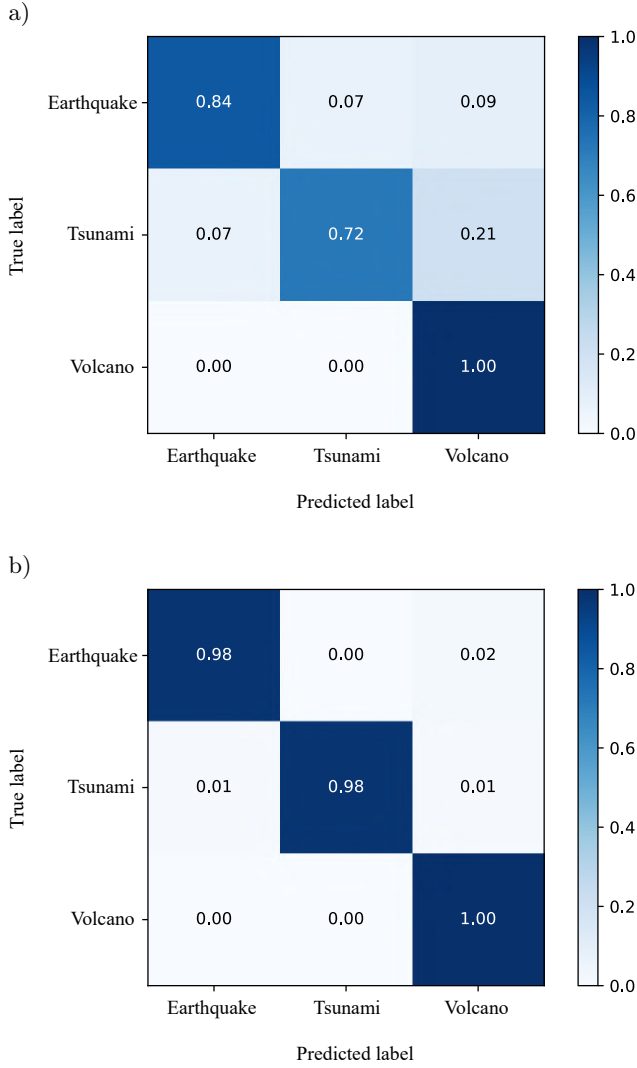


Fig. 8. Confusion matrix of classification results:
a) FFT + SVM; b) ICA + FFT + SVM.

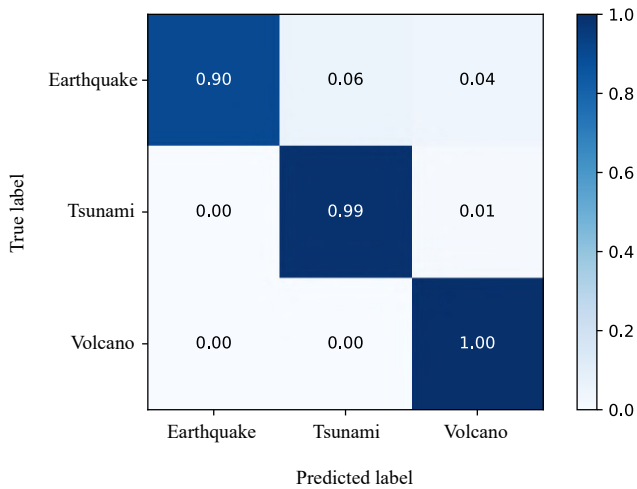


Fig. 9. Confusion matrix of ICA + FFT + KNN classification results.

The confusion matrix of infrasound classification using FFT + SVM is shown in Fig. 8a. The classification accuracy of the three infrasound events is 84.73%. In the classification result of an earthquake by FFT + SVM, five earthquake feature vectors are mistakenly identified as tsunami vectors, and six earthquake feature vectors are mistakenly identified as volcano vectors. In the classification result of a tsunami by FFT + SVM, five tsunami feature vectors are mistakenly identified as earthquake feature vectors, and fifteen tsunami feature vectors are mistakenly identified as volcano feature vectors. Some events are misclassified due to similar characteristics.

The confusion matrix of infrasound classification using ICA + FFT + SVM is depicted in Fig. 8b. The classification accuracy of the three infrasound events is 98.52%, according to the results. In the classification result of an earthquake by ICA + FFT + SVM, one earthquake feature vector is mistakenly identified as a volcano feature vector. In the classification result of a tsunami by ICA + FFT + SVM, one tsunami feature vector is mistakenly identified as an earthquake feature vector, and one tsunami feature vector is mistakenly identified as a volcano feature vector. Some events are misclassified due to similar characteristics.

To compare SVM with other methods, we use the ICA + FFT to extract feature vectors from the same data and then use KNN to classify it, as shown in Fig. 9. A comparative test is utilized to verify the efficiency of the proposed method. Compared with FFT + SVM, the classification accuracy of ICA + FFT + SVM increases by 14% and obtains excellent operating speed, as shown in Table 3. This table shows that the classification result of ICA + FFT + SVM is better than ICA + FFT + KNN, which increases by 2% in accuracy and decreases by 1.5 s in run time. This suggests the SVM method is more suitable for classifying the reduced dimension data by ICA.

Table 3. Comparison results of infrasound signal classification.

Classification scheme	Classification accuracy	Run time [s]
FFT combined with SVM	84.73	5.368
ICA and FFT combined with KNN	96.06	3.625
ICA and FFT combined with SVM	98.52	2.124

As shown in Fig. 3, the source locations of the infrasound stations are distributed widely, but their quantity is small. Due to the limitation of the data, the proposed approach may not be generalized for global hazard monitoring.

4. Conclusion and future work

This research presented a reliable approach for classifying and identifying infrasound signals. ICA separated the source signals of mixed infrasound signals, and then the feature vectors of infrasound events were extracted by FFT. Finally, SVM was used to classify the extracted feature vectors. The experiment results can provide practical solutions for the classification of infrasound signals. The study aimed to improve the accuracy of geophysical monitoring. Due to the limitations of the existing conditions, tests can only use small samples and a few infrasound types, which will affect the reliability of the test results. More infrasound data and infrasonic event types must be evaluated in order to obtain more precise results. For future work, real-time infrasound signal classification will be carried out, and further studies on infrasound types will be performed. Deep learning should be developed for global infrasound signal classification (ALBERT, LINVILLE, 2020).

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 41572347), 2021 Graduate Innovation Fund Project of China University of Geosciences, Beijing (YB2021YC020). The authors would like to thank the Comprehensive Nuclear-Test-Ban Treaty Beijing National Data Center for providing the data.

References

1. ALBERT S., LINVILLE L. (2020), Benchmarking current and emerging approaches to infrasound signal classification, *Seismological Research Letters*, **91**(2A): 921–929, doi: 10.1785/0220190116.
2. AMARNATH M. (2016), Local fault assessment in a helical geared system via sound and vibration parameters using multiclass SVM classifiers, *Archives of Acoustics*, **41**(3): 559–571, doi: 10.1515/aoa-2016-0054.
3. CANNATA A. *et al.* (2011), Clustering and classification of infrasonic events at Mount Etna using pattern recognition techniques, *Geophysical Journal International*, **185**(1): 253–264, doi: 10.1111/j.1365-246X.2011.04951.x.
4. CÁRDENAS-PENÁ D., OROZCO-ALZATE M., CASTELLANOS-DOMÍNGUEZ G. (2013), Selection of time-variant features for earthquake classification at the Nevado-del-Ruiz volcano, *Computers & Geosciences*, **51**: 293–304, doi: 10.1016/j.cageo.2012.08.012.
5. CHERNOGOR L.F., SHEVELEV N.B. (2018), Characteristics of the infrasound signal generated by Chelyabinsk celestial body: Global statistics, *Radio Physics and Radio Astronomy*, **23**(1): 24–35, doi: 10.15407/rpra23.01.024.
6. COOLEY J.W., TUKEY J.W. (1965), An algorithm for the machine calculation of complex Fourier series, *Mathematics of Computation*, **19**(90): 297–301, doi: 10.1090/S0025-5718-1965-0178586-1.
7. CORTES C., VAPNIK V. (1995), Support-vector networks, *Machine Learning*, **20**: 273–297, doi: 10.1007/BF00994018.
8. GI N., BROWN P. (2017), Refinement of bolide characteristics from infrasound measurements, *Planetary and Space Science*, **143**: 169–181, doi: 10.1016/j.pss.2017.04.021.
9. HAM F.M., REKAB K., ACHARYYA R., LEE Y.C. (2008), Infrasound signal classification using parallel RBF Neural Networks, *International Journal of Signal and Imaging Systems Engineering*, **1**(3–4): 155–167, doi: 10.1504/IJSISE.2008.026787.
10. IEZZI A.M., SCHWAIGER H.F., FEE D., HANEY M.M. (2019), Application of an updated atmospheric model to explore volcano infrasound propagation and detection in Alaska, *Journal of Volcanology and Geothermal Research*, **371**: 192–205, doi: 10.1016/j.jvolgeores.2018.03.009.
11. LI M., LIU X.Y., LIU X. (2016), Infrasound signal classification based on spectral entropy and support vector machine, *Applied Acoustics*, **113**: 116–120, doi: 10.1016/j.apacoust.2016.06.019.
12. LIU D., TANG D., ZHANG S., LENG X., HU K., HE L. (2021), Method for feature analysis and intelligent recognition of infrasound signals of soil landslides, *Bulletin of Engineering Geology and the Environment*, **80**: 917–932, doi: 10.1007/s10064-020-01982-w.
13. LIU X.Y., LI M., TANG W., WANG S.C., WU X. (2014), A new classification method of infrasound events using Hilbert-Huang transform and support vector machine, *Mathematical Problems in Engineering*, **2014**(3): 1–6, doi: 10.1155/2014/456818.
14. MAYER S., VAN HERWIJNEN A., ULIVIERI G., SCHWEIZER J. (2020), Evaluating the performance of an operational infrasound avalanche detection system at three locations in the Swiss Alps during two winter seasons, *Cold Regions Science and Technology*, **173**: 102962, doi: 10.1016/j.coldregions.2019.102962.
15. MCKEE K., FEE D., HANEY M., MATOZA R.S., LYONS J. (2018), Infrasound signal detection and back azimuth estimation using ground-coupled airwaves on a seismo-acoustic sensor pair, *Journal of Geophysical Research: Solid Earth*, **123**(8): 6826–6844, doi: 10.1029/2017JB015132.
16. MIKA D., KLECZKOWSKI P. (2011), ICA-based single channel audio separation: new bases and measures of distance, *Archives of Acoustics*, **36**(2): 311–331, doi: 10.2478/v10168-011-0024-x.

17. QIAN G., WANG L., WANG S., DUAN S. (2019), A novel fixed-point algorithm for constrained independent component analysis, *EURASIP Journal on Advances in Signal Processing*, **2019**(1): 28, doi: 10.1186/s13634-019-0622-8.
18. SASTRY A.V., HU A., HECKMANN D., POUDEL S., KAVVAS E., PALSSON B.O. (2021), Independent component analysis recovers consistent regulatory signals from disparate datasets, *PLOS Computational Biology*, **17**(2): e1008647, doi: 10.1371/journal.pcbi.1008647.
19. THÜRING T., SCHOCH M., VAN HERWIJNENA A., SCHWEIZER J. (2015), Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays, *Cold Regions Science and Technology*, **111**: 60–66, doi: 10.1016/j.coldregions.2014.12.014.
20. TSYBUL'SKAYA N.D., KULICHKOV S.N., CHULICHKOV A.I. (2012), Studying possibilities for the classification of infrasonic signals from different sources, *Izvestiya, Atmospheric and Oceanic Physics*, **48**(4): 384–390, doi: 10.1134/S0001433812040147.
21. ZHAO J., LIU Y., YANG J. (2021), 3D matching positioning method for landslide using infrasound signal received by triangular pyramid vector array, based on ray theory, *Bulletin of Engineering Geology and the Environment*, **80**(2): 889–904, doi: 10.1007/s10064-020-01988-4.

Research Paper

Study on the Effectiveness of Monte Carlo Filtering
when Correcting Negative SEA Loss FactorsPaweł NIERADKA^{(1),(2)*}, Andrzej DOBRUCKI⁽¹⁾⁽¹⁾ *Wrocław University of Science and Technology
Department of Acoustics, Multimedia and Signal Processing
Wrocław, Poland*⁽²⁾ *KFB Acoustics, Acoustic Research and Innovation Center
Domasław, Poland*

*Corresponding Author e-mail: p.nieradka@kfb-acoustics.com

(received July 21, 2022; accepted November 18, 2022)

The power injection method (PIM) is an experimental method used to identify the statistical energy analysis (SEA) parameters (called loss factors – LFs) of a vibroacoustic system. By definition, LFs are positive real numbers. However, it is not uncommon to obtain negative LFs during experiments, which is considered a measurement error. To date, a recently proposed method, called Monte Carlo filtering (MCF), of correcting negative coupling loss factors (CLFs) has been validated for systems that meet SEA assumptions. In this article, MCF was validated for point connections and in conditions where SEA assumptions are not met (systems with low modal overlap, non-conservative junctions, strong coupling). The effect of removing MCF bias on the results was also examined. During the experiments, it was observed that the bias is inversely proportional to the damping loss factor of the examined subsystems. The obtained results confirm that the PIM, combined with MCF, allows to determine non-negative SEA parameters in all considered cases.

Keywords: statistical energy analysis; coupling loss factor; Monte Carlo filtering; power injection method.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) <https://creativecommons.org/licenses/by-sa/4.0/> which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Statistical energy analysis (SEA) is a numerical method that allows the behavior of vibroacoustic systems to be modeled using a system of linear equations at medium and high frequencies (LYON, DEJONG, 1995). SEA can deal with problems where acoustic and vibrational fields are coupled. SEA is widely used in the automotive industry (noise, vibration, and harshness – NVH) (CHEN *et al.*, 2012) and also when designing silent ships (HATTORI *et al.*, 1985), trains (JI *et al.*, 2015), airplanes (BORELLO, 2018), and buildings (Craik, 1982). Other applications include: predicting the insertion loss of sound-insulating enclosures (NIERADKA, DOBRUCKI, 2018) and machinery noise (LALOR, 1996), estimating the sound reduction index of partitions (PRICE, CROCKER, 1970) and noise in kitchen appliances (ZARATE *et al.*, 2017), damage detection in joints (PANKAJ, 2019), predicting the in-

sertion loss of lagging in ducts (YOGANANDH *et al.*, 2019), and noise transmission in the spacecraft industry (HWANG, 2002).

The vibroacoustic system is fully identified with regard to SEA when parameters called loss factors (LFs) are known. LFs divide into coupling loss factors (CLFs) and damping loss factors (DLFs). Experimental SEA (E-SEA) is a set of methods allowing LFs to be determined using measurements. The most widely used E-SEA method is the power injection method (PIM) (BIES, HAMID, 1980), which is utilized in this paper. Other methods can also be labeled as E-SEA, e.g., identification based on intensity measurements (MING, 1998), mobility measurements (CACCIOLATI, GUYADER, 1994), input power modulation (FAHY, RUIVO, 1997), and the energy ratio method (GU, SHENG, 2015).

Post-processing E-SEA data can occasionally produce negative LFs, which are considered measurement

errors. Two reasons for the occurrence of negative LFs can be distinguished: measurement uncertainty and the failure to meet SEA assumptions. SEA assumptions are widely discussed in the literature (LAFONT *et al.*, 2014) and are shortly summarized in this article. In order to correct negative LFs, the Monte Carlo filtering method (MCF) can be used (DE LAS HERAS *et al.*, 2020). In MCF, a statistical ensemble of energy matrices is generated based on the mean value and variance of the experimental data. Only those members of the ensemble that give correct results are used in the final calculations. HOPKINS (2002) conducted virtual E-SEA experiments, where he used an alternative approach to MCF. Instead of building a statistical ensemble based on measurement variance, he was varying system dimensions and physical properties, which in turn allowed for positive LFs to be obtained for most members of the ensemble. This approach is possible and robust due to the fact that virtual experiments in the FEM framework allowed for the system dimensions and properties to be quickly alternated. LALOR (1990) proposed an alternative formulation of PIM to obtain lower matrix condition numbers. HODGES *et al.* (1987) suggested using matrix fitting methods to correct negative CLFs.

In this work, experiments are performed on steel plates connected using different structural junctions at different levels of damping. Similar structures have been investigated by others, but the research was mainly focused on the values of coupling loss factors. The dependence of CLFs on the type of joint and the type of material was investigated by MANDALE *et al.* (2016), BHAGWAN, POPURI (2019), and others. The dependence of the CLF value on the plate's thickness ratio and the number of points in point junctions has also been experimentally verified (PANUSZKA *et al.*, 2005). Negative CLFs were not considered in those publications. The effectiveness of MCF was experimentally proved for a plate system suspended in a room (acoustics system) and numerically validated for plate ensembles (DE LAS HERAS *et al.*, 2020). This verification only included the effect of measurement uncertainty because it was performed in conditions where SEA assumptions were met.

The authors of the present publication are unaware of any article where MCF was employed for point junctions and in conditions that violate SEA requirements (low modal overlap, non-conservative junctions). Conditions where SEA assumptions are violated are often encountered in industrial and other real-life scenarios, and it is therefore desirable to know how the MCF method will behave in such situations. Assessing the effectiveness of MCF in non-ideal conditions is the main research problem tackled in this publication.

The paper is organized as follows. Section 2 gives a brief introduction to E-SEA, MCF, and SEA assumptions. Section 3 presents the tested systems and de-

scribes the course of the measurements. In Sec. 4, the relationships between the results of the MCF procedure and the SEA assumptions are presented for different joints and damping levels. Moreover, the effect of MCF bias on the obtained LFs is investigated. Section 5 summarizes the major findings of the study.

2. Foundations of the applied methods

2.1. Experimental statistical energy analysis

In the SEA approach, the structure (system) is subdivided into a set of subsystems. A subsystem is defined as a group of similar modes of vibration. The structure should be physically divided into parts in places with high impedance mismatches (branches, discontinuities, junctions, changes in material type, and thickness) in order to obtain weak coupling conditions. Both acoustic cavities and elastic solids can be treated as a subsystem.

In SEA, the vibroacoustic system is fully described when the set of LFs parameters are known. LFs are subdivided into DLFs and CLFs. DLFs describe the inner energy dissipation of subsystems, while CLFs describe the energy flow between subsystems.

SEA is based on writing down a set of equations that describe the energy balance between subsystems. Figure 1 shows an example of the energy balance between two subsystems – subsystem 1 with mean energy E_1 and subsystem 2 with mean energy E_2 . Furthermore, η_{11} and η_{22} are damping loss factors, η_{12} and η_{21} are coupling loss factors, and P_{in1} and P_{in2} are input powers.

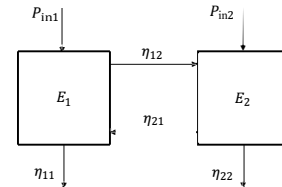


Fig. 1. Energy balance between two SEA subsystems.

The general SEA equation for a system comprising N subsystems can be written down using matrix notation (LYON, DEJONG, 1995):

$$\{P\} = \omega [L] \{E\}, \quad (1)_1$$

where $\{P\}$ is a column vector of input powers, $\{E\}$ is a column vector of the (unknown) subsystems' energies, ω is the center angular frequency of the analyzed band $\Delta\omega$ (i.e., octave or 1/3-octave band), and $[L]$ is the loss factor matrix.

The loss factor matrix $[L]$ is constructed as follows:

$$L_{ij} = \begin{cases} \sum_{u=1}^N \eta_{iu} & \text{if } i = j, \\ -\eta_{ji} & \text{if } i \neq j, \end{cases} \quad (1)_2$$

where $i = 1, \dots, N$, $j = 1, \dots, N$.

When LFs and input powers are known, determining the unknown energies of subsystems is straightforward, as can be seen in Eq. (1). For complex junctions and structures, LFs are often unknown and need to be identified.

Experimental SEA is a set of methods allowing to determine the LFs of a vibroacoustic system. In this work, we focus on a method derived directly from the energy balance equations. Equation (1) can be rewritten and modified so that the LFs (matrix $[L]$) become unknowns:

$$[L] = [P][E]^{-1}/\omega, \quad (2)$$

where $[P]$ is the input power matrix and $[E]$ is the energy matrix. This is called the power injection method (BIES, HAMID, 1980). During the experiments, input powers and the subsystems' energies have to be obtained. Each subsystem is excited separately (the power is injected), and the responses are measured simultaneously on every subsystem in order to match the number of equations with the number of unknowns. The consequence is that vectors $\{E\}$ and $\{P\}$ in Eq. (1) become matrices $[E]$ and $[P]$ in Eq. (2). $[E]$ matrix entries E_{ij} stands for the energy of subsystem i when subsystem j is excited. $[P]$ is a diagonal matrix with entries P_{jj} , which stands for the power injected into subsystem j .

In order to obtain LFs, the energy matrix needs to be inverted. Since the input force has a form of the Dirac impulse in the time domain, actual powers and energies as in Eqs. (1) and (2) cannot be utilized. Instead, equivalent input powers and energies (normalized to force autospectrum) in frequency band $\Delta\omega = \omega_2 - \omega_1$ can be determined based on transfer function measurements as follows:

$$P_{\Delta\omega, eq} = \frac{1}{\omega_0} \text{Im} \left[\int_{\omega_1}^{\omega_2} H_{af}(\omega) d\omega \right], \quad (3)$$

$$E_{\Delta\omega, eq} = M \frac{1}{\omega_0^2} \int_{\omega_1}^{\omega_2} |H_{af}(\omega)|^2 d\omega, \quad (4)$$

where ω_0 is the center frequency of $\Delta\omega$, H_{af} is the transfer function between acceleration and the input force, and M is the mass of the subsystem. From Eqs. (3) and (4), it is only possible to determine the effective powers and energies that are normalized to the force autospectrum and not the actual powers and energies. Note that the imaginary part and the division by ω are used to compute the input power. The reason is that during measurement, acceleration is obtained, but in calculations, velocity is used. A transfer function H_{af} in Eq. (3) is taken between the force and acceleration in the driving (excitation) point. In Eq. (4), acceleration is taken from random points in the structure away from the driving point and subsystem boundaries (in order to avoid contributions of direct and evanescent waves).

It is possible to further normalize energy matrix $[E]$ in Eq. (2) in terms of input power in order to obtain normalized energy matrix $[G] = \omega[E][P]^{-1}$. It was shown that this could lead to a better matrix condition number (DE LAS HERAS *et al.*, 2018). The formula can now be written as:

$$[I] = [L][G], \quad (5)$$

where $[I]$ is an identity matrix. The variant of PIM that uses the $[G]$ matrix is known in the literature as normalized energy matrix inversion (NEMI). The loss factor matrix $[L]$ in NEMI is computed by inverting $[G]$:

$$[L] = [G]^{-1}. \quad (6)$$

After the inversion, LFs can be directly extracted from $[L]$ as follows: 1) coupling loss factor η_{ij} is off-diagonal term L_{ji} multiplied by -1 , and 2) damping loss factor η_{ii} is obtained by summing the elements of column i .

The procedure described above needs to be repeated for each frequency band of interest.

2.2. Monte Carlo filtering

The inverse energy matrix obtained from measurements will not always provide only positive LFs. SEA matrices are very sensitive to measurement uncertainties (DE LAS HERAS, 2018). One of the ways to solve this is to use MCF, which was proposed by DE LAS HERAS (2020). A similar approach, proposed by BOUHAIJ (2017), was utilized to assess CLFs uncertainties. The general idea of MCF is presented in Fig. 2.

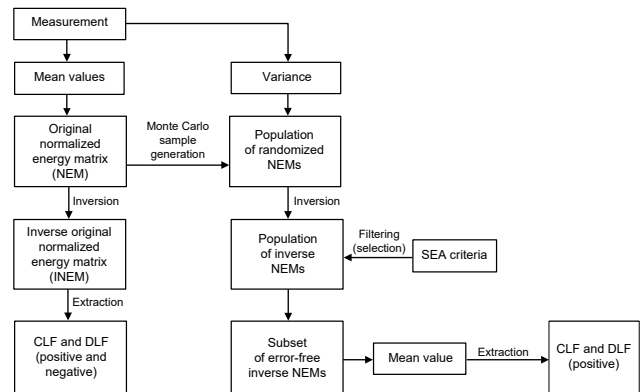


Fig. 2. Monte Carlo filtering procedure.

The classical E-SEA approach (as described in Subsec. 2.1) is presented in the first column of the graph in Fig. 2. In the MCF approach, the population of randomized NEMs is generated by using the original NEM and the measurement variance. The population is then filtered in order to remove all matrices that will produce negative LFs. The final step is to compute the mean value of the obtained LFs. More details

on using this method can be found in reference articles (DE LAS HERAS *et al.*, 2020; BOUHAJ *et al.*, 2017).

This article provides an extended MCF validation for a wide range of technical junctions. The validation includes the following steps: 1) the influence of violating SEA assumptions on MCF results, 2) the influence of the sample bias on MCF results.

2.3. SEA assumptions

The SEA assumptions are as follows (LE BOT, 2015): 1) a high modal overlap factor; 2) weak coupling between subsystems; 3) non-conservative junctions; 4) a high number of modes in frequency bands; 5) a diffuse field; 6) uncorrelated white noise excitations; 7) statistically independent modes; 8) negligible non-resonant transmission.

In this publication, we mainly focus on the first three assumptions. This approach is justified due to the fact that assumption 1 is often referred to as the most crucial requirement (CULLA, SESTIERI, 2006).

Assumption 1 states that:

$$\mu \gg 1, \quad (7)$$

where μ is the modal overlap factor, which is defined as:

$$\mu = n\omega\eta, \quad (8)$$

where n is the modal density, ω is the angular frequency, and η is the DLF. Modal density can be estimated by equation (LE BOT, 2015):

$$n = \frac{2M\langle \text{Re}\{Y\} \rangle_{A,\omega}}{\pi}, \quad (9)$$

where M is the mass of the subsystem, Y is the driving point mobility, and $\langle \dots \rangle_{A,\omega}$ stands for space and frequency averaging. Equation (9) is useful because variables from the right side of the equation can be obtained from modal hammer measurements performed during PIM. A high modal overlap results in a smooth frequency response function, which is not dominated by any resonance. This is crucial because SEA was derived with the assumption of the equipartition of modal energy, where each resonant mode is of equal importance when computing energy flow.

Smith's criterion (SMITH, 1979) can be used to check assumption 2 and states that:

$$\gamma \ll 1, \quad (10)$$

where γ is called the coupling strength indicator, which can be computed by taking the ratio of CLFs to DLFs. In this study, Smith's criterion is used, but it must be noted that other criteria for coupling strength can be found in the literature (FINNVEDEN, 2011).

Non-conservative junctions are an idealization that never exists in practice. One can say that assumption 3

is always violated to some extent because even welding junctions introduce some damping to the system. In practice, this effect is neglected, apart from when the losses associated with the joint are comparable to the losses of the entire system. For example, a joint that does not meet assumption 3 may be the 6 mm thick rubber pads that were used in the experiments.

Assumption 4 can be written as:

$$N_{\Delta\omega} \gg 1, \quad (11)$$

where $N_{\Delta\omega}$ is the number of modes in frequency band $\Delta\omega$, and $N_{\Delta\omega}$ can be counted directly when modes are distinguishable. This is usually not the case for high frequencies, and one can use the approximate formula:

$$N_{\Delta\omega} = n\Delta\omega. \quad (12)$$

Assumption 4 in the present research is fulfilled for all systems.

Assumption 5, in some sense, stands in opposition to assumptions 1 and 2. In assumptions 1 and 2, a high damping loss factor is desirable, but in assumption 5, a low damping loss factor is desirable in order to ensure diffuse field conditions. The following requirement for the normalized attenuation factor \bar{m} can be checked to test diffuse field conditions:

$$\bar{m} \ll 1. \quad (13)$$

The normalized attenuation factor is computed by using the equation:

$$\bar{m} = \frac{\eta\omega\bar{l}}{c_g}, \quad (14)$$

where c_g is the group speed, and \bar{l} is the mean-free-path, which for 2D subsystems is equal to

$$\bar{l} = \pi \frac{A}{S}, \quad (15)$$

where A and S are the surface area and perimeter of the subsystem, respectively.

Inequality (Eq. (13)) assures that rays will experience few reflections before being attenuated by the internal damping mechanisms of the subsystem. Assumption 5 is violated for all the considered high-damped systems above 2500 Hz.

Assumption 6 is always fulfilled in this research by using modal hammer impulses that provide a wideband uniform spectrum in the frequency range of interest.

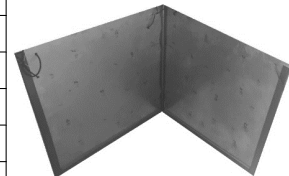
Assumption 7 can be fulfilled by using rain-on-the-roof excitation. In the present research, this excitation is reproduced by performing spatial averaging for point excitations, which is standard practice (CIMERMAN *et al.*, 1997).

Assumption 8, in practice, could be violated in some circumstances (e.g., highly damped systems, mass law transmission), but this effect is not further analyzed here.

3. Experiments

The investigated systems were steel plates connected to each other at a right angle. The mechanical and geometrical parameters of the plates are shown in Table 1. The studied objects contained holes that were made for free suspension. The diameter of the holes was 23 times smaller than the smallest considered flexural wavelength in the system (7 cm), so it can therefore be assumed that the influence of the holes is negligible in the frequency range of interest.

Table 1. Geometric and mechanical parameters of the tested structure.

Geometry		
Thickness	2 mm	
Length	0.49 m	
Width	0.49 m	
Mechanical parameters		
Material	Steel DC03	
Density	7827 kg/m ³	
Young's modulus	205 GPa	
Poisson number	0.3	

The plates were measured for three damping levels (low, medium, and high) in order to vary the modal overlap factor. Rubber magnetic tape was used to introduce damping to the plates (Fig. 3).

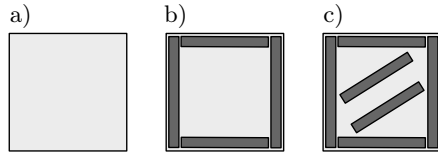


Fig. 3. Tested plate with: a) low damping; b) medium damping; c) high damping.

The plates were connected by different structural junctions (Fig. 4). The line junctions included a line weld junction and a rubber connection. The point junctions used in this research were: point welding, bolts, and rivets. The point junctions consisted of three points evenly distributed along the common edge of the plates.

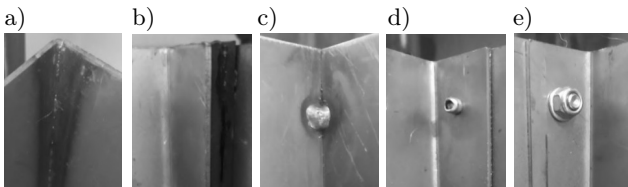


Fig. 4. Technical junctions used to connect the plates: a) line welding; b) rubber connection; c) point welding; d) rivets; e) bolts.

The panels were hung freely on elastic lines. The modal hammer PCB 086C03 was used to excite the struc-

tures, the PCB 356A03 accelerometer was used to measure the source subsystem response in the driving point, and PCB T356A32 accelerometers were used to measure the source and receiver subsystem response away from the excitation points. HEAD acoustics SQuadriga II was used for data acquisition. The sensors were bonded to the structure using wax.

For averaging purposes, three excitation and six response points were randomly chosen on each subsystem. Only the “z” component of acceleration was registered during the experiments (this research focuses on a bending wave transmission, so only out-of-plane wave motion was crucial).

The post-processing script was written in the Python programming language. Based on the obtained signals from the sensors, the input powers and energies were determined using Eqs. (3) and (4). Then, the procedure described in Subsec. 2.1 was followed to obtain the LFs. Some loss factors were negative, which justified the use of the MCF described in Subsec. 2.2.

4. Results and discussion

4.1. Correction of negative CLFs

In this section, a detailed analysis is conducted on the impact of meeting SEA conditions on the effectiveness of the MCF method. A logarithmic scale was used to present the data in graphs, and it was therefore not possible to represent the actual values of negative numbers. Instead, negative CLFs are represented symbolically with gray markers as the mean value of adjacent positive CLFs. The graphs associated with the SEA assumptions also appear in this section. The gray areas appearing on these graphs indicate the frequency ranges where the SEA assumptions are not met. Gray areas indicate regions where $\gamma < 1$ and $\mu > 1$ (we set $\gamma = 1$ and $\mu = 1$ as threshold values). Nevertheless, one must keep in mind that the requirements are $\gamma \ll 1$ and $\mu \gg 1$, and other choices of threshold values are possible (e.g., $\gamma = 0.1$ and $\mu = 10$). In order to avoid complications in the analysis, only the direction from subsystem 1 to subsystem 2 is considered. Loss factors η_{21} and η_{22} are not shown in the plots. This is a simplification because some examined junctions are not fully symmetrical. Complete results, including the LFs for the opposite direction, are included in the summary table in Subsec. 4.1.4.

4.1.1. Line welding

In this section, the results obtained for the line welding junction are analyzed. Figures 5, 6, and 7 show the relation between the sign of the obtained loss factors and SEA assumptions for low, medium, and high damping, respectively. Figure 8 shows the identification results for all the damping levels.

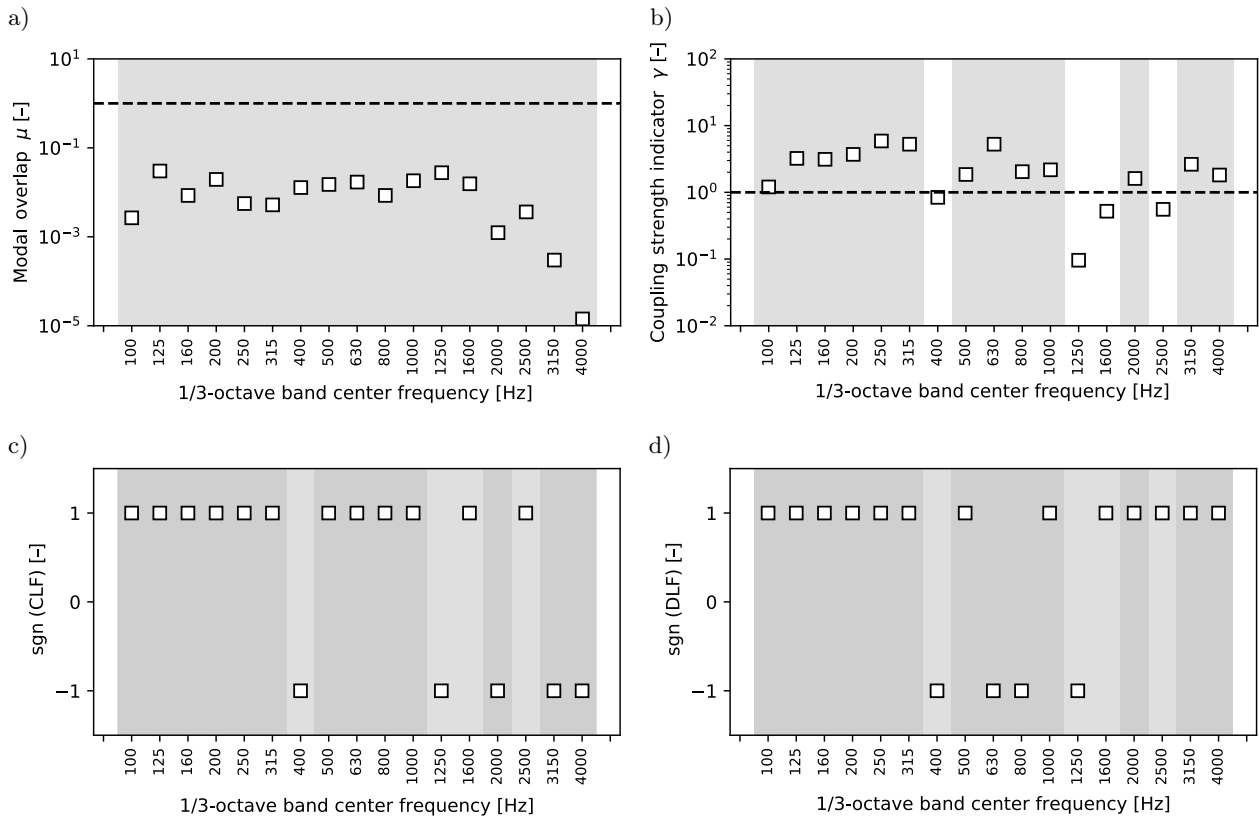


Fig. 5. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the line welding at low damping.

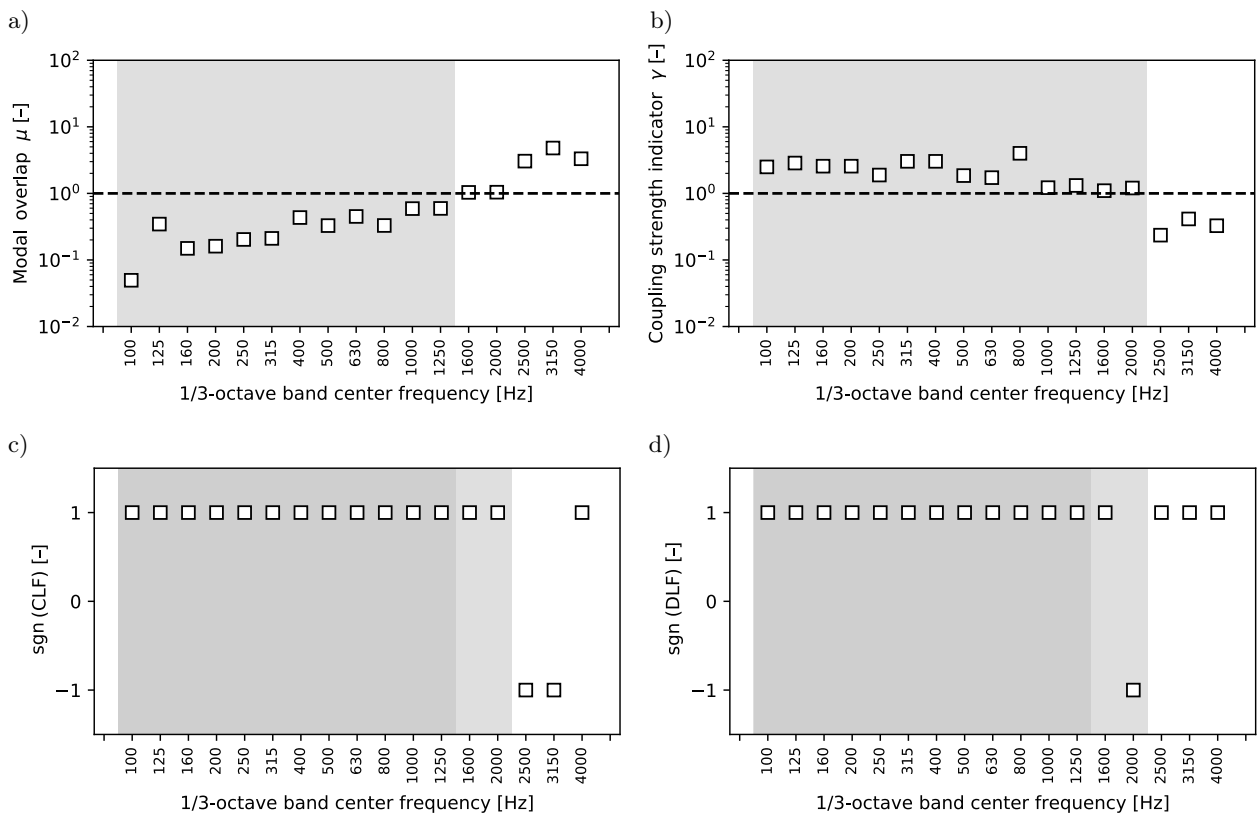


Fig. 6. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the line welding at medium damping.

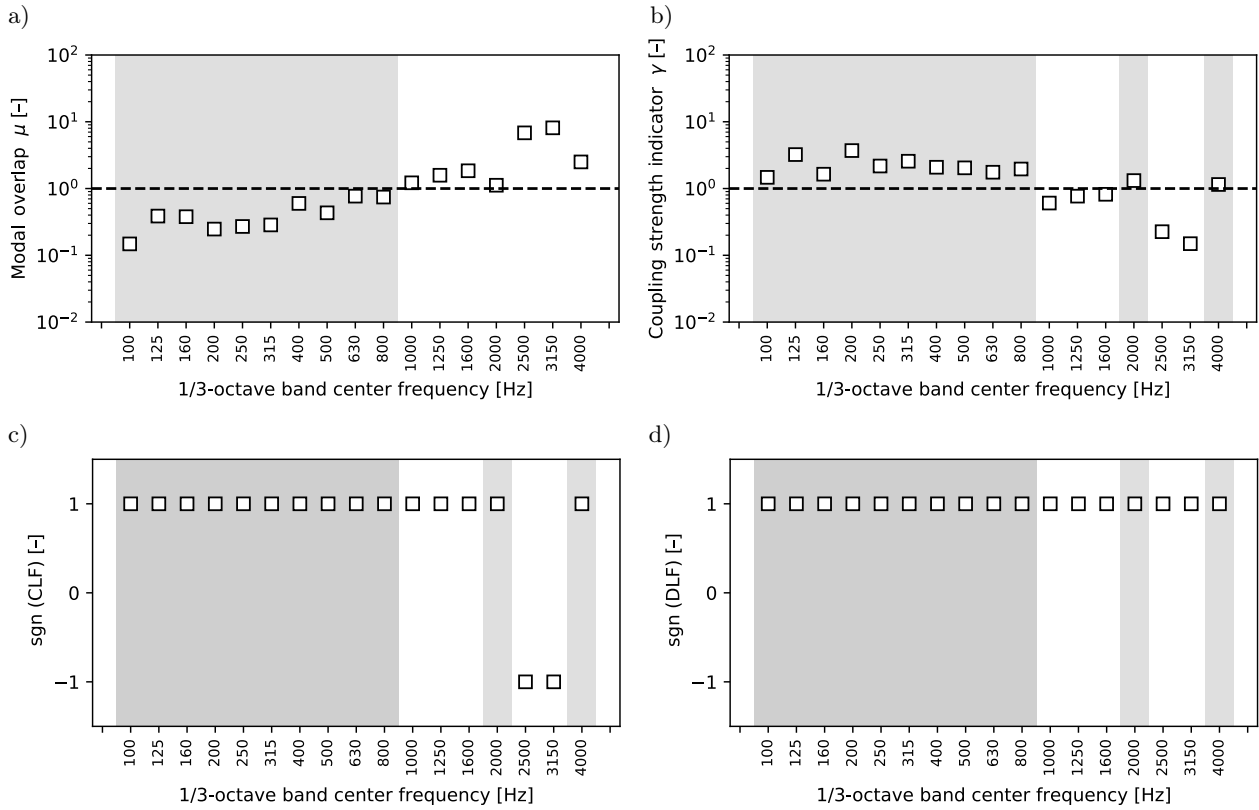


Fig. 7. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the line welding at high damping.

One can notice that as the damping increases, the requirement for the modal overlap is fulfilled in a wider frequency range (Figs. 5a, 6a, 7a). Nevertheless, the requirement for coupling strength is only met in a few bands (Figs. 5b, 6b, 7b). The line weld connects the plates continuously along the common line, which results in high transmission, in turn creating strong coupling conditions (even for the high damped system – Fig. 7b). This allowed (for medium and high damping) the frequency region to be isolated where only one SEA assumption is violated, namely coupling strength.

Negative LFs were obtained for the systems with low, medium, and high damping (Figs. 8a, 8c). All negative LFs were successfully corrected (Figs. 8b, 8d). The biggest amount of negative LFs occurred for the low damped system in the frequency region with a low modal overlap and high coupling strength indicator (Figs. 5c, 5d). This case is difficult to analyze because many unfavorable factors act simultaneously. For the medium and highly damped structures, a few negative LFs were determined within the frequency range for which the SEA assumptions are met (Figs. 6c, 7c). A single DLF was also designated within the frequency region in which the coupling is strong and the modal overlap is high (Fig. 6d). In this case, obtaining negative LFs could result from the measurement uncertainty or strong coupling.

By comparing Fig. 8a with 8b, and 8c with 8d, it can be seen that the MCF procedure used for the low damped system significantly altered the CLF and DLF curves. In turn, for the medium and highly damped system, the MCF provided a smoothing effect while at the same time keeping the values similar.

It is interesting to note that the CLF for the low damping is lower than the CLFs obtained for the medium and high damping. This can be counterintuitive because, in classical SEA, CLFs are independent of DLFs. A lack of dependence on DLFs is assumed in theoretical SEA when CLFs are derived based on wave theory. However, it was shown that CLFs derived using a more precise modal approach do indeed depend on DLFs. This can be observed in the frequency region, where the DLF approaches zero (YAP, WOODHOUSE, 1996).

4.1.2. Rubber connection

In this section, the results obtained for the rubber junction are analyzed. Figures 9, 10, and 11 show the relation between the sign of the obtained loss factor and SEA assumptions for the low, medium, and high damping, respectively. Figure 12 shows the identification results for all the damping levels.

Rubber constitutes a non-conservative junction, so the connection itself violates SEA assumptions. On the

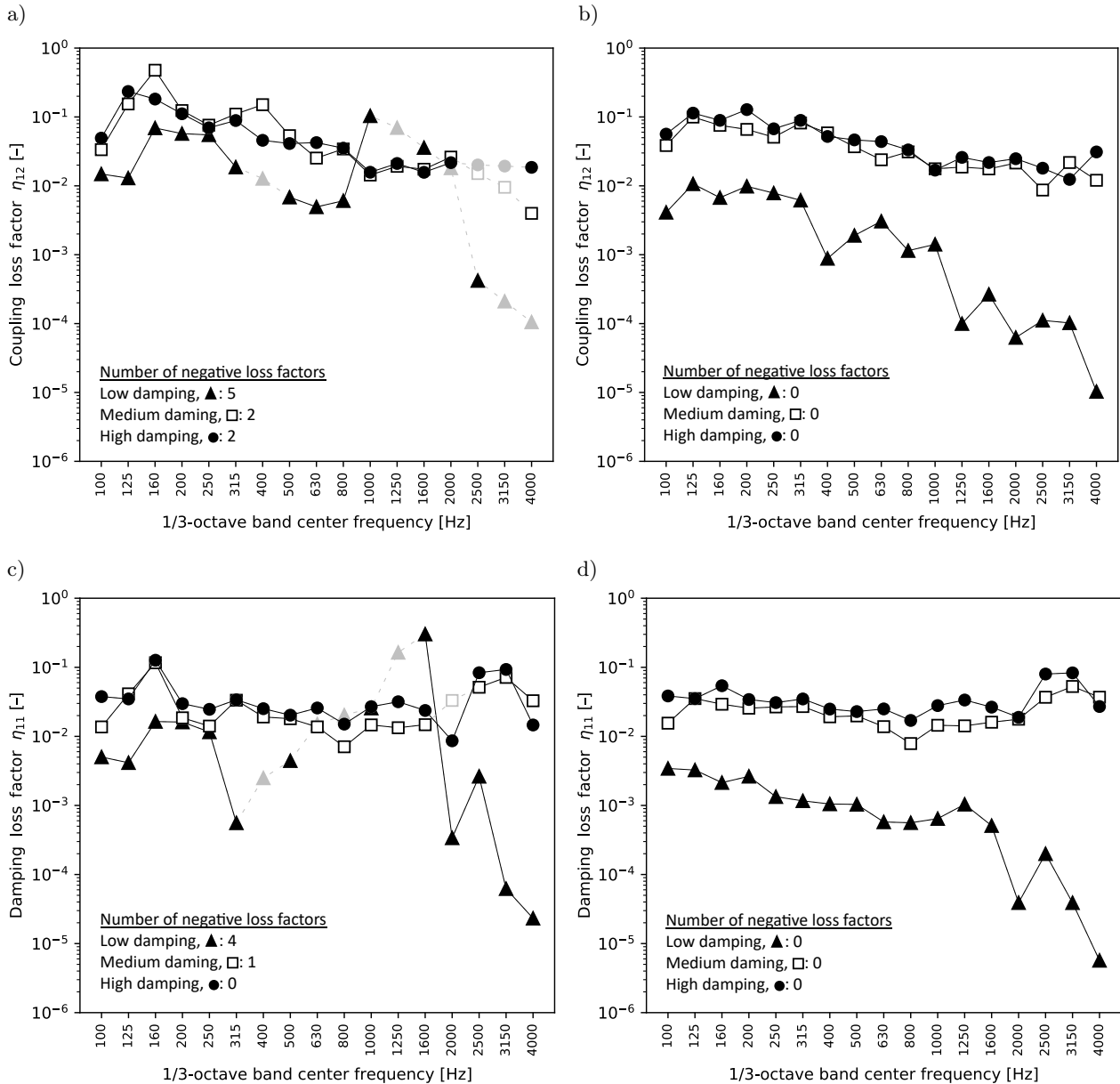


Fig. 8. Identification results for the line welding at all damping levels: a) CLFs without MCF; b) CLF with MCF; c) DLFs without MCF; d) DLF with MCF.

other hand, requirements for the modal overlap and coupling strength were much easier to meet for the rubber junction when compared to the line welding. A non-conservative junction (dissipation of energy in the rubber) resulted in an apparent DLF increase, which helped to obtain a high modal overlap μ . For low damping, μ was higher when compared to the line welding case but still below unity for all the frequency bands (Fig. 9a). For high damping, μ was larger than unity above 315 Hz (in the case of line welding, μ was larger than unity above 800 Hz), as shown in Fig. 11a. The weak coupling condition was easy to obtain thanks to the large compliance of the rubber (Figs. 9b, 10b, 11b), which allowed the frequency range in which only one requirement was not met to be iso-

lated, namely the conservative junction requirement. Nevertheless, no negative LF was determined in this frequency range.

