

eISSN 2300-262X (online)



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

# ARCHIVES of ACOUSTICS

QUARTERLY

Vol. 49, No. 3, 2024

WARSAW





# ARCHIVES of ACOUSTICS

QUARTERLY, Vol. 49, No. 3, 2024

## Research Papers

W. Batko, L. Radziszewski, A. Bąkowski, <i>A scalar measure of acoustic hazard assessment</i> .....	299
E. Zhang, Y. Chen, L. Su, R. Zhonglian, X. Chen, S. Jiang, <i>Evaluation modeling of electric bus interior sound quality based on two improved XGBoost algorithms using GS and PSO</i> .....	307
D. Młynarczyk, J. Wiciak, <i>Virtual reality technology in analysis of the Sarek National Park soundscape in Sweden</i>	319
P. Małecki, J. Stefańska, M. Szydłowska, <i>Assessing spatial audio: A listener-centric case study on object-based and ambisonic audio processing</i> .....	331
K. Kosala, <i>Modelling the acoustic properties of baffles made of porous and fibrous materials</i> .....	345
T. Vilniškis, T. Januševičius, <i>Sound insulation properties of sound-reduction louvers with innovative devulcanized rubber</i> .....	359
A. Klimek, J.F. Łątka, P. Nieradka, A. Dobrucki, <i>Application of cellulose and paper-based products in building acoustics</i> .....	369
A. Rojczyk, A. Porzuczek, <i>How much voicing in voiced geminates? The laryngeal voicing profile of Polish double stops</i> .....	385
Y. Wang, X. Deng, B. Chi, Y. Cui, <i>Study on the impact of drainage noise in residential bathrooms based on finite element simulation</i> .....	399
W.C. Hong, Y.Y. Kim, C.D. Kwon, K.C. So, <i>Effects of ultrasonic power and intensity of mechanical agitation on pretreatment of a gold-bearing arsenopyrite</i> .....	419
I. Trots, J. Tasinkiewicz, A. Nowicki, <i>Mutually orthogonal complementary Golay coded sequences: An in-vivo study</i>	429
A. Salaeh, <i>Influence of ultrasonic cavitation on Botryococcus Braunii growth</i> .....	439
L. Morzyński, M. Podleśna, G. Szczepański, A. Włodarczyk, <i>Ultrasonic haptic devices: Ultrasonic noise assessment</i> .....	447
U. Libal, P. Biernacki, <i>Drone flight detection at an entrance to a beehive based on audio signals</i> .....	459

## 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)

GAMBIN Barbara (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 Krakow, Poland)

Architectural acoustics

KAMISIŃSKI Tadeusz (AGH University of Krakow, Poland)

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

Musical acoustics and psychological acoustics

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

PREIS Anna (Adam Mickiewicz University, Poland)

Underwater acoustics and nonlinear acoustics

MARSZAŁ Jacek (Gdańsk University of Technology, Poland)

Speech, computational acoustics, and signal processing

DRGAS Szymon (Poznan University of Technology)

KOCIŃSKI Jędrzej (Adam Mickiewicz University, Poland)

Ultrasonics, transducers, and instrumentation

GAMBIN Barbara (Institute of Fundamental Technological Research PAS, Poland)

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)

KOMPALA Janusz (Central Mining Institute, Poland)

LENIOWSKA Lucyna (University of Rzeszów, Poland)

PIECHOWICZ Janusz (AGH University of Krakow, Poland)

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

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

### Advisory Editorial Board

**Chairman:** WESOŁOWSKI Zbigniew (Polish Academy of Sciences, Poland)

KOZACZKA Eugeniusz (Polish Academy of Sciences, Poland)

BATKO Wojciech (AGH University of Krakow, Poland)

BLAUERT Jens (Ruhr University, Germany)

BRADLEY David (The Pennsylvania State University, USA)

CROCKER Malcolm J. (Auburn University, USA)

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

HANSEN Colin (University of Adelaide, Australia)

HESS Wolfgang (University of Bonn, Germany)

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 PAS

### 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 in BazTech, Science Citation Index-Expanded (Web of Science Core Collection),

ICI Journal Master List, Scopus, PBN – Polska Bibliografia Naukowa,

Directory of Open Access Journals (DOAJ)

Recognised by The International Institute of Acoustics and Vibration (IIAV)

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

PUBLISHED IN POLAND

Typesetting in L<sup>A</sup>T<sub>E</sub>X: JEZIERSKA Katarzyna (Institute of Fundamental Technological Research PAS, Poland)

## Research Paper

## A Scalar Measure of Acoustic Hazard Assessment

Wojciech BATKO<sup>(1)</sup>, Leszek RADZISZEWSKI<sup>(2)</sup>, Andrzej BAKOWSKI<sup>(2)\*</sup><sup>(1)</sup> State University of Applied Sciences in Krosno  
Krosno, Poland<sup>(2)</sup> Kielce University of Technology  
Kielce, Poland\*Corresponding Author e-mail: [abakowski@tu.kielce.pl](mailto:abakowski@tu.kielce.pl)

(received January 12, 2024; accepted February 26, 2024; published online April 22, 2024)

The article addresses the problem of assessing the impact of road modernization on improving the acoustic environment. It formulates a hypothesis about the advisability of adopting the scalar dimension of the decibel space to describe acoustic hazards. It is proposed to reduce the analysis of changes in sound levels to the analysis of changes in the coefficient of exceedance of the recommended noise levels. Its value is determined by the decibel relation of dividing sound levels. The basis for the assessment of the effectiveness of the adopted solution was the analysis of the statistical characteristics of monitored exceedances of recommended noise levels, considered through the prism of the current and proposed measure. They showed greater sensitivity of the proposed coefficient in the assessment of the improvement of the acoustic climate caused by road modernization. For example, the median noise level for nights before the road modernization was 66.9 dB(A), and after the modernization 65.6 dB(A). However, the coefficient of exceedances decreased by approximately 25 %. Numerical simulations, in accordance with the Cnossos noise model, showed that reducing the speed by 10 km/h will reduce the coefficient of exceedances by approximately 20 %.

**Keywords:** traffic noise; noise hazard; coefficient of exceedances; Cnossos noise model.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The results obtained from measurements of acoustic parameters expressed on the decibel scale, the algebra of their transformations, and the comparison metric should correctly identify the phenomena under study. However, the noise hazard indicators used today – expressed as the Euclidean measure of the difference between the measurement result and the adopted standard values on the decibel scale – do not adequately reflect their perception by humans (WEIDENFELD *et al.*, 2021; VERBEEK, 2018; WUNDERLI *et al.*, 2016; KHAN, BURDZIK, 2023). These indicators may be difficult to understand for the majority of society and may make it difficult to comprehensively assess the harmfulness of noise. Furthermore, the use of the Euclidean evaluation measure also makes it difficult to interpret the obtained results (BATKO *et al.*, 2023; PRZYSUCHA *et al.*, 2021; PETERS, 2020; SAHU *et al.*, 2021).

Therefore, it is advisable to search for a transformation that reduces the process of identifying acoustic hazards and assessing the acoustic effectiveness of completed road modernizations from the Euclidean space of comparisons of measurement results expressed in the decibel scale to a scalar space. Adopting this concept of solving the problem enables to assess the state of acoustic hazard using classic identification algorithms. The solution sought should also meet the inverse condition, allowing the obtained scalar results to be reduced to the decibel space. This calculation method requires the determination of a new measure of exceedance of the recommended noise indicator values, as well as the development of a new algorithm for controlling acoustic hazard to the environment (SÁNCHEZ-FERNÁNDEZ, 2021).

The adopted scalar indicator, integrated with traditional descriptors that are mandatory under regulations, could provide additional information, simplify

and reduce the costs of assessing the acoustic quality of the environment, thereby increasing general awareness of the harmful effects of noise (GRAZIOUSO *et al.*, 2022; WUNDERLI *et al.*, 2016). Some attempts to address this problem were made in the work (SAHU *et al.*, 2021), where it was proposed to use a parameter called “noise exposure index” as an alternative measure for assessing acoustic hazards. Similarly, MOROE and MABASO (2022) proposed to use a pollution modeling technique and a parameter known as the “pollution standard index”. This approach takes into account the values of individual noise parameters, which are then compared to a single quality standard number. In (SMIRAGLIA *et al.*, 2016) normalization was used, consisting of Euclidean division of the recorded hourly noise values by the maximum value. In a similar way, YANG *et al.* (2020) and UPADHYAY *et al.* (2023) calculated the share of sound energy in each frequency band as a percentage of the total acoustic energy emitted by the tested vehicle. The advantage of these models is that the result is a dimensionless number. However, a weakness of these proposed solutions is the incompatibility of the transformations used with the formalism of decibel algebra.

In this article, the authors aimed to assess noise hazards and the acoustic effectiveness of implemented road modernizations by adopting a measure of exceedance of recommended sound levels in accordance with the decibel algebra relations. They propose to correlate this measure with the relation of references of two sound levels to each other, which is one of the basic relations in decibel algebra. The functional properties of such a solution were verified using an example of assessing acoustic hazards on one of the main communication road of the city of Kielce. This assessment was based on the results of continuous noise monitoring in the vicinity of the analyzed road after its modernization. The functional properties of both acoustic hazard estimation solutions were assessed, comparing the classic Euclidean approach to exceedance comparisons of recommended values with the approach proposed by the authors.

## 2. Scalar dimension of decibel numbers

The idea of the solution proposed by the authors is to replace the analysis of the values of acoustic parameters expressed on the decibel scale with the analysis of a coefficient (marked as  $k_i$ ) representing their multiplicity of exceedance relative to reference noise levels values ( $L_{\text{ref}}$ ). This operation can be performed in accordance with the division operation, present in decibel algebra, applied to acoustic disturbances described by sound levels  $L_i$ ,  $i = 0, 1, 2, \dots, n$ , and  $L_{\text{ref}}$ :

$$k_i = \frac{10^{0.1L_i}}{10^{0.1L_{\text{ref}}}} = 10^{0.1(L_i - L_{\text{ref}})}, \quad (1)$$

where  $L_i$  – the designated noise level,  $L_{\text{ref}}$  – the reference noise level value, for example, according to the World Health Organization [WHO] (2018): 53 dB for the day or 45 dB for the night. The result obtained in such an operation is a scalar and dimensionless quantity. This means that it is possible to perform further operations on the obtained result, in a manner typical of operations performed on Euclidean numbers. The inverse transformation from the space of scalar numbers to the space of numbers expressed in decibels is determined by multiplying the sound level by the designated scalar  $k_i$ . We can transform Eq. (1) as follows:

$$L_i = L_{\text{ref}} + 10 \log(k_i). \quad (2)$$

The average value of the sound level can be determined from the relationship:

$$\bar{L} = L_{\text{ref}} + 10 \log \bar{k}, \quad (3)$$

where  $\bar{k}$  is the average value of exceedances of the reference noise level, e.g.,  $\bar{k} = \frac{1}{n} \sum_{i=1}^n k_i$  or the median. The average value of exceeding the reference noise level on the decibel scale can be determined from the relationship:

$$\Delta \bar{L} = \bar{L} - L_{\text{ref}} = 10 \log \bar{k}. \quad (4)$$

After performing simple transformations, we can determine other parameters, e.g.:

- deviation from the reference value:

$$\sigma_L = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (10 \log k_i)^2}; \quad (5)$$

- positional coefficient of variation:

$$V_{Q_{31}} = \frac{Q_{31}}{\text{Med}} \cdot 100\%, \quad (6)$$

where  $Q_{31} = 0.5 \cdot [Q_3 - Q_1]$ ,  $Q_1$ ,  $Q_2$ , and  $Q_3$  – quartiles, corresponding to the first, second, and third quartiles, respectively, and Med – the median of the equivalent sound level;

- dispersion coefficient of quarter deviation:

$$V_{Q_1 Q_3} = \frac{Q_3 - Q_1}{Q_1 + Q_3} \cdot 100\%. \quad (7)$$

Equation (3), based on the analysis of the variability of the  $k_i$  coefficient value, provides conditions for estimating noise hazard assessment. For this reason, the authors decided to use the parameter  $k_i$  defined by Eq. (1) to assess the noise of road transport. This parameter is also used to assess the occupational risk of the harmful impact of vibroacoustic disturbances on employees' health (PLEBAN *et al.*, 2021; SEIXAS *et al.*, 2005; ROBERTS *et al.*, 2018). The analysis of the  $k_i$  coefficient facilitates the comparison of constant components (expected value or median) and variables of

the analyzed signals. The research on the components of the variables contained in the analyzed signals was based, among others, on the analysis of positional coefficients: deviation from the normative value, the coefficient of variation of the quarter deviation (denoted as  $V_{Q_{31}}$ ), and the dispersion coefficient of the quarter deviation (marked as  $V_{Q_1Q_3}$ ).

### 3. Monitoring of traffic noise hazards

The basis for verification analyzes of the suitability of the proposed noise assessment procedure was data recorded by an automatic station for continuous noise and traffic monitoring located at Popiełuszko Avenue in Kielce. Popiełuszki Avenue consists of four lanes separated by a green belt approximately 3 m wide. This avenue is a section of national road no. 73 and serves as the main the exit route from the center of Kielce towards Tarnów. This avenue is used for transit, urban and suburban traffic. The monitoring station includes a road radar, a sound level meter, and a weather station. Traffic flow was measured with the Wavetronix digital radar. Measurements were carried out 24 hours a day from 2011 to 2016. Traffic volume and speed data were recorded every minute (buffered) and the averaged results were reported every hour. However, due to various technical problems, the monitoring station did not always correctly record traffic parameters, and thereby the database was incomplete.

Based on the counts, traffic volume (understood as the sum of the number of light motor vehicles, medium heavy vehicles, heavy vehicles, and powered two-wheelers registered during the specified time period on all four lanes) and speed were calculated, divided into seven days of the week, hours and three sub-periods of the day. Because no data was recorded for some days, only days for which traffic parameters were recorded hourly throughout the entire 24-hour period were included in the analysis. Acoustic measurements were carried out with the SVAN 958A device, which is a four-channel, digital vibration and sound level meter of class 1. A condenser microphone from Microtech Gefell MK 250, prepolarized  $1/2''$ , class 1, with a sensitivity of 50 mV/Pa and an SV 12L preamplifier, was used. Detailed information about the measurement station is presented in the paper (BAKOWSKI, RADZISZEWSKI, 2022). Calculations of the equivalent sound level were performed for three time sub-periods, i.e., day – from 6 a.m. to 6 p.m., evening – from 6 p.m. to 10 p.m., and night – from 10 p.m. to 6 a.m.

The study analyzed the results of acoustic measurements carried out in the period from March 25, 2016 to June 25, 2016. It should be noted that in 2013–2014 Popiełuszki Avenue was thoroughly modernized. In 2016, there was a 13.1 % increase in traffic volume compared to 2011 for all vehicle types. The analysis of the distribution of individual vehicle groups in

the traffic structure in 2016 showed that the predominant group are light motor vehicles, whose share in the traffic flow increased to 71 % and on Sundays to 84 %. The next group were medium heavy vehicles, for which the share in the analyzed period did not change significantly and amounted to 18 % and on Sundays 11 %. The share of heavy vehicles was 8 % and on Sundays 2 %. Furthermore, the speed below which 85 % of vehicles on lanes 2 and 3 traveled was 85 km/h in 2011 and 81 km/h in 2016. In turn, the speed of vehicles in the analyzed period on lanes 1 and 4 was approximately 72 km/h (BAKOWSKI, RADZISZEWSKI, 2022).

Road noise in urbanized areas can be analyzed based on adopted parameters, time intervals and road location within the communication system (LU *et al.*, 2019; JANDACKA *et al.*, 2022; MELLER *et al.*, 2023). Within one year or one week, the variability of road noise is significantly different from the variability on weekdays or weekends and holidays (BAKOWSKI, RADZISZEWSKI, 2022). The problem of traffic noise on weekends is much less frequently analyzed in the literature, which is mainly due to lower traffic flow, especially heavy vehicles. However, as the authors' research has shown, when traffic volume decreases, vehicle speed increases and, as a result, the sound pressure level does not change significantly, especially when the equivalent sound level on the decibel scale is used as a measure (BAKOWSKI, RADZISZEWSKI, 2023). For this reason, the authors decided to conduct a noise analysis on Fridays, when traffic volume is highest, and on Sundays, when it is lowest. The analyzes were carried out for the night time sub-period because, as noise during this time is most annoying to residents (European Environment Agency, 2020).

Figure 1 shows the results of the equivalent sound level (marked  $L_{Ni}$ ) and the  $k_i$  coefficient calculated for subsequent nights, i.e., in the time interval from 10 p.m. to 6 a.m. in 2016.

The values of the  $L_{Ni}$  parameter range from approximately 64 dB(A) to 68 dB(A) and from approximately 62 dB(A) to 64 dB(A) on Sundays. The values of selected  $L_{Ni}$  parameters in 2016, two years after the completion of the road modernization, are presented in Table 1. The median  $L_{Ni}$  in 2016 was 65.6 dB(A) and 66.9 dB(A) in 2011. However, the maximum value of  $L_{Ni}$  in 2016 was 68 dB(A), whereas it was 69.7 dB(A) in 2011. The high variability in noise levels may cause much greater nuisance to local residents than suggested by the median value alone. The type A uncertainty of noise measurements  $u_A$  determined in accordance with the ISO standard (International Organization for Standardization, 2016) is less than 0.2 dB(A). The presented data on the decibel scale suggests that despite the road modernization, the noise level values have only changed slightly.

The analysis of acoustic hazards using the  $k_i$  parameter indicates more suggestively than the decibel

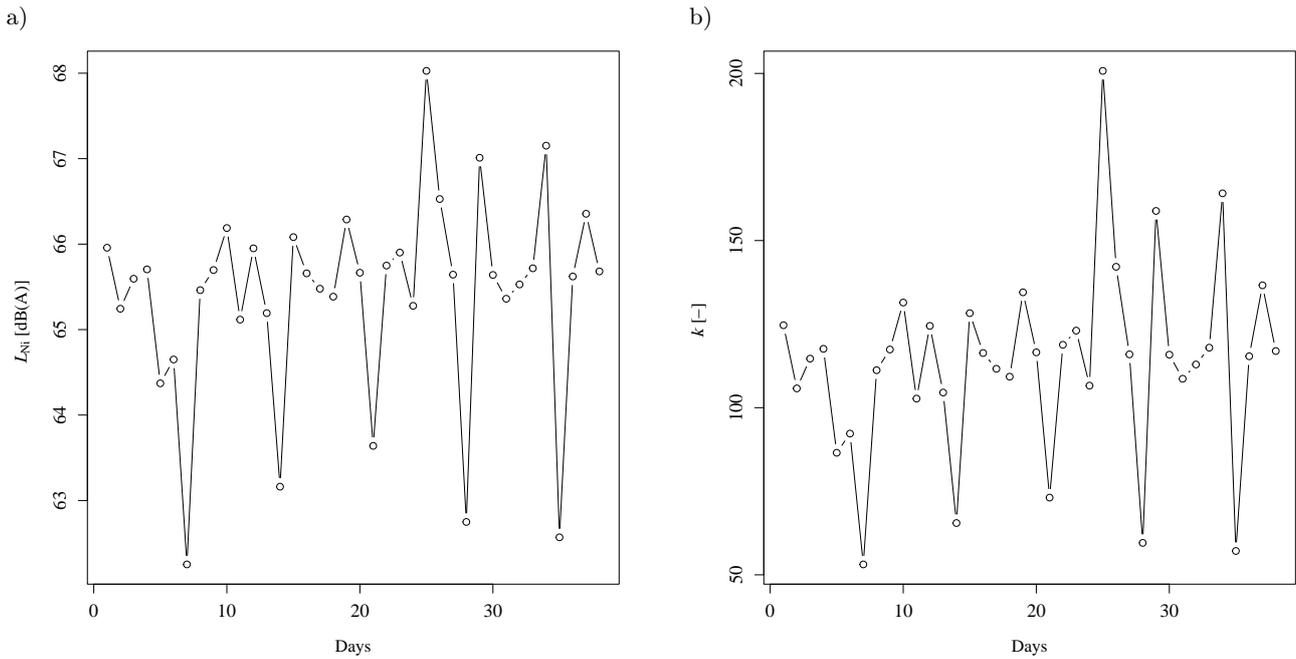


Fig. 1. Noise parameters experimentally determined for individual nights: a) equivalent sound level  $L_{Ni}$  values on the decibel scale; b) coefficient  $k_i$  of multiplicity of exceedance of the recommended noise levels by WHO (45 dB) noise levels.

Table 1. Road noise statistics on the decibel scale and coefficient  $k_i$  of noise level 45 dB recommended by WHO, two years after road modernization.

	$L_{eq}$	$Q_2$	$\bar{k}$	$\sigma$	$Q_3$	$L_i$ [max]	$k_i$ [max]	$u_A$	$V_{Q_{31}}$ [%]	$V_{Q_1 Q_3}$ [%]
$L_N$ [dB(A)]	65.5	65.6	–	1.2	65.9	68.0	–	0.19	7.9	7.9
$\frac{k_N}{L_0 = 45}$ [dB]	–	116	114	29	124	–	201	4.6	7.9	7.9

$$L_{eq} = 10 \log \left( \frac{1}{N} \sum_{i=1}^N 10^{0.1 L_i} \right) - \text{logarithmic mean; } \sigma_{L_{eq}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (L_i - L_{eq})^2} - \text{standard deviation;}$$

$$\sigma_k = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (k_i - \bar{k})^2} - \text{standard deviation; } u_A = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (L_i - L_{eq})^2} - \text{uncertainty;}$$

$$k_i = 10^{0.1(L_{Ni} - 45)} - \text{multiplicity of the sound level recommended by WHO for the day-night time;}$$

$L_N$  – noise level for the sub-period of day-night, on subsequent days.

scale the dangers resulting from exceedance of the recommended road noise levels (see Table 1) and the acoustic effects of road modernization (PLEBAN *et al.*, 2021). The arithmetic mean of this parameter is 114, which means that the noise energy emission is 114 times higher than recommended by WHO. Even on Sundays, the  $k_i$  value ranges from about 50 to 70, as shown in Fig. 1b.

The analysis of the value of the coefficient  $k_i$  after the completion of the road modernization also indicates a reduction in the nuisance of acoustic disturbances. The type A uncertainty of determining  $k_i$  is about 5. The arithmetic mean of the  $k_i$  parameter decreased after the road modernization by about 24 %. However, the maximum value of  $k_i$  decreased by approximately 31 %. It can be seen that the values of the relative variability coefficients  $L_{Ni}$  and  $k_i$  are equal regardless of the calculation method. This confirms the

correctness of the applied calculation procedures for these coefficients. The values of noise variability coefficients were reduced by approximately 54 %, which confirms that vehicle traffic has calmed down to some extent. This contributes to increasing the safety of road users.

The values of noise parameters averaged over the year for individual nights are used in the development of static noise maps and indicate the possibility of health problems among local residents (WHO, 2018; SRAMEK *et al.*, 2022; RANPIS, TANDEL, 2022). Additional information about the risk of acoustic hazards can be obtained by analyzing the values of noise parameters averaged over the year, on different days of the week, and during subsequent hours of the day. Such information may facilitate taking administrative actions aimed at reducing noise nuisance, e.g., limiting speed and movement of certain groups of vehicles at

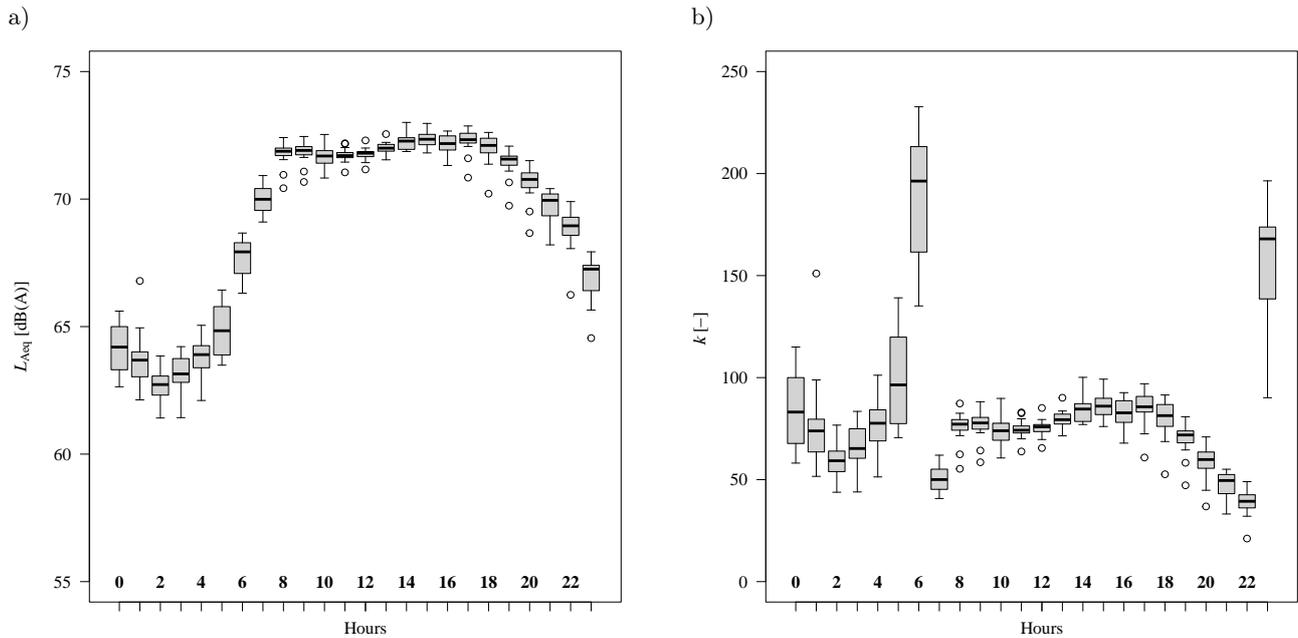


Fig. 2. Box plot of traffic noise on Fridays: a) noise level on the decibel scale in subsequent hours of the day; b) coefficients of exceedance of the noise level recommended by WHO.

certain hours of the day. Figure 2a shows the box plot of the noise level, and Fig. 2b shows the exceedance rates in subsequent hours on Fridays. Comparing both of these drawings, it can be seen that the scalar scale has a higher resolution, which facilitates the analysis of changes and the identification of extreme noise values. Figure 2b shows that exceedances of permissible noise levels in the morning and late at night are the highest for the entire 24-hour period.

A sudden change in the  $k_i$  value can be seen in Fig. 2b at 6 a.m. and 11 p.m. It results from the change in the  $L_{ref}$  value from 45 dB for the night sub-period to 53 dB for the day sub-period (WHO, 2018). The noise level on Sundays increases gradually from 6 a.m. to its maximum value at 3 p.m., and then gradually decreases, as shown in Fig. 3a. In the case of Sundays, it can be seen in Fig. 3b that the maximum value of the  $k_i$  coefficient occurs at 11 p.m.

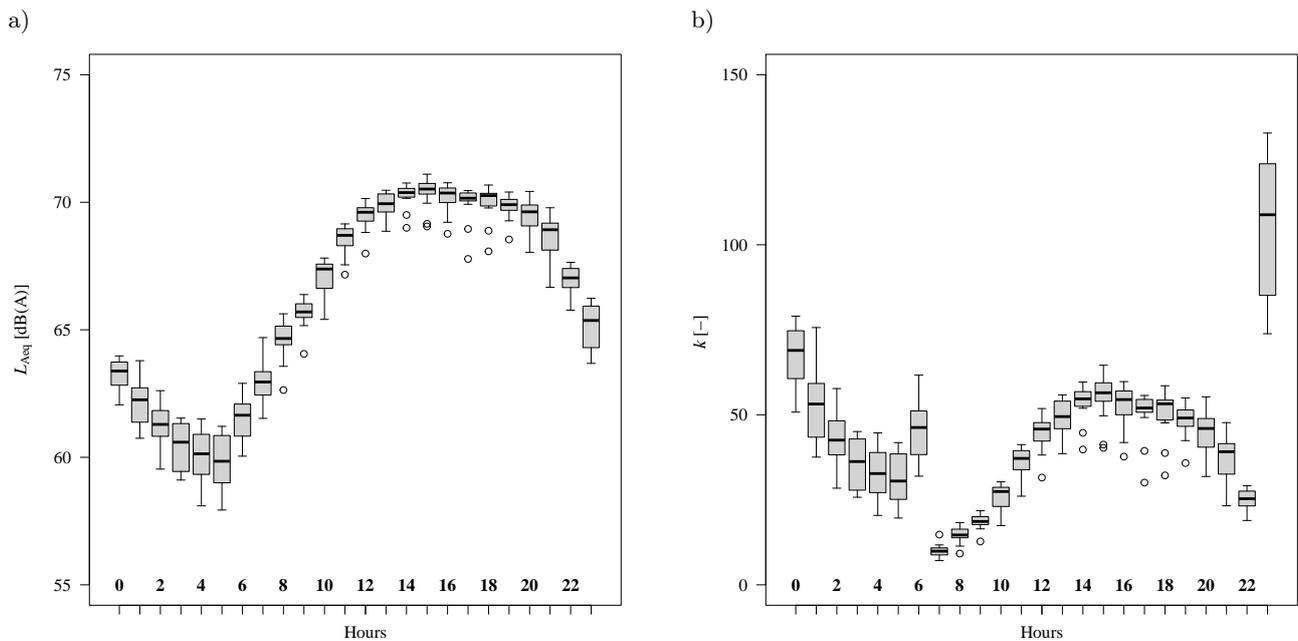


Fig. 3. Box plot of traffic noise on Sundays: a) noise level on the decibel scale in subsequent hours of the day; b) coefficients of exceedance of the noise level recommended by WHO.

The analyses conducted revealed that road modernization contributed to the reduction of acoustic hazards, but this reduction is still not satisfactory. For this reason, in this study, the authors decided to conduct numerical simulations consistent with the Cnossos noise model to assess the impact of selected administrative restrictions on noise parameters. Simulations were performed to evaluate the effects of introducing a ban on heavy vehicles traffic at night on Fridays and Sundays. The calculations indicated that the lack of heavy vehicles would result in a decrease in the value of the  $k_i$  coefficient, e.g., on Fridays at 6 a.m. from approximately 196 to 106, which corresponds to a change in the sound level from 68 dB(A) to 65.25 dB(A). However, on Sundays, e.g., at 6 a.m., this ban would result in a decrease in the value of the  $k_i$  coefficient from approximately 46 to 38, which corresponds to a change in the sound level from 61.65 dB(A) to 60.80 dB(A). Simulations were also carried out on the impact of reducing the speed of each vehicle by 10 km/h and 20 km/h on noise, e.g., at 6 a.m. in the vicinity of the monitoring station. The calculations showed that the speed decreased:

- a) a reduction of 10 km/h on Fridays led to a decrease in the value of the  $k_i$  coefficient from approximately 196 to 158, which corresponds to a reduction in the noise level from 68 dB(A) to 67 dB(A);
- b) a reduction of 20 km/h on Fridays resulted in a decrease in the value of the  $k_i$  coefficient from approximately 196 to 128, which corresponds to a reduction in the noise level from 68 dB(A) to 66 dB(A);
- c) on Sundays, a reduction of 10 km/h caused a decrease in the value of the  $k_i$  coefficient, e.g., from approximately 46 to 37, which corresponds to a reduction in the noise level from 62 dB(A) to 61 dB(A);
- d) on Sundays, a reduction of 20 km/h led to a decrease in the value of the  $k_i$  coefficient from approximately 46 to 29, which corresponds to a reduction in the noise level from 62 dB(A) to 60 dB(A). The results obtained from these simulations indicate that even a slight reduction in the speed of each vehicle results in satisfactory noise reduction effects.

#### 4. Summary and conclusions

The article discusses the problem of the correctness of the applied rules for assessing the state of acoustic hazards to the environment, expressed on the decibel scale, based on the Euclidean measure for assessing exceedances of normative values. It emphasizes the need to modify algorithms for assessing acoustic hazards, taking into account the appropriate metric

for their comparison. The study proposes an original solution for controlling acoustic hazards based on the number of times the recommended values are exceeded. The solution used eliminates methodological and metrological paradoxes of currently used procedures based on the Euclidean measure of exceedance of recommended values. The proposed procedure allows for:

- transformation of the analysis of decibel-based noise hazard assessments into scalar number calculations within the Euclidean space;
- simulations and identification of noise hazards using standard calculation packages, including classic uncertainty estimation procedures regarding exceedances of permissible noise levels.

The article presents the results of the analysis of the proposed solution and the assessment of its usefulness, based on the example of diagnosing the state of noise hazards in the vicinity of one of the main streets of the city of Kielce. Its use enabled an objective assessment of the acoustic effects of a comprehensive road modernization. As research into acoustic hazards in its surroundings has shown, there has been some reduction in the noise level. The analyzes were performed using noise measures on the decibel scale and the proposed scalar measure of the multiplicity of exceeding the recommended sound level. The reduction in noise levels assessed according to the first measure suggests slight changes in the emission of acoustic hazards. For example, the median  $L_{Ni}$  before the road modernization was 66.9 dB(A) and after the modernization 65.6 dB(A). The assessment of the impact of road modernization on traffic noise according to the proposed method, i.e., assessing the value of the coefficient  $k_i$  of multiplicity of exceedances of recommended noise level, showed a decrease of approximately 25 % for the night sub-period. The maximum values of coefficient  $k_i$  after road modernization decreased by 32 %. Analyses of the coefficients of variation of noise level showed a decrease in their values after road modernization by approximately 40 %, regardless of the scale used. The time of calculating the values of the analyzed parameters (for example  $V_{Q_1Q_3}$ ) based on coefficient  $k_i$  is approximately 20 % shorter than when using the decibel scale. The results of the conducted numerical simulations justify the advisability of introducing road signs with variable content to enforce certain limits on the values of vehicle traffic parameters at certain hours of the day.

Overall, the proposed procedure for assessing acoustic hazards to the environment takes into account the premises that support the need to revise the solutions currently used in the practice of environmental control. This procedure may offer a response to the need of implementing new methods for diagnosing acoustic hazards generated by road transport.

## Acknowledgments

The authors thank the Municipal Road Authority in Kielce for providing the measurement data used in this study.

## Additional information

The authors declare that there are no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited, and that appropriate permits were obtained.

## References

- BATKO W., RADZISZEWSKI L., BĄKOWSKI A. (2023), Limitations of decibel algebra in the study of environmental acoustic hazards, [in:] *AIP Conference Proceedings*, **2949**(1): 020001, doi: [10.1063/5.0166002](https://doi.org/10.1063/5.0166002).
- BĄKOWSKI A., RADZISZEWSKI L. (2022), Analysis of the traffic parameters on a section in the city of the national road during several years of operation, *Communications – Scientific Letters of the University of Zilina*, **24**(1): 12–25, doi: [10.26552/com.C.2022.1.A12-A25](https://doi.org/10.26552/com.C.2022.1.A12-A25).
- BĄKOWSKI A., RADZISZEWSKI L. (2023), Urban tidal flow noise-case study, *Vibrations in Physical Systems*, **34**(1), doi: [10.21008/j.0860-6897.2023.1.19](https://doi.org/10.21008/j.0860-6897.2023.1.19).
- European Environment Agency (2020), *Environmental noise in Europe, 2020*, Publications Office, doi: [10.2800/686249](https://doi.org/10.2800/686249).
- GRAZIUSO G., FRANCAVILLA A.B., MANCINI S., GUARNACCIA C. (2022), Application of the Harmonica Index for noise assessment in different spatial contexts, [in:] *Journal of Physics: Conference Series*, **2162**(1): 012006, doi: [10.1088/1742-6596/2162/1/012006](https://doi.org/10.1088/1742-6596/2162/1/012006).
- International Organization for Standardization (2016), *Acoustics Description, measurement and assessment of environmental noise. Part 1: Basic quantities and assessment procedures* (ISO Standard No. 1996-1:2016), <https://www.iso.org/standard/59765.html>.
- JANDACKA D., DECKY M., HODASOVA K., PISCA P., BRILIAK D. (2022), Influence of the urban intersection reconstruction on the reduction of road traffic noise pollution, *Applied Sciences*, **12**(17): 8878, doi: [10.3390/app12178878](https://doi.org/10.3390/app12178878).
- KHAN D., BURDZIK R. (2023), Measurement and analysis of transport noise and vibration: A review of techniques, case studies, and future directions, *Measurement*, **220**: 113354, doi: [10.1016/j.measurement.2023.113354](https://doi.org/10.1016/j.measurement.2023.113354).
- LU X., KANG J., ZHU P., CAI J., GUO F., ZHANG Y. (2019), Influence of urban road characteristics on traffic noise, *Transportation Research Part D: Transport and Environment*, **75**: 136–155, doi: [10.1016/j.trd.2019.08.026](https://doi.org/10.1016/j.trd.2019.08.026).
- MELLER G., DE LOURENÇO W.M., DE MELO V.S.G., DE CAMPOS GRIGOLETTI G. (2023), Use of noise prediction models for road noise mapping in locations that do not have a standardized model: A short systematic review, *Environmental Monitoring and Assessment*, **195**(6): 740, doi: [10.1007/s10661-023-11268-9](https://doi.org/10.1007/s10661-023-11268-9).
- MOROE N., MABASO P. (2022), Quantifying traffic noise pollution levels: A cross-sectional survey in South Africa, *Scientific Reports*, **12**(1): 3454, doi: [10.1038/s41598-022-07145-z](https://doi.org/10.1038/s41598-022-07145-z).
- PETERS R. [Ed.] (2020), *Uncertainty in Acoustics: Measurement, Prediction and Assessment*, CRC Press, doi: [10.1201/9780429470622](https://doi.org/10.1201/9780429470622).
- PLEBAN D., SMAGOWSKA B., RADOSZ J. (2021), Assessment of occupational risk in the case of the ultrasonic noise exposure, *Archives of Acoustics*, **46**(1): 167–175, doi: [10.24425/aoa.2021.136570](https://doi.org/10.24425/aoa.2021.136570).
- PRZYSUCHA B., PAWLIK P., STEPIEŃ B., SUROWIEC A. (2021), Impact of the noise indicators components correlation  $L_d$ ,  $L_e$ ,  $L_n$  on the uncertainty of the long-term day–evening–night noise indicator  $L_{den}$ , *Measurement*, **179**: 109399, doi: [10.1016/j.measurement.2021.109399](https://doi.org/10.1016/j.measurement.2021.109399).
- RANPISE R.B., TANDEL B.N. (2022), Urban road traffic noise monitoring, mapping, modelling, and mitigation: A thematic review, *Noise Mapping*, **9**(1): 48–66, doi: [10.1515/noise-2022-0004](https://doi.org/10.1515/noise-2022-0004).
- ROBERTS B., SEIXAS N.S., MUKHERJEE B., NEITZEL R.L. (2018), Evaluating the risk of noise-induced hearing loss using different noise measurement criteria, *Annals of Work Exposures and Health*, **62**(3): 295–306, doi: [10.1093/annweh/wxy001](https://doi.org/10.1093/annweh/wxy001).
- SAHU A.K., NAYAK S.K., MOHANTY C.R., PRADHAN P.K. (2021), Traffic noise and its impact on wellness of the residents in sambalpur city – A critical analysis, *Archives of Acoustics*, **46**(2): 353–363, doi: [10.24425/aoa.2021.136588](https://doi.org/10.24425/aoa.2021.136588).
- SÁNCHEZ-FERNÁNDEZ L.P. (2021), Environmental noise indicators and acoustic indexes based on fuzzy modelling for urban spaces, *Ecological Indicators*, **126**: 107631, doi: [10.1016/j.ecolind.2021.107631](https://doi.org/10.1016/j.ecolind.2021.107631).
- SEIXAS N., NEITZEL R., SHEPPARD L., GOLDMAN B. (2005), Alternative metrics for noise exposure among construction workers, *The Annals of Occupational Hygiene*, **49**(6): 493–502, doi: [10.1093/annhyg/mei009](https://doi.org/10.1093/annhyg/mei009).
- SMIRAGLIA M., BENOCCI R., ZAMBON G., ROMAN H.E. (2016), Predicting hourly traffic noise from traffic flow rate model: Underlying concepts for the dynamap project, *Noise Mapping*, **3**(1): 130–139, doi: [10.1515/noise-2016-0010](https://doi.org/10.1515/noise-2016-0010).
- SRAMEK J., HODASOVA K., JUHAS M., PITONAK M., DURIS L. (2022), Rutting prediction models in holistic concept to sustainability of semi-rigid pavements, *Civil and Environmental Engineering*, **18**(1): 200–208, doi: [10.2478/cee-2022-0019](https://doi.org/10.2478/cee-2022-0019).

22. UPADHYAY S., PARIDA M., KUMAR B. (2023), Development of a reference energy mean emission level traffic noise models for bituminous pavement for mid-sized cities in India, [in:] *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, **265**(5): 2899–2906, doi: [10.3397/IN\\_2022\\_0408](https://doi.org/10.3397/IN_2022_0408).
23. VERBEEK T. (2018), The relation between objective and subjective exposure to traffic noise around two suburban highway viaducts in Ghent: Lessons for urban environmental policy, *Local Environment*, **23**(4): 448–467, doi: [10.1080/13549839.2018.1428791](https://doi.org/10.1080/13549839.2018.1428791).
24. WEIDENFELD S., SANOK S., FIMMERS R., PUTH M.T., AESCHBACH D., ELMENHORST E.M. (2021), Short-term annoyance due to night-time road, railway, and air traffic noise: Role of the noise source, the acoustical metric, and non-acoustical factors, *International Journal of Environmental Research and Public Health*, **18**(9): 4647, doi: [10.3390/ijerph18094647](https://doi.org/10.3390/ijerph18094647).
25. World Health Organization (2018), *Environmental noise guidelines for the European region*, Regional Office for Europe.
26. WUNDERLI J.M. *et al.* (2016), Intermittency ratio: A metric reflecting short-term temporal variations of transportation noise exposure, *Journal of Exposure Science & Environmental Epidemiology*, **26**(6): 575–585, doi: [10.1038/jes.2015.56](https://doi.org/10.1038/jes.2015.56).
27. YANG W., CAI M., LUO P. (2020), The calculation of road traffic noise spectrum based on the noise spectral characteristics of single vehicles, *Applied Acoustics*, **160**: 107128, doi: [10.1016/j.apacoust.2019.107128](https://doi.org/10.1016/j.apacoust.2019.107128).

## Research Paper

## Evaluation Modeling of Electric Bus Interior Sound Quality Based on Two Improved XGBoost Algorithms Using GS and PSO

Enlai ZHANG<sup>(1).(2)</sup>, Yi CHEN<sup>(1)</sup>, Liang SU<sup>(3)\*</sup>, Ruoyu ZHONGLIAN<sup>(1)</sup>,  
Xianyi CHEN<sup>(1)</sup>, Shangfeng JIANG<sup>(1)\*\*</sup>

<sup>(1)</sup> School of Mechanical and Automotive Engineering, Xiamen University of Technology  
Xiamen, China

<sup>(2)</sup> Xiamen Key Laboratory of Robot Systems and Digital Manufacturing  
Xiamen, China

<sup>(3)</sup> Bus Engineering Research Institute, Xiamen King Long United Automotive Industry Co., Ltd  
Xiamen, China

Corresponding Authors' e-mails: \*slkinglong@163.com, \*\*jiangsfa@163.com

(received October 10, 2023; accepted March 15, 2024; published online April 24, 2024)

There is no doubt that traffic noise has become one of the main sources of urban noise, and the electric bus, as an important means of transport frequently used by people in daily life, has a direct impact on the psychological and auditory health of passengers due to its interior noise characteristics. Consequently, studying electric bus sound quality is an important way to improve vehicle performance and comfort. In this paper, eight electric buses were selected and 64 noise samples were measured. Acoustic comfort was taken as an evaluation index, professionals were organized to complete the subjective evaluation tests for all noise samples based on rank score comparison (RSC). And nine psycho-acoustic objective parameters such as loudness, sharpness and roughness were calculated using Artemis software to establish the sound quality database of electric buses. Aiming at the practical application requirements of high-precision modeling of acoustic comfort in vehicles, this paper presented two improved extreme gradient boosting (XGBoost) algorithms based on grid search (GS) method and particle swarm optimization (PSO), respectively, with objective parameters and acoustic comfort as input and output variables, and established three regression models of standard XGBoost, GS-XGBoost, and PSO-XGBoost through data training. Finally, the calculation results of three indexes of average relative error, square root error and correlation coefficient indicate that the proposed PSO-XGBoost model is significantly better than GS-XGBoost and standard XGBoost, with its prediction accuracy as high as 97.6 %. This model is determined as the evaluation model of interior acoustic comfort for this case, providing a key technical support for future sound quality optimization of electric buses.

**Keywords:** electric bus; sound quality; acoustic comfort; GS-XGBoost; PSO-XGBoost.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As an important form of green transport, electric bus has been developing rapidly domestically in recent decades. However, there is serious homogenization in the same class of models in terms of motor electric control, battery range, vehicle safety, and other conventional performance measures, resulting in increased industry competition. Focusing on internal quality im-

provement and improving user experience have become a differentiation strategy for electric bus development, shifting focus from high speed to high quality. Practice demonstrates that vehicle interior noise quality is the most direct factor affecting people's subjective experiences. An excellent acoustic environment is conducive to the physical and mental health of drivers and passengers, and significantly improves users' satisfaction with the vehicle (STEINBACH, ALTINSOY, 2019;

ZHANG *et al.*, 2021). Therefore, the change of electric bus noise control from noise reduction to sound quality is of great application importance in establishing the core competitiveness of vehicle products.

Because driving motor replaces conventional engine, the noise inside electric bus is more pronounced due to the absence of engine masking effect, and many noise types that are not easily to be detected in the fuel bus appear more prominent in electric buses, such as air conditioning noise, electromagnetic noise, tire noise, and mechanical transmission noise (SHI *et al.*, 2018; ZHANG *et al.*, 2022b). In particular, the motor electromagnetic noise is characterized by high current, variable frequency regulation and high magnetic density. Its spectral characteristics often fall within the sensitive range of human ears to noise, leading to a harsh subjective feeling, which has a significant negative impact on the whole vehicle acoustic comfort (DOLESCHAL, VERHEY, 2022). It can be concluded that it is urgent and imperative for the development of electric buses to improve interior sound quality.

It is well-known that subjective and objective evaluations are the main issues in sound quality research. Subjective evaluation can directly reflect people's auditory feelings, but the evaluation process requires a lot of human and material resources, and its results are susceptible to the psychological and physiological factors of evaluators (ZHANG *et al.*, 2018). Therefore, on the basis of obtaining noise database, using objective parameters as independent variables and subjective evaluation test results as dependent variables, establishing a functional relationship between them through data fitting, namely, a quantifiable mapping model for acoustic quality evaluation, is a hot direction in current vehicle sound quality research.

Sound quality modeling approaches can be broadly classified into two categories: the first is based on mathematical statistics, mainly including multiple linear regression (ZHANG *et al.*, 2018), Kriging model (ZHANG *et al.*, 2020) and grey system theory (CHEN *et al.*, 2012); the second category is based on machine learning algorithms to simulate the ability of human neural networks to extract and process information features. Since the structure of the human ear resembles a nonlinear and complex sound receiver, studies have indicated that the nonlinear mapping relationship between subjective and objective evaluations can more accurately describe real auditory perception (WANG, 2009; POURSEIEDREZAEI *et al.*, 2021), in which the algorithms involved in this category include back-propagation neural networks (BPNN) (POURSEIEDREZAEI *et al.*, 2021; ZHANG *et al.*, 2016), support vector machine (LIANG *et al.*, 2020; DING *et al.*, 2023), deep learning (HUANG *et al.*, 2016; 2021), and extreme gradient boosting (XGBoost) (WANG *et al.*, 2022), etc. For example, HUANG *et al.* (2020) and KIM and LEE (2022), respectively, used convolutional

neural network to establish quantitative evaluation models of in-vehicle annoyance and driving sound quality. In the past two years, XGBoost has been applied to nonlinear modeling of vehicle annoyance due to its excellent fitting performance and generalization ability (WANG *et al.*, 2022; ZHANG *et al.*, 2023b). WANG *et al.* (2022) collected interior noise at different speeds to build a nonlinear XGBoost-based annoyance model using subjective and objective database.

XGBoost, first proposed by prof. Chen in 2016, uses second-order Taylor expansion to optimize the objective function and introduces regularization terms to improve the model's generalization ability and to effectively control its over-fitting problem (CHEN, GUESTRIN, 2016). In our past researches, for 64 electric bus sound quality modeling applications, BPNN and standard XGBoost were utilized to establish two high-precision interior acoustic comfort prediction models with average relative error of 4.35 % and 4.67 %, respectively, (ZHANG *et al.*, 2022a; 2023a). To make the prediction model more accurate and robust, based on the existing standard XGBoost regression model, we explore the combination of XGBoost with grid search (GS) and particle swarm optimization (PSO) to optimize the mapping model parameters. Thus, propose two improved algorithms of GS-XGBoost and PSO-XGBoost, to further expand the applicability of XGBoost in the field of vehicle sound quality.