Negative LFs were only identified for the low damped system (Figs. 12a, 12c), and were successfully corrected (Figs. 12b, 12d). Negative values lie in the region where the condition for coupling dissipation is not met (for visual purposes, in Figs. 9c and 9d, this region is not grayed out) and the modal overlap is below unity. Thus, the obtained negative LFs are probably the effect of the non-conservative junction, measurement uncertainty, or low μ .

By comparing Fig. 12c with 12d, it can be seen that the MCF introduced a slight smoothing effect on the obtained DLFs.

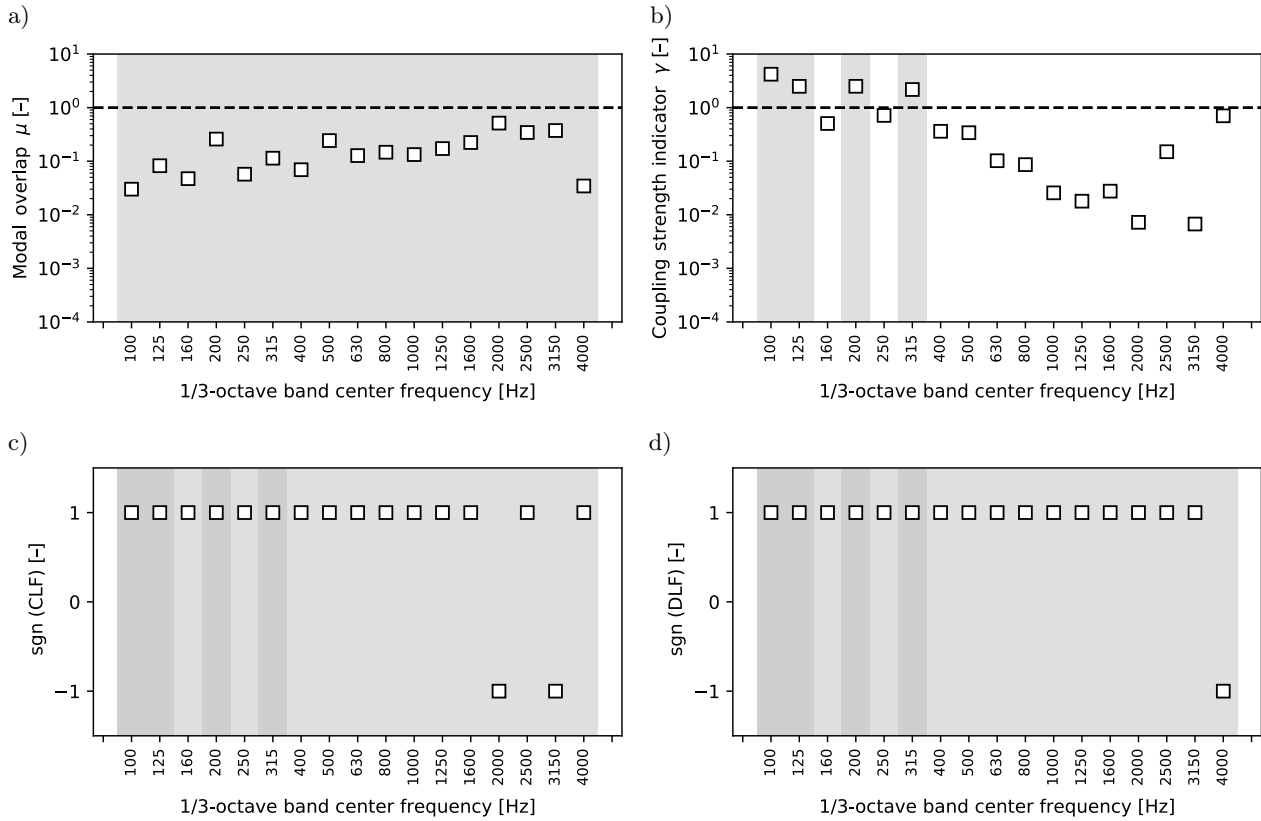


Fig. 9. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the rubber connection at low damping.

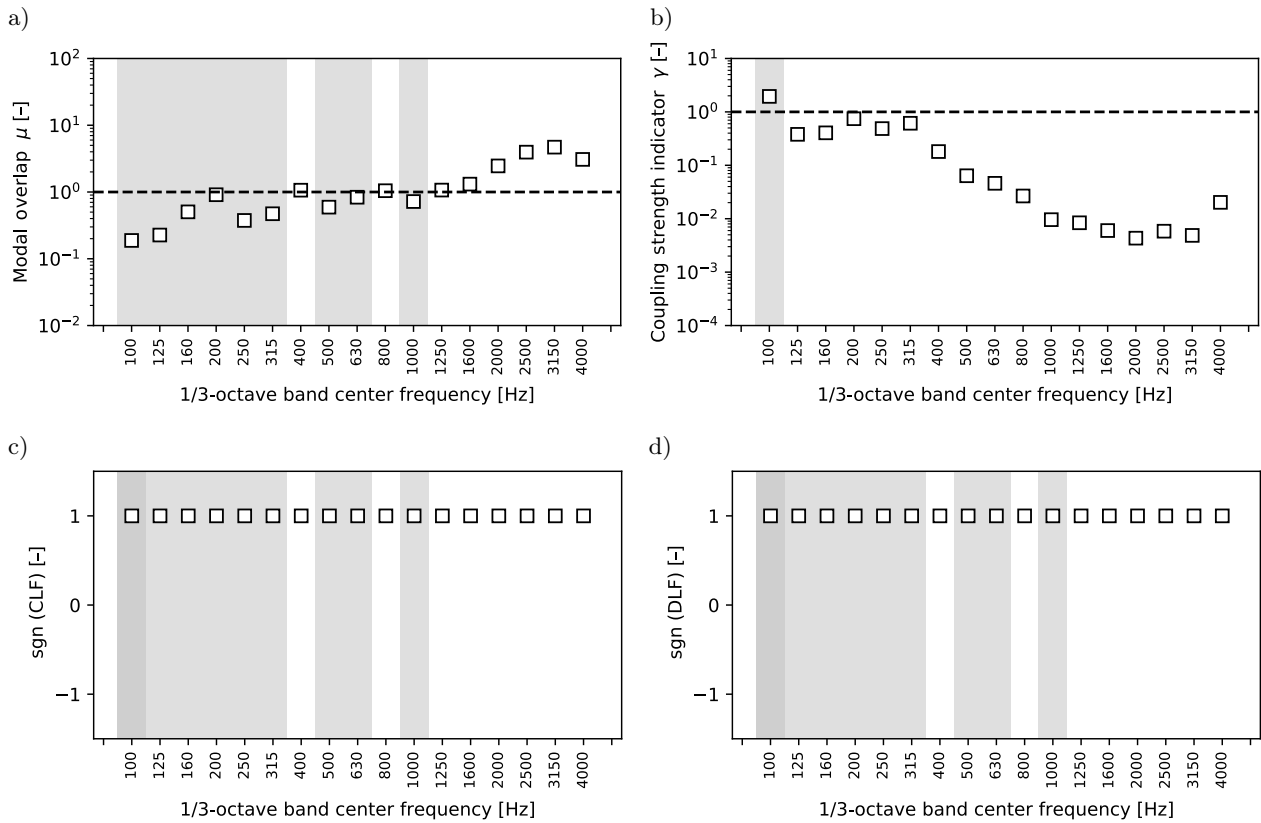


Fig. 10. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the rubber connection at medium damping.

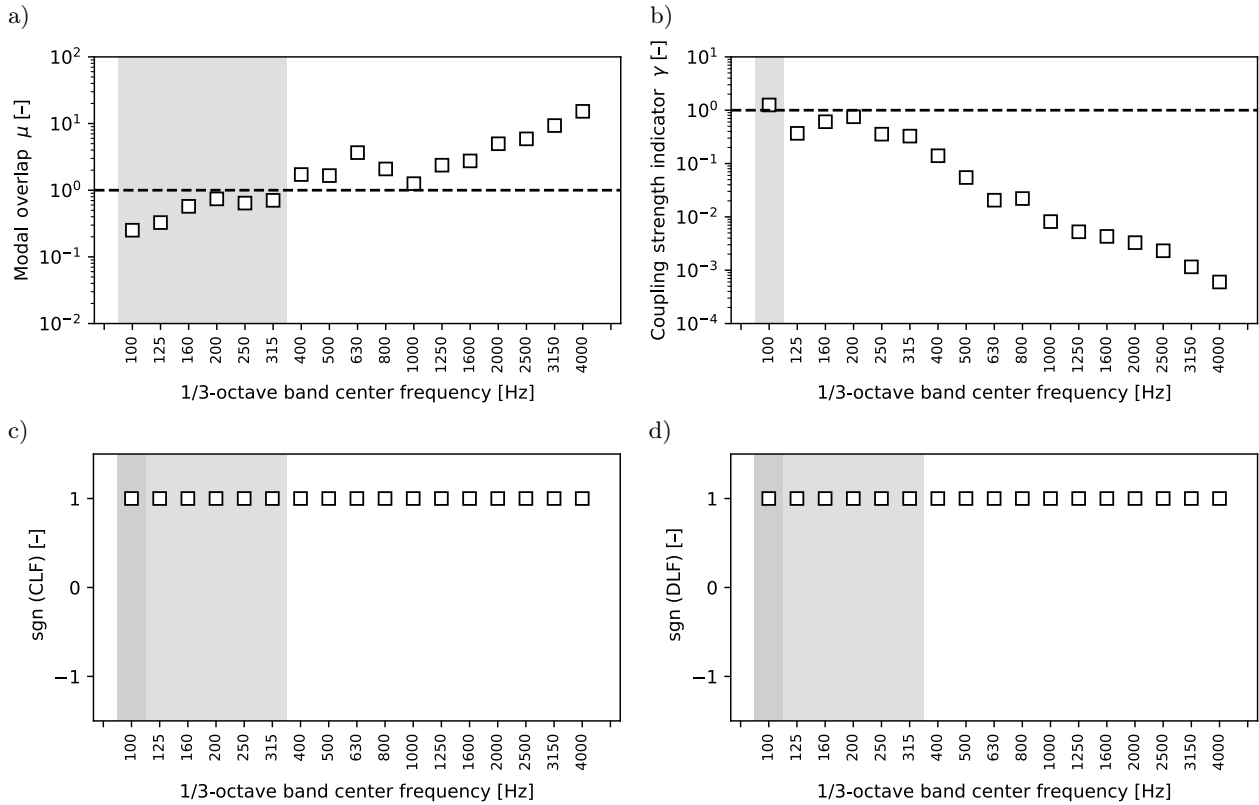


Fig. 11. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the rubber connection at high damping.

The CLFs obtained for the rubber junctions were the smallest ones in this research. Nevertheless, the usefulness of this kind of junction can be seen to be limited due to the fact that much stiffer joints are required in many technical applications.

4.1.3. Point connections

In this section, the results obtained for the point junctions are analyzed. Figures 13, 14, and 15 show the relation between the sign of the obtained loss factor and SEA assumptions for the point welding, rivets, and bolts, respectively. Figure 16 shows the identification results for all the examined point connections at high damping.

The point connections were approximately non-conservative and provided a weak coupling in a wide frequency range (Figs. 13b, 14b, 15b). Weak coupling was easier to obtain when compared to the line junction cases because the systems were only connected at 3 points. Additionally, all the systems with point connections were highly damped, which resulted in a high modal overlap (Figs. 13a, 14a, 15a). Such conditions seem to be ideal in the case of SEA.

However, even under favorable conditions, a single negative CLF was observed for the system with point welding at 2500 Hz (Fig. 16a, gray arrows). This point lies in the region where all the SEA assumptions are met (Fig. 13c), except for assumption 5.

Thus, the error is probably connected with measurement uncertainty or a lack of diffuse field conditions. As shown in Fig. 16b, the negative value was successfully corrected.

By comparing Fig. 16a with 16b, and 16c with 16d, one can state that the MCF did not introduce much distortion in the final results.

The CLF values for all the point connections are similar, but some differences can be noticed. Figure 16b shows that bolted junctions provide the highest vibration transmission in the low-frequency range. Conversely, the biggest vibration reduction in a wide frequency range can be obtained with rivets.

4.1.4. Results in the opposite directions

Subsections 4.1.1, 4.1.2, and 4.1.3 focused on the LFs obtained for the direction from subsystem 1 to subsystem 2. In general, the tendencies described in the previous subsections also apply to the 2–1 direction, even though some systems were not ideally symmetrical. The CLFs obtained for both directions were similar (Fig. 17). The total number of corrected LFs (including both directions) is summarized in Table 2.

4.2. Influence of the MCF bias

In Fig. 18, the result of the MCF simulation is shown. Each dot in the figure represents a randomized

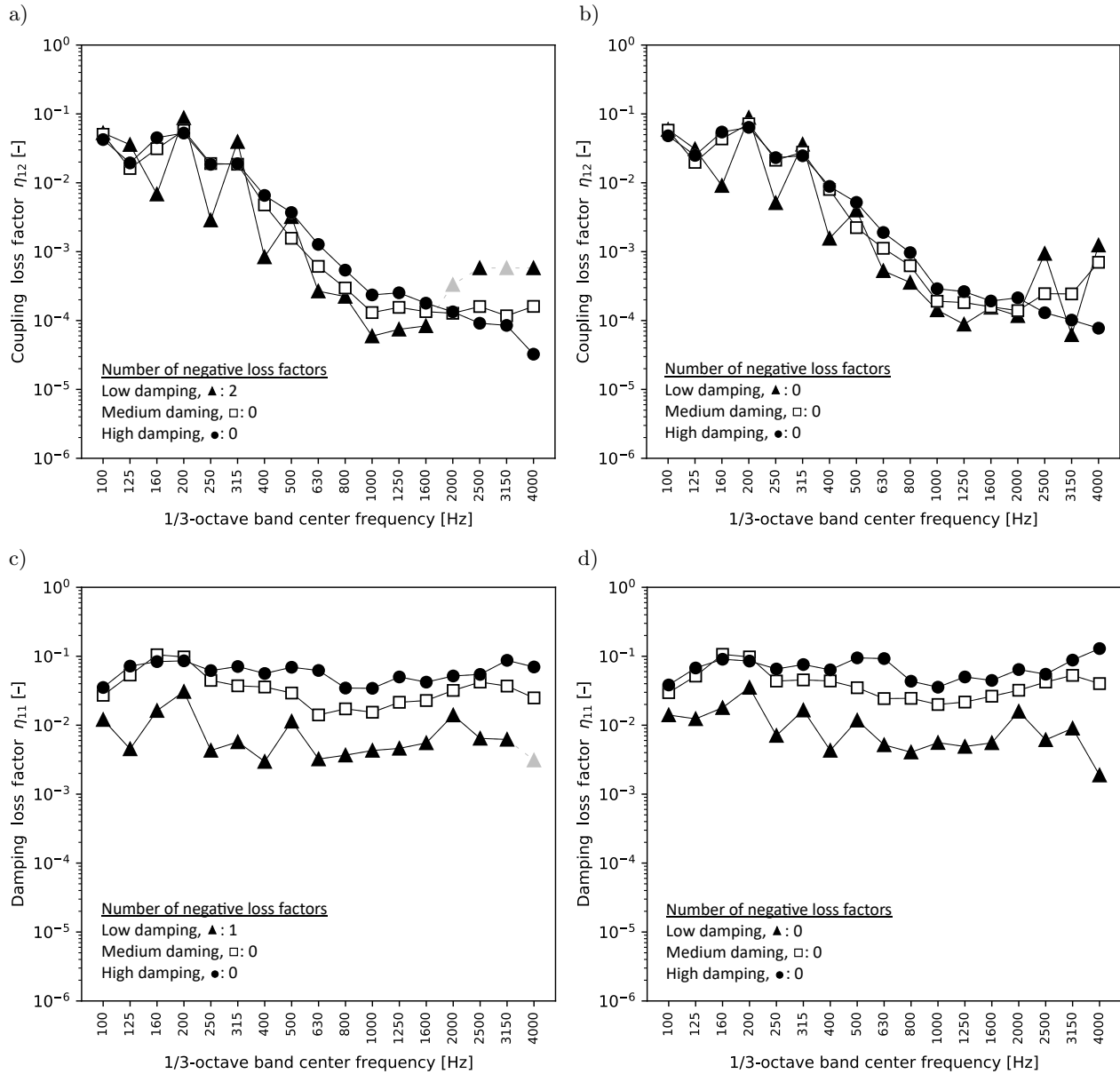


Fig. 12. Identification results for the rubber connection at all the damping levels: a) CLFs without MCF; b) CLF with MCF; c) DLFs without MCF; d) DLF with MCF.

Table 2. Effectiveness of the MCF in correcting negative LFs (both directions included).

Junction type	Damping	Corrected negative CLFs	Corrected negative DLFs
Line weld	Low	7 \rightarrow 0	15 \rightarrow 0
Line weld	Medium	2 \rightarrow 0	4 \rightarrow 0
Line weld	High	2 \rightarrow 0	1 \rightarrow 0
Rubber	Low	2 \rightarrow 0	2 \rightarrow 0
Rubber	Medium	0 \rightarrow 0	0 \rightarrow 0
Rubber	High	0 \rightarrow 0	0 \rightarrow 0
Point welding	High	1 \rightarrow 0	0 \rightarrow 0
Bolts	High	0 \rightarrow 0	0 \rightarrow 0
Rivets	High	0 \rightarrow 0	0 \rightarrow 0

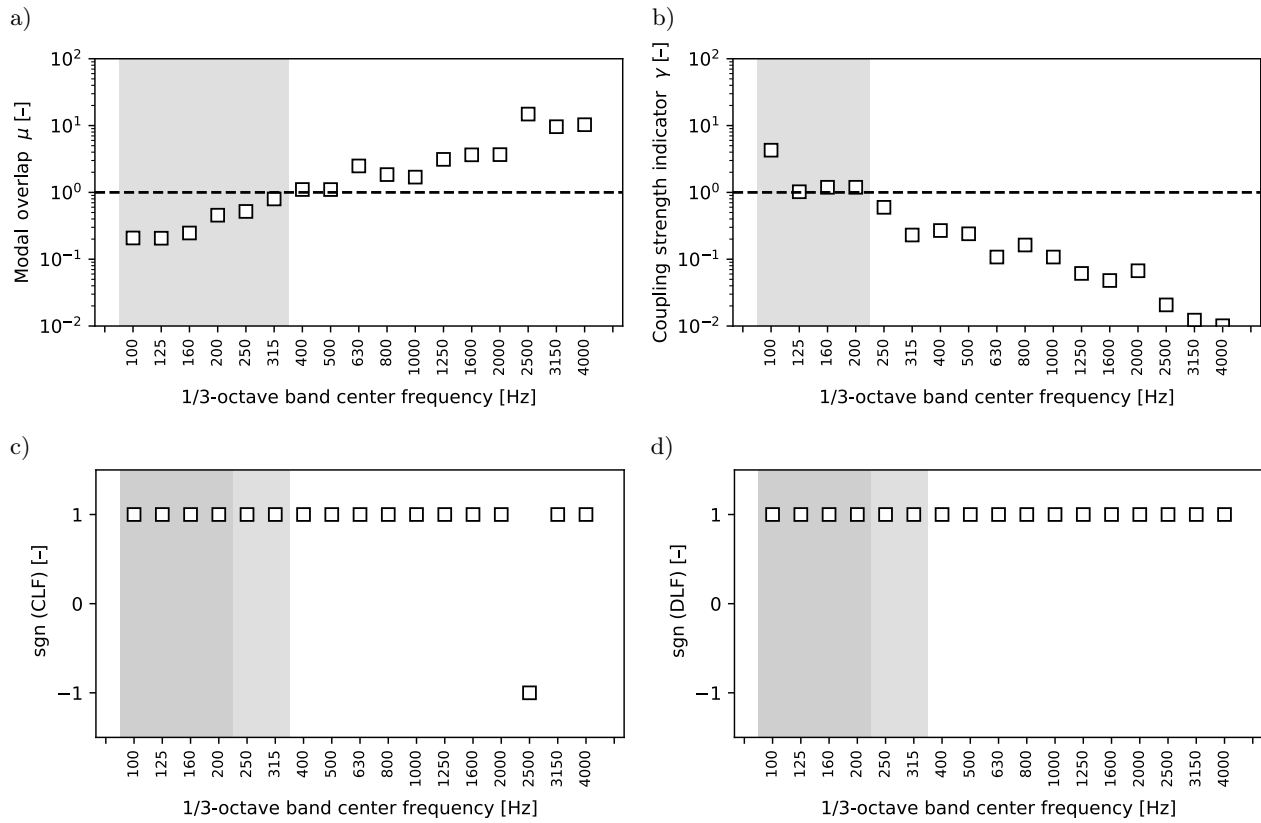


Fig. 13. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the point welding at high damping.

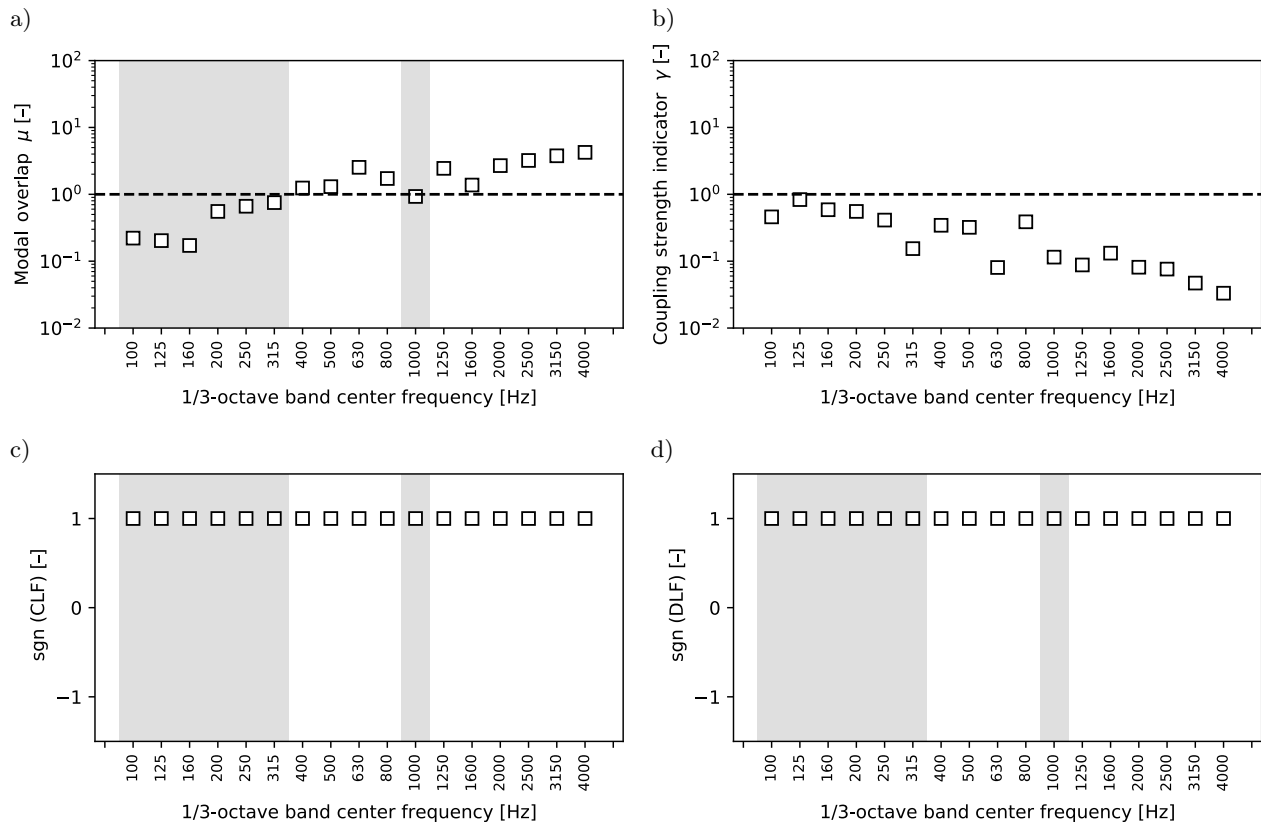


Fig. 14. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the rivets at high damping.

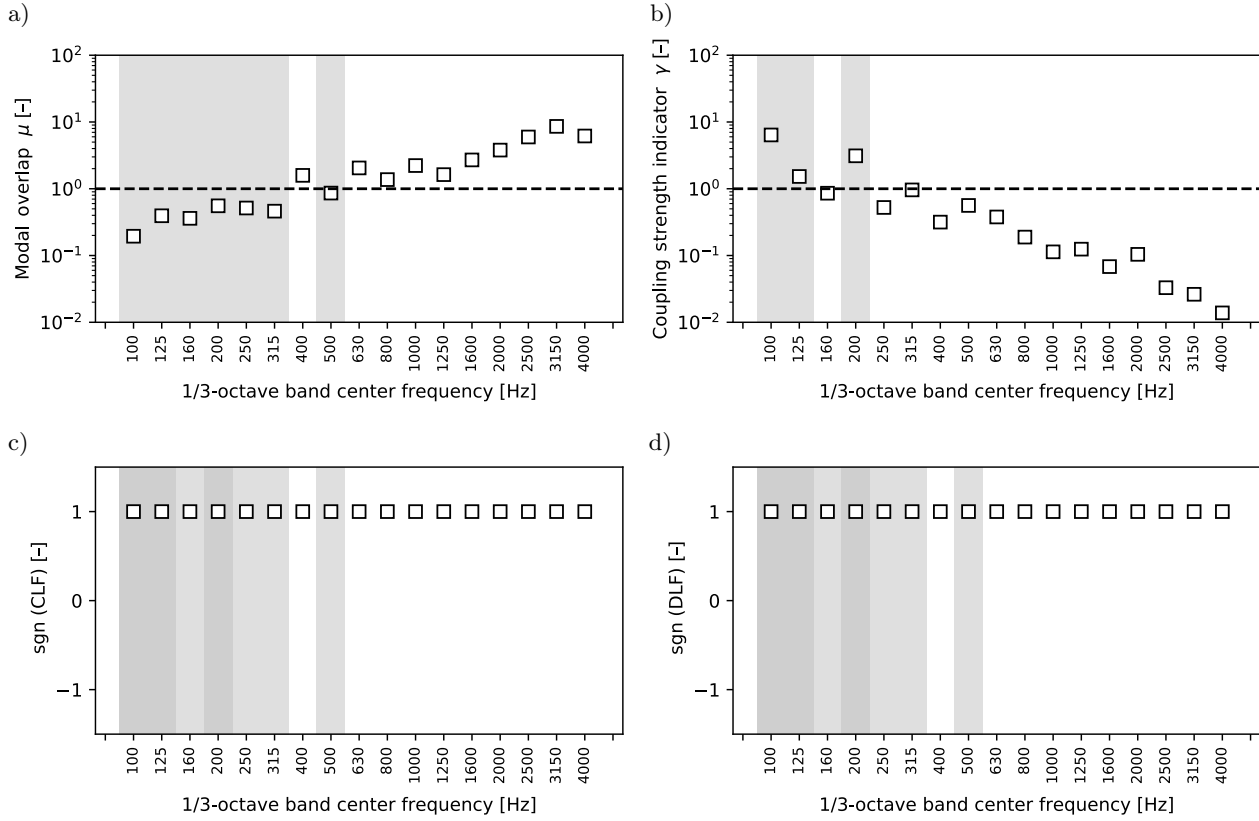


Fig. 15. Influence of the modal overlap (a) and coupling strength (b) on signs of the CLFs (c) and DLFs (d) for the bolts at high damping.

SEA matrix generated in the Monte Carlo simulation. The black dots represent correct SEA matrices, and the gray dots show incorrect matrices (giving negative LFs). The dots that are closer to zero represent the matrices closer (more similar) to the original matrix. Many norms can be used to compute the distance between the original matrix and individual sample matrices. In this work, the Frobenius norm is utilized. The matrix population generated according to instructions from (BOUHAIJ *et al.*, 2017) was based on measurement variance and characterized by the normal distribution to simulate a typical measurement process.

When MCF is performed on experimental data, the so-called bias appears. When the bias is present (Fig. 18a), several incorrect matrices can be found in the original matrix's vicinity (distance close to zero). Such bias is a result of the fact that the measurement is not providing the real expected value but only the approximate mean value. It is possible to remove the bias (Fig. 18b) by performing the second iteration of the MCF.

After removing the bias for the highly damped system with the line weld junction, the CLF decreased slightly in most frequency bands (Fig. 19b). The DLF was less sensitive to this operation (Fig. 19d). The CLF computed with and without removing the bias was compared with the FEM simulation results of this system. Slightly better convergence with the FEM sim-

ulation for most frequency bands was observed for the CLF computed without the bias (Fig. 20). The possible explanation for this can be that after removing the bias, a better average value can be obtained due to the fact that more correct samples are generated. This can be easily observed for the frequencies and systems with larger biases. Figure 21a shows the bias for the 2500 Hz frequency band of the low damped system with line welding. There is only a single "thin line" of correct matrices. Removing bias (Fig. 21b) resulted in significantly more correct matrices that were concentrated around the distance equal to 0. The change in the CLF for a low damped system is also more apparent (Fig. 19a). The same tendencies were observed for the rubber and point junctions, and it therefore seems reasonable to recommend removing the bias when performing MCF.

5. Conclusions

An extended validation of the MCF method was conducted, which can be considered the main contribution of this article. The validation included the following steps: 1) the influence of violating SEA assumptions on MCF results, and 2) the influence of the sample's bias on MCF results. Negative LFs occurred for frequency ranges where SEA assumptions were both

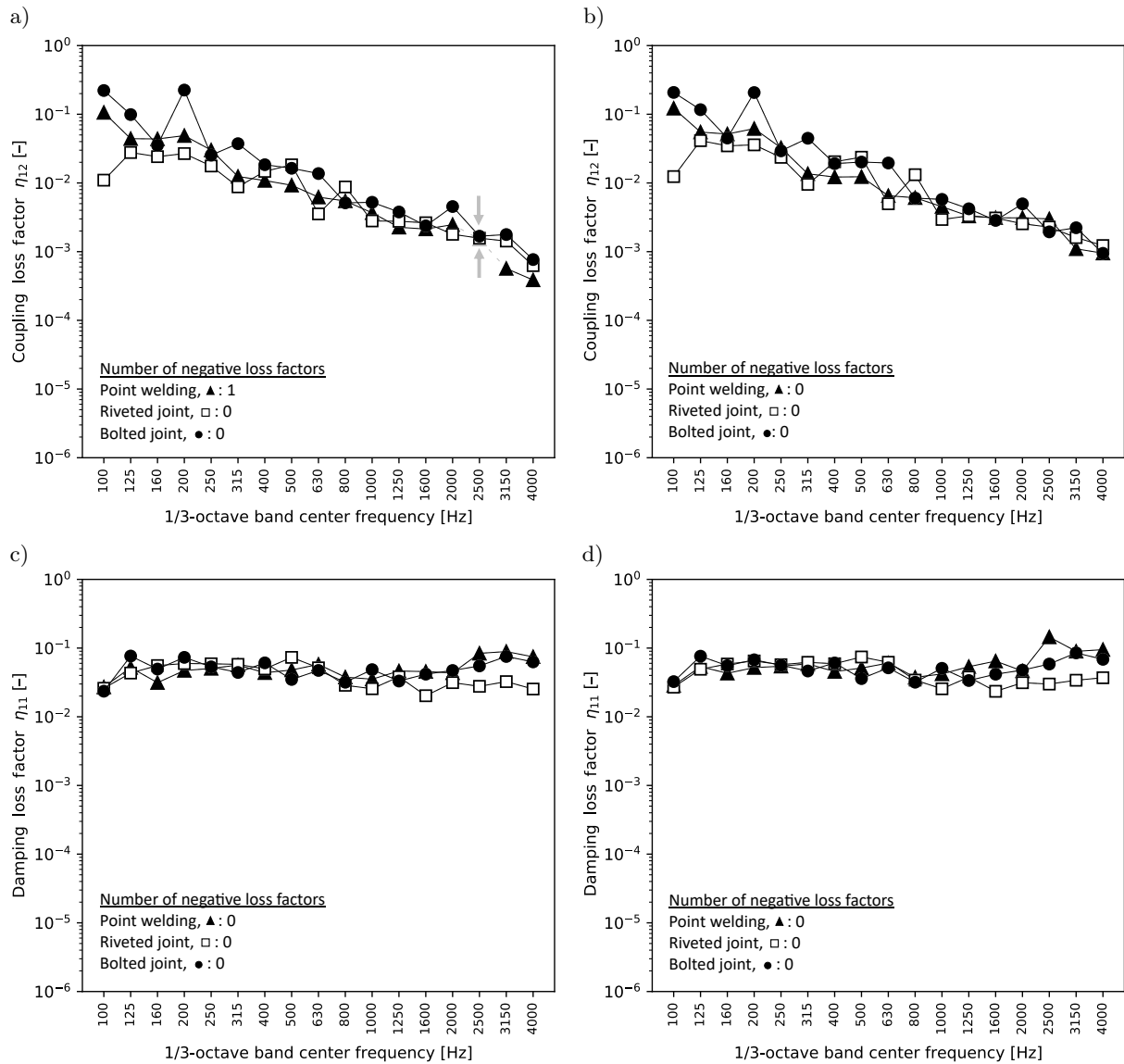


Fig. 16. Identification results for the point junctions: a) CLFs without MCF; b) CLF with MCF; c) DLFs without MCF; d) DLF with MCF.

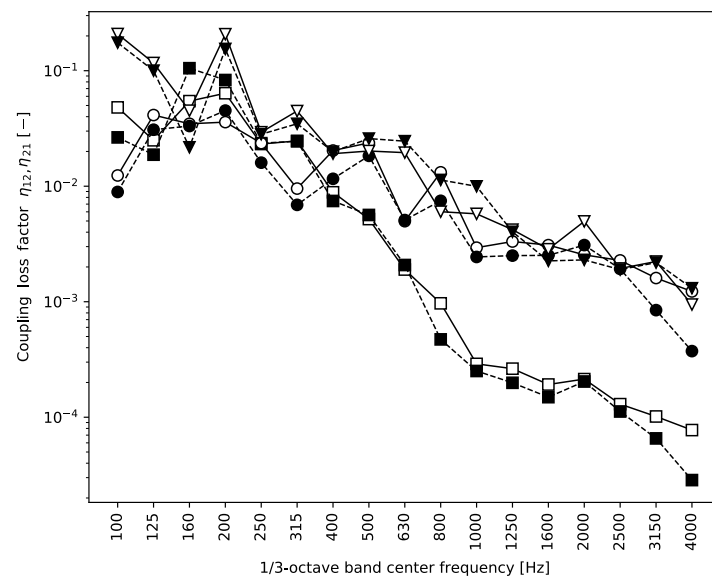


Fig. 17. CLFs in two directions for non-symmetrical systems: rubber (\blacksquare and \square), rivets (\bullet and \circ), bolts (\blacktriangledown and \triangledown).

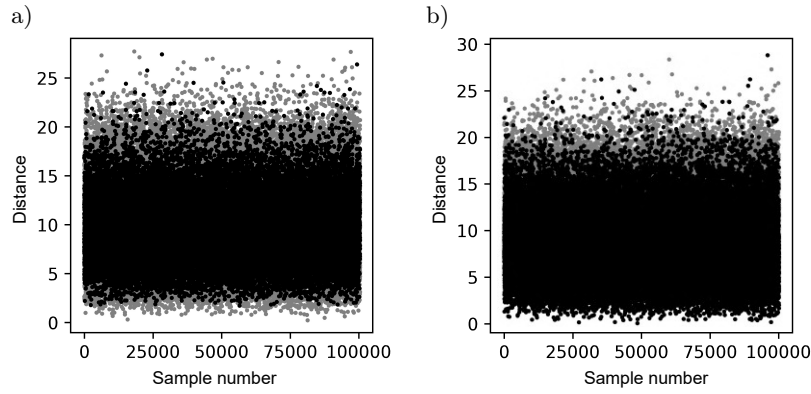


Fig. 18. Distance between the randomized energy matrices and the original energy matrix. Result for the line weld and high damping at 160 Hz: a) with bias; b) with bias removed.

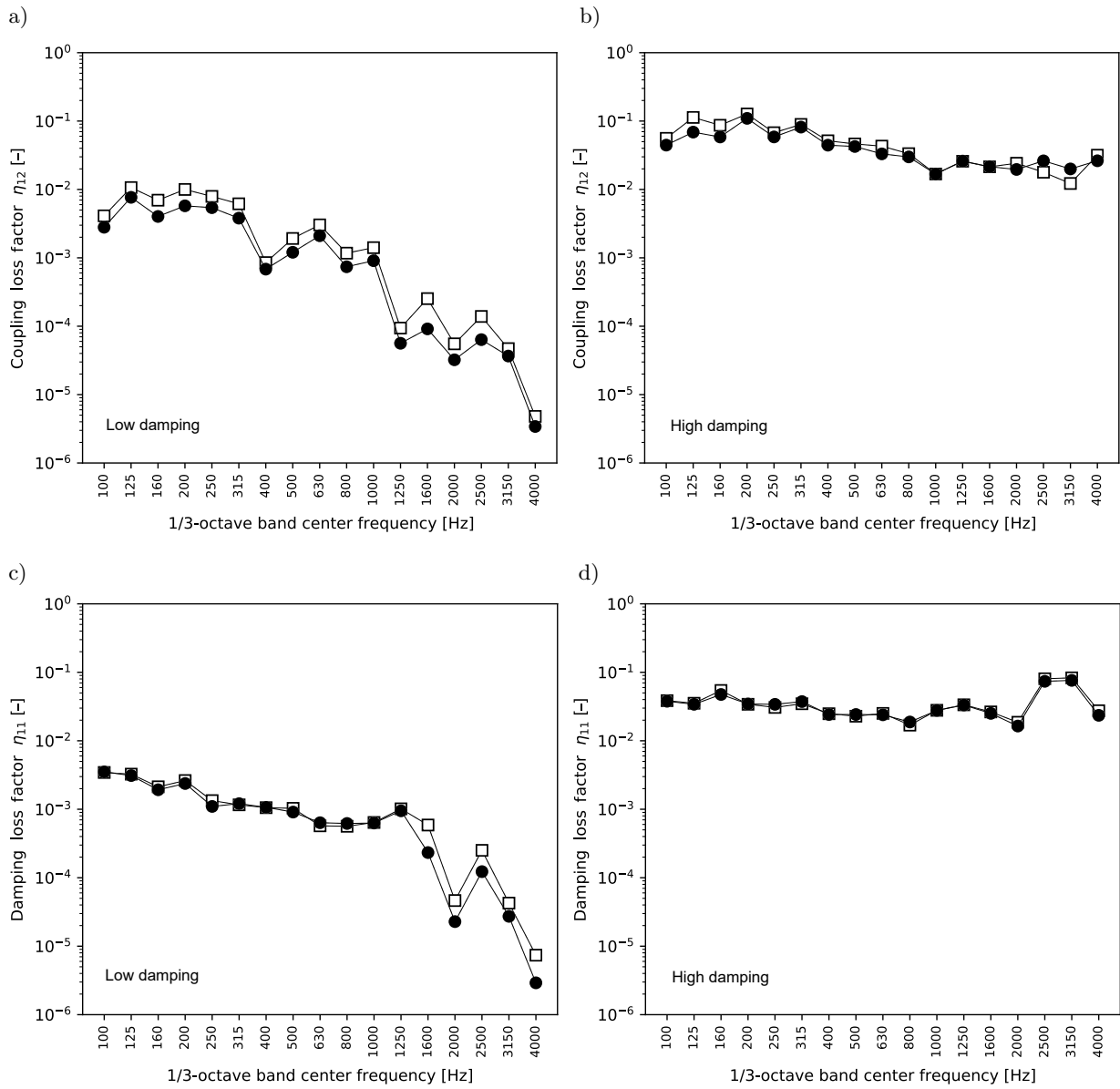


Fig. 19. Effect of removing the bias on the LF values. Results for the line weld: a) CLF at low damping; b) CLF at high damping; c) DLF at low damping; d) DLF at high damping.

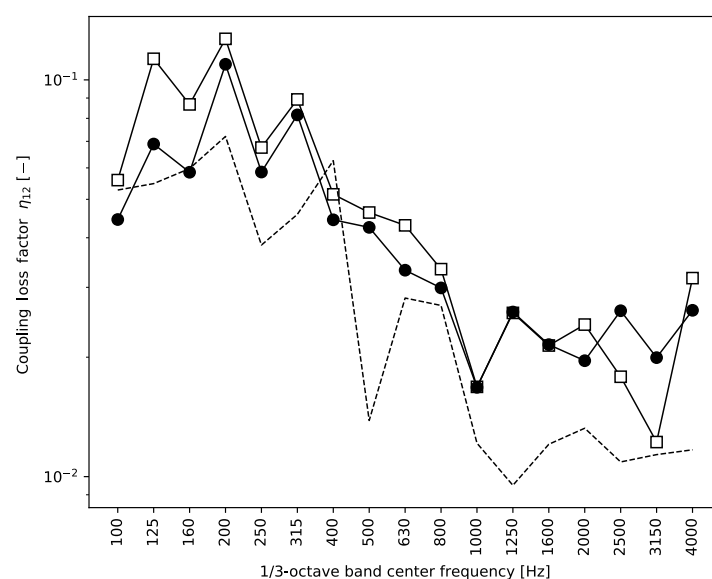


Fig. 20. Results in Fig. 19b compared with the FEM simulation. The experimental result with bias (\square), experimental result without bias (\bullet), FEM simulation (---).

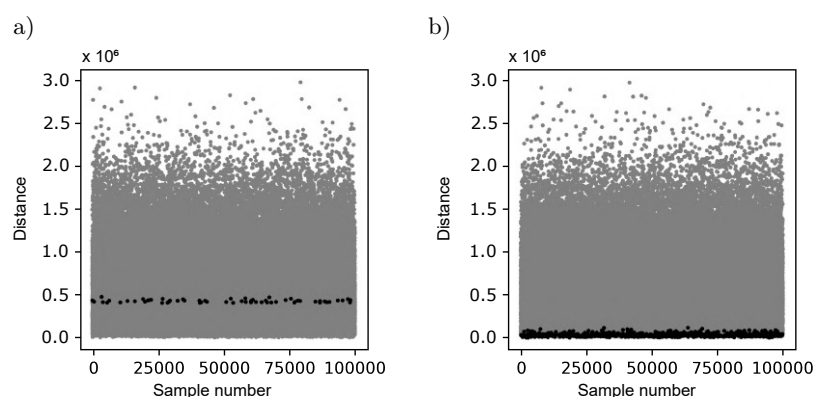


Fig. 21. Distance between the randomized energy matrices and the original energy matrix. Result for the line weld and low damping at 2500 Hz: a) with bias; b) with bias removed.

met and not met. After performing MCF, all negative values were successfully corrected. It seems that MCF can correct negative LFs, irrespective of whether SEA conditions are met or not.

It was observed that the LFs calculated from the bias-free MCF simulation show better agreement with the FEM simulation for frequency bands where the initial bias was significant. For frequency bands where the initial bias was slight, the removal of bias did not significantly alter the LFs. Based on this, it seems reasonable to remove bias each time the MCF procedure is performed.

Acknowledgments

This work was financed by KFB Acoustics and the Department of Acoustics, Multimedia and Signal Processing at Wrocław University of Science and Technology.

References

1. BHAGWAN M.M., POPURI B. (2019), Estimation of coupling loss factors for rectangular plates with different materials and junctions, *Noise & Vibration Worldwide*, **50**(9–11): 306–312, doi: 10.1177/0957456519883264.
2. BIES D.A., HAMID S. (1980), In situ determination of loss and coupling loss factors by the power injection method, *Journal of Sound and Vibration*, **70**(2): 187–204, doi: 10.1016/0022-460X(80)90595-7.
3. BORELLO G. (2018), Prediction of sound transmission in aircraft over the mid and high frequency range, [in:] *Inter-Noise and Noise-Con Congress and Conference Proceedings*, **258**(2): 5115–5124.
4. BOUHAJ M., VON ESTORFF O., PEIFFER A. (2017), An approach for the assessment of the statistical aspects of the SEA coupling loss factors and the vibrational energy transmission in complex aircraft structures: Experimental investigation and methods benchmark, *Journal of Sound and Vibration*, **403**: 152–172, doi: 10.1016/j.jsv.2017.05.028.