## 2. Subjective and objective evaluations

### 2.1. Noise sample collection

Eight different types of electric buses, denoted A to H, were selected for noise testing. As there is no bus sound quality standardized test, the test and evaluation methods for city bus internal noise (XMQT075-2021) were used for reference. During the test, we chose two different working conditions with the air conditioner on and off. Measurements were taken at two positions: the driver's seat and the rear seat, as illustrated in Fig. 1. The sample buses operated at constant speeds of 30 km/h and 50 km/h on a professional track. When the tested bus was running stably, its interior noise signals were acquired by Squadriga II binaural acquisition system with a headset BHS II, as shown in Fig. 2.

### 2.2. Evaluation process

Obviously, in this test, the signal samples collected are transmitted to the human ear after the comprehensive superposition of various noise sources such as tire-road noise and electromagnetic noise. Since the

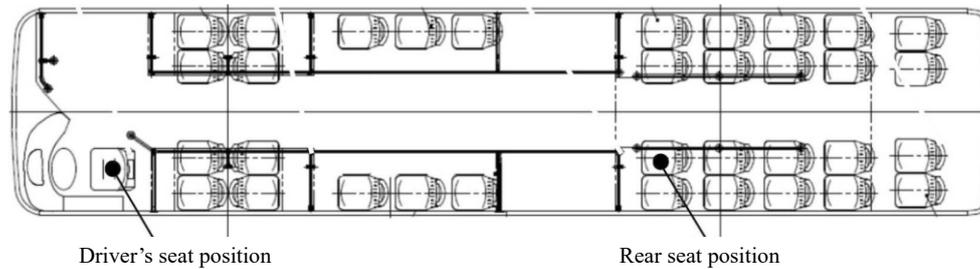


Fig. 1. Distribution of the two measuring points.



Fig. 2. Test scenes and measuring instruments.

human ear may feel auditory fatigue due to exposure to longer noise samples, pre-processing steps such as interception and screening were completed before the evaluation test, and a total of 64 in-vehicle noise samples from electric buses, each with a duration of 5 seconds, were finally acquired (ZHANG *et al.*, 2022a; 2023a). To avoid the potential information influence of measuring points, speeds and working conditions of noise samples on sound quality evaluation tests, all samples were re-coded and randomly ordered. The results are shown in Table 1.

Table 1. Random order of all noise samples.

AFD3	AFD5	AFM3	AFM5	AND3	AND5	ANM3	ANM5
49	18	9	37	1	32	13	21
BFD3	BFD5	BFM3	BFM5	BND3	BND5	BNM3	BNM5
44	27	36	20	46	8	30	41
CFD3	CFD5	CFM3	CFM5	CND3	CND5	CNM3	CNM5
24	19	10	26	54	48	51	2
DFD3	DFD5	DFM3	DFM5	DND3	DND5	DNM3	DNM5
31	55	40	4	62	12	29	57
EFD3	EFD5	EFM3	EFM5	END3	END5	ENM3	ENM5
3	34	15	22	58	11	35	63
FFD3	FFD5	FFM3	FFM5	FND3	FND5	FNM3	FNM5
60	43	7	52	5	16	64	25
GFD3	GFD5	GFM3	GFM5	GND3	GND5	GNM3	GNM5
23	33	28	61	17	59	39	53
HFD3	HFD5	HFM3	HFM5	HND3	HND5	HNM3	HNM5
14	47	6	45	38	50	42	56

Note – the code meaning is as follows: AND3 (number 1) represents the noise sample acquired from an electric bus in the driver's position at 30 km/h with the air conditioner turned on.

The sound quality evaluation process consists of subjective evaluation and objective acoustic parameter calculation. In this case, the rank score comparison (RSC) was used for subjective evaluation tests of electric bus noise samples. Acoustic comfort was taken as the evaluation index, divided into five comfort levels: poor, accepted, satisfied, good, and excellent, and each further added two values for each level. Thus, the acoustic comfort level range within [1, 10] was constituted. The jury for this subjective evaluation test was composed of NVH engineers, drivers and acoustic experts with rich experience in bus noise. It is well-known that different people may yield different evaluation results for the same noise sample due to psychological and emotional influences and cognitive differences. In order to ensure the validity and reliability of the subjective evaluation results, four measures were taken: (1) increase the number of acoustic experts, including university professors and senior engineers, who accounted for nearly half of the jury; (2) to avoid mutual interference between evaluators, only one evaluator was allowed to participate in the subjective evaluation in a specialized testing room; (3) for reducing the discreteness of subjective activities, the evaluation test was divided into two stages, i.e., pre-evaluation and final evaluation. In the pre-evaluation test, acoustic experts conducted subjective evaluation on the two selected comparison noise samples (numbered 34 and 49 in Table 1) and assigned appropriate acoustic comfort values. In the final evaluation test, all evaluators first played the two comparison samples and paid attention to their acoustic comfort values, and then combined them with their subjective feelings to complete the scoring of 64 noise samples in turn; (4) in the

data statistics stage, Spearman correlation coefficient and *K*-means clustering were used to analyze the data characteristics of all the evaluators, and after excluding the data with correlation coefficients lower than 0.7 and low similarity, the data of the remaining evaluators were averaged, that is, the acoustic comfort values of all noise samples were finally obtained, as listed in Table 2.

In terms of objective evaluation, Artemis software by HEAD was applied to calculate objective psychoacoustic parameters of all noise samples, including linear sound pressure level (dB), *A*-weighted sound pressure level (dB(A)), loudness (sone), sharpness (acum), roughness (asper), fluctuation strength (vacil), articulation index (AI, %), impulsiveness (iu), and relative approach (cPa), corresponding to independent variables  $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8,$  and  $x_9,$  respectively. The objective evaluation results of all noise samples are displayed in Table 3.

Table 2. Acoustic comfort values of 64 noise samples after subjective evaluations.

1	2	3	4	5	6	7	8
5.17	3.33	6.42	5.33	4.67	3.25	2.58	3.42
9	10	11	12	13	14	15	16
5.75	5.25	4.58	4.83	4.42	4.67	2.33	5.42
17	18	19	20	21	22	23	24
3.17	7.25	2.92	7.08	4.92	2.08	5.33	3.84
25	26	27	28	29	30	31	32
2.33	3.08	3.92	2.5	4.25	5.08	3.92	2.5
33	34	35	36	37	38	39	40
2.42	4.25	1.75	7.5	5.08	1.42	2.58	6.75
41	42	43	44	45	46	47	48
6.92	2.5	4	5.17	2.33	3.84	4.83	3.08
49	50	51	52	53	54	55	56
7.67	4.12	3.67	2.58	2.67	4.25	5.58	1.92
57	58	59	60	61	62	63	64
5.33	3.58	2.58	5.32	4.33	4.84	1.58	3.13

Table 3. Objective parameter calculation results of 64 noise samples.

Sample	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$
1	105.43	73.11	29.6	1.6	0.105	0.087	43.6	0.398	19.1
2	108.24	77.59	43.2	2.3	0.122	0.108	18.8	0.409	19.7
3	96.73	61.18	15.8	1.34	0.071	0.0722	79.4	0.57	15.7
4	108.53	74.4	34.1	1.84	0.0958	0.112	34.8	0.36	20.2
5	102	67.51	24.5	1.62	0.0827	0.0795	54.6	0.376	17.9
6	89.57	66.71	20.9	1.39	0.131	0.061	62.4	0.286	15.7
7	95.65	67.97	24.3	1.6	0.105	0.0623	55.7	0.334	16.3
8	102.38	74.24	34	1.99	0.12	0.0645	31.4	0.394	18.4
9	100.06	69.54	22.2	1.42	0.0951	0.0793	59.9	0.37	16.3
10	102.46	69.85	25.2	1.73	0.0974	0.0755	48.5	0.408	15.9
11	99.03	72.45	28.6	1.84	0.122	0.0638	41.2	0.343	16.6
12	107.04	72.78	33.4	1.78	0.0943	0.0954	37.9	0.431	20.3
13	100.57	70.76	25.7	1.72	0.106	0.0824	47.5	0.387	16.7
14	95.06	63.03	18.2	1.55	0.0878	0.0727	67.4	0.279	15.6
15	91.63	69.81	22.5	1.28	0.117	0.0638	63.1	0.358	13.7
16	106.93	72.01	33.7	1.74	0.0932	0.0967	41.2	0.465	21.4
17	97.07	69.2	25.3	1.8	0.106	0.0576	46	0.351	16.2
18	105.61	72.5	28.4	1.51	0.1	0.0847	47.2	0.414	19.1
...	...	...	...	...	...	...	...	...	...
50	101.53	71.83	31.6	1.59	0.117	0.0744	43.1	0.264	18.6
51	100.35	71.11	26.9	1.81	0.112	0.0664	44.8	0.393	16
52	101.18	73.57	35.3	1.86	0.123	0.0761	36	0.492	20.6
53	101.12	74.08	35.4	1.78	0.12	0.0649	35.2	0.349	18.4
54	107.55	72.01	31.6	2.01	0.097	0.0795	35.4	0.46	19.7
55	107.61	73.21	32	1.67	0.0889	0.092	41.8	0.474	20.8
56	95.57	74.79	31.5	1.67	0.157	0.0793	36.8	0.225	17.6
57	106.14	75.12	35	2.06	0.102	0.111	30.8	0.363	18.7
58	97.78	70.22	26.4	1.86	0.112	0.0574	44.9	0.343	16
59	103.2	78.46	42.8	1.68	0.104	0.106	35.3	0.322	19.2
60	101.53	65.3	21.7	1.41	0.0728	0.0804	66.5	0.464	18.9
61	100.28	72.34	29.3	1.58	0.1	0.0704	46.4	0.306	17.6
62	101.65	70.62	25.5	1.77	0.0837	0.0845	50.5	0.429	18.4
63	97.08	78.18	38.1	1.96	0.165	0.0779	26.5	0.334	16.3
64	96.98	70.89	29.9	1.72	0.117	0.0626	44.4	0.363	16.7

### 3. XGBoost and its improved algorithms

#### 3.1. XGBoost

XGBoost is implemented based on the gradient boosting decision tree using the integration method of boosting. The modeling idea is to first define an objective function, then find the best tree model to fit the residual error of the previous prediction in each iteration, and pursue to minimize the objective function so that predicted value is as close to targeted value as possible. Its theory and implementation process are detailed in the corresponding reports (WANG *et al.*, 2022; ZHANG *et al.*, 2023b).

XGBoost is a powerful and highly flexible machine learning model that is significantly affected by several structural parameters. These structural parameters are categorized into weak learner parameters and boosting framework parameters, mainly including the maximum depth of the decision tree, the maximum number of leaf nodes, the learning rate, etc. Different combinations of these parameters may result in the model exhibiting very different levels of performance on different tasks. Due to the complex interactions between these

parameters, manually tuning them to achieve optimal performance becomes extremely complex in the high-dimensional parameter space. Therefore, the introduction of intelligent optimization algorithms to efficiently search for optimal combinations of structural parameters with optimal performance is necessary to improve the prediction accuracy of the model.

#### 3.2. Improved XGBoost based on GS

The grid search (GS) method introduced in this section is an intelligent algorithm for structural parameter optimization (WANG *et al.*, 2020). The steps for combining it with XGBoost mainly include: traversing all possible values of specified parameters to form a parameter grid. Then, all parameter combinations in the grid search are put into the model successively for prediction, and the one with the smallest error among all parameter combinations is taken as the best parameter. Finally, a new GS-XGBoost model is established with the best parameters, thus realizing the parameter optimization of XGBoost (DAS *et al.*, 2014). The proposed algorithm of GS-XGBoost is illustrated in Fig. 3.

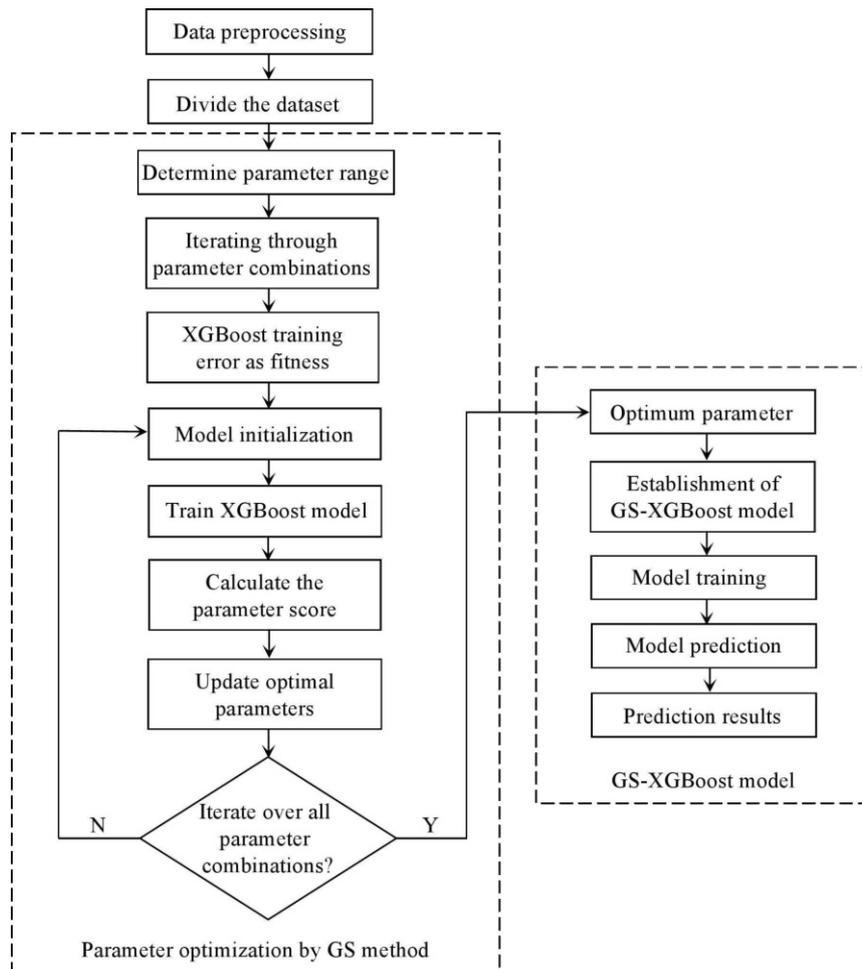


Fig. 3. Flow block diagram of GS-XGBoost algorithm.

Table 4. Structural parameter and range for GS-XGBoost.

Parameter	Value and range	Meaning of structural parameters
Learning_rate	(0.2, 0.3) with step of 0.01	Learning rate
Max_depth	(3, 10) with step of 1	Maximum depth of the tree
Subsample	(0.6, 0.9) with step of 0.01	Percentage of randomly selected samples
Colsample_bytree	(0.6, 0.9) with step of 0.01	Feature random sampling ratio
alpha	(0, 0.01)	$L_1$ canonical term parameters
lambda	(0, 1)	$L_2$ canonical term parameters

According to the above description, GS is an exhaustive method that involves exploring all possible combinations of structural parameters and then using these parameter sets for model prediction. Based on the characteristics of GS, the iteration time for optimizing XGBoost increases significantly with the number and range of parameters. Therefore, in this case, after going through the initial manual search, the parameters to be optimized and their respective ranges are identified. The structural parameters involved and their ranges are listed in Table 4.

### 3.3. Improved XGBoost based on PSO

Particle swarm optimization (PSO) is a population intelligence-based algorithm that optimizes objective function by simulating the collective behavior of birds (ZHANG *et al.*, 2016; MADVARI *et al.*, 2022). In principle, it is assumed that  $n$  particle populations in  $D$ -dimensional space are generated, the position of each particle is solved as  $X_i = (x_i^1, x_i^2, \dots, x_i^D)$ , and the velocity value of the  $i$ -th particle is  $V_i = (v_i^1, v_i^2, \dots, v_i^D)$ , the extreme values of individual and population are  $P_i = (p_i^1, p_i^2, \dots, p_i^D)$  and  $P_g = (p_g^1, p_g^2, \dots, p_g^D)$ , respectively. Then, according to the given fitness function, the fitness of each particle is calculated and the optimal values of individual and population are continuously updated. Finally, the remaining particles update their positions and velocities according to the current extreme values. The position and velocity of a particle are updated by the following equations:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k), \quad (1)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1}, \quad (2)$$

where  $\omega$  is the inertia factor;  $i = 1, 2, 3, \dots, n$ , and  $n$  is the number of particles;  $d = 1, 2, 3, \dots, D$ , and  $D$  is the dimension of the particle swarm;  $k$  is the number of

current iterations;  $V_{id}$  and  $X_{id}$  stand, respectively, for the velocity and position of the particle  $id$ ;  $P_{id}$  and  $P_{gd}$  are the optimal position experienced by particle  $id$  and the entire particle swarm;  $c_1$  and  $c_2$  represent learning factors; and  $r_1$  and  $r_2$  are random numbers distributed in the interval  $[0, 1]$ . The basic parameters of PSO are listed in Table 5.

Compared with the GS method, PSO can improve search efficiency through the collaborative and following behavior of the population, thereby reducing the time and space complexity of parameter search. And thus, for PSO, the influence of the number and range of parameters on optimization time is relatively small, mainly depending on the optimization objectives and constraints set. In this case, in order to construct the PSO-XGBoost regression model with high accuracy and strong generalization capability, the types and ranges of structural parameters are expanded and presented as in Table 6.

The proposed regression modeling process of PSO-XGBoost consists of three steps (see Fig. 4):

- Step 1: preprocess the sound quality data and divide it into training and testing sets, followed by determining the model parameters and their search ranges as shown in Table 6.
- Step 2: take the training error of XGBoost model as fitness, initialize the parameters of PSO shown in Table 5, and obtain the optimal fitness parameters in the solution space. Finally, the convergence condition is whether the maximum number of iterations or accuracy requirements is reached, and the model's best parameters are output when the condition is satisfied; otherwise the algorithm continues to perform the above optimization search steps.
- Step 3: establish a new PSO-XGBoost model based on the optimal parameters, train the model, and output its prediction results.

Table 5. Pre-selected parameters of PSO.

Parameter	Value	Meaning of structural parameters
$\omega$	0.5	Inertia factor
$c_1$	0.5	Learning factor
$c_2$	0.5	Learning factor
swarmsize	20	Number of populations
Max Stalliterations	50	Maximum number of iterations

Table 6. Structural parameter and range for PSO-XGBoost.

Parameter	Value range	Meaning of structural parameters
n_estimators	(50, 100)	Number of trees
Learning_rate	(0.001, 0.5)	Learning rate
Max_depth	(1, 5)	Maximum depth of the tree
Subsample	(0.1, 1)	Percentage of randomly selected samples
Colsample_bytree	(0, 0.8)	Feature random sampling ratio
alpha	(1, 10)	$L_1$ canonical term parameters
lambda	(0, 10)	$L_2$ canonical term parameters
gamma	(0, 15)	Minimum drop value of loss function required for node splitting
Max_delta_step	(0, 10)	Maximum step size for each tree weight change
Min_child_weight	(0, 10)	Sum of sample weights of minimum leaf nodes

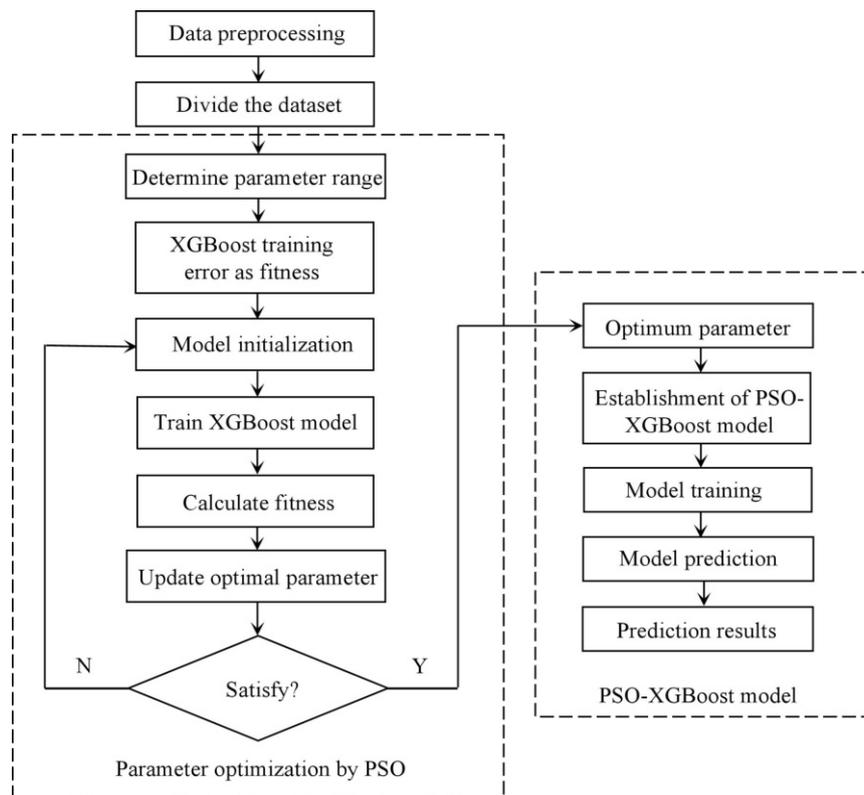


Fig. 4. Flow block diagram of PSO-XGBoost algorithm.

## 4. Acoustic comfort evaluation modeling and prediction

### 4.1. Optimum structural parameters

The purpose of this section is to develop a new model for electric bus acoustic comfort with high accuracy based on the above established sound quality database, where objective psycho-acoustic parameters and acoustic comfort are the input and output dependent variables of the model, respectively, corresponding to  $x_1$  to  $x_9$  in Tables 2 and 3. The database is divided into two data sets by noise samples: the first to 56th sample data for training and the 57th to 64th sample data for testing. The algorithms of standard

XGBoost, GS-XGBoost, and PSO-XGBoost were performed, respectively, on the MATLAB platform, with their corresponding model parameters and imported data were set. After running the codes, the optimal structural parameters of each regression model were obtained. Finally, the results are listed in Table 7.

### 4.2. Evaluation indicators of model accuracy

In order to evaluate and compare the algorithm performance of the above three regression models, average relative error (ARE), root mean square error (RMSE) and correlation coefficient ( $R^2$ ) are taken as the accuracy indexes to measure the models, and their calculation formulas are expressed further, where  $y_k$

Table 7. Optimized structural parameters for two regression models.

Regression model	Structural parameter variable	Optimal structural parameters
GS-XGBoost	Learning_rate	0.285
	Max_depth	3
	Subsample	0.9
	Colsample_bytree	0.8
	alpha	0
	lambda	1
PSO-XGBoost	n_estimators	71
	Learning_rate	0.499
	Max_depth	2
	Subsample	0.739
	Colsample_bytree	0.691
	lambda	7.259
	alpha	1.644
	gamma	0.181
	Max_delta_step	5.689
Min_child_weight	3.024	

and  $z_k$  represent the test and predicted values, respectively.

The smaller the values of ARE and RMSE, describing the degree of deviation between predicted and test values, the higher the model prediction accuracy, and they are defined by the equations:

$$e_{\text{ARE}} = \frac{1}{N} \sum_{k=1}^N \left| \frac{y_k - z_k}{y_k} \right| \times 100\%, \quad (3)$$

$$e_{\text{RMSE}} = \sqrt{\frac{\sum_{k=1}^N (z_k - y_k)^2}{N}}. \quad (4)$$

$R^2$  is a parameter measuring the closeness of two data groups, calculated by Eq. (5), and its higher value indicates the higher model's fitting accuracy:

$$R_{X,Y}^2 = \frac{E((X - EX)(Y - EY))}{\sqrt{D(X)}\sqrt{D(Y)}}, \quad (5)$$

where  $E$  represents the variable's mathematical expectation;  $D$  is the symbol for variance;  $E((X - EX)(Y - EY))$

is the covariance between the random variables  $X$  and  $Y$ .

#### 4.3. Acoustic comfort prediction results and comparison

Based on the testing set, the evaluation indexes of regression models trained by standard XGBoost, GS-XGBoost, and PSO-XGBoost were calculated. The predicted values and accuracy results are obtained and listed in Table 8.

Table 8 visually indicates that the prediction accuracy of standard XGBoost, GS-XGBoost, and PSO-XGBoost models all meet the application requirements of maximum relative error (MRE) and ARE less than 10 % and 5 %, respectively. For the standard XGBoost model, MRE and ARE are 8.53 % and 4.67 %, respectively. For the GS-XGBoost model, MRE is 8.83 %, 0.3 % higher than the standard XGBoost, and ARE is 3.64 %, 1.03 % lower than the standard XGBoost. For the PSO-XGBoost model, MRE and ARE are 6.33 % and 2.30 %, respectively, which are 2.2 % and 2.37 % lower than the standard XGBoost. In addition,

Table 8. Prediction results of three regression models.

Sample	Standard XGBoost (ZHANG <i>et al.</i> , 2023a)			GS-XGBoost		PSO-XGBoost	
	Target value	Predicted value	Relative error [%]	Predicted value	Relative error [%]	Predicted value	Relative error [%]
57	5.33	5.20	2.44	5.24	1.69	5.35	0.38
58	3.58	3.59	0.28	3.53	1.40	3.66	2.23
59	2.58	2.8	8.53	2.54	1.55	2.66	3.10
60	5.32	5.68	6.77	5.79	8.83	5.22	1.88
61	4.33	4.02	7.16	4.04	6.70	4.35	0.46
62	4.84	4.56	5.79	4.79	1.03	4.94	2.06
63	1.58	1.55	1.90	1.67	5.70	1.48	6.33
64	3.13	3.27	4.47	3.06	2.23	3.07	1.92

according to Eq. (4), the RMSEs of PSO-XGBoost and GS-XGBoost models correspond to 0.0843 and 0.2039, which are 0.1365 and 0.0169 smaller than 0.2208 of the standard XGBoost model, respectively. It can be concluded that the prediction accuracy of the three established models can be ranked from high to low as follows: PSO-XGBoost > GS-XGBoost > standard XGBoost.

In terms of model's fitting accuracy, three correlation coefficients of standard XGBoost, GS-XGBoost and PSO-XGBoost are calculated by Eq. (5) as 0.9848, 0.9881, and 0.998. Figures 5–7 show the fitting results of the above three models, in which solid dot is the sample data corresponding to test and predicted values, and orange solid and blue dotted lines represent the best fitting position and the error range of 10%, respectively. It can be seen that the predicted data point distribution of PSO-XGBoost model is closest to the best fitting line, followed by GS-XGBoost model, and the last is the standard XGBoost model, indicating that the three models perform consistently in terms of fitting accuracy with respect to prediction accuracy.

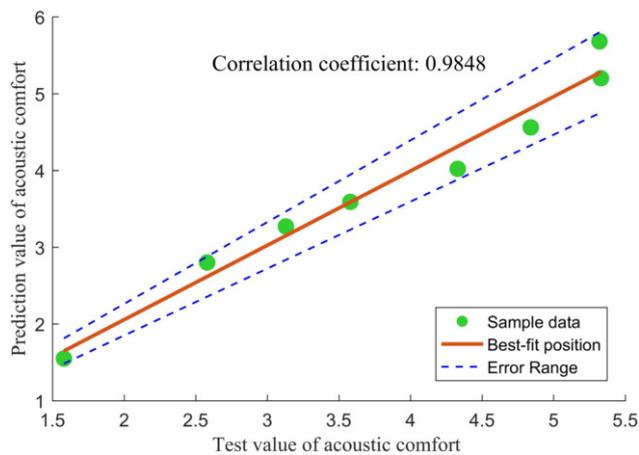


Fig. 5. Fitting result of standard XGBoost model.

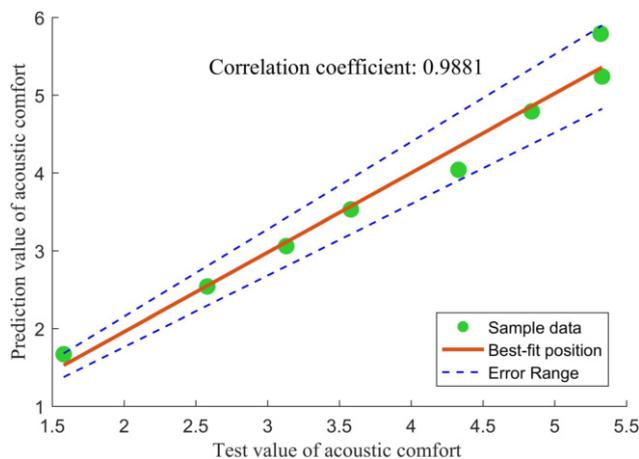


Fig. 6. Fitting result of GS-XGBoost model.

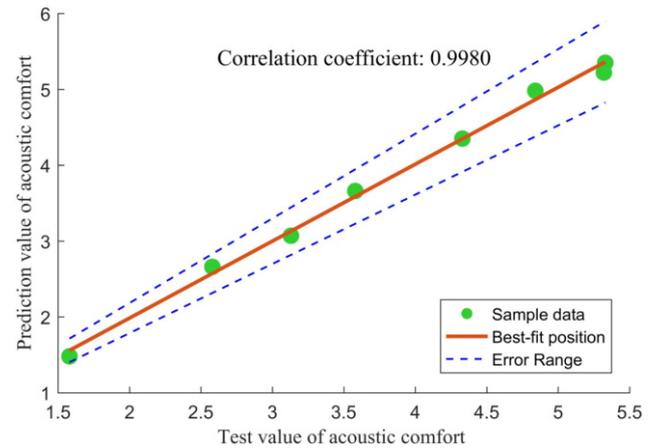


Fig. 7. Fitting result of PSO-XGBoost model.

To sum up, for the acoustic comfort modeling of electric bus, the prediction and fitting accuracy of GS-XGBoost model and PSO-XGBoost model are higher than that of standard XGBoost model, proving the improved XGBoost algorithms have obvious application effectiveness and superiority. Therefore, the established PSO-XGBoost model is finally adopted as the acoustic comfort evaluation model for this case, which provides a new technical support for research on electric bus interior sound quality.

## 5. Conclusions and future work

The electric bus is widely used and its interior sound quality research is an emerging field. Establishing sound quality prediction model aims to overcome the shortcomings of subjective evaluation tests with tedious processes and susceptible results. Based on the subjective and objective evaluation database, for improving model's accuracy and robustness, two intelligent algorithms, GS and PSO, were introduced to improve the XGBoost algorithm by optimizing model parameters. This led to establishing the acoustic comfort evaluation model for electric bus based on GS-XGBoost and PSO-XGBoost. Finally, ARE, RMSE, and  $R^2$  of the predicted results demonstrates that PSO-XGBoost model has the best prediction and fitting accuracy, followed by GS-XGBoost model, proving the effectiveness and applicability of the improved XGBoost algorithms.

Future research will primarily focus on two aspects: one is to collect noise signals in unsteady conditions, and further evaluate the generalization abilities of the improved XGBoost algorithms in characterizing dynamic sound quality; second will consider more observation points for improving sound quality data scale, and combine the proposed algorithms with semi-supervised learning to explore online modeling methods for electric bus interior sound quality.

## Acknowledgments

The authors would like to thank the Bus Engineering Research Institute of Xiamen King Long United Automotive Industry Co., Ltd for providing bus prototypes and signal acquisition instruments, and for organizing engineers to participate in subjective evaluation tests.

## Fundings

The work was supported by the National Natural Science Foundation of China (12004136), the Natural Science Foundation of Fujian Province (2023J011438, 2023J011437), the Key Research and Industrialization Project of Technological Innovation in Fujian Province (2023G013), the Major Educational Research Project of Fujian Province (FBJY20230154), and the Science and Technology Project for High-level Talents of XMUT (YKJ22017R).

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

- CHEN S.M., WANG D.F., LIANG J. (2012), Sound quality analysis and prediction of vehicle interior noise based on grey system theory, *Fluctuation and Noise Letters*, **11**(2): 1250016, doi: [10.1142/S0219477512500162](https://doi.org/10.1142/S0219477512500162).
- CHEN T.Q., GUESTRIN C. (2016), XGBoost: A scalable tree boosting system, [in:] *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- DAS G., PATTNAIK P.K., PADHY S.K. (2014), Artificial neural network trained by particle swarm optimization for non-linear channel equalization, *Expert Systems with Applications*, **41**(7): 3491–3496, doi: [10.1016/j.eswa.2013.10.053](https://doi.org/10.1016/j.eswa.2013.10.053).
- DING F., XIE W.T., XIE X.P. (2023), Research on optimization of car door closing sound quality based on the integration of structural simulation and test, [in:] *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **237**(6): 1378–1390, doi: [10.1177/09544070221088370](https://doi.org/10.1177/09544070221088370).
- DOLESCHAL F., VERHEY J.L. (2022), Pleasantness and magnitude of tonal content of electric vehicle interior sounds containing subharmonics, *Applied Acoustics*, **185**: 108442, doi: [10.1016/j.apacoust.2021.108442](https://doi.org/10.1016/j.apacoust.2021.108442).
- HUANG H.B., HUANG X.R., LI R.X., LIM T.C., DING W.P. (2016), Sound quality prediction of vehicle interior noise using deep belief networks, *Applied Acoustics*, **113**: 149–161, doi: [10.1016/j.apacoust.2016.06.021](https://doi.org/10.1016/j.apacoust.2016.06.021).
- HUANG H.B., WU J.H., LIM T.C., YANG M.L., DING W.P. (2021), Pure electric vehicle nonstationary interior sound quality prediction based on deep CNNs with an adaptable learning rate tree, *Mechanical System and Signal Processing*, **148**: 107170, doi: [10.1016/j.ymssp.2020.107170](https://doi.org/10.1016/j.ymssp.2020.107170).
- HUANG X.R., HUANG H.B., WU J.H., YANG M.L., DING W.P. (2020), Sound quality prediction and improving of vehicle interior noise based on deep convolutional neural networks, *Expert Systems with Applications*, **160**: 113657, doi: [10.1016/j.eswa.2020.113657](https://doi.org/10.1016/j.eswa.2020.113657).
- KIM D., LEE J. (2022), Predictive evaluation of spectrogram-based vehicle sound quality via data augmentation and explainable artificial Intelligence: Image color adjustment with brightness and contrast, *Mechanical Systems and Signal Processing*, **179**: 109363, doi: [10.1016/j.ymssp.2022.109363](https://doi.org/10.1016/j.ymssp.2022.109363).
- LIANG L.Y., CHEN S.M., LI P.R. (2020), The evaluation of vehicle interior impact noise inducing by speed bumps based on multi-features combination and support vector machine, *Applied Acoustics*, **163**: 107212, doi: [10.1016/j.apacoust.2020.107212](https://doi.org/10.1016/j.apacoust.2020.107212).
- MADVARI R.F., SHARAK M.N., TEHRANI M.J., ABBASI M. (2022), Estimation of metal foam microstructure parameters for maximum sound absorption coefficient in specified frequency band using particle swarm optimisation, *Archives of Acoustics*, **47**(1): 33–42, doi: [10.24425/aoa.2022.140730](https://doi.org/10.24425/aoa.2022.140730).
- POURSEIEDREZAEI M., LOGHMANI A., KESHMIRI M. (2021), Development of a sound quality evaluation model based on an optimal analytic wavelet transform and an artificial neural network, *Archives of Acoustics*, **46**(1): 55–65, doi: [10.24425/aoa.2021.136560](https://doi.org/10.24425/aoa.2021.136560).
- SHI W.K., LIU G.Z., SONG H.S., CHEN Z.Y., ZHANG B. (2018), Vibration and noise characteristics of electric bus [in Chinese], *Jilin Daxue Xuebao (Gongxueban)*, **48**(2): 373–379, doi: [10.13229/j.cnki.jdxbgxb20170110](https://doi.org/10.13229/j.cnki.jdxbgxb20170110).
- STEINBACH L., ALTINSOY M.E. (2019), Prediction of annoyance evaluations of electric vehicle noise by using artificial neural networks, *Applied Acoustics*, **145**: 149–158, doi: [10.1016/j.apacoust.2018.09.024](https://doi.org/10.1016/j.apacoust.2018.09.024).
- WANG C., PENG J.X., ZHANG X.W. (2020), A classification method related to respiratory disorder events based on acoustical analysis of snoring, *Archives of Acoustics*, **45**(1): 141–151, doi: [10.24425/aoa.2020.132490](https://doi.org/10.24425/aoa.2020.132490).
- WANG Y., ZHANG S., MENG D.J., ZHANG L.J. (2022), Nonlinear overall annoyance level modeling and interior sound quality prediction for pure electric vehicle with extreme gradient boosting algorithm, *Applied Acoustics*, **195**: 108857, doi: [10.1016/j.apacoust.2022.108857](https://doi.org/10.1016/j.apacoust.2022.108857).
- WANG Y.S. (2009), Sound quality estimation for non-stationary vehicle noises based on discrete wavelet transform, *Journal of Sound and Vibration*, **324**(3–5): 1124–1140, doi: [10.1016/j.jsv.2009.02.034](https://doi.org/10.1016/j.jsv.2009.02.034).
- ZHANG E.L., CHEN X.Y., LI S.Y., WU Q.Q., ZHUO J.M. (2022a), A comprehensive evaluation model of electric

- bus interior acoustic comfort and its application, *International Journal of Acoustics and Vibration*, **27**(4): 361–366, doi: [10.20855/ijav.2022.27.41887](https://doi.org/10.20855/ijav.2022.27.41887).
19. ZHANG E.L., CHEN Y., CHEN X.Y., ZHANG J.B., XUN P.W., ZHUO J.M. (2023a), High-precision modeling and prediction of acoustic comfort for electric bus based on BPNN and XGBoost, *International Journal of Acoustics and Vibration*, **28**(2): 158–164, doi: [10.20855/ijav.2023.28.21922](https://doi.org/10.20855/ijav.2023.28.21922).
  20. ZHANG E.L., HOU L., SHEN C. (2016), Sound quality prediction of vehicle interior noise and mathematical modeling using a back propagation neural network (BPNN) based on particle swarm optimization (PSO), *Measurement Science and Technology*, **27**(1): 015801, doi: [10.1088/0957-0233/27/1/015801](https://doi.org/10.1088/0957-0233/27/1/015801).
  21. ZHANG E.L., ZHANG Q.M., XIAO J.J., HOU L., GUO T. (2018), Acoustic comfort evaluation modeling and improvement test of a forklift based on rank score comparison and multiple linear regression, *Applied Acoustics*, **135**: 29–36, doi: [10.1016/j.apacoust.2018.01.026](https://doi.org/10.1016/j.apacoust.2018.01.026).
  22. ZHANG E.L., ZHUO J.M., HOU L., FU C.H., GUO T. (2021), Comprehensive annoyance modeling of forklift sound quality based on rank score comparison and multi-fuzzy analytic hierarchy process, *Applied Acoustics*, **173**: 107705, doi: [10.1016/j.apacoust.2020.107705](https://doi.org/10.1016/j.apacoust.2020.107705).
  23. ZHANG X.C., CHENG J., LU J.W., YUAN B., JIANG P., SHA W. (2022b), Sound quality evaluation of pure electric vehicle with subjective and objective unified evaluation method, *International Journal of Vehicle Design*, **88**(2–4): 283–303, doi: [10.1504/IJVD.2022.127024](https://doi.org/10.1504/IJVD.2022.127024).
  24. ZHANG Y., MENG T., WANG K. (2020), Subjective evaluation model for interior sound quality and optimization of medium-frequency noise, *Automotive Engineering*, **42**(5): 651–657+664, doi: [10.19562/j.chinasae.qcgc.2020.05.013](https://doi.org/10.19562/j.chinasae.qcgc.2020.05.013).
  25. ZHANG Y., PENG B.T., YANG E.C., REN K.L., OU J. (2023b), Prediction and analysis of vehicle interior sound quality based on XGBoost algorithm [in Chinese], *Noise and Vibration Control*, **43**(3): 161–166+211, doi: [10.3969/j.issn.1006-1355.2023.03.025](https://doi.org/10.3969/j.issn.1006-1355.2023.03.025).



## Research Paper

Virtual Reality Technology in Analysis  
of the Sarek National Park Soundscape in SwedenDorota MŁYNARCZYK\*<sup>id</sup>, Jerzy WICIAK<sup>id</sup>AGH University of Krakow  
Kraków, Poland\*Corresponding Author e-mail: [dorota.mlynarczyk@agh.edu.pl](mailto:dorota.mlynarczyk@agh.edu.pl)

(received July 21, 2023; accepted April 3, 2024; published online July 9, 2024)

The paper presents an in-depth analysis of the soundscape within Sarek National Park, the oldest national park in Europe, situated in Lapland, northern Sweden. The comprehensive acoustic measurements, ambisonic recordings, and 360° video recordings were carried out during a scientific expedition in the summer of 2020. The aim of the paper is to show the soundscape analysis of carefully selected characteristic locations in various parts of the valley. The paper extensively discusses the findings derived from the recorded data using both classical acoustic methods and the soundscape approach. The classic acoustic parameters, commonly employed in environmental acoustics as well as eco-acoustic indices such as: ACI (acoustic complexity index), ADI (acoustic diversity index), AEI (acoustic evenness index), NDSI (normalized difference soundscape index), BIO (energy level of biophony), amplitude index ( $M$ ), and total entropy ( $H$ ) were calculated. To gain further insights, listening tests, facilitated through virtual reality tools, were conducted, enabling participants to engage in soundwalk experiences. By employing a combination of traditional acoustic methods and innovative soundscape approaches, the paper presents a holistic evaluation of the auditory environment in Sarek National Park. The main contribution of the presented research is providing new data from the unique, geographically inaccessible region of the world, the Sarek National Park. This research not only enriches our understanding of the national park's soundscape but also offers valuable insights into the interaction between the natural environment and human perception of sound.

**Keywords:** virtual soundwalk; Lapland; UNESCO; natural soundscape; virtual reality technology.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Unique sounds should be treated as intangible heritage worthy of protection and preservation. The idea of the value of sounds found in the acoustic environment is closely related to the concept of soundscape. The soundscape method was first introduced in the work of authors such as SOUTHWORTH (1969), SCHAFER (1977), KRAUSE (2012), BROWN *et al.* (2015). The data collection methods and soundscape descriptors for the soundscape assessment have been introduced by ALETTA *et al.* (2016; 2019), KANG *et al.* (2016).

The perception of sounds in the environment is a complex creative process, involving the active perception, analysis, and interpretation of sensory phenom-

ena. In this process, incoming sensory signals are processed in a manner consistent with previously acquired experience. Some techniques to scale the perception of specific sounds are described by BERGLUND and NILSSON (2006), DAVIES *et al.* (2013), DE COENSEL and BOTTELDOOREN (2006). A principal component model of the soundscape perception is portrayed by AXELSSON *et al.* (2010). An extensive set of 116 soundscape attributes were well integrated into three basic components or dimensions of the soundscape perception: pleasantness, eventfulness, and familiarity.

In 2008, the Working Group at the International Organization for Standardization [ISO] was established and was named “Perceptual assessment of soundscape quality” (MITCHELL *et al.*, 2022). The results of the work of this team have been published in the form

of standards within the ISO 12913 series on soundscape. Part 1 is a full standard and provides a general framework and definitions of soundscape concepts (ISO, 2014), while part 2 and part 3 are technical specifications and offer guidance on how data should be collected and analysed, accordingly (ISO, 2018; 2019).

Psychoacoustic research on the soundscape is difficult and complex. It is difficult to determine the values of variables that will define not only acoustic preferences, but also the mood and emotional state of an individual person. Different subcultures can be identified in each area, which are further influenced by individual preferences (FARINA, 2014). A soundscape only defined as friendly to some people may be neutral or annoying to others. The determination of what is a friendly and beneficial soundscape and what is not, is also determined by the level of social acceptance (JIANG *et al.*, 2022). The following eco-acoustic indices are currently used: acoustic complexity index (ACI), acoustic diversity index (ADI), acoustic evenness index (AEI), normalized difference soundscape index (NDSI), energy level of biophony (BIO), amplitude index ( $M$ ), and total entropy ( $H$ ) (BRADFER-LAWRENCE *et al.*, 2019; FLOWERS *et al.*, 2021; PIERETTI *et al.*, 2011). These indices are quantitative measures used in environmental research to assess acoustic properties of ecosystems. The calculation of these indices is based on the analysis of the acoustic energy distribution in soundscape recordings. Through monitoring and such analysis, biodiversity patterns, habitat and ecosystem health, and ecological changes can be studied.

The ACI quantifies sound complexity by assessing the variability of intensities between time samples within a specific frequency band. The higher ACI values the higher level of bird or insect activity. The ADI is calculated by applying the Shannon diversity index to the relative proportion of signals occurring in each 1 kHz frequency band. The higher ADI values the greater evenness soundscape and number of occupied frequency. The AEI assesses the balance of sound sources by estimating the Gini coefficient based on the signal proportion in each 1 kHz band. The lower AEI values the less saturated soundscape. The NDSI calculates the ratio of signal power in the frequency bands between 1 kHz–2 kHz (anthrophony) and 2 kHz–8 kHz (biophony). The higher NDSI values the higher level of biophonic activity. The BIO quantifies the signal power specifically within the 2 kHz–8 kHz frequency band, representing the biophony. The higher BIO values the higher levels of biophonic activity in the soundscape. The amplitude index ( $M$ ) is a metric that assesses the variations in signal amplitude within a designated frequency band, providing a measure of the intensity or strength of acoustic signals present in a given environment. The higher  $M$  value, the greater the amplitude variations within the specified frequency band. Total entropy ( $H$ ) quantifies the overall unpredictability and

complexity of acoustic signals in a given system, reflecting the diversity across different frequency bands. The higher  $H$  value, the greater the unpredictability and complexity of acoustic signals, indicating a more diverse sound environment.

The dynamic development of civilisation and the huge number of anthropogenic noise sources in the environment makes silence of particular value. Silence, understood as the audibility of the sounds of nature, is becoming a much sought after value. Numerous scientific studies show that recording acoustic environments captures their phonic richness and the unique sound features of the environment (DE COENSEL, BOTTELDOOREN, 2006; SCARPELLI *et al.*, 2021; BERNAT, 2013; RYCHTÁRIKOVÁ, VERMID; 2013; CZOPEK *et al.*, 2019; MAŁECKI *et al.*, 2020; BORKOWSKI *et al.*, 2021).

This paper presents the results of measurements and analysis of survey data recorded in Sarek National Park, in Lapland in Sweden during a 10-day research expedition (CZOPEK *et al.*, 2022).

The analysis was carried out using the classical method and the soundscape method. Basic classical acoustic parameters used in environmental acoustics (the  $A$ -weighted, equivalent continuous sound level, spectrograms, spectra) were calculated. Also, basic soundscape attributes “pleasant”, “eventful”, “annoying”, “calm”, etc., were used to assess the sound environment. The virtual reality (VR) technology was used to perform listening tests. Furthermore, eco-acoustic indices including: ACI, ADI, AEI, NDSI, BIO,  $M$ , and  $H$  were employed. These research methods facilitated the comparison of multi-temporal acoustic patterns across various sections of the valley, as well as their correlation with the subjective evaluations provided by the listeners.

## 2. Lapland and the Sarek National Park

In present times, Lapland (also named Sápmi) spans across the northern parts of the Scandinavian Peninsula and the Kola Peninsula. This ethno-cultural region boasts vast expanses of land. It covers approximately 390 thousand square kilometres across four countries: Norway, Sweden, Finland, and Russia. The indigenous people living in the Lapland region are the Saami, whose population is estimated around 60 000–90 000. At present, they are a minority among the inhabitants of Lapland – approximately 2.3 million people. Most of the Lapland’s territory is situated north of the Arctic Circle. Its western region encompasses fjords, deep valleys, glaciers, and mountains, including the highest peak, Kebnekaise (2111 m) located in Swedish Lapland. The Swedish part of Lapland is distinguished by the presence of major rivers that flow from the northwest to the southeast.

Sarek National Park was established in 1909. It is located in northern Sweden about one hundred kilo-

metres beyond the Arctic Circle and covers an area of 1970 km<sup>2</sup>. Sarek National Park is approximately circular in shape with a diameter of approximately 50 km. It is adjacent to two other national parks, namely Stora Sjöfallet and Padjelanta National Park. In December 1996, the adjacent national parks of Sarek, Stora Sjöfallet (1278 km<sup>2</sup>), and Padjelanta (1984 km<sup>2</sup>) were inscribed on the UNESCO World Heritage List collectively as the Laponian Area. It met the five demanding criteria (iii, v, vii, viii, ix) for Outstanding Universal Values. The Laponian Area, located in northernmost Sweden, is a magnificent wilderness of high mountains, primeval forests, vast marshes, beautiful lakes, and well-preserved river systems. It contains areas of exceptional beauty such as the snow-covered mountains of Sarek, the large alpine lakes of Padjelanta, and the extensive river delta in the Rapa Valley. Also, the Laponian Area is an outstanding example of traditional land-use, a cultural landscape reflecting the ancestral way of life of the Saami people based around the seasonal herding of reindeer (UNESCO, n.d.).