5. CACCIOLATI C., GUYADER J.L. (1994), Measurement of SEA coupling loss factors using point mobilities, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, **346**(1681): 465–475, doi: 10.1098/rsta.1994.0029.
6. CHEN X., WANG D., MA Z. (2012), Simulation on a car interior aerodynamic noise control based on statistical energy analysis, *Chinese Journal of Mechanical Engineering*, **25**(5): 1016–1021.
7. CIMERMAN B., BHARJ T., BORELLO G. (1997), Overview of the experimental approach to statistical energy analysis, [in:] *SAE Noise and Vibration Conference and Exposition*, doi: 10.4271/971968.
8. CRAIK R.J.M. (1982), The prediction of sound transmission through buildings using statistical energy analysis, *Journal of Sound and Vibration*, **82**(4): 505–516, doi: 10.1016/0022-460X(82)90404-7.
9. CULLA A., SESTIERI A. (2006), Is it possible to treat confidentially SEA the wolf in sheep's clothing?, *Mechanical Systems and Signal Processing*, **20**(6): 1372–1399, doi: 10.1016/j.ymssp.2005.02.007.
10. DE LAS HERAS M.J.F., CHIMENO M., MILLÁN E.R., HIDALGO F.S. (2018), On the influence of the condition number on the resolution of an ESEA model, [in:] *Inter-Noise and Noise-Con Congress and Conference Proceedings*, **257**(1): 153–161.
11. DE LAS HERAS M.J.F., MANGUÁN M.C., MILLÁN E.R., DE LAS HERAS L.J.F., HIDALGO F.S. (2020), Determination of SEA loss factors by Monte Carlo Filtering, *Journal of Sound and Vibration*, **479**: 115348, doi: 10.1016/j.jsv.2020.115348.
12. FAHY F.J., RUIVO H.M. (1997), Determination of statistical energy analysis loss factors by means of an input power modulation technique, *Journal of Sound and Vibration*, **203**(5): 763–779, doi: 10.1006/jsvi.1996.0892.
13. FINNVEDEN S. (2011), A quantitative criterion validating coupling power proportionality in statistical energy analysis, *Journal of Sound and Vibration*, **330**(1): 87–109, doi: 10.1016/j.jsv.2010.08.003.
14. GU J., SHENG M. (2015), Improved energy ratio method to estimate coupling loss factors for series coupled structure, *Journal of Mechanical Engineering*, **45**(1): 37–40, doi: 10.3329/jme.v45i1.24382.
15. HATTORI K., NAKAMACHI K., SANADA M. (1985), Prediction of underwater sound radiated from ship's hull by using statistical energy analysis, *Proceedings of Inter-Noise 85*, Vol. II, p. 645.
16. HODGES C.H., NASH P., WOODHOUSE J. (1987), Measurement of coupling loss factors by matrix fitting: An investigation of numerical procedures, *Applied Acoustics*, **22**(1): 47–69, doi: 10.1016/0003-682X(87)90015-6.
17. HOPKINS C. (2002), Statistical energy analysis of coupled plate systems with low modal density and low modal overlap, *Journal of Sound and Vibration*, **251**(2): 193–214, doi: 10.1006/jsvi.2001.4002.
18. HWANG H.J. (2002), Prediction and validation of high frequency vibration responses of NASA Mars Pathfinder spacecraft due to acoustic launch load using statistical energy analysis, *NASA Technical Reports Server*.
19. JI L., SHENG X., XIAO X., WEN Z., JIN X. (2015), A review of mid-frequency vibro-acoustic modelling for high-speed train extruded aluminium panels as well as the most recent developments in hybrid modelling techniques, *Journal of Modern Transportation*, **23**(3): 159–168, doi: 10.1007/s40534-015-0080-4.
20. LAFONT T., TOTARO N., LE BOT A. (2014), Review of statistical energy analysis hypotheses in vibroacoustics, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **470**(2162): 20130515, doi: 10.1098/rspa.2013.0515.
21. LALOR N. (1990), Practical Consideration for the Measurement of Internal and Coupling Loss Factors on Complex Structures, *ISVR Technical Report*.
22. LALOR N. (1996), Experimental statistical energy analysis: A tool for the reduction of machinery noise, *The Journal of the Acoustical Society of America*, **99**(4): 2568–2574, doi: 10.1121/1.415057.
23. LE BOT A. (2015), *Foundation of Statistical Energy Analysis in Vibroacoustics*, Oxford University Press, United Kingdom.
24. LYON R.H., DEJONG R.G., (1995), *Theory and Application of Statistical Energy Analysis*, Elsevier.
25. MANDALE M.B., BANGARU BABU P., SAWANT S.M. (2016), Statistical energy analysis parameter estimation for different structural junctions of rectangular plates, *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **230**(15): 2603–2610, doi: 10.1177/0954406215615628.
26. MING R. (1998), The measurement of coupling loss factors using the structural intensity technique, *The Journal of the Acoustical Society of America*, **103**(1): 401–407, doi: 10.1121/1.421096.
27. NIERADKA P., DOBRUCKI A. (2018), Insertion loss of enclosures with lined slits, [in:] *Euronoise 2018-Conference Proceedings*.
28. PANKAJ A.C. (2019), *Numerical and Experimental Investigations on Damage Detection in Joints Based on Statistical Energy Analysis like Approach*, Ph.D. Thesis, Department of Mechanical Engineering, National Institute of Technology Karnataka, Surathkal.
29. PANUSZKA R., WICIAK J., IWANIEC M. (2005), Experimental assessment of coupling loss factors of thin rectangular plates, *Archives of Acoustics*, **30**(4): 533–551.
30. PRICE A.J., CROCKER M.J. (1970), Sound transmission through double panels using statistical energy analysis, *The Journal of the Acoustical Society of America*, **47**(3A): 683–693, doi: 10.1121/1.1911951.
31. SMITH JR. P.W. (1979), Statistical models of coupled dynamical systems and the transition from weak to

- strong coupling, *The Journal of the Acoustical Society of America*, **65**(3): 695–698, doi: 10.1121/1.382481.
32. YAP F.F., WOODHOUSE J. (1996), Investigation of damping effects on statistical energy analysis of coupled structures, *Journal of Sound and Vibration*, **197**(3): 351–371, doi: 10.1006/jsvi.1996.0536.
33. YOGANANDH M., NAGARAJA J., VENKATESHAM B. (2019), Prediction of insertion loss of lagging in rectangular duct using statistical energy analysis, *Noise Control Engineering Journal*, **67**(6): 438–446, doi: 10.3397/1/376740.
34. ZARATE R., MATUS E., LOPEZ M., BALLESTEROS L. (2017), Design of quieter kitchen appliances: Sound pressure level modeling and validation of a household refrigerator using statistical energy analysis, [in:] *Proceedings of Meetings on Acoustics*, **30**(1): 030009, doi: 10.1121/2.0000632.

Research Paper

Three-Dimensional Freehand Ultrasound Strain Elastography Based on the Assessment of Endogenous Motion: Phantom Study

Andrius SAKALAUSKAS⁽¹⁾, Rytis JURKONIS^{(2)*}, Arūnas LUKOŠEVIČIUS⁽²⁾

⁽¹⁾ *TELEMED, Ultrasound Medical Systems*
Vilnius, Lithuania

⁽²⁾ *Biomedical Engineering Institute, Kaunas University of Technology*
Kaunas, Lithuania

*Corresponding Author e-mail: rytis.jurkonis@ktu.lt

(received June 16, 2022; accepted November 10, 2022)

The purpose of this paper is to present the results of the pilot experiments demonstrating proof of concept of three-dimensional strain elastography, based on freehand ultrasound for the assessment of strain induced by endogenous motion. The technique was tested by inducing pulsatility in an agar-based tissue mimicking phantom with inclusions having different stiffness and scanning the 1D array with an electromagnetic position sensor. The proof of concept is explored with a defined physical phantom and the adopted algorithm for strain analysis. The agar-based phantom was manufactured with two cylindrical inclusions having different stiffness (7 kPa and 75 kPa in comparison to the background 25 kPa) and scattering properties. The internal strain in the phantom was introduced by mimicking a pulsating artery. The agar mixture displacements were estimated by using the GLUE algorithm. The 3D isosurfaces of inclusion from rendered volumes obtained from the B-mode image set and strain elastograms were reconstructed and superimposed for a quantitative comparison. The correspondence between the B-mode image-based inclusion volume and the strain elastography-based volume was good (the Jaccard similarity coefficient in the range 0.64–0.74). The obtained results confirm the 3D freehand endogenous motion-based elastography as a feasible technique. The visualization of the inclusions was successful. However, quantitative measurements showed that the accuracy of the method in volumetric measurements is limited.

Keywords: strain elastography; endogenous motion; freehand scanning; 3D imaging; tissue mimicking phantom.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Strain elastography is one of the techniques used for the assessment of tissue stiffness and has been shown to be efficient in the diagnostics of breast lesions (DIETRICH *et al.*, 2017), thyroid (COSGROVE *et al.*, 2017), prostate (BARR *et al.*, 2017), and other areas. The strain elastography measures slight tissue deformations (strain range 0.1–2%), which are typically induced by external compression by an ultrasound probe (WELLS, LIANG, 2011). The algorithm calculates the axial displacements between the consecutive RF data frames and the strain estimates are obtained by taking the spatial derivative in the axial direction

(DIETRICH *et al.*, 2017). The obtained strain values are color coded and the two-dimensional (2D) strain image is superimposed on the structural B-mode image, which is based on the reflection magnitude. The scanning specialist interprets both images and measures parameters of the region of interest (ROI), such as the strain ratio or the elasticity score. Overall, the strain elastography is a qualitative technique due to unknown stress, but still provides the relative information about tissue stiffness (DIETRICH *et al.*, 2017). Unfortunately, the strains induced by the probe can be estimated only in the case of tissue, which are relatively superficial (up to few centimeters), and the deeper structures cannot be observed and assessed by the technique, be-

cause uniform strain cannot be induced in deeper layers. Fortunately, there are other ways to implement strain elastography by employing endogenous motion and strain of tissue, which are caused due to the natural pulsatility of blood vessels and the beating heart. The strain elastography based on endogenous motion was pioneered by the authors (DICKINSON, HILL, 1982; WILSON, ROBINSON, 1982; TRISTAM *et al.*, 1986). Endogenously induced strain could also be employed for imaging purposes. The phantom study with internal deformation caused by pressure variation in the vessel is presented in (MAI, INSANA, 2002). The initial experiments of endogenous strain assessment and imaging were performed by our group (ZAMBACEVIČIENĖ *et al.*, 2019; SAKALAUSKAS *et al.*, 2018). Later, it was found that the parameters derived from the two-dimensional images of endogenous strain have a correlation with different stages of liver fibrosis (SAKALAUSKAS *et al.*, 2019) and portal hypertension (GELMAN *et al.*, 2020).

Three-dimensional (3D) elasticity imaging in clinical practice is still limited, manufacturers of ultrasound scanners and the scientific community are still improving this modality. Three-dimensional imaging by using matrix arrays is limited by the frame rate and spatial resolution, which are typically lower for 2D arrays (LEE *et al.*, 2018). The 2D matrix array technology requires very sophisticated electronics to deal with very large number of channels connected to array elements. Fortunately, it is possible to implement 3D volume acquisition by using a conventional one-dimensional (1D) array probe and a mechanical linear scanner (HUANG *et al.*, 2015; HENDRIKS *et al.*, 2016; RICHARDS *et al.*, 2009), or a 3D transducer translating mechanical device (CHEN *et al.*, 2016). A much cheaper and relatively simple method could be implemented only by attaching a position tracking sensor to an ultrasound probe (GILJA *et al.*, 1998; WANG *et al.*, 2013). This enables so-called freehand 3D mode (MOZAFFARI, LEE, 2017). The comparison of simulated 3D strain imaging with the mechanically swept probe and freehand scanning (HOUSDEN *et al.*, 2010) has shown that it is possible to produce a good strain images using the freehand method, which may be preferable in clinical practice. The freehand 3D ultrasound imaging in the B-mode was extensively reviewed by MOZAFFARI and LEE (2017). The most of reviewed B-mode systems employed optical or electromagnetic position tracking solutions. But only several approaches to implement elastography by using 1D array probe and position sensor could be found (LEE *et al.*, 2018; LINDOP *et al.*, 2006). LINDOP *et al.* (2006) used an optical position tracking system for 3D quasistatic elastography, meanwhile LEE *et al.* (2018) proposed to use an electromagnetic one. Both groups successfully conducted 3D freehand strain elastography experiments and demonstrated the feasibility of the method. However, these methods applied external pressure by the operator to

induce strain by the scanning probe. This makes the scanning procedure quite complicated since the probe movements should meet the simultaneous requirements of strain induction and scanning of volume. An external strain induction source based on the use of applied low frequency vibrator was demonstrated on the phantom (BERCOFF *et al.*, 2004). A similar motorized 3D vibro-elastography method also uses induced harmonic motion and synchronous strain measurements (ABEYSEKERA *et al.*, 2015), but both methods are not adapted to clinically preferable freehand scanning.

Fortunately, external deformation or vibration potentially could be substituted by endogenous excitation induced by natural cardiovascular activity. Initial publications on the endogenous cardiovascular motion application showed its potential for liver elasticity imaging (KOLEN *et al.*, 2004), also for thyroid nodule diagnostics using carotid artery pulsation (BAE *et al.*, 2006) and diastolic strain variation (LUO *et al.*, 2009). This approach could have advantage, since 3D scanning using 1D array will not require to apply the uniform compression at the same time, that might be not so easy to accomplish by scanning by the performing operator. In case of endogenous motion, the probe can be moved completely freehand without the requirement to make uniform pushes to induce strain. However, the displacements and, respectively, a strain magnitude for the endogenous excitation are frequently much lower and omnidirectional comparing with the ones induced externally but it is not so easy to measure them. This leads to high requirements for the specificity of the displacement estimation algorithm, especially in the case of 3D imaging of a dynamic structure.

The purpose of this paper is to present the proof of concept of 3D strain elastography, based on strain induced by endogenous motion and freehand 3D ultrasound imaging. To the best of our knowledge there are no published scientific papers, which deal with endogenous motion-based 3D elastography and we see our novelty in the phantom-based analysis of strain endogenous excitation and its application for 3D fused strain and B-mode imaging. The method was tested by inducing pulsatility in the agar-based tissue mimicking phantom with cylindrical inclusions having different stiffness and scattering properties in comparison to the background material. The strain-based volumes of the inclusions were evaluated and compared to volumes obtained from B-mode images for a quantitative assessment.

2. Materials and methods

2.1. Scanner and preset

The ultrasound scanner Ultrasonix Sonix Touch (Analogic Corporation, Peabody, MA, USA), which allows the collection of raw RF signals from all scanning

lines, equipped with the linear array transducer (L14–5/38 GPS, 128 acoustic elements) was used for data collection in this study.

The main parameters of ultrasonic scanning were as follows: scanning depth – 7 cm, ultrasound frequency – 13.3 MHz, transmit single focal depth – 3.5 cm. RF data were recorded in 314 scanning lines per the B-scan frame, the frame rate – 27 Hz. The RF signal digitized with 40 MHz sampling frequency and the analog-to-digital converter resolution 16 bits.

The ultrasound scanner has the ability to measure the position and orientations of the linear probe by the electromagnetic tracking system trakSTAR (Ascension, Yarraville, Australia) incorporated with the ultrasound scanner by a manufacturer. According to the specification from a manufacturer the accuracy of a static position is 1.4 mm RMS and 0.5° RMS of a static orientation angle. Sequences of B-scan images the RF signals and 3D position data were acquired and stored for later offline processing.

2.2. Phantom

An experimental phantom, mimicking tumor foci in the liver, was produced. The liver-mimicking phantom having two inclusions with different stiffness was made from agar mixtures. The surrogate endogenous motion was induced by mimicking a pulsating artery. The background of the phantom was manufactured by mixing 10 g/l agar concentration in distilled water. The agar powder was slowly stirred until it was fully dissolved in boiling water. The total weight of the mixture was controlled at the start and end of the boiling of the mixture. The evaporated amount of water was compensated by weight to ensure the planned concentration. The planned concentration ensures the stiffness of 25 kPa according to the Young modulus as predicted by HALL *et al.* (1997). To obtain scattering the powder prepared from a carbon tablet was used. We dispersed 1.05 g of the carbon powder in 1300 ml of the agar mixture. We continuously stirred the mixture with the carbon powder until it cooled down. Stirring prevented sedimentation of the carbon powder. At the start of gelation, the mixture was poured into the prepared container. Two pieces of metal pipe of 16 mm in diameter and about 90 mm in length were placed in the container. The thin-walled rubber vessel (diameter 7 mm, length 200 mm) was fixed to mimic the artery below these pipes. Both inclusions of approximately half the length of the rubber vessel were arranged along a single axis that is parallel to the axis of a vessel. Both ends of the rubber vessel were connected to flange fixtures in the walls of container, to allow the fluid to pass through. The bottom side of the container was covered with a 10 mm thick layer of foam below the rubber vessel. This foam layer acts as an absorber of ultrasound waves and allows one to

avoid multiple reflections of waves in the phantom. The schematic drawing of the phantom structure is presented in Fig. 1.

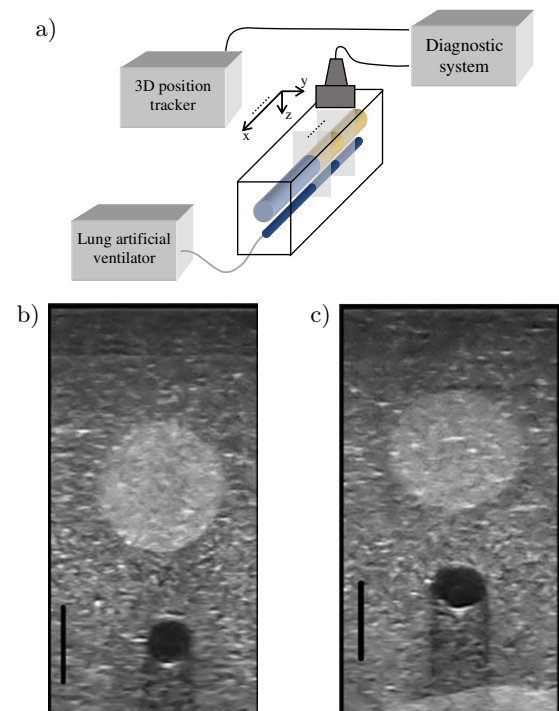


Fig. 1. Sketch of the experimental setup (a) and obtained B-scans (b, c): a) 3D freehand scanning for data recording (B-scan images, RF signals, and position sensor readings) for further volume rendering and 3D imaging of the inclusions; b) and c) B-scan images acquired with fixed probe from inclusions having similar acoustic (scattering) contrasting from background, but different mechanical (strain) properties. The black bar indicates 1 cm length.

The container was kept still to solidify at the refrigerator temperature for 1 hour. The container with a solid mixture of agar was further elaborated by implementing two inclusions. The pieces of metal pipe were filled with hot water to warm up and make the removal of the pipes from the agar mixture more slippery. Inclusions were made of agar concentrations 17.7 and 5 g/l (or predicted the Young modulus 75 kPa and 7 kPa, respectively) according to the same procedure as a background. Two mixtures of 100 ml each were added with 0.27 g of the carbon powder to make the waves scattering from inclusions stronger than the background material. At the beginning of gelation, the first mixture was poured into tunnel-like holes in the background. The phantom was kept still to solidify the agar mixture into the first inclusion. The implementation of the second inclusion was repeating the procedure of preparation of the first inclusion.

The thin-walled rubber vessel is intended to mimic the cardiovascular pulsations. The vessel was filled with distilled water and occluded in one end. The other end (the flange) was connected to the corru-

gated tubing of the artificial lung ventilation output. Pulsations in the elastic rubber vessel were induced by a ventilator operated in a pressure-controlled mode. The pressure pulsation amplitude in the rubber vessel was 32 mmHg, while the pulsing rate was 70 times per minute. Both (the soft and hard) inclusions were excited from the same pulsating vessel.

The access to the prepared tissue mimicking materials in the container was from an open top of the container. The linear array transducer was directed to send ultrasound waves from the top surface of the agar mixture downwards: crossing through the one of inclusions and a rubber vessel (see Fig. 1a). B-scan images were obtained in a cross sections of inclusion and the rubber vessel (see Figs. 1b and 1c).

2.3. Experiment

Two experiments of RF scanning of the phantom were performed:

- 1) RF data acquisition keeping the probe fixed. The probe was fixed in a laboratory stand. The B-scan plane was directed orthogonally to the inclusions, observing to get cross-section images of them. The diameters of 16.3 ± 0.4 mm and 15.3 ± 1.2 mm for hard and soft inclusions correspondingly were measured from B-scan images when the probe was fixed,
- 2) freehand 3D RF scanning. The translation of the probe on a top surface of phantom was manual. The translation was guided by the frame to get a straight-line trajectory, which was approximately aligned in parallel to axis of X coordinates of the electromagnetic position tracker. The B-scan plane direction to the phantom was handled manually, ensuring center of the selected inclusion in the middle of the B-scan image. The probe translation was limited to 21 mm. Freehand translations along only the selected inclusion were used for processing, excluding imaging of other inclusions.

2.4. Collected data

The B-scans (174 frames for each recording) and raw corresponding RF signals were acquired and stored for offline processing. The electromagnetic tracking sensor measurements ($x, y, z, \alpha, \beta, \gamma$) were also stored for each frame. The RF data matrixes were 3648 (samples) \times 314 (scanning lines) \times 174 (frames) in size, meanwhile the B-mode datasets were of $616 \times 820 \times 174$ size with an isotropic pixel size of 0.15 mm. In total 4 records were acquired and analyzed during the study (for both inclusions with a fixed probe and applying 3D freehand scanning).

The flow chart representing the concept of data processing in the study is presented in Fig. 2.

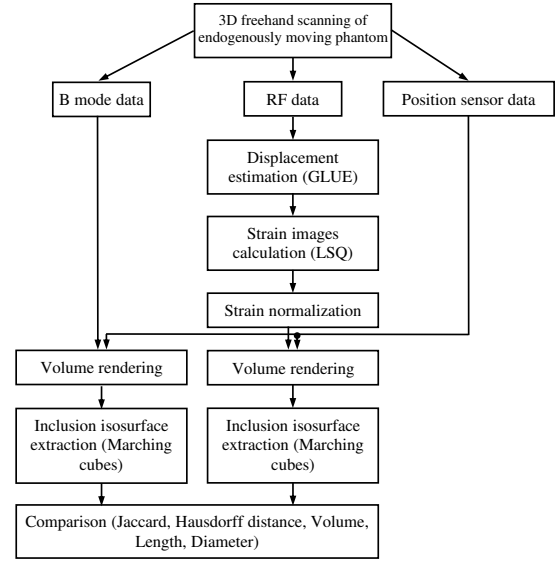


Fig. 2. Flowchart of the presented study: 3D freehand scanned data (B-mode images, RF signals, and position sensor readings) are used for volume rendering and 3D imaging of the inclusions having similar acoustic (scattering), but different mechanical (strain) properties, which are finally compared for the correspondence assessment.

2.5. Displacements estimation

Displacement images for consecutive frame pairs were obtained using the global time-delay estimation technique (GLUE) proposed by Hassan Rivaz group. The implementation and codes are provided by GLUE developers (HASHEMI, RIVAZ, 2017; RIVAZ *et al.*, 2011). GLUE is a continuation of their work on time delay estimation, which at first were solved by the dynamic programming-based approach (RIVAZ *et al.*, 2011). The GLUE algorithm refines the integer displacements estimates obtained by dynamic programming to a subsample accuracy. The subsample accuracy displacements are obtained in the axial and lateral directions. Overall, the displacement estimation algorithm has two stages: 1) estimation of integer displacements using dynamic programming; 2) GLUE: refinement of integer displacements to a subsample accuracy by minimizing the regularized cost function (presented in Eq. (1)).

The cost function for GLUE is formulated for the entire two-dimensional image (HASHEMI, RIVAZ, 2017):

$$\begin{aligned}
 C(\Delta a_{1,1}, \dots, \Delta a_{m,n}, \Delta l_{1,1}, \dots, \Delta l_{m,n}) \\
 = \sum_{j=1}^n \sum_{i=1}^m \left\{ [I_1(i, j) - I_2(i + a_{i,j} + \Delta a_{i,j}, j + l_{i,j} + \Delta l_{i,j})]^2 \right. \\
 + \alpha_1 (a_{i,j} + \Delta a_{i,j} - a_{i-1,j} - \Delta a_{i-1,j})^2 \\
 + \beta_1 (l_{i,j} + \Delta l_{i,j} - l_{i-1,j} - \Delta l_{i-1,j})^2 \\
 + \alpha_2 (a_{i,j} + \Delta a_{i,j} - a_{i,j-1} - \Delta a_{i,j-1})^2 \\
 \left. + \beta_2 (l_{i,j} + \Delta l_{i,j} - l_{i,j-1} - \Delta l_{i,j-1})^2 \right\}, \quad (1)
 \end{aligned}$$

where a is the initial integer axial displacements estimated by dynamic programming, l is the initial integer lateral displacements estimated by dynamic programming, α and β are the regularization terms, I_1 and I_2 are the consecutive RF data frames before and after the deformation, $i = 1, \dots, m$, where m is the number of samples in RF scanning line, $j = 1, \dots, n$, where n is the number of RF scanlines in the RF frame, Δ notes the subsample accuracy values for axial and lateral displacements. Displacements are usually very small, and subsample accuracy must be achieved.

For the presented study, the following values of the regularization coefficients were established: $\alpha_1 = \alpha_2 = 35.8$, and $\beta_1 = \beta_2 = 0.02$. The parameters of the dynamic programming algorithm were set as recommended by GLUE developers (HASHEMI, RIVAZ, 2017; RIVAZ *et al.*, 2011). Only axial displacement images were used for further processing in our 3D approach, because it was found that the lateral displacement does not provide sufficient contrast for the analysis.

We recorded the data to evaluate and verify the frame-to-frame displacement waveform that we induced with pressure pulsations with a fixed array transducer in the first experiment. The spatial mean displacement was calculated from all points in the frame to evaluate the amplitude and shape of the excitation waveform. First, the average value for each displacement image was calculated, obtaining how the mean displacement varies over time. Next, the summed displacement signal was obtained by calculating the cumulative sum of the frame-to-frame displacement waveform. The summed displacement waveform was high-pass filtered (cut-off frequency 0.5 Hz) to remove the baseline drift.

2.6. Strain estimation

Strain images were obtained using the least squares regression (LSQ) technique as proposed by HASHEMI and RIVAZ (2017). The size of the differentiation kernel was 35 samples axially (0.67 mm), and 5 lines laterally (0.43 mm).

2.7. Strain normalization

Strain images were normalized (provided in Eq. (2)) to rescale them and to compensate for the strain differences, which arises due to the nature of endogenous compression that varies over time. The statistical normalization method was used:

$$S_N = \frac{S - \bar{S}}{\sigma_S}, \quad (2)$$

where S is the strain image before normalization, \bar{S} is the mean value of the strain image, and σ_S is the standard deviation of the strain image.

2.8. Volume rendering

The rendering stage could be divided into three steps. Firstly, only in the case of 3D endogenous strain, some of the obtained images are acquired in a phase when the deformation is very low and insufficient to obtain a contrast between inclusion and the background. Such low signal-to-noise ratio (SNR) elastograms must be filtered out because otherwise the 3D volume will be affected by low SNR images. It is the fundamental limitation of the technique using harmonic natural strain, which sometimes becomes very close or even equal to zero. The empirical threshold was used for this purpose, all strain elastograms having mean strain lower than 0.03% were removed from the dataset. The threshold was established after visual revision of the sequences of strain images, when the majority of the selected frames, exceeding the threshold, provided a contrast for inclusions.

Secondly, the selected images (B-scans or strains elastograms) are arranged spatially according to the position sensor readings ($x_s, y_s, z_s, \alpha, \beta, \gamma$). The 3D arrangement was done according to coordinates transformed by equation (LEE *et al.*, 2018):

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} + \mathbf{C} \cdot \mathbf{R}(\alpha, \beta, \gamma) \cdot \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}, \quad (3)$$

where (x_t, y_t, z_t) is the image point coordinates in the arranged 3D space, (x_s, y_s, z_s) is the position coordinates measured by the electromagnetic sensor (incorporated in the linear array), \mathbf{C} is the calibration matrix for electromagnetic sensor and linear array setup (values provided by vendor), $\mathbf{R}(\alpha, \beta, \gamma)$ is the rotation matrix calculated (Eq. (4)) from the angles measured by an electromagnetic sensor, (x_p, y_p, z_p) is the point position in the image (in a coordinate system of linear array) before transform, p is the index of pixels in image data. B-scan images were of 616×820 size and strain elastograms were the same as the RF data frame in size. The rotation matrix is expressed as follows:

$$\mathbf{R}(\alpha, \beta, \gamma) = \begin{bmatrix} \cos(\alpha) \cdot \cos(\beta) & \cos(\alpha) \cdot \sin(\beta) \cdot \sin(\gamma) - \sin(\alpha) \cdot \cos(\gamma) & \cos(\alpha) \cdot \sin(\beta) \cdot \cos(\gamma) + \sin(\alpha) \cdot \sin(\gamma) \\ \sin(\alpha) \cdot \cos(\beta) & \sin(\alpha) \cdot \sin(\beta) \cdot \sin(\gamma) + \cos(\alpha) \cdot \cos(\gamma) & \sin(\alpha) \cdot \sin(\beta) \cdot \cos(\gamma) - \cos(\alpha) \cdot \sin(\gamma) \\ -\sin(\beta) & \cos(\beta) \cdot \sin(\gamma) & \cos(\beta) \cdot \cos(\gamma) \end{bmatrix}, \quad (4)$$

where α – azimuth, β – elevation, and γ – roll angles. So, the points with coordinates (x_p, y_p, z_p) from the B-scan plane are arranged at coordinates (x_t, y_t, z_t) in 3D space.

Thirdly, the arranged planes are mapped to a regular rectangular grid representing voxels. Standard MATLAB processing routine (griddata function) was used for the purpose of applying cubic interpolation for the irregularly spaced image data. The dimensions of the voxel were set to be $0.25 \times 0.25 \times 0.25 \text{ mm}^3$. The obtained 3D matrixes were filtered by the smoothing moving average filter ($9 \times 9 \times 9$ voxels).

2.9. Extraction of inclusion isosurface

Isosurfaces of the phantom inclusions were extracted for the visual assessment and quantitative comparison for both parametric 3D images (obtained from the B-mode and normalized strain). The marching cube algorithm (LORENSEN, CLINE, 1987) was used for extraction. The algorithm requires to set a threshold isovalue for the extraction of the volume enclosed by the isosurface. The threshold values were determined by using these criteria:

- for the B-mode the threshold was set above majority of randomly distributed parasitic reflections which occurs in the background region;
- threshold values were adjusted looking for the best correlation according to quantitative metrics between strain and reflection 3D isosurfaces;
- extracted isosurfaces were evaluated visually and the surface reconstruction was accepted only if there were no strong shape distortions, even if the quantitative results were better.

Threshold values for the B-mode: soft inclusion = 127, hard = 140; for normalized strain images: soft = 0.35 and hard = 0.63 were established. An example illustrating the selected threshold values and the resulting boundaries for these values (red contour) of the inclusions for both types (the B-mode and strain) of images is presented in Fig. 3.

2.10. Quantitative comparison

After volume rendering and isosurface extraction, two types of 3D images were obtained: first reconstructed from the B-mode data and based on acoustical property – reflections, and second reconstructed from the calculated strain elastograms and representing the mechanical property – stiffness. The 3D isosurfaces were superimposed. Five quantitative metrics were used for the assessment of the correspondence between 3D image of the reflections and the 3D image of strain:

- volumes calculated by summing voxels enclosed by isosurface;

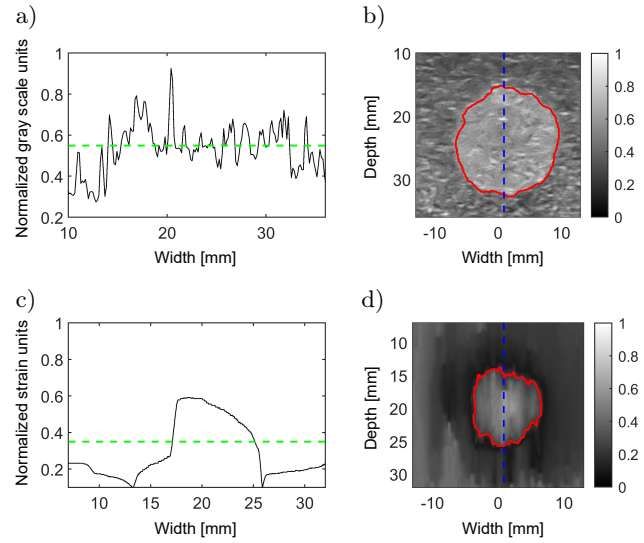


Fig. 3. Threshold selection examples for B-mode image (a–b) and strain elastogram (c–d): a) shows single normalized gray scale intensity line crossing the middle of inclusion (marked by dashed blue line in part (b) superimposed on B-mode image) and threshold level is shown by green dashed line; b) shows B-mode image together with a contour (red color) extracted by using the threshold; c) and d) show exactly the same just for strain elastogram case.

- diameters of the inclusions estimated by the manual measurement (three repetitive measurements);
- lengths of the inclusions estimated by the manual measurement (three repetitive measurements);
- the Jaccard similarity coefficient, which is expressed as follows:

$$\text{Jaccard} = \frac{|A \cap B|}{|A \cup B|}, \quad (5)$$

where A and B are the volumes obtained from images having different properties (the B-mode and endogenous normalized strain);

- the Hausdorff distance – which measures the largest distance between paired points of the extracted isosurfaces. It is found by using the following Eqs. (6) and (7). Firstly, the one-sided Hausdorff distance is found for both isosurfaces:

$$h(\mathbf{A}, \mathbf{B}) = \max_{\mathbf{a} \in \mathbf{A}} \min_{\mathbf{b} \in \mathbf{B}} \|\mathbf{a} - \mathbf{b}\|, \quad (6)$$

here \mathbf{A} and \mathbf{B} are the points cloud coordinates (x, y, z) of the extracted isosurfaces obtained from images of different properties (the B-mode and endogenous normalized strain), $\mathbf{a} = \{x_1, y_1, z_1, \dots, x_n, y_n, z_n\}$, and $\mathbf{b} = \{x_1, y_1, z_1, \dots, x_m, y_m, z_m\}$, m, n are the numbers of points in data clouds. Next, the bidirectional Hausdorff distance H is calculated by Eq. (7):

$$H(\mathbf{A}, \mathbf{B}) = \max(h(\mathbf{A}, \mathbf{B}), h(\mathbf{B}, \mathbf{A})). \quad (7)$$

3. Results

At first, the excitation waveform of endogenous phantom displacements was evaluated from the first experiment when the array transducer was fixed during RF recording. Figure 4 presents the summed mean displacement versus time.

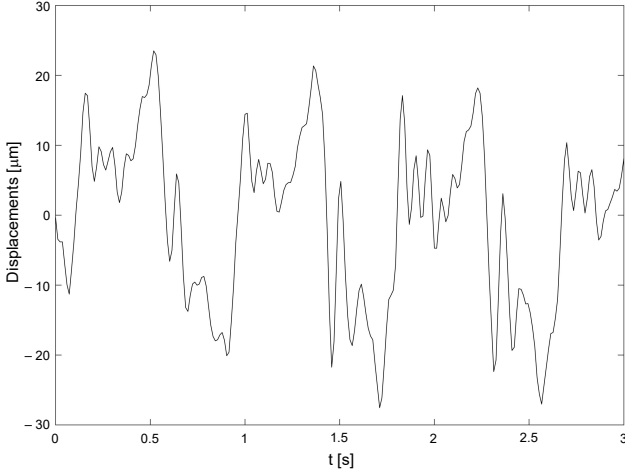


Fig. 4. Comparison of averaged in-space summed phantom displacements (3.5 periods, first 3 seconds). The displacements were estimated from the experiment with fixed array transducer.

The periodicity of the waveform is noticeable. The amplitude of frame-averaged displacements reaches 20 μm . The period of the waveform of the induced displacements matches the rate of the ventilation apparatus (70 times per minute).

Figure 5 presents the examples of normalized 2D strain images (S_N) obtained in the second experiment (freehand 3D scanning) at the time of maximum normalized mean strain.

It could be seen that both inclusions (hard – 75 kPa and soft – 7 kPa) could be clearly identified. The contrast between inclusion and background was a bit larger for a soft inclusion. It is observable that the inclusions are slightly deformed and non-circular, but the contrasts are sufficient for imaging purposes. However, after a visual analysis of all (of time instances) strain images (S_N) in the 3D dataset, it was found that only about 25% of the 2D strain images provided a sufficient contrast (mean strain of the frame $> 0.03\%$) for 3D rendering. Therefore, the mean strain parameter was proposed as a criterion to filter low contrast strain images and remove about 75% of them from the 3D dataset, before 3D rendering.

The cross-sectional images of both inclusions of the phantom rendered volumes are presented in Fig. 6.

Figures 7 and 8 present the extracted isosurfaces of cylindrical inclusions.

Figure 7 shows the inclusions extracted from the 3D datasets of the imitated endogenous strain. It could be observed that the inclusions do not follow the exact

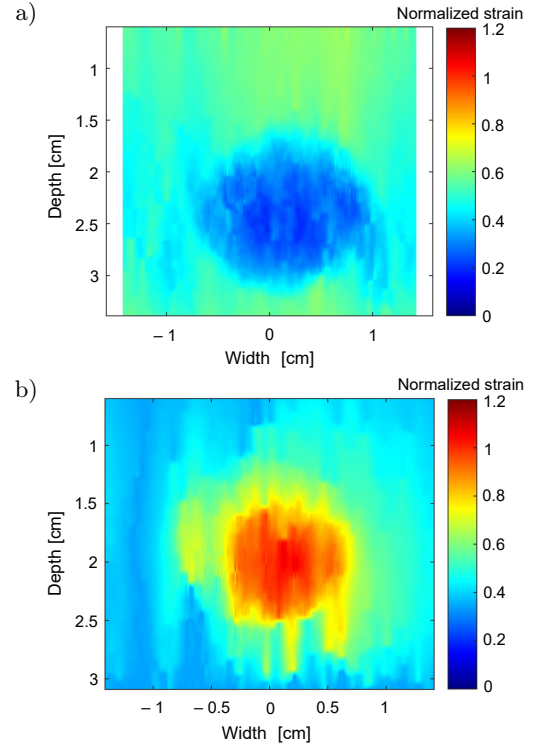


Fig. 5. The 2D examples of normalized strain elastograms of the phantom containing (a) hard inclusion and (b) soft inclusion.

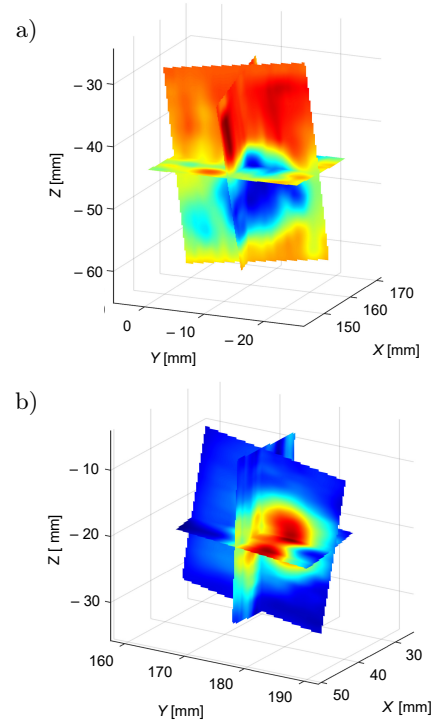


Fig. 6. Cross-sectional images of 3D strain elastograms for the phantom containing (a) hard inclusion and (b) soft inclusion.

shape of the cylinder, but at the same time they are relatively smooth.

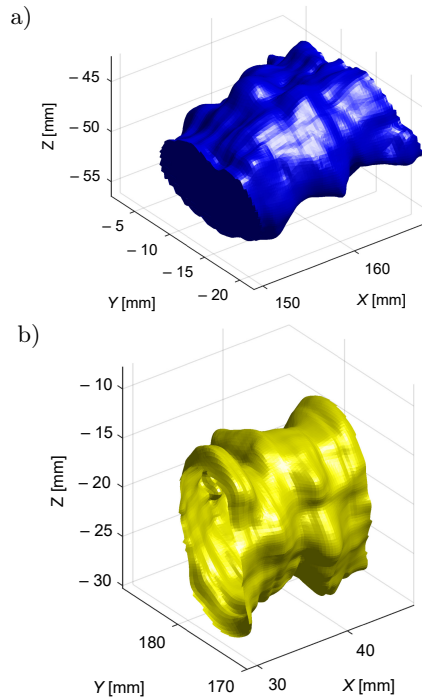


Fig. 7. Results of 3D isosurface imaging for (a) hard and (b) soft inclusions (images obtained from strain elastography data).

A comparison between the isosurfaces extracted from the B-mode image dataset and normalized strain images (Fig. 8) showed that the volume of strain-based 3D image was underestimated. For both inclusions, the volumes were smaller in the case of rendering from strain images. The hard inclusion in Fig. 8a was less correlated with the reference obtained from the B-mode image dataset, in comparison to the soft one (Fig. 8b).

The quantitative results of comparison of inclusions obtained by 3D imaging of reflections and strain are presented in Table 1.

The volumes obtained from strain elastograms were smaller in comparison to those obtained from the B-mode (remarkably about 30% for the hard inclusion and a bit less for the soft one ~15%). The correspondence between the volume of inclusion based on B-mode images and the volume based on strain elastography was good (the Jaccard coefficient within the range 0.64–0.74).

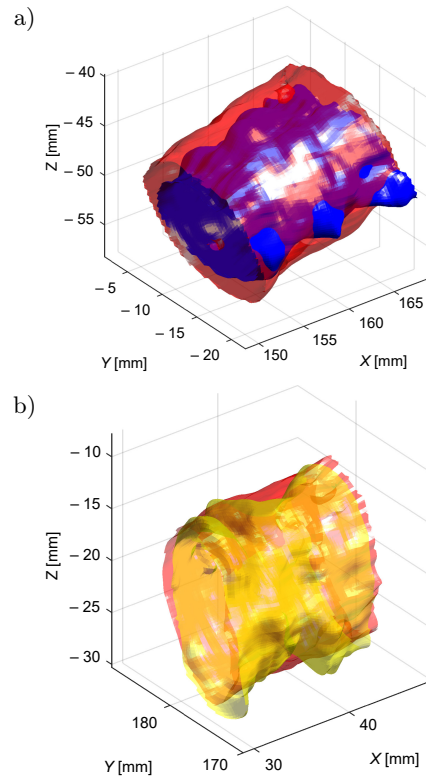


Fig. 8. Comparison of isosurfaces obtained from volumes reconstructed from B-mode (red) and normalized strain images for (a) hard (blue) and (b) soft (yellow) inclusions.

The comparison of diameters and lengths showed that the diameter value is less correlated with the measured from the isosurface obtained from B-mode images than the length for both inclusions. The Jaccard similarity coefficient and the Hausdorff distance shows that the isosurface of the soft inclusion was in better correlation to the isosurface obtained from the B-mode dataset.

4. Discussion

The obtained results confirm the endogenous motion-based elastography as a feasible technique. The 3D freehand visualization of the inclusions by the strain was successful. Thresholding of strain dataset proved as sufficient with elastograms of at least 0.03% of the mean strain. To the best of our knowledge, there

Table 1. Quantitative comparison of obtained dimensions and volumes.

		Volume rendered from 3D data [mm ³]	Diameter from 3D data [mm]	Length [mm]	Diameter from static data [mm]	Jaccard similarity coefficient	Hausdorff distance [mm]
Hard inclusion [75 kPa]	Reflections	2832.2	13.92 ± 1.65	14.10 ± 0.47	16.30 ± 0.40	0.64	4.85
	Strain	1970.0	13.02 ± 2.66	15.53 ± 0.21	15.98 ± 1.27		
Soft inclusion [7 kPa]	Reflections	2744.1	15.30 ± 1.77	16.08 ± 0.29	15.30 ± 1.20	0.74	3.77
	Strain	2361.5	14.46 ± 2.94	15.26 ± 0.48	12.34 ± 1.20		

is no published data about the results of endogenous strain 3D imaging, so it is hard to discuss in a direct comparison and to make a comparison with an existing technique. The range of strain for 3D imaging was evaluated by HAVRE *et al.* (2017), but for the case of real-time elastography. In their study, sufficient strains levels were found in range of approximately 0.1% to 2% for the by hand deformations of phantom. Our in vitro pilot achievement is obtained not by external sources of deformations, but by internal sources and with an even smaller strain of only $>0.03\%$. Our findings show that soft inclusion was in better correspondence with the reference obtained from B-scans in comparison to hard inclusion. However, the discrepancy of quantitative estimates shows the limitations of our method studied. In general, the freehand 3D imaging with ultrasound is found to result in discrepancies as reviewed by MOZAFFARI and LEE (2017). They reviewed that with the electromagnetic positioning system the volume estimation accuracy errors from 1.1% to 5.44% are possible. These accuracies are reviewed for the cases with B-scan images obtained in vitro. In our case, only discrepancies of estimates between B-scan images and strain images are possible, as we used in-lab-made phantoms of non-verified dimensions. Both inclusions were found to be smaller when the isosurfaces and volumes extracted from the reflection image data and the strain image data were compared. The volumes of elasticity obtained were smaller by up to 30% compared to volumes reconstructed with B-scan. This total discrepancy includes the level of about 5% of volume estimation accuracy that is a common feature for 3D ultrasonic imaging with an electromagnetic positioning system. So about 25% discrepancy we attribute to the strain image reconstruction. Strain volumes were estimated to be smaller than volumes reconstructed from B-scans, and this is contrary to (LEE *et al.*, 2018). The 3D freehand ultrasound elastography study, where the measured volumes were highly correlated with reference values in a commercial phantom. The absolute values of the volumes and dimensions (provided in Table 1) used for comparison were estimated from the measurements based on the scanning data. Although ultrasound velocity in the phantom base material and inclusions might be different and not exactly known, the scanner setting for velocity is standard – 1540 m/s, as always in clinical practice. Therefore, the estimates calculated assuming the uniform speed 1540 m/s, have deviation from the real values. From B-scan images, obtained with a fixed probe, diameters of inclusions, we found a little distorted from a round shape: the hard inclusion dimensions in X and Y directions were 16.3 ± 0.4 mm, that is 2.5% distortion from the round shape, and for soft inclusion 15.3 ± 1.2 mm the 7.8% distortion from the round shape. The variations of a sound speed in agar mixtures of the similar magnitude were found

(BURLEW *et al.*, 1980; MADSEN *et al.*, 1998). So, we attribute the diameter discrepancies to uncertainty of the sound speed in agar mixtures. Distortions from the round shape of inclusion diameters in strain images are much more significant, that we originate from strain imaging peculiarities which have not been discovered yet.

The preparation of inclusions intended a threefold difference in stiffness of agar mixtures. The soft inclusion was 7 kPa versus 25 kPa stiffness of the background material, or 3.6 times difference. This possibly resulted the contrast between inclusion and background was a bit larger for soft inclusion in 2D elastograms. The hard inclusion was 75 kPa versus 25 kPa, so the stiffness ratio was 3 times.

The endogenous motion is relatively weak in comparison to compressions induced by the ultrasound probe, but has potential for the assessment of deeper layers of tissue. It should be noted that the amplitude of the imitated displacements reached only 20 μm (Fig. 4), meanwhile we have found it to be close to 100 μm in the specific segment of the liver, close to the beating heart, in a previous study (SAKALAUSKAS *et al.*, 2016). The better contrast of the strain elastogram could be expected for larger displacement amplitude. Researchers (LEE *et al.*, 2018) have indicated that the 3D freehand elastography using probe compressions suffers from respiratory and cardiac motion artefacts, which is not the case for the presented technique where the cardiac activity will serve as an excitation source.

5. Limitations of the study

We have identified few limitations of the study. Firstly, the inclusions in the laboratory-prepared phantom had a relatively high stiffness difference in comparison to the background (75 kPa and 7 kPa in contrast to 25 kPa). The approach of phantom preparation requires a relatively stiff background material that can hold its form. The very soft (less 25 kPa) background material does not allow the implementation of inclusions as hanging structures above the rubber vessel. We have plans to manufacture a phantom of a different structure, having the soft background material and inclusions closer to the background in stiffness. In future studies, the new structure phantom would enable evaluating the potential stiffness contrasting possibilities of the technique. The presented study was just a first step proof of a concept, confirming the method as promising.

Secondly, the trajectory of a freehand probe motion in our case was approximately linear. Only axial displacements were evaluated in our study, the lateral and respectively strain in that direction do not provide sufficient contrast for 3D rendering. Using other, more complicated trajectories possibly will introduce differ-

ent 3D reconstruction errors, although the 3D rendering method (Eqs. (3) and (4)) allows free trajectories of any shape. We hypothesize, that the lateral sampling (more scanning lines) could be employed in the future studies. As well the lateral displacements, that are possible to detected with the GLUE algorithm, could provide an extra contrast in the elastogram from lateral strain.

Thirdly, the low strain elastograms were filtered by using a relatively simple technique – the mean strain value threshold, and it was noticed that some elastograms selected after filtering do not provide high contrast of the inclusion. These elastograms affected the final precision of the isosurface of the inclusion dimensions and the volume estimates. We have shown the technique to be feasible, however, we have to admit that the endogenous motion of the cylindrical shape of source is still too simple an approximation of the in vivo case. The properties of rubber vessels were not controlled in the experiments, while the stiffness of blood vessels has its own physiological behavior in vivo. Time waveforms of endogenous displacements and their frequency spectrum features which have been found related to a waveform pattern of an excitation pressure. The surrogate waveform induced in the phantom needs adjustments to the waveform features of in vivo displacements. Thus, living tissue requires further investigation, which can allow for a more reliable selection of frames for 3D rendering. It was found that the low mean strain of the frame is not always followed by a low contrast elastogram and vice versa. Such findings are still difficult to explain and will be investigated more widely in future studies. Repeatability of the method was not evaluated. However, we still recall that this pilot study was conducted just to evaluate the feasibility of performing 3D imaging of endogenous strain.

6. Conclusion

The results obtained confirm the endogenous motion-based 3D elastography as a feasible technique. The freehand 3D visualization of the inclusions by the strain was successful. However, quantitative measurements showed that the accuracy of the method in volumetric measurements is limited.

Declarations

The authors declare that they have no conflicts of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

Acknowledgments (funding)

This work was funded by Research Council of Lithuania in the frame of Researchers group project “Investigation of radiofrequency based ultrasonic

strain elastography algorithms (ELASTUS)” Reg. No.: S-MIP-19-8.

References

1. ABEYSEKERA J., ROHLING R., SALCUDAN S. (2015), Vibro-elastography: Absolute elasticity from motorized 3D ultrasound measurements of harmonic motion vectors, [in:] *2015 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, doi: 10.1109/ULTSYM.2015.0201.
2. BAE U., DIGHE M., SHAMDASANI V., MINOSHIMA S., DUBINSKY T., KIM Y. (2006), 6F-6 thyroid elastography using carotid artery pulsation: A feasibility study, [in:] *2006 IEEE Ultrasonics Symposium*, pp. 614–617, doi: 10.1109/ULTSYM.2006.160.
3. BARR R.G. *et al.* (2017), WFUMB guidelines and recommendations on the clinical use of ultrasound elastography: Part 5. Prostate, *Ultrasound in Medicine and Biology*, **43**(1): 27–48, doi: 10.1016/j.ultrasmedbio.2016.06.020.
4. BERCOFF J., SINKUS R., TANTER M., FINK M. (2004), 3D ultrasound-based dynamic and transient elastography: First in vitro results, [in:] *IEEE International Ultrasonics Symposium*, **1**: 28–31, doi: 10.1109/ULTSYM.2004.1417660.
5. BURLEW M.M., MADSEN E.L., ZAGZEBSKI J.A., BANJAVIC R.A., SUM S.W. (1980), A new ultrasound tissue-equivalent material, *Radiology*, **134**(2): 517–520, doi: 10.1148/radiology.134.2.7352242.
6. CHEN Z., CHEN Y., HUANG Q. (2016), Development of a wireless and near real-time 3D ultrasound strain imaging system, *IEEE Transactions on Biomedical Circuits and Systems*, **10**(2): 394–403, doi: 10.1109/TBCAS.2015.2420117.
7. COSGROVE D. *et al.* (2017), WFUMB guidelines and recommendations on the clinical use of ultrasound elastography: Part 4. Thyroid, *Ultrasound in Medicine & Biology*, **43**(1): 4–26, doi: 10.1016/j.ultrasmedbio.2016.06.022.
8. DICKINSON R.J., HILL C.R. (1982), Measurement of soft tissue motion using correlation between A-scans, *Ultrasound in Medicine & Biology*, **8**(3): 263–271, doi: 10.1016/0301-5629(82)90032-1.
9. DIETRICH C.F. *et al.* (2017), Strain elastography – How to do it?, *Ultrasound International Open*, **3**(4): E137–E149, doi: 10.1055/s-0043-119412.
10. GELMAN S. *et al.* (2020), Endogenous motion of liver correlates to the severity of portal hypertension, *World Journal of Gastroenterology*, **26**(38): 5836–5848, doi: 10.3748/wjg.v26.i38.5836.
11. GILJA O.H., HAUSKEN T., OLAFSSON S., MATRE K., ØDEGAARD S. (1998), In vitro evaluation of three-dimensional ultrasonography based on magnetic scan head tracking, *Ultrasound in Medicine & Biology*, **24**(8): 1161–1167, doi: 10.1016/S0301-5629(98)00098-2.
12. HALL T.J., BILGEN M., INSANA M.F., KROUSKOP T.A. (1997), Phantom materials for elastography, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **44**(6): 1355–1365, doi: 10.1109/58.656639.

13. HASHEMI H.S., RIVAZ H. (2017), Global time-delay estimation in ultrasound elastography, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **64**(10): 1625–1636, doi: 10.1109/TUFFC.2017.2717933.
14. HAVRE R.F. *et al.* (2008), Freehand real-time elastography: Impact of scanning parameters on image quality and in vitro intra- and interobserver validations, *Ultrasound in Medicine & Biology*, **34**(10): 1638–1650, doi: 10.1016/j.ultrasmedbio.2008.03.009.
15. HENDRIKS G.A.G.M., HOLLÄNDER B., MENSSEN J., MILKOWSKI A., HANSEN H.H.G., DE KORTE C.L. (2016), Automated 3D ultrasound elastography of the breast: A phantom validation study, *Physics in Medicine & Biology*, **61**(7): 2665–2679, doi: 10.1088/0031-9155/61/7/2665.
16. HOUSDEN R.J., GEE A.H., TREECE G.M., PRAGER R.W. (2010), 3-D ultrasonic strain imaging using freehand scanning and a mechanically-swept probe – correspondence, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **57**(2): 501–506, doi: 10.1109/TUFFC.2010.1431.
17. HUANG Q., XIE B., YE P., CHEN Z. (2015), Correspondence – 3-D ultrasonic strain imaging based on a linear scanning system, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **62**(2): 392–400, doi: 10.1109/TUFFC.2014.006665.
18. KOLEN A.F., MILLER N.R., AHMED E.E., BAMBER J.C. (2004), Characterization of cardiovascular liver motion for the eventual application of elasticity imaging to the liver in vivo, *Physics in Medicine & Biology*, **49**(18): 4187–4206, doi: 10.1088/0031-9155/49/18/001.
19. LEE F.-F., HE Q., LUO J. (2018), Electromagnetic tracking-based freehand 3D quasi-static elastography with 1D linear array: A phantom study, *Physics in Medicine & Biology*, **63**(24): 245006, doi: 10.1088/1361-6560/aaefae.
20. LINDOP J.E., TREECE G.M., GEE A.H., PRAGER R.W. (2006), 3D elastography using freehand ultrasound, *Ultrasound in Medicine & Biology*, **32**(4): 529–545, doi: 10.1016/j.ultrasmedbio.2005.11.018.
21. LORENSEN W.E., CLINE H.E. (1987), Marching cubes: A high resolution 3D surface construction algorithm, [in:] *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, **21**(4): 163–169, doi: 10.1145/37401.37422.
22. LUO S., KIM E.-H., DIGHE M., KIM Y. (2009), Screening of thyroid nodules by ultrasound elastography using diastolic strain variation, [in:] *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4420–4423, doi: 10.1109/IEMBS.2009.5332744.
23. MADSEN E.L., FRANK G.R., DONG F. (1998), Liquid or solid ultrasonically tissue mimicking materials with very low scatter, *Ultrasound in Medicine & Biology*, **24**(4): 535–542, doi: 10.1016/S0301-5629(98)00013-1.
24. MAI J.J., INSANA M.F. (2002), Strain imaging of internal deformation, *Ultrasound in Medicine & Biology*, **28**(11–12): 1475–1484, doi: 10.1016/S0301-5629(02)00645-2.
25. MOZAFFARI M.H., LEE W.-S. (2017), Freehand 3-D ultrasound imaging: A systematic review, *Ultrasound in Medicine & Biology*, **43**(10): 2099–2124, doi: 10.1016/j.ultrasmedbio.2017.06.009.
26. RICHARDS M.S., BARBONE P.E., OBERAI A.A. (2009), Quantitative three-dimensional elasticity imaging from quasi-static deformation: A phantom study, *Physics in Medicine & Biology*, **54**(3): 757–779, doi: 10.1088/0031-9155/54/3/019.
27. RIVAZ H., BOCTOR E.M., CHOTI M.A., HAGER G.D. (2011), Real-time regularized ultrasound elastography, *IEEE Transactions on Medical Imaging*, **30**(4): 928–945, doi: 10.1109/TMI.2010.2091966.
28. SAKALAUSKAS A., JURKONIS R., GELMAN S., LUKOŠEVIČIUS A., KUPČINSKAS L. (2016), Initial results of liver tissue characterization using endogenous motion tracking method, [in:] *Conference Proceedings Biomedical Engineering*, pp. 132–137, <http://biomed.ktu.lt/index.php/BME/article/viewFile/3392/150>.
29. SAKALAUSKAS A., JURKONIS R., GELMAN S., LUKOŠEVIČIUS A., KUPČINSKAS L. (2018), Development of radiofrequency ultrasound based method for elasticity characterization using low frequency endogenous motion: Phantom study, [in:] *IFMBE Proceedings*, Eskola H., Väisänen O., Viik J., Hyttinen J. [Eds.], Springer Singapore, pp. 474–477, doi: 10.1007/978-981-10-5122-7_119.
30. SAKALAUSKAS A., JURKONIS R., GELMAN S., LUKOŠEVIČIUS A., KUPČINSKAS L. (2019), Investigation of radiofrequency ultrasound-based fibrotic tissue strain imaging method employing endogenous motion: Endogenous motion-based strain elastography, *Journal of Ultrasound in Medicine*, **38**(9): 2315–2327, doi: 10.1002/jum.14925.
31. TRISTAM M., BARBOSA D.C., COSGROVE D.O., NASIRI D.K., BAMBER J.C., HILL C.R. (1986), Ultrasonic study of in vivo kinetic characteristics of human tissues, *Ultrasound in Medicine & Biology*, **12**(12): 927–937, doi: 10.1016/0301-5629(86)90061-x.
32. WANG Y., SPANGLER C.H., TAI B.L., SHIH A.J. (2013), Positional accuracy and transmitter orientation of the 3D electromagnetic tracking system, *Measurement Science and Technology*, **24**(10): 105105–105113, doi: 10.1088/0957-0233/24/10/105105.
33. WELLS P.N.T., LIANG H.-D. (2011), Medical ultrasound: Imaging of soft tissue strain and elasticity, *Journal of the Royal Society Interface*, **8**(64): 1521–1549, doi: 10.1098/rsif.2011.0054.
34. WILSON L.S., ROBINSON D.E. (1982), Ultrasonic measurement of small displacements and deformations of tissue, *Ultrasonic Imaging*, **4**(1): 71–82, doi: 10.1016/0161-7346(82)90006-2.
35. ZAMBACEVIČIENĖ M., JURKONIS R., GELMAN S., SAKALAUSKAS A. (2019), RF ultrasound based estimation of pulsatile flow induced microdisplacements in phantom, [in:] *World Congress on Medical Physics and Biomedical Engineering 2018*, Lhotska L., Sukupova L., Lacković I., Ibbott G.S. [Eds.], Springer Singapore, pp. 601–605, doi: 10.1007/978-981-10-9035-6_112.

Research Paper

Effect of Rotation on the Piezoelectric Wave Impedance Characteristics

Xiaoguang YUAN*, Chaoyu HAO, Quan JIANG

*School of Transportation and Civil Engineering, Nantong University
Nantong, China*

*Corresponding Author e-mail: xgyuan@gmail.com

(received May 24, 2022; accepted November 20, 2022)

The dependence of piezoelectric wave impedance on the rotation speed is investigated theoretically and numerically. The Coriolis force due to rotation is introduced into the piezoelectric motion equations, which is solved by the harmonic plane wave solution. It is shown that the wave impedance variations of longitudinal and transverse waves due to rotation are clearly different. The longitudinal wave impedance continuously increases with a small rotation ratio and one transverse wave impedance is almost irrespective of a rotation ratio. In contrast, the rotation applies a big impact on the other transversal wave impedances in the piezoelectric crystal which decreases monotonically with the rotation speed. Such characteristics are significant in piezoelectric transducers and sensors.

Keywords: wave impedance; Coriolis acceleration; piezoelectric crystal; wave velocity.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Under ordinary conditions, the solid-to-gas ultrasonic transmission encounters the problem of large acoustic impedance mismatch and makes it difficult to transfer the ultrasonic wave energy commutatively.

It is due to the close to zero transmission coefficient of an equation expression in the form of wave impedance, for the linear wave field, defined by dividing wave pressure with the particle velocity at some position. In an unbounded medium, the wave impedance is independent of the position and is numerically equal to the product of density and wave velocity, i.e., $Z = \rho v$ [$\text{Pa} \cdot \text{sec} \cdot \text{m}^{-1}$] known as “the specific or characteristic wave impedance”.

Consequently in the gas sensor field, it is difficult to use directly the piezoelectric crystal device to detect gas features, etc., (LYNNWORTH, 1965). One may turn to the medium of slower density, like P(VDF/TrFE) (TAKAHASHI, OHIGASHI, 2009), or other technologies, i.e., adding the chemical reaction film (WANG *et al.*, 2021), the transition layer (SURAPPA *et al.*, 2018), or a resonance cavity (DONG *et al.*, 2003) for the specific kind of gas.

In essence, the fundamental solution is to slow the piezoelectric wave impedance in order to interact between solid and gas, which is a problem of equal scientific and practical significance as well as helpful for the above technology.

Unfortunately no one attempts to slow the wave velocity to obtain smaller wave impedance in the acoustic field. In the field of solid mechanics, the authors of references (YUAN, 2016; 2019; YUAN, JIANG, 2017; YUAN *et al.*, 2016; 2020; YUAN, LI, 2015a; 2015b) investigated the propagation, refraction and transmission processes of bulk waves in the rotating piezoelectrics and pyroelectrics, from which it is found that the Coriolis force due to rotation can change the bulk wave velocities.

Accordingly, the objective of the present study consists in investigating the piezoelectric wave impedance characteristics in the acoustic field using the Coriolis force.

2. Problem formulation and results

The object of investigation is an infinite piezoelectric medium (Fig. 1), in which the computational coordinates (x_1, x_2, x_3) , n_i , θ , and Ω_i are indicated.

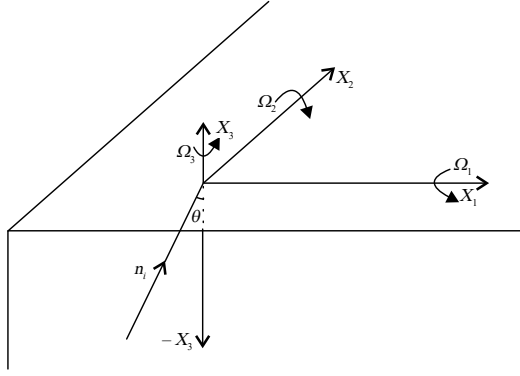


Fig. 1. Wave propagation in the infinite piezoelectric crystal: n_i wave propagation vector, θ wave propagation angle, Ω_i rotation speed vector.

We introduce the Coriolis force K_i due to rotation (SUCHKOV *et al.*, 2011; YUAN, JIANG, 2017)

$$K_i = 2\rho\varepsilon_{irk}\Omega_r \frac{\partial u_k}{\partial t} \quad (1)$$

to the piezoelectric motion equation and obtain:

$$C_{ijkl}u_{k,li} + e_{kij}\phi_{,ki} = \rho \left[\frac{\partial^2 u_j}{\partial t^2} + 2\varepsilon_{jik}\Omega_i \frac{\partial u_k}{\partial t} \right], \quad (2)$$

$$-\epsilon_{ij}\phi_{,ji} + e_{ikl}u_{k,li} = 0,$$

where ρ represents mass density, t is time, u_j is the displacement vector, ϕ is the electric potential, ε_{jik} , C_{ijkl} , e_{kij} , ϵ_{ij} indicate the tensors of permutation, elasticity, piezoelectricity, and permittivity constant. Ω_i is the rotation speed vector whose unit is the same with the wave frequency ω , so the dimensionless quantity of the rotation ratio $\eta_i = \Omega_i/\omega$ is used in the following.

Without loss of generality, we seek a harmonic plane wave solution in the form:

$$\{u_1, u_2, u_3, \phi\} = \{U_1, U_2, U_3, \Phi\} e^{I\omega(s_i x_i - t)}, \quad (3)$$

where s_i is the slowness vector, ω is the wave frequency and I equals to $\sqrt{-1}$. (U_i, Φ) are amplitudes of the displacement and the electric potential $\{u_1, u_2, u_3, \phi\}$.

After substituting Eq. (3) in (2), we obtain:

$$\begin{bmatrix} s_l s_i \overset{\circ}{C}_{irkl} - \rho \delta_{rk} - 2I\varepsilon_{rik}\eta_i & s_j s_i \overset{\circ}{e}_{jir} \\ s_l s_i \overset{\circ}{e}_{ikl} & -s_j s_i \overset{\circ}{e}_{ijj} \end{bmatrix} \begin{Bmatrix} U_k \\ \Phi \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \quad (4)$$

where

$$s_i = n_i/v, \quad (5)$$

n_i denotes the wave propagation vector (Fig. 1), and v is the wave velocity. Replacing the slowness vector in Eq. (4) by Eq. (5) and considering the condition of non-vanishing (U_i, Φ) , we obtain an associated characteristic polynomial equation about v .

As an computational example, the material parameters of piezoelectric medium lithium niobate are listed in Table 1 and their Voigt notation matrices were presented in (LEDBETTER *et al.*, 2004).

The computation was carried out under the following conditions: the rotation ratio $\eta_i = \frac{\Omega_i}{\omega}$, in which the subscript i indicates that the rotation axis is the x_i axis; ω – wave frequency – ($= 1$ MHz) and the wave propagation vector of $(\sin \theta, 0, \cos \theta)$ lies on the x_1 – x_3 plane without loss of generality.

According to the computation results of Eqs. (4) and (5), it is found that when the rotation ratio about any axis is above 0.00235, the real part of velocity of the longitudinal wave vanishes, which implies that it does not propagate and becomes the harmonic form of vibration in the angular frequency ω .

Therefore, in Fig. 2, a dimensionless quantity of $Z(\eta_i = 0.002)/Z_0$ is defined as the ratio between wave impedances in the case of $\eta_i = 0.002$ and the case of $\eta_i = 0$.

Figure 2 exhibits the variations of wave impedance ratios of L, T1, and T2 waves as function of the propagation angle θ in the case of $\eta_1, \eta_2, \eta_3 = 0.002$, respectively. We see that the increases of L wave impedance due to rotation are different with θ especially in the case of η_3 ; in contrast, the wave impedance of T2 wave drops remarkably below the half of the case of without rotation almost at any θ . The feature of T1 wave is similar to T2 wave in the case of η_2 and shows a big change with θ in the case of η_3 .

Table 1. Lithium niobate (LiNbO₃) material parameters (JAMES, 1975).

Elasticity tensor C_{ij} [GPa]	C_{11}	C_{33}	C_{44}	C_{12}	C_{13}	C_{14}
	203	245	60	53	75	9
Piezoelectric stress tensor e_{ij} [C/m ²]	e_{15}	e_{22}	e_{31}	e_{33}		
	3.7	2.5	0.2	1.3		
Normalized permittivity tensor ϵ_{ij}/ϵ_0	ϵ_{11}/ϵ_0	ϵ_{33}/ϵ_0				
	44	29				
Vacuum permittivity ϵ_0 [F/m]	8.854×10^{-12}					
Density ρ [kg/m ³]	4700					

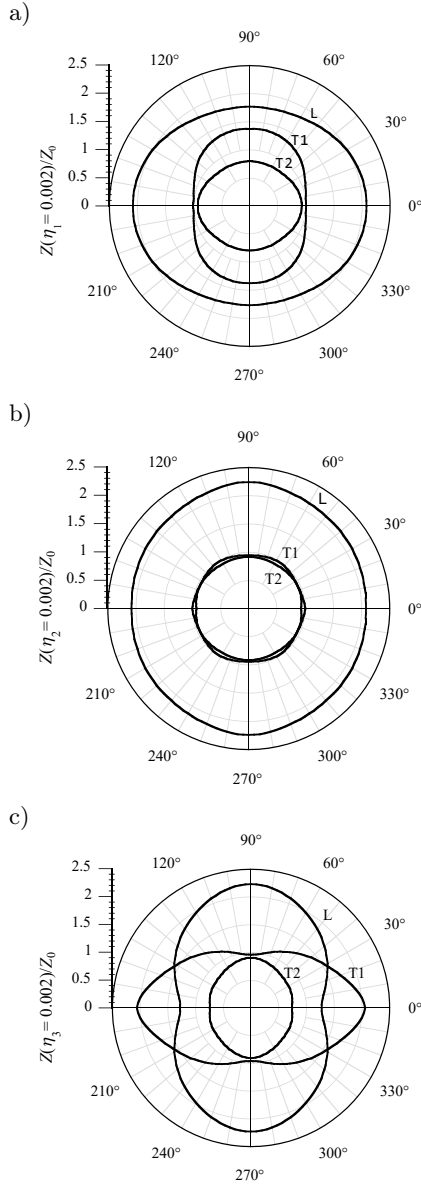


Fig. 2. Wave impedance ratio variations with wave propagation angle: a) indicates the case of rotation about the axis x_1 ; b) the axis x_2 ; c) the axis x_3 . L represents the longitudinal wave, and T1, T2 the transverse waves.

Next, setting $\theta = \frac{\pi}{2}$, we can investigate the wave impedances as a function of the rotation ratio as shown in Figs. 3, 4, and 5.

The L wave only exists for a low rotation ratio, i.e., $\eta < 0.00235$. The changes of L wave impedance initially are not obvious and then become large especially in the case of rotating about x_2 , x_3 as shown in Fig. 3.

Figure 4 shows the variation of the wave impedance ratio of T1 wave as a function of the rotation ratio. We see that T1 wave is not sensitive to rotation except for small η : the wave impedance ratio of η_1 grows rapidly to be in the vicinity of 1.4, and then remain unchanged; in the cases of η_2 and η_3 their wave impedances are similar.

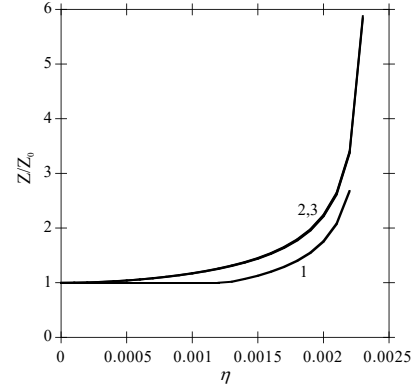


Fig. 3. Wave impedance ratio of L wave rotating about x_1 , x_2 , and x_3 axis, 1 indicates axis x_1 , 2 – x_2 , 3 – x_3 .

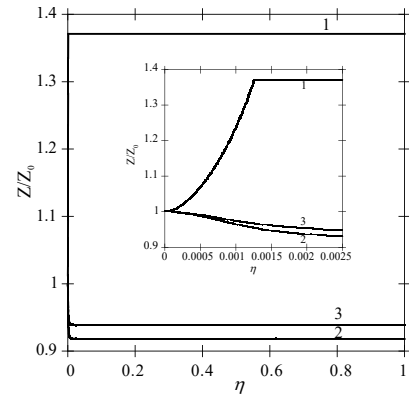


Fig. 4. Wave impedance ratio of T1 wave rotating about x_1 , x_2 , and x_3 axis, 1 indicates axis x_1 , 2 – x_2 , 3 – x_3 .

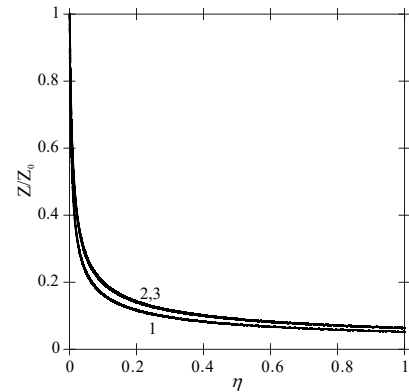


Fig. 5. Wave impedance ratio of T2 wave rotating about x_1 , x_2 , and x_3 axis, 1 indicates axis x_1 , 2 – x_2 , 3 – x_3 .

From Fig. 5, we see that the wave impedance ratio of T2 wave is a monotone decreasing function of the rotation ratio of η_1 , η_2 , and η_3 .

Analyzing the curves of Figs. 3–5, we assert the following: the numerical results presented indicate that the rotation influence on wave impedance is remarkable; however, wave impedance variations of longitudinal and transverse waves due to rotation differ distinctly. For instances, the longitudinal wave impedance continuously increases within a small rotation ratio.

For transverse waves, one transverse wave impedance is almost irrespective of the rotation ratio η except for the small interval of the rotation ratio η , out of which it acquires a horizontal line irrespective of the rotation ratio. On the other hand, T2 wave impedance of the other transverse wave between the interval of $\eta = 0$ and 1 drops monotonically (Fig. 5). When the rotation ratio η is 1, its wave impedance decreases by 95%. Thus, the application of this wave of a transverse mode is particularly advantageous to acquire small wave impedance in the associated technologies.

3. Conclusions

The wave impedances in the rotating piezoelectric crystal are studied theoretically and numerically. We note that, characteristically, the Coriolis force induced by rotation is justified to significantly alter wave impedances, particularly to that of the transverse wave. As the rotation increases, the wave impedance of T2 wave become small monotonically, this, in essence, increases the transmission coefficient and accordingly enhances the interaction between gas and the piezoelectric crystal. Such distinct attributes are very valuable from the viewpoint of the practical utilization, i.e., in the acoustic electronic devices like the piezoelectric transducer and acoustic sensors, which have never been reported and put into application before.

Acknowledgments

The author expresses his appreciation to Jiangsu Province Natural Science Foundation (BK20191450), Jiangsu Province Higher Education Institutions Natural Science Foundation (18KJB130004), and Nantong Basic Science Foundation (JC22022101).

Ethical standards and conflict of interest

The named authors confirm that all the research meets the ethical guidelines, including compliance with the legal requirements of the country of study. The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xiaoguang Yuan reports financial support was provided by Jiangsu Province Natural Science Foundation and Jiangsu Province Higher Education Institutions Natural Science Foundation.

References

1. DONG S., BAI F., LI J.-F., VIEHLAND D. (2003), A piezoelectric-sound-resonance cavity for hydrogen gas detection, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **50**(9): 1105–1113, doi: 10.1109/tuffc.2003.1235322.
2. JAMES B. (1975), Acoustic fields and waves in solids, volumes I and II: B.A. Auld. Published by J. Wiley and Sons, Ltd. 1973, Book review, *Applied Acoustics*, **8**(1): 72–73, doi: 10.1016/0003-682X(75)90008-0.
3. LEDBETTER H., OGI H., NAKAMURA N. (2004), Elastic, anelastic, piezoelectric coefficients of monocrystal lithium niobate, *Mechanics of Materials*, **36**(10): 941–947, doi: 10.1016/j.mechmat.2003.08.013.
4. LYNNWORTH L.C. (1965), Ultrasonic impedance matching from solids to gases, *IEEE Transactions on Sonics and Ultrasonics*, **12**(2): 37–48, doi: 10.1109/t-su.1965.29359.
5. SUCHKOV S.G., NIKOLAEVTSEV V.A., NIKITOV V. (2011), The influence of rotation on a phase deviation in surface acoustic wave devices, *Journal of Communications Technology and Electronics*, **56**: 1017, doi: 10.1134/S1064226911080110.
6. SURAPPA S., TAO M., LEVENT DEGERTEKIN F. (2018), Analysis and design of capacitive parametric ultrasonic transducers for efficient ultrasonic power transfer based on a 1-D lumped model, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **65**(11): 2103–2112, doi: 10.1109/tuffc.2018.2866058.
7. TAKAHASHI S., OHIGASHI H. (2009), Ultrasonic imaging using air-coupled P(VDF/TrFE) transducers at 2 MHz, *Ultrasonics*, **49**(4): 495–498, doi: 10.1016/j.ultras.2008.10.020.
8. WANG B., SUN L., SCHNEIDER-RAELOW M., LANG K.-D., NGO H.-D. (2021), Recent advances and challenges of nanomaterials-based hydrogen sensors, *Micromachines*, **12**(11): 1429, doi: 10.3390/mi12111429.
9. YUAN X. (2016), Effects of rotation and initial stresses on pyroelectric waves, *Archive of Applied Mechanics*, **86**(3): 433–444, doi: 10.1007/s00419-015-1038-z.
10. YUAN X. (2019), Electric bias dependence of piezoelectric wave reflection upon a rotating half-plane, *ZAMM – Journal of Applied Mathematics and Mechanics*, **99**(10): e201900021, doi: 10.1002/zamm.201900021.
11. YUAN X., JIANG Q. (2017), Reflection of plane waves from rotating pyroelectric half-space under initial stresses, *ZAMM – Journal of Applied Mathematics and Mechanics*, **97**(3): 365–374, doi: 10.1002/zamm.201500223.
12. YUAN X., JIANG Q., YANG F., (2016), Wave reflection and transmission in rotating and stressed pyroelectric half-planes, *Applied Mathematics and Computation*, **289**: 281–297, doi: 10.1016/j.amc.2016.05.016.
13. YUAN X., LI L. (2015a), Wave reflection and refraction in rotating and initially-stressed piezoelectric crystals, *Acta Mechanica*, **226**(10): 3243–3261, doi: 10.1007/s00707-015-1377-4.
14. YUAN X., LI L. (2015b), Waves in a rotating pyroelectric body, *Journal of Thermal Stresses* **38**(4): 399–414, doi: 10.1080/01495739.2015.1015838.
15. YUAN X., SHI Q., CAO Y. (2020), Constant electric bias dependence of wave propagation in a rotating piezoelectric crystal, *Acta Mechanica*, **231**(3): 1209–1215, doi: 10.1007/s00707-019-02596-4.

Research Paper

Research on Scattering Feature Extraction of Underwater Moving Cluster Targets Based on the Highlight Model

Yang YANG, Jun FAN*, Bin WANG

*Key Laboratory of Marine Intelligent Equipment and System of the Ministry of Education
Shanghai Jiao Tong University
Shanghai, China*

*Corresponding Author e-mail: fanjun@sjtu.edu.cn

(received October 5, 2022; accepted November 28, 2022)

In detecting cluster targets in ports or near-shore waters, the echo amplitude is seriously disturbed by interface reverberation, which leads to the distortion of the traditional target intensity characteristics, and the appearance of multiple targets in the same or adjacent beam leads to fuzzy feature recognition. Studying and extracting spatial distribution scale and motion features that reflect the information on cluster targets physics can improve the representation accuracy of cluster target characteristics. Based on the highlight model of target acoustic scattering, the target azimuth tendency is accurately estimated by the splitting beam method to fit the spatial geometric scale formed by multiple highlights. The instantaneous frequencies of highlights are extracted from the time-frequency domain, the Doppler shift of the highlights is calculated, and the motion state of the highlights is estimated. Based on the above processing method, target highlights' orientation, spatial scale and motion characteristics are fused, and the multiple moving highlights of typical formation distribution in the same beam are accurately identified. The features are applied to processing acoustic scattering data of multiple moving unmanned underwater vehicles (UUVs) on a lake. The results show that multiple small moving underwater targets can be effectively recognized according to the highlight scattering characteristics.

Keywords: motion small cluster targets; feature fusion; azimuth trend scale; azimuth trend Doppler.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

The early warning and defence of ports, islands, and large floating platforms is mainly conducted by small targets such as frogman teams and underwater robot clusters (JIANG *et al.*, 2009; HUANG *et al.*, 2020). In recent years, many countries have invested great efforts in studying the detection technology for small underwater targets. Diver detection sonar (DDS) is an important equipment for small underwater target detection (SARANGAPANI *et al.*, 2005; LO, FERGUSON, 2004). The Sentinel intruder detection sonar system by the American Sonardyne Company has a working frequency of 70 kHz, a bandwidth of 20 kHz, and an operating range of 600–900 m. The AquaShield detection sonar by the Israeli DSIT Company has a receiving array length of 1250 mm, a working frequency of 60 kHz, and an operating range of 700–1000 m. The X-type underwater monitoring system of the British

Oceanscan Company can automatically track multiple targets at the same time; the operating frequency is 100 kHz, the operating range is 500–1000 m, and the direction-finding accuracy is 1°. The centre frequency of most diver detection sonar is mainly concentrated in the range of 60~100 kHz, the bandwidth is 20 kHz, and the detection range is hundreds of metres to kilometres. Diver detection sonar is a high-resolution image sonar that takes small targets as detection objects. The higher the working frequency is, the more detailed the description of the targets; however, the higher the frequency is, the closer the operating distance and its design are restricted by the resolution and operating distance. In the measurement test of small targets at sea, at 100 kHz, the target intensity of diver is −25~−20 dB, and the average target intensity of the frogman's exhalations (bubbles) is −14 dB (HOLLETT *et al.*, 2006). In addition, the average travelling speed of the frogman under load is approximately 1.8 km/h,

and the driving speed of the frogman carrier under load can reach 4.5–5.4 km/h. “Tuna 9” is a hand-held unmanned vehicle that is used to perform coastal and port monitoring and anti-mine warfare tasks in shallow water environments, and the maximum speed of the “Tuna 21” – a submersible craft of the same series – is 4.5 knots (ZHANG, TONG, 2008). It can be seen that the target intensity of small underwater targets is low, the echo is weak, and the shallow sea interface reverberation makes the signal mixing ratio of the received signal low. A small target is mobile and flexible, but its moving speed is slow, the echo Doppler shift is low, and the classical frequency domain filtering makes it difficult to complete the high-resolution frequency shift estimation. The size and shape of different kinds of small targets lead to different spatial distributions of the target intensity. For an underwater cluster composed of multiple small targets, the multiple acoustic scattering interaction between the targets increases the complexity of the physical characteristics of the cluster targets. Therefore, it is important to study objective acoustic features that are easy to extract and apply to describe underwater moving cluster targets. The characteristics of the target acoustic scattering echo are related to the natural vibration mode of the target and the properties of the incident sound wave. The approximate calculation method of the target acoustic scattering field is commonly used in engineering, and the generation point of the scattering echo is regarded as a highlight (TANG, 1994; CHEN *et al.*, 2013). The acoustic scattering highlights of small cluster targets with a certain distribution law will cover a large scale in space and will be equivalent to large-scale volume targets. However, due to the different spatial distribution, motion speed, and motion state of individual small moving targets constituting the cluster, the target echo is reflected in the modulation of Doppler; that is, the Doppler shift of the highlight of individual small targets relative to the receiving array is different. Although the clear echo structure of the target cannot be obtained using active sonar to detect small cluster targets, when the interface interference is serious or multiple targets are in the same or adjacent beam, the scattered waves of the strong highlights of the target still contain rich target feature information. Aiming to identify large-scale targets such as underwater crafts and small cluster targets in port monitoring, this paper develops an extraction method of the spatial scale and motion parameter features of the highlights of the target. Based on the highlight model theory, the acoustic scattering echo model of the relative motion between multiple highlights and the receiving hydrophone array is established. The azimuth tendency of the target is estimated by the high-resolution split beam method, and the spatial geometric scale covered by multiple highlights is calculated. The target multi-highlight structure is obtained in the Wigner-Weill dis-

tribution (WVD) time-frequency domain (LATIF *et al.*, 2003; RODRÍGUEZ *et al.*, 2004; SABRA, ANDERSON, 2014), the reverberation and cross-term interference in the time-frequency domain are removed by morphological filtering (RONSE *et al.*, 2005; BOUAYNAYA *et al.*, 2008), and the instantaneous frequency of the underwater target highlight is extracted. According to the derived instantaneous frequency spectral line slope and frequency shift relationship, the broadband spectral Doppler shift of the highlight is calculated, and the motion state of the highlight is analysed. Combined with the target orientation, scale and motion information, a multiparameter feature space is established to describe the essential attributes of the target. The effectiveness of the features for identifying small cluster targets is verified by simulations and lake experiments.

2. Multitargets resolution method and theoretical derivation

For active sonar target detection, the spatial spectrum estimation method (LI *et al.*, 2014) is often used in engineering to suppress background interference and determine target orientation. When different targets are located in different scanning beams, they can be easily distinguished. However, when the azimuth of multiple targets is less than the beam resolution, multiple targets appear in the same beam, and only a bright spot of multi-target energy superposition is displayed in the angle-distance spectrum, it is difficult to distinguish different target highlights. On the basis of spatial spectrum estimation results, this paper studies and extracts the acoustic scattering characteristics of the beam domain echo, and identifies the target through the characteristics of the strong scattering highlights superimposed in the echo. For an underwater vehicle with a certain scale, the reflected echo generated by it is an extended body with a certain distribution in the range and orientation. This extended body is approximately equivalent to the sum of several “highlight” signals with a certain distribution in space (YANG, LI, 2016). Small cluster targets work together underwater in a planned formation, and the acoustic scattering highlights of multiple small targets cover large-volume targets with a certain scale in space. For large-scale targets such as warships and submarines, the motion parameters of each strong scattering highlight on the same target are the same. However, for the volume target echo composed of multiple small targets, due to the different spatial distribution and motion state of each target, the phase and Doppler shift of each highlight echo are also different. So the spatial orientation distribution, dispersion scale structure features and motion features of highlights are extracted and fused to identify multiple underwater moving targets in this paper. The method flow chart is shown in Fig. 1.

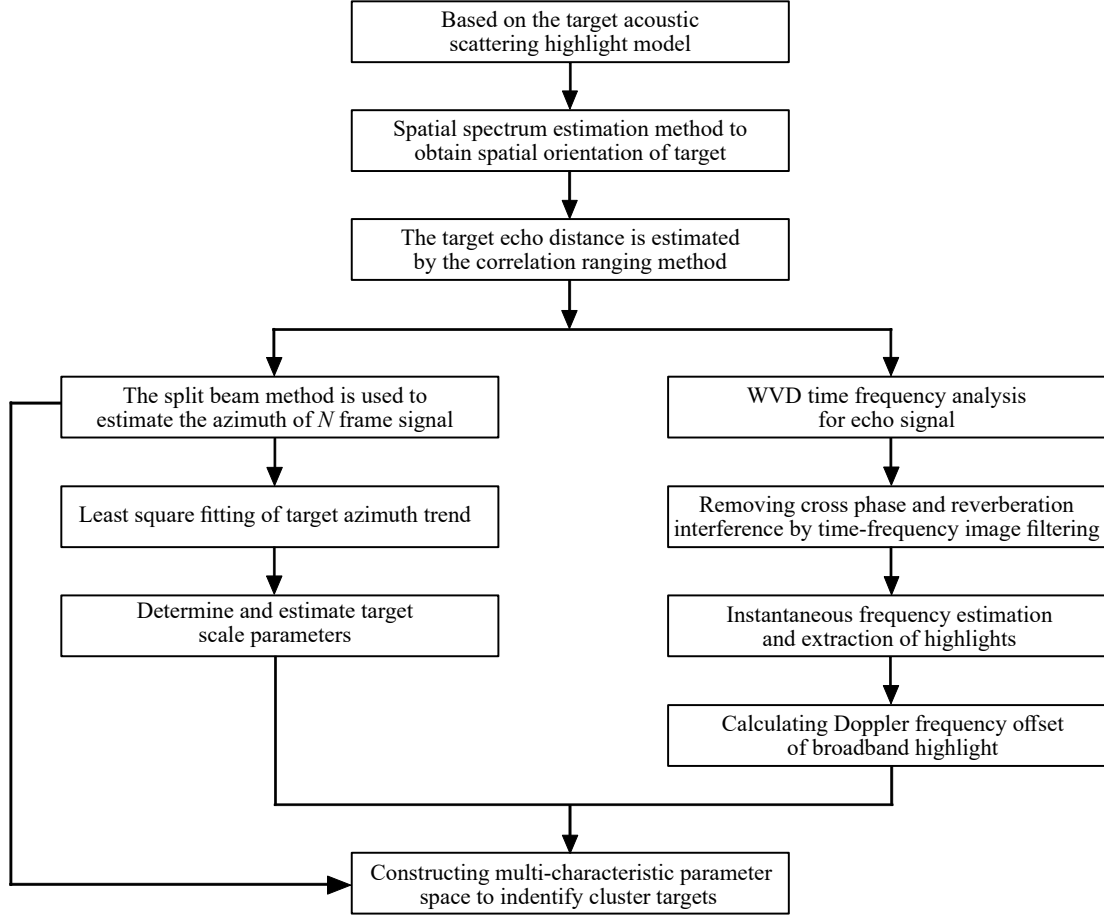


Fig. 1. Flow chart of multi-target resolution method.

First, a spatial spectrum estimation method is used to obtain the target azimuth. Next, the scale of the target is obtained by estimating the target azimuth trend, and then Doppler frequency shift of the highlights is estimated by extracting the instantaneous frequency of the target in the time-frequency domain. Finally, the multiparameter joint feature space is established to identify different targets in the same beam. The relevant theories and derivation of azimuth trend target scale estimation and Doppler parameter estimation are given in Subsecs. 2.1 and 2.2, respectively.

2.1. Azimuth trend target scale estimation

Target scale estimation based on the target azimuth trend is an effective scale recognition method (Ma *et al.*, 2004; Zhou, Yi, 2004; Jiang, Xiang, 2019) that has been applied in practical engineering. The basic idea of azimuth trend scale recognition is as follows: the target echo signal is divided into N subframes. According to the signals received by multiple array elements, the split beam method is used to obtain the azimuth estimation θ_i ($i = 1, 2, \dots, N$) of each subframe signal with N azimuth estimation values in total, and then the least squares linear fitting is used to obtain

the azimuth trend value, that is, the linear estimation of the maximum angle range of the echo distribution over the whole target coverage space within the echo duration. Then, the target azimuth trend T_r is estimated.

It is assumed that the signals received by the two array elements with d space are $x(t)$ and $y(t) = x(t+\tau)$, and the relative delay is $\tau = d \sin \theta / c$. The two signals are Fourier transformed, and the cross-power spectrum of the sequence is calculated as follows:

$$\begin{aligned}
 P_{xy} &= X(\omega)Y^H(\omega) \\
 &= X(\omega) [X(\omega)e^{j\omega\tau}]^H = |X(\omega)|^2 e^{-j\omega\tau}. \quad (1)
 \end{aligned}$$

It can be seen from the above equation that the phase angle of the cross-power spectrum of the output sequence of the two array elements is the phase difference generated by the incident signal on the two array elements, and the azimuth θ of the target can be calculated according to the phase angle of the cross-power spectrum. If the azimuth θ_i ($i = 1, 2, \dots, N$) of each subframe is obtained by splitting beam accurate azimuth estimation, it is assumed that the fitting equation is:

$$\theta = a + b \cdot i. \quad (2)$$

The least square minimizes the value in Eq. (3):

$$\Theta = \sum_{i=1}^N (\theta_i - \theta)^2 = \sum_{i=1}^N (\theta_i - (a + b \cdot i))^2. \quad (3)$$

We can obtain two equations that determine a and b :

$$\begin{aligned} a &= \frac{6}{N(N+1)} \left(\frac{2N+1}{3} A - B \right), \\ b &= \frac{12}{N(N^2-1)} \left(B - A \frac{N+1}{2} \right). \end{aligned} \quad (4)$$

In Eq. (4), assume $\sum_{i=1}^N \theta_i = A$, $\sum_{i=1}^N \theta_i i = B$.

The azimuth trend scale value T_r is obtained as follows:

$$\begin{aligned} T_r &= \theta(N) - \theta(1) = b(N-1) \\ &= 12 \left(\frac{B}{N(N+1)} - A \frac{1}{2N} \right). \end{aligned} \quad (5)$$

Although azimuth trend scale identification has been applied in practical engineering, it is worth studying how the general law of the azimuth trend changes when cluster targets are incident in different spaces. In this paper, the azimuth trend values under different incident azimuth angles are obtained by simulating the superimposed multi-highlight echo. The general variation law is given, and the identification angle range with good effect can be obtained by using the azimuth trend for scale estimation.

2.2. Azimuth trend Doppler parameter estimation

In active detection, when there is relative motion between the sonar system and the target, the carrier frequency of the secondary transmitted wave received by the sonar system will produce a Doppler effect, which is manifested in the echo signal expanding in the time domain and shifting in the frequency domain (BOASHASH, 1992). Broadband transmitting signals are often used for detection. The echo signal contains more target attribute information and can improve the time-frequency resolution. Analysing and deriving the expression of the Doppler shift of wideband echo is the basis of studying the feature extraction method of the Doppler shift of target acoustic scattering highlights.

As shown in Fig. 2, the combined transceiver mode is adopted. The acoustic source emits a signal $s(t)$ in the far field. The point source target moves at a uniform speed from starting position A at speed μ , the angle relative to the combined transceiver system is θ , and v is the radial velocity component of the target. Assuming that the target moves towards the sonar, the pulse width of the transmitted signal is T . De-

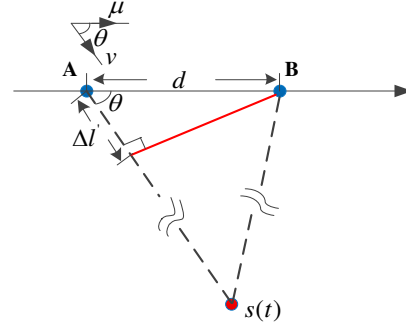


Fig. 2. Target motion model.

rived from the geometric relationship, the received signal pulse width (SHENG *et al.*, 2020) is:

$$T_1 = \frac{1 - \frac{v}{c}}{1 + \frac{v}{c}} T = s \cdot T. \quad (6)$$

The speed of the underwater vehicle is markedly smaller than the speed of sound $v \ll c$, assuming

$$s = \frac{1 - \frac{v}{c}}{1 + \frac{v}{c}} = \frac{c - v}{c + v} \approx 1 - \frac{2v}{c} = 1 - \Delta,$$

$\Delta = \frac{2v}{c}$ will be called the Doppler factor, $f_d = \frac{2v}{c} f_0$ represents Doppler frequency shift in the frequency domain, and f_0 is the transmission signal frequency.

For broadband signals, different frequency points correspond to different Doppler shifts. The starting frequency of the linear frequency modulation (LFM) signal is set as f_L , the cut-off frequency is f_H , and the frequency modulation slope is $k = \frac{f_H - f_L}{T}$. Then, after the Doppler effect occurs, the initial frequency of the echo signal is set to $f_{Lecho} = f_L + \Delta f_L$, and the cut-off frequency is set to $f_{Hecho} = f_H + \Delta f_H$. According to Eq. (6), the frequency modulation slope of the frequency shift is:

$$\begin{aligned} k_{echo} &= \frac{f_{Hecho} - f_{Lecho}}{T_1} \\ &= \frac{f_H + \Delta f_H - f_L - \Delta f_L}{sT} = \frac{1 + \Delta}{s} \cdot k. \end{aligned} \quad (7)$$

After reorganization:

$$\Delta = \frac{2v}{c} = \frac{k_{echo} - k}{k_{echo} + k}. \quad (8)$$

According to Eq. (8), the Doppler factor of the broadband signal is related to the frequency modulation slope of the transmitted signal and the received highlight. As long as the frequency modulation slope of the received highlight is detected, the Doppler shift of the target can be estimated, and then the radial motion velocity of the target can be obtained.

3. Numerical simulation and analysis

The performance of azimuth trend estimation, Doppler parameter estimation of moving multiple

highlight targets, and recognition of typical formation motion highlights in feature space are verified through three simulation experiments.

The target model composed of three small balls is established. The coordinate system of the transceiver array and the target model is shown in Fig. 3. The target model and the L-shaped array are located in the xoz plane, in which the distance l_1 between ball 1 and ball 2 is 23 m, and the distance l_2 between ball 2 and ball 3 is 53 m. The 28-element receiving array is placed perpendicular to the z -axis, its first array element (set as the reference array element) is located along the z -axis, the no. 2 small ball is located at the coordinate $(0, 0, 0)$, that is, the no. 0 beam is facing the no. 2 small ball, and the transmitting array of the L-shaped array is located near the reference array element, so the azimuth angle of the no. 2 small ball relative to the reference array element is 0° . The distance between the first array element of the array and the centre of the model is set as R , the target is located in the far field, and the distance taken in the simulation is $R = 600$ m. The angle between the longitudinal line of the model and the z -axis direction is the incident angle φ , and ϕ_1 and ϕ_2 are the azimuth angles of the reference array element corresponding to ball 1 and ball 3, respectively. The propagation time of the target model echo reaching the reference array element can be calculated according to the spatial position coordinates of the transmitting point and the spatial position coordinates of the receiving array element, and the echo signal received by the reference array element can be obtained. The echo signal of other array elements can be calculated through the array popularity matrix. The frequency range of LFM signal transmission is from $f_L = 100$ kHz to $f_H = 200$ kHz, and the signal pulse width is 20 ms.

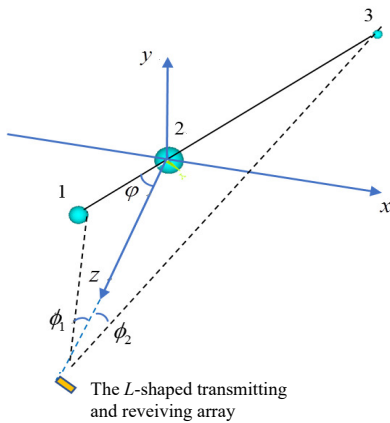


Fig. 3. Three-ball composition model.

Simulation analysis I: Azimuth trend estimation of single and multiple highlight targets

The azimuth trend estimation of single highlight and three highlight models at a certain angle and the

omnidirectional incidence is studied. The physical properties of azimuth trend estimation are explained with this trend estimation. The validity of azimuth trend estimation affects the accuracy of target scale estimation. The three highlight echo is composed of the superposition of highlight components with different time delays, amplitudes and phases, and the amplitude corresponding to the three components is set to 0.8:1.2:0.6. The results are shown in Fig. 4.