Sarek National Park is distinguished by its breathtaking alpine scenery, featuring majestic mountain ranges, narrow valleys, glaciers, swift rivers with abundant rapids and waterfalls. Six of Sweden's thirteen highest mountains, nearly 100 glaciers, and long, deep, and narrow valleys are located in the park. It is renowned for its diverse wildlife, including large elk and numerous predators. The meltwater from the numerous glaciers feeds the main Sarek river, the Ráhpaaädnö, forming the most renowned delta in the alpine world. However, the park does not boast a wide variety of plant species. This is primarily attributed to the fact that most of the park lies above the tree line, approximately 500 m in altitude, limiting the presence of coniferous vegetation. Within the park, there are no

tourist facilities, marked trails, or shelters for overnight stays. The landscape of the area bears the unmistakable imprint of the ancient ice sheet, and traversing Sarek is akin to embarking on a journey through Sweden's geological past. In Sarek National Park, many places have names derived from Sami languages. The most common Sami names for locations or objects in the park are *tjåkkå* or *tjåkko* (mountain), *vagge* (valley), *jåkkå* or *jåkko* (stream), *lako* (plateau), and *ätno* (river), e.g., Rapaätno means Rapa River.

Sarek is a challenging measurement and research environment. The terrain and the weather can be very hard and changing. Measurements and recording of soundscapes were carried out in 8 places located between: Aktse and Skárjá. Among which, the soundscapes captured in three places of the Sarek National Park: Tjasskávárvasj, by the Rådnik, Skárjá were analysed further (Figs. 1 and 2). At each point, 30-minute ambisonic and 360° video recordings were made and, except for Akste, where a 10-day sound recording was carried out.

Tjasskávárvasj – the location on the left bank of the Rapadalen Valley, near a section of the Kungsleden (King's Trail) leading to the summit of Skierffe, famous for its magnificent view of the Valley of the Rapaätno River and its picturesque delta. The measurement site is already within the boundaries of Sarek National Park, shortly after branching off from the Kungsleden Trail and heading up the Rapaätno River.

By the Rådnik – the location on the left bank of the Rapadalen Valley, northwest of the Rådnik peak. From the slope, there is a view extending over the valley.

Skárjá (Smaila Moot) – the central location of the park, was previously used as a pasture by the Saami. Now, there is a small cabin equipped with an emergency phone. The cabin is not open to hikers but

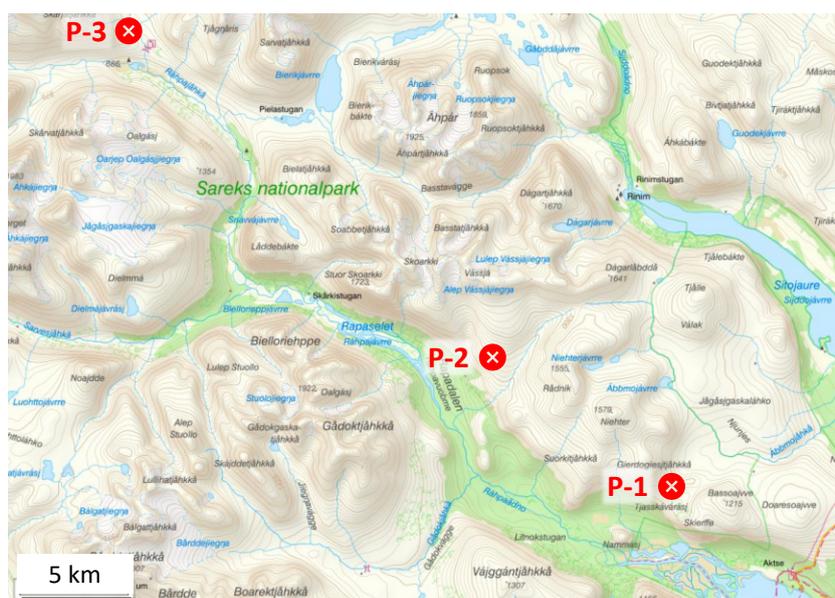


Fig. 1. Location of measuring points in Sarek National Park: P-1 Tjasskávárvasj, P-2 by the Rådnik, P-3 Skárjá.



Fig. 2. Sarek National Park landscape measuring points: a) Tjasskávárvasj; b) by the Rådnik; c) Skárjá.

can be used as an emergency shelter. It is a place where many hikers can be found. There is also a bridge over the Smailajákk canyon which allows hikers to cross the stream. The bridge is removed every winter and put back in the spring, after the spring flood.

There were minimal anthropogenic sounds in the soundscapes used to recreate and simulate the natural environment. At the initial location, Tjasskávárvasj, the primary sound sources are birds with quite high frequencies. There is also sound of wind and streams. At the site named “by the Rådnik” the wind is the predominant sound source. The soundscape of Skárjá is the only one with anthropophony. Here, the main source of sound emanates from the water, accompanied by the whispers of the wind, the gentle rustling of grass, and occasional sounds of human origin. In this rich audio tapestry, one can also discern the singing birds.

### 3. Virtual reality method and measurement

The analysis of the recorded soundscapes was carried out using the method of recording the acoustic environment and later recreating it in a laboratory with a sound system and using the VR technology (LAVALLE, 2019; WANG, 2020). Virtual reality is created with the use of hardware and software. After donning the goggles and headphones, our senses are cut off from the outside world and we experience the phenomenon of immersion, i.e., full immersion in a virtual world. This allows us to experience simulations of unusual places, objects or activities. Environmental recordings were made using an ambisonic microphone, 360° camera and sound analyser SVAN957. Laboratory tests were carried out using a VR set – Oculus Quest 2. The realisation of a suitable VR project requires the spatial image recording and the sound recording made with an ambisonic microphone to be converted and synchronised due to the different formats of the recorded data. The obtained files were rendered and then submitted to a final conversion in Spatial Media Metadata Injector, resulting in 360° videos with spatialised sound. The next stage of the work was to calibrate the playback level so that it was possible to play the films on the Oculus Quest 2 at the same sound level as in real life.

Calibration was carried out in an anechoic chamber with the use of the HATS (Head and Torso Simulator) simulator with built-in microphones in the ears by Bruel&Kjaer 1/2" type 4189 and the DAW – Pro Tools software (Fig. 3).

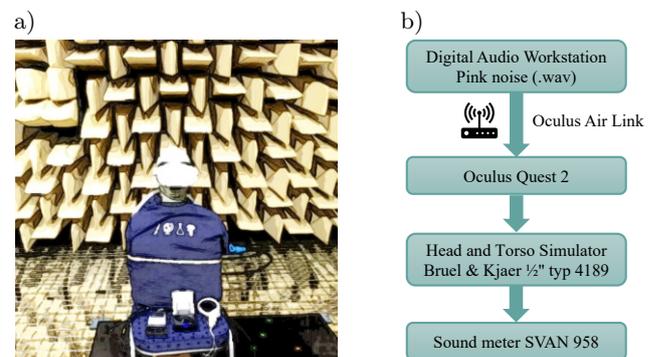


Fig. 3. Calibration of the Oculus's sound level for listening tests: a) photograph of the equipment in an anechoic chamber; b) block diagram of the calibration process.

The sound level was similar – around 30 dBA – for the by the Rådnik and Tjasskávárvasj soundscapes. For the third soundscape (Skárjá) used in simulations, the sound pressure level was over 50 dBA.

The qualitative evaluation of the soundscape was carried out with the use of questionnaires (ISO, 2018). The questions concerned the issues of loudness assessment, general impression, adjusting the sound to the environment, describing emotions experienced

at a given moment, but also specifying the existing sound sources. There were questions verifying the context and considering the extent to which the studied environment is perceived as monotonous, dynamic, varied, unvaried, irritating, pleasant, calm or chaotic. The survey included basic personal information and seven questions. The Likert scale was used for all survey questions. Numerical values from 1 to 5 were assigned to the responses increasing with the nature of the characteristic under investigation.

Psychoacoustic research was carried out on a group of 30 people using a questionnaire that was completed by each participant after simulating the soundscape in the auralization laboratory. The study participants included 13 men and 17 women. The age of the respondents ranged from 21 to 23 years old, the largest part of which were 22-year-old. Each examined person completed the questionnaire three times.

The research was carried out in the auralization laboratory of the Department of Mechanics and Vibroacoustics of the AGH University of Krakow. The examined person sat on a swivel seat that allowed him to turn freely during the simulation. Then, the test participant put on the VR goggles – Oculus Quest 2 and received one of the two controllers compatible with the VR headset (Fig. 4).



Fig. 4. Tested person wearing VR goggles during an auditory test at the Auralization Laboratory at AGH University of Krakow.

The controller allowed each person to set individual characteristics for each played file. Navigating through cyberspace and switching on subsequent films was also made possible by the controller. Playback levels were calibrated individually for each participant and recording, so that each test participant perceived the same sound level as was actually present in the location.

The examined person was presented three spherical films with spatial sound depicting three different sound landscapes. Each simulation lasted three minutes and was completed with a survey. The movies were played in the following order: Tjasskávárvasj, by the Rådnik, Skárjá the same sound level that was actually present in the given place.

#### 4. Results

The following four figures show the results of the assessment of the surrounding environment. Figure 5 summarises the comparative characteristics of the variation in responses regarding the overall assessment of the studied acoustic environments.

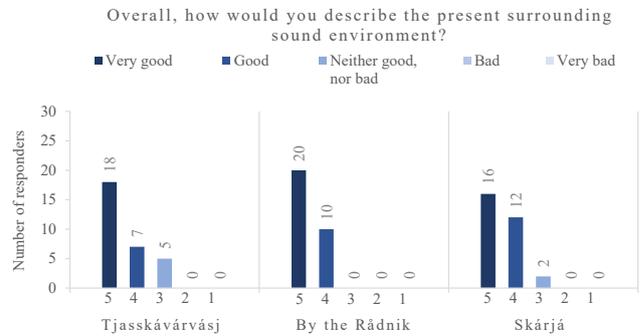


Fig. 5. Distribution of votes for the assessment of the surrounding sound environment of the tested sites.

It is noticeable that none of the soundscapes were generally perceived as bad or very bad by the survey participants. The significant majority of respondents rated the soundscapes presented as good or very good. A very small number of participants described the perception of the acoustic environment occurring at the sites: Tjasskávárvasj and Skárjá as neither good nor bad. The median values calculated for all those three places are the same – very good. Comparing the distribution of votes for Tjasskávárvasj and Skárjá, it can be seen that there is very little difference between the overall perception of the studied soundscapes. It is noteworthy that the respondents rated the soundscape of Rådnik as the best, resulting in the highest number of votes for a very good assessment and no votes for a neither good, nor bad, bad or very bad. In contrast, the acoustic environment recorded at the Tjasskávárvasj site was generally received as the worst.

Distribution of responses assessing appropriateness of the surrounding sound environment of the tested sites is shown in Fig. 6.

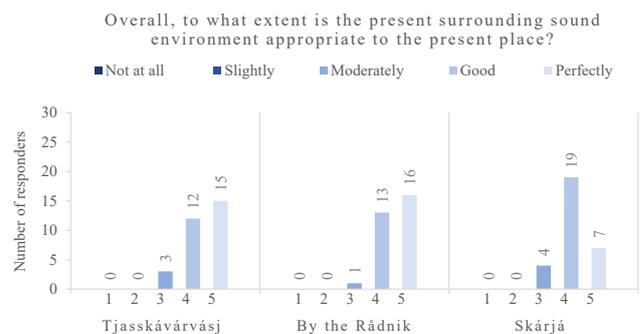


Fig. 6. Distribution of votes assessing appropriateness of the surrounding sound environment of the tested sites.

None of the sites analysed were rated by respondents as matching the surroundings slightly or not at all. A small number of respondents rated the sound fit at each analysed site as moderate.

The significant majority described the fit of the soundscapes as very good or ideal. The best match between sound and surroundings with median value equals 5.0 was recorded for the soundscape of by the Rådник. The soundscape of Tjasskávárvasj received a slightly lower rating, with a median of 4.5. The acoustic environment documented for Skárjá was the worst in this comparison and had a median value equal 4.0.

The distribution of votes regarding the loudness of a site is shown in Fig. 7.

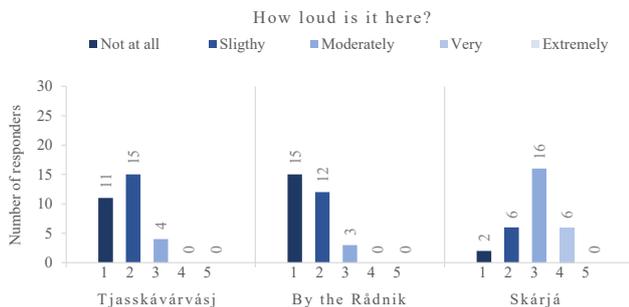


Fig. 7. Distribution of votes assessing a loudness of the sound environment.

None of the three analysed soundscapes were rated as extremely noisy. Most people described the places surveyed as slightly or moderately noisy or quiet. Out of those surveyed, only the soundscape of Skárjá was rated as noisy by six respondents.

When comparing the median values of the loudness ratings of the surveyed soundscapes, it is clearly noticeable that Skárjá's soundscape was described as the loudest. The median value is 3.0, indicating that the soundscape was considered moderately loud. The other two landscapes were rated very similarly – as not at all or slightly noisy. The median value for Tjasskávárvasj is 2.0 and for by the Rådnik is 1.5.

The distribution of votes regarding willingness to return to the surveyed places is shown in Fig. 8.

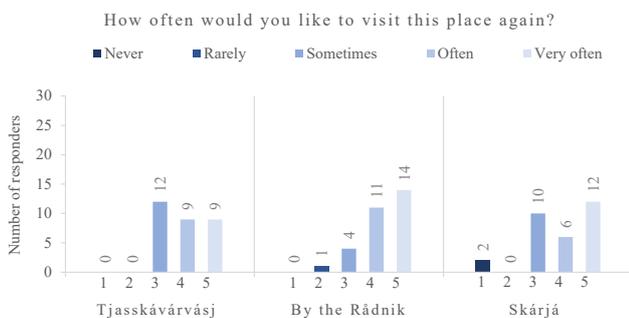


Fig. 8. Distribution of votes assessing a willingness to return to the surveyed places.

The great majority of respondents would like to be able to visit the presented places again. The median for all locations is 4.0, indicating that respondents would like to visit all places with similar frequency, and often. Only two respondents would not like to visit the Skárjá landscape again. Analysing the distribution of votes, it is noticeable that by the Rådnik was considered by far the most visitable place.

Figures 9 and 10 present selected results of the survey related to the eight perceptual attributes of the acoustic environment of the three tested places.

The analysis of the results presented in Fig. 9 shows that none of the respondents described the studied acoustic environments as definitely unpleasant. The Tjasskávárvasj and by the Rådnik soundscapes were also not rated as rather unpleasant. Only three of the study participants characterized Skárjá's soundscape as rather unpleasant, and two as moderate unpleasant. The second landscape by the Rådnik was definitely the most pleasantly perceived, the median of which was 5.0. For the other two locations – the median was 4.0. Generally, all soundscapes were found to be pleasant.

The next question of the survey questionnaire examined the extent to which the analysed soundscape is perceived as annoying. The soundscape of by the Rådnik was rated by all 30 respondents as definitely non-annoying. None of the three locations was identified as being clearly annoying. The third soundscape Skárjá was found to be the most annoying place. The median for Tjasskávárvasj soundscape is 2.0 and for by the Rådnik and Skárjá is 1.0, which means that the respondents defined a small or very small degree of compliance with the examined statement.

As in the previous question, all of 30 respondents described the by the Rådnik soundscape as calm. Only one person rated the last soundscape Skárjá as definitely restless. The median of compliance was the highest for the second soundscape – 5.0, followed by the first – 4.5 and third – 3.0 – soundscapes.

None of the soundscapes was considered chaotic by the respondents. Twenty-six of respondents indicated the soundscape of by the Rådnik as the least chaotic. The median was the highest for the third soundscape, which of the three was considered to be the most consistent with the examined statement.

Figure 10 shows respondents' answers regarding to what extent the surveyed acoustic environments are perceived as monotonous, vibrant, uneventful and eventful. None of the landscapes were rated as definitely monotonous, with the exception of only three people describing by the Rådnik's acoustic environment in this way. The median was the highest for this landscape. On the other hand, the soundscape of Skárjá was considered to be the least monotonous.

The most vibrant soundscape was found to be the last, third soundscape, with a median of 3.0. By the Rådnik's soundscape was rated as the least vi-

For each of the 4 scale below, to what extent do you agree or disagree that the present surrounding sound environment is ... [number of responders].

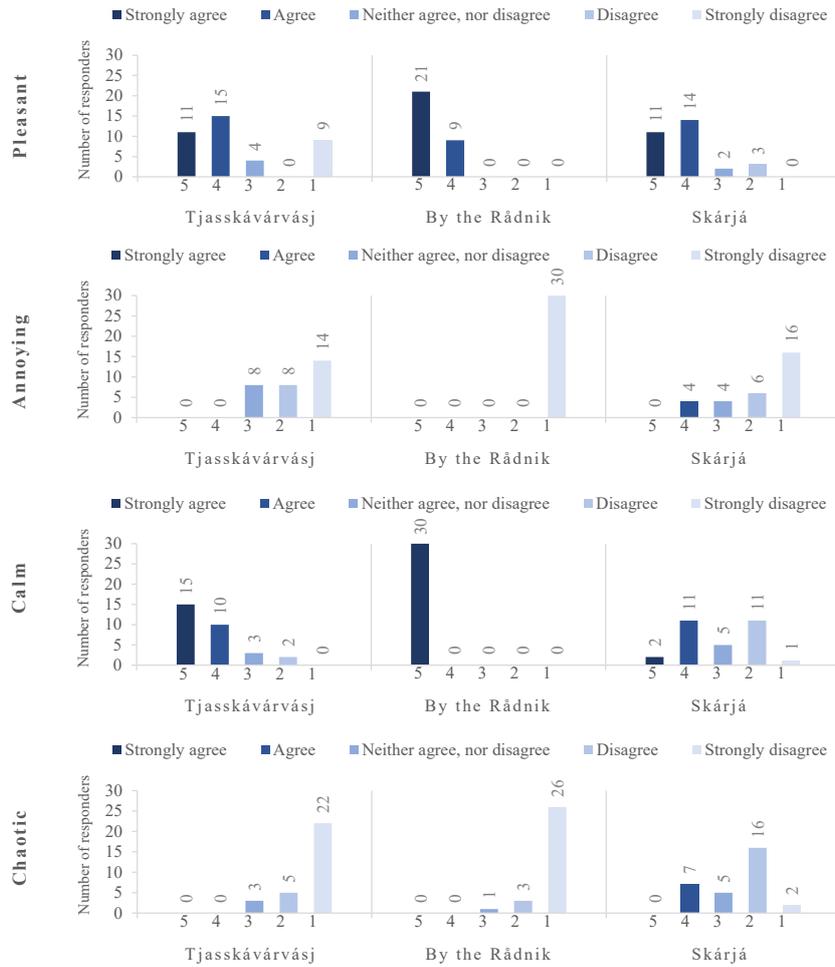


Fig. 9. Distribution of votes regarding the perception of the surveyed place as pleasant, annoying, calm, and chaotic.

brant – median is 1.0. The highest median – 4.0 – was achieved by the soundscape of by the Rådrik receiving the title of the most uneventful soundscape. None of the sites surveyed were identified as definitely eventful. Feedback from participants identified the third landscape as the most eventful.

Results shown in Figs. 9 and 10 could be reported in a two-dimensional scatter plot with coordinates for the two dimensions: pleasantness and eventfulness (Fig. 11). In order to calculate the position of the individual soundscapes, values from 5 (strongly agree) to 1 (strongly disagree) are taken for the individual perception assessments (ISO, 2019). The coordinates for the pleasantness axis –  $P$  – are then obtained by substituting the mean values for pleasant ( $p$ ), annoying ( $a$ ), calm ( $ca$ ), chaotic ( $ch$ ), vibrant ( $v$ ), and monotonous ( $m$ ) into Eq. (1). The coordinates for the eventfulness axis –  $E$  – are obtained by substituting the mean values for eventful ( $e$ ), uneventful ( $u$ ), chaotic ( $ch$ ), calm ( $ca$ ), vibrant ( $v$ ), and monotonous ( $m$ ) into Eq. (2):

$$P = (p - a) + \cos 45^\circ \cdot (ca - ch) + \cos 45^\circ \cdot (v - m), \quad (1)$$

$$E = (e - u) + \cos 45^\circ \cdot (ch - ca) + \cos 45^\circ \cdot (v - m). \quad (2)$$

Comparing results is much easier if the range of values is  $\pm 1$ . Therefore, the resulting coordinates can be normalised by dividing by  $(4 + \sqrt{32})$ .

All tested acoustic environments were described as pleasant. The soundscape of by the Rådrik turned out to be the most calm and uneventful of all. The reason for this could be the small number of sound sources present. The sounds of wind and grass were predominant in the soundscape. The singing of the birds was inaudible when listening with a VR system.

The soundscape of Skárjá, which is much more eventful and vibrant due to the presence of more varied sounds, is different from the rest. In the acoustic environment, not only sounds of natural origin, but also those of human origin appear. The more sound sources, the more vibrant and eventful the place is.

In Fig. 12, the comparison of National Park Sarek's soundscape with a sample of London's soundscape is presented (MITCHELL *et al.*, 2021). The values obtained in the park closely resemble those obtained

For each of the 4 scale below, to what extent do you agree or disagree that the present surrounding sound environment is ... [number of responders].

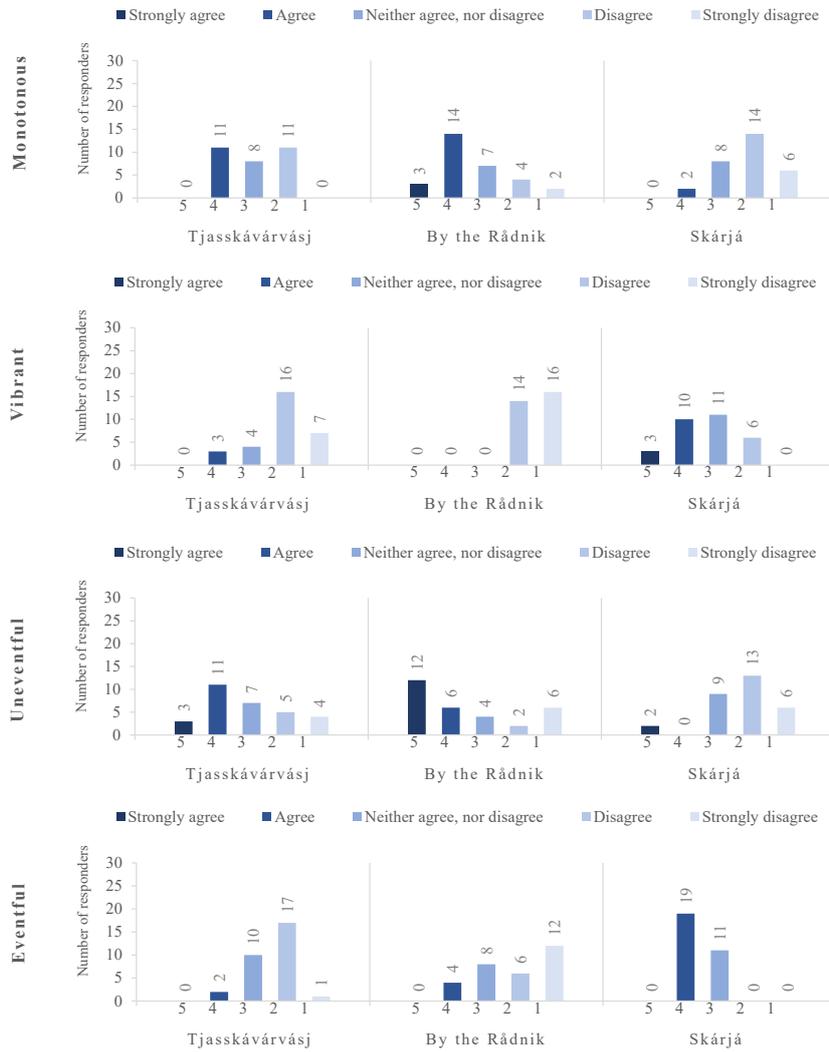


Fig. 10. Distribution of votes regarding the perception of the surveyed place as monotonous, vibrant, uneventful, and eventful.

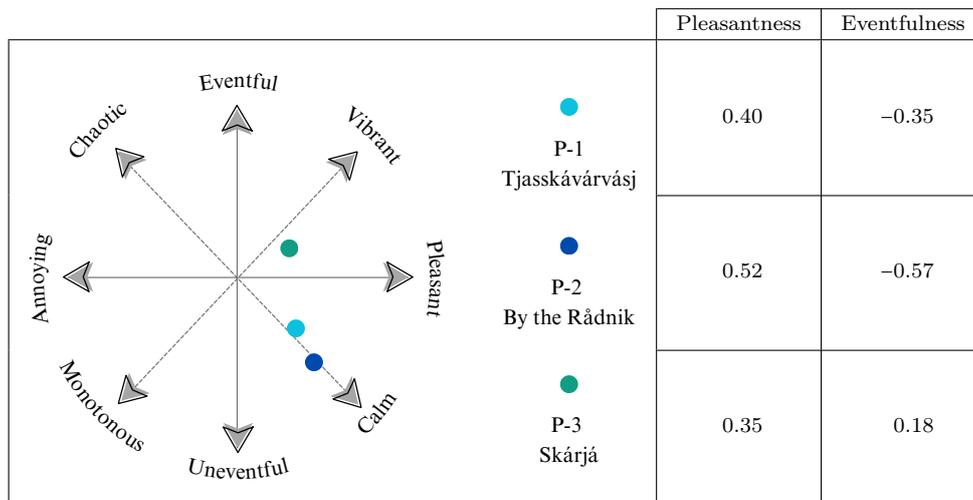


Fig. 11. Graphic classification of pleasantness and eventfulness of Sarek’s soundscapes calculated using ISO Eqs. (1) and (2). The values have been normalised (by dividing the coordinates by  $(4 + \sqrt{32})$ ).

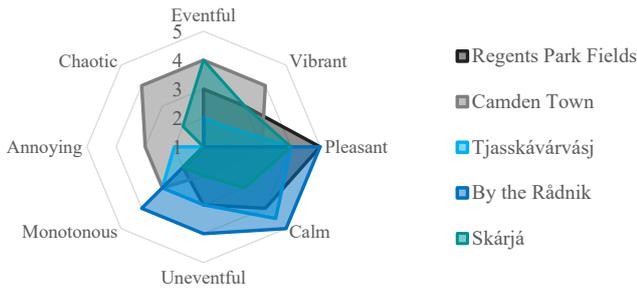


Fig. 12. Comparison of Sarek’s soundscapes (Tjasskávárvasj, by the Rådник, Skárjá) with soundscape of London park (Regents Park Fields) and square (Camden Town) on a radar plot of median value of perceptual attribute ratings on the Likert scales from 5 to 1 (MITCHELL *et al.*, 2021).

in National Park Sarek. However, it is essential to note that different research groups participated in the studies, and they were not conducted simultaneously. The assessment of Camden Town’s soundscape is significantly shifted towards the Eventful coordinate axis. This location received relatively high ratings for chaotic, eventful, and vibrant.

Skárjá’s landscape was louder than the other two landscapes (Fig. 13), which may have made it perceived as less pleasant. In Skarja’s landscape, anthropogenic noises such as tourists’ footsteps and conversations disrupt the natural harmony of the environment, leading to a more negative perception and a decreased desire to return frequently. All studied landscapes are generally classified as pleasant, calm and not very eventful.

The *A*-weighted, equivalent continuous sound levels, spectra and spectrograms – classical acoustic parameters commonly used in environmental acoustics – were used to objectively characterise the soundscapes studied. Figure 13 shows the average spectra (left) and spectrograms (right) measured at the study sites.

In Fig. 13a, the bi-modal energy distribution in the frequency domain at Tjasskávárvasj can be observed. The highest SPL values are found at frequencies 160 Hz and 2000 Hz. The equivalent *A*-weighted sound pressure level was 32 dB. At by the Rådник acoustic energy is accumulated in the low-frequency range, here there is a rapid decrease in value from 42 dB to 12 dB, in the frequency band from 50 Hz

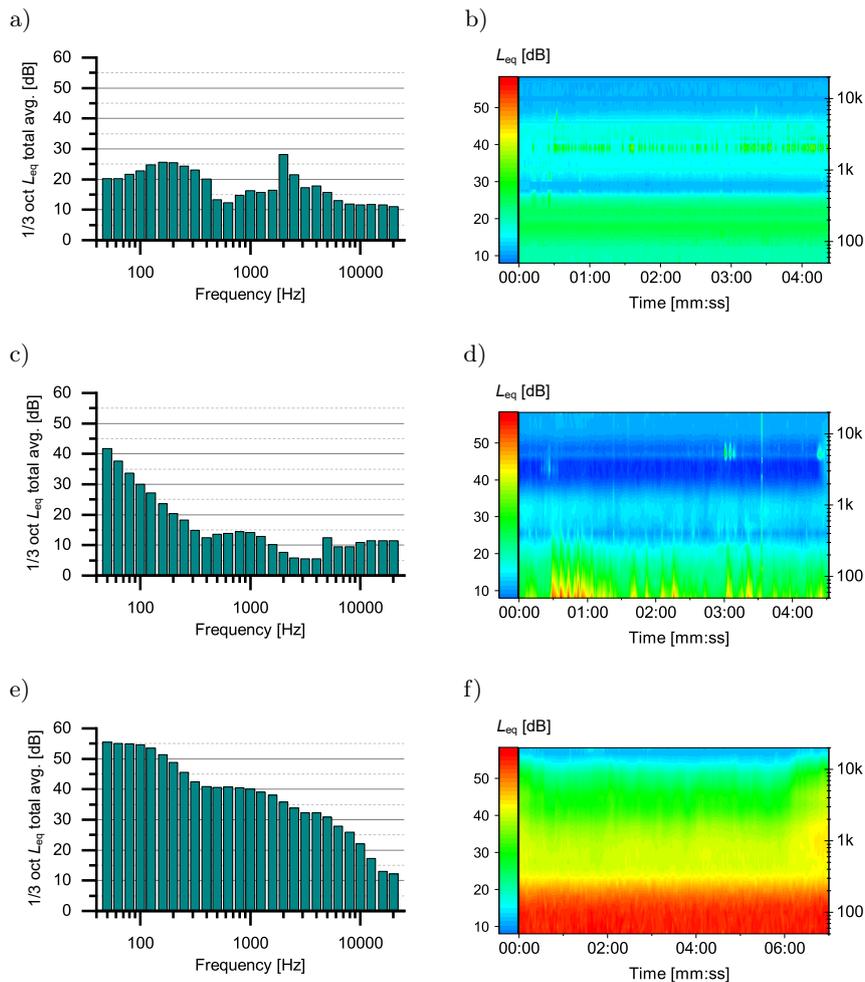


Fig. 13. The Tjasskávárvasj’s ((a) and (b)), by the Rådник’s ((c) and (d)) and Skárjá’s ((e) and (f)) sound level: average spectrum (on the left) and spectrogram (on the right).

Table 1. Basic eco-acoustic indices calculated for 300 s recording window.

Indicator	Place		
	Tjasskávárvasj	By the Rådник	Skárjá
	Index value		
Acoustic complexity index (ACI)	1785.92	1823.25	1794.31
Acoustic diversity index (ADI)	0.85	0.003	0.78
Acoustic evenness index (AEI)	0.81	0.90	0.82
Normalized difference soundscape index (NDSI)	0.69	0.11	0.51
Energy level of biophony (BIO)	3.39	1.53	4.76
Amplitude index ( $M$ )	0.042	0.016	0.023
Total entropy ( $H$ )	0.75	0.35	0.66

to 400 Hz (Fig. 13c). In the remaining frequency range 500 Hz–20 000 Hz,  $1/3$  octave SPL values oscillate around 7 dB–12 dB. The equivalent  $A$ -weighted sound pressure level was 24 dB. Figure 13e shows a broadband distribution of acoustic energy at Skárjá, with a drop in  $1/3$  octave SPL values from 70 dB to approximately 14 dB as the frequency increases. The equivalent  $A$ -weighted sound pressure level was 50 dB.

For comparison purposes, calculations of basic eco-acoustic indices were also carried out. The calculation results are presented in Table 1.

The calculated values of the indicators confirm that the acoustic activity is not high at the presented measurement points. High NDSI values, around 0.8, at Skárjá and Tjasskávárvasj indicate that signals do not contain anthrophony.

Slightly higher values of ACI and BIO factors (4.76 in Skárjá and 3.39 in Tjasskávárvasj) indicate a slightly higher activity of birds comparing to by the Rådник. Slightly higher values of the ADI index (Skárjá, Tjasskávárvasj) indicate the occurrence of higher wind comparing to by the Rådник. Also, low ADI values (by the Rådник) show that the soundscape is not containing many vocalizing species.

Higher values of the NDSI coefficient (Tjasskávárvasj) indicate greater biophonic activity and minimal anthrophonic noise in 1 kHz–2 kHz and indicate higher levels of biophonic activity in the soundscape. Similar values of the AVE coefficient show that all three soundscapes are similarly saturated.

## 5. Summary

The conducted research allowed for the mapping of three unique soundscapes in the VR technology. The VR technology allows for a very realistic representation of any acoustic environment. In total, approximately 160 hours of audio recordings were gathered, including 100 hours of ambisonic audio and 360° video recordings in Sarek National Park. Sound level measurements accompanied all the recordings. Modern VR devices can significantly facilitate the organization of soundscape research, or even enable it. The executed recordings facilitate subsequent soundscape research in

inaccessible regions such as Sarek National Park. This affords the opportunity to conduct comparative studies and investigate the influence of natural soundscapes on humans' wellbeing within laboratory environments, thereby opening new avenues for research.

In the conducted study, respondents generally classified the three tested soundscapes as pleasant and calm. This was confirmed by both the analysis of the voice distribution and the median on the radar chart. Graphic classification of pleasantness and eventfulness of Sarek's soundscapes shows that all three soundscapes surveyed were classified as pleasant and peaceful. However, the louder and more varied the soundscape (e.g., presence of pulse biophones) the more annoying and eventful the rating.

## Acknowledgments

Authors would like to express deepest gratitude to Sandra Lichoń for her invaluable input throughout the research process. Moreover, authors are deeply grateful to all auditory tests' volunteer participants. This work was partly funded by the Department of Mechanics and Vibroacoustics of the AGH University of Science and Technology in Kraków, Poland (the grant number 16.16.130.942). The research project partly supported by the program "Excellence Initiative – Research University" for the AGH University of Science and Technology (AGH University of Krakow).

## References

1. ALETTA F., KANG J., AXELSSON O. (2016), Soundscape descriptors and a conceptual framework for developing predictive soundscape models, *Landscape Urban Planning*, **149**: 65–74, doi: [10.1016/j.landurbplan.2016.02.001](https://doi.org/10.1016/j.landurbplan.2016.02.001).
2. ALETTA F., GUATTARI C., EVANGELISTI L., ASDRUBALI F., OBERMAN T., KANG J. (2019), Exploring the compatibility of "Method A" and "Method B" data collection protocols reported in the ISO/TS 12913-2:2018 for urban soundscape via a soundwalk, *Applied Acoustics*, **155**: 190–203, doi: [10.1016/j.apacoust.2019.05.024](https://doi.org/10.1016/j.apacoust.2019.05.024).

3. AXELSSON Ö., NILSSON M.E., BERGLUND B. (2010), A principal components model of soundscape perception, *The Journal of the Acoustical Society of America*, **128**(5): 2836–2846, doi: [10.1121/1.3493436](https://doi.org/10.1121/1.3493436).
4. BERGLUND B., NILSSON M.E. (2006), On a tool for measuring soundscape quality in urban residential areas, *Acta Acustica united with Acustica*, **92**(6): 938–944.
5. BORKOWSKI B., SUDER-DĘBSKA K., WICIAK J., SZLACHTA A.M. (2021), Acoustic panels inspired by nature, *Archives of Acoustics*, **46**(1): 135–146, doi: [10.24425/aoa.2021.136567](https://doi.org/10.24425/aoa.2021.136567).
6. BRADFER-LAWRENCE T., GARDNER N., BUNNEFELD L., BUNNEFELD N., WILLIS S.G., DENT D.H. (2019), Guidelines for the use of acoustic indices in environmental research, *Methods in Ecology and Evolution*, **10**(1): 1796–1807, doi: [10.1111/2041-210X.13254](https://doi.org/10.1111/2041-210X.13254).
7. BERNAT S. (2013), Awareness of noise hazards and the value of soundscapes in polish national parks, *Archives of Acoustics*, **38**(4): 479–487, doi: [10.2478/aoa-2013-0057](https://doi.org/10.2478/aoa-2013-0057).
8. BROWN A.L., GJESTLAND T., DUBOIS D. (2015), Acoustic environments and soundscapes, [in:] *Soundscape and the Built Environment*, Kang J., Schulte-Fortkamp B. [Eds.], pp. 1–16, CRC press.
9. CZOPEK D., MAŁECKI P., PIECHOWICZ J., WICIAK J. (2019), Soundscape analysis of selected landforms on Spitsbergen, *Archives of Acoustics*, **44**(3): 511–519, doi: [10.24425/aoa.2019.129266](https://doi.org/10.24425/aoa.2019.129266).
10. CZOPEK D., LICHON S., MAŁECKI P., WICIAK J. (2022), The Soundscape of the Sarek National Park – Case study, [in:] *Proceedings of the 24th International Congress on Acoustics*.
11. DAVIES W.J. *et al.* (2013), A Perception of soundscapes: An interdisciplinary approach, *Applied Acoustics*, **74**(2): 224–231, doi: [10.1016/j.apacoust.2012.05.010](https://doi.org/10.1016/j.apacoust.2012.05.010).
12. DE COENSEL B., BOTTELDOOREN D. (2006), The quiet rural soundscape and how to characterize it, *Acta Acustica United with Acustica*, **92**: 887–897.
13. FARINA A. (2014), *Soundscape Ecology: Principles, Patterns, Methods and Applications*, Dordrecht, Springer.
14. FLOWERS C., LE TOURNEAU F-M., MERCHANT N., HEIDORN B., FERRIERE R., HARWOOD J. (2021), Looking for the -scape in the sound: Discriminating soundscapes categories in the Sonoran Desert using indices and clustering, *Ecological Indicators*, **127**: 107805, doi: [10.1016/j.ecolind.2021.107805](https://doi.org/10.1016/j.ecolind.2021.107805).
15. International Organization for Standardization (2014), *Acoustics – Soundscape – Part 1: Definition and conceptual framework* (ISO Standard No. 12913-1:2014), <https://www.iso.org/standard/52161.html>.
16. International Organization for Standardization (2018), *Acoustics – Soundscape – Part 2: Data collection and reporting requirements* (ISO/TS Standard No. 12913-2:2018), <https://www.iso.org/standard/75267.html>.
17. International Organization for Standardization (2019), *Acoustics – Soundscape – Part 3: Data analysis* (ISO/TS Standard No. 12913-3:2019), <https://www.iso.org/standard/69864.html>.
18. JIANG L. *et al.* (2022), Ten questions concerning soundscape valuation, *Building and Environment*, **2019**: 109231, doi: [10.1016/j.buildenv.2022.109231](https://doi.org/10.1016/j.buildenv.2022.109231).
19. KANG J. *et al.* (2016), Ten questions on the soundscapes of the built environment, *Building and Environment*, **108**: 284–294, doi: [10.1016/j.buildenv.2016.08.011](https://doi.org/10.1016/j.buildenv.2016.08.011).
20. KRAUSE B. (2012), *The Great Animal Orchestra: Finding the Origins of Music in the World's Wild Places*, New York, Little, Brown and Company.
21. LAVALLE S.M. (2019) *Virtual Reality*, Cambridge University Press.
22. MANCINI S., MASCOLO A., GRAZIUSO G., GUARNACCIA C. (2021), Soundwalk, questionnaires and noise measurements in a university campus: A soundscape study, *Sustainability*, **13**(2): 841, doi: [10.3390/su13020841](https://doi.org/10.3390/su13020841).
23. MAŁECKI P., CZOPEK D., PIECHOWICZ J., WICIAK J. (2020), Acoustic analysis of the glacier caves in Svalbard, *Applied Acoustics*, **165**: 107300, doi: [10.1016/j.apacoust.2020.107300](https://doi.org/10.1016/j.apacoust.2020.107300).
24. MITCHELL A. *et al.* (2021), The International Soundscape Database: An integrated multimedia database of urban soundscape surveys – Questionnaires with acoustical and contextual information, <https://zenodo.org/records/10672568> (access: 15.02.2024).
25. MITCHELL A., ALETTA F., KANG J. (2022), How to analyse and represent quantitative soundscape data, *JASA Express Letters*, **2**(3): 037201, doi: [10.1121/10.0009794](https://doi.org/10.1121/10.0009794).
26. PIERETTI N., FARINA A., MORRI D. (2011), A new methodology to infer the singing activity of an avian community: The acoustic complexity index (ACI), *Ecological Indicators*, **11**(3): 868–873, doi: [10.1016/j.ecolind.2010.11.005](https://doi.org/10.1016/j.ecolind.2010.11.005).
27. RYCHTÁRIKOVÁ M., VERMEIR G. (2013), Soundscape categorization on the basis of objective acoustical parameters, *Applied Acoustics*, **74**(2): 240–247, doi: [10.1016/j.apacoust.2011.01.004](https://doi.org/10.1016/j.apacoust.2011.01.004).
28. SCARPELLI M.D.A., RIBEIRO M.C., TEIXEIRA C.P. (2021), What does Atlantic Forest soundscapes can tell us about landscape?, *Ecological Indicators*, **121**: 107050, doi: [10.1016/j.ecolind.2020.107050](https://doi.org/10.1016/j.ecolind.2020.107050).
29. SCHAFER R.M. (1977), *The soundscape: Our Sonic Environment and the Tuning of the World*, Destiny Books, Rochester, VT.
30. SOUTHWORTH M. (1969), The sonic environment of cities, *Environment and Behavior*, **1**(1): 49–70, doi: [10.1177/001391656900100104](https://doi.org/10.1177/001391656900100104).
31. UNESCO (n.d.), *Laponian Area*, <https://whc.unesco.org/en/list/774/> (access: 21.07.2023).
32. WANG M. (2020), Social VR: A new form of social communication in the future or a beautiful illusion?, *Journal of Physics: Conference Series*, **1518**: 012032, doi: [10.1088/1742-6596/1518/1/012032](https://doi.org/10.1088/1742-6596/1518/1/012032).



## Research Paper

# Assessing Spatial Audio: A Listener-Centric Case Study on Object-Based and Ambisonic Audio Processing

Paweł MAŁECKI\*, Joanna STEFAŃSKA, Maja SZYDŁOWSKA

*Department of Mechanics and Vibroacoustics, AGH University of Krakow  
Kraków, Poland*

\*Corresponding Author e-mail: [pawel.malecki@agh.edu.pl](mailto:pawel.malecki@agh.edu.pl)

(received July 18, 2023; accepted March 20, 2024; published online June 11, 2024)

The research explores the production and critical evaluation of two distinct mixes of “Dancing Ends”, a musical composition by Łukasz Pieprzyk. These mixes were engineered using two cutting-edge spatial sound technologies: Dolby Atmos and Ambisonics. The recording process incorporated overdub and multitrack recording techniques. Once created, the mixes were evaluated using a method of direct rating, based on an average rank system from 1 to 5, adhering strictly to the (ITU-R, 2015) BS.1116-3 and (ITU-R, 2019) BS.1284-2 standards. Evaluation criteria included factors such as mix selectivity, depth, width, and height of the sound stage, sound envelopment, tonal brightness, and quality of source localization. Additionally, some criteria were specifically tailored to evaluate characteristics unique to the composition. The evaluations were performed on three different listening systems and environments: surround systems of 5.1 and 7.1.4, and binaural listening. Although Ambisonics’ mix received higher ratings in several categories, Dolby Atmos’ mix was preferred across all listening environments. The results underscore the potential benefits of employing spatial sound technologies in music production and evaluation, offering insight into the capabilities of Dolby Atmos and Ambisonics.

**Keywords:** spatial sound technologies; Dolby Atmos; Ambisonics; music production; sound evaluation.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The origins of spatial sound experiments date back to the 1930s (SPORS *et al.*, 2013). Presently, with the advancement in sound engineering, extraordinary possibilities are being achieved in the field of three-dimensional sound technology. Modern technologies not only ensure precise reproduction in the frequency domain but also allow for faithful representation of the sound space. Nowadays, immersive audio format is available in a binaural standard for the mass user (Apple, 2023). The binaural format enables users to experience three-dimensional audio over standard headphones, enhancing the listening experience in everyday use.

In 2012, a new spatial sound format, Dolby Atmos, was introduced, which uses audio objects (KELLY *et al.*, 2020). This innovative format added a new dimension of height to surround sound systems, offering a more immersive and realistic sound experience. The

technology comprises two basic elements: bed tracks and objects. Bed tracks are channel-based buses that can be decoded to various standard configurations such as 2.0, 5.1 or 7.1 but all with fixed locations, strictly defined by the speaker layout. On the other hand, objects refer to sound elements individually mapped on a hemisphere, independent of the reproduction system. These consist of an audio stream that is sent to the Dolby Atmos Renderer and a metadata stream that carries panning information to determine the location in the space (Dolby Laboratories, n.d.). The Dolby Atmos Renderer application is a pivotal component of any Dolby Atmos mixing system. In configuration with a digital audio workstation (DAW), it generates positional metadata that enables spatial representation of an audio mix in a playback environment. The number of input channels that can be configured depends on the renderer itself and the sampling frequency of the session. When operating at 48 kHz, the format supports 128 monophonic input channels, while at a sam-

pling frequency of 96 kHz, it handles 64 input channels. By default, channels 1–10 are configured as a 7.1.2 bed track, and channels 11–128 are objects; 7.1.2 notation refers to a speaker layout with seven main channels around the listener, one subwoofer channel for low-frequency effects, and two overhead or height channels.

Apart from Dolby Atmos, AURO-3D has emerged as another significant immersive audio technology. AURO-3D enhances the sound field by adding an additional “height” layer, creating a three-dimensional experience. The original AURO-3D approach was channel-based, with an emphasis on vertical sound layering to produce a more enveloping audio experience. This format has been recognized for its ability to produce a natural and realistic listener experience. Subsequent iteration AURO-Cx, have introduced a versatile engine that supports not only the original channel-based approach but also object-based and scene-based (Ambisonics) audio, along with scalable channel-based configurations, thereby broadening the potential applications of AURO-3D technology in various listening environments (AURO-3D, 2023).

Wave field synthesis (WFS), on the other hand, is a spatial audio rendering technique that uses many speakers to recreate an acoustic environment. It enables the synthesis of sound waves to form a continuous wave front, creating a sound field that can simulate sounds both inside and outside the listener’s space, offering a heightened sense of realism. While this approach to spatial sound provides listeners with an exceptional level of immersion by accurately reproducing the way sound interacts with the environment, it is accompanied by technical and financial challenges. Implementation is complex and costly, and it can be prone to truncation error and spatial aliasing, which are limitations that can affect sound quality and spatial accuracy (WITTEK, 2013).

Ambisonics is a sophisticated spatial audio technology that enables the encoding and decoding of sound fields in a full-sphere around the listener. It is grounded in the principles of spherical harmonics, which are used to mathematically represent complex sound fields (ZOTTER, FRANK, 2019). Contrary to channel-based audio systems that transmit signals to designated speakers, Ambisonics represents the sound field in a speaker-independent format, utilizing the physical properties of the sound to create a scene-based audio experience. It encodes sound waves in a way that captures their complete directional information, which can be decoded by a corresponding array of loudspeakers.