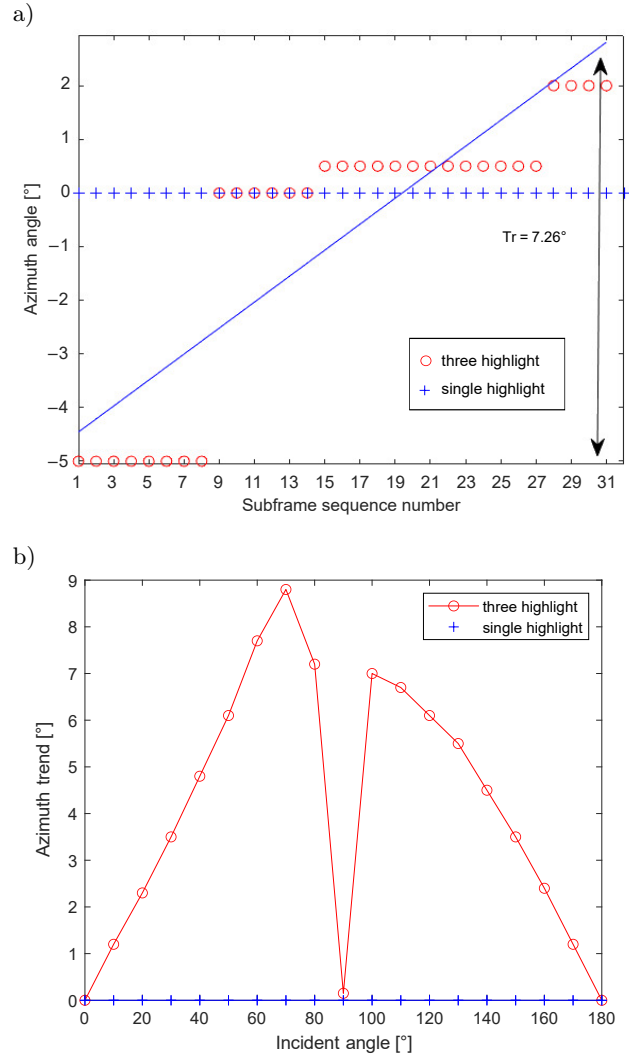


Fig. 4. Echo azimuth trend estimation: a) azimuth trend of 60° incident angle; b) azimuth trend of different incident angles.

Taking the incident angle $\varphi = 60^\circ$, Fig. 4a shows that for single highlight targets, the azimuth trend value is very small, which is basically close to 0. Multiple highlights have a certain angle and distance distribution in space. The subframe segmentation in the azimuth trend estimation process distinguishes each highlight from the echo time or distance, accurately estimates the formation angle of each highlight by splitting the beam, and then obtains the maximum angle distribution range of the target highlight by the least

squares linear fitting. The estimated target azimuth trend $Tr = 7.6^\circ$, $R1 = 593.1$ m, and $R2 = 604.4$ m. The calculated coverage scale of the three highlights is 76.6 m, which is consistent with the simulation condition of 76 m.

Figure 4b shows that the azimuth trend of a single highlight changes little with the incident angles and is also close to 0. The azimuth trend of multiple highlights with different incident angles is different. This is mainly because the relative positions of multiple highlights change when the incident angles are different, and the angular distribution of their echoes in space is different. However, the azimuth trend near the positive cross tends to 0 because the time delay difference of each ball echo is very small, most of the echoes overlap, and it is difficult to distinguish the highlights through subframe segmentation.

From the above analysis, we can see that the essence of the so-called azimuth trend is to distinguish the highlights through distance segmentation within the echo duration and accurately estimate the azimuth angle of the highlights by splitting the beam to obtain a linear estimation of the maximum angular range value of the echo distribution on the spatial distance covered by multiple highlights. The azimuth trend is an important feature that distinguishes multiple highlight targets from point targets and can be used for target scale recognition.

Simulation analysis II: Doppler parameter estimation of moving multiple highlight targets

To analyse the effect of the Doppler shift estimation method of the moving highlights, the instantaneous frequency distribution of the highlight is extracted in the time-frequency domain, and the Doppler shift is estimated by the relationship between the received highlight time-frequency spectrum slope and the transmitted signal frequency modulation slope. The initial state of the three highlights is shown in Fig. 3, and those are set to move at constant speeds of 1.2, 0, and -0.8 m/s. According to the derivation in Subsec. 2.2, the expression of the echo signal obtained from the moving highlight model is:

$$x(t) = \sum_{i=1}^3 x_i(t) = \sum_{i=1}^3 A_i \cos(2\pi(f_l + f_{di})t - \pi k_{di}t^2),$$

$$k_{di} = \frac{B}{\left(1 - \frac{f_{di}}{f_j}\right)T}, \quad (9)$$

where B is the bandwidth, T is the pulse width of the transmission signal, f_{di} is the Doppler frequency shift of the i -th highlight, and $f_j \in [f_L f_H]$.

To calculate the Doppler frequency shift of the moving target, it is necessary to obtain the instantaneous frequency of the highlight. In this paper, WVD is used to map the echo data to the time-frequency domain. The WVD of the LFM signal has the best energy

aggregation in the time-frequency distribution method, and the signal energy is concentrated on a straight line that can reflect the instantaneous frequency change of the signal. However, there is cross-term interference in the multicomponent WVD. Morphological filtering is used to filter out cross-term interference by selecting appropriate structural elements, and the instantaneous frequency distribution of echo highlights is extracted without changing the time-frequency resolution and signal energy.

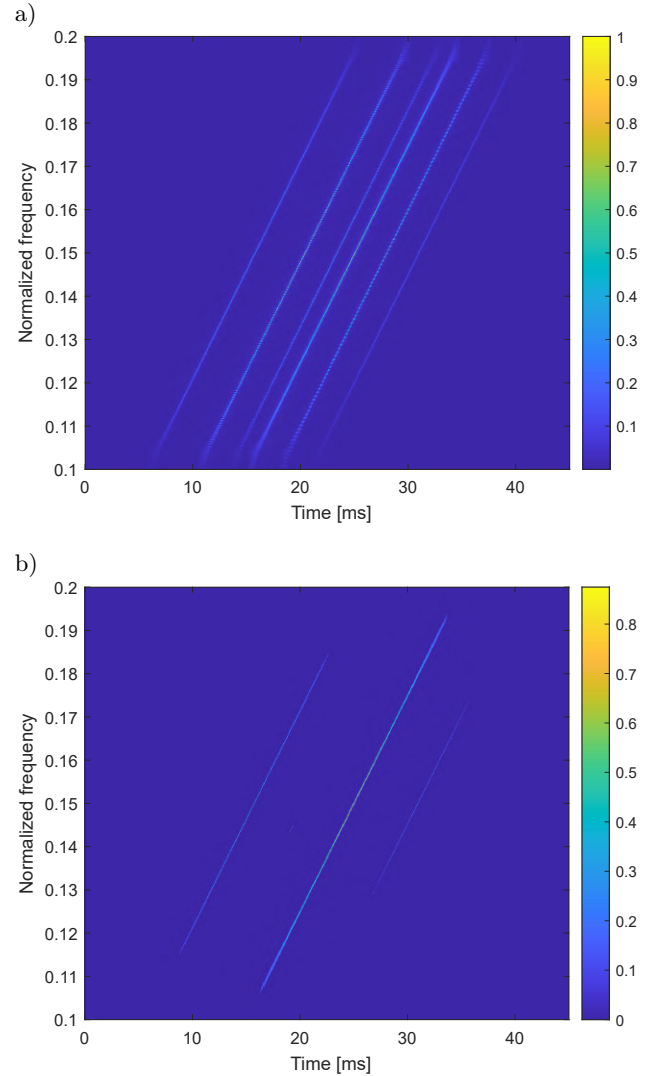


Fig. 5. Multiple highlight time-frequency distribution: a) time-frequency distribution of multiple highlights; b) highlight distribution after removing cross-term interference.

Figure 5a shows the time-frequency distribution of three highlights. The horizontal axis represents time, and the vertical axis represents normalized frequency. The normalized frequency is the ratio of signal frequency to the sampling frequency. In the experiment, the sampling frequency is 1 MHz. Theoretically, the time-frequency distribution should be three continu-

ous line spectra in the time-frequency domain, but the number of line spectra in the figure is greater than three, which is due to cross-term interference generated by the WVD algorithm. According to the morphological difference between highlight echo and cross-item interference, here, an inclined line with the same slope as the transmitted signal, is selected as the structural element, its length is half of the pulse width of the transmitted signal, and image objects of corrosion and expansion operations are used. The processing result is shown in Fig. 5b. The cross-term interference in the time-frequency plane is removed by morphological filtering, but the weak energy at the beginning and end of the highlight line spectrum is corroded, which does not affect the extraction and calculation of the line spectrum slope.

According to the time delay energy distribution of highlights, the line spectrum components of three highlights are searched and extracted, and the instantaneous frequency line spectrum of each highlight is obtained, as shown in Fig. 6. The slope of each highlight line spectrum is calculated as $5.0145e6$, $4.9972e6$, and $4.9917e6$ by using the method of overlapping and averaging, and the slope of the transmitted signal is $5.0e6$. The spectral slope of the no. 1 sphere increases relative to the transmitted signal, indicating that the highlight moves towards the transceiver transducer. The spectral slope of the no. 3 sphere decreases relative to the transmitted signal, indicating that the highlight moves to the transceiver transducer in a reverse manner. The no. 2 sphere is located on the centreline of the transceiver transducer, and its radial velocity in the direction of the transceiver transducer is 0, so the line spectrum slope is closer to the spectrum slope of the transmitted signal. According to the instantaneous frequency slope and Eq. (9), the radial velocity corresponding to the line spectrum is calculated as shown

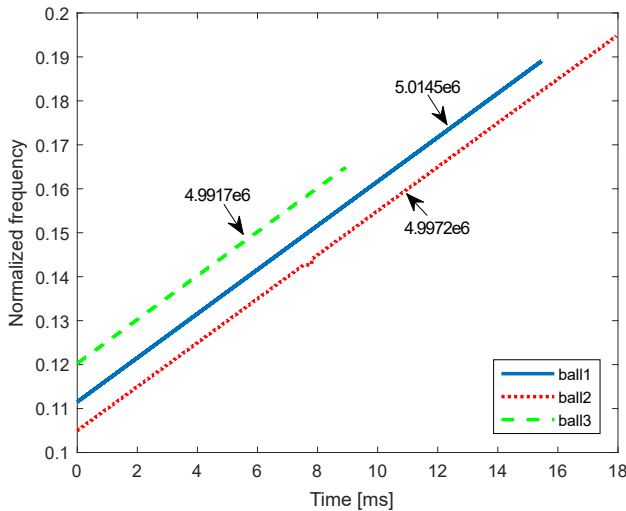


Fig. 6. Multiple highlight instantaneous frequency extraction.

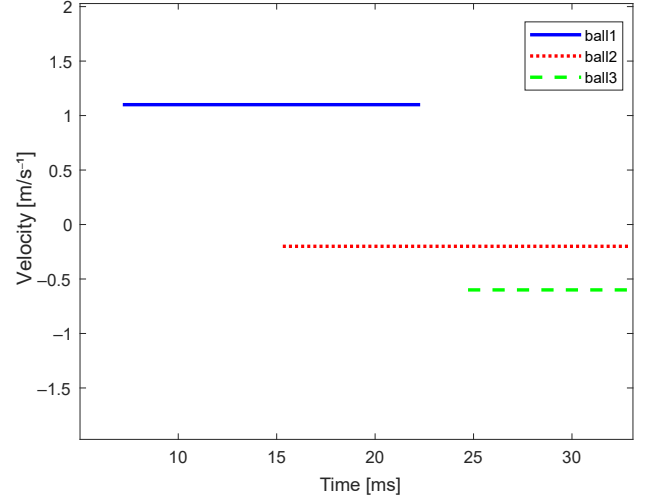


Fig. 7. Speed estimation.

in Fig. 7, which is 1.1, 0.2, and -0.63 m/s, respectively, which is basically consistent with the simulation conditions. The speed estimation accuracy is related to the calculation accuracy of the frequency modulation slope of the highlight, and it depends on the resolution of the instantaneous frequency of the highlight. The frequency resolution is inversely proportional to the effective length of the analysis signal. The simulation analysis signal length is 20 ms, and the corresponding frequency resolution is 50 Hz. According to the Eq. (8), the radial speed error is approximately 0.2 m/s.

Simulation analysis III: Recognition of typical formation motion highlights in feature fusion space

Verify the effectiveness of highlight identification in feature fusion space. The three-dimensional feature space of the target highlights is constructed by fusing the spatial orientation, velocity, and azimuth trend scale information, and the feature space distribution under the typical formation of multiple highlights is studied to identify the cluster targets. The spatial resolution of the receiving array is approximately 1° , and the three highlights move parallel to the x -axis with uniform speeds of 0.8, 1.6, and 2.2 m/s.

Assuming that the three highlights are arranged in a horizontal line, the vertical distance from the transceiver system is 600 m, and the spatial azimuth angles of the three highlights are -28° , -30° , and -31.5° , as shown in Fig. 8a. The orientation, azimuth trend scale, and Doppler parameters of the three highlights are extracted, the motion speed and coverage space scale of the echo are calculated and obtained, and the feature space is constructed, as shown in Fig. 8b. The spatial orientation of the three highlights is the same as the simulation setting, with speeds of 0.68, 1.53, and 2.1, respectively, and the coverage space scale is 48.2 m.

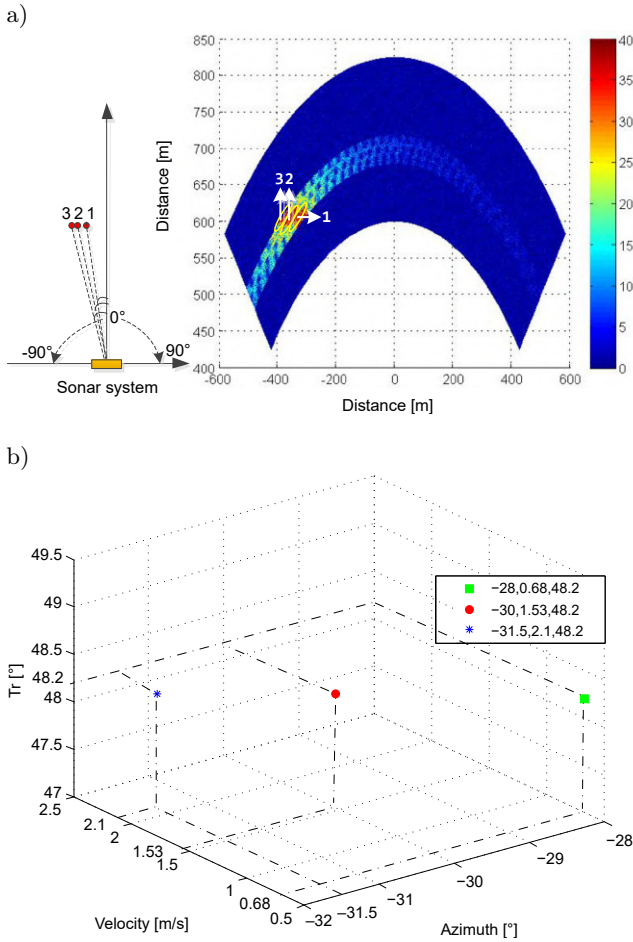


Fig. 8. Spatial characteristic distribution of the highlights in a line: a) spatial distribution of the highlights in a line; b) characteristic space of three highlights in a line.

Assuming that the three highlights are arranged in a straight line, the vertical distances from the transceiver system are 600, 603, and 608 m, and the vertical distances from the three highlights to the y -axis are all 10 m, as shown in Fig. 9a. The feature quantity of the three highlights is extracted and calculated, and the feature space is constructed, as shown in Fig. 9b. The spatial orientation of the three highlights is -30° because the spatial orientation distribution difference of the highlights is less than the resolution of the receiving array, resulting in the inability to accurately distinguish. The estimated velocities of the three highlights are 0.69, 1.45, and 2.14 m/s. The orientation estimation error is within the range of 1° , which has little impact on the velocity accuracy and can be ignored. The spatial scale covered by the three highlights is 8.3 m.

Assuming that the three highlights are arranged in a triangular formation, the vertical distances from the transceiver system are 600, 610, and 600 m, and the vertical distances from the three highlights to the y -axis are 5, 10, and 15 m, as shown in Fig. 10a. The con-

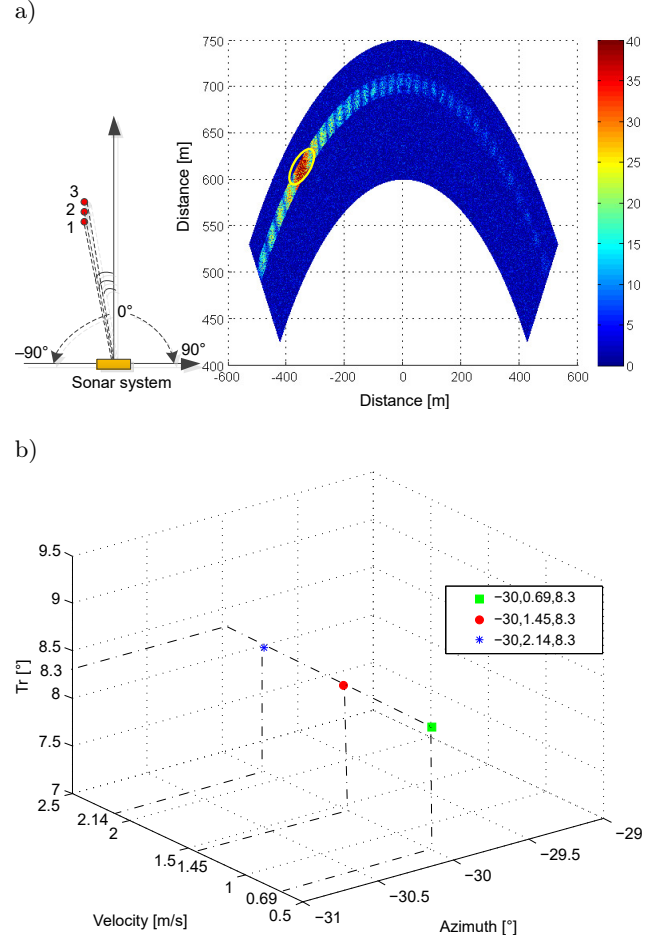


Fig. 9. Spatial characteristic distribution of three highlights in a column: a) spatial distribution of the highlights in a column; b) characteristic space of three highlights in a column.

struction feature space is shown in Fig. 10b. The spatial orientation of the three highlights is approximately -28.5° , the speeds are 0.68, 1.43, and 2.06 m/s, and the spatial scale covered is 10.4 m.

In summary, multiple highlights can be covered in large-scale targets in space. For clusters, due to the different motion states of small targets constituting clusters, the motion characteristics of multiple highlights in space are also different. Therefore, large-scale targets such as underwater crafts and cluster targets can be distinguished.

The simulation assumes that the target moves at a horizontal uniform speed; however, the target attitude and motion state of the actual cluster are usually unknown, so the actual motion speed of the target cannot be obtained, but the radial velocity of highlights still can be estimated according to the frequency modulation slope, and it can be determined that they belong to different targets according to the difference of the radial velocity in the same wave beam. In addition, during long-term underwater monitoring, the movement speed of the cluster also changes, and we

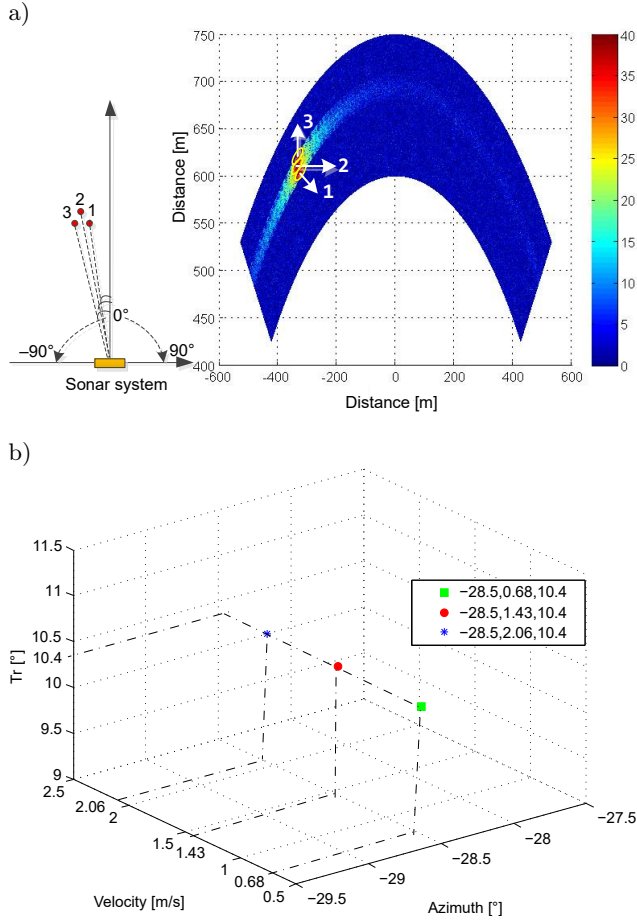


Fig. 10. Spatial characteristic distribution of three highlights in a triangle formation: a) spatial distribution of three highlights in a triangle formation; b) characteristic space of three highlights in a triangle formation.

can divide and analyse the length of the received data according to the resolution requirements, calculate the characteristic parameters of different data segments, and obtain real-time dynamic cluster characteristics for cluster target recognition.

4. Experimental analysis and verification

To verify the effectiveness of the multi-feature fusion method in identifying cluster targets, the experiment uses two small-scale UUVs moving underwater, with a length of 2.3 metres. The layout of the experiment is shown in Fig. 11. Using the combined transceiver detection mode, the pulse width of the transmitted linear frequency modulated signal is 2 ms, the frequency band is 100~200 kHz, the sampling rate is 1 MHz, the reception is a 28-element linear array, and the angular resolution of the conventional beamforming is 3.6° . The test conditions are as follows: one UUV and two UUVs sail autonomously on the side of the measuring ship at different speeds. The echo data of one UUV moving in a straight line at the side

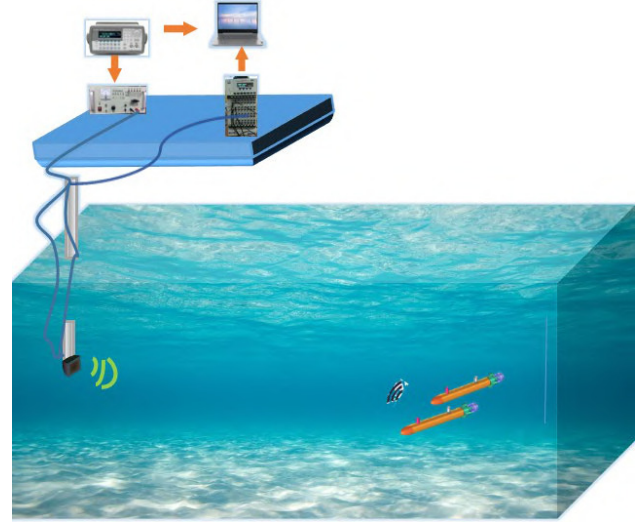


Fig. 11. Experimental layout.

of the measuring platform at a speed of 1.2 m/s and the echo data of two UUVs moving in circles at the side of the platform at speeds of 1.2 and 0.8 m/s are selected for data analysis and processing. The treatment results of typical working conditions are as follows.

5. Experimental results and analysis

(1) A single UUV sails in a straight line with a speed of 1.2 m/s parallel to the ship's side underwater. The time distance diagram of the target echo is shown in Fig. 12. Taking the 220th echo data for analysis, the spatial orientation of the target is shown in Fig. 13, and the target appears in the -12.4° orientation of the transceiver. Then, WVD time-frequency analysis is performed on the target echo, and the time-frequency distribution results are shown in Fig. 14a. Because there is interference in WVD, image morphological filtering is used to remove the cross term and the

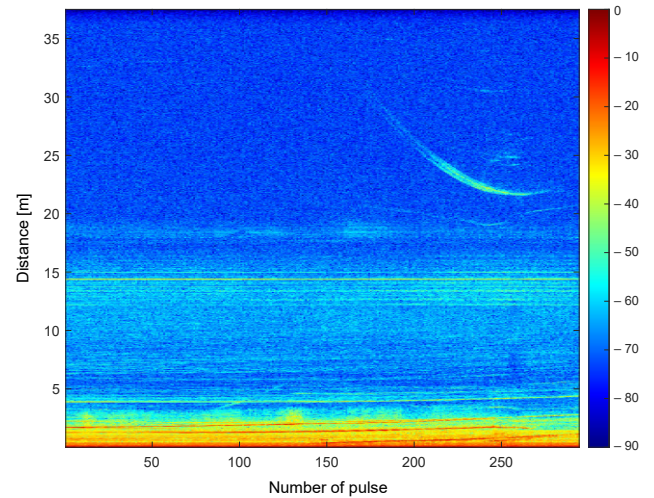


Fig. 12. Target time-distance diagram.

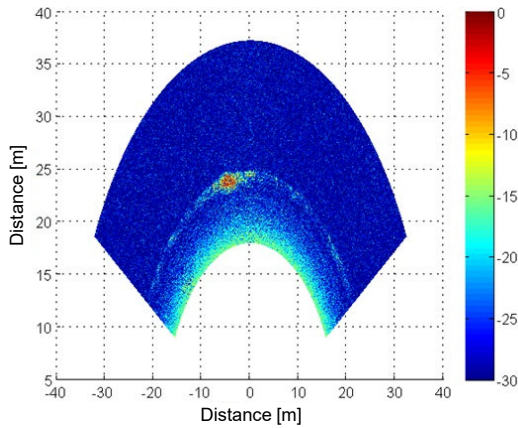


Fig. 13. Target space azimuth diagram.

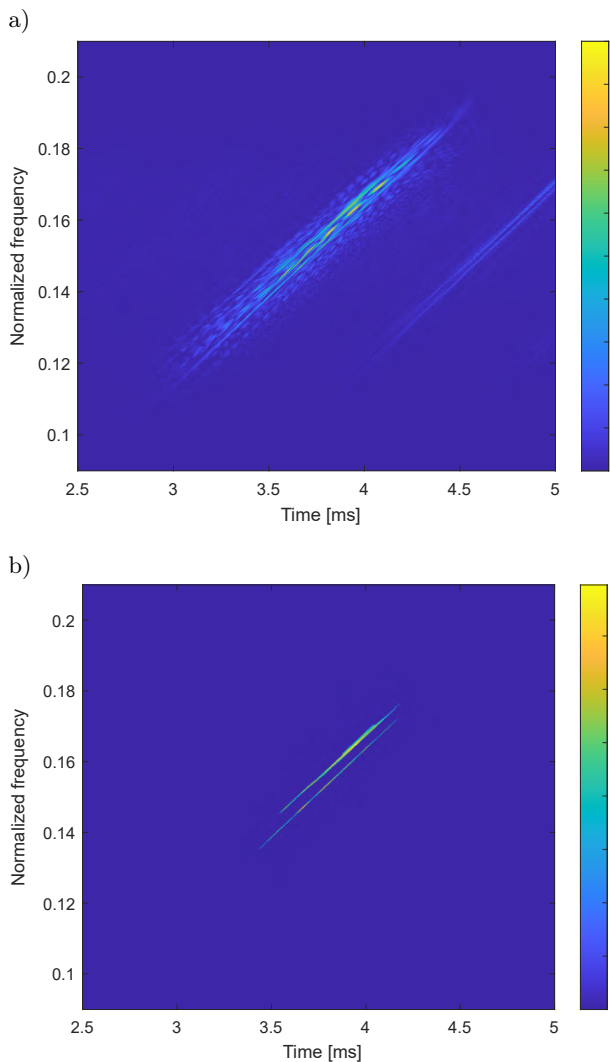


Fig. 14. Time-frequency distribution of target highlights: a) time-frequency distribution of the target echo; b) time-frequency distribution of the echo with cross-term interference removed.

reverberation. The results are shown in Fig. 14b. It can be seen that the target echo contains two scattering highlight components, which are the echoes scattered

by the UUV cylinder. The instantaneous frequency distribution of highlights obtained by energy search in the time-frequency domain is shown in Fig. 15. It is estimated that the slopes of the instantaneous frequency spectral lines of the two highlights are $5.0036e7$ and $5.0033e7$, which are slightly larger than the slope of the transmitted signal of $5e7$. The target moves towards the transceiver, and the calculation speeds are 1.23 and 1.13 m/s, respectively, which are approximately equal within the error range. It is estimated that the azimuth trend of the target $Tr = 4.8^\circ$, as shown in Fig. 16, that is, the opening angle of the reference hydrophone array element relative to the scatterer is 4.8° . Using the maximum detection method, the distance between the first segment echo and the 32nd segment echo relative to the reference array element is 24.4 and 24.1 m, respectively. According to the triangular relationship, the target scale covered by the two highlights can be estimated to be approximately 2.1 m. Because the velocity

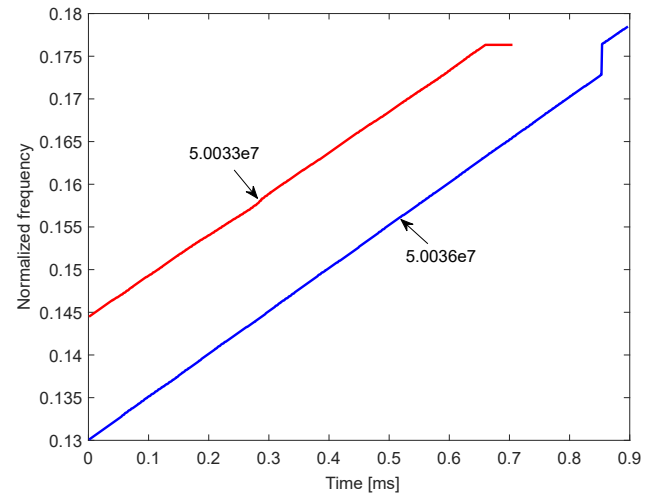


Fig. 15. Instantaneous frequency extraction of target highlights.

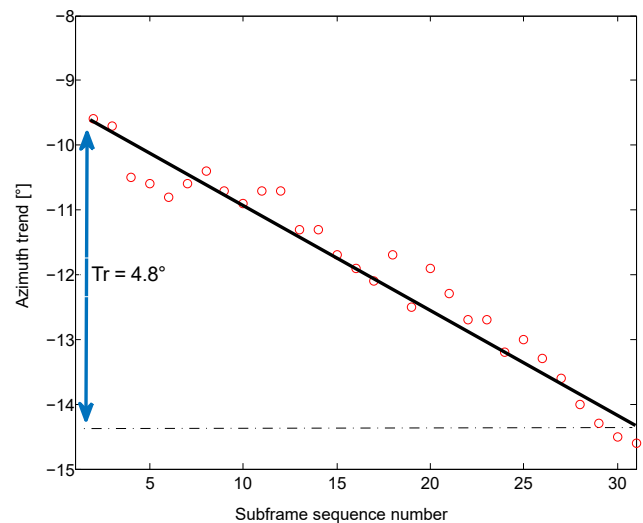


Fig. 16. Target azimuth trend diagram.

of the two highlights is approximately the same, different highlight components scattered by a target can be analysed, which is consistent with the test situation.

(2) Two UUVs move in circles on one side of the platform at speeds of 1.2 and 0.8 m/s. The time distance diagram of the target echo is shown in Fig. 17. We take the 903rd echo data for analysis. At this time, the no. 1 UUV passes the positive horizontal position and moves away from the transceiver along the circumference. The no. 2 UUV moves towards the positive horizontal position of the transceiver along the circumference. The spatial orientation of the target relative to the transceiver is shown in Fig. 18. Both UUVs appear in the 3.5° orientation of the transceiver. Then, the WVD time-frequency analysis is performed on the target echo, and the time-frequency distribution results are shown in Fig. 19a. Image morphological filtering is used to remove the interference, and the results are shown in Fig. 19b. The target echo contains three scattering highlights, and the instantaneous frequency distribution of the highlight echo is found in the time-frequency plane, as shown in Fig. 20. It is estimated that the instantaneous frequency line spectral slopes of the three highlights are $4.992\text{e}7$, $5.004\text{e}7$,

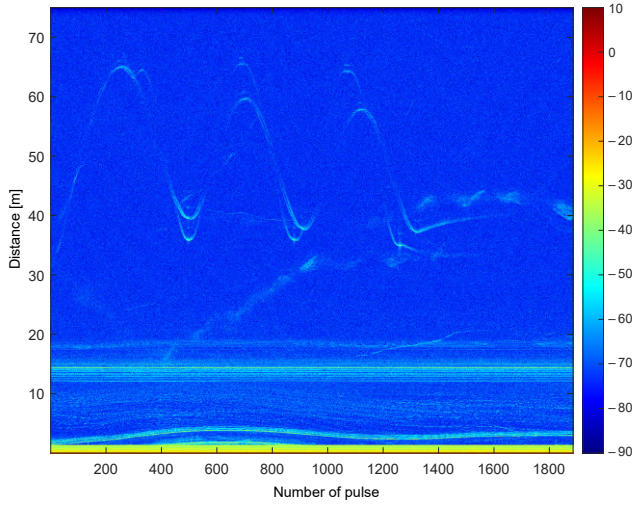


Fig. 17. Target-time distance diagram.

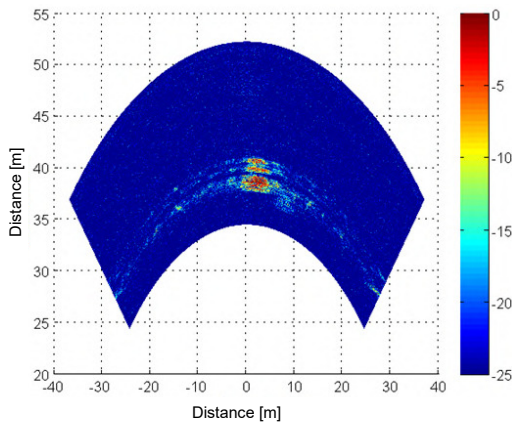


Fig. 18. Target space azimuth diagram.

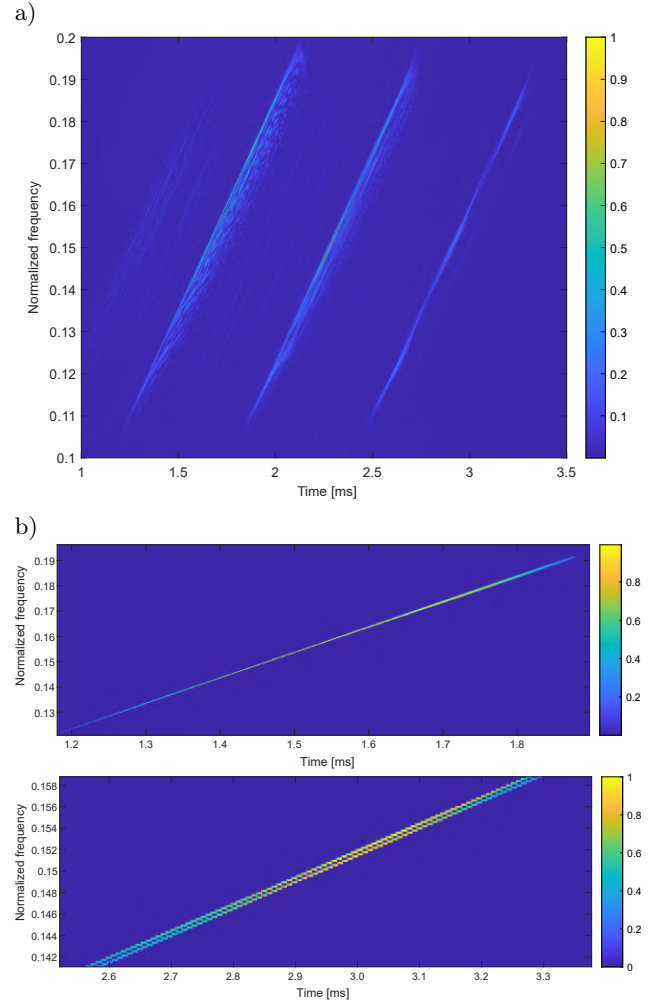


Fig. 19. Time-frequency distribution of target highlights: a) time-frequency distribution of the target echo; b) time-frequency distribution of the echo with cross-term interference removed.

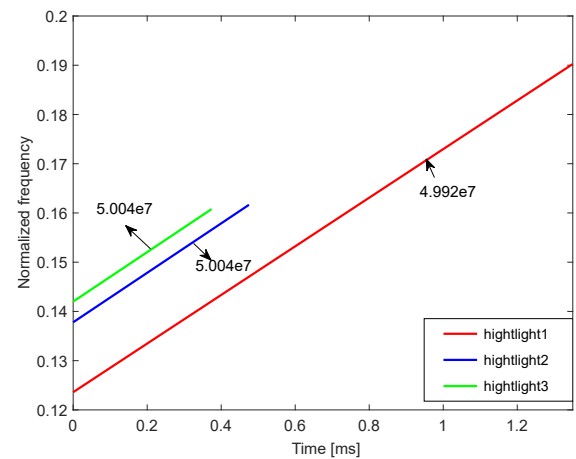


Fig. 20. Instantaneous frequency extraction of target highlights.

and $5.004\text{e}7$, and the corresponding radial velocities are -0.6 , 0.3 , and 0.3 m/s. Because the motion attitude of the target is not accurately known, that is,

the angle between the target velocity and the radial velocity is unknown, the target motion velocity cannot be estimated from the radial velocity, but from the radial velocity, highlights 2 and 3 are the scattering components of the same target, moving towards the transceiver, and highlight 1 is far away from the transceiver. The estimation results of the target azimuth trend are shown in Fig. 21. Highlight 1 is the echo component scattered by the tail of the no. 1 UUV, and the azimuth trend of the single highlight is 0. Highlights 2 and 3 are the echo components scattered by the cylinder structure of the no. 2 UUV, and their azimuth trend is 2.1° . According to the geometric relationship, the distance between the echo and the reference array element is 39.8 and 41.1 m, respectively, and the target scale covered by the two highlights can be estimated to be approximately 1.9 metres. According to the analysis, highlight 1 is the echo of the no. 1 UUV, and highlights 2 and 3 are the echoes of the no. 2 UUV. Integrating the target orientation, azimuth trend and Doppler characteristics, it is recognized that the echo in a beam is composed of two UUV scattering components, which is consistent with the test situation.

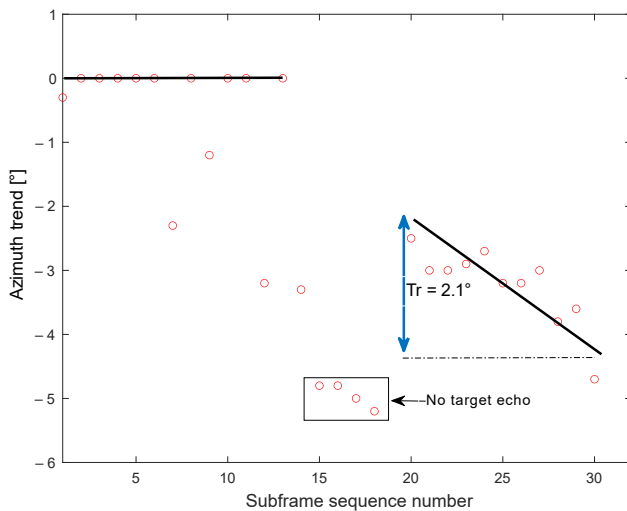


Fig. 21. Target azimuth trend diagram.

6. Conclusions

From the application perspective of underwater cluster target recognition, this paper studies and establishes the multi-feature space of echo highlights under typical formation and motion statuses. Through the analysis, the following conclusions can be drawn:

(1) The combined transceiver scattering characteristics of underwater cluster small targets are extracted, and the spatial orientation, geometric distribution and motion state differences of cluster targets are converted into physical features that are easy to use for sonar recognition;

(2) The objective of the azimuth trend is to distinguish each highlight through distance segmentation within the echo duration and obtain a linear estimation of the maximum angle range of the echo distribution on the spatial distance covered by multiple highlights. The azimuth trend is an important feature that distinguishes multi-highlight scale targets from point targets and can be used for target scale recognition.

(3) The corresponding relationship between the slope of the instantaneous frequency spectral line of highlights and the Doppler frequency shift is derived. The instantaneous frequency of multiple highlights of the target is extracted from the WVD time-frequency domain, and the motion information is calculated and extracted to describe the motion state of the highlights.

(4) The three-dimensional feature space of target highlight orientation, spatial scale and motion information is constructed, which realizes the effective analysis of multiple targets and solves the problem of fuzzy identification of multiple targets in the same beam. It provides effective and stable cluster echo recognition features for the early warning and monitoring of underwater ports and shallow waters.

Acknowledgments

The authors would like to acknowledge the China Postdoctoral Foundation Science Foundation (Grant no. 2020M681303) and the National Natural Science Foundation of China (Grant no. 61901079).

References

1. BOASHASH B. (1992), Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals, *Proceedings of the IEEE*, **80**(4): 520–538, doi: 10.1109/5.135376.
2. BOUAYNAYA N., CHARIF-CHEFCHAOUNI M., SCHONFELD D. (2008), Theoretical foundations of spatially variant mathematical morphology, Part I: Binary images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**(5): 823–836, doi: 10.1109/TPAMI.2007.70754.
3. CHEN Y.-F., LI G.-J., WANG Z.-S., ZHANG M.-W., JIA B. (2013), Statistical feature of underwater target echo highlight [in Chinese], *Acta Physica Sinica*, **62**(8): 084302, doi: 10.7498/aps.62.084302.
4. HOLLETT R.D., KESSEL R.T., PINTO M. (2006), At-sea measurements of diver target strengths at 100 kHz: Measurement technique and first results, [in:] *UDT-Europe 2006, Undersea Defence Technology Conference and Exhibition*, Germany.
5. HUANG Y.S. *et al.* (2020), Review on the development of diver detection sonar system [in Chinese], *Journal of Unmanned Undersea Systems*, **28**(1): 1–8.

6. JIANG C.O., XIANG X.M. (2019), The analysis of the azimuth trend sequence distortion in the target scale identification [in Chinese], *Acta Acustica*, **38**(2): 275–277.
7. JIANG L.J., YANG J., XU F. (2009), Technological progress of diver detection sonar [in Chinese], *Science in China Press*, **54**(3): 269–272.
8. LATIF R., AASSIF E., MOUDDEN A., FAIZ B. (2003), High-resolution time–frequency analysis of an acoustic signal backscattered by a cylindrical shell using a modified Wigner–Ville representation, *Measurement Science and Technology*, **14**(7): 1063–1065, doi: 10.1088/0957-0233/14/7/322.
9. LI X.K., LI T.T., GU X.Y. (2014), Array gain of fourth-order cumulants beamforming under typical probability density background [in Chinese], *Acta Acustica*, **39**(5): 557–564, doi: 10.15949/j.cnki.0371-0025.2014.05.010.
10. LO K.W., FERGUSON B.G. (2004), Automatic detection and tracking of a small surface watercraft in shallow water using a high-frequency active sonar, *IEEE Transactions on Aerospace and Electronic Systems*, **40**(4): 1377–1388, doi: 10.1109/TAES.2004.1386890.
11. MA G.Q., XU D.M., WANG X.X. (2004), A study of simulation for underwater target scale identification with the technique of target azimuth tendency [in Chinese], *Ship Science and Technology*, **26**(3): 39–42.
12. RODRÍGUEZ M.A., SAN EMETERIO J.L., LÁZARO J.C., RAMOS A. (2004), Ultrasonic flaw detection in NDE of highly scattering materials using wavelet and Wigner–Ville transform processing, *Ultrasonics*, **42**(1–9): 847–851, doi: 10.1016/j.ultras.2004.01.063.
13. RONSE C., NAJMAN L., DECENCIÈRE E. [Eds.] (2005), *Mathematical Morphology: 40 Years On. Proceedings of the 7th International Symposium on Mathematical Morphology, April 18–20, 2005*, Springer Dordrecht.
14. SABRA K.G., ANDERSON S.D. (2014), Subspace array processing using spatial time-frequency distributions: Applications for denoising structural echoes of elastic targets, *The Journal of the Acoustical Society of America*, **135**(5): 2821–2835, doi: 10.1121/1.4871183.
15. SARANGAPANI S., MILLER J.H., POTTY G.R., REEDER D.B., STANTON T.K., CHU D. (2005), Measurements and modeling of the target strength of diver, [in:] *Europe Oceans 2005*, **2**: 952–956, doi: 10.1109/OCEANSE.2005.1513185.
16. SHENG X.L., MU M.F., YIN J.W., YANG C., LIU T. (2020), Sparse decomposition-based estimation method of Doppler frequency shift of underwater moving target, *Journal of Harbin Engineering University*, **41**(10): 1429–1435, doi: 10.11990/jheu.202007055.
17. TANG W.L. (1994), Highlight model of echoes from sonar targets, *Acta Acustica*, **19**(2): 92–100.
18. YANG Y., LI X.-K. (2016), Blind source extraction based on time-frequency characteristics for underwater object acoustic scattering [in Chinese], *Acta Physica Sinica*, **65**(16): 164301, doi: 10.7498/aps.65.164301.
19. ZHANG Y.-M., TONG S. (2008), Tactics appliance of anti-diver technique in harbor defenses [in Chinese], *Ship Science and Technology*, **30**(6): 168–171.
20. ZHOU D.B., YI H. (2004), A study of submarine target bearing trend measurement and target discrimination [in Chinese], *Torpedo Technology*, **12**(4): 24–28.

Review Paper

Review of Methodologies in Recent Research of Human Echolocation

Michał BUJACZ*, Bartłomiej SZTYLER, Natalia WILEŃSKA,
Karolina CZAJKOWSKA, Paweł STRUMIŁŁO

Institute of Electronics, Lodz University of Technology
Łódź, Poland

*Corresponding Author e-mail: michal.bujacz@p.lodz.pl

(received November 26, 2022; accepted December 14, 2022)

The presented review discusses recent research on human echolocation by blind and sighted subjects, aiming to classify and evaluate the methodologies most commonly used when testing active echolocation methods. Most of the reviewed studies compared small groups of both blind and sighted volunteers, although one in four studies used sighted testers only. The most common trial procedure was for volunteers to detect or localize static obstacles, e.g., discs, boards, or walls at distances ranging from a few centimeters to several meters. Other tasks also included comparing or categorizing objects. Few studies utilized walking in real or virtual environments. Most trials were conducted in natural acoustic conditions, as subjects are marginally less likely to correctly echolocate in anechoic or acoustically dampened rooms. Aside from live echolocation tests, other methodologies included the use of binaural recordings, artificial echoes or rendered virtual audio. The sounds most frequently used in the tests were natural sounds such as the palatal mouth click and finger snapping. Several studies have focused on the use of artificially generated sounds, such as noise or synthetic clicks. A promising conclusion from all the reviewed studies is that both blind and sighted persons can efficiently learn echolocation.

Keywords: echolocation; blindness; testing methodology.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Echolocation is the ability of humans and some animals to locate objects basing on reflected sounds. The research on the ability of humans to echolocate has come a long way since first studies that had to clear up misconceptions about the visually impaired using “facial vision” or “obstacle sense” (SUPA *et al.*, 1944). By now, numerous experiments demonstrated the effectiveness of localizing obstacles using various reflected sounds.

Research no longer focuses on proving that echolocation works, but more on how it works, especially from the neurological perspective (FIEHLER *et al.*, 2015; THALER *et al.*, 2011), and on the ways to teach or improve echolocation skills (Fundacja Instytut Rozwoju Regionalnego [FIRR], 2019; TONELLI *et al.*, 2016). Because the consequence of blindness is a serious sensory deprivation one should exploit any possible cues to enhance safe mobility capabilities among the visually

impaired. Learning and mastering echolocation skills should be an important part of any rehabilitation programme for the visually impaired. Such programmes might benefit if the mechanisms of echolocation abilities and their limitations are well understood. One can observe an increasing number of publications devoted to human echolocation as shown in Fig. 1.

The methodologies in the recent echolocation studies vary greatly – some researchers conducted their trials predominantly with sighted volunteers (ARIAS, RAMOS, 1997; RYCHTARIKOVA *et al.*, 2017; TONELLI *et al.*, 2016), others with various sized groups of blind volunteers (FLANAGIN *et al.*, 2017; THALER, GOODALE, 2016; TIRADO *et al.*, 2019), some including or limiting the studies to echolocation experts (FIEHLER *et al.*, 2015; NORMAN, THALER, 2018). Some trials were in natural (BUJACZ *et al.*, 2018) or anechoic (SCHENKMAN, NILSSON, 2010) conditions, while others utilized recordings (ARIAS, RAMOS, 1997), synthesized echoes (WALLMEIER, WIEGREBE, 2014) or vir-

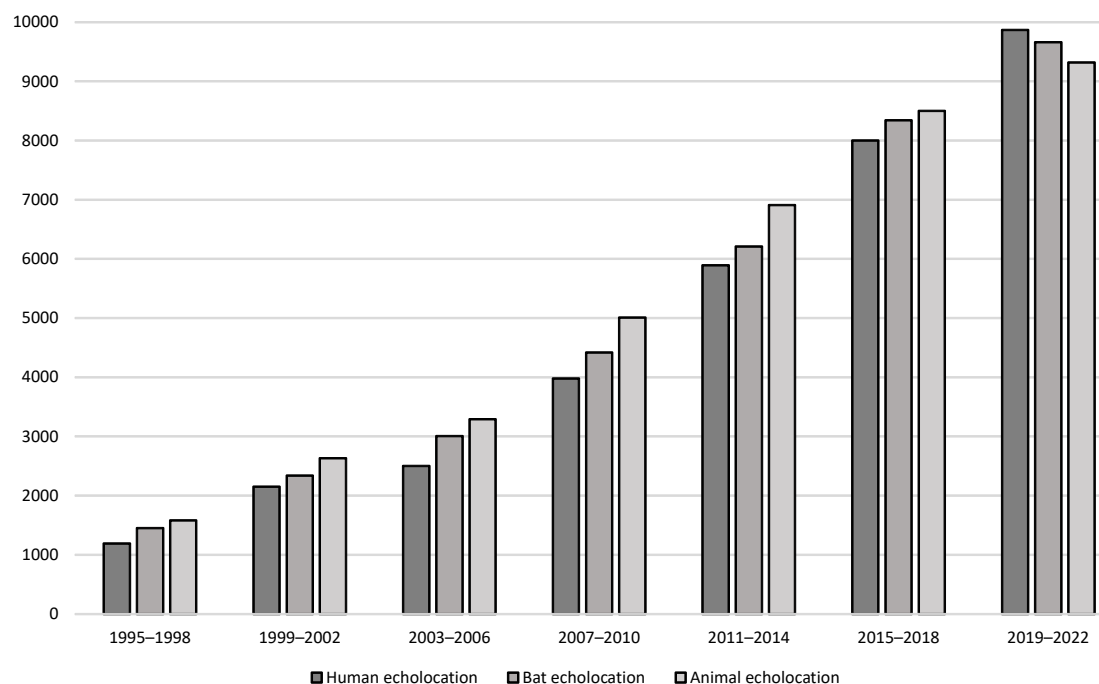


Fig. 1. Number of Google Scholar search results for echolocation related articles and patents.

tual reality environments (DODSWORTH *et al.*, 2020). Some studies let volunteers generate their own sound cues (THALER *et al.*, 2020b) or focused on analyzing those sound cues (ROJAS *et al.*, 2009), while others used recordings (FLANAGIN *et al.*, 2017) or examined the effectiveness of various artificial sounds (TIRADO *et al.*, 2019). A full list of compared studies is available in Table 1, then further sections contain smaller summary tables comparing key aspects of the studies.

An emerging issue with human echolocation research is that there has been no common methodology for studying its effectiveness, making it very difficult to compare the outcomes of various studies. Some researchers prefer to use real life tests with obstacles of various sizes (EKKEK *et al.*, 2017) and in different environments (BUJACZ *et al.*, 2022a) (e.g., anechoic or semi-anechoic chambers), others synthesize virtual scenes (ARIAS *et al.*, 2012) or utilize binaural recordings (SCHENKMAN, NILSSON, 2010). Most studies use static tests (THALER *et al.*, 2018; TIRADO *et al.*, 2019) in which a subject just identifies the presence (NILSSON, SCHENKMAN, 2016) or location of obstacles (TONELLI *et al.*, 2016), some studies on the other hand contain dynamic scenarios (in virtual (DODSWORTH *et al.*, 2020) or real life (FIEHLER *et al.*, 2015) settings) in which participants detected the approach to walls (BUJACZ *et al.*, 2022b), obstacles (SCHENKMAN *et al.*, 2016) or navigate simple mazes (DODSWORTH *et al.*, 2020). In this review we analyze these different aspects of the methodologies and wherever possible compare and judge the different approaches.

In the last years, dozens of papers on the subject have been published and a growing interest in human echolocation has been observed (Fig. 1). The most recent extensive reviews of human echolocation research have been proposed by ARIAS *et al.* (2012), KOLARIK *et al.* (2014), THALER and GOODALE (2016). A notable mention is an older review by KISH (2003), probably the currently most known echolocator in the world, who reviewed a large number of the earliest echolocation research. Our review is a continuation and extension of the earlier reviews in the following aspects:

- we provide an up-to-date review of new studies that have been published during the most recent years;
- we include a subdivision of the echolocation studies with respect to a number of different criteria and present them in a tabular form for better browsing through fields by the reader;
- we provide a discussion on and compare different methodologies applied for studying human echolocation.

This paper began as part of a project the goal of which is to compare the usefulness of various artificial and natural sounds for human echolocation. Earlier, we completed echolocation trials for the Echovis project aimed at developing a mobile game for teaching echolocation (BUJACZ *et al.*, 2018; 2021; 2022) and planned to continue the trials in a way that would allow comparison with other previous studies.

Our previous area of research – virtual sound localization and obstacle sonification – has very similar

methodology issues. Many studies tested the influence of various factors, such as personalized Head Related Transfer Functions (HRTFs) or blindness of test participants (DOBRUCKI *et al.*, 2010), on sound externalization and accuracy of source localization, but it was difficult to compare the results of very different methodologies. The subject complexity is also similar – there can be numerous factors influencing the accuracy of sound localization, just as the accuracy of echolocation. We can often confirm that some factors have little influence on the sound localization or echolocation task, but it may be difficult to objectively measure any specific factor's strength considering the overall large variances. This issue is particularly complex in echolocation studies, because echolocation skills vary greatly between individuals (ARIAS, RAMOS, 1997) and most studies use very small groups of participants (even single subjects to represent expert echolocators (WALLMEIER, WIEGREBE, 2014)).

This manuscript is structured to allow a reader to find easily papers that address specific aspects of echolocation. We start by presenting a summary of the collected research (Sec. 2), then go on to compare trials used for the evaluation of echolocation accuracy in static and dynamic scenarios (Sec. 3). Next, we provide an overview of studies analyzing various man-made and synthetic sounds used as echolocation cues (Sec. 4). In Sec. 5, we review research that discusses comparisons of echolocation skills of sighted, inexperienced blind and experienced blind echolocators. Further, we compare the results of the two approaches to echolocation studies (Sec. 6), i.e., in which the researchers conduct live trials and also aid the studies with pre-recorded sounds or renders. Finally, we appraise the review carried out and summarize state of the art of the human echolocation studies.

2. Review of approaches to echolocation research

The selection of scientific papers for the review was an organic process. We searched the main online tools (scholar.google.com, sciencedirect.com, core.ac.uk, and ieeexplore.ieee.org) for research that included testing of echolocation skills or analysis of signals used in human echolocation. Initially, we included only research papers published after 2015, to not repeat information from other reviews, such as (KOLARIK *et al.*, 2016). However, many of the test methods or signal analyses were only found in older papers, so we expanded the search back to 2010, as well as added several key earlier studies that were most frequently cited by the reviewed articles.

For all the reviewed echolocation studies we prepared a short summary of the main methodology, utilized sounds and environments, participants and key

conclusions. This data is presented in Table 1 with the following cells for each paper:

Cell 1: the cited reference;

Cell 2: the title of the study and a brief summary outlining the key results and the most important conclusions;

Cell 3: category of echolocation trial – static (S) or dynamic (D), or if the study concerned only analysis (A) of echolocation sounds. As well as the utilized obstacle sizes, distances, and types of tasks;

Cell 4: subdivides the studies into three categories: (L) live trials that were carried out in real life indoor or outdoor environments, e.g., with obstacles intentionally positioned at different locations versus the tester, (R) trials with pre-recorded or synthesized sounds, e.g., sounds that were first recorded in real environments using a binaural mannequin and then played-back on headphones for the testers in a laboratory environment or generated by a computer, and finally (V) virtual trials in which the echo-sounds were not simply played back, but were a part of a continuously generated virtual environment usually using HRTF filtering. Quite a few studies combined both live (L) and recording (R) tests;

Cell 5: informs how the sound sources were generated, i.e., whether they were synthesized artificially (A) by an electronic device or in a natural (N) manner by the testers themselves, e.g., the mouth-clicks, finger snaps, footsteps or cane taps;

Cell 6: reports on the number of trial participants and categorizes them primarily into blind (B) and sighted (S) participants, though some studies also distinguished early blind (EB) and late blind (LB) persons. Several studies reported participation of echolocation experts (EE), and although no common definition has been given at what level of experience an echolocator becomes one, their skills clearly stood out from the average novice participant.

To the best of our knowledge the table contains the reported studies on human echolocation with special attention focusing on recent reported studies up to the date of submission of this manuscript, i.e., early 2022.

Recommended review papers on human echolocation and auditory perception of the blind are presented in a separate Table 2. Short reviews of the history of echolocation research can also be found in (COOPER *et al.*, 2020; STOCK, 2022).

Table 1. Summary table of reviewed echolocation studies.

1. Author(s), publication date	2. Title – Summary of results and conclusions			
	3. Type of trial: static (S), dynamic (D), analysis (A), not applicable (–)	4. Sound playback: live sounds (L), recordings/ synthesized (R), virtual reality (V)	5. Sound: artificial (A), natural (N)	6. Number of blind (B), sighted (S), early blind (EB) or expert echolocators (EE)
SCHENKMAN, JANSSON (1986)	“The Detection and Localization of Objects by the Blind with the Aid of Long-Cane Tapping Sounds” – Accuracy and detection distance improved along with the obstacle size (from 0.2 to 0.75 m ²), but not for the largest objects (1.5 m ²); – Variance in the tapping sound spectra had no impact on efficacy; – It was difficult to use cane tapping sounds alone without additional sources for echoes.			
	D – walking a path with cardboard obstacles (sized 50 × 30 cm to 1.5 × 1 m) at face level	L – the participants generated cane tapping sounds	N – long-cane tapping sound	3B
ARIAS, RAMOS (1997)	“Psychoacoustic Tests for the Study of Human Echolocation” – Echolocation seems to depend on perception of a virtual pitch that appears from the difference between the outgoing and incoming sounds, this pitch is more easily perceivable when presented with repeated trains of sounds; – Musical training did not influence the subjects’ performance in these pitch discrimination tests; – Noise signals yielded better echolocation results than click sounds when using recordings of real echoes, but the difference was less significant when the echoes were synthesized.			
	S – testers listen to stimuli on headphones	R – synthetic echoes (2–5 ms delay and –3.5 dB) and recorded echoes (50 cm disk at 35 and 80 cm distance)	A – click-sounds, white noise	30S + 1B
ROSENBLUM <i>et al.</i> (2000)	“Echolocating Distance by Moving and Stationary Listeners” – Participants echolocating more accurately while moving than being stationary; – A follow-up confirmed that this moving advantage was not a function of a specific type of training or the multiple stationary positions available during moving echolocation; – The moving advantage might be a function of echoic time-to-arrival information.			
	S/D – echolocating a 91 × 182 cm wall outdoor while standing/moving	L – the participants generated sounds	N – oral sounds of choice	26S
ROJAS <i>et al.</i> (2009)	“Physical Analysis of Several Organic Signals for Human Echolocation: Oral Vacuum Pulses” – From the three compared sound types (oral “ch”, lip “ch”, oral clicks) the palatal clicks were significantly clearer and more intense than alveolar ones and did not interfere with breathing.			
	A – computer analysis	L – the participants generated sounds with their mouths	N – oral “ch”, lip “ch”, oral clicks	10S
ROJAS <i>et al.</i> (2010)	“Physical Analysis of Several Organic Signals for Human Echolocation: Hand and Finger Produced Pulses” – The knuckle vacuum pulse was judged as best due to its high frequency and “interesting symmetry”, containing similar characteristics of palatal clicks with an even richer content in the high frequency part of the spectrum.			
	A – computer analysis	L – the participants generated sounds with their hands	N – knuckle vacuum pulse, hand clap, finger snap	10S + 1B

Table 1. [Cont.].

SCHENKMAN, NILSSON (2010)	“Human Echolocation: Blind and Sighted Persons’ Ability to Detect Sounds Recorded in the Presence of a Reflecting Object” <ul style="list-style-type: none"> – Blind participants performed significantly better than sighted participants; – All participants performed well in locating objects at distances of less than 2 m; – Detection increased with longer signal durations (up to 500 ms noise burst); – Performance was slightly better in an ordinary room than in an anechoic chamber. 			
	S – 0.5 m disk at distances 0.5 m to 5 m	R – participants listened to binaural recordings taken in an ordinary room and an anechoic chamber	A – 5, 50, 500 ms noise bursts	10S + 10B
SCHENKMAN, NILSSON (2011)	“Human Echolocation: Pitch versus Loudness Information” <ul style="list-style-type: none"> – Participants listened to original and altered binaural recordings, which artificially removed pitch or loudness information from the echo signal; – All altered recordings worsened the echolocation correctness, but removal of pitch information affected it more than loudness; – When the pitch information was removed the difference between blind and sighted participants disappeared. 			
	S – 0.5 m diameter disk at distances 1 m to 3 m	R – participants listened to binaural recordings taken in an ordinary room, some with the pitch or loudness information artificially removed	A – 500 ms noise burst	12B + 25S
THALER <i>et al.</i> (2011)	“Neural Correlates of Natural Human Echolocation in Early and Late Blind Echolocation Experts” <ul style="list-style-type: none"> – Processing of click-echoes recruits brain regions typically devoted to vision rather than audition in both early and late blind echolocation experts; – Brain activation was stronger when listening to echoes reflected from moving targets. 			
	S – listening to sounds via headphones in fMRI	R – recordings played back in an MRI machine	A – trains of click sounds with or without echoes	2EE
TENG, WHITNEY (2011)	“The Acuity of Echolocation: Spatial Resolution in Sighted Persons Compared to the Performance of an Expert Who is Blind” <ul style="list-style-type: none"> – Some, but not all novices quickly learned to echolocate small obstacles at short distances at a level comparable to a blind expert; – The paper additionally presents a short review of the numbers of blind participants in 23 echolocation studies from 1950 to 2010 and only in 5 of them there were more than 10 blind participants. 			
	S – sitting 33–75 cm from vertical pair of 5–23 cm disks, judging which is the larger one	L – in a sound-proof, echo-damped room	N – oral clicks	8S + 1EE
SMITH, BAKER (2012)	“Human Echolocation Waveform Analysis” <ul style="list-style-type: none"> – The mouth click waveform is wideband and complex, with spectrum peaks near 3 kHz and 11 kHz and a high fractional bandwidth; – Spectra of early and late blind echolocators’ clicks differ – LB has a wider central peak, but lower side lobes; – The mouth click of the late blind echolocator seems to contain a Doppler-like frequency shift without actual movement. 			
	S – spectral analysis of recorded sounds	R/L – analysis of recorded tongue generated sounds	N – tongue clicks	2B (1 early blind and 1 late blind)

Table 1. [Cont.].

SCHÖRNICH <i>et al.</i> (2012)	<p>“Discovering Your Inner Bat: Echo–Acoustic Target Ranging in Humans”</p> <ul style="list-style-type: none"> – Most participants preferred to use relatively loud, short, broadband tongue clicks with peak frequencies between 5 and 10 kHz (which was noted as much higher than other studies of echolocators’ mouth clicks); – Participants utilized temporal, timbre and spatial cues to assess the distance to a wall; – When comparing consecutive sounds, the sighted participants were able to detect changes of 20–30 cm in the distance to a wall. 			
	S – judging distance changes from a wall at 1.7 to 6.7 m distance	R – artificially generated binaural recordings of echoes with one or two reflective walls	N – tongue clicks	5S
GORI <i>et al.</i> (2014)	<p>“Impairment of Auditory Spatial Localization in Congenitally Blind Human Subjects”</p> <ul style="list-style-type: none"> – Auditory spatial localization along the horizontal axis was found to be severely impaired in the early blind in Bisection tasks (hearing three sound sources in order, then determining whether the middle source was spatially closer to the first or last one); – There was no significant difference between early blind and sighted participants in minimum audible angle resolution tasks (hearing two sounds and determining which one was more to the right). 			
	S – participants sat 180 cm from a perimeter of 23 speakers	R – sound was generated by a bank of speakers	A – 500 Hz tone	27S + 9EB
MILNE <i>et al.</i> (2014)	<p>“The Role of Head Movements in the Discrimination of 2-D Shape by Blind Echolocation Experts”</p> <ul style="list-style-type: none"> – Head movements made while echolocating are necessary for the correct identification of 2-D shape; – Expert echolocators’ performance dropped to chance level when forced to remain still; – Not only experts can use echolocation to successfully identify 2-D shapes. 			
	S – recognizing four geometric shapes 16–100 cm in size at distance 40 or 80 cm	L – sounds generated by the participants in an anechoic chamber or echo-dampened room. Head and torso movements were either allowed or forbidden	N – tongue click, finger snap, speech, hand clap	6EE + 10B + 10S
WALLMEIER, WIEGREBE (2014)	<p>“Ranging in Human Sonar: Effects of Additional Early Reflections and Exploratory Head Movements”</p> <ul style="list-style-type: none"> – Distance discrimination threshold was below 1 m for all reference distances (0.75–4 m) with the best results (20 cm) for the smallest reference distance; – Distance discrimination in complex environments can be improved by allowing free head rotation, but head movements provide no significant advantage over static echolocation from an optimal single orientation. 			
	S/D – distance discrimination from a wall 0.75 m to 4 m	VR – echo generated in virtual echo-acoustic space from participants’ own mouth sounds	N – chosen by a participant	6S + 1B
VERCILLO <i>et al.</i> (2014)	<p>“Enhanced Auditory Spatial Localization in Blind Echolocators”</p> <ul style="list-style-type: none"> – In similar tests as to (GORI <i>et al.</i>, 2014) the blind participants showed much poorer performance than sighted participants in space bisection tasks, but similar performance in minimum auditory angle tasks; – Blind echolocators showed better performance in the spatial bisection tasks than non-echolocating blind participants, showing that the use of echolocation improves auditory spatial localization. 			
	S – discriminating between two of 23 speakers at 180 cm distance	R – sound was generated by a bank of speakers	A – 500 Hz tones, 75 ms, 60 dB (SPL)	11S + 9B

Table 1. [Cont.].

NILSSON, SCHENKMAN (2016)	<p>“Blind People Are More Sensitive Than Sighted People to Binaural Sound”</p> <ul style="list-style-type: none"> – Blind persons show an enhanced sensitivity to inter-aural level difference (ILDs) tests when presented with click pairs in both the leading and the lagging component; – Blind testers showed an increased ability to unsuppress information in lagging clicks. 			
	S – listening to synthetic clicks on headphones	R – sounds composed of 125 ms rectangular pulses (clicks) played over headphones	A – 125 μ s clicks, alone or as pairs spaced 2 ms apart	23B + 65S
FIEHLER, THALER (2015)	<p>“Neural Correlates of Human Echolocation of Path Direction During Walking”</p> <ul style="list-style-type: none"> – All participants were able to differentiate between echo and no-echo stimuli; – Expert blind echolocators performed worse when presented with pre-recorded stimuli during MRI scan; – The observed neural activity suggests that while blind participants processed echo directional meaning automatically, sighted participants had to process information consciously. 			
	D – navigating a corridor and stating its shape, S – listening to recorded sounds during fMRI	L – in indoor and outdoor setup (only 3 blind experts), R – pre-recorded, binaural stimuli	N – mouth clicks	6B + 3S
TONELLI <i>et al.</i> (2016)	<p>“Depth Echolocation Learnt by Novice Sighted”</p> <ul style="list-style-type: none"> – When judging the distance to obstacles the errors in judgements fell from 35 to 10 cm over the course of two one-hour sessions; – Errors were significantly smaller in the reverberant room than in an anechoic chamber; – Participants who used tongue clicks were marginally more accurate than those using finger snaps. 			
	S – subjects sat in front of one of five bars (40–180 cm high and 6–27 cm wide) at five different distances (from 30 cm to 150 cm)	L – the echolocation sound was naturally produced, using no external device	N – tongue clicks + finger snaps	18S
THALER, CASTILLO-SERRANO (2016)	<p>“People’s Ability to Detect Objects Using Click-Based Echolocation – A Direct Comparison between Mouth-Clicks and Clicks Made by a Loudspeaker”</p> <ul style="list-style-type: none"> – Success rates at determining the presence of an obstacle were similar or higher when using a head-worn loudspeaker; – Accuracy in detecting the object was higher at 1 m distance as compared to 2 m; – Sighted participants showed significant improvement in two consecutive sessions. 			
	S – sitting 1 m or 2 m from a 60 cm disk	L/R – in a sound-insulated and echo-acoustic dampened room, participants either generated mouth clicks by themselves or the experimenters generated clicks from a head-worn loudspeaker	N – mouth clicks, A – 4 kHz clicks played through a head-worn loudspeaker	27S + 2B
SCHENKMAN <i>et al.</i> (2016)	<p>“Human Echolocation - Acoustic Gaze for Burst Trains and Continuous Noise”</p> <ul style="list-style-type: none"> – When the obstacle was at 1 m distance the mean accuracy of detecting echoes by blind participants increased with the burst rate (from roughly 60% at 1 burst/500 ms to 80% at 64 bursts/500 ms) and was highest for continuous noise; – For sighted participants and for blind participants at a longer distance of 1.5 m the accuracy was largest at a rate of 32 bursts/500 ms and fell for higher rates; – Of the 38 participants in the study top 5 were blind. 			
	S – 0.5 diameter aluminum disk as the obstacle at 1 m and at 1.5 m	R – binaural echo recordings were made in a lecture hall with reverberations	A – 5 ms noise trains, 1 to 64 bursts per 500 ms versus 500 ms continuous noise	12B + 26S

Table 1. [Cont.].

RYCHTARIKOVA <i>et al.</i> (2017)	“Auditory Recognition of Surface Texture with Various Scattering Coefficients”			
	– From numerous wall shapes tested, two were most likely to be recognized by participants: parabolic (due to sound focusing) and a staircase (due to a chirp-like echo).			
	S – standing at 1.5 m or 10 m from a virtual obstacle	R – synthesized and spatialized echoes played over headphones	A – artificial clicks	16S
KOLARIK <i>et al.</i> (2017)	“Blindness Enhances Auditory Obstacle Circumvention: Assessing Echolocation, Sensory Substitution, and Visual-Based Navigation”			
	– Blind non-echolocators navigated more effectively than blindfolded sighted individuals with fewer collisions; – All participants except the blind echolocation expert navigated better with a sensory substitution device than with echolocation.			
	D – navigating around an obstacle 0.6×2 m	L – participants walked by an obstacle that was directly on or 25 cm off a path. Comparing vision, echolocation and a vibrating distance sensor	N – mouth clicks	10S + 8B + 1EE
EKKEKEL <i>et al.</i> (2017)	“Learning to Echolocate in Sighted People”			
	– A statistically significant improvement was achieved after four days of 1-hour sessions; – The chance to correctly echolocate the position of the larger disk grew proportionally with an angular size difference from 50% (random) for most similar disks to 70% when one disk was 5 cm and the second 25 cm in diameter; – Test participants that did not move their heads during experiments had chance-level results; – The improvement in echolocation ability was positively correlated with performance in an attention PASAT test (Paced Auditory Serial Addition Task), but there was no correlation for spatial cognition and memory tests.			
	S – sitting 50 cm from two disks of different diameters 5–25 cm, determining the position of the larger disk	L – in a soundproof room with sounds generated by a head-mounted small speaker	A – 10 ms white noise pulse (80 dB). As a control, guessing without any sound was also performed	23S
FLANAGIN <i>et al.</i> (2017)	“Human Exploration of Enclosed Spaces through Echolocation”			
	– Participants produced clicks of the length between 3 and 37 ms and absolute sound pressure levels (SPL) between 88 and 108 dB SPL; – Active vocalization was associated with better accuracy of the room size classification; – Visual and parietal activity was observed both in the sighted participants and the blind echolocation expert while performing echolocation.			
	S – listening to synthetic echoes to judge room size changes A – analysis of fMRI during active and passive echolocation	R – participants’ own vocalizations were recorded and convolved with BRIR measurements of a small chapel with highly reflective surfaces	N – mouth clicks recorded for each participant	11S + 1B
HELLER <i>et al.</i> (2017)	“Evaluating Two Ways to Train Sensitivity to Echoes to Improve Echolocation”			
	– Participants were divided into three groups, two trained echo sensitivity using a lab procedure or an app, and the third was a control group; – Pre and post training tests involved localization of a 0.6×1.2 m board at distances from 0.9 to 2.7 m; – Both training groups showed similar improvement after 15 hours of training, although supervised psychoacoustic training in the lab was marginally better.			
	S – listening to synthetic echoes for training and localizing a 0.6×1.2 m board for pre and post tests	R – synthetic echo sounds were used for training L – mouth clicks were used in live pre and post tests	N – recorded mouth clicks selected to meet optimal characteristics (ROJAS <i>et al.</i> , 2009)	13S

Table 1. [Cont.].

THALER <i>et al.</i> (2017)	“Mouth-clicks Used by Blind Expert Human Echolocators – Signal Description and Model Based Signal Synthesis” <ul style="list-style-type: none"> – Analyzed mouth clicks were wideband (up to 10 kHz), consistently very brief (~3 ms duration) with peak frequencies in the range of 2–4 kHz, and maximum energy at 10 kHz; – MATLAB code to synthesize the model clicks was made available in the supplementary material and has been utilized in a number of later echolocation studies (BUJACZ <i>et al.</i>, 2018; DODSWORTH <i>et al.</i>, 2020; FLANAGIN <i>et al.</i>, 2017; RYCHTARIKOVA <i>et al.</i>, 2017; THALER <i>et al.</i>, 2020a; TIRADO <i>et al.</i>, 2019). 			
	A – analysis of expert mouth clicks	L – experts generated clicks in an echo-dampened room	N – mouth clicks	3EE
THALER, FORESTEIRE (2017)	“Visual Sensory Stimulation Interferes with People’s Ability to Echolocate Object Size” <ul style="list-style-type: none"> – Visual stimulation (white light) decreased the sighted participants’ echolocation performance; – Tactile stimulation (skin electrode) had no effect on echolocation performance in sighted and blind people; – The same areas of the brain seem to be involved in processing of both the visual stimuli and echo sounds. 			
	S – sitting 50 cm from two disks 5–25 cm, determining the position (top/bottom) of the larger disk spaced 27 cm apart	L – carried out in a sound-insulated, and echo-acoustic damped room	N – mouth clicks	44S + 3B
NORMAN, THALER (2018)	“Human Echolocation for Target Detection is More Accurate with Emissions Containing Higher Spectral Frequencies, and This is Explained by Echo Intensity” <ul style="list-style-type: none"> – Echolocation was more accurate using emissions with higher spectral frequencies – this advantage was eliminated when the intensity of the echoes was artificially equated to correct for the higher reflectivity of the tested object in the higher spectral range. 			
	S – listening to binaural recordings of reflections from 0.5 m diameter disc at distances 1–3 m	R – recordings made in an anechoic chamber using a custom binaural mannequin	A – synthetic clicks or noise bursts with 9 dB bursts of 3.5–4.5 Hz frequencies	12S
THALER <i>et al.</i> (2018)	“Human Echolocators Adjust Loudness and Number of Clicks” <ul style="list-style-type: none"> – Echolocators accumulate information from multiple samples; – To locate objects off to the sides, the echolocators increased loudness and numbers of clicks; – Echolocation in the Frontal Hemisphere is Better than in the Rear. 			
	S – locating a 17.5 cm disk at 100 cm distance and 0–180° azimuth angles	L – Participants generated clicks by themselves in a noise insulated and echo dampened room	N – mouth clicks	8B
TONELLI <i>et al.</i> (2018)	“How Body Motion Influences Echolocation While Walking” <ul style="list-style-type: none"> – Head exploration (i.e., changing head rotation angle while producing sounds) is crucial for acquiring spatial data; – Echolocation accuracy depends on the distance to an obstacle and the frequentness of head movements during sound emission; – Average velocity, motion duration, and time of the task completion do not significantly influence the correctness of the echolocation task. 			
	D – walking a 4 m long, 1.1 m wide corridor and stating its shape (closed or open to left or right)	L – participants generated clicks by themselves in a larger high-ceiling room with a corridor build from plastic panels	N – mouth clicks	9S

Table 1. [Cont.].

ANDRADE <i>et al.</i> (2018)	<p>“Echo-House: Exploring a Virtual Environment by Using Echolocation”</p> <ul style="list-style-type: none"> – Echolocation provided information on orientation and sense of space that would not otherwise be available; – Echolocation itself did not allow participants to navigate in this environment without additional support, but it did help in locating objects and exploring the environment. 			
	D – controlled an avatar in a virtual environment	V – participants controlled an avatar placed in virtual space	A – footsteps, mouth-clicks, hand clapping	5B
THALER <i>et al.</i> (2019)	<p>“Human Click-Based Echolocation of Distance: Superfine Acuity and Dynamic Clicking Behaviour”</p> <ul style="list-style-type: none"> – Echolocators made more intense and more frequent clicks when dealing with weaker reflections (i.e., the same object at a farther distance, or a smaller object at the same distance); – Number and intensity of clicks were adjusted independently from one another; – Experienced echolocators reliably detected changes in distance of roughly 5% (3 cm at 50 cm, and 7 cm at 150 cm distance). 			
	S – localizing change of distance to disks (28.5 cm or 80 cm diameter) placed at 50 cm or 150 cm	L – Participants generated clicks by themselves. A noise insulated and echo dampened room	N – mouth clicks	8B
TIRADO <i>et al.</i> (2019)	<p>“The Echobot: An Automated System for Stimulus Presentation in Studies of Human Echolocation”</p> <ul style="list-style-type: none"> – A 50 cm reflecting disk was correctly detected at distances 1 to 3.3 m, with an average of 2 m; – Participants showed a small, but steady improvement over 12 echolocation sessions lasting 6–10 min. each, but only when a synthetic clicker was used; – Participants using their own mouth sounds showed no changes in their detection thresholds. 			
	S – sitting in front of a 50 cm aluminum disc repositioned by an automated sled to distances 1–4 m	L – in sound-proofed and padded listening lab	A – synthesized click (THALER <i>et al.</i> , 2017) N – mouth clicks (3 participants)	15S
THALER <i>et al.</i> (2020b)	<p>“The Flexible Action System: Click-Based Echolocation May Replace Certain Visual Functionality for Adaptive Walking”</p> <ul style="list-style-type: none"> – Echolocation experts walked just as fast as sighted participants using vision; – Participants who made clicks with higher spectral frequency content and higher clicking rates walked faster; – The use of echolocation significantly decreased the frequency of collisions with obstacles at head height, but not at ground level. 			
	D – walking across a room and around obstacles	RL – participants generated clicks by themselves in a padded room with two obstacles (80 × 80 cm) at head and ground level	N – mouth clicks	10B + 7EB + 24S
DODSWORTH <i>et al.</i> (2020)	<p>“Navigation and Perception of Spatial Layout in Virtual Echo-Acoustic Space”</p> <ul style="list-style-type: none"> – Sighted people after 10-week training in virtual mazes increased their ability to judge the spatial layout of obstacles through sound, avoid collisions and find safe passage; – Blind echolocators performed at a very high level without any training. 			
	D – navigation with a computer keyboard	V – passing through virtual mazes with walls 75 cm apart	A – synthesized click (THALER <i>et al.</i> , 2017)	20S + 3B

Table 1. [Cont.].

SCHENKMAN, GIDLA (2020)	“Detection, Thresholds of Human Echolocation in Static Situations for Distance, Pitch, Loudness and Sharpness” <ul style="list-style-type: none"> – The repetition pitch was useful for detection at shorter distances and was determined from the peaks in the temporal profile of the autocorrelation function; – At shorter distances loudness provides echolocation information, but at longer distances, timbre aspects, such as sharpness, might be used to detect objects; – Results suggest that blind persons may detect objects at lower values for loudness, pitch strength and sharpness and at further distances than sighted persons. 			
	S – recorded reflections from a 0.5 m disk at distances from 0.5 to 5 m	R – binaural recordings in an ordinary conference room and an anechoic chamber played back over headphones	A – 5, 50, and 500 ms noise burst from a loudspeaker	10B + 10S
NORMAN, THALER (2020)	“Stimulus Uncertainty Affects Perception in Human Echolocation: Timing, Level, and Spectrum” <ul style="list-style-type: none"> – When there was certainty in the acoustic properties of the echo relative to the emission, either in temporal onset, spectral content or level, people detected the echo more accurately; – Participants were more accurate when the emission’s spectral content was certain, but surprisingly, not when either its level or temporal onset was certain. 			
	S – recorded reflections from a 50 cm disc or a 28 cm bowl at 1.2 or 3 m	R – binaural recordings	A – clicks and 500 ms white noise bursts from a loudspeaker	4EE + 20B + 24S
TONELLI <i>et al.</i> (2020)	“Early Visual Cortex Response for Sound in Expert Blind Echolocators, But Not in Early Blind Non-Echolocators” <ul style="list-style-type: none"> – Activation in the posterior area of the scalp while echolocating for the sighted was similar to the one observed in early blind experts; – This activity was associated to sound stimulation and is contralateral to the sound localization in space. 			
	S – participants sat in front of the set-up	L – live played sound via 23 speakers	A – 500 Hz 60 dB pure tone, duration of 75 ms	10B + 5S
TIRADO <i>et al.</i> (2021)	“Comparing Echo-Detection and Echo-Localization in Sighted Individuals” <ul style="list-style-type: none"> – Distinct individual differences in echo-detection and echo-localization abilities; – Better performance in the echo-detection than the echo-localization task; – It may be relevant for echolocation training programs to focus separately on the detection and localization. 			
	S – 50 cm disk at distances from 1 m to 4.25 m	R – synthetic expert mouth clicks played over a loudspeaker in an echo-dampened room	A – synthesized click (THALER <i>et al.</i> , 2017)	10S
ANDRADE <i>et al.</i> (2021)	“Echolocation as a Means for People with Visual Impairment (PVI) to Acquire Spatial Knowledge of Virtual Space” <ul style="list-style-type: none"> – Various techniques were used to describe the virtual space, including perimeter recognition tactics, listing elements and describing holistic map models; – People with Visual Impairment could distinguish whether a virtual room was covered with carpet, wood or metal, identify the relative size of a virtual room, and detect the presence of 90° turns to the left or right on average 70% of the time; – Working with PVI and learning from their lived experience is the most successful way to gain knowledge of technologies accessible to PVI. 			
	D – using the Xbox controller to explore the virtual space	V – travel through virtual world	A – pre-recorded sound, echo generated by the footprint of the avatar	12B

Table 1. [Cont.].

CASTILLO-SERRANO (2021)	<p>“Increased Emission Intensity Can Compensate for the Presence of Noise in Human Click-Based Echolocation”</p> <ul style="list-style-type: none"> – The emission intensity increased so that the spectral power of echoes exceeded the spectral power of noise by 12 dB in 4 kHz and 5 kHz frequency bands; – A potential strategy to deal with noise while echolocating is to increase emission intensity to maintain the signal-to-noise ratio of certain spectral components of the echoes. 			
	S – recordings of 17.5 cm or 26.5 cm disk at 1, 2, or 3 m	R – binaural recordings made in an echo-acoustic dampened room played through headphones	A – synthetic click (a 4.5 kHz sinusoid multiplied by a decaying exponential)	8B + 3EE + 20S
KRITLY <i>et al.</i> (2021)	<p>“Discrimination of 2D Wall Textures by Passive Echolocation for Different Reflected-to-Direct Level Difference Configurations”</p> <ul style="list-style-type: none"> – The discriminability is larger for the walls reflecting with a higher spectral coloration; – Enhancing the reflections as well as removing the direct sound are beneficial to differentiate textures; – The flat wall and the circular wall are the most difficult textures to discriminate, the wall with aperture and the staircase are the most distinguishable textures. 			
	S – synthesized reflection from six different wall shapes at distances from 0.8 to 5 m	R – recordings played through headphones	A – a single anechoically recorded click sound with the synthesized echo	14S
NORMAN, THALER (2021)	<p>“Perceptual Constancy With a Novel Sensory Skill”</p> <ul style="list-style-type: none"> – Blind expert echolocators have higher constancy ability than sighted and blind persons novices to echolocation; – Sighted participants improved their capabilities through training; that suggests that constancy also occurs in a domain with which the respondent has had no previous experience. 			
	S – recorded reflections from a 50 cm disc or a 28 cm bowl at 1, 2 or 3 m	R – recordings played through headphones	A – variations in the click’s peak spectrum were used: 3.5, 4.0, and 4.5 kHz	10S + 17B + 3EE
NORMAN <i>et al.</i> (2021)	<p>“Human Click-Based Echolocation: Effects of Blindness and Age, and Real-Life Implications in a 10-Week Training Program”</p> <ul style="list-style-type: none"> – Training improved performance of both sighted and blind participants, but neither group reached the level of experienced experts; – Some sighted participants performed better than the blind novices after the same training, though this can be attributed to younger age and/or superior binaural hearing; – The ability to learn click-based echolocation is not strongly limited by age or level of vision. 			
	S – discriminating disc size (THALER, FORESTEIRE, 2017) or orientation D – navigating a simple virtual T, U or Z maze and a real natural environment	V – virtual mazes with recorded clicks L – live tasks with participant mouth clicks in an echo-dampened room	N – mouth clicks (live and prerecorded)	14S + 12B + 7EB
BUJACZ <i>et al.</i> (2022a)	<p>“Echovis – A Collection of Human Echolocation Tests Performed by Blind and Sighted Individuals: A Pilot Study”</p> <ul style="list-style-type: none"> – Better results were achieved for outdoor tests than indoors and the worst in a padded room; – Additional signal emissions marginally helped in determining an obstacle’s direction, but not a distance; – Blind and sighted participants performed similarly in most tests, statistically significant difference was found only for determining the distance to an obstacle; – A high correlation between certainty in answers and their real correctness was noted for all adult participants, but not for blind children; – In dynamic trials the average click rate when using a mechanical clicker was once every 2 seconds. 			

Table 1. [Cont.].

	S – localizing a 2 m wooden wall at distances 1–3 m, D – approaching a wall, walking parallel to a wall, localizing an off-the path object	L – similar tests performed outdoors and indoors; static indoor tests were compared in an empty room and in an acoustically padded room, as well as with binaural recordings (R) in the same environments	A – mechanical clicker or synthesized expert click from (THALER <i>et al.</i> , 2017)	10B + 10S (+ 10B children)
BUJACZ <i>et al.</i> (2022b)	<p>“Comparison of Echolocation Abilities of Blind and Normally Sighted Humans using Different Source Sounds”</p> <ul style="list-style-type: none"> – Almost all blind and sighted participants performed significantly above random; – Blind participants performed significantly better than the sighted ones; however, the difference disappeared once the blind participants were analyzed as two separate groups – totally blind vs visually impaired; – Legally blind participants that retained any level of light sensitivity performed on average the same as sighted participants; – From the ten analyzed sounds pink and blue noises along with 3 kHz and 4 kHz percussion were significantly best for accuracy of the echolocation. 			
	S – localizing a 1 × 2 m vertical wall at distances 1–3 m and directions –45° to 45°	L – outdoors using ten different sounds generated by the participant or played from a BT speaker at waist-height	N – mouth clicks or hand clapping A – 1–5 kHz percussion, pink and blue noise, mechanical clicker, synthesized expert click (THALER <i>et al.</i> , 2017)	12B + 14S

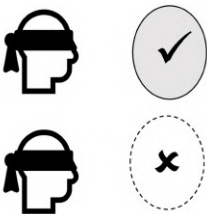

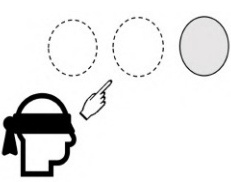
Table 2. Summary table of recent review papers.

KISH (2003)	<p>“Sonic Echolocation: A Modern Review and Synthesis of the Literature”</p> <ul style="list-style-type: none"> – Paper written by a blind echolocation expert; – Extensive review of the early literature on echolocation, including early misconceptions about “facial vision” from the first half of the XX century and many practical experiments from the 60s and 70s; – Review of studies testing various aspects of echolocation including the use of different targets and different sonic sources; – Review of studies on the learning of echolocation by sighted subjects and proposals of training programmes for the blind.
ARIAS <i>et al.</i> (2012)	<p>“Echolocation An Action-Perception Phenomenon”</p> <ul style="list-style-type: none"> – Review paper presenting a historical categorisation of the main studies concerning echolocation; – The authors conclude that echolocation is a “closed-loop perception-action behaviour, in which the subject modulates action (self-generated echolocation signals, exploratory head movements) to control perception (auditory Gestalts learned through implicit learning)”.
KOLARIK <i>et al.</i> (2016)	<p>“Auditory Distance Perception in Humans: A Review of Cues, Development, Neuronal Bases, and Effects of Sensory Loss”</p> <ul style="list-style-type: none"> – A review paper focusing on four aspects of auditory distance perception: cue processing, development, consequences of visual and auditory loss, and neurological bases; – Blind individuals often manifest supra-normal abilities to judge relative distance but show a deficit in absolute distance judgments; – Following hearing loss, the use of an auditory level as a distance cue remains robust, while the reverberation cue becomes less effective.
THALER, GOODALE (2016)	<p>“Echolocation in Humans: an Overview”</p> <ul style="list-style-type: none"> – A review paper summarizing the history of echolocation studies, analyzing the typical mission signal; – An assessment of distance, direction and size discrimination is provided from several studies; – A large review of neural underpinnings of echolocation, especially the plasticity of the brain to adapt “visual” areas to process echolocation signals.

Table 2. [Cont.].

KOLARIK <i>et al.</i> (2021)	<p>“A Framework to Account for the Effects of Visual Loss on Human Auditory Abilities”</p> <ul style="list-style-type: none"> – The paper reviews numerous studies related to the impact of vision loss on spatial and non-spatial auditory perception; – Authors propose a framework comprising a set of nine principles that can be used to predict and explain why given auditory abilities are enhanced or degraded after the loss of vision; – Effects of early, late, partial and full visual loss are also discussed; – The framework includes a Perceptual Restructuring Hypothesis that posits utilization of available cortical resources to provide the most accurate and useful information, sometimes at a loss of some auditory abilities.
---------------------------------	--

Table 3. Static echolocation tests.

Binary – state the presence or absence of an obstacle	
<p>Examples:</p> <p>A disc (50 cm diameter) placed at 1 or 1.5 m (SCHENKMAN <i>et al.</i>, 2016), 0.5–5 m (SCHENKMAN, NIESSEN, 2010) or at 1–3 m (SCHENKMAN, NIESSEN, 2011; NORMAN, THALER, 2018);</p> <p>A disc (60 cm diameter) placed directly at 1 to 2 m (THALER, CASTILLO-SERRANO, 2016), or 1, 2 or 3 m (NORMAN, THALER, 2021);</p> <p>A disc (17.5 cm diameter) placed 1 m at different azimuth angles (from 0° – directly in front to 180° – directly behind) (THALER <i>et al.</i>, 2018);</p> <p>A disc (50 cm diameter) placed from 0.7 to 3.9 m and moved further or closer based on the correct or incorrect answer (TIRADO <i>et al.</i>, 2019).</p>	 <p>e.g., “Is there an object in front of you?”</p>
Distinguish between objects	
<p>A reference disc (diameter 25.4 cm) and 5 comparison discs (diameter 5.1–22.9 cm) placed at different distance 0.33 m, 0.5 m or 0.75 m (TENG, WHITNEY, 2011);</p> <p>Four geometrical shapes: rectangle 100 × 16 cm vertically or horizontally, square 40 cm, triangle 52 cm wide and 45 cm high (MILNE <i>et al.</i>, 2014);</p> <p>A reference disc (diameter 25.4 cm) and 5 comparison discs (diameter 5.1–22.9 cm) placed 0.5 m away (EKKEK <i>et al.</i>, 2017);</p> <p>Two distinct architectural structures from a distance 1.5 m or 10 m (RYCHTARIKOVA <i>et al.</i>, 2017);</p> <p>Distinguish which wall was more reflective (KRITLY <i>et al.</i>, 2021).</p>	 <p>e.g., “Which is the larger object?”</p>
Determine direction and/or distance to obstacle:	
<p>A wall (1.83 m × 0.914 m × 1.27 cm) placed at 0.91 m, 1.83 m, 2.74 m or 3.66 m from the starting point (ROSENBLUM <i>et al.</i>, 2000);</p> <p>A virtual reflective surface placed 1.7–6.8 m in front or 1.7 m at an angle 15–45° (SCHÖRNICH <i>et al.</i>, 2012);</p> <p>Rectangular bars (length 40–180 cm, width 6–27 cm) placed at 0.3–1.5 cm (depending on obstacle size) (TONELLI <i>et al.</i>, 2016);</p> <p>A disk (28.5 cm or 80 cm diameter) placed 0.5 m or 1.5 m from a participant (THALER <i>et al.</i>, 2019);</p> <p>The 1 × 2 m wall at distances 1 to 3 m (BUJACZ <i>et al.</i>, 2018; 2022b);</p> <p>The 60 × 120 cm board at a distance 90 to 270 cm (HELLER <i>et al.</i>, 2017).</p>	 <p>e.g., “Where is the object?”</p>

3. Static versus dynamic trials

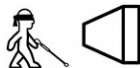
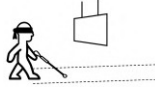
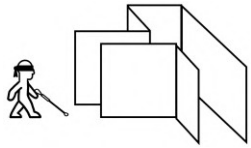
A good way to subdivide test methodologies are static and dynamic trials. In the static trials the test participant is not moving and localizes real or virtual targets of different types (the most common being circular disks 50 cm in diameter) at different distances (from 30 cm up to 5 m) or directions. In moving trials the echolocator travels through a simple controlled environment localizing one or more obstacles or navigating simple mazes. Tables 3 and 4 summarize the most common types of tests.

Most of the studies devoted to human echolocation are based on static experiments. This is because such tests are more straightforward to plan, carry out, and the results are simpler to analyze and interpret. The participants sit or stand and provide answers about the direction and distance of objects positioned in the environment. Trials that utilize recordings or renders can also be generally regarded as static, though they are discussed in a separate section.

The static tests can be divided into four main categories: binary tests, distance, location or size/type discrimination tests (THALER, GOODALE, 2016). In binary tests, participants simply state the presence

of an obstacle or the lack of thereof. A frequently used object for detection is a disc, e.g., 50–100 cm in diameter, and placed 1–2 m in front of test subjects who produce the echolocation sound themselves (THALER, CASTILLO-SERRANO, 2016) or only listen to the recordings (SCHENKMAN *et al.*, 2016). The disc in the binary test is usually not removed entirely, but rotated 90° as to present a narrow, non-reflecting edge to the participant. The binary test can be modified by placing a disk at an angle to the participants. While an obstacle displacement up to 90° does not affect the overall performance significantly, there was a sudden accuracy decrease observed at 135° (THALER *et al.*, 2018). Another modification to a binary test was implemented in the study by (TIRADO *et al.*, 2019). A distance to an obstacle was modified based on the accuracy of the participants' answers. An obstacle was not removed from a setup, only turned perpendicular to a test subject (non-reflective mode). Correct identification of a reflective mode increased the distance by 0.25 m, correct identification of a non-reflective mode did not change the distance. False-negative identification decreased the distance by 0.25 m and false-positive identification decreased the distance by 0.5 m. While simple in design, the binary tests provide

Table 4. Dynamic trials.