The evaluation of spatial sound quality comprises several different components, the attributes of which include source location, perceived source width, and listener envelopment (POWER, 2015). According to (RUMSEY *et al.*, 2005), spatial attributes account for over one-third of all quality ratings in listening tests and are therefore crucial in determining the quality

of a system. In the work of (FRANCOMBE *et al.*, 2017), an evaluation of spatial sound reproduction methods was conducted, and it was found that the listening test results are influenced by the test subject’s experience with multichannel formats. The most common attributes used by experienced and inexperienced listeners to describe auditory impressions were identified. Experienced listeners used the depth of the sound field, surroundings, and spectral clarity, while inexperienced listeners determined the position of the sound source, its transparency, and the space. ORAMUS and NEUBAUER (2020) conducted studies comparing an object-based and channel-based panning models. Tests were conducted with 127 subjects to compare the perceived positions of six audio samples, each of which was reproduced in 5.1, 7.1, and Dolby Atmos. The results did not show an increase in spatial location precision when using object sound, however, listeners demonstrated greater confidence in determining the position of the sound object compared to conventional channel-based playback. CENGARLE (2013) compared the Ambisonic technique and 5.1 surround – he stated that first-order Ambisonics is suitable for diffuse sounds. Furthermore, he concluded that the higher-order Ambisonic technique used in a spatial system enhances perceived realism compared to the 5.1 system. KLECZKOWSKI *et al.* (2015) examines the perceptual effect of the separation of the components of direct and reflected sound impulse responses in multichannel systems, using phantom sound sources. The findings reveal a more consistent perceptual advantage of separation, particularly among experienced listeners. Many research papers point to the significant superiority of spatial sound over stereo productions. In a prior study by the authors of the current work (MALECKI *et al.*, 2020), the focus was on electronic music. This earlier study involved creating a spatial remix of a stereophonic composition using Ambisonics. This was followed by a subjective comparative analysis between the original stereophonic version and the spatially remixed Ambisonic version. The primary objective was to explore the potential of spatial dimensions and an extended music scene. The subjective evaluation involved a group of experts and predicted playback in stereo and Ambisonic configurations, as well as binaural listening. The subjects evaluated aspects such as spatiality, selectivity, timbre, dynamics, and overall impression. On the basis of the listening tests conducted, a preference for spatiality and selectivity of the Ambisonic production was established. On the basis of a comparison of the stereophonic and binaural render of the Ambisonic mix, a clear preference for spatiality in the binaural version was demonstrated.

The purpose of this study is to produce and conduct auditory evaluations of music mixes created using two distinct spatial sound technologies: Ambisonics and Dolby Atmos. This comparative analysis aims

to understand the nuances in listener perception and audio quality between these advanced sound reproduction methods. By doing so, the study explores the effectiveness and immersive qualities of each technology in the context of classical music production.

In the first stage, music mixes were created using the two technologies. The first mix was prepared using the Dolby Atmos Production Suite. The Pro Tools Ultimate digital workstation was used to create the mix. The Dolby Atmos Renderer application, which communicates with the DAW software, was used to generate positional metadata that allows for precise spatial reproduction of the mix. The second mix was performed using the Ambisonic technique in REAPER software. To compare the final materials, a key task was to reproduce the first production. The [IEM Plug-in Suite \(2023\)](#), which includes a set of open-source Ambisonic plug-ins, was used to create the mix. The next section of the study describes the survey conducted of the participants in terms of auditory impressions. It was investigated how listeners perceive the selected audio formats in terms of spatiality and sound quality in various loudspeaker system configurations and in binaural listening. The following chapter presents a statistical analysis of the results of listening tests in terms of technology preference and listening system.

The article represents a significantly expanded continuation of the work ([MALECKI et al., 2023](#)) that was presented and discussed at a conference. This current paper encompasses a considerably more detailed description of the conducted experiments, additional results, and an in-depth statistical analysis. The manuscript emphasizes its novel contributions through new results and expanded analysis, clearly delineating its incremental advancements.

## 2. Production of spatial sound mixes

The composition selected for spatial mixing and subsequent evaluation is “Dancing Ends for Symphony Orchestra and Piano” by [PIEPRZYK \(2023\)](#). This score belongs to the genre of film music, originally produced in a stereophonic format. Łukasz Pieprzyk is an alumnus of the Krakow Academy of Music, where he studied composition under the tutelage of Professors Zbigniew Bujarski and Krzysztof Penderecki. The production phase aimed at a natural representation of the musical stage, striving for a realistic placement of a symphony orchestra ensemble. It was assumed that the listener’s position as shown in Fig. 1 within the sound space would mirror that of a conductor in a concert hall, characterized by a typical reverberation time of approximately 2 s. No dynamic compression was used during the mixing process. Only slight timbre equalizations and the manual adjustments of volume levels over time within DAW software were done. This tech-

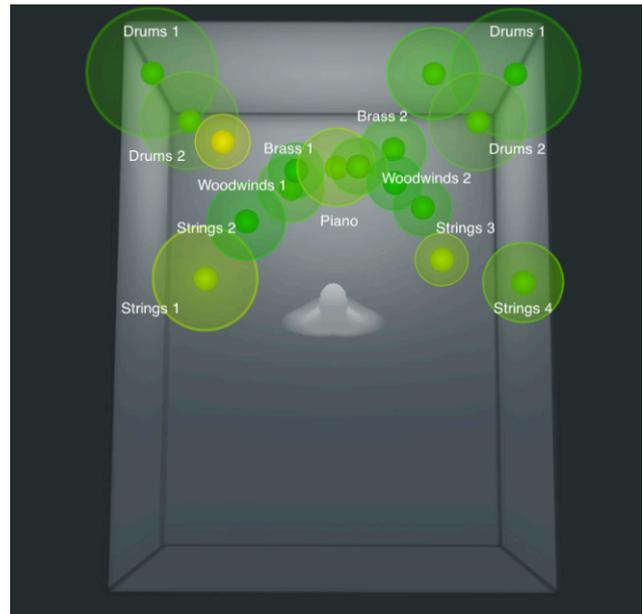


Fig. 1. Main source panning as represented in Dolby Atmos Panner.

nique was employed to maintain the appropriate balance of the orchestra and to accurately reflect the dynamic changes in the music. These adjustments were made track-by-track and were crucial for preserving the natural dynamics and expression of the orchestral performance.

The composition was produced in the early period of the SARS-COV-2 pandemic, and the recording was made during its ongoing course; thus the recording was made using the overdub method at the [Kotłownia Recording Studio \(2023\)](#) of the AGH University of Krakow. Each individual section or instrument of the orchestra was recorded separately to a guide track (pilot) provided by the composer. The ensemble included a diverse array of percussion instruments – a casa, daiko, snare drum, toms, tam-tam, kettledrums, along with various types of bells, cymbals, and smaller percussive elements, cumulatively forming two distinct percussion sets, each recorded on separate stem together with individual mono tracks with main percussion elements. Also, the orchestration featured an array of woodwinds and brass, including flutes, oboes, clarinets, bassoons, French horns, trumpets, tubas, and trombones. The string section comprised first violins, solo violin, second violins, violas, cellos, and double basses, complemented by a grand piano. During these sessions, the musicians wore one-ear headphones, through which the pilot track was played. This setup enabled the musicians to hear the guide track in one ear while still maintaining a natural perception of their own instrument. Additionally, a conductor was present in front of the musicians during the recording. The conductor also wore headphones to follow the pilot track and led the musicians through their performance, en-

suring coherence and musicality akin to a traditional orchestral recording.

In the recording of each instrumental section performing in unison (*tutti*), the Blumlein pair microphone technique was utilized, involving a pair of Neumann U87 microphones. This method was selected to ensure the capture of the rich acoustic detail and spatial characteristics of the ensemble's performance. The placement of the microphones relative to the instruments was determined based on the Recording Angle of the stereo pair, which provided a balance between direct sound and ambient reflections and avoided the proximity effect typically associated with "close miking". For soloists or smaller sections, such as tubas and trombones, individual microphones were used as necessary, with options including the AKG C414 or Schoeps MK4. Drum instruments were similarly captured individually and were further enhanced by an overhead stereo microphone configuration. The entirety of the recording process took place in the live room of Kottownia Studio, which boasts roughly 80 m<sup>2</sup> of the floor space and an average ceiling height of 8 m. The studio's reverberation time is around 1 s for the acoustic bandwidth, thereby providing a controlled yet resonant acoustic setting ideal for high-fidelity recordings.

The preparation of stems is a very important element in the process of creating a spatial mix. Stems are mono- or stereo-audio files that create subgroups of similar sound sources and represent specific elements of the mix. Typically, a stem represents a group of instruments, such as strings, percussion, or woodwinds. However, this is not a strict rule, and instruments are grouped depending on the genre of material being produced. By default, stems take into account the signal processing chain applied to their components. When played together, they create a full musical mix. Stems have found widespread use in the film industry and are typically divided into dialogues, music, and sound effects. For the execution of the surround mixes, encompassing both Dolby Atmos and Ambisonic formats, a total of 27 stereo and mono stems were rendered from the original Pro Tools stereo mix session. In alignment with our production concept, the rendering of these stems was deliberately executed without incorporating any previously applied equalization, dynamic processing, or reverberation effects. This approach ensured the preservation of the raw, unaltered essence of each instrument group, allowing for greater precision and creativity in the subsequent spatial mixing phase.

### 2.1. Dolby Atmos mix

The initial plan was to create a first mix using Dolby Atmos technology. Once this was established, a parallel approach was utilized to process the material using Ambisonics, based on the preliminary decisions derived from the Dolby Atmos execution. The first

part of the composition contained only percussion instruments. Due to the relatively small number of sources, sound objects were widely panned as shown in Fig. 2. In the second part, the entire symphony orchestra participated. The sound space was divided into plans where individual sections of the instruments were placed. Sections of string instruments, woodwind and brass were distinguished. According to the standard orchestra layout, string instruments were placed at the front, followed by woodwinds and then brass. Figure 1 shows the panning position of the main instrument sections. The figure does not show reverb panning or the drums during the introduction, nor the sound effects of the composition's outro. To achieve greater selectivity and separation between objects, the height of the sources was added; the wind instruments were positioned slightly above the string section. Assigning varied sizes to the instruments aided in unifying their sound while also imparting a distinctive character to each.

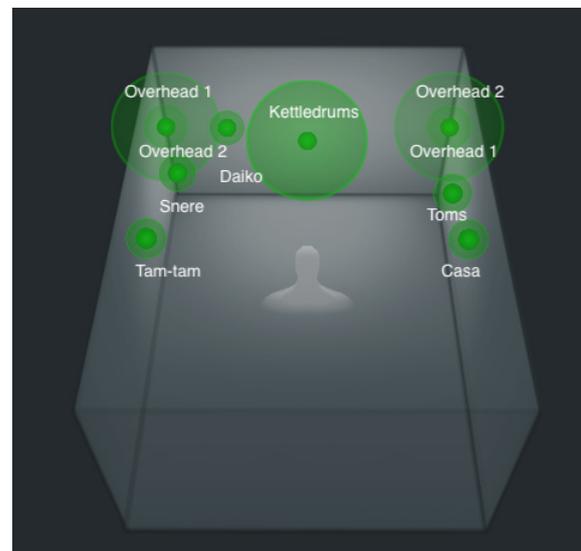


Fig. 2. Percussion instruments panning as represented in Dolby Atmos Panner.

Once the mixing process was completed, files were generated to facilitate playback on the intended systems. The renders were made based on the main Master File. It was created in real time during the recording process from Pro Tools to Dolby Renderer. The Master File contained three files with extensions: .atmos, .atmos.metadata, and .audio. The first of them, the top-level file, provided basic information about the project. The second contained all the 3D position coordinates for the object sound in the .audio file, while the last contained audio data for all bed track signals and objects. The rendering was performed using Dolby Atmos presets for the 5.1, 7.1.4, and binaural formats. The binaural audio was rendered statically, without any additional diffuse-field or free-field equalization applied. Furthermore, no head tracking

was employed in the rendering process, implying that the binaural audio output remained consistent regardless of the listener's head movements.

The preliminary version of the spatial mix in Dolby Atmos technology was made in the AGH Music Studio Kotłownia (Fig. 3). The listening room system is based on a 5.1 configuration, standardized by the (ITU-R, 2022) BS.775 standard. The system includes the following speaker models: Genelec 1034 BM (*L*, *R* channels), Genelec 1034 BC (*C* channel), Genelec 1038B (*Ls* and *Rs* channels) and Genelec 7360 (LFE channel).



Fig. 3. Control room of the Kotłownia recording studio of the AGH University of Krakow (Kotłownia Recording Studio, 2023).

The validation and additional spatial enhancement of the mix were carried out in the ATMOS Sound Truck (Fig. 4), equipped with an SSL T80 unit and a Dolby Atmos 7.1.4 monitoring system, based on Genelec speakers. The setup included Genelec 8351 (*L*, *R* channels), Genelec 8331 (*C* channel), Genelec 8320 (*Ls*, *Rs*, *Lrs*, *Rrs*, *Ltf*, *Rtf*, *Ltr*, *Rtr* channels) and Genelec 7360 (LFE channel) (Fig. 5). The pro-



Fig. 4. 120 dB ATMOS Sound Truck (120db Sound Engineering, n.d.).

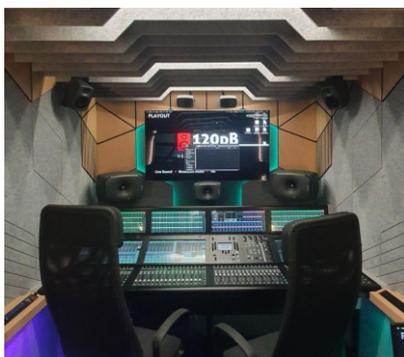


Fig. 5. Inside of 120 dB ATMOS Sound Truck (120db Sound Engineering, n.d.).

cess involved a thorough check of several key aspects to ensure the mix's quality and spatial accuracy. Selectivity and localization precision, ensuring that each sound element was clearly distinguishable and accurately positioned within the sound field. The balance between direct sound and reverberation. This involved fine-tuning the mix to achieve the right blend of clarity and spatial depth, ensuring that the reverberation did not overpower the direct sound but rather complemented it to enhance the overall spatial impression. The balance of levels across the mix was meticulously adjusted that all elements were at appropriate levels relative to each other, maintaining a harmonious and cohesive soundstage.

After finishing the first mix and preparing it for different listening setups, work on an Ambisonics version of the audio has been started. This phase involved carefully directing the audio signals and using special tools for encoding and decoding.

## 2.2. Ambisonic mix

The Ambisonic mix was carried out in the Auralization Laboratory of the Department of Mechanics and Vibroacoustics at the AGH University of Science and Technology. The room is equipped with a 16-channel system arranged in a spherical layout (Fig. 6). The system consists of sixteen Genelec 6010 speakers set in a radius of 1.5 m from the center of the sphere.

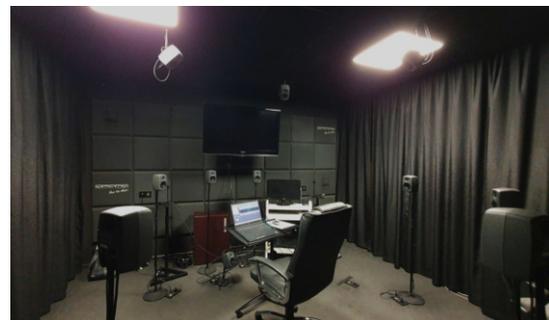


Fig. 6. Auralization Laboratory at the AGH University of Krakow.

The loudspeakers are arranged in three layers relative to the listener's ear level. In the horizontal plane, eight loudspeakers (channels 1 to 8) are positioned at ear level, at azimuthal angles of  $45^\circ$  increments. Above the listener, four loudspeakers are placed at an elevation angle of  $45^\circ$ , also spaced at  $90^\circ$  intervals azimuthally. Similarly, below the listener, four loudspeakers are situated at an elevation angle of  $-45^\circ$ .

The configuration of the IEM AllRAD (IEM Plugin Suite, 2023) decoding plugin mirrored the physical speaker setup with basic decoding of third-order Ambisonics.

The Ambisonic mix was fundamentally intended to mirror the mix in Dolby Atmos technology. This en-

tails that specialized Ambisonic tools were employed to craft the optimal mix, ensuring adherence to the assumptions and the results obtained from the Dolby Atmos mix. Therefore, from the Dolby Atmos session, 27 stems were generated, incorporating signal processing elements such as automatic volume, correction, and compression. This procedure was carried out to ensure the coherence of sound and dynamics of the signals. Based on the created stems, a spatial Ambisonic mix was created. In order to achieve better source separation and spatiality, some instruments were positioned slightly below the listener's head level, which could not be accomplished using Dolby Atmos technology.

The positioning of signals in the space and their simultaneous encoding in the Ambisonic domain was achieved using the StereoEncoder plugin (IEM Plug-in Suite, 2023). Instruments were intended to be placed as similar as possible to the positioning in the Dolby Atmos mix. To adapt the created Ambisonic mix to the 5.1 and 7.1.4 playback systems, the AllRADecoder plugin was used to decode the third-order Ambisonic signal to selected arrangements. To appropriately design the decoder, JSON configuration files were created for the 5.1 and 7.1.4 setups, containing information on the coordinates of all speakers and their corresponding channel numbers. For the binaural version of the Ambisonics mix, an HRTF set from the Neumann KU 100 dummy head was used by Binaural Decoder (IEM Plug-in Suite, 2023). No additional headphone correction was applied.

### 3. Listening evaluation

#### 3.1. Tests in a 5.1 surround sound system

The subsequent phase of the research involved the preparation of a protocol for subjective tests. Two listening rooms were chosen for the study: the AGH Kotłownia Music Studio and the AGH Auralization Laboratory. The former room was used to perform tests on the surround 5.1 system. The placement of individual channels was designed according to the (ITU-R, 2022) BS.775 standard. The speaker setup consisted of the following models: Genelec 1034 BM ( $L$ ,  $R$  channels), Genelec 1034 BC ( $C$  channel), Genelec 1038B ( $Ls$  and  $Rs$  channels) and Genelec 7360 (LFE channel) as shown in Fig. 7. The distance from each speaker to the sweet spot equals 2.8 m. The system is based on the AVID HDX PCIe Card sound card. The room is characterized by complete acoustic adaptation, ensuring appropriate conditions in the listening space by RFZ (reflection free zone) solution (Fig. 3). The room is symmetrical with respect to the vertical plane and the floor surface has a trapezoidal shape. The room's area meets the requirements specified for a multichannel system, measuring over  $35\text{ m}^2$  according to (ITU-R, 2015) BS.1116 standard.

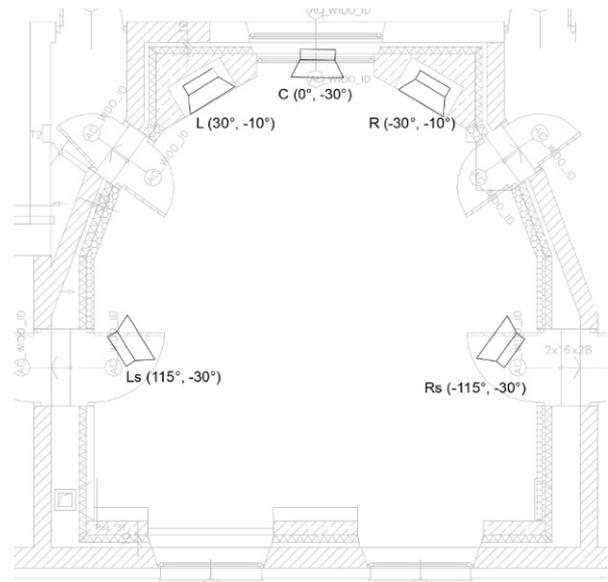


Fig. 7. Speaker placement in the control room of the Kotłownia Recording Studio (2023).

The mean measured reverberation time in third-octave bands from 200 Hz to 4 kHz falls within the range of 0.2 s to 0.4 s according to (EBU, 2004) Tech 3276 S1.

#### 3.2. Tests in 7.1.4 surround sound system

The 7.1.4 configuration for listening tests was implemented in the AGH Auralization Laboratory. The system was equipped with 11 Genelec 6010 speakers, arranged in a radius of 1.5 m from the sweet spot as shown in Fig. 8, and a PSI Sub A225-M subwoofer, all according to Dolby Atmos recommendations. The room has basic acoustic adaptation. The average RT20 (reverberation time) is 0.15 s, calculated for 500 Hz and 1000 Hz. The dimensions of the laboratory are  $3.9\text{ m} \times 6.7\text{ m} \times 2.8\text{ m}$ . The room meets most of the (ITU-R, 2022) BS.775 standard or has parameters very close to recommended. It meets the criteria for floor area and the ratio of dimensions. The RT is much

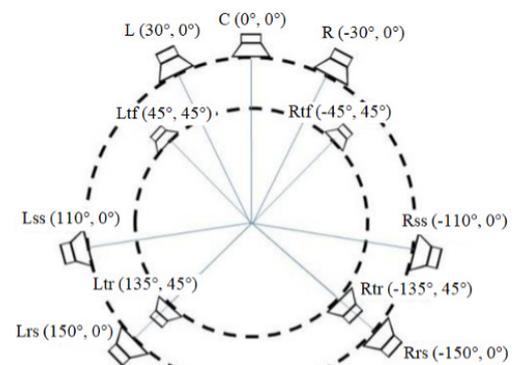


Fig. 8. Speaker placement in the control room of the Auralization Laboratory.

shorter than the recommended 0.56 s and the noise level is 19 dB(A) and meets the most stringent rating, NC15.

### 3.3. Binaural tests

For binaural listening, a set-up that incorporates a computer equipped with DAW software, Focusrite Scarlett 8i6 audio interface, and the Beyerdynamic DT770 Pro 250 Ohm headphones were used.

### 3.4. Subjective tests protocol

To evaluate the audio samples, both absolute (direct) and relative (comparative) evaluation methods were implemented, guided by the [ITU-R \(2019\) BS.1284](#) and [ITU-R \(2015\) BS.1116](#) standards. This included a global assessment of the overall quality or differences in the given objects and a parametric examination of individual sound attributes such as clarity, spatiality, and timbre. The test plan incorporated three distinct evaluation methods: detection, ordinal, and assignment procedures, each chosen according to the nature of the question. The detection method involved questions related to compatibility evaluation (determining if samples are identical or different) or situations that required a choice (identifying the differing sample). Ordinal evaluation was used for questions regarding ranking (intensity of a certain feature), preferences (better/worse), and similarity (most similar/different). The assignment method allowed for numerical estimation across different types of scales. Listening tests were organized as surveys, in which participants analyzed the material and made choices between the music excerpts presented. The scales for rating were discrete, graphical, and accompanied by labels. Using Google Forms, a survey consisted of eight questions:

- 1) Rate the following sound properties:
  - a) selectivity
  - b) depth of soundstage,
  - c) width of soundstage,
  - d) height of soundstage,
  - e) sound immersion,
  - e) clarity of sound,
  - f) localization quality.
- 2) Evaluate in which mix you can better locate the flute?
- 3) Assess whether the piano's position aligns precisely in both mixes?
- 4) Evaluate in which mix you rate the balance between the string section and the brass section better?
- 5) Assess if any of the mixes more realistically represented the placement of musicians in the space?
- 6) Which mix is more balanced in terms of frequency?
- 7) Which mix do you prefer?
- 8) What aspects differentiate these mixes the most? (choose two):
  - a) selectivity,
  - b) timbre,
  - c) source location,
  - d) listener's perspective,
  - e) width of the soundstage,
  - f) sound envelopment.

The sound samples for each question were carefully selected to ensure signal diversity and reduce listener fatigue. Each music excerpt was designed to be directly related to a specific question. To optimize the accuracy of the results, each listener was tested individually with an interactive signal presentation that allowed for unlimited repetitions of each excerpt. All the sound samples were logically arranged to avoid abrupt endings and presented in random order. The first question followed a single signal presentation principle (parameter evaluation), whereas the subsequent questions used a paired comparison (preference or difference evaluation). In the first and the last questions, signals were played one after another, whereas in the other questions, switching between signals was enabled.

Following the guidelines of [ITU-R \(2019\) BS.1284](#) for conducting subjective tests, the participant group comprised a so-called “expert group” of at least ten individuals. For this experiment, twelve people were involved for the 5.1 system and binaural listening, and ten people for the 7.1.4 configuration. Each participant had a higher education degree in acoustics, had basic skills in sound production, and previous experience in listening tests. Some participants also did first- or second-degree music education. All were otologically normal, which means that they were free of diagnosed diseases or pathologies of the auditory system.

To minimize the potential influence of the participants' emotions and attitudes on their judgment, questions were precisely articulated, signals were equalized to the same level (SPL *A*-weighted equal to 80 dB), and all listeners received training prior to the listening tests. Initial preparation included familiarization with the survey structure, rules for presenting music excerpts, and methods of answering individual questions. Further clarification of the evaluation parameters of the first question was provided in a document at the beginning of the study, minimizing any misunderstanding of the applied concepts.

## 4. Results

Listening tests were conducted with the intent to compare the Ambisonic technique and Dolby Atmos in a selected speaker configurations. On the basis of the characteristics of the constructed questions, a distinction was made between qualitative and quantitative variables. Depending on the features under examination, the data were classified on an interval or a nominal scale.

### 4.1. Question 1

The construction of the first question indicated quantitative variables assigned to the interval scale. Following this assumption, it was necessary to investigate whether the results demonstrated characteristics of a normal distribution (the Shapiro–Wilk test), determine if variables were correlated, and verify if the variances of variables across populations were equal (Levene’s test). Depending on the final assignment of data, an independent  $t$ -Student test or a Mann–Whitney U test was performed (Table 1). When verifying the statistical hypotheses, a significance level  $\alpha$  of 0.05 was adopted in all tests. Assessments were made based on the responses collected for eight distinct questions.

The averaged parameter ratings (question 1) did not show significant differences between technologies. The Shapiro–Wilk test, carried out to assess the *selectivity* parameter, reached statistical significance of a normal distribution only for Dolby Atmos technology in the 7.1.4 system. In other tests, the null hypothesis was rejected. To evaluate statistical significance, Mann–Whitney tests were performed. It was found that the evaluation of *selectivity* of the mixes presented in the 5.1, 7.1.4 systems and in binaural listening, did not show significant differences depending on technology.

For both technologies, Levene’s tests proved that the variances of the *depth of soundstage* parameter are homogeneous so the null hypothesis was accepted. Also, based on the  $t$ -Student test, no significant differences depending on the technology used was found.

The distribution of the *width of soundstage* parameter was found to be normal only in the 5.1 system. The  $t$ -Student test also showed no significant differences in terms of the *width of soundstage* but within binaural and the 7.1.4 configuration, the performed tests showed significant differences in terms of the evaluated property within the two technologies. In both cases, Ambisonic signals was rated higher than Dolby Atmos as shown in Fig. 9. The evaluation of the *Height of the soundstage* and *sound immersion* also did not show significant differences between the compared technologies. The distribution of the results for the *clarity of sound* was not normal for any listening technology. In the 7.1.4 configuration, the results showed significant differences depending on whether the Dolby Atmos or the Ambisonics was presented to the listeners, in favor of the Ambisonic system (Fig. 10). No statistically significant differences were obtained in the remaining systems.

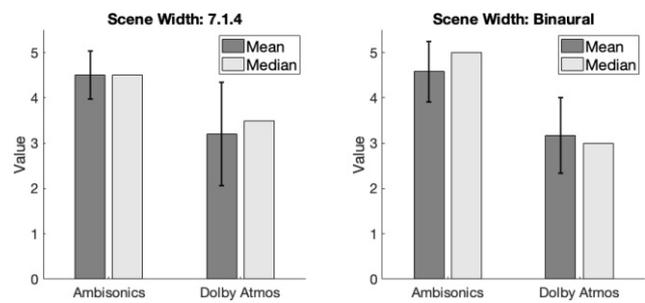


Fig. 9. *Width of soundstage* parameter rating for 7.1.4 system and binaural listening. Error bars represent one standard deviation from the mean.

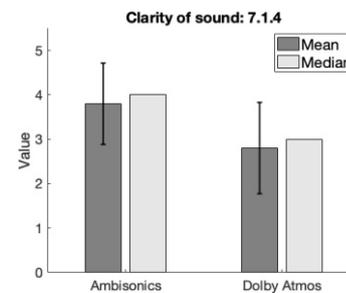


Fig. 10. *Clarity of sound* parameter rating for 7.1.4 system. Error bars represent one standard deviation from the mean.

Table 1. Significance test results for all listening configurations for question 1.

Parameter	5.1		7.1.4		Binaural	
	Test	p-value	Test	p-value	Test	p-value
Selectivity	M–W	0.667	M–W	0.155	M–W	0.116
Depth of soundstage	$t$ -test	1.000	$t$ -Test	0.492	$t$ -test	0.292
Width of soundstage	$t$ -test	0.239	M–W	<b>0.010*</b>	M–W	<b>0.001*</b>
Height of soundstage	M–W	0.976	M–W	0.345	M–W	0.707
Sound immersion	M–W	0.951	M–W	0.097	M–W	0.066
Clarity of sound	M–W	0.206	M–W	<b>0.030*</b>	M–W	0.763
Localization quality	M–W	0.140	M–W	<b>0.018*</b>	M–W	0.233

M–W – Mann–Whitney test,  $t$ -test –  $t$ -Student test.

\*Significant values ( $p < 0.05$ ).

The Shapiro–Wilk test, conducted to assess the *localization quality* parameter, reached statistical significance of the normal distribution assumption for Dolby Atmos technology in the 5.1 and 7.1.4 systems. However, for none of the systems in Ambisonics technology, a normal distribution was obtained for the examined feature. The Mann–Whitney significance test, performed for the 5.1 system and binaural listening, showed that the characteristic analyzed does not show significant differences depending on the technology. The analysis for the 7.1.4 configuration showed significant differences in the assessment of the *localization quality* and Ambisonic coding was rated better than Dolby Atmos (Fig. 11).

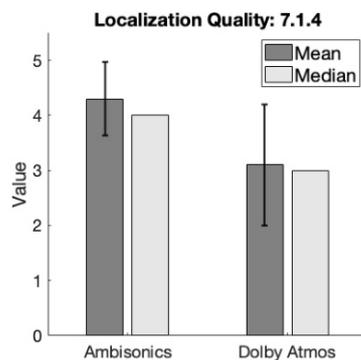


Fig. 11. *Localization quality* parameter rating for 7.1.4 system. Error bars represent one standard deviation from the mean.

In addition, a statistical comparison of the two technologies was conducted in terms of the significance of the system used on the results for all the evaluations received for the 7 perceptual parameters in the question 1. The evaluation results of a given parameter were significantly dependent on the listening system when the  $p$ -value was less than the significance level  $\alpha$  set to 0.05. Ambisonic mix evaluations for the 5.1 system compared to the 7.1.4 system showed significant differences only for the *width of soundstage* parameter ( $p = 0.011$ , the Mann–Whitney test). For evaluations of Dolby Atmos mix, statistical significance tests did not show differences between listening in 5.1 and 7.1.4.

When compared against binaural listening with respect to the 7.4.1 system, the evaluations of the Ambisonic system differed statistically for the *height of the soundstage* parameter ( $p = 0.034$ , the Mann–Whitney test). When comparing binaural listening to the 5.1 system, statistically significant differences were only obtained for the *width of soundstage* parameter ( $p = 0.005$ , the Mann–Whitney test).

For the mix in Dolby Atmos technology, the assessment of the *sound clarity* parameter depended on whether the hearing tests were conducted in binaural listening or in the 7.1.4 system ( $p = 0.042$ , the Mann–Whitney test). When comparing headphone listening with the 5.1 system for Dolby Atmos mix, no statisti-

cally significant differences were demonstrated for any parameter.

#### 4.2. Questions 2-7

The results for questions 2–7 are shown in Fig. 12. In the question 2, study participants were asked to select the technology in which they could more accurately locate the flute. In each listening system, a larger percentage of respondents indicated that Dolby Atmos technology allows for more precise localization. During tests conducted in the 5.1 system, 83 % of respondents chose the mix made in Dolby Atmos, in the 7.1.4 configuration the same answer was indicated by 70 % of people, whereas during binaural listening – 58 %.

The next question concerned the placement of the piano. The subjects determined whether the location of the instrument matched in both presented pieces of music. According to the majority of listeners, it did not – respectively, 75 % and 80 % of respondents for the 5.1 and 7.1.4 systems. In the binaural listening, those who responded that the location of the instrument had changed were in the minority – 33 %. The listeners' answers indicated that there is a discrepancy between the mixes in the location of the piano, even though the instrument was positioned directly in front of the listener (Fig. 12).

In the question number 4, listeners chose the technology in which a better balance between the string section and the brass section was obtained. In the 5.1 configuration, 58 % of respondents chose Dolby Atmos and 42 % chose Ambisonics as the technology providing a better balance between the selected sections. The same percentage results were obtained for binaural listening, where the majority of respondents chose Dolby Atmos technology. The opposite situation occurred for the 7.1.4 system, where the balance of the Ambisonic mix was better assessed – 80 % of listeners pointed out this technology. It was noticed that the preference for a given technology was associated with the listening room in which this technology was implemented (Fig. 12).

In the following question, the respondents were asked whether any mix reflected the arrangement of musicians in the space, in a more realistic way. The answers obtained in the question 5 were very diverse. According to the respondents, in the 5.1 system, the Dolby Atmos mix reflected the arrangement of musicians in a more realistic way, while in the 7.1.4 configuration, the Ambisonic mix did. Ambisonics was better assessed in the room where the mix in this technology was implemented, similarly for Dolby Atmos. In binaural listening, no technology was distinguished that would enable arranging musicians in a more realistic way (Fig. 12).

In evaluating the presented musical materials for frequency balance, the majority of respondents in both

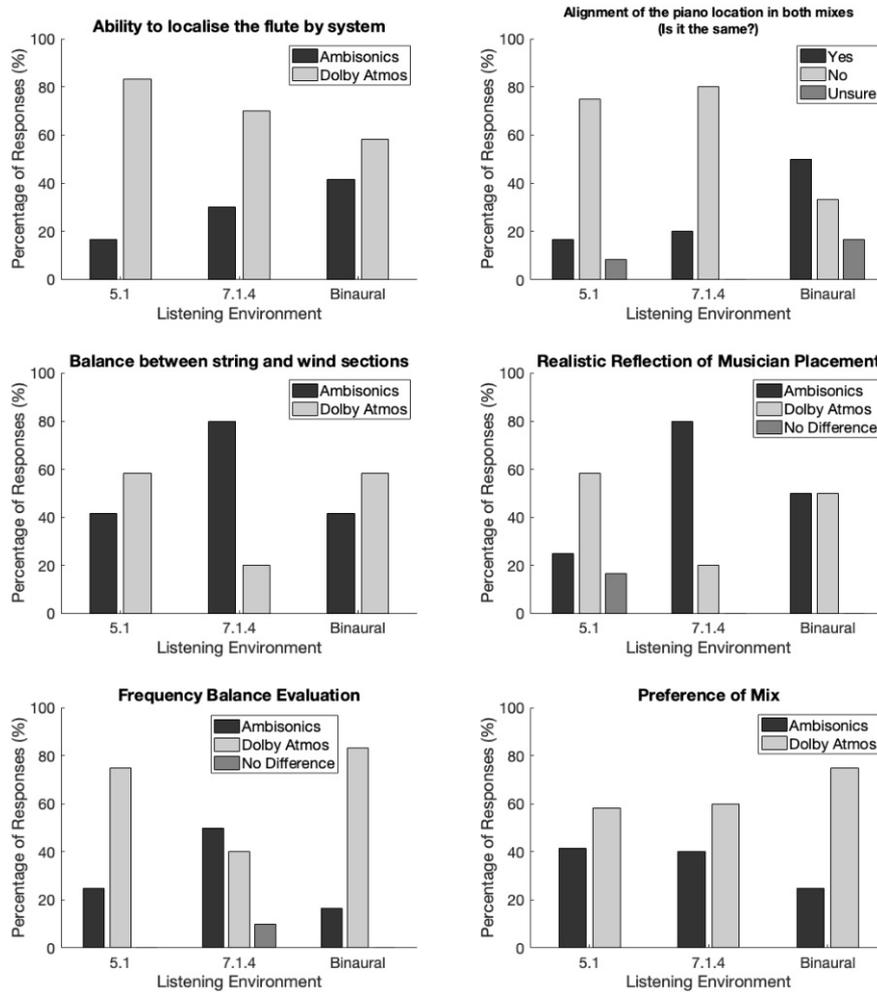


Fig. 12. Results from questions 2–7: Q2. Which mix provides better flute localization? Q3. Does the piano’s position align in both mixes? Q4. Which mix offers better balance between the string and brass sections? Q5. Which mix more realistically represents the musicians’ spatial placement? Q6. Which mix is more balanced in terms of frequency? Q7. Which mix do you prefer?

5.1 surround and binaural listening selected Dolby Atmos technology (Fig. 12). The significant advantage of one technology over another in headphone listening might have been influenced by the use of different binaural rendering algorithms. In the 7.1.4 setup, 10 % of participants indicated that they heard no difference between the mixes, 40 % chose the Dolby Atmos mix, and 50 % chose the Ambisonic mix.

The last question of this set of the results (the question 7) aimed to collect information about listener preferences. In all listening systems, most respondents decided that they prefer the Dolby Atmos mix. The largest advantage of the Dolby Atmos mix over the Ambisonic mix was obtained during binaural listening 75 % (Fig. 12).

#### 4.3. Question 8

The question 8 required the identification of two aspects that most differentiated the mixes made in two different technologies. When identifying the most

differentiating aspects between the samples, listeners most often chose *sound envelopment* for the 5.1 system and binaural listening, and *width of the soundstage* for 7.1.4 arrangement (Fig. 13). According to the respondents, significant differences between the mixes in the

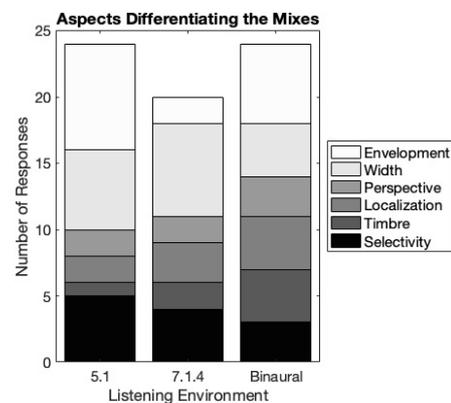


Fig. 13. Results from question 8: Which two aspects differentiate these mixes the most?

5.1 setup were also shown in the *selectivity* and *Width of the soundstage* parameters. In contrast to the 5.1 channel listening and binaural listening, in the 7.1.4 configuration, a small number of respondents voted for the *sound envelopment* parameter. In headphone listening, the respondents' answers were diverse, and there was no second dominating feature differentiating the presented musical materials.

Based on the conducted  $\chi^2$  independence test, it was found that the variables are independent, and the listening system does not affect the preference for a given technology ( $p > \alpha$ ). In statistical terms, the evaluations were not dependent on either the technology or the speaker configuration. For questions 2 through 8, independent  $\chi^2$  tests (variables unlinked in the nominal scale) were performed and shown in Table 2. The adopted null hypothesis for  $p > \alpha$  stated that the listening system does not affect the preference for a given technology and there is no significant relationship between the variables. Based on the calculations made, a decision was taken to reject the null hypothesis.

Table 2. Results of the  $\chi^2$  independence test for questions 2 to 8.

Question	$\chi^2$ value	$p$ -value	D.F.
2	1.809	0.405	2
3	6.865	0.143	4
4	4.163	0.125	2
5	8.929	0.063	4
6	6.246	0.182	4
7	0.867	0.648	2
8	7.311	0.696	10

## 5. Discussion

The implementation of this study involved making a sequence of critical decisions that unquestionably influenced the outcomes obtained. The process of creating a musical mix in each instance was deeply subjective and rooted in individual auditory perceptions. Numerous stages of material representation work were contingent on the personal judgment of the sound engineer, thus introducing an element of subjectivity.

This is noteworthy that the ratings gathered in the subjective tests were influenced by a variety of uncontrolled variables. These variables could be anything from the time of day when the tests were conducted, the listener's mood, or their prior experience with spatial audio. Such variables, though not directly controlled or manipulated in the study, could still exert significant effects on the results.

Additionally, another crucial decision, that was in essence arbitrary but had a potential bearing on the results, was the initial choice of starting the mix with Dolby Atmos technology instead of Ambisonics. Fol-

lowing this, there was an attempt to replicate the effect achieved with Dolby Atmos using Ambisonics. This approach, although logical in its structure, may have inadvertently introduced a bias towards the Dolby Atmos technology.

Furthermore, the basic mix was initially crafted in the 5.1 system and was subsequently examined and upgraded in a system specifically dedicated to Atmos. The selection of the 5.1 system as the starting point, followed by enhancement in the Atmos-specific system, was yet another decision that could have a substantial impact on the final outcome of the study. This sequence of decisions reinforces the fact that the results of the study, although comprehensive, are influenced by subjective choices and uncontrolled variables.

In the study, unconventional loudspeaker configurations, such as 5.1 or 7.1.4, were employed for the reproduction of Ambisonic recordings since it was intended to compare with Dolby Atmos system that is limited to standardized layouts. These configurations can impact the accurate rendering of the Ambisonic field due to their irregular spacing and positioning, which may not align with the standard Ambisonic decoding formats that are designed for uniform speaker layouts. Such irregular setups could potentially introduce spatial anomalies, especially when reproducing higher-order Ambisonics that rely on precise speaker placement to convey detailed spatial information. Converting third-order Ambisonic recordings to a 5-loudspeaker array, might result in spatial aliasing or spatial distortions. This is because the downmixing process does not preserve the higher resolution of spatial cues encoded in the third-order Ambisonic format, leading to a less accurate sound field reproduction.

## 6. Summary

The focal point of this research was a meticulous comparative evaluation of spatial audio mixing executed in two contemporary technologies – Ambisonic and Dolby Atmos. The study involved conducting an auditory examination of musical materials processed by these technologies. The breadth of the work was extensive and encompassed a detailed narrative of mix realizations, implementation of listening tests, in-depth statistical analysis, and a comprehensive interpretation of the data collected from the research.

The study was fundamentally rooted in the subjective analysis of the participants. The methodology involved executing surveys about the psychoacoustic impressions of the respondents across two different speaker configurations and binaural listening. These tests and the resultant feedback painted an interesting picture about the relative efficacy of these technologies.

Despite the Ambisonic mix scoring higher on many critical criteria (width of soundstage, clarity of sound,

localization quality, the Mann–Whitney test ( $p < 0.05$ ) for evaluating the quality of the musical material, the subjective tests pointed towards a general preference for the mix produced by Dolby Atmos technology across all the listening systems (preferences: 59 % for 5.1 and 7.1.4; 75 % for binaural). This preference resonated irrespective of the speaker setup and was observed even in binaural listening.

Delving into the statistical aspect of the research, interesting results were observed. In the case of the 5.1 system, no many significant differences were noted between examined variables. However, when it came to the 7.1.4 setup, the data showed substantial disparities in the evaluation of the scene width, sound clarity, and localization quality (the Mann–Whitney test, ( $p = 0.01, p = 0.03, p = 0.018$ )). Moreover, in binaural listening, the Scene width was marked with significant differences between loudspeaker systems ( $p = 0.001$ ).

Another very important conclusion is that the ambiguity of the results obtained suggests that the difference between the systems is not significant. This opens up possibilities for the production of high-quality spatial signals using open technology and free tools. Also, it can be stated that the purpose of the study was achieved. Although not all conclusions were statistically confirmed, the study successfully identified general trends in the auditory evaluation of the two technologies. Furthermore, this investigation underscores the high quality of both Ambisonics and Dolby Atmos technologies, highlighting their respective strengths and capabilities in spatial audio reproduction. The insights gained from this comparative analysis provide valuable contributions to the field of spatial sound and its application in music production.

Currently, spatial sound is on the rise, experiencing significant technological evolution and witnessing increased utilization in various fields. The subject matter addressed in this study warrants a broader examination considering different music genres and a range of listening systems. The findings of this research can serve as a robust foundation for further, more diversified analyses in the field of spatial audio technology.

### Acknowledgments

The research presented in this paper was conducted as part of the statutory activities of the Faculty of Mechanical Engineering and Robotics at AGH University of Science and Technology (AGH University of Krakow), under the project number 16.16.130.942. We would like to extend our sincere gratitude to the anonymous reviewers for their meticulous and insightful feedback, which greatly contributed to the enhancement of this manuscript. Their expertise and constructive criticism have been invaluable in refining our work and guiding its development to meet the high standards of academic rigor.

### References

- 120db Sound Engineering (n.d.), 120dB ATMOS Sound Truck, <https://www.120db.pl/atmos> (access: 12.06.2023).
- Apple (2023), About Spatial Audio with Dolby Atmos in Apple Music, <https://support.apple.com/en-us/HT212182> (access: 12.06.2023).
- AURO-3D. (2023), AURO<sup>®</sup>-CX<sup>™</sup>. Advanced next generation audio codec, NEWAURO BV, <https://www.auro-3d.com/wp-content/uploads/2023/08/Auro-Cx-White-Paper-rev1-20230714.pdf> (access: 30.11.2023).
- CENGARLE G. (2013), *3D audio technologies: Applications to sound capture, post-production and listener perception*, Unpublished Ph.D. Thesis, Universitat Pompeu Fabra.
- Dolby Laboratories (n.d.), The Dolby Atmos essentials course, <https://learning.dolby.com/course/info.php?id=191> (access: 12.06.2023).
- European Broadcasting Union (2004), *Listening conditions for the assessment of sound programme material*, EBU Tech 3276-E, Supplement 1.
- FRANCOMBE J., BROOKES T., MASON R., WOODCOCK J. (2017), Evaluation of Spatial Audio reproduction methods (Part 2): Analysis of listener preference, *Journal of the Audio Engineering Society*, **65**(3): 212–225, doi: 10.17743/jaes.2016.0071.
- IEM Plug-in Suite (2023), <https://plugins.iem.at/> (access: 12.06.2023).
- International Telecommunication Union (2015), *Methods for the subjective assessment of small impairments in audio systems*, Recommendation ITU-R BS.1116-3.
- International Telecommunication Union (2019), *General methods for the subjective assessment of sound quality*, Recommendation ITU-R BS.1284-2.
- International Telecommunication Union (2022), *Multichannel stereophonic sound system with and without accompanying picture*, Recommendation ITU-R BS.775-4.
- KELLY J., WOSZCZYK W., KING R. (2020), Are you there?: A literature review of presence for immersive music reproduction, [in:] *Audio Engineering Society Convention 149*.
- KLECZKOWSKI P., KRÓL A., MAŁECKI P. (2015), Reproduction of phantom sources improves with separation of direct and reflected sounds, *Archives of Acoustics*, **40**(4): 575–584, doi: 10.1515/aoa-2015-0057.
- Kotłownia Recording Studio (2023), <https://kotlownia.agh.edu.pl/> (access: 12.06.2023).
- MAŁECKI P., PIOTROWSKA M., SOCHACZEWSKA K., PIOTROWSKI S. (2020), Electronic music production in ambisonics-case study, *Journal of the Audio Engineering Society*, **68**(1/2): 87–94, doi: 10.17743/jaes.2019.0048.

16. MAŁECKI P., STEFAŃSKA J., SZYDŁOWSKA M., TĘCZYŃSKA KĘSKA M. (2023), A listening test evaluation of spatial sound technologies in music production: Dolby Atmos and ambisonics, [in:] *Audio Engineering Society Conference: AES 2023 International Conference on Spatial and Immersive Audio*.
17. ORAMUS T., NEUBAUER P. (2020), Comparison of perception of spatial localization between channel and object based audio, [in:] *Audio Engineering Society Convention 148*.
18. PIEPRZYK Ł. (2021), Dancing Ends for Symphony Orchestra and Piano (feat. Gajusz Keska), Pedagogical University in Cracow, <https://open.spotify.com/track/3fCCmIQvxLeAb0WdIuSgtj> (access: 12.06.2023).
19. POWER P.J. (2015), *Future spatial audio: Subjective evaluation of 3D surround systems*, University of Salford.
20. RUMSEY F., ZIELIŃSKI S., KASSIER R., BECH S. (2005), On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality, *The Journal of the Acoustical Society of America*, **118**(2): 968–976, doi: [10.1121/1.1945368](https://doi.org/10.1121/1.1945368).
21. SPORS S., WIERSTORF H., RAAKE A., MELCHIOR F., FRANK M., ZOTTER F. (2013), Spatial sound with loudspeakers and its perception: A review of the current state, [in:] *Proceedings of the IEEE*, **101**(9): 1920–1938, doi: [10.1109/JPROC.2013.2264784](https://doi.org/10.1109/JPROC.2013.2264784).
22. WITTEK H. (2013), *Perceptual differences between wavefield synthesis and stereophony*, Ph.D. Thesis (unpublished), University of Surrey.
23. ZOTTER F., FRANK M. (2019), *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer Nature, doi: [10.1007/978-3-030-17207-7](https://doi.org/10.1007/978-3-030-17207-7).



## Research Paper

## Modelling the Acoustic Properties of Baffles Made of Porous and Fibrous Materials

Krzysztof KOSAŁA

*Faculty of Mechanical Engineering and Robotics, Department of Mechanics and Vibroacoustics  
AGH University of Krakow*

Kraków, Poland; e-mail: [kosala@agh.edu.pl](mailto:kosala@agh.edu.pl)

*(received February 24, 2023; accepted January 21, 2024; published online April 12, 2024)*

The research described in the article addresses the problem of measurement, prediction and practical use of the acoustic properties of materials determined in an impedance tube. The aim of the research was to develop a simple calculation model for the insertion loss of small machinery enclosures, based on the normal incidence sound transmission loss and the normal incidence sound absorption coefficient of porous and fibrous materials. Both experimental and model tests were carried out on materials such as mineral wool, melamine foam and rebonded polyurethane foam.