Approach a wall or an obstacle	
<p>Detected a wall ($1.8\text{ m} \times 0.9\text{ m} \times 1.3\text{ cm}$) placed at 0.9 m, 1.8 m, 2.7 m or 3.7 m from the starting point while moving along a guide string (ROSENBLUM <i>et al.</i>, 2000);</p> <p>Approached a wall from a random distance (3, 4 or 5 m) to stop at touch distance (BUJACZ <i>et al.</i>, 2022a).</p>	 <p>e.g., “Walk to the wall and stop before it”</p>
Travel a path and detect obstacles on or off the path	
<p>Navigated the length of the large room ($24 \times 15\text{ m}$) with four obstacles ($1.46 \times 1.03\text{ m}$, $0.73 \times 1.03\text{ m}$, $0.515 \times 0.73\text{ m}$, and $0.365 \times 0.51\text{ m}$ pieces of cardboard, suspended from metal racks so that the obstacle midpoint was placed at 1.44 m from the ground), placed 7, 11, 15, and 19 m from the starting point (SCHENKMAN, JANSSON, 1986);</p> <p>Navigated a corridor built of wooden panels ($1.85 \times 1.1\text{ m}$) of different shapes (opened to the left or right, closed from both sides) (FIEHLER <i>et al.</i>, 2015);</p> <p>Navigated the length of the room ($5.8 \times 9 \times 3\text{ m}$) with two obstacles ($80 \times 80\text{ cm}$ polystyrene blocks), placed 2.1–4 m from a starting plane at a different height (THALER <i>et al.</i>, 2020a);</p> <p>Localized obstacles placed off the path (car, streetlight, open door, end of wall) (BUJACZ <i>et al.</i>, 2022a).</p>	 <p>e.g., “Walk the path and avoid the face level obstacle”.</p>
Navigate in an artificial “maze” or other environment	
<p>Navigated a corridor built of wooden panels ($1.85 \times 1.1\text{ m}$) of different shapes (opened to the left or right, closed from both sides) (FIEHLE <i>et al.</i>, 2015);</p> <p>Navigated a corridor built of poly-methyl methacrylate panels ($4 \times 1.1\text{ m}$) of different shapes (opened to the left or right, closed from both sides) (TONELLI <i>et al.</i>, 2018);</p> <p>Navigate a virtual maze (DODSWORTH <i>et al.</i>, 2020; NORMAN, THALER, 2021).</p>	 <p>e.g. “Find the corridor that is not a dead end”.</p>

information not only on the range and resolution of effective echolocation, but they also allow to collect information on optimal echolocation sound parameters under controlled conditions (THALER, CASTILLO-SERRANO, 2016).

Another type of a static test concerns distinguishing between two types of objects, e.g., big and small. The size discrimination test usually takes the form of two-alternative forced-choice task. The two objects are placed at the same distance and presented to a participant simultaneously. The test subject must indicate where the bigger object is located (TENG, WHITNEY, 2011). Alternatively, in the study by RYCHTARIKOVA *et al.* (2017) participants were asked to differentiate between two distinct architectural structures (staircase and different types of walls: parabolic, sinusoid, periodic squared, broad, narrow, convex circular, and a narrow wall with an aperture).

As far as distance discrimination is concerned, the participants are presented with an obstacle placed at a different distance. Their task is to report the relative distance to an object. The obstacles of different sizes can be utilized in this type of test, with the object size increasing along with the distance (TONELLI *et al.*, 2016). Two obstacles of the same size can be also used, the first one as a reference and the second one placed at an angle (SCHÖRNICH *et al.*, 2012).

The types, sizes and distances of objects/obstacles are listed in cells 3. of Table 1 for the various echolocation studies as well as summarized in Table 2. Most obstacles/objects range 20–60 cm in size and 1–2 m in distance from the observer, though large walls or panels are also sometimes used.

An important methodology question has been whether to conduct echolocation studies in echoic or anechoic environments (KOLARIK *et al.*, 2014). On the one hand, it can be expected that in an anechoic environment a subject could better focus only on the single reflection from an object or obstacle used during the test. On the other, anechoic environments are very unnatural to humans, make loudness judgements more difficult, and provide no background to perceive an “acoustic shadow” – the blocking of more distant echoes (BUJACZ *et al.*, 2018). Luckily, this matter has more or less been settled, as a number of studies have demonstrated that obstacle detection in anechoic or acoustically dampened settings is marginally (BUJACZ *et al.*, 2018) or even significantly worse than in natural environments (TONELLI *et al.*, 2016).

An important observation was made by MILNE *et al.* (2014) who noticed that expert echolocators could determine the shape of objects with exceptional accuracy when they were allowed to make head movements. These results can be explained by other studies that noted that blind people are more sensitive to interaural level (ILD) differences than the sighted individuals (NILSSON, SCHENKMAN, 2016). Also, WALLMEIER

and WIEGREBE (2014) observed that when it comes to distance discrimination, head movements in a static position did not much improve echolocation performance. On the other hand, when the tester changed its reference positions the distance discrimination of objects has improved.

Here, we can state that, although the static tests have brought important insight into human echolocation abilities, they are far from real live situations in which the visually impaired would use echolocation in practice. The dynamic echolocation tests were carried out mainly with participation of expert echolocators.

An interesting approach to testing echolocation abilities in dynamic settings was proposed by DODSWORTH *et al.* (2020) who underlined the importance of “active” navigation tasks for safe mobility and wayfinding. They made binaural acoustic recordings in real environments that were later replayed to test participants, who moved in the replicated virtual spaces. Such an approach is worth further studies because the results show that sighted people after 20 virtual navigation training sessions acquire and generalize navigation abilities using echo-acoustics. Also, the three blind echolocator experts were able to complete similar virtual navigation tasks without any training.

Another recent study by TONELLI *et al.* (2018) has been the first to investigate the influence of the body motion in real environments on echolocation abilities. The authors of the study built a corridor of complex geometries composed of sound-reflecting panels and asked the blindfolded sighted individuals, without prior echolocation experience, to move in such model spaces. The trial participants used mouth clicks to explore the space. The results confirm that kinematic activity of an individual such as walking and a stopping pattern and also head movements allow him/her to successfully navigate in new environments by the use of self-generated echoes.

We can conclude that from numerous studies we have acquired a good understanding of human echolocation abilities confirmed in the static experiments. However, studies of human-echolocation in dynamic experiments, i.e., while the test participant actively explores the environment, are sparse and few. We see two prospective research directions in this context. First, echolocation while moving in virtual reality environments, although difficult to simulate, can be a good solution (DODSWORTH *et al.*, 2020). Second, the research initiated by TONELLI *et al.* (2018) should be expanded and concentrate on echolocation abilities while the trial participant is in motion in real environments. Results of such studies can bring new insights into the interrelation between the body motion and space exploration capabilities of the visually impaired.

The key observations from the static echolocation trials carried out with blind and sighted participants are the following:

- echolocation can be learnt and trained by sighted people (NORMAN, THALER, 2021);
- experienced echolocators significantly outperform novices (NORMAN, THALER, 2020; VERCILLO *et al.*, 2014);
- expert echolocators can detect changes in a distance of 3 cm at a reference distance of 50 cm, and a change of 7 cm at a reference distance of 150 cm (THALER *et al.*, 2019).

The conclusions from a few dynamic echolocation trials are the following (THALER *et al.*, 2020b):

- echolocation experts walked just as fast as sighted participants using vision;
- participants who made clicks with a higher spectral frequency content and higher clicking rates walked faster;
- the use of echolocation significantly decreased collision occurrences with obstacles at head height, but not at ground level.

4. Sound sources – artificial versus natural

There are numerous ways to produce sound sources that serve as the origin signal for the echoes used in echolocation. Early echolocation research in the first half of the XX century had to verify experimentally that the blind participants of their tests were using sounds (e.g., of their own footsteps or cane taps) to detect obstacles (KISH, 2003). Now that the phenomenon of echolocation is much better understood, there has been a growing interest in determining the influence of a sound source on echolocation, trying to analyze and even potentially optimize it (THALER *et al.*, 2017).

Currently, the list of sounds used by the blind for echolocation is quite long: there are mouth or hand-made sounds (such as clicks, finger snaps, clapping or

knuckle vacuum pulses), mechanical sounds (cane taps, mechanical clickers or castanets) and artificially synthesized sounds played from speakers, such as modelled clicks, white or pink noise bursts or rectangular pulses. Table 5 summarizes this division and in this section we discuss key studies related to testing or analyzing sound sources used for echolocation.

All signals that could be used in human echolocation can be categorized into the two main groups: artificial and natural sounds. Research on natural sounds can be divided into mouth and hand-made signals. ROJAS *et al.* (2009) have examined many natural generated sounds such as palatal clicks, oral “ch” (sound of tongue moving backwards from teeth), lip “ch” (quick munching), finger snapping and hand clapping, an “iu” sound vocalization or whistling to imitate bat chirps. These natural sounds were analyzed with respect to usability, reproducibility and intensity. The results suggest that the oral produced click is the most suitable for human echolocation. Its spectrum consists of clearly separated frequency bands. The signal energy concentrates on average at a frequency of 1.15 kHz, although the study only tested 10 sighted volunteers. In a follow-up study it was shown that the oral clicks are effective in the presence of ambient noise (ROJAS *et al.*, 2010).

In a different study, SMITH and BAKER (2012) report that the tongue-click generated by an expert echolocator is a complex sound and feature a wide spectrum band. In their group of that the spectrum peak of a tongue-click is located at 3 kHz, and its bandwidth is located within the range of 1.5 kHz to 4.5 kHz. The authors also conclude that it is the large fractional bandwidth (spectrum width) of the click that gives it great range resolution.

Results from the study conducted by THALER and CASTILLO-SERRANO (2016) show a difference in detec-

Table 5. Commonly tested natural and artificial sound.

Natural	Artificial
Mouth-made sounds: – tongue clicks (FIEHLE <i>et al.</i> , 2015, 2015; HELLER <i>et al.</i> , 2017; ROJAS <i>et al.</i> , 2008; SMITH, BAKER, 2012; TENG, WHITNEY, 2011; THALER <i>et al.</i> , 2017, 2018, 2019; THALER, CASTILLO-SERRANO, 2016; TONELLI <i>et al.</i> , 2016, 2018); – oral “ch”, lip “ch”, whistling (ROJAS <i>et al.</i> , 2008); – unvoiced consonant “s” (SCHÖRNICH <i>et al.</i> , 2012).	Mechanical-made sounds: – cane taps (ARIAS, RAMOS, 1997; SCHENKMAN, JANSSON, 1986); – mechanical clickers (ARIAS, RAMOS, 1997; BUJACZ <i>et al.</i> , 2018).
Hand-made sounds: – finger snapping (ROJAS <i>et al.</i> , 2008); – hand clapping (ROJAS <i>et al.</i> , 2010; TONELLI <i>et al.</i> , 2016); – knuckle vacuum pulses (ROJAS <i>et al.</i> , 2010).	Computer-made sounds: – synthetic clicks (BUJACZ <i>et al.</i> , 2018; 2022b; DODSWORTH <i>et al.</i> , 2020; HELLER <i>et al.</i> , 2017; NILSSON, SCHENKMAN, 2015; THALER <i>et al.</i> , 2011, 2017 2020a; THALER, CASTILLO-SERRANO, 2016; TIRADO <i>et al.</i> , 2019); – noise (white or pink) (ARIAS, RAMOS, 1997; EKKEL <i>et al.</i> , 2017; GORI <i>et al.</i> , 2014; SCHENKMAN <i>et al.</i> , 2016); – transient trains (ARIAS, RAMOS, 1997); – short noise bursts (ARIAS, RAMOS, 1997; NILSSON, SCHENKMAN, 2016; SCHENKMAN <i>et al.</i> , 2016).

tion accuracy between the sounds generated by a tongue and artificially generated clicks produced by a head-worn speaker in a sighted participant group. During echolocation sessions with the use of a loudspeaker and an obstacle positioned at a distance of 1 m, echolocators were more accurate in locating an obstacle ($M = 0.653$, $SD = 0.161$) than in sessions in which natural sounds were generated with a tongue ($M = 0.579$, $SD = 0.093$). However, while performing the same tests at a distance of 2 m object localization accuracies were comparable, with slightly better results obtained with the use of artificially generated clicks. When the tests were repeated, the echolocation precision of the testers improved, with significantly better results for the speaker-generated echolocation sounds.

THALER and CASTILLO-SERRANO (2016) tested the echolocation abilities of two blind echolocators. The first subject with a longer experience performed perfectly in each trial. The second person was less accurate, but still performed much better than the sighted participants. This person preferred using tongue generated sounds.

EKKEL *et al.* (2017) conducted trials with twenty-three sighted participants in a soundproof room 2 to examine peoples' ability to discriminate size of objects by using echolocation techniques. Among all the tests, they compared results with no sound generated and with the use of white noise produced by a small speaker that was attached to participants' foreheads. Obstacles were positioned at different angular directions. Although, the echolocation results with white noise were better than chance, the authors concluded that the differences were not statistically significant ($p = 0.052$).

In a recent study by TIRADO *et al.* (2019) several participants have attempted tests both with synthetic clicks played from a loudspeaker and with their own mouth clicks. The authors observed that sighted participants novices to echolocation generally did better with the synthetic sounds, while the blind participants performed equally well with mouth clicks and with the sound played from speakers. The key might be a lower ability of the inexperienced echolocators to produce repeatable "efficient vocalizations", while loudspeaker-generated sounds are perfectly repeatable.

There is a lack of a clear answer as to the usefulness of noise sounds for echolocation. One of the few studies that compared different types of sounds (ARIAS, RAMOS, 1997) showed that white noise resulted in more correct echolocation answers than click sounds for a group of sighted volunteers in a test with recordings of real echoes, but not with synthetic echoes. On the other hand, in other studies (EKKEL *et al.*, 2017) white noise was a worse sound when compared to clicks, or there was no statistically significant difference between sound types (NORMAN, THALER, 2020).

None of the sound-related studies used large numbers of participants, so many conclusions may not be significant; however, the general agreement is that sounds optimal for echolocation should be relatively wide-band with at least some energy in the higher 5–10 kHz range, but with a peak frequency in a range of 1–4 kHz. This is not only because of the sensitivity of the human ears, but also due to the reflectivity of various surfaces in the environment (NORMAN, THALER, 2018). Conclusions from older studies (KISH, 2003) show that higher frequencies are the key to localizing objects that are smaller and/or further away, but are not necessary for large and nearby objects. Similar conclusions have been drawn from bat echolocation studies, showing that bats use higher frequency ultrasound for localizing small insects, while lower frequencies for large obstacles and walls (GRIFFIN, 1958, pp. xviii, 413).

Also, the familiarity of the echolocator with the sound, especially its spectral content, plays a key role, as demonstrated by NORMAN and THALER (2020). This is likely why repeatability of an echolocation signal is important, and why inexperienced echolocators may prefer artificial sounds over untrained mouth clicks, which vary significantly in spectrum (BOGUS, BUJACZ, 2021).

A final observation from other studies (THALER, CASTILLO-SERRANO, 2016) and the authors' own experiences (BUJACZ *et al.*, 2021) is that for experienced echolocators the sound source type seems to make little or no difference; however, for novice blind echolocators and sighted persons there are sounds that can give a significant improvement in echolocation accuracy, i.e., sounds with appropriately wide and predictable spectral content.

5. Blind versus sighted testers

From the 42 echolocation studies with volunteer participants reviewed in this paper, 31 were conducted with involvement of blind echolocators and 13 tested only normally sighted volunteers. Only 11 studies had more than 30 participants, while 14 had less than 10 participants. The first thing evident from the review is that the testing groups are usually very small, often too small to draw strong statistically significant conclusions, which has been noticed by previous meta reviews (TENG, WHITNEY, 2011). The usual textbook advice for parametric tests that expect probabilistic distributions of results is to collect a minimum of 30 samples (CORDER, FOREMAN, 2009). The average number of blind participants in the reviewed studies was 8 and sighted participants 19. It was even more difficult to find experiments with a group of experienced echolocators larger than 3.

Several studies compare the listening abilities of blind and sighted with mixed results. On the one

hand, the binaural localization accuracy of blind listeners has been shown to be worse with virtual sources (DOBRUCKI *et al.*, 2010), which can be attributed to the lack of audio-visual feedback training their perception. On the other hand, the visually impaired are definitely more experienced in interpreting sounds occurring naturally thus their sense of hearing is more trained, increasing the sensitivity to monaural or binaural cues (NILSSON, SCHENKMAN, 2016) as well as localization abilities in peripheral (LESSARD *et al.*, 1998) and far-space (VOSS *et al.*, 2004). In the two studies (NILSSON, SCHENKMAN, 2016; SCHENKMAN *et al.*, 2016) 23 and 12 blind testers took part in echolocation experiments, respectively and twice the number of sighted testers. The studies showed that blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences (ILDs). The authors of the study suggest that this observation may be related to the blind person's experience of localizing reflected sounds, for which ILDs may be more efficient than the inter-aural time differences (ITDs). The latter study also shows that, on average, the blind outperforms the sighted testers (noise and bursting type sounds were used). It was also noted, however, that the three best sighted echolocators performed significantly above the mean performance of all the blind participants.

Quick learning capabilities of untrained novices in echolocation were also noted in the studies reported by TENG and WHITNEY (2011). These sighted testers were able to detect size and location of objects with a surprising precision. A majority of studies (BUJACZ *et al.*, 2018; THALER, CASTILLO-SERRANO, 2016) confirm that blind echolocators perform generally better than the sighted participants, while some show a significant difference only in specific conditions, e.g., when using mouth clicks – compared to a loudspeaker (THALER, CASTILLO-SERRANO, 2016). Finally, a recent study with 17 blind testers conducted by THALER *et al.* (2020b) have showed remarkable abilities of expert echolocators, who walked in test environments as fast as sighted (and not blindfolded) participants.

The main conclusion from the reviewed studies is that the main factor in echolocation ability is not blindness or sight, but the experience with the use of echolocation, even if untrained. Research has shown that echolocation skills can be quickly learned by sighted individuals, even to a level that outperforms blind individuals (NORMAN, THALER, 2021). This observation suggests that effective echolocation training programmes can be worked out for novice echolocators (FIRR, 2019; HOLMES, 2011).

5.1. Learning to echolocate

Several of the reviewed papers focused on the process of learning to echolocate and all came to the conclu-

sion that sighted persons can acquire and demonstrate this skill just as efficiently (THALER, CASTILLO-SERRANO, 2016) or even better than the blind (EKKELE *et al.*, 2017; TENG, WHITNEY, 2011; TONELLI *et al.*, 2016), especially better than novice blind children (BUJACZ *et al.*, 2018) or blind seniors (NORMAN *et al.*, 2021). By appropriate echolocation training, both the blind and sighted people can learn to confidently detect the presence and/or location of objects of up to distances of 3–4 m and thus use echolocation for obstacle avoidance and to aid in orientation.

Several publications have been aimed at developing a curriculum for echolocation training (FIRR, 2019; KISH, HOOK, 2017; NORMAN *et al.*, 2021). Typical exercises in such training programs involve first improving awareness of echoes, as our brain intuitively ignores them. Daniel Kish has referred to this step as “unlocking”. Other preparatory exercises involve practicing general sound recognition and localization skills to improve overall hearing. Then the practice moves on to the sound source signals (usually mouth clicks) to make them as repeatable as possible and as loud as necessary.

Recently a valuable active echolocation training curriculum for people with visual impairment has been elaborated within the Erasmus+ EU programme titled: Echolocation for people with visual impairment (FIRR, 2019) in which three countries have participated, i.e., Poland, Denmark, and Lithuania. This open access (under a Creative Commons License) curriculum is dedicated to Orientation & Mobility (O&M) instructors as an educational aid for teaching active echolocation. It consists of four parts: basic theoretical information on echolocation, learning to produce tongue-click in basic exercises in using active echolocation inside buildings, active echolocation exercises in an outdoor environment, and finally the use of complex active echolocation skills, and the methods of on route problem solving.

A very recent paper on a 10-week echolocation training of 14 sighted and 12 blind participants (NORMAN *et al.*, 2021) has made some interesting observations. Throughout the course that included both live and VR exercises, the sighted participants performed better than the majority of the blind. This may be because many of the exercises and tests included virtual sounds unfamiliar to both groups and because the sighted group was overall younger.

6. Conclusions

With the ongoing research we understand the phenomenon of echolocation more and more. Myths of “facial vision” and “obstacle sense” are a thing of the past (STOCK, 2022). It is a well-documented auditory based phenomenon that both blind and sighted people can learn with practice (NORMAN *et al.*, 2021). Since most

sounds reflected from the environment fall below the delay threshold to be consciously recognized as separate auditory events, the echolocation skill must be implicitly learned through repeated use (ARIAS *et al.*, 2012). Neurological studies of blind echolocation experts show that the extremely flexible human brain will start to utilize regions previously responsible for vision to process sounds of environmental echoes (THALER *et al.*, 2011).

Testing of echolocation performance primarily consists of volunteer subjects determining the presence of nearby objects based upon emission of a source sound. In the majority of studies the subjects are stationary, the objects are disks 1 m or smaller in diameter and at distances from several centimeters up to 4 meters. The simplest tests require declaring the presence or absence of an obstacle (which for ease of procedure is usually a surface rotated to show either the flat or edge “view”), while the more complex ones also ask about the direction or distance, or have participants discriminate between different objects. The tests are best conducted in naturally reflective environments as echolocation performance in anechoic or acoustically dampened rooms is usually lower (BUJACZ *et al.*, 2022a; SCHENKMAN, NILSSON, 2010). Although the use of binaural recordings or virtual reality with spatial audio is much more efficient for conducting experiments, the echolocation effectiveness when compared to real-life trials is significantly lower. This doesn’t invalidate the results, but lower correctness rates are to be expected in research with recordings than in live experiments.

The sounds most frequently used in echolocation and echolocation-related experiments are oral palatal clicks made by the echolocators, or when using loudspeaker generated sounds either artificial clicks, percussive sounds or short noise bursts. Generally, the ideal sounds for echolocation should be familiar to the echolocator, repeatable, have a peak frequency near the human optimal hearing range (2–5 kHz), but also have a high fractional bandwidth (components in a wider spectrum around the center frequency). New research suggests, the high frequencies may produce better effects simply due to higher intensities of reflected sounds from typical surfaces used in experiments (NORMAN, THALER, 2018).

Many of the reviewed studies had a common weak point – a low number of participants. This is understandable due to difficulties in finding visually impaired volunteers, especially those experienced in echolocation. However, this can be remedied using various statistical tools, such as repeated tests for different subgroups (VAN DE SCHOOT, MIOČEVIĆ, 2020) and calculating the minimum detectable effect sizes for the utilized sample sizes (NORMAN *et al.*, 2021).

A promising conclusion is that both blind and sighted persons can efficiently learn echolocation. After comparable training courses sighted blindfolded

novices outperform inexperienced blind echolocators (NORMAN, THALER, 2021). This may be a strong argument to begin echolocation training by persons at high risk of losing eyesight, such as those with progressing cataract or glaucoma.

Acknowledgments

The presented research was financed by the Polish National Science Center grant OPUS 2019/33/B/ST7/02813. This article has been completed while the second and third author were the Doctoral Candidates in the Interdisciplinary Doctoral School at the Lodz University of Technology, Poland.

References

1. ANDRADE R., BAKER S., WAYCOTT J., VETERIE F. (2018), Echo-house: Exploring a virtual environment by using echolocation, [in:] *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, pp. 278–289, doi: 10.1145/3292147.3292163.
2. ANDRADE R., WAYCOTT J., BAKER S., VETERIE F. (2021), Echolocation as a means for people with visual impairment (PVI) to acquire spatial knowledge of virtual space, *ACM Transactions on Accessible Computing*, **14**(1): 1–25, doi: 10.1145/3448273.
3. ARIAS C., BERMEJO F., HÜG M.X., VENTURELLI N., RABINOVICH D., SKARP A.O. (2012), Echolocation: An action-perception phenomenon, *New Zealand Acoustics*, **25**(2): 20–27.
4. ARIAS C., RAMOS O.A. (1997), Psychoacoustic tests for the study of human echolocation ability, *Applied Acoustics*, **51**(4): 399–419, doi: 10.1016/S0003-682X(97)00010-8.
5. BOGUS M., BUJACZ M. (2021), Analysis of mouth click sounds used in echolocation, [in:] *2021 Signal Processing Symposium (SPSymposium)*, pp. 23–25, doi: 10.1109/SPSymposium51155.2020.9593698.
6. BUJACZ M. *et al.* (2018), EchoVis: Training echolocation using binaural recordings – Initial benchmark results, [in:] *Computers Helping People with Special Needs. ICCHP 2018. Lecture Notes in Computer Science*, Miesenberger K., Kouroupetroglou G. [Eds.], Vol. 10897, pp. 102–109, doi: 10.1007/978-3-319-94274-2_15.
7. BUJACZ M., GÓRSKI G., MATYSIK K. (2021), Mobile game development with spatially generated reverberation sound, [in:] *Advances in Systems Engineering. ICSEng 2021. Lecture Notes in Networks and Systems*, Borzowski L., Selvaraj H., Świątek J. [Eds.], Vol. 364, pp. 69–78, Springer, doi: 10.1007/978-3-030-92604-5_7.
8. BUJACZ M., KRÓLAK A., GÓRSKI G., MATYSIK K., WITEK P. (2022a), Echovis – A collection of human echolocation tests performed by blind and sighted individuals: A pilot study, *British Journal of Visual Impairment*.

9. BUJACZ M., SKULIMOWSKI P., KRÓLAK A., SZTYLER B., STRUMILLO P. (2022b), Comparison of echolocation abilities of blind and normally sighted humans using different source sounds, *Vibrations in Physical Systems*, **33**(2), doi: 10.21008/j.0860-6897.2022.2.13.
10. CASTILLO-SERRANO J.G., NORMAN L.J., FORESTEIRE D., THALER L. (2021), Increased emission intensity can compensate for the presence of noise in human click-based echolocation, *Scientific Reports*, **11**: 1750, doi: 10.1038/s41598-021-81220-9.
11. COOPER S., VELAZCO P., SCHANTZ H. (2020), Navigating in darkness: Human echolocation with comments on bat echolocation, *Journal of the Human Anatomy and Physiology Society*, **24**(2): 36–41, doi: 10.21692/haps.2020.016.
12. CORDER G.W., FOREMAN D.I. (2009), *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*, John Wiley & Sons, Inc.
13. DOBRUCKI A., PLASKOTA P., PRUCHNICKI P., PEC M., BUJACZ M., STRUMILLO P. (2010). Measurement system for personalized head-related transfer functions and its verification by virtual source localization trials with visually impaired and sighted individuals, *Journal of The Audio Engineering Society*, **58**(9): 724–738.
14. DODSWORTH C., NORMAN L.J., THALER L. (2020), Navigation and perception of spatial layout in virtual echo-acoustic space, *Cognition*, **197**: 104185, doi: 10.1016/j.cognition.2020.104185.
15. EKKEL M.R., VAN LIER R., STEENBERGEN B. (2017), Learning to echolocate in sighted people: A correlational study on attention, working memory and spatial abilities, *Experimental Brain Research*, **235**: 809–818, doi: 10.1007/s00221-016-4833-z.
16. FIEHLER K., SCHÜTZ I., MELLER T., THALER L. (2015), Neural correlates of human echolocation of path direction during walking, *Multisensory Research*, **28**(1–2): 195–226, doi: 10.1163/22134808-00002491.
17. FLANAGIN V.L. et al. (2017), Human exploration of enclosed spaces through echolocation, *Journal of Neuroscience*, **37**(6): 1614–1627, doi: 10.1523/JNEUROSCI.1566-12.2016.
18. Fundacja Instytut Rozwoju Regionalnego [FIRR] (2019), *Training Curriculum. Active Echolocation for People with Visual Impairment*.
19. GORI M., SANDINI G., MARTINOLI C., BURR D.C. (2014), Impairment of auditory spatial localization in congenitally blind human subjects, *Brain: A Journal of Neurology*, **137**: 288–293, doi: 10.1093/brain/awt311.
20. GRIFFIN D.R. (1958), *Listening in the Dark: The Acoustic Orientation of Bats and Men*, Yale University Press.
21. HELLER L.M., SCHENKER A., GROVER P., GARDNER M. (2017), Evaluating two ways to train sensitivity to echoes to improve echolocation, [in:] *The 23rd International Conference on Auditory Display (ICAD 2017)*, pp. 159–166, doi: 10.21785/icad2017.053.
22. HOLMES N. (2011), An Echolocation Training Package, *International Journal of Orientation & Mobility*, **4**(1): 84–91.
23. KISH D. (2003), *Sonic Echolocation: A Modern Review and Synthesis of the Literature*.
24. KISH D., HOOK J. (2017), *Echolocation and Flash Sonar*, American Printing House.
25. KOLARIK A.J., CIRSTEA S., PARDHAN S., MOORE B.C.J. (2014), A summary of research investigating echolocation abilities of blind and sighted humans, *Hearing Research*, **310**: 60–68, doi: 10.1016/j.heares.2014.01.010.
26. KOLARIK A.J., MOORE B.C.J., ZAHORIK P., CIRSTEA S., PARDHAN S. (2016), Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss, *Attention, Perception, & Psychophysics*, **78**(2): 373–395, doi: 10.3758/s13414-015-1015-1.
27. KOLARIK A.J., PARDHAN S., MOORE B.C.J. (2021), A framework to account for the effects of visual loss on human auditory abilities, *Psychological Review*, **128**(5): 913–935, doi: 10.1037/rev0000279.
28. KOLARIK A.J., SCARFE A.C., MOORE B.C.J., PARDHAN S. (2017), Blindness enhances auditory obstacle circumvention: Assessing echolocation, sensory substitution, and visual-based navigation, *PLOS ONE*, **12**(4): e0175750, doi: 10.1371/journal.pone.0175750.
29. KRITLY L., SLUYS Y., PELEGRÍN-GARCÍA D., GLORIEUX C., RYCHTARIKOVA M. (2021), Discrimination of 2D wall textures by passive echolocation for different reflected-to-direct level difference configurations, *PLOS ONE*, **16**(5): 10.1371/journal.pone.0251397.
30. LESSARD N., PARÉ M., LEPORE F., LASSONDE M. (1998), Early-blind human subjects localize sound sources better than sighted subjects, *Nature*, **395**: 278–280, doi: 10.1038/26228.
31. MILNE J.L., GOODALE M.A., THALER L. (2014), The role of head movements in the discrimination of 2-D shape by blind echolocation experts, *Attention, Perception, & Psychophysics*, **76**: 1828–1837, doi: 10.3758/s13414-014-0695-2.
32. NILSSON M.E., SCHENKMAN B.N. (2016), Blind people are more sensitive than sighted people to binaural sound-location cues, particularly inter-aural level differences, *Hearing Research*, **332**: 223–232, doi: 10.1016/j.heares.2015.09.012.
33. NORMAN L.J., DODSWORTH C., FORESTEIRE D., THALER L. (2021), Human click-based echolocation: Effects of blindness and age, and real-life implications in a 10-week training program, *PLOS ONE*, **16**(6): e0252330, doi: 10.1371/journal.pone.0252330.
34. NORMAN L.J., THALER L. (2018), Human echolocation for target detection is more accurate with emissions containing higher spectral frequencies, and this is explained by echo intensity, *I-Perception*, **9**(3), doi: 10.1177/2041669518776984.

35. NORMAN L.J., THALER L. (2020), Stimulus uncertainty affects perception in human echolocation: Timing, level, and spectrum, *Journal of Experimental Psychology: General*, **149**(12): 2314–2331, doi: 10.1037/xge0000775.
36. NORMAN L.J., THALER L. (2021), Perceptual constancy with a novel sensory skill, *Journal of Experimental Psychology: Human Perception and Performance*, **47**(2): 269–281, doi: 10.1037/xhp0000888.
37. ROJAS J.A.M., HERMOSILLA J.A., MONTERO R.S., ESPÍ P.L.L. (2009), Physical analysis of several organic signals for human echolocation: Oral vacuum pulses, *Acta Acustica United with Acustica*, **95**(2): 325–330, doi: 10.3813/AAA.918155.
38. ROSENBLUM L., GORDON M.S., JARQUIN L. (2000), Echolocating distance by moving and stationary listeners, *Ecological Psychology*, **12**(3): 181–206, doi: 10.1207/S15326969ECO1203_1.
39. RYCHTARIKOVA M., ZELEM L., KRITLY L., GARCIA D.P., CHMELÍK V., GLORIEUX C. (2017), Auditory recognition of surface texture with various scattering coefficients, *The Journal of the Acoustical Society of America*, **141**(5): 3452–3452, doi: 10.1121/1.4987157.
40. SCHENKMAN B.N., GIDLA V.K. (2020), Detection, thresholds of human echolocation in static situations for distance, pitch, loudness and sharpness, *Applied Acoustics*, 163: 107214, doi: 10.1016/j.apacoust.2020.107214.
41. SCHENKMAN B.N., JANSSON G. (1986), The detection and localization of objects by the blind with the aid of long-cane tapping sounds, *Human Factors*, **28**(5): 607–618.
42. SCHENKMAN B.N., NILSSON M., GRBIC N. (2016), Human echolocation: Acoustic gaze for burst trains and continuous noise, *Applied Acoustics*, **106**: 77–86, doi: 10.1016/j.apacoust.2015.12.008.
43. SCHENKMAN B.N., NILSSON M.E. (2010), Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object, *Perception*, **39**(4): 483–501, doi: 10.1068/p6473.
44. SCHENKMAN B.N., NILSSON M.E. (2011), Human echolocation: Pitch versus loudness information, *Perception*, **40**(7): 840–852, doi: 10.1068/p6898.
45. SCHÖRNICH S., NAGY A., WIEGREBE L. (2012), Discovering your inner bat: Echo-acoustic target ranging in humans, *Journal of the Association for Research in Otolaryngology: JARO*, **13**(5): 673–682, doi: 10.1007/s10162-012-0338-z.
46. SMITH G.E., BAKER C.J. (2012), Human echolocation waveform analysis, [in:] *IET International Conference on Radar Systems (Radar 2012)*, doi: 10.1049/cp.2012.1595.
47. STOCK R. (2022), Hearing echoes as an audile technique, [in:] Schillmeier M., Stock R., Ochsner B. [Eds.], *Techniques of Hearing. History, Theory and Practices*, pp. 55–65, Routledge, doi: 10.4324/9781003150763-6.
48. SUPA M., COTZIN M., DALLENBACH K.M. (1944), “Facial vision”: The perception of obstacles by the blind, *The American Journal of Psychology*, **57**(2): 133–183, doi: 10.2307/1416946.
49. TENG S., WHITNEY D. (2011), The acuity of echolocation: Spatial resolution in the sighted compared to expert performance, *Journal of Visual Impairment & Blindness*, **105**(1): 20–32.
50. THALER L. *et al.* (2017), Mouth-clicks used by blind expert human echolocators – signal description and model based signal synthesis, *PLOS Computational Biology*, **13**(8): e1005670, doi: 10.1371/journal.pcbi.1005670.
51. THALER L., ANTONIOU M., ZHANG X., KISH D. (2020a), The flexible action system: Click-based echolocation may replace certain visual functionality for adaptive walking, *Journal of Experimental Psychology: Human Perception and Performance*, **46**(1): 21–35, doi: 10.1037/xhp0000697.
52. THALER L., ARNOTT S.R., GOODALE M.A. (2011), Neural correlates of natural human echolocation in early and late blind echolocation experts, *PLOS ONE*, **6**(5): e20162, doi: 10.1371/journal.pone.0020162.
53. THALER L., CASTILLO-SERRANO J. (2016), People's ability to detect objects using click-based echolocation: A direct comparison between mouth-clicks and clicks made by a loudspeaker, *PLOS ONE*, **11**(5): e0154868, doi: 10.1371/journal.pone.0154868.
54. THALER L., DE VOS H.P.J.C., KISH D., ANTONIOU M., BAKER C.J., HORNIKX M.C.J. (2019), Human click-based echolocation of distance: Superfine acuity and dynamic clicking behaviour, *Journal of the Association for Research in Otolaryngology, JARO*, **20**(5): 499–510, doi: 10.1007/s10162-019-00728-0.
55. THALER L., DE VOS R., KISH D., ANTONIOU M., BAKER C., HORNIKX M. (2018), Human echolocators adjust loudness and number of clicks for detection of reflectors at various azimuth angles, *Proceedings of the Royal Society B: Biological Sciences*, **285**(1873): 20172735, doi: 10.1098/rspb.2017.2735.
56. THALER L., FORESTEIRE D. (2017), Visual sensory stimulation interferes with people's ability to echolocate object size, *Scientific Reports*, **7**(1): 13069, doi: 10.1038/s41598-017-12967-3.
57. THALER L., GOODALE M.A. (2016), Echolocation in humans: An overview, *WIREs Cognitive Science*, **7**(6): 382–393, doi: 10.1002/wcs.1408.
58. THALER L., ZHANG X., ANTONIOU M., KISH D.C., COWIE D. (2020b), The flexible action system: Click-based echolocation may replace certain visual functionality for adaptive walking, *Journal of Experimental Psychology: Human Perception and Performance*, **46**(1): 21–35, doi: 10.1037/xhp0000697.

59. TIRADO C., GERDFELDTER B., KÄRNEKULL S.C., NILSSON M.E. (2021), Comparing echo-detection and echo-localization in sighted individuals, *Perception*, **50**(4): 308–327, doi: 10.1177/03010066211000617.
60. TIRADO C., LUNDÉN P., NILSSON M.E. (2019), The Echobot: An automated system for stimulus presentation in studies of human echolocation, *PLOS ONE*, **14**(10): e0223327, doi: 10.1371/journal.pone.0223327.
61. TONELLI A., BRAYDA L., GORI M. (2016), Depth echolocation learnt by novice sighted people, *PLOS ONE*, **11**(6): e0156654, doi: 10.1371/journal.pone.0156654.
62. TONELLI A., CAMPUS C., BRAYDA L. (2018), How body motion influences echolocation while walking, *Scientific Reports*, **8**(1): 15704, doi: 10.1038/s41598-018-34074-7.
63. TONELLI A., CAMPUS C., GORI M. (2020), Early visual cortex response for sound in expert blind echolocators, but not in early blind non-echolocators, *Neuropsychologia*, **147**: 107617, doi: 10.1016/j.neuropsychologia.2020.107617.
64. VAN DE SCHOOT R., MIOČEVIĆ M. [Eds.] (2020), *Small Sample Size Solutions a Guide for Applied Researchers and Practitioners*, Routledge, doi: 10.4324/9780429273872.
65. VERCILLO T., MILNE J.L., GORI M., GOODALE M.A. (2014), Enhanced auditory spatial localization in blind echolocators, *Neuropsychologia*, **67**: 35–40, doi: 10.1016/j.neuropsychologia.2014.12.0.
66. VOSS P., LASSONDE M., GOUGOUX F., FORTIN M., GUILLEMOT J.-P., LEPORE F. (2004), Early- and late-onset blind individuals show supra-normal auditory abilities in far-space, *Current Biology*, **14**(19): 1734–1738, doi: 10.1016/j.cub.2004.09.051.
67. WALLMEIER L., WIEGREBE L. (2014), Ranging in human sonar: Effects of additional early reflections and exploratory head movements, *PLOS ONE*, **9**(12): e115363, doi: 10.1371/journal.pone.0115363.

Review Paper

A Study of the Romanian Framework and the Challenges in Implementing the Noise Mapping Legislation

Diana Ioana POPESCU

Technical University of Cluj-Napoca

Cluj-Napoca, Romania; e-mail: Diana.Popescu@mep.utcluj.ro

(received June 5, 2022; accepted December 20, 2022)

The Environmental Noise Directive (END), published in 2002, was transposed into Romanian local law in 2005, and it was the starting point for the first urban noise mapping exercises, initially conducted in nine Romanian cities. This paper presents the main evolutionary aspects of the noise assessment process, the development of strategic noise maps, and action plans, dealing with both the legislative side and the practical approach and results obtained. The study considers the specific regulations established by the European Commission regarding environmental noise assessment and deals with the global context at the country level, in which they have been implemented and applied.

Keywords: environmental noise; noise assessment; noise mapping; annoyance; noise exposure; action plans.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Environmental noise is becoming an increasingly acute problem worldwide, and concerns about reducing its harmful effects have intensified over the last twenty years. It is very possible that many people do not realize its effect on health. Noise is one of the hazards that threatens the health and well-being of urban areas inhabitants. Its consequences may not be immediate; however, it causes disease in the human body over time through gradual changes that affect the auditory organs and induce negative effects in the cardio-respiratory, metabolic and nervous systems (BERGLUND *et al.*, 1999; COBZEANU *et al.*, 2019; World Health Organization [WHO], 2018). The following primary annoying symptoms can be acquired due to daily living in a noisy environment: decreased attention and ability to concentrate, decreased sleep quality and a permanent state of fatigue and stress.

At the level of the European Union, the concern related to the reduction of environmental noise has intensified starting with the Fifth Environment Action Program of the European Commission from 1993, which stated: “no person should be exposed to noise levels which endanger health and quality of life” (European Environment Agency [EEA], 2014, p. 6). In 1996, the European Commission strengthened the idea of prior-

itizing noise assessment as one of the main factors of environmental pollution through the document “Future Noise Policy – EU Green Paper” (Commission of the European Communities [CEC], 1996). Subsequently, the document that marked a new perspective on the noise policy and a unified approach among Member States was Directive 2002/49/EC (2002), also known as The Environmental Noise Directive (END).

Romania became a member of the European Union on January 1, 2007. The signing of the EU Accession Treaty on April 25, 2005, was followed by a period of preparation for accession and monitoring, during which Romania was concerned with the transposition of European legislation related to noise. In 2006, the preparation of the first noise maps for the main urban agglomerations began.

The aim of this study is to describe the specific framework and review the main aspects related to how the issue of environmental noise in Romania was approached in the context of alignment with the requirements of the European family. The focus of the study is especially on urban noise mapping.

2. Legislative issues on environmental noise

At the time of initiating the process of accession to the European Union, the main law for the control of

noise pollution in Romania was The Law for Environmental Protection, no. 137/1995 (The Law No. 137, 1995), which established the general framework of the national policy on health protection of the population exposed to harmful environmental factors.

In 2005, Romania transposed into its national legislation the content of The Environmental Noise Directive (Commission Directive 2002/49/EC, 2002), by the Government Decision HG 321 (2005)⁷ on the assessment and management of environmental noise. The document defined environmental noise as: “unwanted or harmful sound from the environment, created by human activities, which includes noise emitted by means of transport, road, rail, air traffic and from locations where industrial activities are carried out” (HG 321, 2005). It introduced the noise indicators L_{den} (day-evening-night noise level) and L_{night} (night-time noise level), which could be evaluated by interim calculation methods or measurement during the noise mapping process. To avoid, prevent and reduce the harmful effects caused by the exposure of the population to environmental noise, including discomfort, the law required the implementation of action plans for areas affected by excessive noise levels. HG 321 (2005) was subsequently amended and completed by the following Government Decisions: HG 674 (2007), HG 1260 (2012), and HG 944 (2016). The law has also been supplemented by a series of Ministerial Orders, to clarify its implementation.

Based on the obligations arising from its membership in the European Community, Romania has subsequently adapted the legislation in the field, so that currently the assessment and management of ambient noise is regulated by The Law No. 121 (2019). The new law contains the provisions corresponding to the Commission Directives 2002/49/EC (2002) and 2015/996 (2015), establishing the obligation to use common noise assessment methods at the EU level. Thus, the new law transposes into national legislation the consolidated version of Commission Directive 2002/49/EC (2002) by setting out: the scope, reporting/drafting/approval obligations of strategic noise maps and action plans under the responsibility of various authorities, evaluation methods for determining the noise indicators, provisions regarding the abrogation of the Government Decision HG 321 (2005) with subsequent amendments and completions, the agglomerations to which it applies, and the contraventions. According to Law no. 121 (2019), the limit date for updating all strategic noise maps in Romania is June 30, 2022. The strategic noise maps must present the situation from the previous year for all agglomerations, roads, railways and the country’s main airports.

The Romanian government approved (on February 25, 2022) a draft normative act proposed by the Ministry of Environment, which amends and completes

Law No. 121 (2019), on the assessment and management of ambient noise. The draft law mainly contains the provisions of Commission Directive 2020/367 (2020) of March 4, 2020, amending Annex III to Directive 2002/49/EC (2002) of the European Parliament and of the Council, as regards the establishment of methods for assessing the harmful effects of environmental noise and the responsibilities of public administration authorities.

3. The environmental noise context in Romania. Practical approach

According to the END (Commission Directive 2002/49/EC, 2002, p. 14), noise mapping means: “the presentation of data on an existing or predicted noise situation in terms of a noise indicator, indicating breaches of any relevant limit value in force, the number of people affected in a certain area, or the number of dwellings exposed to certain values of a noise indicator in a certain area”. For the first noise mapping exercise, in 2007, Romania had to provide noise exposure information for urban agglomerations with more than 250 thousand inhabitants, for major roads with more than 6 million vehicles per year, major railways with more than 60 thousand trains per year and major airports with more than 50 thousand air traffic movements per year. As it did not have national calculation methods for noise indicators, the interim calculation methods recommended by the European Commission were used, in accordance with the requirements of HG 321 (2005). Noise maps were prepared for 9 cities: Bucharest, Braşov, Cluj-Napoca, Constanţa, Craiova, Galaţi, Iaşi, Ploieşti, Timişoara, for 31 sections of national roads totaling 274.4 km, for 2 main railway sections: Bucharest North – Chitila, Saligny – Palas, and for the railway stations: Ploieşti Sud, Arad, Simeria Călători, representing a total of 68 km. Regarding the noise due to air traffic, Henri Coandă International Airport in Bucharest was analyzed, which had registered 55 430 air traffic movements during 2006. Noise mapping for Bucharest Băneasa, Cluj-Napoca, Iaşi, and Craiova airports was treated as part of the respective urban agglomerations.

It can be said that this first round of noise mapping in Romania was a pioneering one, even if some previous studies related to the noise inside urban agglomerations have been made before 2006, and also there were periodic assessments and measurements made by the environmental agencies. The main challenges encountered were:

- the relatively short time for transposing European legislation and adapting national legislation;
- lack of a national methodology and strategy for assessing and predicting noise in agglomerations and its impact on the population;

- lack of digital maps for most areas to be analyzed;
- finding the material, financial, and human resources to be allocated to this action;
- the need to train people involved in the use of software for noise mapping and more.

This first round opened a profitable market for companies specialized in consulting and noise analysis in Romania and the EU. The strategic noise maps, and later the action plans for the different objectives, were made by various companies by contracting services. They used specialized software for noise prediction and sets of measurements to validate the results of the calculation models. Although their approaches have a common base, they have been pretty much different.

Strategic noise maps produced in the context of the END are generated every five years, the aim being to obtain an updated and retrospective representation of the environmental noise climate. Thus, 2012 is linked with the second round of noise mapping, whose main feature was the expansion of the analysis area and data required. The threshold for agglomerations went down from 250 thousand to 100 thousand inhabitants and the number of target cities increased from 9 to 19, as specified in (HG 1260, 2012). The main roads targeted were those with over 3 million vehicles per year, totaling 3269 km in 270 road segments. The legislation (HG 1260, 2012) also indicates the other areas to be analyzed: 51.4 km of railways, Henri Coandă International Airport in Bucharest, and 9 other urban airports identified within the targeted agglomerations.

The last modification and completion of HG 321 (2005) were made by Government Decision HG 944 (2016). It clarified the responsibilities for drawing up strategic noise maps and action plans for the railways inside the agglomerations and the deadline by which interim calculation methods can be used. The agglomerations, main roads, main railways, and major airports associated with the third round of noise mapping were specified, for which the completion dead-

lines were set: June 30, 2017 for strategic noise maps and July 18, 2018 for action plans (HG 944, 2016). This third round was characterized in Romania by stability. A sufficiently good knowledge of the working procedures and a high enough level of expertise were reached, which generated a certain continuity and confidence in the obtained results. The areas targeted for noise analysis remained approximately the same: 20 urban agglomerations, 261 sections of main roads totaling 3382.6 km, 488.2 km of railway, Henri Coandă International Airport in Bucharest, and 10 other urban airports identified within the selected agglomerations.

4. Data on noise exposure

The strategic noise maps and the action plans that followed them were available for a while on the websites of the local public administrations and the authorities responsible for their elaboration. Gradually the documents were replaced with their new variants, now with those related to the third round of noise mapping. The numerical data presented as follows were obtained from the reported data on noise exposure covered by Directive 2002/49/EC (2021), provided by the EEA. They are part of the data submitted by EEA member countries until January 1, 2021. Using the available data, the EEA has conducted comprehensive analyzes of the noise exposure situation in Europe associated with the three rounds of noise mapping (EEA, 2014; 2019).