Assessing the absorption properties of the tested porous and fibrous materials was performed using selected theoretical models, relating the calculations of the normal incidence sound absorption coefficient to measurements of this parameter conducted using an impedance tube. The application of the modified Allard and Champoux model brought the best results with the smallest discrepancies of the obtained results in relation to the experimental tests.

Assessing the sound-insulating properties of the tested mineral wool was carried out using the proposed calculation model for the normal incidence sound transmission loss, relating the obtained results to measurements conducted using an impedance tube. The assessment of the sound-insulating properties of porous and fibrous materials was performed using the proposed calculation model for insertion loss, which was validated using two prototype test stands for determining the insertion loss of cubic enclosures, in this case with walls made of porous and fibrous materials. Satisfactory results were obtained for engineering applications in the calculation results using the proposed models with respect to measurements. The results may have practical applications in assessing the effectiveness of acoustic enclosures, in which the basic construction material is an appropriate porous or fibrous plate, selected to have both sound-absorbing and sound-insulating properties.

**Keywords:** sound absorption coefficient; insertion loss; sound transmission loss; acoustical enclosures; porous materials.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## Acronyms

$\alpha$  – random incidence sound absorption coefficient,  
 $\alpha_f$  – normal incidence sound absorption coefficient,  
 $f$  – frequency,  
 $h$  – specimen thickness,  
 IL – insertion loss,  
 $m_p$  – surface density,  
 nTL – normal incidence sound transmission loss,  
 PCC – Pearson’s linear correlation coefficient,  
 $r$  – airflow resistivity,  
 $R$  – sound reduction index,  
 RMSE – root mean square error,  
 $R_S$  – airflow resistance.

## 1. Introduction

In issues related to the construction of anti-noise protection, such as acoustic barriers, enclosures or shields of noisy sound sources, both materials resistant to the penetration of sound waves and absorbing sound energy are used (ENGEL, SIKORA, 1997; FIEBIG, DĄBROWSKI, 2020; MORZYŃSKI, SZCZEPAŃSKI, 2018; VER, BERANEK, 2006). The basic acoustic parameter determining the properties of sound-insulating or sound-absorbing and insulating baffles, from which these protections are constructed, is the sound insulation.

Porous and fibrous materials are used in anti-noise protection, mainly as cores or sound-absorbing linings

in the construction of sound-absorbing and insulating baffles. The use of single materials, such as felt, can be found, among others, in the screening of office spaces at workstations in open-space rooms. Usually, the study of the acoustic properties of porous and fibrous materials consists of determining the sound absorption coefficient for the case of a sound wave incident perpendicularly on the specimen, which takes place during preliminary tests, especially for new materials, or for the random incidence of the sound wave on a specimen of relatively large size. Less frequently, studies are conducted to determine the extent to which these materials are resistant to sound wave penetration. The effectiveness of porous and fibrous materials in this respect is associated with increased density, and thus the airflow resistance, and with their appropriately large thickness, when insulation for lower frequency bands is desired.

The normal incidence sound absorption coefficient ( $\alpha_f$ ) of porous and fibrous materials can be determined using the semi-phenomenological and macroscopic empirical calculation models (COX, D'ANTONIO, 2017). Modelling sound propagation using a semi-phenomenological approach is more complicated and harder to use than an empirical one. Calculation of the characteristic impedance and propagation wavenumber requires knowledge of many material properties, such as porosity, airflow resistivity, tortuosity, viscous and thermal characteristic lengths, which are obtained by considering the microscopic propagation within the pores. An example of the semi-phenomenological approach is the model developed by ALLARD and CHAMPOUX (1992). The best-known empirical models include the model developed by DELANY and BAZLEY (1970) and models of other researchers such as MIKI (1990), MECHEL (1988), and QUNLI (1988). In these models, the impedance and wavenumber were determined empirically. To predict the absorption of sound-absorbing materials using these models, it is necessary to know the specific airflow resistivity and material thickness.

Research on materials of the porous and fibrous structures, used as linings or sound-absorbing cores in baffles, generally focuses on determining their absorbing properties. The problems of sound insulation properties of such materials are also investigated, although on a smaller scale (BIES, HANSEN, 2009). Sound insulation and sound absorption parameters of materials and baffles are determined primarily by laboratory tests (BATKO *et al.*, 2017; BERARDI, IANNACE, 2015; NURZYŃSKI, 2022; PUTRA *et al.*, 2015), while, with material data, they can also be estimated using calculation models (DAVY, 2009; SHARP, 1973; TRINH *et al.*, 2022; KOSAŁA *et al.*, 2020a).

The article focuses on the research into a certain group of sound-absorbing materials in the form of plates, which also have sound-insulating properties, especially when they are high-density materials, above

100 kg/m<sup>3</sup>, and with a large thickness of at least 50 mm.

While increasing the thickness of porous and fibrous materials, better absorption towards lower frequencies is obtained. An increase in the density of the material causes a decrease in the maximum absorption, which for this type of material occurs in higher frequency bands, and an increase in absorption in the region of lower frequencies, for which the sound absorption coefficient is generally low.

The sound insulation of baffles is defined by the sound reduction index  $R$  (International Organization for Standardization [ISO], 2021), which is determined in laboratory conditions, with the specimen placed in the measurement window separating coupled reverberation, transmitting and receiving rooms.

The sound reduction index  $R$  (also known as the sound transmission loss, TL) is defined as:

$$R = 10 \log_{10} \left( \frac{1}{\tau} \right) = 10 \log_{10} \left( \frac{W_i}{W_t} \right), \quad (1)$$

where  $\tau$  is the transmission coefficient defined by the sound power, the ratio of the transmitted power  $W_t$  and the power  $W_i$  incident on the specimen.

For laboratory measurement using sound pressure,  $R$  is calculated from the equation:

$$R = L_1 - L_2 + 10 \log_{10} \left( \frac{S}{A} \right), \quad (2)$$

where  $L_1$  and  $L_2$  are the average energy sound pressure level in dB in the source and receiving room, respectively,  $S$  is the area in m<sup>2</sup> of the free test opening in which the specimen is installed, and  $A$  is the equivalent sound absorption area in the receiving room, also in m<sup>2</sup>.

It is assumed that the sound fields are diffuse and that only sound radiated into the receiving room is from the specimen.

The parameter insertion loss (IL) is used to determine the effectiveness of acoustic enclosures, as the difference between the sound power levels of the unenclosed and enclosed sound sources (VER, BERANEK, 2006), according to the formula,

$$IL = 10 \log \left( \frac{W_0}{W_E} \right) = L_{W0} - L_{WE}, \quad (3)$$

where  $W_0$  and  $W_E$  are the sound power radiated by the unenclosed and enclosed sources, respectively, while  $L_{W0}$  and  $L_{WE}$  are the corresponding sound power levels.

Laboratory methods for the determination of IL of small machine enclosures are specified in (ISO, 2009). Evaluating the IL of acoustic enclosures with lined slits, using this standard, is shown in one work (NIE-RADKA, DOBRUCKI, 2018). Research on the properties of sound-insulating and sound-absorbing and insulating enclosures using a prototype test stand developed for this purpose is shown in some works (KOSAŁA *et al.*, 2020b; KOSAŁA, 2022).

In the case of porous or fibrous structures, the sound insulation properties of such materials, for conditions in which the sound wave falls perpendicularly on the specimen, can be determined using an impedance tube (JIANG *et al.*, 2017; KUNIO *et al.*, 2009; KOSALA, 2021), determining the normal incidence sound transmission loss parameter (nTL) (ASTM E2611-19, 2019), which is defined as ten times the common logarithm of the normal incidence sound transmission coefficient's reciprocal.

Two different parameters, nTL and  $R$ , cannot be compared with each other because they concern the normal or random incidence of sound on the specimen, respectively. Currently, there are no known methods to determine the correlation between these parameters. The method of determining nTL is used to compare the insulating properties of small specimens.

The purpose of the research described in the article is to determine the acoustic properties of single baffles made of porous and fibrous materials, which relate to the absorption and resistance of these materials to sound penetration. Acoustic tests were carried out using an impedance tube and two prototype stands for determining the IL of cubic enclosures, in this case with walls made of porous and fibrous materials.

Calculation models for the nTL of mineral wools and for the IL of an acoustic enclosure in the form of a cube using plates with the porous and fibrous structures were proposed. The calculation model for the IL for porous and fibrous materials can be used when their basic acoustic parameters in the form of the  $\alpha_f$  and the nTL are known. For baffles made of fibrous materials, in the form of mineral wool, material data, including airflow resistance, is sufficient to determine the IL of the enclosure. Validation of the calculation models was carried out using two developed prototype stands for determining the acoustic properties of enclosures.

## 2. Specimens' material data

Three high-density materials, above  $100 \text{ kg/m}^3$ , were tested, i.e., mineral wool and rebonded polyurethane foam, and a material with very low density and high absorption, i.e., melamine foam. All materials had comparable thicknesses of 50 mm or 60 mm. Figure 1 shows specimens of materials whose round shape resulted from the tests carried out on them, which were tests of air flow resistance (discs with a diameter of

100 mm) and the  $\alpha_f$  and the nTL (discs with a diameter of 34.9 mm).



Fig. 1. Material specimens of mineral wool (MW151), mineral wool with glass fleece (MW100), melamine foam (ME), and rebonded polyurethane foam (PU).

Table 1 shows the material data of the specimens made of the same materials as shown in Fig. 1, but in the form of plates, whose square shape was associated with their use as walls of acoustic enclosures in the shape of cubes with wall dimensions of  $0.55 \text{ m} \times 0.55 \text{ m} \times 0.55 \text{ m}$  and  $0.7 \text{ m} \times 0.7 \text{ m} \times 0.7 \text{ m}$ .

## 3. Testing methods and test facilities

### 3.1. The airflow resistance

In order to determine the airflow resistance of materials  $R_S$  (ISO, 2020), the Norsonic Nor1517A stand was used, which is the equipment of the laboratory for testing the acoustic properties of materials and structures at the Department of Mechanics and Vibroacoustics of the AGH University of Science and Technology. The tests were carried out using alternating airflow methods (ISO, 2020), using material specimens with a diameter of 100 mm, as shown in Fig. 1. The measurements of the  $R_S$ , were carried out for the atmospheric conditions prevailing in the laboratory, i.e., at a temperature of  $22 \text{ }^\circ\text{C}$  and an atmospheric pressure of 1002 hPa.

Due to the fact that the theoretical calculation models for the  $\alpha_f$ , used in further sections of the article, do not use the  $R_S$  values obtained directly from the measurements, but the airflow resistivity, the values of this parameter were also calculated. The airflow resistivity ( $r$ ) is defined as the airflow resistance ( $R_S$ ) per unit length, which is the specimen thickness ( $h$ ):

$$r = \frac{R_S}{h}. \quad (4)$$

The obtained values of  $R_S$  and  $r$  for the four analysed materials are shown in Table 1.

Table 1. Specimens' material data.

ID	Material	Density [ $\text{kg/m}^3$ ]	The size of the side of a square plate [m]	Plate thickness ( $h$ ) [m]	The airflow resistance ( $R_S$ ) [ $\text{Pa} \cdot \text{s} \cdot \text{m}^{-1}$ ]	The airflow resistivity ( $r$ ) [ $\text{Pa} \cdot \text{s} \cdot \text{m}^{-2}$ ]
MW151	Mineral wool	151.4	0.55	0.06	6417.8	106963.3
MW100	Mineral wool with glass fleece	100.2	0.7	0.05	1811.0	36220.8
ME	Melamine foam	9	0.7	0.05	669.2	13384.8
PU	Rebonded polyurethane foam	214.4	0.55	0.05	2394	47880

### 3.2. Normal incidence sound absorption coefficient and the normal incidence sound transmission loss

The experimental tests were carried out in the laboratory for testing the acoustic properties of materials and structures, in accordance with the relevant standards: (ISO, 2001) for the  $\alpha_f$ , and (ASTM E2611-19, 2019) for the nTL. Both parameters were determined using a Mecanum Inc. impedance tube, which enabled the determination of the values of these parameters in the lower and higher frequency ranges for material specimens of one diameter, 34.9 mm. The measuring apparatus was identical to the tests described in (KOSAŁA, 2021).

The specificity of the laboratory stand used, with a Mecanum Inc. impedance tube, the Siemens LMS SCADAS Mobile analyzer, a computer with Simcenter Testlab software, and type 378A14 PCB measuring microphones, spaced 65 mm or 29 mm apart, respectively, for the low and high frequency ranges, allowed the acoustic parameters to be determined, after averaging the results, in the range of 50 Hz to 5700 Hz.

### 3.3. Insertion loss of enclosure

In the Department of Mechanics and Vibroacoustics, a prototype stand for testing the acoustic properties of baffles and enclosures was developed (KOSAŁA *et al.*, 2020b). So far, single rigid homogeneous baffles of various thicknesses (steel, aluminium, plastics) have been tested on this stand, which, together with an omnidirectional sound source placed inside, imitating a noisy machine or device, constituted a sound-insulating enclosure (KOSAŁA *et al.*, 2020b; KOSAŁA, 2022), as well as two-layer baffles made of a rigid plate and a layer of material in the form of mineral wool, constituting the sound-absorbing and insulating enclosure. This stand allows baffles with a thickness of 1 mm to 90 mm and external dimensions of 0.7 m × 0.7 m to be tested.

Another developed stand, also consisting of a steel frame, enables the testing of baffles with the same thickness range, but with external dimensions of 0.55 m × 0.55 m. The construction of the new stand for testing baffles and enclosures is schematically shown in Fig. 2. Five identical baffles form an acoustic enclosure. As in the case of the former enclosure, an omnidirectional sound source placed centrally was used for acoustic tests (Fig. 2).

The walls of the enclosure are pressed against the steel frame using mechanisms and clamping frames, similarly to the solution described in detail in some works (KOSAŁA *et al.*, 2020b; KOSAŁA, 2022), as shown in Fig. 3, where material made of rebonded polyurethane foam (PU) was used as the walls of the enclosure.

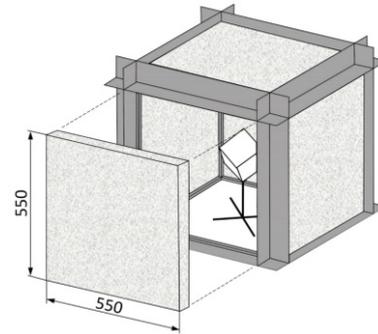


Fig. 2. The scheme of the enclosure frame with a sound source and five walls.

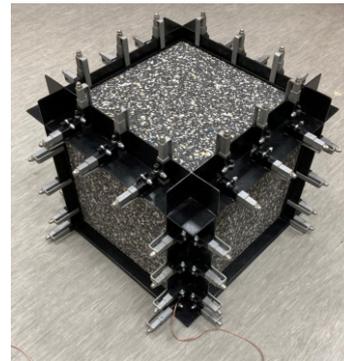


Fig. 3. Enclosure with walls made of PU boards with dimensions of 0.55 m × 0.55 m × 0.05 m.

To test the acoustic properties of single homogeneous materials – plates made of porous and fibrous materials, both stands were used, designed for walls with dimensions of 0.7 m × 0.7 m and 0.55 m × 0.55 m. The purpose of using these two stands was to check whether, using the theoretical computational models described in the article, it is possible to estimate the spectral characteristics of ILs, taking into account enclosures of different dimensions.

In order to determine the effectiveness of an acoustic enclosure built of five identical walls for the four tested materials, the IL of the enclosure was determined using Eq. (3). The tests of the sound power level were carried out in the measurement conditions and with the use of measurement equipment identical to those described in (KOSAŁA, 2022). The sound power levels were determined in a room with a capacity of 79 m<sup>3</sup>, using the survey method in accordance with the standard (ISO, 2010).

## 4. Results of experimental tests

### 4.1. Normal incidence sound absorption coefficient and the normal incidence sound transmission loss

Due to the fact that the frequency range for the IL of the acoustic enclosure made of the tested materials

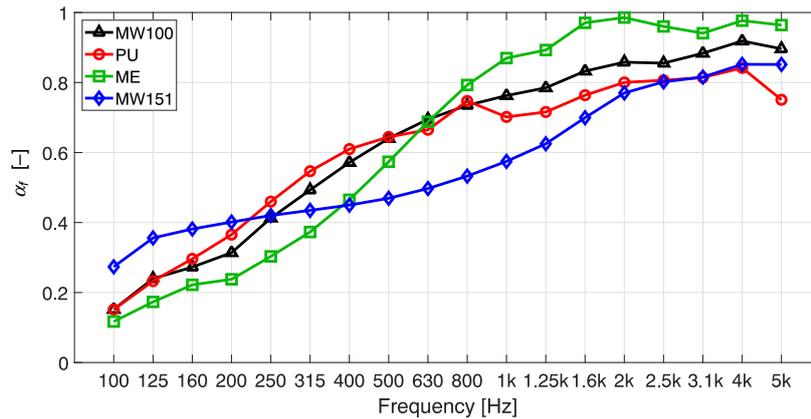


Fig. 4. Normal incidence sound absorption coefficient ( $\alpha_f$ ) of materials: MW100, PU, ME, and MW151.

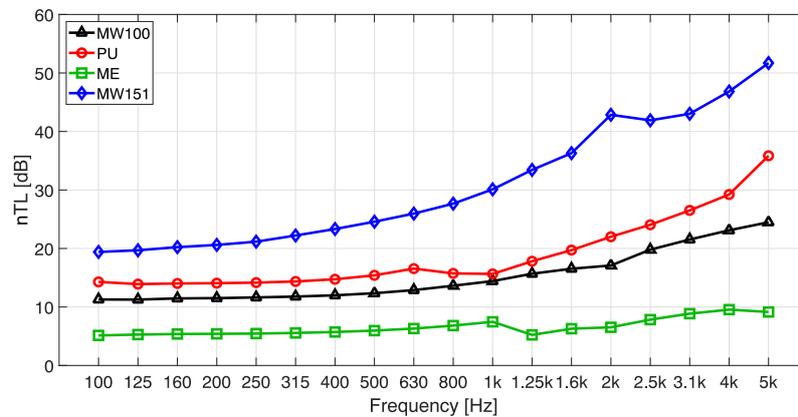


Fig. 5. Normal incidence sound transmission loss (nTL) of materials: MW100, PU, ME, and MW151.

was 100 Hz to 5000 Hz, the values of the  $\alpha_f$  and the nTL were also presented in the  $1/3$  octave frequency bands for the same frequency range as shown in Figs. 4 and 5, respectively.

The highest values of the nTL, amounting to 0.8–0.98, were found in the material with the lowest density ( $9 \text{ kg/m}^3$ ), which was the melamine foam for the frequency range from 800 kHz to 5 kHz (Fig. 4). The mineral wool with the highest density, MW151, had the lowest values of the absorption coefficient among the tested materials, but it was the best at absorbing sounds for low frequencies, 100 Hz to 200 Hz. Usually, increased absorption in the lower frequency range, at the cost of decreased absorption in the higher frequency range, is obtained when the airflow resistivity is too high. For this reason, for mineral wool with the highest airflow resistivity ( $r = 106.9 \text{ kPa} \cdot \text{s} \cdot \text{m}^{-2}$ ), MW151, the shape of the spectral characteristics of the  $\alpha_f$  differs significantly from the characteristics of other materials, as shown in Fig. 4.

As expected, the material with the lowest density (ME) showed the lowest values of the nTL, which for the  $1/3$  octave frequency bands do not exceed 10 dB (Fig. 5). With the increase in frequency, the increase in the nTL for the ME specimen was the smallest

among the tested materials. The tests showed that mineral wool with a density of  $151 \text{ kg/m}^3$  had the best resistance to sound penetration with the perpendicular sound wave incidence on the specimen. The rebonded polyurethane foam with a much higher density ( $214 \text{ kg/m}^3$ ) had weaker sound insulation properties compared to this mineral wool, depending on the frequency of values, by about 5 dB to 15 dB. It should be taken into account here that the rebonded polyurethane foam is a less homogeneous material than wool.

#### 4.2. Insertion loss of the enclosure

Figures 6 and 7 show spectral characteristics of the IL in  $1/3$  octave frequency bands for materials in the form of baffles with dimensions of  $0.7 \text{ m} \times 0.7 \text{ m}$ , which were MW100 and ME, and baffles with dimensions of  $0.55 \text{ m} \times 0.55 \text{ m}$ , which were MW151 and PU.

Comparing the IL of the enclosure with walls made of MW100 and ME plates, dimensions  $0.7 \text{ m} \times 0.7 \text{ m}$  (Fig. 6), it can be seen that the differences are on average about 6 dB, for frequencies lower than 800 Hz, and the range from 4 dB to 8 dB for a given centre frequency of the  $1/3$  octave band. Above a frequency of 800 Hz, the IL difference increases with frequency from

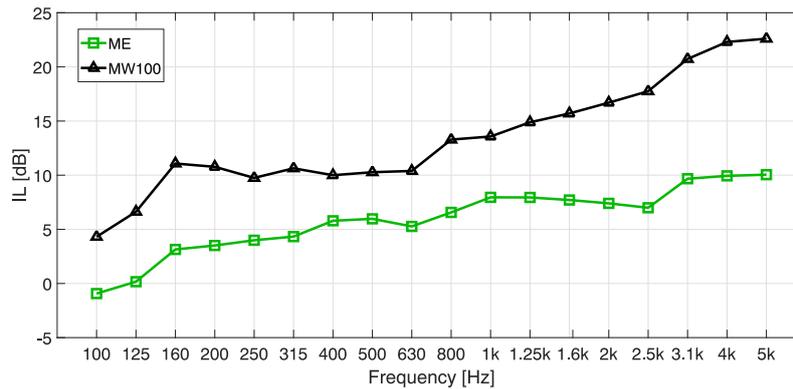


Fig. 6. Insertion loss (IL) of an acoustic enclosure with walls measuring 0.7 m  $\times$  0.7 m made of MW100 and ME materials.

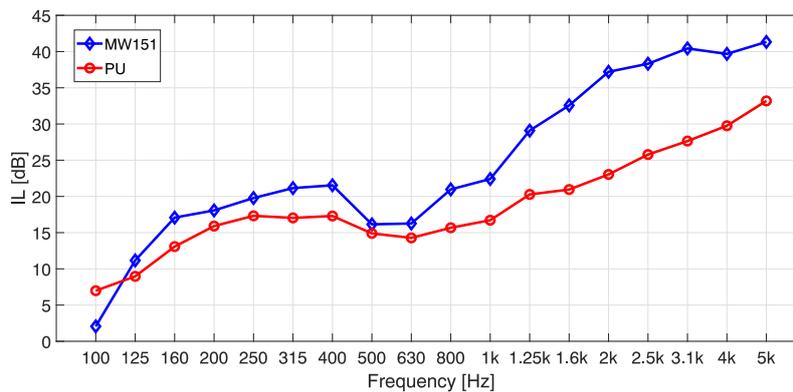


Fig. 7. Insertion loss (IL) of an acoustic enclosure with walls measuring 0.55 m  $\times$  0.55 m made of MW151 and PU materials.

5.6 dB up to 12.5 dB. The average value of the difference for this frequency range is 9.5 dB.

In Fig. 7, which concerns IL for an enclosure with smaller wall dimensions (0.55 m  $\times$  0.55 m), made of MW151 and PU plates, the curve of this parameter has a shape with a visible reduction in the IL value for the frequency of 630 Hz. The reason for this phenomenon may be the resonance of the enclosure cavity, whose impact on the spectral characteristics of IL is much more visible when we are dealing with rigid plates constituting the walls of the sound-insulating enclosures, which was considered in a previous work (KOSAŁA, 2022). The shape of the spectral characteristics shown in Fig. 7 may be influenced by a certain stiffness of the enclosure walls, which causes the tested porous and fibrous materials (PU and MW151) to behave like stiff plates. It can be assumed that the nature of the IL curves for MW151 and PU materials (Fig. 7) is similar only in the frequency range from 100 Hz to 630 Hz. In this frequency range, the differences in IL values for individual centre frequencies are relatively small and average about 3 dB, with a range from about 1 dB to 5 dB. A significantly higher IL value for the MW151 mineral wool compared to the rebounded polyurethane foam PU can be observed above the frequency of 630 Hz. With the exception of the centre frequencies of 4 kHz and 5 kHz, as the frequency

increases, the IL of the MW151 mineral wool also increases, on average by about 10 dB, oscillating from a difference of about 5 dB to about 14 dB.

## 5. Modelling of acoustic properties of porous and fibrous materials

### 5.1. Calculation models of the normal incidence sound absorption coefficient for porous and fibrous materials

As part of the research described in this article, it was checked which of the empirical models such as Qunli, Mechel, Miki, and Delany and Bazley, is best suited for the prediction of the tested porous and fibrous materials. In addition, the Allard and Champoux model, modified by OLIVA and HONGISTO (2013) was also used. This model, among the eight empirical ones tested in (OLIVA, HONGISTO, 2013), achieved the best prediction accuracy, determined by comparing the predicted and measured absorption coefficients of 82 mineral wool configurations.

The results of calculations of the  $\alpha_f$  using empirical models in relation to the results of laboratory tests, described in Subsec. 3.2, are shown in Fig. 8.

In order to determine the discrepancies between the calculations using the models and the results ob-

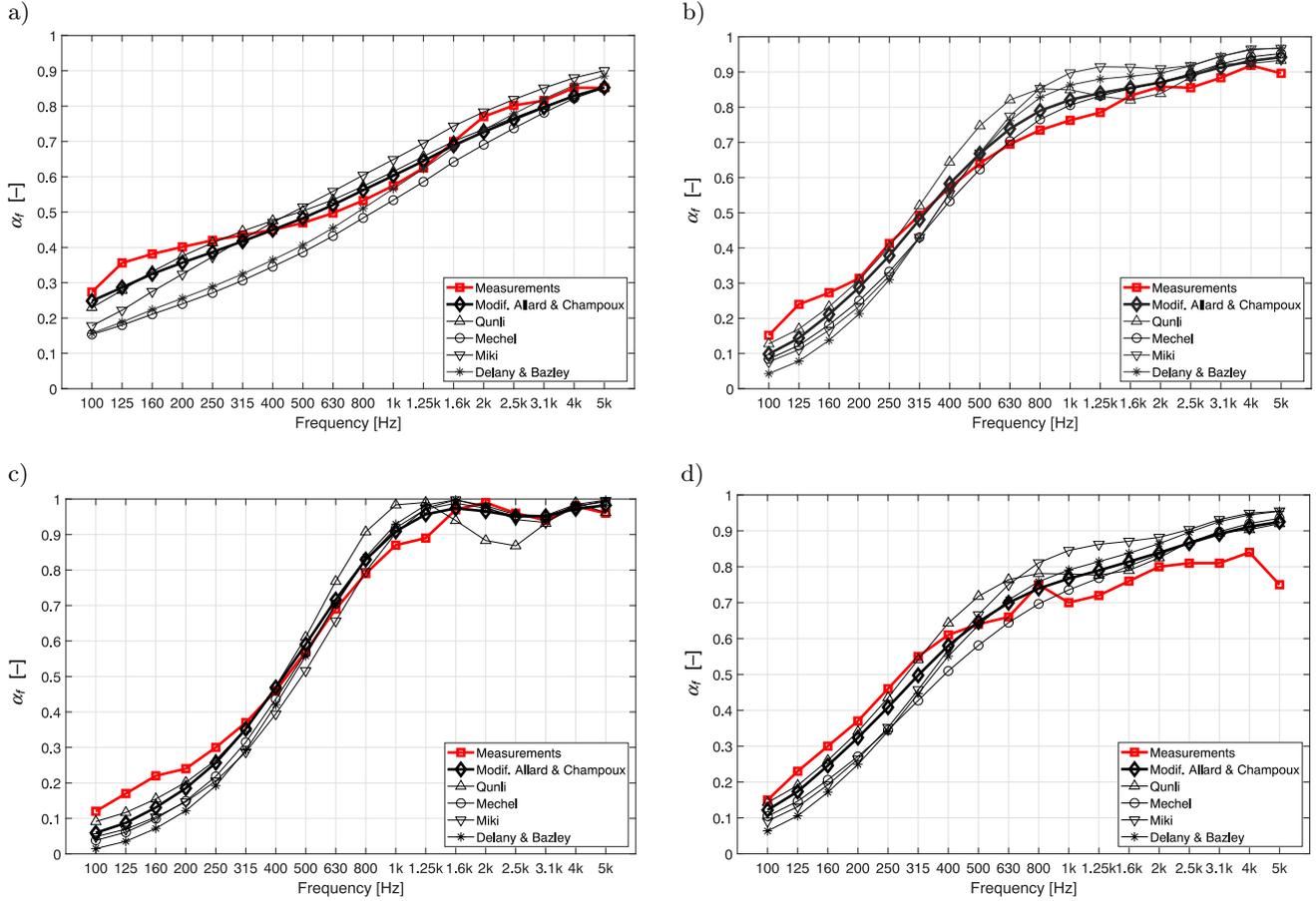


Fig. 8. Normal incidence sound absorption coefficient ( $\alpha_f$ ) determined from measurements and using empirical models for specimens: a) MW151; b) MW100; c) ME; d) PU.

tained from the measurements, root mean square errors (RMSE) were calculated for each tested material using the equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}}, \quad (5)$$

where  $x_i$  and  $\hat{x}_i$  are the values of the measured and predicted acoustic parameter for the  $i$ -th centre frequency of the  $1/3$  octave bands, respectively, and  $N$  is the total number of the centre frequencies of the  $1/3$  octave bands.

The calculated values of RMSE, determining the discrepancies between  $\alpha_f$  calculations using empirical

models and  $\alpha_f$  obtained from measurements, are presented in Table 2.

As can be seen from Table 2, the application of the modified Allard and Champoux model brought about the best results in the form of the lowest RMSE values for all tested specimens, even though the tested materials had different structures – porous or fibrous.

### 5.2. Proposal of a calculation model for normal incidence sound transmission loss for mineral wools

Based on the results of the nTL of 12 material specimens in the form of mineral wool with a thickness of 20 mm, 40 mm, and 50 mm, and a density of

Table 2. RMSE for four tested materials, determining the discrepancies between  $\alpha_f$  calculations using the empirical models and  $\alpha_f$  obtained from measurements.

	RMSE [dB]				
	Modified Allard and Champoux	Qunli	Mechel	Miki	Delany and Bazley
MW151	<b>0.0642</b>	0.0648	0.0845	0.1068	0.1007
MW100	<b>0.0432</b>	0.0647	0.0601	0.0641	0.0755
ME	<b>0.0327</b>	0.0350	0.1007	0.0646	0.0863
PU	<b>0.0442</b>	0.062	0.055	0.086	0.086

25 kg/m<sup>3</sup>, 70 kg/m<sup>3</sup>, 110 kg/m<sup>3</sup>, and 150 kg/m<sup>3</sup> published in (KOSALA, 2021) and, additionally, mineral wool with a thickness of 60 mm and a density of 151.4 kg/m<sup>3</sup> (Fig. 1), it is possible to obtain a calculation model for nTL, which is a function of the frequency  $f$  and the surface density of the mineral wool specimen  $m_p$ .

The nTL spectral characteristics of the tested mineral wool can be determined using a linear function:

$$nTL = a \cdot f + b. \tag{6}$$

The dependence of the slope ( $a$ ) on the surface mass ( $m_p$ ) is shown in Fig. 9, however, dependence of the intercept ( $b$ ) on the surface mass ( $m_p$ ) is shown in Fig. 10.

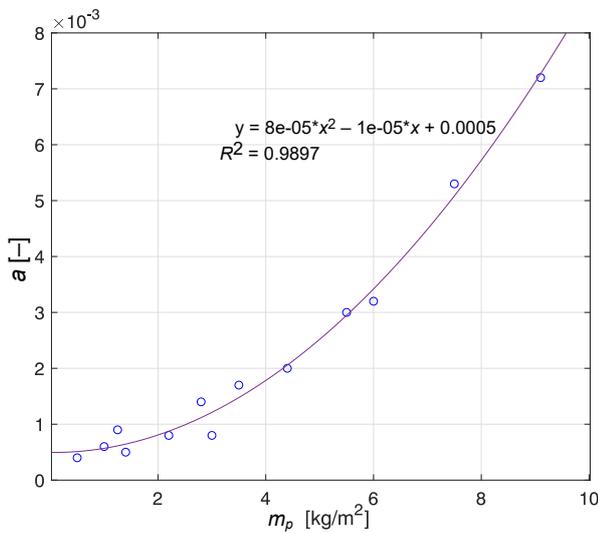


Fig. 9. Dependence of the slope ( $a$ ) on the surface mass ( $m_p$ ) for 13 mineral wool specimens.

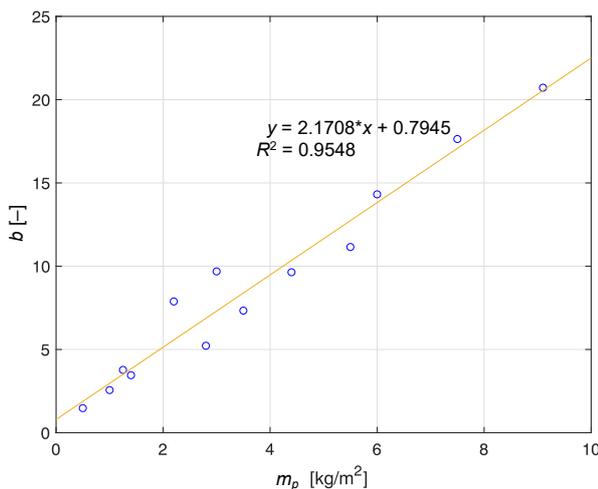


Fig. 10. Dependence of the intercept ( $b$ ) on the surface mass ( $m_p$ ) for 13 mineral wool specimens.

By substituting the polynomial function, shown in Fig. 9, for  $a$  and the linear function, shown in Fig. 10, for  $b$  in Eq. (6), the equation for nTL was obtained:

$$nTL = f(0.00008m_p^2 - 0.00001m_p + 0.0005) + 2.1708m_p + 0.7945 \text{ [dB]}, \tag{7}$$

where  $f$  is the frequency, and  $m_p$  is the surface density of the mineral wool.

The nTL values calculated from Eq. (7) for the 1/3 octave frequency bands in the range of 100 Hz to 5 kHz for 13 mineral wools with a density ( $d$ ) from 25 kg/m<sup>3</sup> to 151 kg/m<sup>3</sup> and a thickness ( $h$ ) from 20 mm to 60 mm are shown in Fig. 11.

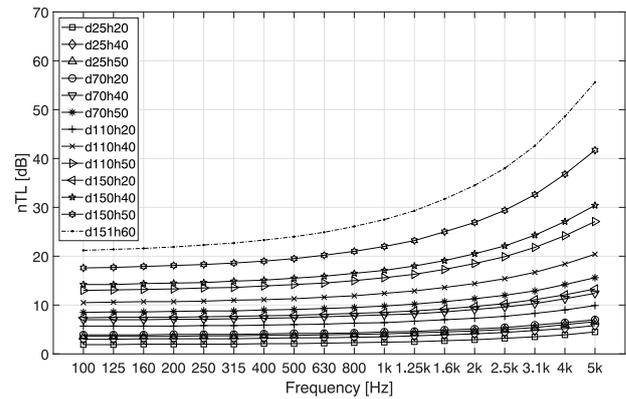


Fig. 11. Normal incidence sound transmission loss values for mineral wools with density  $d$  [kg/m<sup>3</sup>] and thickness  $h$  [mm] calculated on the basis of the surface density of the specimens.

In order to check the accuracy regarding the extent to which the nTL calculations using Eq. (7) differ from the results of experimental tests of this parameter carried out in a previous study (KOSALA, 2022) for 12 mineral wool specimens and in this work (the mineral wool MW151, corresponding to the designation d151h60 in Fig. 11), RMSE was calculated using Eq. (5). Calculations of nTL using the proposed Eq. (7) resulted in small discrepancies for the majority of mineral wool specimens compared to nTL results obtained from measurements carried out in the impedance tube. As shown in Fig. 12, RMSE  $\cong$  1–2 dB.

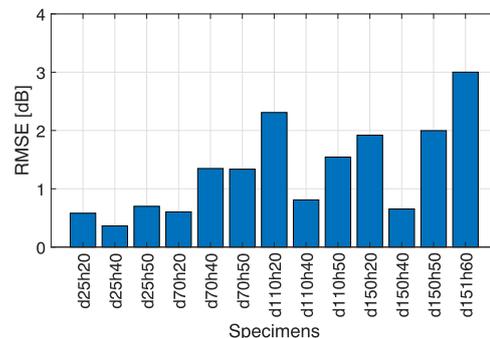


Fig. 12. RMSE for nTL calculations in relation to the values measured for thirteen mineral wool specimens.

The exception was the specimen with the highest density and thickness (d151h60), for which RMSE was equal to 3 dB.

### 5.3. Proposal of a calculation model for IL for a cubic enclosure with walls made of porous and fibrous materials

The IL for enclosures composed of sound-absorbing and insulating walls can be calculated using a known model based on the random incidence sound absorption coefficient of the lining material ( $\alpha$ ) and the sound reduction index ( $R$ ) of the entire panel (enclosure wall) (VER, BERANEK, 2006), using the equation:

$$IL = 10 \log(\alpha) + R \text{ [dB]}. \quad (8)$$

The sound reduction index ( $R$ ) of the sound-absorbing and insulating baffles constituting the enclosure wall can be obtained on the basis of laboratory tests (ISO, 2021), provided that the dimensions of the baffle are adjusted to the dimensions of the measurement window of the coupled reverberation rooms. In an approximate manner,  $R$  can be determined using theoretical calculation models appropriate to the given construction and material solution of the baffle. The validation of the model (Eq. (8)) was shown in the calculations of the IL of the prototype sound-absorbing and insulating enclosure in one work (KOSALA *et al.*, 2020b).

To calculate the IL for a cubic enclosure of an omnidirectional sound source, the walls of which are single baffles made of porous or fibrous materials, a similar formula was proposed, in which the simplification in the form of the nTL of the baffle is used instead of  $R$ :

$$IL = 10 \log(\alpha) + nTL \text{ [dB]}. \quad (9)$$

In the case of the porous and fibrous materials in question, most of which can be roughly described as locally reacting materials (ALLARD, 1992; COX, D'ANTONIO, 2017), the attenuation of sound in the material is so high that it limits the lateral transmission. The propagation direction within these materials is normal to the surface, even for oblique incidence

sound, because of refraction (COX, D'ANTONIO, 2017), therefore the surface impedance is independent of the incident wave nature. Hence, in Eq. (9), a simplification was made by replacing  $R$  with nTL.

The normal incidence sound transmission loss of such materials can be determined by testing a sample in an impedance tube, as described in Subsec. 3.2, or using a calculation model for a material of infinite extent in the lateral direction, since the boundary conditions resulting from the dimensions of the enclosure wall in the case of the materials in question are less important. Such a model for calculating nTL, based on surface mass, was proposed using Eq. (7).

The random incidence sound absorption coefficient ( $\alpha$ ) is determined according to the standard (ISO, 2005) in laboratory reverberation conditions on a material specimen of 10–12 m<sup>2</sup>. However,  $\alpha$  can also be determined using the approximate relationship between this coefficient and the  $\alpha'_f$ , as given in (EVEREST *et al.*, 2013) in the shape of a graph. This relationship can be calculated using approximation by a second-degree polynomial with the coefficient of determination equal to  $R^2 = 0.9994$  from the equation:

$$\alpha = -0.97(\alpha'_f)^2 + 1.97\alpha'_f. \quad (10)$$

The normal incidence sound absorption coefficient ( $\alpha_f$ ) can be determined by measurement using an impedance tube, as described in Subsec. 3.2, or by using theoretical calculation models, as described in Subsec. 5.1.

#### 5.3.1. Validation of the proposed model for porous and fibrous materials

The validation of the proposed model (Eq. (9)) was carried out based on the results of experimental tests of  $\alpha_f$  and nTL, carried out for the tested material specimens with the use of an impedance tube. The test results of these parameters are presented in Subsec. 4.1. The values of the ( $\alpha$ ) coefficient were calculated on the basis of the  $\alpha_f$  values from Eq. (10). Figures 13–16

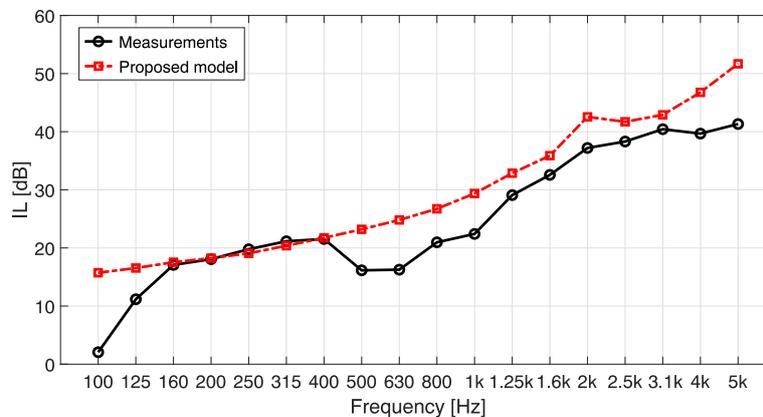


Fig. 13. Insertion loss (IL) of the enclosure made of MW151 baffles obtained from measurements and calculations using the proposed model.

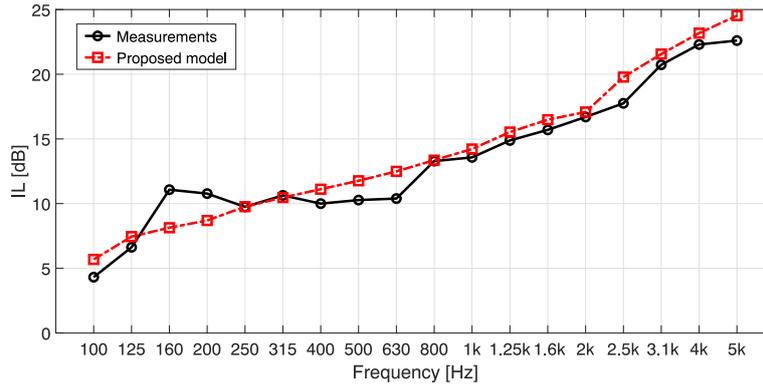


Fig. 14. Insertion loss (IL) of the enclosure made of MW100 baffles obtained from measurements and calculations using the proposed model.

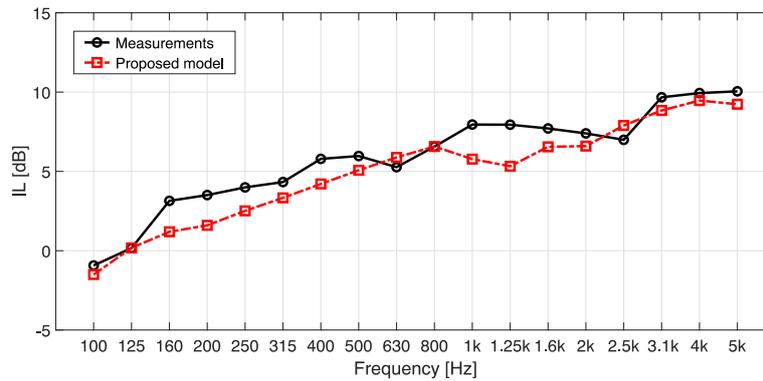


Fig. 15. Insertion loss (IL) of the enclosure made of ME baffles obtained from measurements and calculations using the proposed model.

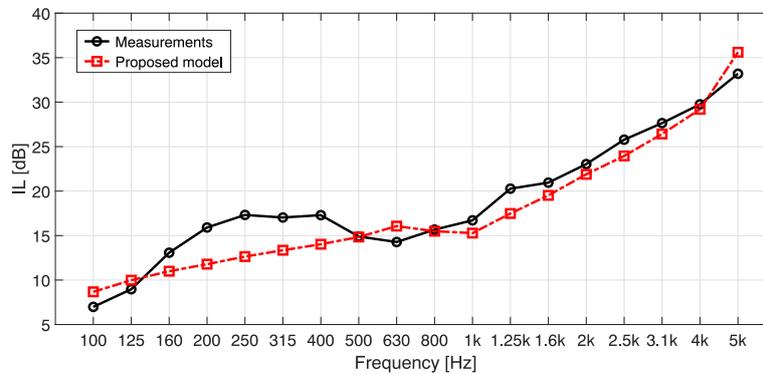


Fig. 16. Insertion loss (IL) of the enclosure made of PU baffles obtained from measurements and calculations using the proposed model.

show the IL values of enclosures with walls made of MW151, MW100, ME, and PU baffles, obtained from measurements and using the proposed IL calculation model, defined by Eq. (9).

Table 3 shows the calculated values of Pearson’s linear correlation coefficient (PCC) and RMSE, which were used to estimate the discrepancies of the results obtained from the IL calculations for the four tested materials using the proposed model (Eq. (9)) and those obtained from measurements.

For all tested materials, there is a very strong linear correlation between the results obtained from the pro-

Table 3. PCC and RMSE for four tested materials, determining the discrepancies between IL calculations using the proposed model and IL obtained from measurements.

	PCC	RMSE [dB]
MW151	0.9392	6.01
MW100	0.9734	1.38
ME	0.9551	1.30
PU	0.9581	2.34

posed model and those obtained from measurements. For the MW151 specimen, the highest RMSE value

was obtained among all specimens, which was mainly due to large discrepancies at the lowest ( $f = 100$  Hz) and highest centre frequencies of the  $1/3$  octave bands ( $f = 5$  kHz).

5.3.2. Validation of the proposed model for mineral wool specimens

Due to the wider scope of research on the acoustic properties of mineral wool materials (KOSALA, 2021), compared to porous materials, it is possible to calculate IL, defined by Eq. (9), without the need to use an impedance tube. The knowledge of the material data is sufficient to estimate the input parameters for calculating IL, i.e.,  $\alpha$  and nTL.

The first needed parameter  $\alpha$  is calculated on the basis of  $\alpha_f$  from Eq. (10), while  $\alpha_f$  can be determined using one of the known calculation models when the  $r$  value of mineral wool of a given thickness is available. For this purpose, it is proposed to use the model developed by Allard and Champoux, modified

by OLIVA and HONGISTO (2013) into an empirical model, which is much easier to use than the original semi-phenomenological model, because there is no need to determine the necessary properties of microscopic material. The second parameter, nTL, can be determined from the  $m_p$  of the mineral wool plate constituting the wall of the enclosure, using Eq. (7), as proposed in Subsec. 5.2.

The validation of the proposed model of IL (Eq. (7)) for enclosures with walls made of mineral wool plates, based on material data, was carried out using MW151 with the dimensions of  $0.55 \text{ m} \times 0.55 \text{ m} \times 0.06 \text{ m}$  and MW100 with dimensions of  $0.7 \text{ m} \times 0.7 \text{ m} \times 0.05 \text{ m}$ . The IL values calculated and obtained from measurements in  $1/3$  octave frequency bands are shown in Figs. 17 and 18.

Table 4 shows a comparison of the application of the proposed calculation model of IL, which was verified for fibrous materials in the form of MW151 and MW100 for two cases. The first concerned the use of

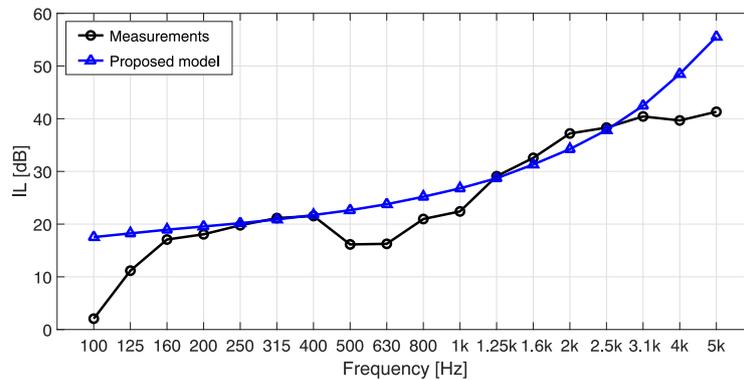


Fig. 17. Insertion loss (IL) of the enclosure made of MW151 (wall size  $0.55 \text{ m} \times 0.55 \text{ m} \times 0.06 \text{ m}$ ) determined from measurements and calculations using the proposed model based on material data.

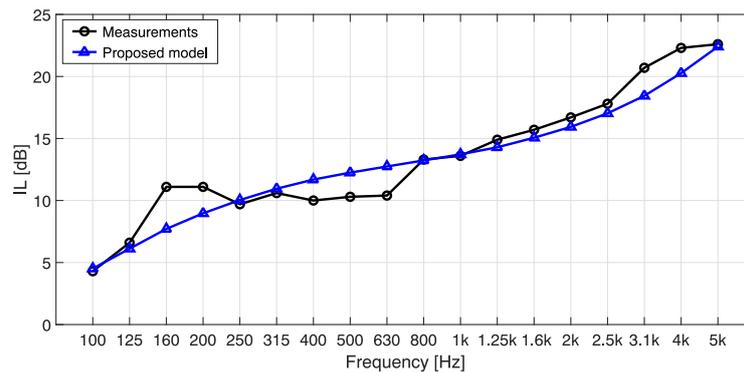


Fig. 18. Insertion loss (IL) of the enclosure made of MW100 (wall size  $0.7 \text{ m} \times 0.7 \text{ m} \times 0.05 \text{ m}$ ) determined from measurements and calculations using the proposed model based on material data.