In the data presented in Tables 1 to 3, one may observe that in Romania, most people are exposed to the noise generated by road traffic inside urban areas as it is observed in other European Union countries. The highest number of people are exposed to the two lowest noise bands mapped: 55–59 dB and 60–64 dB L_{den} , 50–55 dB and 55–59 dB L_{night} . This is a predictive result, as far as the used calculation models agree on the rate and type of noise attenuation during propagation from the source, implying the increase of the

Table 1. Population exposed to noise from different sources in Romania, 2007.

	Number of people exposed to different noise bands (L_{den})					Number of people exposed to different noise bands (L_{night})				
	55–59	60–64	65–69	70–74	>75	50–55	55–59	60–64	65–69	>70
Inside agglomerations										
Roads	943 800	1 052 300	607 300	223 900	51 700	1 021 500	664 800	285 400	71 300	13 500
Railways	153 300	100 600	17 300	1 700	200	144 200	53 000	7 900	700	200
Airports	9 400	10 000	5 600	400	400	14 300	9 500	3 200	400	300
Industry	42 200	24 000	13 000	10 300	900	28 300	8 200	4 200	9 600	100
Outside urban areas										
Major roads	22 700	15 400	21 900	5 700	1 100	17 900	13 200	8 000	2 600	300
Major railways	3 900	1 000	0	0	0	3 400	700	0	0	0
Major airports	2 400	500	100	0	0	3 900	1 300	100	100	0

Table 2. Population exposed to noise from different sources in Romania, 2012.

	Number of people exposed to different noise bands (L_{den})					Number of people exposed to different noise bands (L_{night})				
	55–59	60–64	65–69	70–74	>75	50–55	55–59	60–64	65–69	>70
Inside agglomerations										
Roads	543 800	843 400	775 800	268 700	79 400	790 200	830 100	329 000	124 900	16 700
Railways	120 400	118 100	32 900	4 000	0	119 000	98 500	21 500	2 600	0
Airports	30 000	10 000	4 000	300	0	19 000	6 000	1 300	0	0
Industry	141 600	73 100	15 200	700	0	56 800	6600	500	0	0
Outside urban areas										
Major roads	356 800	275 300	254 900	149 300	38 100	301 500	264 100	209 300	86 300	8 100
Major railways	19 000	4 500	1 300	400	0	8 100	1 700	600	100	0
Major airports	6 400	100	0	0	0	200	0	0	0	0

Table 3. Population exposed to noise from different sources in Romania, 2017.

	Number of people exposed to different noise bands (L_{den})					Number of people exposed to different noise bands (L_{night})				
	55–59	60–64	65–69	70–74	>75	50–55	55–59	60–64	65–69	>70
Inside agglomerations										
Roads	1 076 600	828 600	583 400	230 800	33 300	1 005 400	599 000	271 800	76 600	3 900
Railways	48 300	41 000	7 100	300	0	40 800	46 800	5 800	100	0
Airports	3 300	200	0	0	0	11 600	1 600	100	0	0
Industry	13 200	2 300	300	0	0	5 400	1 400	300	0	0
Outside urban areas										
Major roads	350 200	326 300	325 500	160 000	34 200	324 900	331 700	215 700	63 900	4 500
Major railways	28 700	16 000	5 200	300	0	24 700	13 300	3 300	200	0
Major airports	14 700	600	100	0	0	5 300	200	0	0	0

exposed people over the propagation distance as the noise level decreases. This intuitive result is confirmed by the numerical values presented for industrial noise in 2007, 2012, and 2017 for all noise bands. In the case of the other three categories of noise sources, there are some discrepancies: the greater number of exposed people is in the band of 60–64 dB than in 55–59 dB L_{den} in Table 1 (roads and airports noise) and Table 2 (roads noise), the greater number of exposed people is in the band of 55–59 dB than in 50–55 dB L_{night} in Table 2 (roads noise) and Table 3 (railways noise). It should also be mentioned that similar observations can also be made, for some other EU agglomerations, by studying the previously mentioned database provided by the EEA.

The results presented in Tables 1, 2, and 3 cannot be compared between years in terms of the number of people affected at the country level because the analyzed objectives were not identical, and the analysis methods were approached quite differently in the three rounds of noise mapping.

The comparative diagram in Fig. 1 presents, by percentage, the noise exposure results for the three rounds in noise bands. The calculations were made considering

the number of inhabitants in urban agglomerations for each noise mapping round. We may say that the situation has improved in recent years for exposure to noise over 60 dB. However, in the noise band of 55–59 dB L_{den} and 50–55 dB L_{night} , the percentage of the exposed population increased in 2017. Figures 2 and 3 show the percentage of reported results on the traffic noise exposure for each agglomeration. One may observe that there are quite large differences between the individual agglomerations and that some of them have missing data.

In some of the agglomerations, the fluctuating evolution of the reported data is noticeable:

- decrease in 2012 compared to 2007 and then increase in 2017 (Figs. 2 and 3) – L_{den} and L_{night} for Cluj-Napoca and Timișoara;
- increase in 2012 compared to 2007 and then decrease in 2017 (Figs. 2 and 3) – L_{den} and L_{night} for Craiova, Iași, and Ploiești.

The main reason for discrepancies is the lack of continuity and the lack of connection between the three stages of the noise mapping process conducted in Romanian urban areas (POPESCU *et al.*, 2017). Like any

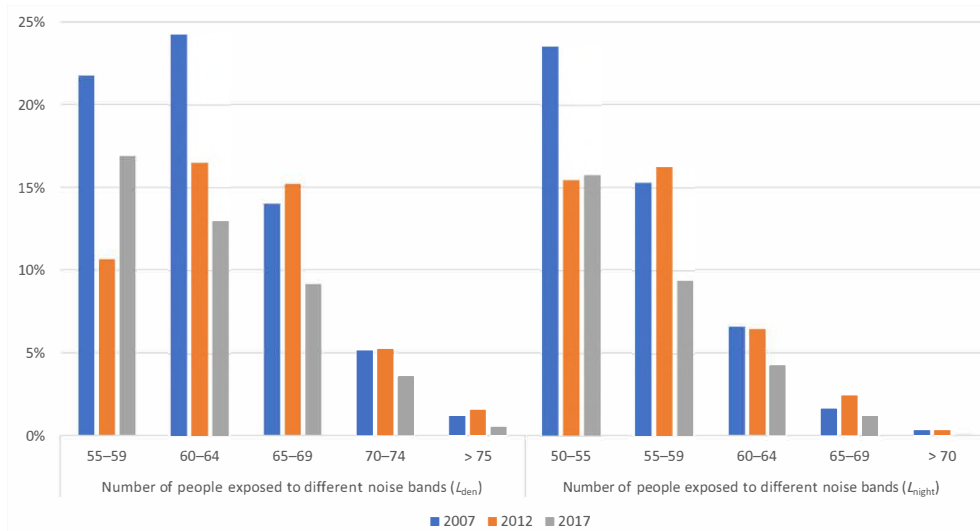


Fig. 1. Percentage of people exposed to different noise bands from roads inside agglomerations.

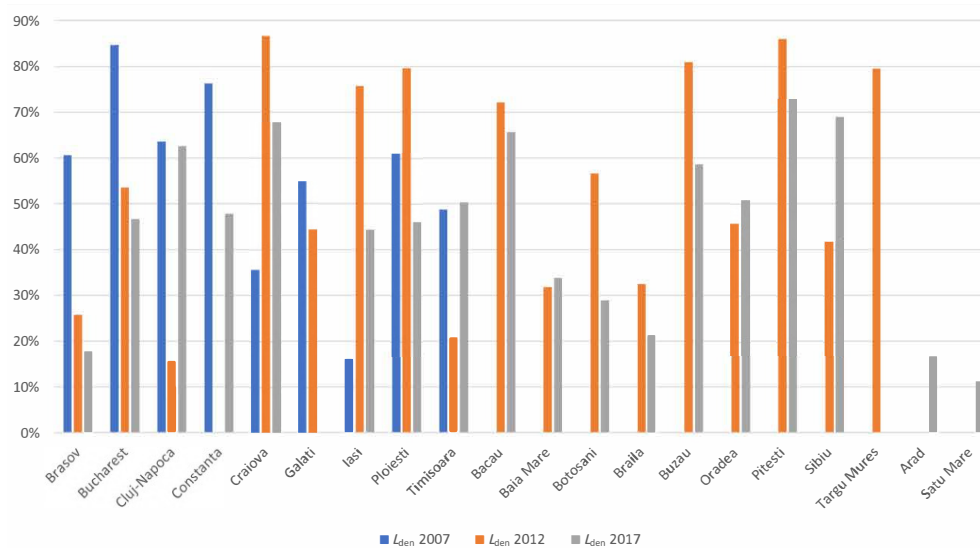


Fig. 2. Percentage of people in agglomerations exposed to road traffic noise, $L_{den} \geq 55$ dB.

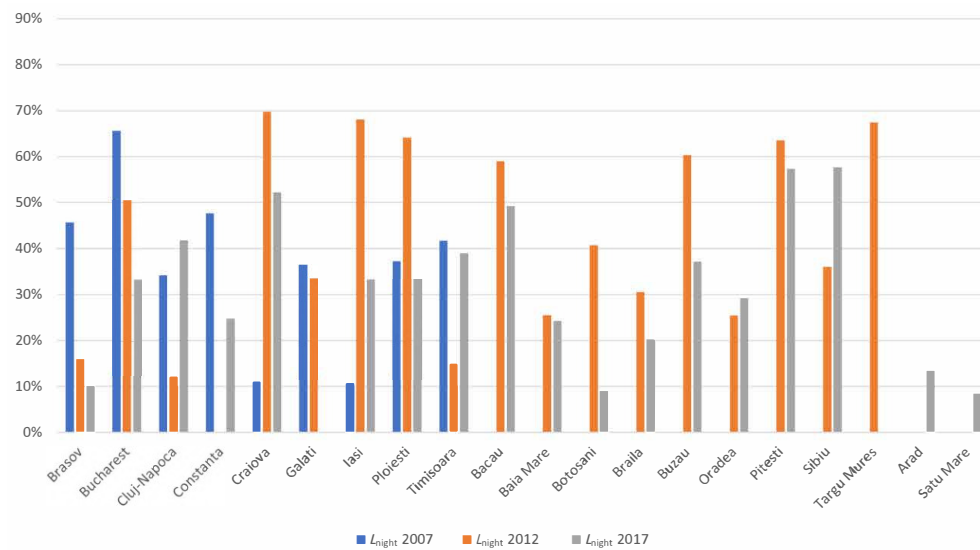


Fig. 3. Percentage of people in agglomerations exposed to road traffic noise, $L_{night} \geq 50$ dB.

analysis, those related to noise in urban agglomeration should have started from a better knowledge of the existing situation and the previously obtained results. This would have allowed the newly obtained results to be correlated with the previous ones or the correction of the previous ones, accompanied by appropriate explanations of the changes that occurred.

This lack of connection between stages generated the summation of a number of methodological factors and different approaches regarding noise mapping software, which could induce the distortion of the results (BENNETT *et al.*, 2010; PROBST, 2005), especially if the problem is the lack of experience in use or less careful coordination. Some of the aspects whose different approaches, from case to case, could influence the generated models leading to the irregular aspects mentioned above, are the following:

- the sources used for the input data, their complexity and accuracy, estimation and approximation of missing data;
- the volume of input data, which also determined the complexity of the model obtained in each case, with influence on the need for computing resources, working time and accuracy of results;
- the punctual way of interpreting some situations from reality that are not covered by the unitary work methodology;
- significant changes in actual conditions compared to the estimation made due to uncontrollable factors;
- some uncertainties introduced by the numerical methods used by different noise prediction software, model parameters, and software calibration (WIERZBICKI, BATKO, 2008).

5. Noise abatement measures included in the action plans

According to the EEA's report (EEA, 2019), various measures have been taken in each EU country to reduce and manage noise levels, but their benefits in terms of positive health outcomes are difficult to assess. Here also comes the fact that the opportunity of the proposed measures is supported mainly by theoretical arguments, and they are not formulated based on previous analyzes or research.

On the other hand, the measures proposed in the noise reduction plans must have financial support to be implemented, and the allocation of the budget is a matter that depends more on the policy area, being controlled by each local government or entity responsible for enforcing noise legislation.

For urban agglomerations in Romania, noise reduction plans included the following main categories of measures:

a) Traffic management measures aimed to streamline and calm traffic and clear road congestion, such as: intelligent traffic lights and traffic control systems that consider the variations of the flow of traffic throughout the day, the reconfiguration of routes and directions on different streets in correlation with the size of the road and traffic needs, Park & Ride systems developed on the main access roads in urban areas, clearing the streets of cars parked on the roadway and sidewalks, redesign of the road network by modifying and adapting the configuration of the cross-section of the street (if there is an available area). It is worth mentioning the need to develop an overall traffic management concept adapted to each urban agglomeration based on the initial analysis of the conditions in which road traffic can be obtained with the maximum flow, with the highest possible operating speed, in conditions of maximum safety and minimum pollution.

b) Measures aimed at reducing noise and creating quiet areas, such as: creating pedestrian areas by banning road traffic, imposing speed limitation measures on certain streets, finding alternative routes for heavy traffic, and establishing a program of access for cars that supply markets and shops or perform various urban works.

c) Development and modernization of public transport services to reduce travel by personal car between home and locations with major functions: work, school, shopping areas, and recreation areas. There are four main directions of action:

- improving the quality and comfort of public transport vehicles;
- increase the speed of the transportation service by introducing dedicated lanes and traffic lights in intersections with priority for public transport;
- reducing noise at the source by purchasing modern and quiet means of transport (e.g., electric buses);
- increasing the quality of roads by repairing damaged areas and using materials with special sound-absorbing qualities.

d) Promoting the use of non-motorized means of transport in the urban area, even if they are personal or with public access. This target involves the following actions:

- set up dedicated routes to provide traffic safety for non-motorized vehicles (e.g., bicycles);
- an appropriate adaptation of urban transport legislation;
- arrangement of parking spaces for non-motor vehicles;
- creating a fleet of vehicles available on loan and providing the service for their maintenance.

e) Noise reduction on the transmission route through measures such as sound rehabilitation of

buildings, placement of sound-absorbing panels, and curtains of vegetation.

f) Redesigning the general urban plan bearing in mind the goal of solving the problem of noise exposure of the inhabitants and following the desire to prevent the aggravation of noise pollution in the coming years. The general idea is to separate the noise sources from the residential areas, as well as to separate the local traffic from the transit ones.

g) Increasing the information and awareness of the population regarding the effects of noise exposure and sustaining education in the spirit of anti-pollution attitudes and behavior.

6. Conclusions

Among the first actions Romania took during the preparation for accession to the European Union was the transposition of European legislation regarding the assessment of environmental noise. The action of making the strategic noise maps was practically started in autumn of 2006, a few months before the finalization of the accession process. The first round of noise mapping meant a stage of learning and accommodation, and the results could be analyzed together with those from the rest of the European countries (EEA, 2014). Road traffic has been identified as the main source of environmental noise in Romania and the most significant exposure corresponds to urban areas, where pollution has increased rapidly in recent years.

From the reports prepared by the EEA as well as by analyzing the EEA databases containing information on noise exposure covered by the END, one may conclude that there were deficiencies in data reporting and that the situation of exposure to environmental noise in Romania was not significantly improved between the 2007 and 2017 noise mapping rounds. Certain fluctuations in the existing data at the level of urban agglomerations may be due to different approaches and the lack of continuity between the three stages of noise analysis. In order to resolve the observed discrepancies, it is important that the noise mapping at the level of each agglomeration includes a mandatory preliminary study of the reports and results on the previous noise assessments, which would allow the correlation of the working methodology. The newly obtained results should be presented compared to the previous ones, accompanied by explanations and interpretations of the changes that have occurred, including relation to the established action plans.

Regarding the implementation of the noise reduction action plans, it has been partially done. It is obvious that the planned measures need financial support and must be correlated with the zonal development policy, and this involves actions that may extend over a longer period of time. The unfinished activities were reconsidered in the phase of updating the noise reduc-

tion plans with the necessary modifications and adaptations.

For the fourth round of noise mapping, in 2022, the new methods provided by European legislation will be applied to assess the harmful effects induced by exposure to air, road, and rail noise. Their unitary application by EU Member States will lead to results that can be easily compared with each other. Noise pollution is expected to increase in the future (EEA, 2019) due to the tendencies of urban sprawl and mobility demand. This means increasing attention to be paid to both noise assessments and noise reduction plans.

References

1. BENNETT G.J., KING E.A., CURN J., CAHILL V., BUSTAMANTE F., RICE H.J. (2010), Environmental noise mapping using measurements in transit, [in:] *ISMA2010 – International Conference on Noise and Vibration Engineering*, Katholieke Universiteit Leuven, Belgium, pp. 1795–1809.
2. BERGLUND B., LINDVALL T., SCHWELA D.H. [Eds.] (1999), *Guidelines for Community Noise*, World Health Organization.
3. COBZEANU B., BUTNARU C., LUNGU A., POENARU M., HAINĂROȘIE R., RĂDULESCU T. (2019), Environmental noise: Health and policy – An up to date minireview, *Environmental Engineering and Management Journal*, **18**(3): 749–755.
4. Commission Directive 2002/49/EC (2002), Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise, *Official Journal of the European Communities*.
5. Commission Directive 2015/996 (2015), Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council, *Official Journal of the European Union*.
6. Commission Directive 2020/367 (2020), Commission Directive (EU) 2020/367 of 4 March 2020 amending Annex III to Directive 2002/49/EC of the European Parliament and of the Council as regards the establishment of assessment methods for harmful effects of environmental noise, <https://eur-lex.europa.eu/eli/dir/2020/367/oj>.
7. Commission of the European Communities (1996), *Future Noise Policy – European Commission Green Paper*, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:51996DC0540&from=PT> (access: 16.05.2022).
8. European Environment Agency (2019), Environmental noise in Europe – 2020, Report No 22/2019, <https://www.eea.europa.eu/publications/environmental-noise-in-europe> (access: 06.05.2022).

9. European Environment Agency [EEA] (2014), *Noise in Europe 2014*, Report No. 10/2014, <https://www.eea.europa.eu/publications/noise-in-europe-2014> (access: 06.05.2022).
10. HG 1260 (2012), Government Decision on the Modification and Completion of the HG 321/2005 on the Assessment and Management of Ambient Noise [in Romanian: Hotărârea Guvernului nr. 1260/2012 pentru modificarea și completarea Hotărârii Guvernului nr. 321/ 2005 privind evaluarea și gestionarea zgomotului ambiant], *Official Monitor of Romania*, Number 15/09.01.2013.
11. HG 321 (2005), Government Decision Related to the Assessment and Management of the Environmental Noise [in Romanian: Hotărârea Guvernului nr. 321/ 2005 privind evaluarea și gestionarea zgomotului ambiant], *Official Monitor of Romania*, Part I, Number 358/27.04.2005.
12. HG 674 (2007), Decision no. 674 of June 28, 2007 for the amendment and completion of Government Decision no. 321/2005 on the assessment and management of ambient noise [in Romanian: Hotărâre nr. 674 din 28 iunie 2007 pentru modificarea și completarea Hotărârii Guvernului nr. 321/2005 privind evaluarea și gestionarea zgomotului ambiant], *Official Monitor of Romania*, Number 485/19.07.2007.
13. HG 944 (2016), Government Decision on the Modification and Completion of the HG 321/2005 on the Assessment and Management of Ambient Noise [in Romanian: Hotărârea Guvernului nr. 944/2016 Pentru Modificarea și Completarea Hotărârii Guvernului nr. 321/ 2005 Privind Evaluarea și Gestionarea Zgomotului Ambiant], *Official Monitor of Romania*, Number 1056/28.12.2016.
14. POPESCU D.I., URSU-FISCHER N., MOHOLEA I.F. (2017), Road traffic noise in Cluj-Napoca city – Ten years after the first strategic noise map, *Acta Technica Napocensis, Series: Applied Mathematics, Mechanics and Engineering*, **60**(4): 515–520.
15. PROBST W. (2005), Uncertainties in the prediction of environmental noise and in noise mapping, *Acoustique & Technique*, **40**: 34–39.
16. Reported data on noise exposure covered by Directive 2002/49/EC (2021), *European Environment Agency*, <https://www.eea.europa.eu/data-and-maps/data/data-on-noise-exposure-8>.
17. The Law No. 121 (2019), Law on the assessment and management of ambient noise [in Romanian: Lege nr. 121/2019 privind evaluarea și gestionarea zgomotului ambiant], *Official Monitor of Romania*, Number 604/23.07.2019.
18. The Law No. 137 (1995), Law for Environmental Protection [in Romanian: Lege nr. 137 din 29 decembrie 1995 privind protecția mediului], *Official Monitor of Romania*, Part I, Number 304/30.12.1995.
19. WIERZBICKI J., BATKO W. (2008), Uncertainty of noise mapping software, [in:] *Acoustics 08 Paris*, <http://www.conforg.fr/acoustics2008/cdrom/data/articles/002379.pdf> (access: 06.11.2022).
20. World Health Organization [WHO], *Environmental Noise Guidelines for European Region* (2018), <https://www.euro.who.int/en/health-topics/environment-and-health/noise/publications/2018/environmental-noise-guidelines-for-the-european-region-2018> (access: 15.04.2022).

Technical Note

Corrected Long-Term Time-Average Sound Level of Amplitude-Modulated Wind Turbine Noise

Rufin MAKAREWICZ, Maciej BUSZKIEWICZ*

*Faculty of Physics, Adam Mickiewicz University
Poznań, Poland*

*Corresponding Author e-mail: maciej.buszkiewicz@amu.edu.pl

(received August 20, 2022; accepted November 15, 2022)

Amplitude-modulated noise from a single wind turbine is considered. The time-varying modulation depth D_m and the short time-average sound level $L_{Aeq,\tau}$ (with $\tau = 20$ s) are measured at the reference distance d_* . Due to amplitude modulation, a penalty has to be added to $L_{Aeq,\tau}$. The paper shows how to calculate the corrected long-term time-average sound level $\hat{L}_{Aeq,T}$ (with $T \gg 20$ s), which accounts for amplitude modulation at any distance $d \neq d_*$ from the wind turbine. The proposed methodology needs to be tested by research.

Keywords: wind turbine noise; amplitude modulation; annoyance.



Copyright © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>) which permits use, distribution, and reproduction in any medium, provided that the article is properly cited. In any case of remix, adapt, or build upon the material, the modified material must be licensed under identical terms.

1. Introduction

Noise generation and the propagation of noise from a single wind turbine are analyzed. Under certain conditions, the A-weighted sound pressure level L_{pA} of wind turbine noise (WTN) is modulated with the blade

passing frequency f_m . Figure 1 shows pulses of duration $\tau \approx 1/f_m$ and the time-varying modulation depth $-D_m$. In two studies (VON HÜNERBEIN, PIPER, 2015; RenewableUK, 2013), D_m is defined as the difference between the mean peak and the mean trough in the A-weighted RMS time series for any consecutive group

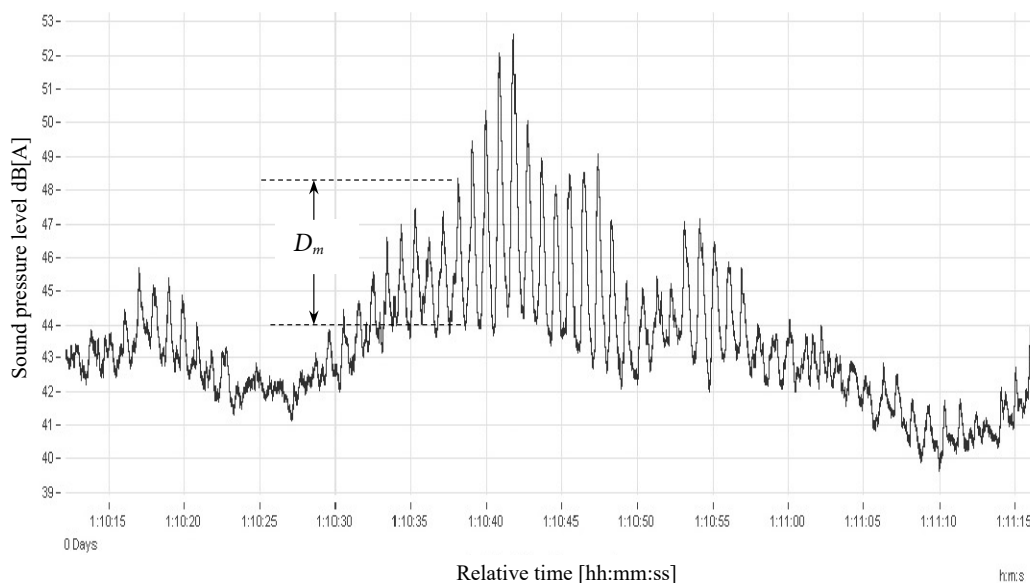


Fig. 1. Amplitude modulation of WTN recorded in a far field location (Di Napoli, 2011).

of 12 pulses that occur during the $\tau = 20$ s time interval.

Suppose every WTN segment of $\tau = 20$ s duration is characterized by the short term time-average sound level $L_{\text{Aeq},\tau}$ and modulation depth D_m . Recent studies prove that amplitude modulation increases WTN annoyance (ALAMIR *et al.*, 2021; LOTINGA, 2021). Therefore, a penalty k [dB] is to be added to the measured or calculated value of the short-term time-average sound level (MAKAREWICZ, 2022):

$$\widehat{L}_{\text{Aeq},\tau} = L_{\text{Aeq},\tau} + k \text{ [dB]}. \quad (1)$$

Figure 2 shows the old penalty scheme published in (RenewableUK, 2013). For a small amplitude depth, $0 < D_m < 3$ dB, there is no penalty at all, $k = 0$. When $3 \text{ dB} < D_m < 10$ dB, the penalty k increases linearly from 3 dB to 6 dB. Finally, for D_m exceeding 10 dB, the penalty equals 6 dB. Table 1 contains discrete values of D_m and k . An alternative penalty scheme is proposed by VIRJONEN *et al.* (2019). Problem of WTN modulation was also discussed in numerous articles (BASS *et al.*, 2016; BOWDLER *et al.*, 2018; HANSEN *et al.*, 2018).

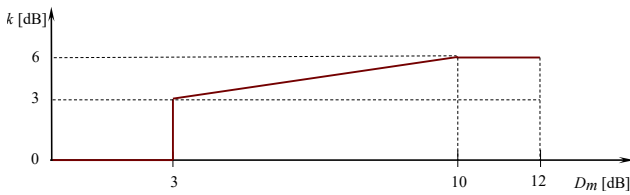


Fig. 2. The penalty scheme published in (RenewableUK, 2013).

Table 1. Discrete values of modulation depth D_m and corresponding penalty k .

D_m [dB]	0	1	2	3	4	5	6	7	8	9	10	11	12
k [dB]	0	0	0	3	$3\frac{3}{7}$	$3\frac{6}{7}$	$4\frac{2}{7}$	$4\frac{5}{7}$	$5\frac{1}{7}$	$5\frac{4}{7}$	6	6	6

2. Noise measurements

When T denotes the time duration of WTN, $N = T/\tau$ measurements at the distance d_* of the short-term A-weighted time-average sound levels $L_{\text{Aeq},\tau}^{(i)}$ and modulation depths $D_m^{(i)}$ (Fig. 3) provide complete information on WTN. It is assumed that $n \ll N$ values of $L_{\text{Aeq},\tau}^{(i)}$, $D_m^{(i)}$ yield not complete but “good enough” information. The methodology of $L_{\text{Aeq},\tau}$ and D_m measurements one finds in (HANSEN *et al.*, 2017; RenewableUK, 2013; IEC 61400-11, 2012). Figure 4 shows the turbine-receiver arrangement.

Figures 5 and 6 present the results of measurements with the upper and lower limits of the short-term A-weighted time-average sound levels

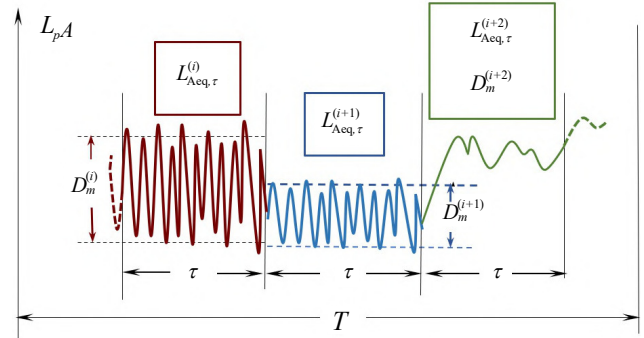


Fig. 3. Time variation of A-weighted sound level L_{pA} characterized by series of pairs, $L_{\text{Aeq},\tau}^{(i)}$, $D_m^{(i)}$.

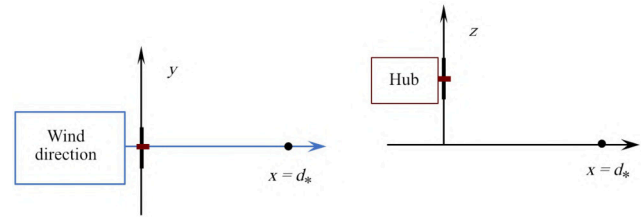


Fig. 4. Measurements of $L_{\text{Aeq},\tau}^{(i)}$ and $D_m^{(i)}$ at the reference distance d_* .

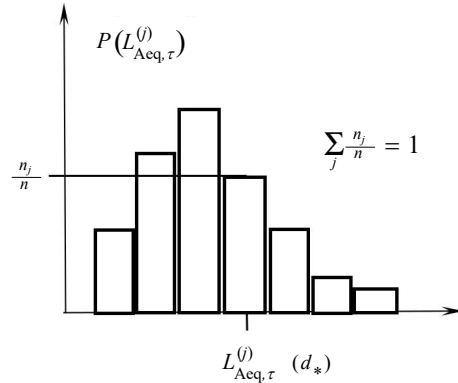


Fig. 5. Histogram of short-term A-weighted time-average sound levels at the reference distance d_* .

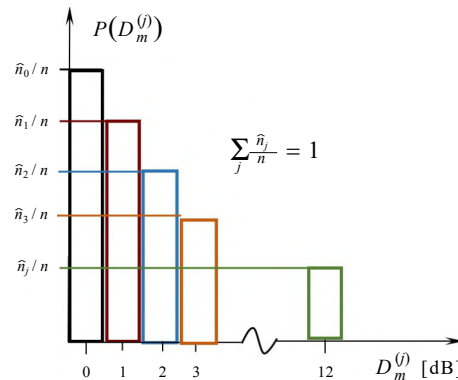


Fig. 6. Histogram of modulation depths. The probability $P = \hat{n}_0/n$, obtained for $D_m = 0$, corresponds to non-modulated intervals τ (Fig. 3).

(e.g., $L_{\text{Aeq},\tau}^{(j)} - 2.5$ dB; $L_{\text{Aeq},\tau}^{(j)} + 2.5$ dB) and modulation depths (e.g., $D_m^{(j)} - 0.5$ dB; $D_m^{(j)} + 0.5$ dB).

3. Theory

The long-term A-weighted time-average sound level is defined as (American National Standards Institute [ANSI], 2020; ISO 1996-1, 2016):

$$L_{\text{Aeq},T} = 10 \log \frac{\langle p_A^2 \rangle_T}{p_o^2}, \quad p_o = 20 \text{ } \mu\text{Pa}, \quad (2)$$

where (Fig. 5)

$$\frac{\langle p_A^2 \rangle_T}{p_o^2} = \frac{1}{T} \int_0^T \frac{p_A^2}{p_o^2} dt = \sum_j P(L_{\text{Aeq},\tau}^{(j)}) \cdot 10^{0.1L_{\text{Aeq},\tau}^{(j)}} \quad (3)$$

represents the long-term A-weighted time-average squared sound pressure for the interval $T \gg \tau$ (e.g., one or a few hours). The two above expressions yield $L_{\text{Aeq},T}$ at the reference distance d_* , where the measurements $L_{\text{Aeq},\tau}^{(i)}$ have been performed (Fig. 5). How can these results be used for $L_{\text{Aeq},T}$ prediction at any distance, $d \neq d_*$? When the distance d between the turbine and the receiver exceeds the double rotor disc diameter, the turbine blades could be replaced by a point source at the hub height (MAKAREWICZ, 2011; ECOTIÈRE, 2015). The momentary A-weighted squared sound pressure can be written as (FORSSÉN *et al.*, 2010; Australian Standard AS 4959, 2010; PLOVSING, SØNDERGAARD, 2011):

$$\frac{p_A^2(t)}{p_o^2} = \frac{W_A(t)}{W_o} \cdot F(d), \quad W_o = 10^{-12} \text{ [W]}, \quad (4)$$

where $W_A(t)$ and $F(d)$ represent the A-weighted sound power and propagation function, respectively. When noise measurements are carried out with a microphone on the perfectly reflecting board and noise propagation is governed by geometrical spreading and air absorption, then (IEC 61400-11, 2012):

$$F(d) = 4 \cdot \frac{d_o^2}{4\pi d^2} \exp\left(-\alpha \frac{d}{d_o}\right), \quad d_o = 1 \text{ m}. \quad (5)$$

Symbol α [1/m] denotes the representative air absorption coefficient for the noise spectrum, which equals $\alpha \approx 1.15 \cdot 10^{-3}$ [1/m] (Swedish Environmental Protection Agency [SEPA], 2001). More complicated propagation functions $F(d, \dots)$ are discussed by HANSEN *et al.* (2017).

For the reference distances d_* and the distance $d \neq d_*$, the short-term A-weighted time average squared sound pressure takes the form (Eq. (4)):

$$\frac{\langle p_A^2 \rangle_{d_*,\tau}}{p_o^2} = \frac{\langle W_A \rangle_\tau}{W_o} \cdot F(d_*), \quad \frac{\langle p_A^2 \rangle_\tau}{p_o^2} = \frac{\langle W_A \rangle_\tau}{W_o} \cdot F(d), \quad (6)$$

where $\langle W_A \rangle_\tau$ represents the short-term A-weighted time-average sound power. Consequently, Eq. (3) yields

the long-term A-weighted time-average squared sound pressure at the distance $d \neq d_*$,

$$\frac{\langle p_A^2 \rangle_T}{p_o^2} = \frac{F(d)}{F(d_*)} \cdot \sum_j P(L_{\text{Aeq},\tau}^{(j)}) \cdot 10^{0.1L_{\text{Aeq},\tau}^{(j)}}, \quad (7)$$

and Eqs. (2), (4), (7) combine into the long-term A-weighted time-average sound level:

$$L_{\text{Aeq},T}(d) = 10 \log J_* + 10 \log \frac{F(d)}{F(d_*)}, \quad d \neq d_*. \quad (8)$$

Here (Fig. 5)

$$J_* = \sum_j P(L_{\text{Aeq},\tau}^{(j)}) \cdot 10^{0.1L_{\text{Aeq},\tau}^{(j)}(d_*)}. \quad (9)$$

For the simplest case (Eqs. (5) and (8)):

$$L_{\text{Aeq},T}(d) = 10 \log J_* - 20 \log \frac{d}{d_*} - 4.34\alpha \cdot (d - d_*). \quad (10)$$

Unfortunately, the value of J_* (Eq. (9)) does not take into account the annoyance increase due to amplitude modulation (ALAMIR *et al.*, 2021; LOTINGA, 2021). So, let us replace $L_{\text{Aeq},\tau}^{(j)}$ with $L_{\text{Aeq},\tau}^{(j)} + k_j$ (Eq. (1)) and write Eq. (9) again in a new form:

$$\widehat{J}_* = \sum_j P(L_{\text{Aeq},\tau}^{(j)}, k_j) \cdot 10^{0.1[L_{\text{Aeq},\tau}^{(j)} + k_j]}. \quad (11)$$

Symbol $P(L_{\text{Aeq},\tau}^{(j)}, k_j)$ in Eq. (11) represents the unknown joint probability of co-occurrence of the short-term time-average sound level $L_{\text{Aeq},\tau}^{(j)}$ and penalty k_j , which depends on modulation depth $D_m^{(j)}$ (Table 1, Fig. 2). It is known that $L_{\text{Aeq},\tau}^{(j)}$ increases with wind speed (HANSEN *et al.*, 2017). On the other hand, D_j was found to be not related to wind speed (PAULRAJ, VÄLISUO, 2017). Accordingly, penalty k_j does not depend on wind speed either. Because the joint probability of independent events is calculated as the probability of event A multiplied by the probability of event B, Eq. (11) can be written as follows:

$$\widehat{J}_* = \sum_i P(L_{\text{Aeq},\tau}^{(i)}) \cdot 10^{0.1L_{\text{Aeq},\tau}^{(i)}(d_*)} \cdot \sum_j P(k_j) \cdot 10^{0.1k_j}. \quad (12)$$

Then, in view of the identities (ANS, 2020; ISO 1996-1, 2016):

$$\begin{aligned} \sum_i P(L_{\text{Aeq},\tau}^{(i)}) \cdot 10^{0.1L_{\text{Aeq},\tau}^{(i)}} &= \frac{1}{n} \sum_{i=1}^n 10^{0.1L_{\text{Aeq},\tau}^{(i)}} \\ &= 10^{0.1L_{\text{Aeq},T}}, \end{aligned} \quad (13)$$

the combination of Eqs. (8) and (11) results in the corrected long-term A-weighted time-average sound level:

$$\widehat{L}_{\text{Aeq},T} = L_{\text{Aeq},T}(d) + \Delta \widehat{L}, \quad (14)$$

where $L_{\text{Aeq},T}(d)$ represents the calculated long-term A-weighted time-average sound level at the distance $d \neq d_*$ (Eqs. (8) and (9)), and

$$\Delta\widehat{L} = 10 \log \left\{ \sum_j \frac{n_j}{n} 10^{0.1k_j} \right\} \quad (15)$$

represents the correction due to amplitude modulation (Fig. 6). Table 2 contains the measured values $D_m^{(j)}$, corresponding penalties k_j (Table 1) and measured probabilities $P(k_j)$ (Fig. 6).

Table 2. Modulation depth $D_m^{(j)}$ with the corresponding penalty k_j and probability $P(k_j)$ of its occurrence.

$D_m^{(j)}$ [dB]	0	1	2	3	4	5	6	7	8	9	10	11	12
k_j	0	0	0	3	$3\frac{3}{7}$	$3\frac{6}{7}$	$4\frac{2}{7}$	$4\frac{5}{7}$	$5\frac{1}{7}$	$5\frac{4}{7}$	6	6	6
$P(k_j)$	$\frac{n_0}{n}$	$\frac{n_1}{n}$	$\frac{n_2}{n}$	$\frac{n_3}{n}$	$\frac{n_4}{n}$	$\frac{n_5}{n}$	$\frac{n_6}{n}$	$\frac{n_7}{n}$	$\frac{n_8}{n}$	$\frac{n_9}{n}$	$\frac{n_{10}}{n}$	$\frac{n_{11}}{n}$	$\frac{n_{12}}{n}$

For example, let us determine the corrected one-hour A-weighted time-average sound level, $\widehat{L}_{\text{Aeq},1h}$ (Eq. (13)):

$$\widehat{L}_{\text{Aeq},1h} = L_{\text{Aeq},1h} + \Delta\widehat{L}. \quad (16)$$

Making use of $n = 180$ values of the measured short-time A-weighted average sound level, $L_{\text{Aeq},20s}^{(j)}$, and $n = 180$ values of the modulation depth $D_m^{(j)}$ (Table 3), Eq. (15) gives $\Delta\widehat{L} = 2.5$ dB.

Table 3. Modulation depth $D_m^{(j)}$ with the corresponding penalty k_j and probability $P(k_j)$ of its occurrence for 180 values.

$D_m^{(j)}$ [dB]	0	1	2	3	4	5	6	7	8	9	10
k_j	0	0	0	3	$3\frac{3}{7}$	$3\frac{6}{7}$	$4\frac{2}{7}$	$4\frac{5}{7}$	$5\frac{1}{7}$	$5\frac{4}{7}$	6
$P(k_j)$	$\frac{35}{180}$	$\frac{28}{180}$	$\frac{25}{180}$	$\frac{22}{180}$	$\frac{19}{180}$	$\frac{16}{180}$	$\frac{13}{180}$	$\frac{10}{180}$	$\frac{7}{180}$	$\frac{4}{180}$	$\frac{1}{180}$

When $L_{\text{Aeq},\tau}^{(j)}(d_*)$ (Fig. 5) is measured at the distance d_* with the microphone on the perfectly reflecting board, then in the simplest case of propagation (Eqs. (9) and (10)) the corrected long-term A-weighted time-average squared sound pressure at $d \neq d_*$ is determined by:

$$\begin{aligned} \widehat{L}_{\text{Aeq},T}(d) = 10 \log \sum_j \frac{n_j}{n} 10^{0.1L_{\text{Aeq},\tau}^{(j)}(d_*)} \\ - 20 \log \frac{d}{d_*} - 4.34\alpha \cdot (d - d_*) + \Delta\widehat{L}, \end{aligned} \quad (17)$$

where $\alpha \approx 1.15 \cdot 10^{-3}$ [1/m] (SEPA, 2001).

4. Conclusions

The annoyance caused by WTN increases with modulation depth (ALAMIR *et al.*, 2021; LOTINGA, 2021).

Input data of the methodology proposed here consists of the short-term time-average sound level $L_{\text{Aeq},\tau}$ (with $\tau = 20$ s) and D_m (Fig. 3) measured at the reference distance d_* . To find the corrected long-term time-average sound level $\widehat{L}_{\text{Aeq},T}$ (with $T \gg \tau$), which takes into account annoyance increases due to modulation, one can use Eqs. (8), (9) or (14). When propagation depends mainly on geometrical spreading and air absorption, Eq. (17) is recommended. The proposed methodology needs to be tested by research.

References

1. ALAMIR M.A., HANSEN K.L., CATCHESIDE P. (2021), Penalties applied to wind farm noise: current allowable limits, influencing factors, and their development, *Journal of Cleaner Production*, **295**: 126393, doi: 10.1016/j.jclepro.2021.126393.
2. American National Standards Institute (2020), *American National Standard Acoustical Terminology*, S1.1-2013, Acoustical Society of America.
3. Australian Standard AS 4959 (2010), *Acoustics – Measurements, Prediction and Assessment of Noise from Wind Turbine Generators*, Standards Australia.
4. BASS J. *et al.* (2016), *Method for Rating Amplitude Modulation in Wind Turbine Noise*, Institute of Acoustics, IOA Working Group, UK.
5. BOWDLER D.B., CAND M., LEA H., HAYES M., IRVINE G. (2018), Wind turbine noise amplitude penalty considerations, [in:] *Proceedings of the Institute of Acoustics*, **40**: 253–261.
6. DI NAPOLI C. (2011), Long distance amplitude modulation of wind turbine noise, [in:] *Fourth International Meeting on Wind Turbine Noise*, Rome, Italy.
7. ECOTIÈRE D. (2015), Can we really predict wind turbine noise with only one point source?, [in:] *6th International Meeting on Wind Turbine Noise*, Glasgow.
8. FORSSÉN F., SCHIFF M., PEDERSEN E., WAYE K.P. (2010), Wind turbine noise propagation over flat ground: measurements and predictions, *Acta Acustica united with Acustica*, **96**(4): 753–760, doi: 10.3813/AAA.918329.
9. HANSEN C.H., DOOLAN C.J., HANSEN K.L. (2017), *Wind Farm Noise: Measurements, Assessment and Control*, John Wiley & Sons Ltd.
10. HANSEN K.L., ZAJAMSEK B., HANSEN C.H. (2018), Towards a reasonable penalty for amplitude modulated wind turbine noise, *Acoustics Australia*, **46**(1): 21–25, doi: 10.1007/s40857-018-0132-0.
11. VON HÜNERBEIN S., PIPER B. (2015), Affective response to amplitude modulated wind turbine noise, *Proceedings of Wind Turbine Noise 2015*, <https://usir.salford.ac.uk/id/eprint/35663>.

12. IEC 61400-11 (2012), *Wind Turbines – Part 11: Acoustic Noise Measurement Techniques*, International Standard.
13. ISO 1996-1 (2016), *Acoustics – Description, Measurement and Assessment of Environmental Noise – Part 1: Basic Quantities and Assessment Procedures*, ISO Standards.
14. LOTINGA M.J.B. (2021), Subjective responses to wind turbine noise amplitude modulation: Pooled analysis of laboratory listening studies and synthesis of an AM character rating penalty, [in:] *9th International Conference on Wind Turbine Noise, Remote from Europe*.
15. MAKAREWICZ R. (2011), Is a wind turbine a point source?, *The Journal of Acoustical Society of America*, **129**(2): 579–581, doi: 10.1121/1.3514426.
16. MAKAREWICZ R. (2022), Proposed methodology for the annoyance penalty of amplitude modulated wind turbine noise, *Archives of Acoustics*, **47**(1): 125–128, 2022, doi: 10.24425/aoa.2022.140738.
17. PAULRAJ T., VÄLISUO P. (2017), Effect of wind speed and wind direction on amplitude modulation of wind turbine noise, [in:] *Inter-Noise and Noise-Con Congress and Conference Proceedings*, **255**(2): 5479–5489, Hong Kong.
18. PLOVSING B., SØNDERGAARD B. (2011), Wind turbine noise propagation: Comparison of measurements and predictions by a method based on geometrical ray theory, *Noise Control Engineering Journal*, **59**(1): 10–22, doi: 10.3397/1.3548871.
19. RenewableUK (2013), *Template Planning Condition on Amplitude Modulation*, Noise Guidance Notes.
20. Swedish Environmental Protection Agency (2001), *Noise from Wind Turbines, Report 6241*, Stockholm.
21. VIRJONEN P., HONGISTO V., RADUN J. (2019), Annoyance penalty of periodically amplitude-modulated wide-band sound, *The Journal of Acoustical Society of America*, **146**(6): 4159–4170, doi: 10.1121/1.5133478.

AIMS AND SCOPE

Archives of Acoustics was founded in 1976 and since then publishes original research papers from all areas of acoustics including:

- acoustical measurements and instrumentation,
- musical acoustics,
- acousto-optics,
- architectural, building, and environmental acoustics,
- bioacoustics,
- electroacoustics,
- linear and nonlinear acoustics,
- noise and vibration,
- psychoacoustics,
- physical and chemical effects of sound,
- physiological acoustics,
- speech processing and communication systems,
- speech production and perception,
- sonochemistry,
- transducers,
- ultrasonics,
- underwater acoustics,
- quantum acoustics.

INFORMATION FOR AUTHORS

- Manuscripts intended for publication in *Archives of Acoustics* should be submitted in pdf format by an on-line procedure using **New Submission** link in **User Home** menu.
- Manuscript should be original and should not be submitted either previously or simultaneously elsewhere, neither in whole, nor in part.
- Submitted papers must be written in good English and proofread by a native speaker.
- In principle the papers should not exceed 40 000 typographic signs (about 20 typewritten pages).
- A brief abstract must be included. The keywords and subject classification for the article are required.
- Postal addresses, affiliations and email addresses for each author are required.
- The submitted manuscript should be accompanied by a cover letter containing the following information:
 - why the paper is submitted to *Archives of Acoustics*,
 - suggestion on the field of acoustics related to the topic of the submitted paper,
 - the statement that the manuscript is original, the submission has not been previously published, nor was it sent to another journal for consideration,
 - 3–5 names of suggested reviewers together with their affiliations, full postal and e-mail addresses; at least 3 suggested reviewers should be affiliated with other scientific institutions than the affiliations of the authors,
 - author's suggestion to classification of the paper as the research paper, review paper or technical note.
- Information regarding the preparation of manuscripts is provided on the journal's website <https://acoustics.ippt.pan.pl>.
- Reviews of manuscript is available online.
- Full information about the journal are available at <https://acoustics.ippt.pan.pl>.

Each manuscript is initially evaluated by the Editor and then allocated to an Section Editor according to specific subject area. Upon receiving the Section Editor sends it to reviewers. The Editor and the Section Editors make the final recommendation. Their opinions are crucial for the Editorial Committee to decide whether the paper can be accepted for publication or not.

The Word or L^AT_EX source and pdf files of the final version of articles will be requested after acceptance for publication. High-quality illustrations (format bitmap (gif, jpg) or PostScript, with resolution no lower than 300 dpi) should be sent along with the manuscript.

Editorial Board
Archives of Acoustics