Table 4. PCC and RMSE for enclosures with walls made of mineral wools, determining discrepancies between IL calculated and measured for  $\alpha_f$  and nTL obtained with the use of an impedance tube and calculation models.

	Method of determining $\alpha_f$ and nTL			
	Impedance tube		Calculation models	
	PCC	RMSE [dB]	PCC	RMSE [dB]
MW151	0.9392	6.01	0.8918	6.35
MW100	0.9734	1.38	0.9574	1.49

an impedance tube to determine  $\alpha_f$  and nTL of the tested mineral wools. In the second case,  $\alpha_f$  and nTL were obtained using the calculation models, the Allard and Champoux, and those proposed by Eq. (7), respectively.

As shown in Table 4, in the case of MW151, the IL discrepancies in the form of RMSE obtained using the calculation models for  $\alpha_f$  and nTL are about 6.3 dB and are similar to the case when these parameters were obtained from the impedance tube. For MW100, the obtained linear correlation coefficients have similarly high values for both cases. For the two analysed methods of determining  $\alpha_f$  and nTL needed for IL calculations, the discrepancies in the form of RMSE are similar and relatively low (around 1.5 dB).

## 6. Conclusions

As part of the research, two developed stands were used to enable the construction of cubic enclosures made of walls of the tested materials, intended for determining the acoustic efficiency from the penetration of sounds coming from a source placed inside. The measure of acoustic efficiency was the IL of the enclosure. The research has shown that it is possible, based on known empirical models for the sound absorption coefficient, the airflow resistance and the thickness of material specimens, to calculate the IL of an enclosure with walls made of a fibrous material, such as mineral wool. In the case of materials with a different structure, such as porous ones, for example polyurethane foam, to calculate the IL using the proposed model, it is necessary to use an impedance tube to determine the input data. This data is in the form of the  $\alpha_f$  and the normal incidence transmission loss of the material. The results of calculations of the spectral characteristics of the IL, satisfactory for engineering purposes, were obtained for the four tested materials. The study also showed that among the empirical models such as Delany and Bazley, Miki, Mechel, Qunli, and Allard and Champoux, the latter model yielded the best results in calculating the  $\alpha_f$  for mineral wools, melamine foam and rebonded polyurethane foam, in the form of the smallest discrepancies obtained compared to the results obtained from laboratory tests.

The second proposed calculation model for the nTL, using the surface density of a mineral wool specimen, can be used not only for calculations of the IL of the enclosure, but also for the initial estimation of nTL for mineral wools, without the use of an impedance tube. The results of the experimental and model tests obtained as part of the article in the form of acoustic parameters determining the acoustic properties of single porous and fibrous materials can also be used in further research on the modelling of layered baffles.

## Acknowledgments

This research was funded by a research subvention supported by the Polish Ministry of Science and Higher Education (grant no. 16.16.130.942).

## References

1. ALLARD J.F., CHAMPOUX Y. (1992), New empirical equations for sound propagation in rigid frame fibrous materials, *The Journal of the Acoustical Society of America*, **91**(6): 3346–3353, doi: [10.1121/1.402824](https://doi.org/10.1121/1.402824).
2. ASTM E2611-19 (2019), *Normal Incidence Determination of Porous Material Acoustical Properties Based on the Transfer Matrix Method*, ASTM International.
3. BATKO W., PAWLIK P., WSZOLEK G. (2017), Sensitivity analysis of the estimation of the single-number sound absorption evaluation index  $DL\alpha$ , *Archives of Acoustics*, **42**(4): 689–696, doi: [10.1515/aoa-2017-0071](https://doi.org/10.1515/aoa-2017-0071).
4. BERARDI U., IANNACE G. (2015), Acoustic characterization of natural fibers for sound absorption applications, *Building and Environment*, **94**(2): 840–852, doi: [10.1016/j.buildenv.2015.05.029](https://doi.org/10.1016/j.buildenv.2015.05.029).
5. BIES D.A., HANSEN C.H. (2009), *Engineering Noise Control. Theory and Practice*, 4th ed., p. 675, CRC Press Taylor & Francis Group.
6. COX T.J., D'ANTONIO P. (2017), *Acoustic Absorbers and Diffusers. Theory, Design and Application*, 3rd ed., CRC Press Taylor & Francis Group.
7. DAVY J.L. (2009), Predicting the sound transmission loss of cavity walls, [in:] *Proceedings of Interior Noise Climates: National Conference of the Australian Acoustical Society*, p. 1–16, Perth.
8. DELANY M.E., BAZLEY E.N. (1970), Acoustical properties of fibrous absorbent materials, *Applied Acoustics*, **3**(2): 105–116, doi: [10.1016/0003-682X\(70\)90031-9](https://doi.org/10.1016/0003-682X(70)90031-9).
9. ENGEL Z., SIKORA J. (1997), *Sound-Absorbing and Insulating Enclosures. The Basics of Design and Application* [in Polish: *Obudowy Dźwiękochłonna-Izolacyjne. Podstawy Projektowania i Stosowania*], Wydawnictwa AGH, Kraków.
10. EVEREST F.A., POHLMANN K.C., KURYLAK W. (2013), *Acoustic Handbook* [in Polish: *Podręcznik Akustyki*], Wydawnictwo Sonia Draga, Katowice.
11. FIEBIG W., DĄBROWSKI D. (2020), Use of acoustic camera for noise sources localization and noise reduction in the industrial plant, *Archives of Acoustics*, **45**(1): 111–117, doi: [10.24425/aoa.2020.132487](https://doi.org/10.24425/aoa.2020.132487).
12. International Organization for Standardization (2001), *Acoustics, Determination of sound absorption coefficient and impedance in impedance tubes. Part 2 Transfer-function method*, (ISO Standard No. ISO 11534-2).
13. International Organization for Standardization (2005), *Acoustics. Measurements of sound absorption in a reverberation room* (ISO Standard No. 354).

14. International Organization for Standardization (2009), *Acoustics, Determination of sound insulation performances of enclosures. Part 1: Measurements under laboratory condition (for declaration purposes)* (ISO Standard No. ISO 11546-1).
15. International Organization for Standardization (2010), *Acoustics, Determination of sound power levels and sound energy levels of noise sources using sound pressure, Survey method using an enveloping measurement surface over a reflecting plane* (ISO Standard No. 3746).
16. International Organization for Standardization (2020), *Acoustics, Determination of the airflow resistance, Part 2: Alternating airflow method* (ISO Standard No. ISO 9053-2), <https://www.iso.org/standard/76744.html>.
17. International Organization for Standardization (2021), *Acoustics, Laboratory measurement of sound insulation of building elements. Part 2: Measurement of airborne sound insulation* (ISO Standard No. ISO 10140-2), <https://www.iso.org/standard/79487.html>.
18. JIANG Y., CHEN S., WANG D., CHEN J. (2017), Multi-objective optimization of acoustical properties of PU-bamboo-chips foam composites, *Archives of Acoustics*, **42**(4): 707–714, doi: [10.1515/aoa-2017-0073](https://doi.org/10.1515/aoa-2017-0073).
19. KOSALA K. (2021), Experimental tests of the acoustic properties of sound-absorbing lining and cores of layered baffles, *Vibrations in Physical Systems*, **32**(1): 2021107, doi: [10.21008/j.0860-6897.2021.1.07](https://doi.org/10.21008/j.0860-6897.2021.1.07).
20. KOSALA K. (2022), Experimental tests and prediction of insertion loss for cubical sound insulating enclosures with single homogeneous walls, *Applied Acoustics*, **197**: 108956, doi: [10.1016/j.apacoust.2022.108956](https://doi.org/10.1016/j.apacoust.2022.108956).
21. KOSALA K., MAJKUT L., MLECZKO D., OLSZEWSKI R., WSZOŁEK T. (2020a), Accuracy of prediction methods for sound insulation of homogeneous single baffles, *Vibration in Physical Systems*, **31**(2): 2020210, doi: [10.21008/j.0860-6897.2020.2.10](https://doi.org/10.21008/j.0860-6897.2020.2.10).
22. KOSALA K., MAJKUT L., OLSZEWSKI R. (2020b), Experimental study and prediction of insertion loss of acoustical enclosures, *Vibrations in Physical Systems*, **31**(2): 2020209, doi: [10.21008/j.0860-6897.2020.2.09](https://doi.org/10.21008/j.0860-6897.2020.2.09).
23. KUNIO J., YOO T., JOU K., BOLTON J.S., ENOK J. (2009), A comparison of two and four microphone standing wave tube procedures for estimating the normal incidence absorption coefficient, [in:] *Proceedings of 38th International Congress and Exposition on Noise Control Engineering*, pp. 1057–1065.
24. MECHEL F.P. (1988), Design charts for sound absorber layers, *The Journal of the Acoustical Society of America*, **83**(3): 1002–1013, doi: [10.1121/1.396045](https://doi.org/10.1121/1.396045).
25. MIKI Y. (1990), Acoustical properties of porous materials – modifications of Delany–Bazley models, *Journal of the Acoustical Society of Japan (E)*, **11**(1): 19–24, doi: [10.1250/ast.11.19](https://doi.org/10.1250/ast.11.19).
26. MORZYŃSKI L., SZCZEPAŃSKI G. (2018), Double panel structure for active control of noise transmission, *Archives of Acoustics*, **43**(4): 689–696, doi: [10.24425/aoa.2018.125162](https://doi.org/10.24425/aoa.2018.125162).
27. NIERADKA P., DOBRUCKI A. (2018), Insertion loss of enclosures with lined slits, [in:] *Proceedings of 11th Euronoise 2018*, pp. 893–898.
28. NURZYŃSKI J. (2022), Sound insulation of lightweight external frame walls and the acoustic effect of additional thermal insulation, *Applied Acoustics*, **190**(15): 08645, doi: [10.1016/j.apacoust.2022.108645](https://doi.org/10.1016/j.apacoust.2022.108645).
29. OLIVA D., HONGISTO V. (2013), Sound absorption of porous materials – Accuracy of prediction methods, *Applied Acoustics*, **74**(12): 1473–1479, doi: [10.1016/j.apacoust.2013.06.004](https://doi.org/10.1016/j.apacoust.2013.06.004).
30. PUTRA A., KHAIR F.A., NOR M.J.M. (2015), Utilizing hollow-structured bamboo as natural sound absorber, *Archives of Acoustics*, **40**(4): 601–608, doi: [10.1515/aoa-2015-0060](https://doi.org/10.1515/aoa-2015-0060).
31. QUNLI W. (1988), Empirical relations between acoustical properties and flow resistivity of porous plastic open-cell foam, *Applied Acoustics*, **25**(3): 141–148, doi: [10.1016/0003-682X\(88\)90090-4](https://doi.org/10.1016/0003-682X(88)90090-4).
32. SHARP B.H. (1973), *A study of techniques to increase the sound installation of building elements*, US Department of Commerce, National Technical Information Service.
33. TRINH V.H., NGUYEN T.V., NGUYEN T.H.N., NGUYEN M.T. (2022), Design of sound absorbers based on open-cell foams via microstructure-based modeling, *Archives of Acoustics*, **47**(4): 501–512, doi: [10.24425/aoa.2022.142894](https://doi.org/10.24425/aoa.2022.142894).
34. VER I.L., BERANEK L.L. (2006), *Noise and Vibration Control Engineering: Principles and Applications*, John Wiley & Sons, Inc., Hoboken, New Jersey.



## Research Paper

Sound Insulation Properties of Sound-Reduction Louvers  
with Innovative Devulcanized Rubber

Tomas VILNIŠKIS\*, Tomas JANUŠEVIČIUS

*Department of Environmental Protection and Water Engineering  
Vilnius Gediminas Technical University  
Vilnius, Lithuania*\*Corresponding Author e-mail: [tomas.vilniskis@vgtu.lt](mailto:tomas.vilniskis@vgtu.lt)*(received June 7, 2021; accepted January 22, 2024; published online April 5, 2024)*

The growing amount of used tires presents many environmental challenges around the world. Therefore, new ways to reuse used tires are being sought. One of the uses of waste tires is in sound-insulating constructions. Waste tires can be shredded into granules, which can be further devulcanized to increase their porosity. These granules can then be glued to panels and used in sound-insulating structures. Acoustic louvers were investigated in this study, with the louvers' plates covered with rubber granule panels. Sound absorption parameters of rubber granule panels were tested across frequency bands ranging from 160 Hz to 5000 Hz. The results showed the normal incidence sound absorption coefficient reached 0.87–0.96 at 3150 Hz for 12 mm rubber granule plates. Measurements were conducted in a semi-anechoic chamber. The study has shown that rubber louvers can reduce the sound pressure level by 8 dB–12 dB, depending on the composite of the rubber granule panels and the tilt angle of the louvers' plates.

**Keywords:** acoustical louvers; rubber; sound absorption; insertion loss.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Globally, approximately 0.7–1 billion new tires are produced each year due to the growing number of vehicles. Today, several approaches are taken to address the end-of-life tire (ELT) problem: reuse (5 %–23 %), recycling (3 %–15 %), and recovery (25 %–60 %) (DISANAYAKE *et al.*, 2021). However, more than 30 % of used tires still end up in landfills or are discarded as untreated waste (KARAKURT, 2015). Due to increasing environmental pollution and the European Union's directives banning the disposal of tires in landfills, there is a growing interest in finding ways to reuse or recycle tires to achieve circular economy goals (NACIF *et al.*, 2013; CORREDOR-BEDOYA *et al.*, 2017).

The accumulation of untreated tires, a non-biodegradable substance, is a major concern in many countries. Long-term storage of tires in landfills affects ecosystems, pollutes the environment, and creates a high risk of fire (GÜNEYISI, 2010; SIDDIQUE, NAIK, 2004). Recycled rubber has good acoustic properties due to its porous structure and can therefore be

recycled and used in construction materials, industry and various structures. This helps in protecting land resources and maintaining ecological balance (GANDOMAN, KOKABI, 2015).

Noise is considered to be one of the most important elements of environmental pollution, caused by different sources: cars, air transport, and industrial machinery (DISSANAYAKE *et al.*, 2021). Noise does not only cause inconvenience in everyday life but is also harmful to our health. Noise is a form of energy that travels through solid bodies, liquids and gases, forcing particles to vibrate in longitudinal waves. To reduce the environmental impact of used tires, ways to recycle them efficiently and use them in production are being pursued. Granules, steel wires, and tire fibers are typical materials derived from ELT treatment (LANDI *et al.*, 2018). One of the most popular methods is the incorporation of rubber granules into concrete mixes, which allows to develop new cement products for the construction of civil buildings. However, such composites are found to be weaker than pure concrete (KARAKURT, 2015). Nevertheless, according

to the researchers, different strength characteristics of the products can be obtained by using different rubber granules. Products with smaller particle granules have better strength properties than products with larger granules (SU *et al.*, 2015). Studies show that cement incorporating rubber granules absorbs sound more efficiently than conventional cement products, due to its higher porosity (KHALOO *et al.*, 2008).

There are many researchers who investigated the acoustic properties of rubber, since good sound attenuation of rubber is one of its characteristic advantages over many other materials available for acoustic applications. SWIFT *et al.* (1999) investigated noise barriers made of recycled rubber granulates, finding that the adhesive part of the composite does in fact reduce sound absorption. Low-frequency acoustic properties of honeycomb silicone rubber acoustic metamaterials were investigated by GAO and HOU (2017). MADERUELO-SANZ *et al.* (2012) developed a sound absorber by using waste tire rubber. Further investigations were conducted on acoustic metamaterials combining waste rubber (fibers or particles from waste tires or other products) with other substances such as plant flours or fibers, polypropylene or polyethylene, where rubber acts as an acoustic reinforcement unit (XU *et al.*, 2018). Rubber granulometry has been found to influence sound-absorbing behaviours. BUJOREANU *et al.* (2017) investigated an experimental acoustic system with rubber particles by additionally using different backing plates including plasterboard, OSB, and polystyrene. Segura-Alcaraz investigated rubber-fiber-rubber layered construction panels as sound absorbers by using rubber from scrap tires and recycled fibers (JULIÁ *et al.*, 2013). HONG *et al.* (2007) mixed waste rubber particles with polymer porous foams. KOSALA (2019) investigated the sound insulation properties of two-layer baffles consisting of rubber and steel, which exhibited good sound insulation properties.

Sound-absorbing materials can also be used in sound insulation constructions. VIVEIROS and GIBBS (2003) investigated the performance of acoustic louvers and found that, at frequencies where transmission is greatly unaffected by absorption (in this case below 1 kHz), the impulse measurement of transmission loss agrees well with insertion loss. This occurs due to the fact that, in this frequency region, insertion loss remains unaffected by the geometry of the louver and the angle of transmission. However, at frequencies where absorption and interference effects come into play (above 1 kHz), insertion loss becomes angularly dependent. MARRIOTT (2012) concluded that the maximum sound insulation efficiency of acoustic blinds is achieved at high frequencies, potentially reaching up to 10 dB–12 dB.

In this work, we focus on innovative acoustic panels made of rubber granules that have good sound-absorbing properties. Recycling waste tires through

mechanical and chemical devulcanization processes increases the porosity of the rubber granules, thereby improving their sound absorption properties. The granules are glued together with an innovative two-component polyurethane glue, enabling the creation of panels that can be affixed to the structure. Combining them with steel plates allows to make sound-absorbing plates, which can be used for sound absorbing and insulating barriers such as acoustic louvers. The aim of our work is to use such panels in constructions designed to reduce the noise generated by engineering equipment that requires ventilation. By covering gaps in the structure with rubber granulate panels, we anticipate achieving effective acoustic absorption and insulation.

## 2. Methodology

### 2.1. Research object

In this research, acoustic louvers consisting of seven metal steel plates covered on both sides with different composite 12 mm thick rubber panels were made and fixed onto a wooden frame (with adjustable panels angle) (Fig. 1). This louver design was intended for equipment that requires ventilation and air exhaust. The gaps between the panels allow for air to be removed, while also reducing the sound level when the panels are covered with sound-absorbing material. The structure was designed to have hydrophobic properties, be heat resistant, and absorb sound.

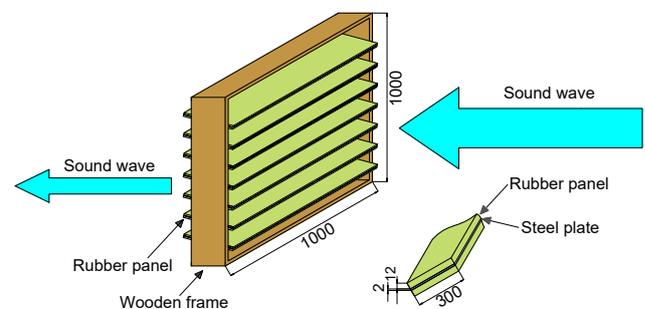


Fig. 1. Principal construction of acoustic louvers.

The plates were angled from  $0^\circ$  (horizontal) to  $45^\circ$  towards and away from the noise source (Fig. 2) in increments of  $15^\circ$  steps for each test angle increase.

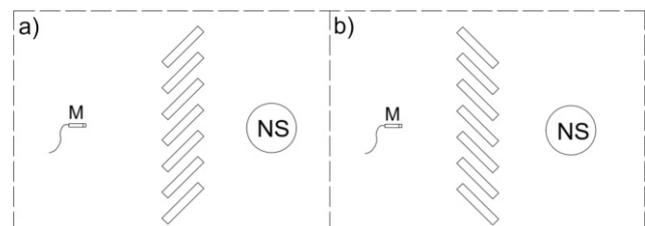


Fig. 2. Panels tilted in the following way: a) tilted towards the noise source; b) tilted away from the noise source.

## 2.2. Tested materials

Rubber granule panels were made from rubber granules obtained through ozonation after the rubber was separated from the tire structure. Two types of rubber granules were used for the production of rubber panels. The first type comprised mechanically removed primary tire tread, with the rubber granule size ranging from 4 mm to 15 mm (Fig. 3a). The second type consisted of mechanically removed and chemically treated tread layer, with the rubber granule size ranging from 1 mm to 3 mm (Fig. 3b).

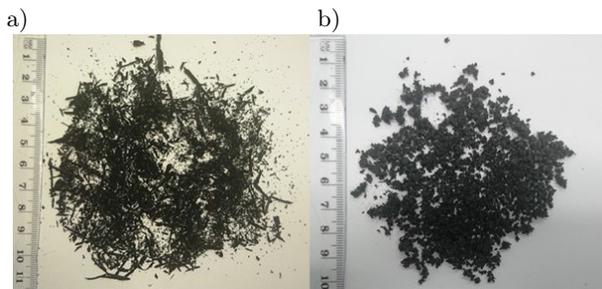


Fig. 3. a) Mechanically removed rubber; b) chemically devulcanized rubber.

The chemical treatment of the rubber granules resulted in a higher porosity and a partially fibrous structure, thereby producing higher sound absorption. The difference between mechanically and chemically devulcanized rubber granules is shown in Fig. 4.

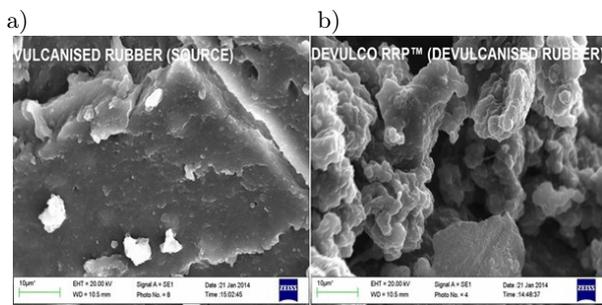


Fig. 4. a) Mechanically removed rubber; b) chemically devulcanized rubber.

During the chemical devulcanization of rubber (converting it into devulcanized rubber flour), mechanical shear causes stresses in the bridges (S–S)

between the rubber chains, while devulcanizing agent compounds promote the delocalization of these bridges (S–S), which results in the inhibition of rubber particle adhesion. This, in turn, results in a higher surface porosity of the rubber granules. The mechanically separated rubber granules form a straw-like structure, serving a reinforcing function, and corresponding to the properties of fibrous material in the rubber plate. Three types of rubber composite plates were produced using two types of rubber granules (Fig. 5). Plate no. 1 was made of mechanically removed rubber granules with a fibrous structure from elongated rubber granules. Plate no. 2 was made from chemically devulcanized rubber granules. Plate no. 3 was made by mixing both types of rubber granules in equal proportions. All three plates were bonded together using a special polymer glue.

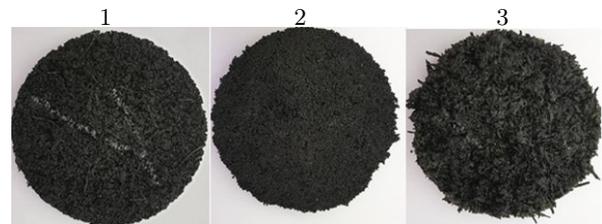


Fig. 5. Rubber granulate panels of different mixtures.

## 2.3. Sound absorption research methodology

Study of the sound absorption properties of rubber panels was carried out in the impedance tube according to International Organization for Standardization [ISO] (1998). The diameter of the impedance tube was 30 mm. Tests were carried out on round samples of 30 mm in diameter and 12 mm in thickness, placed in an impedance tube. Three microphones were used simultaneously to calculate two transfer functions  $H_{13}$  and  $H_{23}$  for measuring sound absorption properties, with the samples being rigidly backed. This method allowed to measure the normal incidence sound absorption coefficient from 160 Hz to 5000 Hz. The schematic experimental setup is shown in Fig. 6. The distance between microphone no. 1 and no. 2  $X_{12} = 120$  mm, between microphone no. 2 and no. 3  $X_{23} = 20$  mm, and the distance from the closest microphone to the sample  $X_{3S} = 60$  mm.

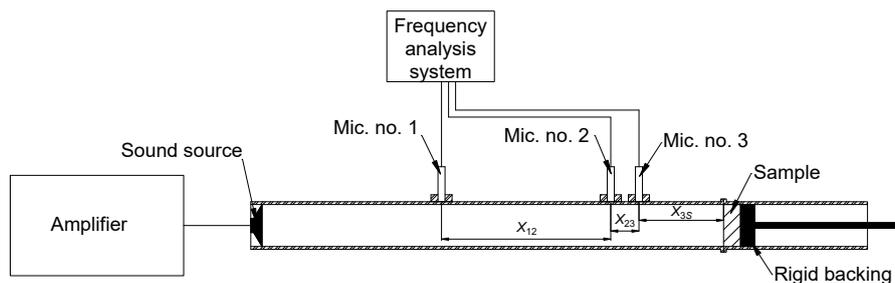


Fig. 6. Schematic experimental setup of impedance tube.

A sample was placed in the tube according to the procedures specified in (ISO, 1998). The transfer functions  $H_{13}$  and  $H_{23}$  in Eq. (1) between microphone positions were calculated as the pressure ratio between both microphones. The complex acoustic transmission function for the incident wave  $H_I$  and the reflected wave  $H_R$  was calculated according to Eqs. (2) and (3) (ISO, 1998):

$$H_{13} = \frac{p_3}{p_1}, \quad H_{23} = \frac{p_3}{p_2}, \quad (1)$$

$$H_{I(160-1000 \text{ Hz})} = \frac{p_{3I}}{p_{1I}} = e^{-jk_0(x_{12}+x_{23})}, \quad (2)$$

$$H_{I(1-5 \text{ kHz})} = \frac{p_{3I}}{p_{2I}} = e^{-jk_0(x_{23})},$$

$$H_{R(160-1000 \text{ Hz})} = \frac{p_{3R}}{p_{1R}} = e^{jk_0(x_{12}+x_{23})}, \quad (3)$$

$$H_{R(1-5 \text{ kHz})} = \frac{p_{3R}}{p_{2R}} = e^{jk_0(x_{23})}.$$

From Eqs. (2) and (3), the reflection coefficient was calculated according to Eqs. (4) and (5) (ISO, 1998):

$$R_{(160-1000 \text{ Hz})} = \frac{H_{13} - H_{I(160-1000 \text{ Hz})}}{H_{R(160-1000 \text{ Hz})} - H_{13}} e^{2jk_0(X_{12}+X_{23}+X_{3S})}, \quad (4)$$

$$R_{(1-5 \text{ kHz})} = \frac{H_{23} - H_{I(1-5 \text{ kHz})}}{H_{R(1-5 \text{ kHz})} - H_{23}} e^{2jk_0(X_{23}+X_{3S})}, \quad (5)$$

where  $R$  is the reflection coefficient and  $k_0$  is the wave-number in the air.

The normal incidence sound absorption coefficient was calculated according to Eq. (6) (ISO, 1998):

$$\alpha = 1 - |R|^2. \quad (6)$$

Results were presented as one-third octave frequency bands.

#### 2.4. Methodology of measuring insertion loss

The study of insertion loss was performed in a semi-anechoic chamber. The chamber consisted of a solid floor, with walls and ceilings covered with acoustic foam to create an anechoic environment. The chamber had separate source and receiver rooms, each with dimensions of 2000 mm × 2500 mm × 2500 mm (L × W × H). A wall separating the two rooms had a test opening of 1000 mm × 1000 mm, where the acoustic louvers were mounted. The equipment used for measurements included an omnidirectional loudspeaker (Bruel & Kjaer Type 4292), an analyzer (Bruel & Kjaer 2270), and a microphone (Bruel & Kjaer Type 4189).

Two sound source positions in the source room and five microphone positions in the receiver room were

designated. Insertion loss was calculated as the difference between the sound pressure level without the louver and the sound pressure level with the louver. Insertion loss was calculated according to Eq. (7) (as specified in (ISO, 2020)):

$$IL = 20 \log \left( \frac{p_{\text{without}}}{p_{\text{with}}} \right) = L_{p1} - L_{p2}, \quad (7)$$

where  $L_{p1}$  is the sound pressure level without the structure [dB] and  $L_{p2}$  is the sound pressure level with structure [dB].

Initially, the sound pressure level was measured without any sample, and then the tests were repeated with the test sample. Results were calculated according to Eq. (7).

The following conditions were maintained during the tests:

- distance between source and microphone: >0.5 m;
- measurement time: 30 s;
- white noise was used as the sound signal.

### 3. Results

#### 3.1. Ray tracing

Ray tracing is a computational method that allows the prediction of wave propagation paths based on the shape of a structure and the properties of materials involved in absorbing and reflecting sound. This method allows for the analysis of how sound waves may change their direction of propagation or reflect off structures. In this study, wave propagation was analyzed during using Odeon 16.0 Auditorium software. The principle behind acoustical louvers relies on the ability to alter the tilt angle of the plates, thereby changing the direction of sound wave propagation. Additionally, using sound-absorbing materials, it becomes possible to absorb a portion of the sound energy. In this part of the research, the propagation of waves through the structural plates was analyzed by changing the tilt angle of the construction plates from 0° to 45°. The angle was adjusted every 15°, with the plates tilted both towards and away from the noise source.

The wave propagation analysis (Fig. 7) shows that when the plates were in the horizontal 0° position, the main proportion of sound waves was reflected from the plates only once. When increasing the tilt angle of the plates to 15°, it was observed that the tilt direction allowed for control over the direction of wave propagation. When the tilt angle was 15°, the largest proportion of transmitted waves reflected 1–2 times. When the plates were tilted at an angle of 30°, depending on the angle direction of the plates, most waves were reflected from 2 to 8 times. Meanwhile, when the panels were tilted at 45°, waves reflected from 4 to >10 times. With each reflection, the sound energy decreased depending on the absorption properties of the

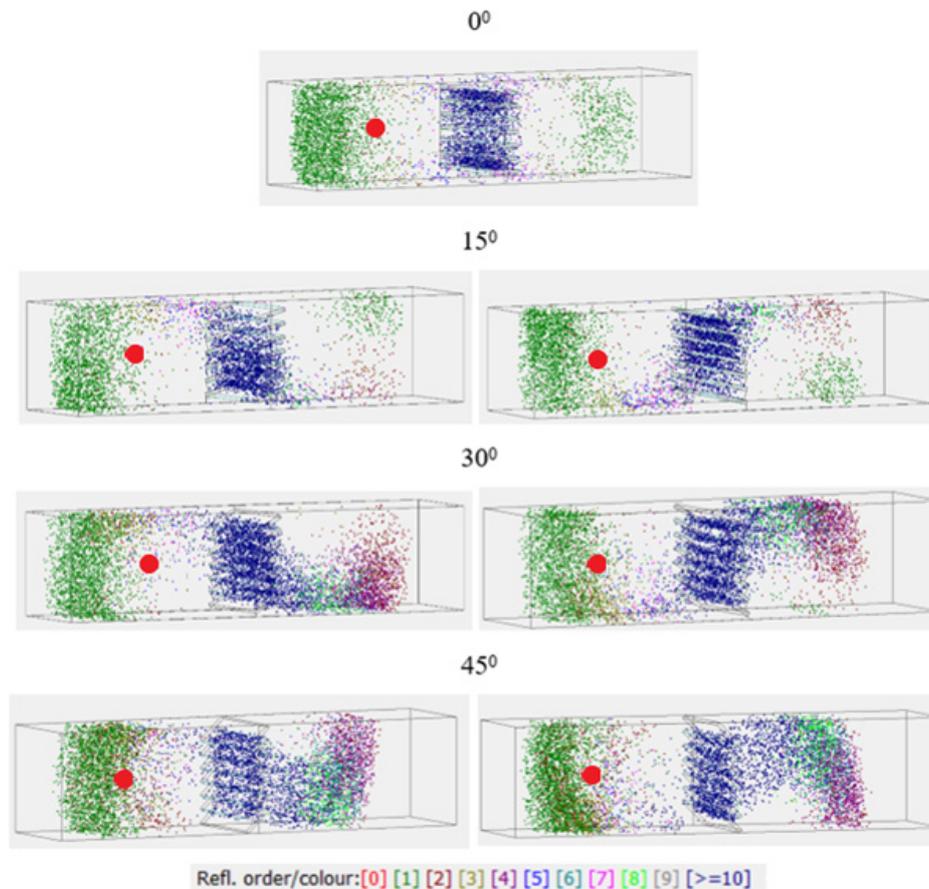


Fig. 7. Ray tracing of louvers, plates tilted at different angles.

material. The sound energy level on the other side of the louver was computed according to the following formula (ISO, 2010):

$$L_J = 10 \log \left( \frac{J}{J_0} \right) \text{ [dB]}, \quad (8)$$

where  $J$  is the sound energy, and  $J_0$  is the reference sound energy level ( $10^{-12}$  J).

According to the aforementioned equation, reducing the energy level by two times results in sound level being reduced by approximately 3 dB. Similarly, in order to reduce the sound pressure level by 10 dB, the energy level must be reduced by 10 times. So, if at a certain frequency the sound wave loses 50 % of its energy, indicating a sound absorption of 0.5 for the rubber panel, then with each reflection the sound level can be reduced by 3 dB. This means that at the 10th reflection, the sound level would be reduced by around 30 dB.

### 3.2. Sound absorption studies of rubber granule plates

Sound absorption was measured in a specially constructed device for the studies for material absorption – the impedance tube. Rubber panel samples, prepared according to the method described in Subsec. 2.3, were

30 mm in diameter and 12 mm in thickness. The composition of all the samples used in the studies is detailed in Table 1.

Table 1. Characteristics of rubber granulate panels.

Rubber granule panel	Particle size range [mm]	Rubber mass fraction [%]	Glue content [%]	Density [g/cm <sup>3</sup> ]
No. 1	4–15	94.9	5.1	0.601
No. 2	1–3	95.5	4.5	0.539
No. 3	1–15	95.7	4.3	0.666

Figure 8 shows the results of the sound absorption of rubber granule samples of compositions no. 1–3. The results indicate that the normal incidence sound absorption coefficient in the frequency band from 160 Hz to 630 Hz was approximately 0.1–0.15, due to relatively small sample thickness. However, from 800 Hz, the values of the normal incidence sound absorption coefficient started to rise and reached the maximum values of 0.79–0.96 at 3150 Hz. Notably, sample no. 2 had exhibited the highest normal incidence sound absorption coefficient (up to 0.96), consisting of 100 % devulcanized rubber granules. It was also observed that by reducing the amount of devulcanized rubber granules in

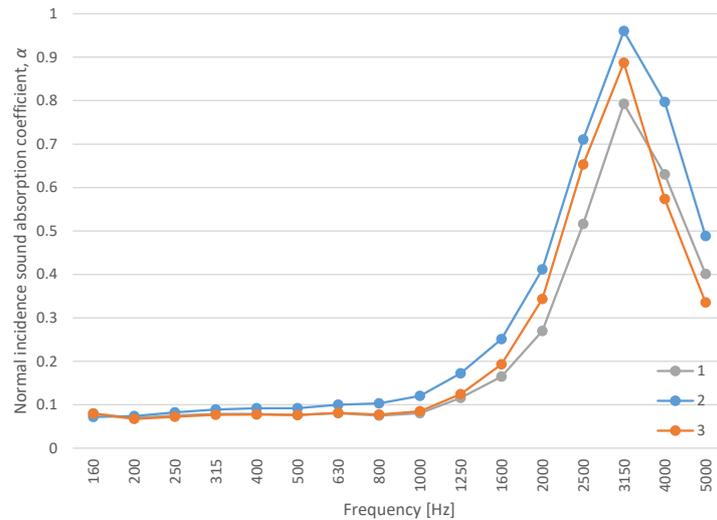


Fig. 8. Sound absorption results of rubber granule plates.

the composition, the normal incidence sound absorption coefficient decreased. Additionally, it was observed that the sound absorption of the rubber sample no. 3 reached up to 0.88, while the sound absorption of the rubber sample of composition no. 1 was 0.79. Based on the results, it can be stated that the panels made of 100 % devulcanized rubber granules had better sound absorption, as this type of panel exhibits the highest porosity. However, if the panel consists of rubber granules of a larger fraction, its porosity decreases, consequently leading to reduced sound-absorbing properties.

### 3.3. Research on the construction of acoustical lowers in a semi-anechoic chamber

Sound insertion loss studies of three structures with steel plates covered with different types of rubber granule panels were performed in a semi-anechoic chamber. Each structure was tested by tilting the louver plates from 0° to 45° in 15° steps. The plates were angled towards and away from the noise source (Fig. 9).

Tests conducted with constructions where plates were tilted towards the noise source revealed that at low frequencies ranging from 100 Hz to 500 Hz, the efficiency of all structures was similar. This similarity is attributed to low-frequency waves with wavelengths of 0.7 m to 3.3 m passing through the structure, creating resonances between the plates of the structure. However, due to the long wavelength, the louvers' construction had little attenuation of the sound pressure level. Since the width of the construction plates was 0.3 m, it was found that in all cases the critical frequency was determined at 1000 Hz. This depended on the thickness of the structure, since the wavelength of 1000 Hz was also close to 0.3 m. From 1000 Hz, the efficiency of structures began to increase rapidly, and in all cases reached its peak at 2000 Hz–3150 Hz.

At high frequencies, differences in the tilt angle of the plates became apparent. Our study showed that increasing the tilt angle of the plates also increased the insertion loss. Specifically, with the plates in 0° horizontal position, the sound level was reduced by 14 dB–17 dB at high frequencies. Changing the angle to 15° reduced the sound level by 17 dB–21 dB, while, further increasing the tilt angle to 30° and 45° resulted in sound pressure level reduction of 22 dB–26 dB and 24 dB–27 dB, respectively. The results showed that increasing tilt angle of louver plates by 15° increased sound reduction at high frequencies increased by an average of 2 dB–4 dB. It was also found that the efficiency also varied with different rubber granule plates. The best efficiency was obtained by using rubber granule plates no. 2, while the worst sound insulation properties was given by using rubber granule plates no. 1. The difference between plate no. 2 and plate no. 1 averaged 3 dB.

In the second case, where structures had plates tilted away from the noise source (Fig. 10), it was observed that at low frequencies ranging from 100 Hz to 500 Hz, the efficiency of all structures was similar because low-frequency sound waves with wavelengths of 0.7 m to 3.3 m, passed through the structure and generated resonances between the plates of the structure. As in the first case, critical frequencies were measured at 1000 Hz where the sound wavelength coincided with the width of the structure. The efficiency of the structures started to increase rapidly from 1000 Hz and reached its peak at high frequencies between 2500 Hz and 3150 Hz.

At high frequencies, differences in the tilt angle of the plates became apparent. It was observed that increasing the angle of the construction plates increased insertion loss. Specifically, with the plates in 0° horizontal position, the sound level was reduced by

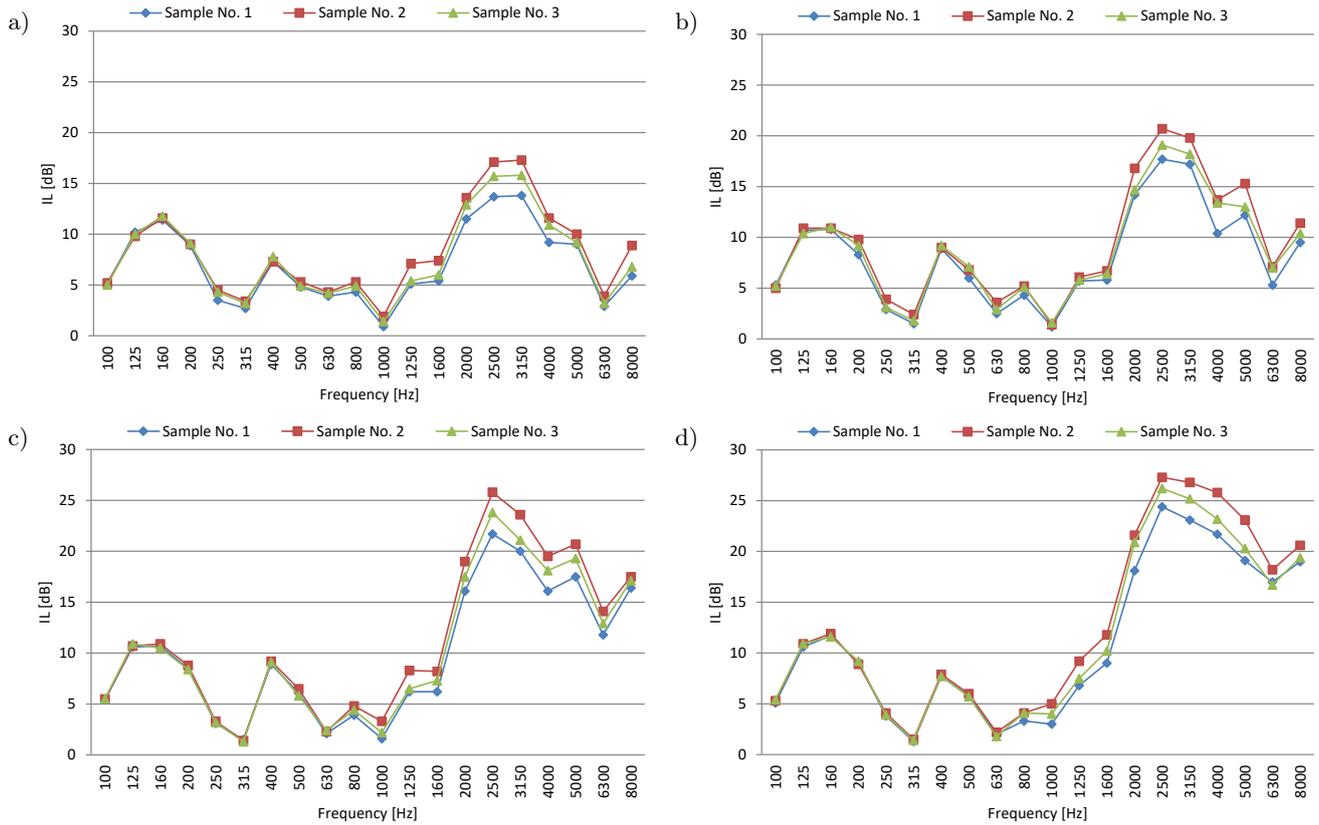


Fig. 9. Results of the insertion loss of the louvers with different angles of the plates tilting towards the source: a) 0°; b) 15°; c) 30°; d) 45°.

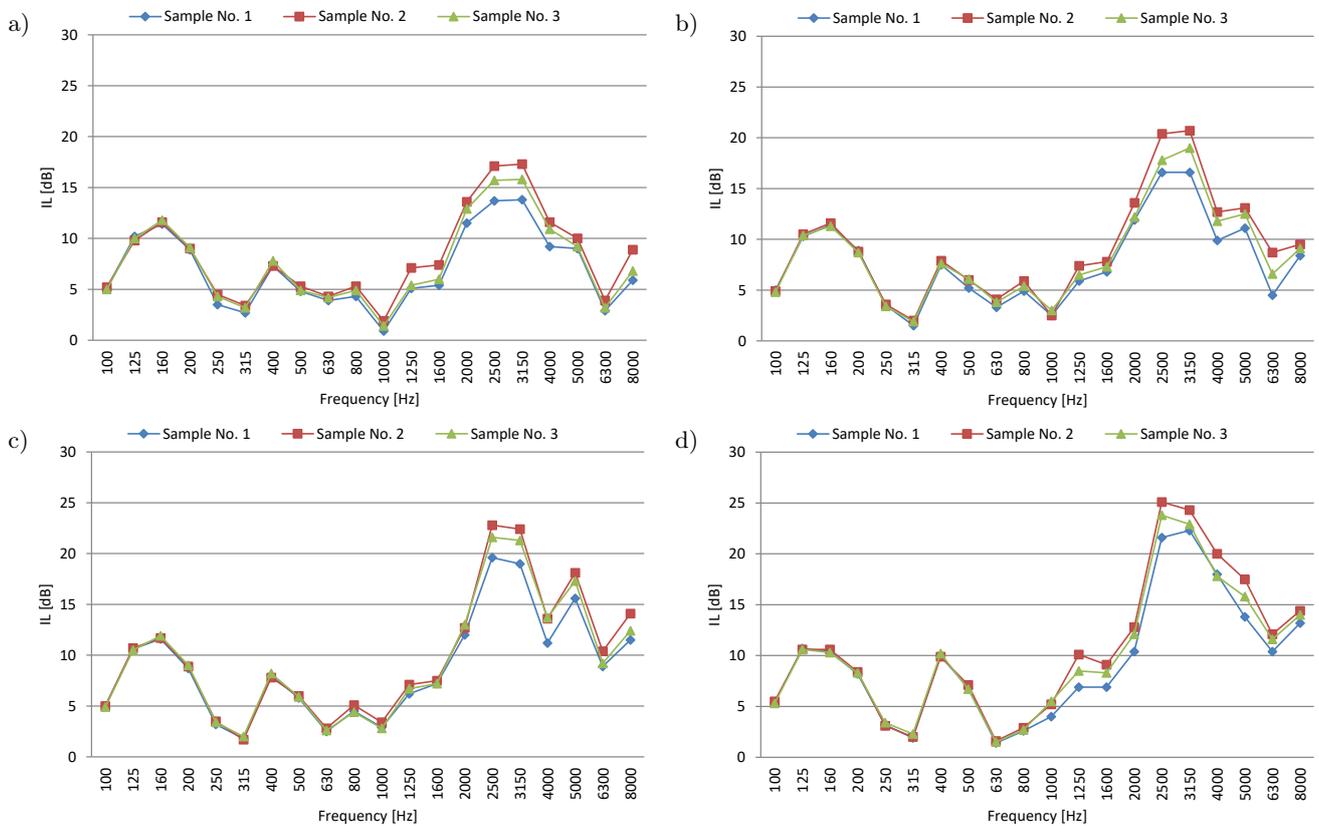


Fig. 10. Results of louvers' insertion loss with different tilt angles of plates away from the source: a) 0°; b) 15°; c) 30°; d) 45°.

Table 2. Equivalent reduction of the sound pressure level  $L_{Aeq}$  [dB].

Louvers plate tilt angle	Equivalent reduction of the sound pressure level $L_{Aeq}$ [dB]		
	Plates covered with rubber granulate no. 1	Plates covered with rubber granulate no. 2	Plates covered with rubber granulate no. 3
0°	6.9	9.4	8.2
15° towards the source	8.4	10.5	9.6
30° towards the source	9.3	12.2	11.7
45° towards the source	9.9	13.5	11.9
15° away from the source	7.7	10.5	9.3
30° away from the source	8.8	12.2	11.0
45° away from the source	9.2	12.9	11.8

14 dB–17 dB at high frequencies. Increasing the angle to 15° resulted in insertion loss values ranging from 16 dB to 21 dB at 2500 Hz. Further changing the tilt angle to the 30° led to sound pressure level reduction by 20 dB–24 dB, while, at an angle of 45°, the sound pressure level was reduced by 22 dB–25 dB. As in the first case, increasing the angle by 15° steps resulted in sound pressure level reduced by an average of 2 dB–3 dB. The dependences on the rubber granulate panels were also determined during the research. Our study showed that acoustic louvers covered with rubber granulate plate no. 2 exhibited a 2 dB greater reduction in sound pressure level compared to those in plate no. 3, and up to 6 dB better reduction compared to those using rubber granule panel no. 1.

In the study, it was also important to determine the equivalent decrease in the sound pressure level, which is given in Table 2.

The results of the equivalent sound reduction showed the same trends. The research results showed that the best reduction in sound pressure level occurred when the tilt angle of the plates was 45°. The increase in angle was found to improve the results by 0.5 dB to 1 dB. In addition, the best results were obtained by covering the construction plates with no. 2 rubber granule panel. When comparing different rubber granule plates, variations in results were noted, with differences of up to 2 dB.

#### 4. Discussion

The innovative design of the louver construction was based on the principle of sound absorption. According to ray-tracing simulation data, it was found that by changing the angle of the plates, sound waves could be reflected up to 10 times. Sound wave theory states that with each reflection, the sound wave loses some of its energy. To increase the efficiency of acoustic louvers, the plates were covered with sound-absorbing materials. Since such constructions were used under rather difficult conditions, in the presence of high heat, humidity, etc., it is especially important to select the

right materials. One of the best solutions to implement the principles of circular economy is to recycle and reuse waste tire rubber. In the studies, waste tires were shredded into two different fractions, with a particle size of 1 mm–3 mm and 4 mm–15 mm, respectively. To further increase porosity, rubber granules were devulcanized. Three type of rubber granule plates were created. The normal incidence sound absorption coefficient of no. 1 rubber granule plate reached 0.79, no. 2 – 0.96, and no. 3 – 0.88 at 3150 Hz.

The results of the research showed that the best sound absorption values were achieved at 3150 Hz, whereas at low frequencies (160 Hz–630 Hz) the absorption reached 0.1–0.15. Similar trends were seen in the insertion loss study. The sound reduction of louvers began to increase from 1000 Hz, along with increasing sound absorption. From this frequency, the insertion loss values reached 15 dB–25 dB according to the angle of louver plates. The sound insulation began to decrease from 3150 Hz, which was also compared to the sound absorption results. Therefore, in summation of the results, it was found that there was a direct relationship between the sound insulation efficiency of the acoustic louvers and the absorption properties of the material since the best results were obtained using rubber granule panels with the highest sound absorption. Also, the change in angle of the louver plates increased the number of reflections of sound waves. When construction plates are in the 45° position, the number of reflections can reach up to 10 times, which means that with each reflection, more than half of the energy, depending on material sound absorption, could be absorbed. The theory states, that if at a certain frequency the sound wave loses 50 % of its energy, then with each reflection, the sound level can be reduced by 3 dB, which means that at the 10th reflection, the sound level is reduced by around 30 dB and this compares with the results of the insertion loss. The highest values reached 25 dB at 2500 Hz–3150 Hz.

According to the literature, in general, all acoustic louvers provide about 10 dB–12 dB of sound insulation. In our case, a structure with 45° angled plates achieved up to 11.9 dB of equivalent reduction of the

sound pressure level. So, it can be stated that the use of materials made of used tires can provide a good sound-insulating structure for equipment that needs to ensure good air circulation.

## 5. Conclusions

The study has shown that at frequencies from 100 Hz to 315 Hz, all materials uniformly reduced the sound level, even by changing the tilt angle of the plates at different angles. Particularly, at 125 Hz–160 Hz, a greater decrease in sound level was found to be 11 dB–12 dB. At 200 Hz–315 Hz, the insertion loss dropped to 3 dB.

At frequencies from 400 Hz to 1250 Hz, all materials also similarly reduced noise, regardless of the changing tilt angle of the plates. In this frequency band, the decrease in sound level reached the lower limits. At 630 Hz–1000 Hz, the reduction in sound level was only 1 dB.

Differences in the reduction of sound level occurred only at high frequencies, where the best efficiency of the louvers was determined. The best efficiency was found to be achieved at frequencies of 2500 Hz–3150 Hz, where, depending on the angle and the material used, the reduction in sound level reached 14 dB–27 dB.

The study has shown that changing the angle of the louvers plates in 15° increments increases sound insertion loss by 2 dB–4 dB with all materials at each increment.

In addition, the best level of sound reduction was achieved with acoustic louvers whose plates were covered with rubber panel no. 2, made of chemically treated rubber granules of 1 mm–3 mm in size. The louver, which was covered with rubber panels no. 1 (granule size 4 mm–15 mm), produced 3 dB–4 dB lower sound level reduction than the panels no. 2. Construction with plates covered with rubber granule plate no. 3, which was made by mixing both types of granules, yielded sound level reduction on average 2 dB better than the structure covered with a rubber panel of mechanically processed rubber granules, but on average 2 dB worse than the structure covered with a board of chemically processed granules.

## References

- BUJOREANU C., NEDEFF F., BENCHEA M., AGOP M. (2017), Experimental and theoretical considerations on sound absorption performance of waste materials including the effect of backing plates, *Applied Acoustics*, **119**: 88–93, doi: [10.1016/j.apacoust.2016.12.010](https://doi.org/10.1016/j.apacoust.2016.12.010).
- CORREDOR-BEDOYA A.C., ZOPPI R.A., SERPA A.L. (2017), Composites of scrap tire rubber particles and adhesive mortar – Noise insulation potential, *Cement and Concrete Composites*, **82**: 45–66, doi: [10.1016/j.cemconcomp.2017.05.007](https://doi.org/10.1016/j.cemconcomp.2017.05.007).
- DISSANAYAKE D.G.K., WEERASINGHE D.U., THEBUNAWANAGE L.M., BANDARA U.A.A.N. (2021), An environmentally friendly sound insulation material from post-industrial textile waste and natural rubber, *Journal of Building Engineering*, **33**: 101606, doi: [10.1016/j.job.2020.101606](https://doi.org/10.1016/j.job.2020.101606).
- GANDOMAN M., KOKABI M. (2015), Sound barrier properties of sustainable waste rubber/geopolymer concretes, *Iranian Polymer Journal*, **24**(2): 105–112, doi: [10.1007/s13726-014-0304-1](https://doi.org/10.1007/s13726-014-0304-1).
- GAO N., HOU H. (2017), Low frequency acoustic properties of a honeycomb-silicone rubber acoustic metamaterial, *Modern Physics Letters B*, **31**(11): 1750118, doi: [10.1142/S0217984917501184](https://doi.org/10.1142/S0217984917501184).
- GÜNEYİSİ E. (2010), Fresh properties of self-compacting rubberized concrete incorporated with fly ash, *Materials and Structures*, **43**: 1037–1048, doi: [10.1617/s11527-009-9564-1](https://doi.org/10.1617/s11527-009-9564-1).
- HONG Z., BO L., GUANGSU H., JIA H. (2007), A novel composite sound absorber with recycled rubber particles, *Journal of Sound and Vibration*, **304**(1–2): 400–406, doi: [10.1016/j.jsv.2007.02.024](https://doi.org/10.1016/j.jsv.2007.02.024).
- International Organization for Standardization (1998), *Acoustics. Determination of sound absorption coefficient and impedance in impedances tubes. Part 2: Transfer-function method* (ISO Standard No. 10534-2:1998), <https://www.iso.org/standard/22851.html>.
- International Organization for Standardization (2010), *Acoustics. Determination of sound power levels and sound energy levels of noise sources using sound pressure. Precision methods for reverberation test rooms* (ISO Standard No. 3741:2010), <https://www.iso.org/standard/52053.html>.
- International Organization for Standardization (2020), *Acoustics. Measurement of insertion loss of ducted silencers without flow. Laboratory survey method* (ISO Standard No. 11691:2020), <https://www.iso.org/standard/69859.html>.
- JULIÁ E., SEGURA J., NADAL A., GADEA J.M., CRESCO J.E. (2013), Study of sound absorption properties of multilayer panels made from ground tyre rubbers, *Annals of the Oradea University. Fascicle of Management and Technological Engineering*, **XXII**(XII): 1–13.
- KARAKURT C. (2015), Microstructure properties of waste tire rubber composites: An overview, *Journal of Material Cycles and Waste Management*, **17**: 422–433, doi: [10.1007/s10163-014-0263-9](https://doi.org/10.1007/s10163-014-0263-9).
- KHALOO A.R., DEHESTANI M., RAHMATABADI P. (2008), Mechanical properties of concrete containing a high volume of tire-rubber particles, *Waste Management*, **28**(12): 2472–2482, doi: [10.1016/j.wasman.2008.01.015](https://doi.org/10.1016/j.wasman.2008.01.015).
- KOSALA K. (2019), Sound insulation properties of two-layer baffles used in vibroacoustic protection, *Applied Acoustics*, **156**: 297–305, doi: [10.1016/j.apacoust.2019.07.028](https://doi.org/10.1016/j.apacoust.2019.07.028).

15. LANDI D., GIGLI S., GERMANI M., MARCONI M. (2018), Investigating the feasibility of a reuse scenario for textile fibres recovered from end-of-life tyres, *Waste Management*, **75**: 187–204, doi: [10.1016/j.wasman.2018.02.018](https://doi.org/10.1016/j.wasman.2018.02.018).
16. MADERUELO-SANZ R., NADAL-GISBERT A.V., CRESPO-AMORÓS J.E., PARRÉS-GARCÍA F. (2012), A novel sound absorber with recycled fibers coming from end of life tires (ELTs), *Applied Acoustics*, **73**(4): 402–408, doi: [10.1016/j.apacoust.2011.12.001](https://doi.org/10.1016/j.apacoust.2011.12.001).
17. MARRIOTT K. (2012), An update on acoustics designs for HVAC (Engineering), [in:] *Proceedings of the Acoustics 2012 Nantes Conference*, pp. 3823–3829.
18. NACIF G.L., PANZERA T.H., STRECKER K., CHRISTOFORO A.L., PAINE K. (2013), Investigations on cementitious composites based on rubber particle waste additions, *Materials Research*, **16**(2), doi: [10.1590/S1516-14392012005000177](https://doi.org/10.1590/S1516-14392012005000177).
19. SIDDIQUE R., NAIK T.R. (2004), Properties of concrete containing scrap-tire rubber – An overview, *Waste Management*, **24**(6): 563–569, doi: [10.1016/j.wasman.2004.01.006](https://doi.org/10.1016/j.wasman.2004.01.006).
20. SU H., YANG J., LING T.C., GHATAORA G.S., DIRAR S. (2015), Properties of concrete prepared with waste tyre rubber particles of uniform and varying sizes, *Journal of Cleaner Production*, **91**: 288–296, doi: [10.1016/j.jclepro.2014.12.022](https://doi.org/10.1016/j.jclepro.2014.12.022).
21. SWIFT M.J., BRIŠ P., HOROSHENKOV K. V. (1999), Acoustic absorption in re-cycled rubber granulate, *Applied Acoustics*, **57**(3): 203–212, doi: [10.1016/S0003-682X\(98\)00061-9](https://doi.org/10.1016/S0003-682X(98)00061-9).
22. VIVEIROS E.B., GIBBS B.M. (2003), An image model for predicting the field performance of acoustic louvres from impulse measurements, *Applied Acoustics*, **64**(7): 713–730, doi: [10.1016/S0003-682X\(03\)00008-2](https://doi.org/10.1016/S0003-682X(03)00008-2).
23. XU X., WANG H., SUN Y., HAN J., HUANG R. (2018), Sound absorbing properties of perforated composite panels of recycled rubber, fiberboard sawdust, and high density polyethylene, *Journal of Cleaner Production*, **187**: 215–221, doi: [10.1016/j.jclepro.2018.03.174](https://doi.org/10.1016/j.jclepro.2018.03.174).

## Research Paper

# Application of Cellulose and Paper-Based Products in Building Acoustics

Aleksandra KLIMEK<sup>(1)\*</sup> , Jerzy F. ŁĄTKA<sup>(2)</sup> ,  
Paweł NIERADKA<sup>(1),(3)</sup> , Andrzej DOBRUCKI<sup>(1)</sup> 

<sup>(1)</sup> *Department of Acoustics, Multimedia and Signal Processing*  
*Wrocław University of Science and Technology*  
Wrocław, Poland

<sup>(2)</sup> *Department of Architecture and Visual Arts*  
*Wrocław University of Science and Technology*  
Wrocław, Poland

<sup>(3)</sup> *KFB Acoustics*  
Wrocław, Poland

\*Corresponding Author e-mail: [Aleksandra.Klimek@pwr.edu.pl](mailto:Aleksandra.Klimek@pwr.edu.pl)

(received July 5, 2023; accepted January 19, 2024; published online April 11, 2024)

This article presents a comprehensive acoustic study of paper-based building products: cellulose wool, paperboard, corrugated cardboard, and honeycomb panels. The material configurations included the intact form as well as the various modifications, i.e., density variation, multiple-layered staking, perforation or acoustic metamaterial setup. Tests covered acoustic absorption and insulation properties, with the last examined under excitation of both a plane wave and a diffused field. Additionally, the cellulose wool is provided with the characteristic impedance and propagation wavenumber results; and the paperboard was tested for its dynamic elastic and damping properties. The paper-based products, giving their weight, prove to be a convincing replacement for conventional materials by both absorptive and insulation performance. The maximum acquired sound reduction index, for exceptionally lightweight ( $2.2 \text{ kg/m}^2$ ) paper double-wall metamaterial structure, reached 26 dB.

**Keywords:** paper; cardboard; insulation; absorption; acoustic metamaterial; honeycomb.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

### Acronyms

$C$ ,  $C_{tr}$  – sound reduction index correction factors,

$C_1$  – coefficient for 1st beam mode,

$E$  – Young's modulus,

$H$  – spacing between the double walls,

$H_b$  – beam height in the vibration direction,

$H_R$  – height of the Helmholtz resonator,

$K$  – bulk modulus of air,

$K_{\text{eff}}$  – effective bulk modulus,

$S_R$  – cross section surface of the Helmholtz resonator neck,

SRI – sound reduction index,

TL – transmission loss,

$\mathcal{X}_{1,2}$  – angle related normalized wall impedance,

$X_{1,2}$  – normalized wall impedance,

$V_R$  – cavity volume of the Helmholtz resonator,

$\mathcal{Z}$  – angle related effective impedance,

$Z_{\text{eff}}$  – effective impedance,

$c$  – speed of sound in air,

$d_1$  – absorbent layer thickness,

$d_R$  – diameter of the Helmholtz resonator neck,

$f\omega$  – frequency, pulsation of wave,

$f_{01,2}$  – resonance frequency of double wall with the Helmholtz resonators system,

$f_{\text{DW}}$  – resonance frequency of a double wall,

$f_R$  – resonance frequency of the Helmholtz resonator,

$\Delta f$  – half-power bandwidth,

$k$  – wavenumber,

$k'$  – complex propagation wavenumber,

$k_{\text{eff}}$  – effective wavenumber,

$l$  – beam length,

$l_R$  – length of the Helmholtz resonator neck,

$m''_{1,2}$  – surface masses of double walls,

$z_1$  – rigid back absorbent surface impedance,  
 $z_c$  – complex characteristic impedance,  
 $\alpha$  – sound absorption coefficient,  
 $\zeta_R$  – the Helmholtz resonator damping ratio,  
 $\eta$  – damping loss factor,  
 $\theta, \psi$  – incident angle of acoustic wave,  
 $\theta_{\text{eff}}$  – angle of refraction,  
 $\tau$  – transmission coefficient,  
 $\phi_R$  – volumetric fill ratio of Helmholtz resonators,  
 $\rho$  – volumetric density,  
 $\rho_{\text{eff}}$  – effective density.

## 1. Introduction

Building industry is widely known as one of the most environmentally hazardous. It is responsible for 40 %–50 % of world greenhouse gas emissions (ABD RASHID, YUSOFF, 2015), 24 % material extraction from the lithosphere (ZABALZA BRIBIÁN *et al.*, 2011) and 40 % of global energy consumption (WANG *et al.*, 2023). Therefore, the research concerning building industry in the last decades emphasised new materials that minimise the environmental impact. On the other hand, more and more architectural structures and infill (such as interior partitions) are built for the limited period of time. After their life span, they should be recycled, upcycled or utilised in a way that reduces the environmental impact.

In the search of environmentally friendly and bio-based building materials, paper among materials such as wood, clay, adobe bricks, hempcrete, recycled plastics and glass, certified timber seems promising and fulfilling the building code requirements.

Paper is a material of natural origin, of which the main building component is cellulose, the most common natural polymer in the world. Its resources are considered inexhaustible (KLEMM *et al.*, 2005). Paper as a material formed by the network of cellulose fibres, which create a hydrogen bonds between each other, may be given different forms and characteristics. For building application, whether it is a structural element or partition, the following paper-based products are mostly often implemented (LATKA, 2017a; DIARTE, SHAFFER, 2021):

- paper tubes as linear and structural elements. Those products were vastly examined by Japanese architect Shigeru Ban, who implemented them in both temporary and permanent structures (MIYAKE *et al.*, 2009);
- cardboard profiles (*U*- and *L*-shapes), similarly like paper tubes are suitable as structural rod elements (LATKA, 2017b);
- corrugated cardboard, the most popular product used in packaging, which can serve as planar elements, i.e., wall, floor and roof panels. This product, due to its internal structure had the best mechanical properties along the corrugation (WOLF *et al.*, 2021);
- honeycomb panels composed of two liners and a honeycomb core between them have a suitable structure for planar elements loaded perpendicularly to their surface (CRIPPS, 2004);
- paper board, best for use as an element that reinforces the properties of a building envelope or its finish;
- cellulose wool, and industrialised material providing thermal insulation to buildings, commonly used as a replacement for polystyrene and mineral wool.

As linear elements (paper tubes and cardboard profiles) are used as structural elements, their most significant feature is their mechanical properties, including bending and compression resistance. On the other hand, planar elements, that serve as partitions (external and internal) are subjected to loads, but also should have sufficient thermal and sound insulation.

The most critical properties of paper based products are their resistance to water and moisture, and incombustibility. However, there is a vast research conducted on those aspects of paper-based structures (KNAACK *et al.*, 2023). One of the least explored issues in the use of paper products in building structures is their acoustic properties. This is due to the focus of researchers on properties related to strength, fire and moisture effects and thermal insulation.

When comparing paper-based building partitions with traditional solutions, such as timber frame wall, it can be clearly seen that with similar thermal properties, products made of paper have a lower environmental impact (JASIOLEK *et al.*, 2023).

Paper can be therefore applied in both interior and exterior partitions. However, next to their structural, thermal and ecological properties, the sound insulation is one of the key features.

### 1.1. Acoustic characterisation of building materials

The fundamental acoustic attribute of regular building partitions is its potential impact on decreasing the noise imissions – either from external or internal sources. A complementary architectural design must fulfil the insulation requirements using components with known acoustic properties. The general building components used in partitions are not only sound insulators but also sound-absorbing materials applied as fillers of the structure. Each of these groups could be described by different parameters with corresponding measurement methods. Unfortunately, sustainable projects based on paper-based products are exacting because of inaccessibility to such data (LATKA *et al.*, 2022). Only sound-absorbing cellulose wool, as a material with growing use in both acoustic and thermal insulation, is sufficiently well-researched (ARENAS *et al.*, 2014). Hitherto published works on paper-based products included sound absorption and insulation tests of

corrugated cardboard (ASDRUBALI *et al.*, 2015; KANG *et al.*, 2021; LATKA, 2017a), honeycomb panels (SECHI *et al.*, 2016), paperboard (NERI *et al.*, 2021), and wastepaper (RICCIARDI *et al.*, 2014). Mentioned research, while valuable, could have a limited application in acoustic projects, either because the paper-based materials are a part of a bigger component, or the test is performed with a small sample impedance tube method.

This publication presents the acoustic parameters for cellulose wool and paper-based products: paperboard, corrugated cardboard, and honeycomb panels (unperforated and perforated). The conducted research employs three methods. First – the impedance tube transfer matrix method for the sound absorption coefficient  $\alpha$ , transmission loss (TL), the complex characteristic impedance  $z_c$ , and complex propagation wavenumber  $k$ , when possible. Second – the reverberation room method for the laboratory-tested weighted sound reduction index (SRI). Additionally, the beam resonance test was performed for paperboard exclusively to obtain its dynamic mechanical properties: Young's modulus  $E$  and the damping loss factor  $\eta$ .

## 2. Materials

The presented study examines the products (Fig. 1):

- cellulose wool made from recycled paper, with density varying from  $30 \text{ kg/m}^3$  to  $90 \text{ kg/m}^3$ ;
- recycled paperboard with the grammage of  $2000 \text{ g/m}^2$  and thickness of 3 mm;
- corrugated cardboard, types: BC, EE, EB, E, and B, as given in Table 1. The cardboard was measured in the configuration of a single plate or multiple plates:
  - a) five plates glued with polyvinyl acetate (PVA) adhesive (tested types are: EE, EB, and BC);
  - b) five plates stacked without any adhesive (tested types are: EE, BC, EE, and BC alternately).
- honeycomb panels with the different cells size, thicknesses and grammage, as listed in Table 2. The specific honeycomb geometry allows for enhancement of the absorption or insulation capabilities simply by perforating of the liner. Effectively, each cell transforms to the Helmholtz resonator with a resonance frequency dependent on the cell size, liner thickness and perforation diameter. Subsection 2.1 describes the concept of such a modification.

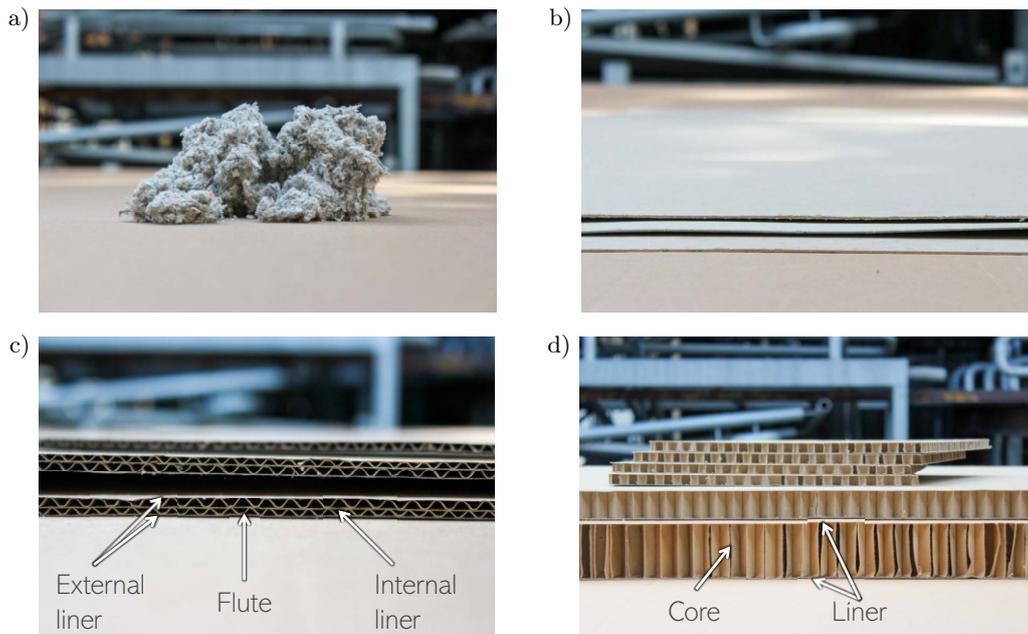


Fig. 1. Materials under test with their structure denotations: a) cellulose wool; b) paperboards; c) corrugated cardboards; d) honeycomb panels.

Table 1. Tested corrugated cardboard types and parameters.

Corrugated cardboard					
Thickness [mm]	Type	Grammage [ $\text{g/m}^2$ ]	Flute paper [ $\text{g/m}^2$ ]	External liner paper [ $\text{g/m}^2$ ]	Internal liner paper [ $\text{g/m}^2$ ]
6.1	Double wall BC	573	Wellenstoff 95	Testliner 110	Testliner 100
2.9	Double wall EE	555	Wellenstoff 95	Testliner 110	Testliner 100
4.0	Double wall EB	556	Wellenstoff 95	Testliner 110	Testliner 100
1.6	Single wall E	338	Wellenstoff 95	Testliner 110	–
2.7	Single wall B	342	Wellenstoff 95	Testliner 110	–

Table 2. Tested honeycomb panel types and parameters.

Honeycomb panels					
Symbol	Thickness [mm]	Cell size [mm]	Grammage [g/m <sup>2</sup> ]	Core paper [g/m <sup>2</sup> ]	Liner paper [g/m <sup>2</sup> ]
H50C22	50	22	1024	Testliner 120	Testliner 120
H50C14	50	14	1620	WB fluting 140	Testliner 120
H25C22	25	22	632	Testliner 120	Testliner 120
H25C14	25	14	930	WB fluting 140	Testliner 120
H10C10	10	10	590	WB fluting 140	Testliner 120
H50C22 KRAFT	50	22	1515	Testliber 140	Kraftliber 300
H25C22 KRAFT	25	22	1057	Testliber 140	Kraftliber 300

2.1. Honeycomb panels perforation

The honeycomb perforation creates the grid of Helmholtz resonators resulting in an exceedingly high absorption coefficient for the resonant frequency. Thus, the perforated honeycomb panels may be used as a filler for the existing structure to increase its sound insulation. May the double panel be the structure under examination in this case.

The resonance frequency  $f_R$  of the single Helmholtz resonator with a round neck is given by LANGFELDT *et al.* (2020):

$$f_R = \frac{c_0}{2\pi} \sqrt{\frac{S_R}{(l_R + \frac{\pi d_R}{4}) V_R}}, \tag{1}$$

where  $c_0$  is the speed of sound in air,  $S_R$ ,  $l_R$ ,  $d_R$ , and  $V_R$  are the resonator dimensions: cross-section surface, length, diameter of the neck, and volume of the cavity, respectively.

The mass-air-mass resonance frequency of air-filled double wall is (LANGFELDT *et al.*, 2020):

$$f_{DW} = \frac{1}{2\pi} \sqrt{\frac{K}{H} \left( \frac{1}{m''_1} + \frac{1}{m''_2} \right)}, \tag{2}$$

where  $K$  is the bulk modulus of air and  $H$  is the spacing between the walls with surface masses of  $m''_{1,2}$ .

LANGFELDT *et al.* (2020) proved that the resonance system of a double wall with the insertion of Helmholtz resonators has two resonance frequencies  $f_{01}$  and  $f_{02}$ , both different from but simultaneously dependent on  $f_R$  and  $f_{DW}$ :

$$f_{01,2}^2 = \frac{1 + \frac{f_R^2}{f_{DW}^2} \pm \sqrt{\left(1 - \frac{f_R^2}{f_{DW}^2}\right)^2 + 4\phi_R \frac{f_R^2}{f_{DW}^2}}}{2(1 - \phi_R)} f_{DW}^2 \tag{3}$$

with  $\phi_R$  being the volumetric fill ratio of Helmholtz resonators, in this case  $\phi_R = H_R/H$ , where  $H_R$  is

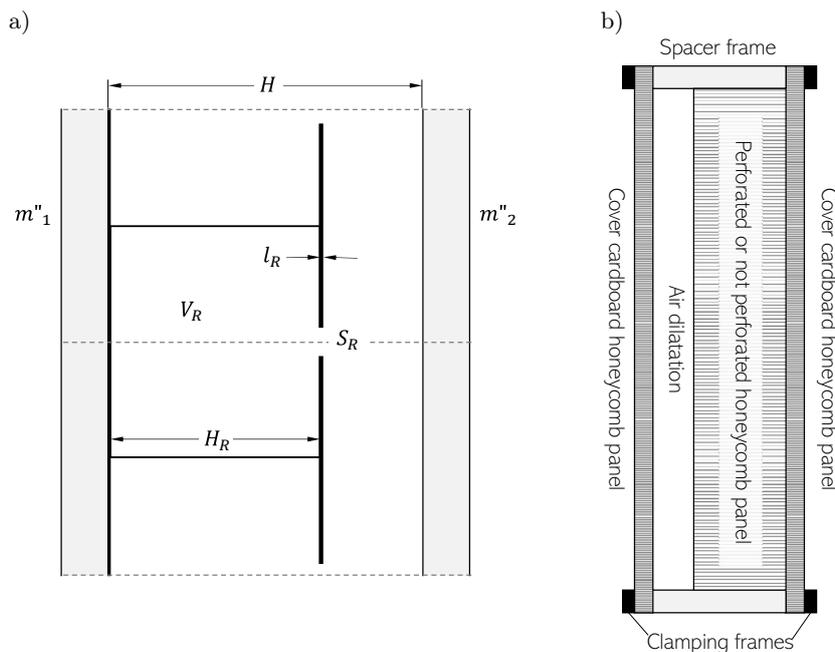


Fig. 2. Detailed scheme of the Helmholtz resonator in double wall (a) and schematic representation of a double wall with perforated or not perforated honeycomb cardboard insertion (b).

Table 3. Resonances frequencies of honeycomb panel perforation.

Honeycomb panel symbol	Assessed case	Neck diameter $d_R$ [mm]	Helmholtz resonance frequency $f_R$ [Hz]	Filling ratio $\phi_R$	Double wall resonance frequency $f_{DW}$ [Hz]	Resonance frequency $f_0$ (Eq. (3))	
						$f_{01}$ [Hz]	$f_{02}$ [Hz]
H50C22	Single panel	3.5	730	–	–	–	–
H25C22	Single panel	5	1260	–	–	–	–
H50C22 KRAFT	Double wall	1.8	316	0.725	346	247	725
H25C22 KRAFT	Double wall	1.6	416	0.625	474	325	1158

the height of the Helmholtz resonator cell (Fig. 2a). When  $f_R$  and  $f_{DW}$  are equal, Eq. (3) simplifies to:

$$f_0^2 = \frac{1 \pm \sqrt{\phi_R}}{1 - \phi_R} f_{DW}^2.$$

A significant transmission loss (TL) increase can be acquired in the bandwidth lying between given frequencies. The amplitude of this TL growth and its bandwidth depend on  $f_R$ ,  $f_{DW}$  and  $\phi_R$ . In case of the plane wave at an incidence angle  $\theta$ , the TL is represented by:

$$\begin{aligned} \text{TL} &= -10 \log \tau \\ &= -10 \log \left| \frac{2Z \sin(k_{\text{eff}} H \cos \theta_{\text{eff}})}{\mathcal{X}_1 \mathcal{X}_2 \sin^2(k_{\text{eff}} H \cos \theta_{\text{eff}}) + Z^2} \right|^2, \end{aligned} \quad (4)$$

where the effective wavenumber  $k_{\text{eff}} = 2\pi f \sqrt{\rho_{\text{eff}}/K_{\text{eff}}}$ , effective density  $\rho_{\text{eff}} \approx \rho_{\text{air}}(2 + \phi_R)/2(1 - \phi_R)$ , the angle of refraction  $\theta_{\text{eff}} = \arcsin(\sin \theta \sqrt{K_{\text{eff}}/\rho_{\text{eff}}/c_0})$ .

The angle related effective impedance is:

$$Z = Z_{\text{eff}} \sec \theta_{\text{eff}} / Z_{\text{air}} \sec \theta,$$

with the effective impedance  $Z_{\text{eff}} = \sqrt{K_{\text{eff}} \rho_{\text{eff}}}$ .

$$\mathcal{X}_{1,2} = X_{1,2} \cos \theta + 1 - Z \cot(k_{\text{eff}} H \cos \theta_{\text{eff}}),$$

where  $X_{1,2} = i\omega m''_{1,2}/Z_{\text{air}}$  are the normalized wall impedances. The factor  $K_{\text{eff}}$  is the effective bulk modulus of the Helmholtz resonator in closed volume as it is in the case under examination. For the damped ( $\zeta_R$  – resonator damping ratio) and air-filled system, the ratio  $K_{\text{eff}}$  to the air bulk modulus  $K$  is equal to:

$$\frac{K_{\text{eff}}}{K} = \frac{1}{1 - \phi_R + \frac{\phi_R}{1 + 2i\zeta_R \frac{\omega}{\omega_R} - \frac{\omega^2}{\omega_R^2}}}, \quad (5)$$

which means, that for certain frequencies  $f = \omega/2\pi$  above the Helmholtz resonance frequency  $f_R = \omega_R/2\pi$ , the  $K_{\text{eff}}$  could be lower below 0. This attribute – the negative bulk modulus (or the negative density, or both) is used in the acoustic metamaterial design, among other things, to increase sound insulation.

The diffused-field transmission coefficient could be then integrated over the incident angles from 0 to the maximum angle of  $\theta_l$ :

$$\bar{\tau} = \frac{\int_0^{\theta_l} \tau \cos \theta \sin \theta \, d\theta}{\int_0^{\theta_l} \cos \theta \sin \theta \, d\theta}. \quad (6)$$

The executed tests covered two issues: the sound absorption of a single perforated or non-perforated honeycomb panel and the sound reduction index of a double wall with the insertion of a perforated or non-perforated honeycomb panel, as in Fig. 2b. To match the Helmholtz resonance slightly below the double wall resonance, light honeycomb panels also form the covers. The dense honeycomb grid (H10C10 from Table 1) was chosen for this purpose, providing a small mass combined with high stiffness; and inhomogeneities, which do not influence the structure's vibration properties significantly. Table 3 lists all assessed structures along with their key parameters.

The analytically obtained results may differ in the practical applications. The cardboard production process does not guarantee the formation of equal cell volumes, which has two consequences: the change in width and position of the resonance (especially the width due to the resonance blur) and the technical difficulties in maintaining a constant opening position in relation to the cell. Both consequences do not exclude the honeycomb from potential utilization.

### 3. Methods

This paper employs three measurement methods:

- impedance tube two-loaded transfer function method (based on [ASTM E2611-19 \(2019\)](#)) for sound absorption coefficient  $\alpha$ , transmission loss TL, characteristic impedance  $z_c$ , and propagation wavenumber  $k$ ;
- laboratory measurements of sound insulation in a reverberation chamber (based on [International Organization for Standardization \(2000\)](#))

Table 4. Tested materials and used measurement methods with assessed parameters.

No.	Material	Sample thickness [mm]	Impedance tube measurement				Reverb. chamber measurement SRI (or $R_{I,M,W}$ )	Beam resonance measurement			
			$\alpha$	TL	$z_c$	$k'$		$E$	$\eta$		
1.	Cellulose wool made from recycled paper [kg/m <sup>3</sup> ]	Density 30	200	+	+	+	+	–	–	–	
		Density 40	200	+	+	+	+	–	–	–	
		Density 50	200	+	+	+	+	–	–	–	
		Density 60	200	+	+	+	+	–	–	–	
		Density 70	200	+	+	+	+	–	–	–	
		Density 80	200	+	+	+	+	–	–	–	
		Density 90	200	+	+	+	+	–	–	–	
2.	Paperboard – grammage 2000 g/m <sup>2</sup>	3.0	+	–	–	–	+	+	+		
3.	Corrugated cardboard	Type BC	6.1	–	–	–	–	+	–	–	
		Type EE	2.9	–	–	–	–	+	–	–	
		Type EB	4.0	–	–	–	–	+	–	–	
		Type E	1.6	–	–	–	–	+	–	–	
		Type B	2.7	–	–	–	–	+	–	–	
4.	Corrugated cardboard – five PVA glued layers	Type BC	30.5	–	–	–	–	+	–	–	
		Type EE	14.5	–	–	–	–	+	–	–	
		Type EB	20	–	–	–	–	+	–	–	
5.	Corrugated cardboard – five stacked layers	Type BC	30.5	–	–	–	–	+	–	–	
		Type EE	14.5	–	–	–	–	+	–	–	
		Type BC/EE	24.1	–	–	–	–	+	–	–	
6.	Honeycomb panel	H50C22	50	–	–	–	–	+	–	–	
		H50C14	50	–	–	–	–	+	–	–	
		H25C22	25	–	–	–	–	+	–	–	
		H25C14	25	–	–	–	–	+	–	–	
		H10C10	10	–	–	–	–	+	–	–	
		H50C22 KRAFT	50	+	+	–	–	–	+	–	–
		H25C22 KRAFT	25	+	+	–	–	–	+	–	–
7.	Perforated honeycomb panel*	H50C22 KRAFT	50	+	+	–	–	–	–	–	
		H25C22 KRAFT	25	+	+	–	–	–	–	–	
8.	Double wall with plain and perforated honeycomb panel*	H50C22	90	–	–	–	–	+	–	–	
		H25C22	60	–	–	–	–	+	–	–	

\* The perforation details are provided in Table 3.

for weighted modified intensity sound reduction index ( $R_{I,M,W}$ );

- measurement of vibration-damping properties with the beam resonance method (based on [ASTM E756-05 \(2005\)](#)) for dynamic Young's modulus  $E$  and the dynamic damping loss factor  $\eta$ .

Table 4 contains all tested materials assigned to the measurements performed upon them, along with assessed parameters.

### 3.1. Impedance tube two-loaded transfer function method

Two-loaded transfer function measurement was performed in an impedance tube system. The circular cross-section tube with an internal diameter of 101.7 mm was excited with the white noise. The tube parameter constrains the upper limit of the measured frequency range to 1600 Hz, thereby the results are provided for  $1/3$  octave bands in the range

50 Hz–1600 Hz. The digital audio pressure time processing determined the transfer function between the four individual microphones and created a transfer matrix. The Hanning window function was used to analyse the 10 s signal, resulting in a frequency resolution equal to 1 Hz. The used equipment were: power amplifier Atoll AM100SE, voltage output/input modules: NI 9260/NI 9234, and microphones G.R.A.S 46BD.

Three different specimens of each material were precisely trimmed to a diameter of 101.7 mm with a laser cutter, when possible. This operation allowed mounting them freely in the tube and consequently avoiding the emergence of membrane resonances in the range of the tested frequencies. Cellulose wool has been immobilized by two steel grids not affecting the acoustic field (Fig. 3a). Any leaks between the pipe and the plates were sealed with butyl.

The [ASTM E2611-19 \(2019\)](#) refers to the normal incidence sound transmission loss measurement, however it also outlines procedures for measuring other

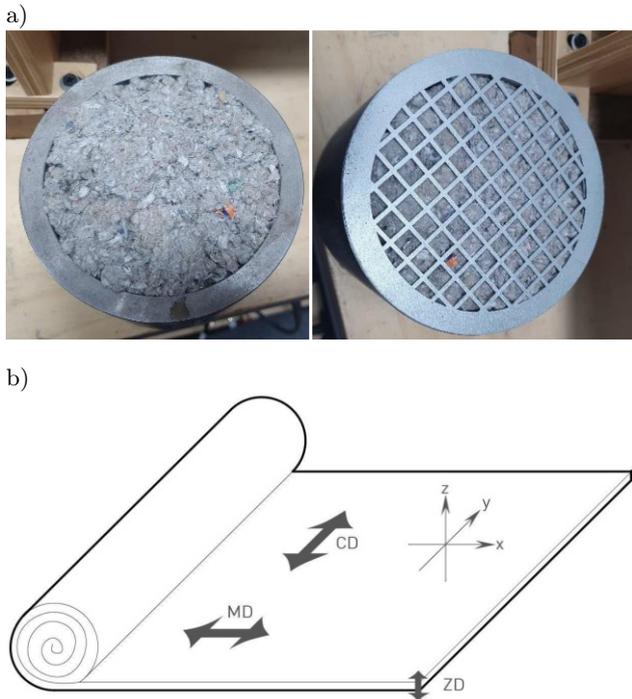


Fig. 3. a) Cellulose wool mounting; b) production directions of paperboard: machine direction (MD); cross-machine direction (CD); thickness direction (ZD) (NISKANEN, 2011).

material properties, employing the transfer matrix method. As a result, the relationship for determining the hard-backed absorption coefficient is provided as:

$$\alpha = 1 - |R|^2 = 1 - \left| \frac{T_{11} - \rho_0 c_0 T_{21}}{T_{11} + \rho_0 c_0 T_{21}} \right|^2, \quad (7)$$

where  $R$  – the hard backed reflection coefficient,  $\rho_0$ ,  $c_0$  – density and sound speed in air,  $T_{ii}$  – measured transfer matrix elements (see ASTM E2611-19 (2019) for details).

Correspondingly, the complex propagation wave-number  $k'$  and the complex characteristic impedance  $z_c$  are given by

$$k' = \frac{1}{d} - \cos^{-1} T_{11} \quad (8)$$

and

$$z = \sqrt{\frac{T_{12}}{T_{21}}}, \quad (9)$$

where  $d$  is the thickness of a sample.

### 3.2. Laboratory measurements of sound insulation in a reverberation chamber

The subsequent reverberation chamber measurement delivered the resulting weighted modified intensity sound reduction index  $R_{I,M,W}$  along with the correction factors  $C$  and  $C_{tr}$ . The measurement consisted in clamping a sample with dimensions of

1162 mm × 865 mm in the window of the reverberation transmitting chamber with a volume of 67 m<sup>3</sup> and non-parallel walls. The joint between the clamping frame and the sample was sealed with butyl and a layer of EPDM rubber. The chamber was excited by a reference sound source Norsonic 276 positioned to create a diffuse sound field, whose average sound pressure level was then measured by a sound level meter SV 22 at locations representative of the sound energy impacting the sample. The sound intensity probe measured the signal outside the chamber with a sweeping method. The Gras 40GK 1/2'' microphones being part of the sound intensity probe were located at a 100 mm distance from the sample surface. The used probe microphone separator was 50 mm.

The measurements were unfortunately confined to 1/3 octave bands in the range of 100 Hz–1600 Hz. The lower limit was a result of resonances occurring within the reverberation chamber, leading to its misalignment with the low-frequency range. The upper limit aligns with the range of the utilized probe separator. Consequently, the SRI values for frequencies between 1600 Hz and 3150 Hz are absent from the following graphs, though they have been factored into the weighted SRI and spectrum adaptation terms. This decision was taken for three main reasons: first, the observed values followed the trend established by the previous frequency bands and adhered to pertinent physical laws (e.g., mass law). Second, the PI index condition was satisfied for all results. Finally, even significant fluctuations in these three bands' values (by ±3 dB) did not induce changes in the calculated parameters exceeding 1 dB. Nonetheless, it should be noted that the concerned parameters are only approximated.

### 3.3. Beam resonance method for vibration-damping properties

The dynamic mechanic properties are essential to perform numerous computer acoustic analyses, i.e., widely used the finite element analyses (FEA). Paperboard, as a non-linear material, has vastly different values for dynamic properties than static ones. What is more, due to the production process, paperboard is an orthogonal material with machine direction (MD) properties incomparable to the cross-machine direction (CD) (Fig. 3b). Thus, the tests apply to both MD and CD directions and for several vibration frequency points between 30 Hz and 160 Hz.

The beam resonance method measures the frequency and the quality factor of cantilever beam resonance. In this examination, the used method is completely non-contact, meaning the beam vibration velocity was obtained with a Polytec PSV-400 laser vibrometer, while the beam itself was excited by a Brüel & Kjær MM002 magnetic transducer with chirp signal.

The transducer affected a metal plate with a negligible mass, attached to the beam by synthetic wax.

This configuration makes it possible to determine the dynamic Young modulus  $E$ , for the resonance frequency of the 1st mode  $f$ , as in (ASTM E756-05 (2005)):

$$E = \frac{(12\rho l^4 f^2)}{(H_b^2 C_1^2)}, \quad (10)$$

and the damping loss factor as:

$$\eta = \frac{\Delta f}{f}, \quad (11)$$

where  $\rho$  – volumetric density,  $l$  – beam length,  $H_b$  – beam height in the vibration direction,  $C_1$  – coefficient for the mode 1 of the clamped-free beam, equal 0.55959,  $\Delta f$  – half-power bandwidth.

## 4. Results and discussion

This section presents the results divided by the material groups, as in Table 4, of the methods described in Sec. 3:

1. cellulose wool made from recycled paper;
2. paperboard;
- 3.–5. corrugated cardboard – single plates and multiple layers;
- 6.–8. honeycomb panel – single plates and structures with and without perforation.

### 4.1. Cellulose wool

Cellulose wool is a loose material widely used in building insulation as an alternative to mineral wool or fibreglass. The wool under examination is a commercial product used for this purpose, i.e. it contains flame retardant and anti-moisture additives. Wool, as a fibrous material, has been tested for a wide range of acoustic properties. Apart from the sound absorption and transmission loss, the characteristic impedance  $z_c$ , and the propagation wavenumber  $k'$  are also the subject of this study.

Figure 4 shows the first two parameters' results for several wool densities. For the lowest density of  $30 \text{ kg/m}^3$  and material thickness equal to 20 cm, the sound absorption coefficient reaches 1, by a frequency of 250 Hz. For the mentioned thickness, the sound absorption worsens above the density of  $50 \text{ kg/m}^3$ . The cellulose wool has similar absorption properties to conventional materials, which similarity is particularly visible in comparison to fibreglass in Fig. 5. Simultaneously, recycled materials such as cellulose wool have a significantly lower environmental impact. For instance – cellulose fibre has over 2.5 times smaller primary energy demand, over 1.5 smaller water demand, and a similar Global Warming Potential to rock wool with the same mass (ZABALZA BRIBIÁN *et al.*, 2011).

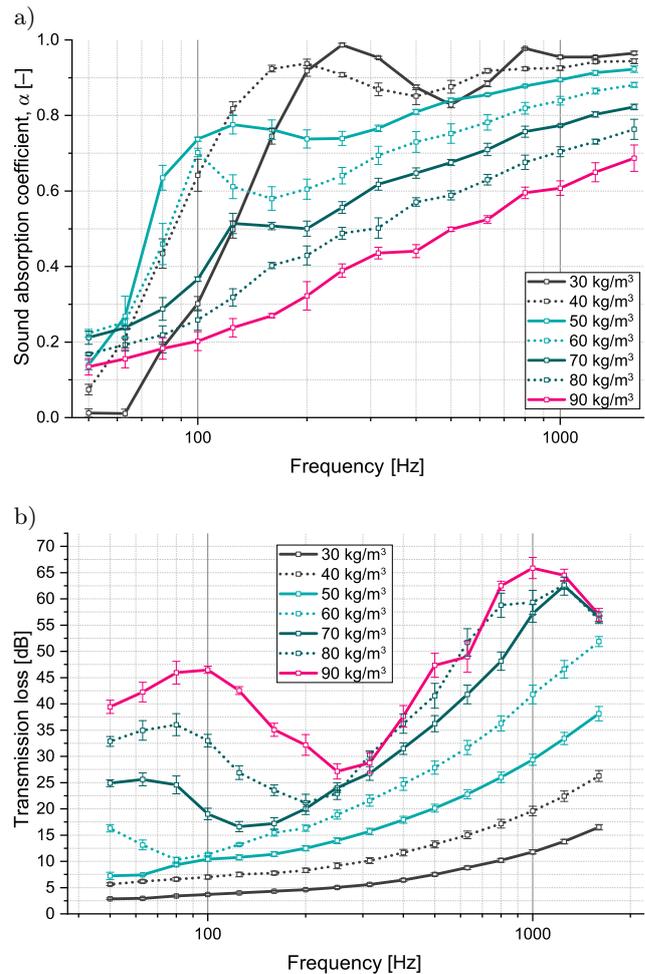


Fig. 4. Results of the sound absorption coefficient (a) and transmission loss (b) of different densities recycled cellulose wool. Horizontal lines mark standard error of measurements.

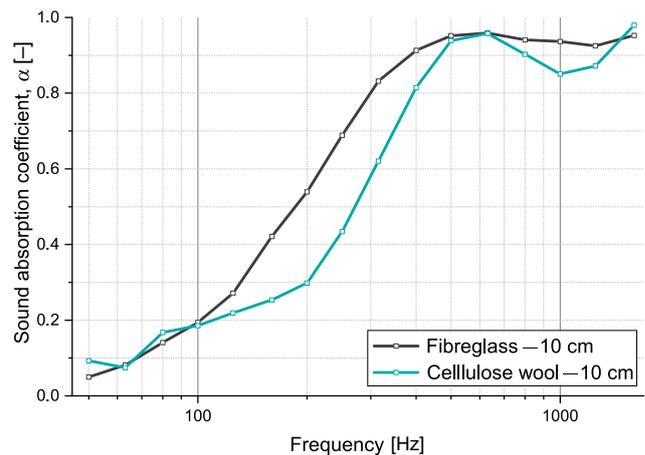


Fig. 5. Sound absorption coefficient comparison of the same density and thickness recycled cellulose wool (blue) and fibreglass (black) ( $30 \text{ kg/m}^3$ ). The results for cellulose wool were calculated with the data available in Table 5.

Table 5 presents the values of the characteristic impedance  $z$  and propagation factor  $k'$ . The data is

Table 5. Results of the complex characteristic impedance  $z_c$  [kg/m<sup>2</sup>s] and complex propagation wavenumber  $k'$  [1/m] of recycled cellulose wool.

Density [kg/m <sup>3</sup> ]	Parameter	The centre frequency of the 1/3 octave band [Hz]																		
		50	63	80	100	125	160	200	250	315	400	500	630	800	1000	1250	1600			
30	$\Re(z_c) \pm \sigma$	926 ± 14	837 ± 9	825 ± 7	799 ± 7	770 ± 5	735 ± 5	705 ± 5	676 ± 5	652 ± 7	646 ± 10	669 ± 14	667 ± 11	618 ± 9	602 ± 12	603 ± 15	598 ± 14			
	$\Im(z_c) \pm \sigma$	-344 ± 18	-329 ± 13	-259 ± 15	-241 ± 11	-213 ± 9	-185 ± 8	-161 ± 7	-140 ± 6	-118 ± 4	-92 ± 4	-78 ± 2	-84 ± 5	-74 ± 6	-56 ± 8	-22 ± 6	-12 ± 9	-12 ± 9		
	$\Re(k') \pm \sigma$	7.3 ± 0.1	7.3 ± 0.1	7.3 ± 0.1	7.3 ± 0.1	7.3 ± 0.1	7.3 ± 0.1	7.4 ± 0.1	8.8 ± 0.1	10.6 ± 0.1	13 ± 0.1	13 ± 0.1	15.7 ± 0.1	19.1 ± 0.1	23.5 ± 0.1	28.5 ± 0.2	34.5 ± 0.2	42.9 ± 0.2	42.9 ± 0.2	
	$\Im(k') \pm \sigma$	-2.2 ± 0.1	-2.2 ± 0.1	-2.2 ± 0.1	-2.2 ± 0.1	-2.2 ± 0.1	-2.2 ± 0.1	-2.2 ± 0.1	-2.6 ± 0.1	-3 ± 0.1	-3.5 ± 0.1	-3.5 ± 0.1	-4.1 ± 0.2	-4.8 ± 0.2	-5.7 ± 0.2	-6.7 ± 0.2	-7.8 ± 0.2	-9.4 ± 0.3	-9.4 ± 0.3	
	$\Re(z_c) \pm \sigma$	1176 ± 29	1097 ± 22	1066 ± 19	1025 ± 12	967 ± 12	890 ± 27	832 ± 33	772 ± 41	737 ± 39	735 ± 26	746 ± 15	736 ± 11	695 ± 9	675 ± 12	662 ± 14	653 ± 19	653 ± 19	653 ± 19	
40	$\Im(z_c) \pm \sigma$	-621 ± 26	-577 ± 23	-497 ± 24	-466 ± 34	-430 ± 37	-378 ± 32	-330 ± 22	-278 ± 12	-221 ± 11	-180 ± 21	-164 ± 20	-145 ± 18	-123 ± 20	-113 ± 19	-72 ± 16	-55 ± 18	-55 ± 18	-55 ± 18	
	$\Re(k') \pm \sigma$	6.2 ± 0.3	6.2 ± 0.3	6.2 ± 0.3	6.2 ± 0.3	6.2 ± 0.3	6.2 ± 0.3	6 ± 0.3	4.6 ± 0.2	2.4 ± 0.2	0.3 ± 0.2	3.4 ± 0.2	7.3 ± 0.2	12.3 ± 0.2	17.9 ± 0.3	24.6 ± 0.3	33.9 ± 0.3	33.9 ± 0.3	33.9 ± 0.3	
	$\Im(k') \pm \sigma$	-4.2 ± 0.3	-4.2 ± 0.3	-4.2 ± 0.3	-4.2 ± 0.3	-4.2 ± 0.3	-4.2 ± 0.3	-4.3 ± 0.3	-4.9 ± 0.4	-5.5 ± 0.4	-6.4 ± 0.4	-7.2 ± 0.4	-8.3 ± 0.4	-9.6 ± 0.5	-11 ± 0.5	-12.6 ± 0.6	-14.9 ± 0.6	-14.9 ± 0.6	-14.9 ± 0.6	
	$\Re(z_c) \pm \sigma$	1017 ± 13	1246 ± 77	1403 ± 55	1345 ± 67	1234 ± 62	1135 ± 45	1033 ± 34	946 ± 16	887 ± 17	854 ± 15	839 ± 4	819 ± 12	772 ± 21	742 ± 28	721 ± 29	706 ± 30	706 ± 30	706 ± 30	706 ± 30
	$\Im(z_c) \pm \sigma$	-967 ± 158	-714 ± 60	-706 ± 22	-740 ± 43	-686 ± 59	-610 ± 51	-546 ± 56	-462 ± 47	-391 ± 24	-340 ± 4	-301 ± 17	-251 ± 23	-205 ± 22	-183 ± 18	-125 ± 17	-88 ± 13	-88 ± 13	-88 ± 13	-88 ± 13
50	$\Re(k') \pm \sigma$	3.6 ± 0.2	3.6 ± 0.2	3.6 ± 0.2	3.6 ± 0.2	3.6 ± 0.2	3.6 ± 0.2	3.3 ± 0.2	1.5 ± 0.2	1 ± 0.3	4.2 ± 0.3	7.7 ± 0.3	12 ± 0.3	17.5 ± 0.3	23.7 ± 0.3	31.1 ± 0.2	41.3 ± 0.2	41.3 ± 0.2	41.3 ± 0.2	
	$\Im(k') \pm \sigma$	-6.2 ± 0.3	-6.2 ± 0.3	-6.2 ± 0.3	-6.2 ± 0.3	-6.2 ± 0.3	-6.2 ± 0.3	-6.3 ± 0.3	-7.2 ± 0.3	-8.4 ± 0.4	-9.7 ± 0.4	-11 ± 0.4	-12.6 ± 0.5	-14.5 ± 0.5	-16.5 ± 0.6	-18.8 ± 0.6	-21.6 ± 0.7	-21.6 ± 0.7	-21.6 ± 0.7	
	$\Re(z_c) \pm \sigma$	1327 ± 92	1211 ± 77	1339 ± 107	1615 ± 157	1569 ± 120	1373 ± 92	1255 ± 94	1182 ± 77	1096 ± 68	1032 ± 52	994 ± 38	959 ± 26	899 ± 17	858 ± 18	833 ± 19	808 ± 20	808 ± 20	808 ± 20	808 ± 20
	$\Im(z_c) \pm \sigma$	-2089 ± 140	-1593 ± 146	-992 ± 144	-898 ± 17	-948 ± 40	-886 ± 60	-724 ± 50	-638 ± 52	-561 ± 51	-504 ± 55	-446 ± 55	-368 ± 51	-295 ± 45	-261 ± 40	-185 ± 35	-133 ± 30	-133 ± 30	-133 ± 30	-133 ± 30
	$\Re(k') \pm \sigma$	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.5 ± 0.5	2 ± 0.6	5 ± 0.6	8.7 ± 0.5	12.8 ± 0.5	17.6 ± 0.5	23.8 ± 0.5	30.2 ± 0.5	38.8 ± 0.6	44.3 ± 8.7	44.3 ± 8.7	44.3 ± 8.7	44.3 ± 8.7
60	$\Im(k') \pm \sigma$	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-8 ± 0.2	-9.4 ± 0.3	-11.2 ± 0.5	-13.1 ± 0.6	-15 ± 0.6	-17.3 ± 0.7	-20.1 ± 0.8	-23.3 ± 1	-26.1 ± 0.9	-29.3 ± 0.5	-29.3 ± 0.5	-29.3 ± 0.5	
	$\Re(z_c) \pm \sigma$	1959 ± 70	1920 ± 77	1671 ± 114	1529 ± 121	1568 ± 90	1588 ± 78	1579 ± 71	1367 ± 58	1300 ± 46	1219 ± 33	1178 ± 21	1144 ± 11	1078 ± 7	1035 ± 8	1005 ± 9	972 ± 10	972 ± 10	972 ± 10	972 ± 10
	$\Im(z_c) \pm \sigma$	-3146 ± 136	-2783 ± 139	-2284 ± 71	-1793 ± 32	-1393 ± 49	-1084 ± 54	-1033 ± 46	-933 ± 49	-741 ± 54	-687 ± 55	-591 ± 51	-479 ± 47	-369 ± 43	-326 ± 37	-223 ± 33	-154 ± 30	-154 ± 30	-154 ± 30	-154 ± 30
	$\Re(k') \pm \sigma$	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2.4 ± 0.7	5.3 ± 0.7	8.8 ± 0.7	13.8 ± 0.7	18.3 ± 0.7	24.1 ± 0.7	31.1 ± 0.7	40.1 ± 0.7	60.1 ± 4.3	58.6 ± 15.7	58.6 ± 15.7	58.6 ± 15.7	58.6 ± 15.7
	$\Im(k') \pm \sigma$	-8.5 ± 0.3	-8.5 ± 0.3	-8.5 ± 0.3	-8.5 ± 0.3	-8.5 ± 0.3	-8.5 ± 0.3	-8.8 ± 0.3	-11.3 ± 0.3	-12.8 ± 0.5	-15.7 ± 0.6	-18.3 ± 0.6	-21.5 ± 0.8	-25.4 ± 0.8	-30.6 ± 0.9	-33.6 ± 0.8	-31.1 ± 0.5	-31.1 ± 0.5	-31.1 ± 0.5	-31.1 ± 0.5
70	$\Re(z_c) \pm \sigma$	2370 ± 106	2436 ± 123	2210 ± 119	2019 ± 133	1806 ± 113	1670 ± 88	1647 ± 85	1712 ± 69	1533 ± 54	1421 ± 41	1393 ± 29	1350 ± 22	1271 ± 17	1217 ± 20	1171 ± 23	1128 ± 15	1128 ± 15	1128 ± 15	1128 ± 15
	$\Im(z_c) \pm \sigma$	-4111 ± 138	-3660 ± 73	-3131 ± 87	-2661 ± 38	-2179 ± 46	-1647 ± 58	-1208 ± 51	-1088 ± 56	-1090 ± 55	-863 ± 55	-798 ± 51	-629 ± 49	-476 ± 41	-425 ± 35	-292 ± 33	-188 ± 26	-188 ± 26	-188 ± 26	-188 ± 26
	$\Re(k') \pm \sigma$	0.7 ± 0.4	0.7 ± 0.4	0.7 ± 0.4	0.7 ± 0.4	0.7 ± 0.4	0.8 ± 0.5	1.6 ± 0.5	8 ± 0.5	15 ± 0.5	17.9 ± 0.5	26.7 ± 0.4	32.3 ± 0.4	40.4 ± 0.5	51.7 ± 3.5	64.3 ± 8.2	61.1 ± 16.9	61.1 ± 16.9	61.1 ± 16.9	61.1 ± 16.9
	$\Im(k') \pm \sigma$	-8.4 ± 0.2	-8.4 ± 0.2	-8.4 ± 0.2	-8.4 ± 0.2	-8.4 ± 0.2	-8.4 ± 0.2	-8.4 ± 0.2	-9.6 ± 0.2	-14.5 ± 0.3	-17.4 ± 0.3	-20.4 ± 0.4	-26.1 ± 0.5	-30.6 ± 0.7	-31.3 ± 0.6	-34.3 ± 0.5	-30.6 ± 0.2	-30.6 ± 0.2	-30.6 ± 0.2	-30.6 ± 0.2
	$\Re(z_c) \pm \sigma$	3347 ± 88	3630 ± 101	3420 ± 98	3312 ± 110	3043 ± 92	2696 ± 70	2383 ± 71	2142 ± 58	1950 ± 35	1781 ± 25	1807 ± 19	1645 ± 16	1594 ± 17	1543 ± 18	1435 ± 13	1435 ± 13	1435 ± 13	1435 ± 13	1435 ± 13
80	$\Im(z_c) \pm \sigma$	-5038 ± 154	-4757 ± 78	-4126 ± 85	-3645 ± 39	-3116 ± 55	-2558 ± 52	-2079 ± 54	-1577 ± 54	-1279 ± 57	-1242 ± 57	-1002 ± 53	-848 ± 44	-596 ± 40	-531 ± 35	-364 ± 33	-234 ± 25	-234 ± 25	-234 ± 25	-234 ± 25
	$\Re(k') \pm \sigma$	2.6 ± 0.5	2.6 ± 0.5	2.6 ± 0.5	2.6 ± 0.5	2.6 ± 0.5	2.6 ± 0.5	2.2 ± 0.5	3.1 ± 0.6	13.2 ± 0.5	23 ± 0.5	27.5 ± 0.6	38.2 ± 0.5	51.2 ± 0.6	51.8 ± 0.6	73.4 ± 2.6	55.8 ± 9.2	55.8 ± 9.2	55.8 ± 9.2	55.8 ± 9.2
	$\Im(k') \pm \sigma$	-13.4 ± 0.3	-13.4 ± 0.3	-13.4 ± 0.3	-13.4 ± 0.3	-13.4 ± 0.3	-13.4 ± 0.3	-13.1 ± 0.4	-10.5 ± 0.5	-11.3 ± 0.7	-16.3 ± 0.8	-22.1 ± 0.8	-22.6 ± 0.9	-31.9 ± 1.1	-33.2 ± 1.1	-33.5 ± 0.7	-29.7 ± 0.4	-29.7 ± 0.4	-29.7 ± 0.4	-29.7 ± 0.4
	$\Re(z_c) \pm \sigma$	1327 ± 92	1211 ± 77	1339 ± 107	1615 ± 157	1569 ± 120	1373 ± 92	1255 ± 94	1182 ± 77	1096 ± 68	1032 ± 52	994 ± 38	959 ± 26	899 ± 17	858 ± 18	833 ± 19	808 ± 20	808 ± 20	808 ± 20	808 ± 20
	$\Im(z_c) \pm \sigma$	-2089 ± 140	-1593 ± 146	-992 ± 144	-898 ± 17	-948 ± 40	-886 ± 60	-724 ± 50	-638 ± 52	-561 ± 51	-504 ± 55	-446 ± 55	-368 ± 51	-295 ± 45	-261 ± 40	-185 ± 35	-133 ± 30	-133 ± 30	-133 ± 30	-133 ± 30
90	$\Re(k') \pm \sigma$	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.8 ± 0.5	0.5 ± 0.5	2 ± 0.6	5 ± 0.6	8.7 ± 0.5	12.8 ± 0.5	17.6 ± 0.5	23.8 ± 0.5	30.2 ± 0.5	38.8 ± 0.6	44.3 ± 8.7	44.3 ± 8.7	44.3 ± 8.7	44.3 ± 8.7
	$\Im(k') \pm \sigma$	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-7.8 ± 0.3	-8 ± 0.2	-9.4 ± 0.3	-11.2 ± 0.5	-13.1 ± 0.6	-15 ± 0.6	-17.3 ± 0.7	-20.1 ± 0.8	-23.3 ± 1	-26.1 ± 0.9	-29.3 ± 0.5	-29.3 ± 0.5	-29.3 ± 0.5	-29.3 ± 0.5
	$\Re(z_c) \pm \sigma$	1959 ± 70	1920 ± 77	1671 ± 114	1529 ± 121	1568 ± 90	1588 ± 78	1579 ± 71	1367 ± 58	1300 ± 46	1219 ± 33	1178 ± 21	1144 ± 11	1078 ± 7	1035 ± 8	1005 ± 9	972 ± 10	972 ± 10	972 ± 10	972 ± 10
	$\Im(z_c) \pm \sigma$	-3146 ± 136	-2783 ± 139	-2284 ± 71	-1793 ± 32	-1393 ± 49	-1084 ± 54	-1033 ± 46	-933 ± 49	-741 ± 54	-687 ± 55	-591 ± 51	-479 ± 47	-369 ± 43	-326 ± 37	-223 ± 33	-154 ± 30	-154 ± 30	-154 ± 30	-154 ± 30
	$\Re(k') \pm \sigma$	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2 ± 0.6	2.4 ± 0.7	5.3 ± 0.7	8.8 ± 0.7	13.8 ± 0.7	18.3 ± 0.7	24.1 ± 0.7	31.1 ± 0.7	40.1 ± 0.7	60.1 ± 4.3	58.6 ± 15.7	58.6 ± 15.7	58.6 ± 15.7	58.6 ± 15.7

tabular, thanks to which it can be managed with ease. The possible applications are, among others, acoustic computer modelling of the equivalent fluid model or estimating the dimension dependent acoustic parameters. In this manner, the sound absorption coefficient  $\alpha$  for the thinner, 10 cm cellulose wool layer, visible in Fig. 4. was calculated with (COX, D'ANTONIO, 2016):

$$\alpha = 1 - \left| \frac{\frac{z_1}{\rho_0 c_0} \cos \psi - 1}{\frac{z_1}{\rho_0 c_0} \cos \psi + 1} \right|^2, \quad (12)$$

where  $z_1$  – rigid back absorbent surface impedance,  $z_1 = jz_c \cot(kd_1)$ ,  $d_1$  – layer thickness,  $\rho_0 c_0$  – air acoustic impedance,  $\psi$  – incidence angle of the wave.

#### 4.2. Paperboard

The single sheet of paperboard with a grammage of 2000 g/m<sup>2</sup> and a thickness of 3 mm has SRI equal to 13 dB and sound absorption not exceeding 0.05 over the whole frequency range (Fig. 6).

The dynamic Young Modulus is approximately 2.3 GPa for the machine direction and 1.1 GPa for the

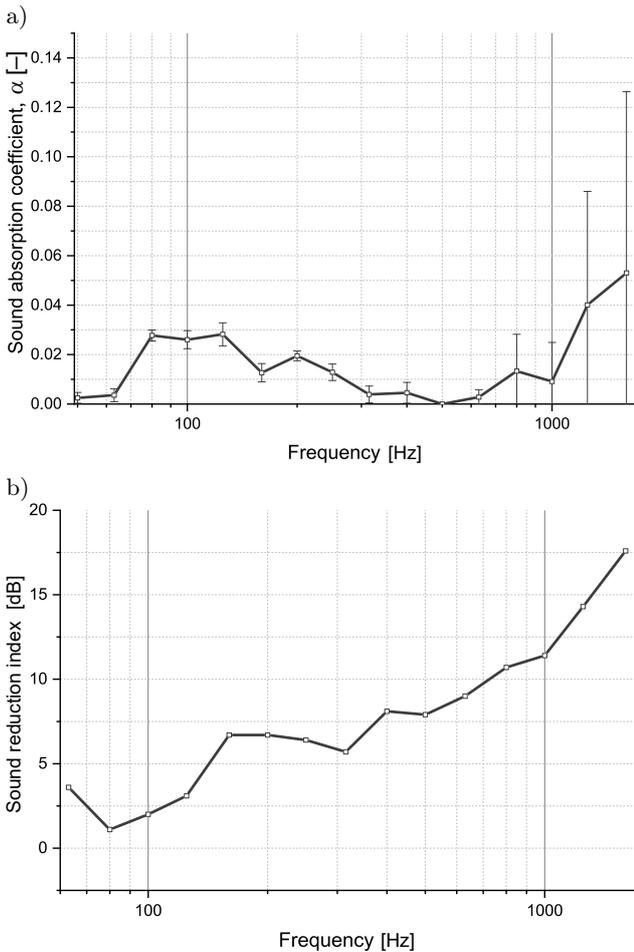


Fig. 6. Results of 2000 g/m<sup>2</sup> paperboard acoustic properties: a) sound absorption coefficient with standard error marked with horizontal lines; b) sound reduction index.

cross-machine direction and decreases slightly with frequency (Fig. 7). The disproportion between the directions is also visible in the repeatability of the results – the CD measurements are stable, which cannot be claimed in the second case, especially for lower frequencies. Both observations are easily explainable by the papermaking process: the pressure forces fibres to align parallelly to the machine direction, leading to higher elasticity and greater density unevenness. The literature provides higher values for static Young's Modulus – 5.4 GPa for MD, 1.9 GPa for CD (NISKANEN, 2011) or 2 GPa–20 GPa for MD, 0.5 GPa–10 GPa for CD (SCHÖNWÄLDER, ROTS, 2008). Unfortunately, a direct comparison is impossible, as there is no data about the paperboard dimensions or density, which may greatly impact the values. Damping properties are less correlated to frequency, e.g., for CD, the Pearson correlation coefficient is  $-0.67$  (while for Young's Modulus, it reaches  $-0.98$ ). The ratio is also direction-independent, i.e., the mean value equals 0.034 for MD and 0.036 for CD.

The acquired values suggest that in an accurate dynamic model of the paperboard, Young's Modulus

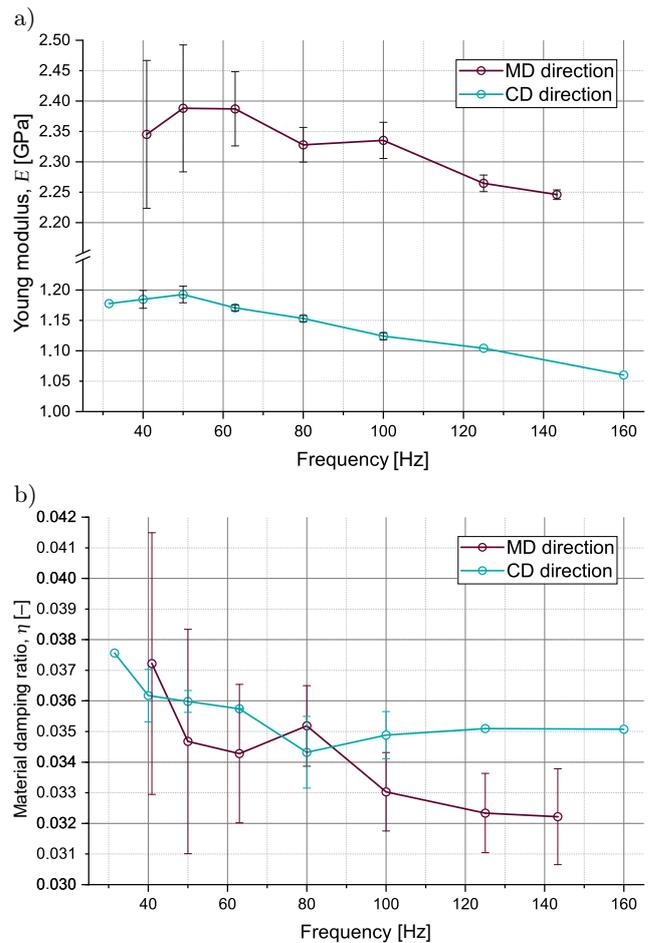


Fig. 7. Results of dynamic Young's Modulus (a) and damping ratio (b) of paperboard with grammage of 2000 g/m<sup>2</sup>. Horizontal lines mark standard error of measurements.

values should relate to the frequency, while damping can be represented by a single value.

### 4.3. Corrugated cardboard

A single wall corrugated cardboard panel have a SRI of 7 dB (for type B) and 9 dB (for type E). Despite the comparable surface mass, the second type reaches a noticeably larger value, most likely due to its more compact structure. This difference disappears when introducing the additional wall – the EE, EB, and BC plates have SRI equal to 11 dB, 12 dB, and 12 dB, respectively (Fig. 8).

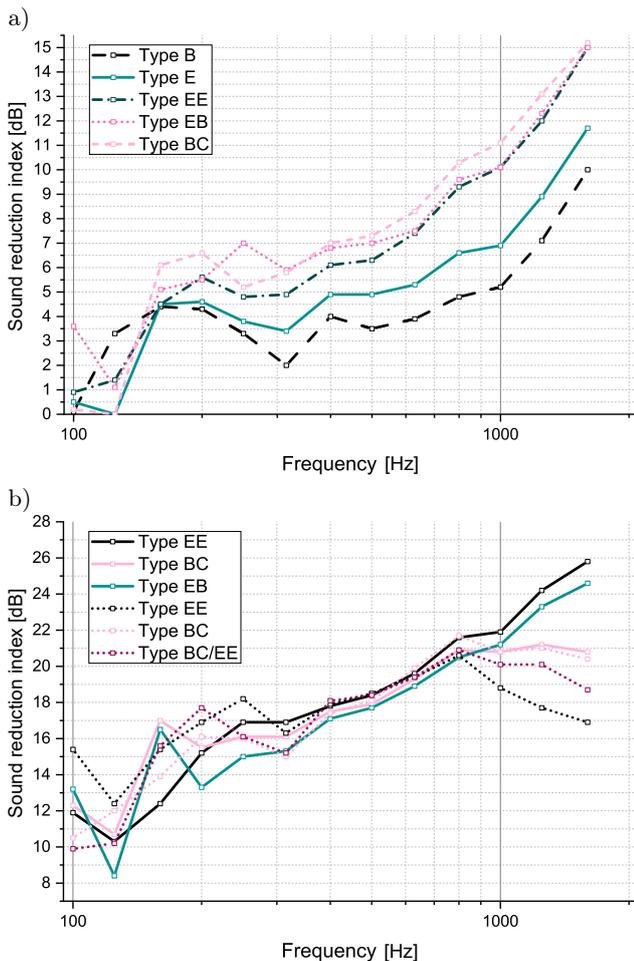


Fig. 8. Results of the sound reduction index of corrugated cardboard: a) the single plate with 3 layers (dashed lines) or 5 layers (solid lines); b) composition of 5 plates stacked without an adhesive (dotted lines) or glued with PVA (solid lines).

The multiple layered samples were arranged with flutes oriented parallel to each other. The stacked panels (Fig. 8) had no adhesive on any part of the surface. The amount of PVA adhesive used in the second case was  $150 \text{ g/m}^2$ , which had an imperceptible effect on the total weight after drying. The SRI in each case is remarkably similar and ranges from 19 dB to 22 dB (see Table 6). As with the double-wall panel, the flute

type has negligible impact on the results, even if the most compact EE layers are arranged alternately with thick BC layers. The only distinction between the results lies above the 1000 Hz, where the SRI of stacked panels decreases rapidly.

The overall insulation property of a 5-layered structure is relatively high, especially when compared to a very low surface mass of the whole structure, which is about  $2.8 \text{ kg/m}^2$ . The resulting SRI could match the properties of heavier walls, for example, a double 3 mm beaverboard panel with a filling of 30 mm Styrofoam and SRI equal to 21 dB has more than twice the surface mass of  $6 \text{ kg/m}^2$  (ŁATKA et al., 2022).

### 4.4. Honeycomb panels

The SRI of honeycomb panels is consistent with the law of mass and similar for almost all cases (Table 6). Heavier Kraftliber sheets have a weighted index higher by 2 dB than the corresponding panels of

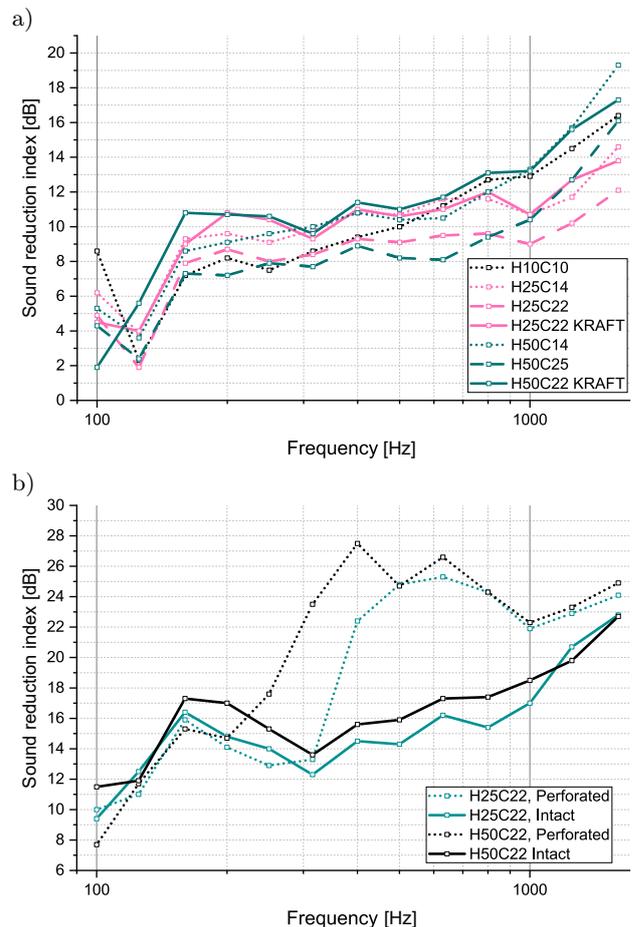


Fig. 9. Sound reduction index of honeycomb panels: a) comparison of different single panel types. Dotted lines: dense structures with cell size of 10 mm and 14 mm, dashed lines: lightweight structures with cell size of 22 mm, solid lines: Kraftliber structures with cell size of 22 mm; b) the influence of perforations on the double wall insulation. Dotted lines: perforated honeycomb insertion, solid lines: unperforated honeycomb insertion.

Table 6. Tested materials and used measurement methods with assessed parameters.

No.	Material	Thickness [mm]	Surface mass [kg/m <sup>2</sup> ]	Weighted SRI* [dB]	Spectrum adaption terms [dB]		Average sound absorption coefficient**	
					C*	C <sub>tr</sub> *		
1.	Cellulose wool made from recycled paper [kg/m <sup>3</sup> ]	Density 30	200	6.0	–	–	–	0.93
		Density 40	200	8.0	–	–	–	0.91
		Density 50	200	10.0	–	–	–	0.84
		Density 60	200	12.0	–	–	–	0.76
		Density 70	200	14.0	–	–	–	0.69
		Density 80	200	16.0	–	–	–	0.61
		Density 90	200	18.0	–	–	–	0.52
2.	Paperboard	3.0	2.0	<b>13</b>	–1	–3	0.03	
3.	Corrugated cardboard	Type BC	6.1	0.5	<b>12</b>	–1	–3	–
		Type EE	2.9	0.6	<b>11</b>	0	–2	–
		Type EB	4.0	0.6	<b>12</b>	–1	–3	–
		Type E	1.6	0.3	<b>9</b>	–1	–2	–
		Type B	2.7	0.3	<b>7</b>	0	–2	–
4.	Corrugated cardboard – five PVA glued layers	Type BC	30.5	2.7	<b>20</b>	–1	–1	–
		Type EE	14.5	2.8	<b>21</b>	–1	–3	–
		Type EB	20	2.8	<b>22</b>	–1	–3	–
5.	Corrugated cardboard – five stacked layers	Type BC	30.5	2.7	<b>21</b>	–1	–2	–
		Type EE	14.5	2.8	<b>19</b>	0	–1	–
		Type BC/EE	24.1	2.8	<b>20</b>	–1	–2	–
6.	Honeycomb panel	H50C22	50	1.0	<b>13</b>	–1	–3	–
		H50C14	50	1.6	<b>15</b>	0	–3	–
		H25C22	25	0.6	<b>11</b>	0	–1	–
		H25C14	25	0.9	<b>13</b>	0	–2	–
		H10C10	10	0.6	<b>14</b>	–1	–2	–
		H50C22 KRAFT	50	1.5	<b>15</b>	0	–2	0.14
		H25C22 KRAFT	25	1.1	<b>13</b>	0	–2	0.19
7.	Perforated honeycomb panel	H50C22 KRAFT	50	1.5	–	–	–	0.59
		H25C22 KRAFT	25	1.1	–	–	–	0.44
8.	Double wall with intact and perforated honeycomb panel	H50C22 Intact	90	2.2	<b>20</b>	0	–2	–
		H25C22 Intact	40	1.8	<b>19</b>	0	–2	–
		H50C22 Perforated	90	2.2	<b>26</b>	–2	–5	–
		H25C22 Perforated	40	1.8	<b>23</b>	–1	–3	–

\* Weighted sound reduction index and spectrum adaption terms are approximated. See Subsec. 3.2 for explanation.

\*\* For 1/3 octave bands 200 Hz ÷ 1600 Hz.

the same geometry (15 dB versus 13 dB for a thickness of 50 mm and 13 dB versus 11 dB for a thickness of 25 mm). The exact rule applies to the core mass increase. By reducing the cell size, the surface gains an additional 1/3 of its mass and 2 dB in insulation. The panel with the finest structure (H10C10) breaks this dependency with the SRI equal to 14 dB. This value is close to the performance of almost three times heavier panels (H50C22 with Kraftliber and H50C14). The frequency-dependent increase of SRI is more rapid for the H10C10 panel, thus it reaches an exceptionally high value for frequencies above 500 Hz (Fig. 9). This increase may correlate to the resonances arising in a denser structure.

As described in Subsec. 2.1, the honeycomb perforation efficiently improves its acoustic parameters.

The average sound absorption coefficient in the 1/3 octave band for a single panel (Fig. 10a), reaches the values:

- 0.98 for the H50C22 KRAFT sample. The calculated frequency of 730 Hz fits the reported band of 800 Hz;
- 0.90 for the H25C22 KRAFT sample. In this case, the frequency of 1260 Hz also corresponds to the observed 1250 Hz.

The perforation induced a slight reduction in transmission loss, equalling 2 dB–3 dB in the whole frequency band (Fig. 10b).

Figure 9b displays the improvement in the sound reduction index of the honeycomb double wall with the perforated insertion over the intact one (as proposed in Fig. 2b). In the case of panel H25C22, the rapid insulation increase is observable from the 400 Hz frequency

band and results in an overall 4 dB gain in full-band weighted SRI (from 19 dB to 23 dB Table 6). The improvement of H50C22 insertion is even more substantial. The growth is visible from the 315 Hz frequency band and causes a weighted SRI shift from 20 dB to 26 dB. Unfortunately, since the gain is achieved only in part of the whole frequency range, both Spectrum Adaption Terms (C and Ctr) deteriorate by 1 dB–3 dB.

The SRI increase is compared to the analytical result in Fig. 11. The calculations were performed with the following assumptions:

- the insertion moves with the cover panel (its weight is calculated to the cover wall);
- the distance between the double walls is reduced by insertion panel thickness in an intact case;
- the Helmholtz resonator damping ratio is 0.2;
- the probe was affected by the diffused field, with waves incident at the angle range from 0° to 80°.

The analytical model aligns very well with the experiment result, proving that the perforated honeycomb panel can be successfully represented by the proposed analytical model.

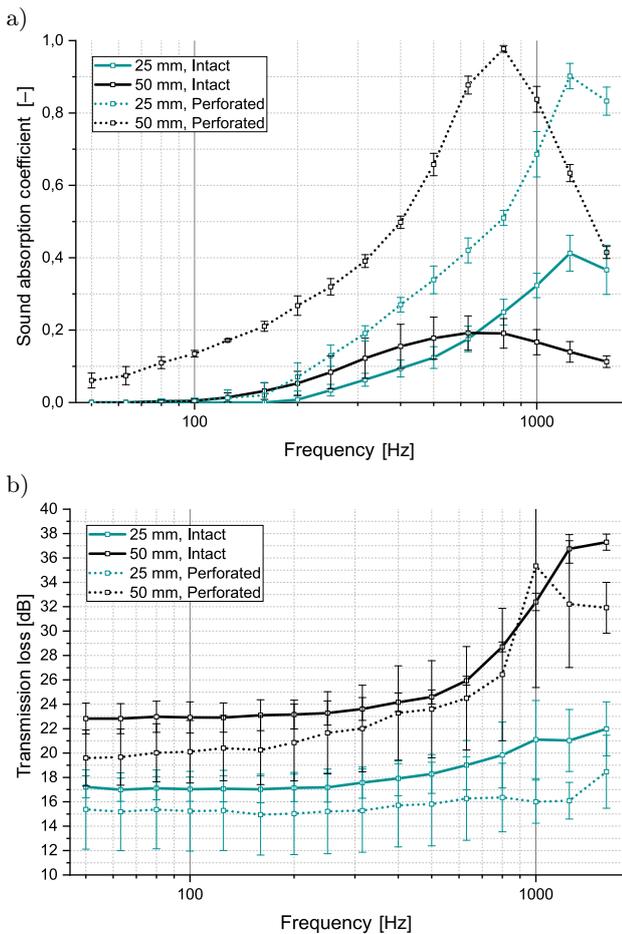


Fig. 10. The results of the sound absorption coefficient (a) and transmission loss (b) of single plate honeycomb panel, with (dotted lines) or without (solid lines) perforation. Horizontal lines mark standard error of measurements.

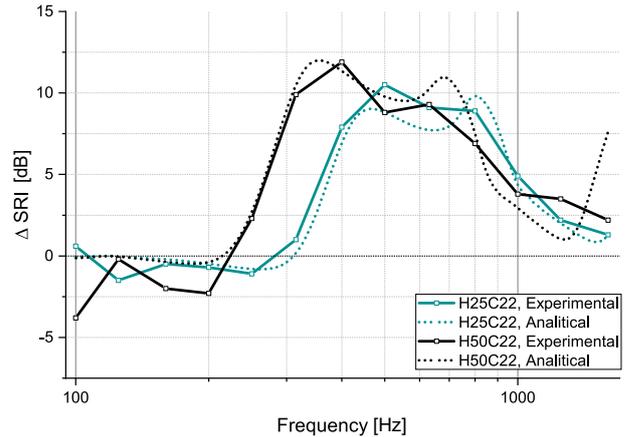


Fig. 11. Gain in sound reduction coefficient due to perforating the honeycomb insertion plate. Solid line – experimental results, dotted line – analytical results calculated with Eq. (4).

#### 4.5. Summarised results

Table 6 recapitulates the main results from the performed tests. The sound insulation and sound absorption parameters are compared along with the surface mass and thickness of the sample.

### 5. Conclusions

This work is a comprehensive study of paper and paper-based products’ acoustic parameters, including the absorption, insulation, and dynamic mechanical properties.

Paper and paper-based products prove to be promising substitutes for conventional construction materials, especially concerning the wide manufacture selection on the market. The primary findings are individual for every material type:

- cellulose wool is widely used as a thermal and acoustic insulator by the time mentioned. Its sound absorption coefficient reaches values similar to fibreglass;
- the dynamic mechanical properties of paperboard depend on the machine manufacture direction and frequency. The SRI of the plate with a surface mass of 2 kg/m<sup>2</sup> equals 13 dB;
- the SRI of multiple layered corrugated cardboard may approach relatively high values. Five layered (10 walls) structure has an SRI of 19 dB–22 dB while maintaining a low surface mass – below 3 kg/m<sup>2</sup>. The insulation does not differ significantly on the wall type but can be improved, in high frequencies, by gluing the layers;
- the honeycomb panels’ structure gives a possibility for easy acoustic parameters enhancement. Perforation produces a grid of Helmholtz resonators, which triggers an increase in the sound

absorption coefficient. When placed in a double-panel, the construction becomes Acoustic Meta-material with a negative bulk modulus, which results in SRI shift up to 6 dB. Such structures have insulation comparable to that of over three times heavier, single 12.5 mm plasterboard (27 dB (ŁĄTKA *et al.*, 2022)).

### Acknowledgments

This research was supported by the National Centre for Research and Development of Poland (the grant number Lider/60/0250/L-11/19/NCBR/2020).

The author Jerzy F. Łątka was supported by the Polish National Agency for Academic Exchange under the Bekker NAWA Programme (the grant number PPN/BEK/2018/1/00383).

The authors thank the KFB Acoustics for providing with the laboratory equipment, necessary to conduct reliable tests.

### References

1. ABD RASHID A.F., YUSOFF S. (2015), A review of life cycle assessment method for building industry, *Renewable and Sustainable Energy Reviews*, **45**: 244–248, doi: [10.1016/j.rser.2015.01.043](https://doi.org/10.1016/j.rser.2015.01.043).
2. ARENAS J.P., REBOLLEDO J., DEL REY R., ALBA J. (2014), Sound absorption properties of unbleached cellulose loose-fill insulation material, *BioResources*, **9**(4): 6227–6240, doi: [10.15376/biores.9.4.6227-6240](https://doi.org/10.15376/biores.9.4.6227-6240).
3. ASDRUBALI F., PISELLO A.L., D’ALESSANDRO F., BIANCHI F., CORNICCHIA M., FABIANI C. (2015), Innovative cardboard based panels with recycled materials from the packaging industry: Thermal and acoustic performance analysis, *Energy Procedia*, **78**: 321–326, doi: [10.1016/j.egypro.2015.11.652](https://doi.org/10.1016/j.egypro.2015.11.652).
4. ASTM E756-05 (2005), *Standard test method for measuring vibration-damping properties of materials*, ASTM Standards (Reapproved 2017).
5. ASTM E2611-19 (2019), *Standard test method for normal incidence determination of porous material acoustical properties based on the transfer matrix method*, ATM International.
6. CRIPPS A. (2004), Cardboard as a construction material: A case study, *Building Research & Information*, **32**(3): 207–219, doi: [10.1080/09613210410001686273](https://doi.org/10.1080/09613210410001686273).
7. COX T.J., D’ANTONIO P. (2016), *Acoustic Absorbers and Diffusers: Theory, Design and Application*, 3rd ed., pp. 91–131, Spon Press.
8. DIARTE J., SHAFFER M. (2021), Cardboard architecture. Eight decades of exploration in academic research and professional practice 1940–2019, *Enquiry The ARCC Journal for Architectural Research*, **18**(1): 17–40, doi: [10.17831/enqarcc.v18i1.1103](https://doi.org/10.17831/enqarcc.v18i1.1103).
9. International Organization for Standardization (2000), Acoustics. Measurement of sound insulation in buildings and of building elements using sound intensity. Part 1: Laboratory measurements, (ISO Standard No. 15186-1:2000), <https://www.iso.org/standard/26097.html>.
10. JASIOLEK A., NOSZCZYK P., ŁĄTKA J.F. (2023), Paper-based building envelopes – Thermal and environmental properties of original envelope designs, *Energy and Buildings*, **289**: 113062, doi: [10.1016/j.enbuild.2023.113062](https://doi.org/10.1016/j.enbuild.2023.113062).
11. KANG C.-W., KIM M.K., JANG E.-S. (2021), An experimental study on the performance of corrugated cardboard as a sustainable sound-absorbing and insulating material, *Sustainability*, **13**(10): 5546, doi: [10.3390/su13105546](https://doi.org/10.3390/su13105546).
12. KLEMM D., HEUBLEIN B., FINK H.P., BOHN A. (2005), Cellulose: Fascinating biopolymer and sustainable raw material, *Angewandte Chemie – International Edition*, **44**(22): 3358–3393, doi: [10.1002/anie.200460587](https://doi.org/10.1002/anie.200460587).
13. KNAACK U., BACH R., SCHABEL S. [Eds.] (2023), *Building with Paper Architecture and Construction*, Birkhäuser.
14. LANGFELDT F., HOPPEN H., GLEINE W. (2020), Broadband low-frequency sound transmission loss improvement of double walls with Helmholtz resonators, *Journal of Sound and Vibration*, **476**: 115309, doi: [10.1016/j.jsv.2020.115309](https://doi.org/10.1016/j.jsv.2020.115309).
15. LATKA J.F. (2017a), House of Cards – design and implementation of a paper house prototype, [in:] *Interfaces: Architecture. Engineering. Science*, pp. 1–10.
16. LATKA J.F. (2017b), *Paper in Architecture: Research by Design, Engineering and Prototyping*, A+BE | Architecture and the Built Environment.
17. ŁĄTKA J.F. *et al.* (2022), Properties of paper-based products as a building material in architecture – An interdisciplinary review, *Journal of Building Engineering*, **50**: 104135, doi: [10.1016/j.jobe.2022.104135](https://doi.org/10.1016/j.jobe.2022.104135).
18. MIYAKE R., LUNA I., GOULD L.A. [Eds.] (2009), *Shigeru Ban: Paper in Architecture*, Rizzoli International Publications.
19. NERI M., LEVI E., CUERVA E., PARDO-BOSCH F., ZABALETA A.G., PUJADAS P. (2021), Sound absorbing and insulating low-cost panels from end-of-life household materials for the development of vulnerable contexts in circular economy perspective, *Applied Sciences*, **11**(12): 5372, doi: [10.3390/app11125372](https://doi.org/10.3390/app11125372).
20. NISKANEN K. [Ed.] (2011), *Mechanics of Paper Products*, De Gruyter, doi: [10.1515/9783110254631](https://doi.org/10.1515/9783110254631).

21. RICCIARDI P., BELLONI E., COTANA F. (2014), Innovative panels with recycled materials: Thermal and acoustic performance and Life Cycle Assessment, *Applied Energy*, **134**: 150–162, doi: [10.1016/j.apenergy.2014.07.112](https://doi.org/10.1016/j.apenergy.2014.07.112).
22. SCHÖNWÄLDER J., ROTS J.G. (2008), Mechanical behaviour of cardboard in construction, *Cardboard in Architecture*, **7**: 131, doi: [10.3233/978-1-58603-820-5-131](https://doi.org/10.3233/978-1-58603-820-5-131).
23. SECCHI S., ASDRUBALI F., CELLAI G., NANNIPIERI E., ROTILI A., VANNUCCHI I. (2016), Experimental and environmental analysis of new sound-absorbing and insulating elements in recycled cardboard, *Journal of Building Engineering*, **5**: 1–12, doi: [10.1016/j.jobee.2015.10.005](https://doi.org/10.1016/j.jobee.2015.10.005).
24. WANG Y., JIANG Z., LI L., QI Y., SUN J., JIANG Z. (2023), A bibliometric and content review of carbon emission analysis for building construction, *Buildings*, **13**(1): 1–22, doi: [10.3390/buildings13010205](https://doi.org/10.3390/buildings13010205).
25. WOLF A., REBECCA B., NIHAT K., KNAACK U., WILFINGER M. (2021), A full performance paper house, *Journal of Facade Design and Engineering*, **9**(1): 117–130, doi: [10.7480/jfde.2021.1.5533](https://doi.org/10.7480/jfde.2021.1.5533).
26. ZABALZA BRIBIÁN I., VALERO CAPILLA A., ARANDA USÓN A. (2011), Life cycle assessment of building materials: Comparative analysis of energy and environmental impacts and evaluation of the eco-efficiency improvement potential, *Building and Environment*, **46**(5): 1133–1140, doi: [10.1016/j.buildenv.2010.12.002](https://doi.org/10.1016/j.buildenv.2010.12.002).



## Research Paper

## How Much Voicing in Voiced Gemimates? The Laryngeal Voicing Profile of Polish Double Stops

Arkadiusz ROJCZYK\*<sup>id</sup>, Andrzej PORZUCZEK<sup>id</sup>Speech Processing Laboratory, University of Silesia in Katowice  
Sosnowiec, Poland; e-mail: andrzej.porzuczek@us.edu.pl

\*Corresponding Author e-mail: arkadiusz.rojczyk@us.edu.pl

(received May 15, 2023; accepted March 5, 2024; published online May 29, 2024)

Gemimates (such as the double /k/ in Polish *lekki* “light”) form a group of consonants that are mainly characterized by longer durations than the corresponding singletons. Most of the research has concentrated on durational and spectral properties of gemimates in contrast to singletons. Much less attention has been paid to the realization of the voicing contrast in gemimates and whether it is differently implemented than in singletons. In the current study, we contribute to this research with the data from Polish stop gemimates. To this end, a total of 49 native speakers of Polish produced all stop gemimates and corresponding singletons in wordforms of the same phonological make-up. The measurements included closure duration, voicing ratio, duration, and mean intensity of the release burst. The results showed that the voicing ratio was 0.69, classifying Polish stop gemimates as mildly devoiced. There was a significant speaker-dependent variability in that some speakers devoiced all gemimates, while others either partially devoiced or never devoiced. The analysis of interactions between gemimates and singletons revealed that gemimates cancelled voicing cues observed in singletons such as longer durations and lower intensity of the release burst. We discuss the current results in terms of voicing implementation in Polish and in relation to other geminating languages.

**Keywords:** gemimates; Polish; voicing; stops; speech production.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

### 1. Introduction

Gemimates are a group of consonants that are characterized by longer constrictions than singletons and for this reason they are frequently referred to as long consonants (HANKAMER *et al.*, 1989). For example, in Polish, gemimates contrast with their corresponding singleton counterparts in pairs such as *lekki* “light” vs. *leki* “medicines” or *ranna* “wounded” (fem.) vs. *rana* “a wound”. The term encompasses both “true” (monomorphemic) and “fake” (assimilated or concatenated) gemimates (e.g., OH, REDFORD, 2012). Most of the research on gemimates in the world’s languages (reviews in (KAWAHARA, 2015; KUBOZONO, 2017)) has concentrated on temporal properties of the singleton/geminate contrast itself (e.g., AMANO, HIRATA, 2010; ESPOSITO, DI BENEDETTO, 1999; HAMZAH *et al.*, 2016; KINGSTON *et al.*, 2009; KOTZOR *et al.*, 2016; TAGLIAPIETRA, MCQUEEN, 2010), as well as of the neighboring vowels (e.g., IDEMARU, GUION, 2008;

KAWAHARA, 2006; KINGSTON *et al.*, 2009; LAHIRI, HANKAMER, 1988; LOCAL, SIMPSON, 1999; OHALA, 2007; PICKETT *et al.*, 1999; PORT *et al.*, 1987). Much less attention has been devoted to voicing in geminate consonants. Sustaining voicing in obstruents is an articulatory challenge since it necessitates sufficient transglottal air pressure drop. This challenge appears to be especially pronounced in the case of gemimates since they have a much longer closure (OHALA, 1983), which may be the reason why cross-linguistically voiced gemimates are less frequent than their voiceless counterparts (HAYES, STERIADE, 2004) and are classified as marked (AL-TAMIMI, KHATTAB, 2018; BLEVINS, 2004; HAYES, STERIADE, 2004; OHALA, 1983; WESTBURY, KEATING, 1986).

Because longer consonant closures are typically linked to voiceless (fortis) articulations (CATFORD, 1977; JAEGER, 1978), a phonetic question emerges of how and to what extent “already-long” gemimates organize their voiced/voiceless contrast both durationally

and spectrally. A number of studies have investigated this issue with different results. BUTCHER (2004) found that in Italian stops voiced and voiceless singletons differed significantly in duration and pressure, but geminates differed only in pressure. On the other hand, AL-TAMIMI and KHATTAB (2018) reported that in Lebanese Arabic geminate duration was the most important correlate for distinguishing the four-way contrast between voicing and gemination. When investigating the amount of voicing in closure, defining it as proportion of detectable low-energy periodic activity to the total closure duration, some languages tend to devoice geminates partially or even fully, e.g., Japanese (FUJIMOTO, KATAOKA, 2016; HIROSE, ASHBY, 2007; HUSSAIN, SHINOHARA, 2019; KAWAHARA, 2006), Tashlhiyt Berber (RIDOUANE, 2010), while other languages such as Buginese (COHN *et al.*, 1999) or Egyptian Arabic (KAWAHARA, 2006) appear to maintain voicing throughout the geminate closure. Much evidence here comes from Japanese. For example, KAWAHARA (2006) reported that Japanese speakers devoiced 60 % to 70 % of geminate closures, which was in line with a later study by HIROSE, ASHBY (2007), who found 47 % of devoicing in the total closure duration in Japanese geminates. The most recent study by HUSSAIN, SHINOHARA (2019) showed that devoicing of Japanese may be even more robust at the level of 75 % to 80 % of the closure. Another issue is whether geminates follow singletons in changes in closure duration as a function of the voicing contrast in that voiced consonants are typically shorter than their voiceless counterparts. This effect has been found for Tokyo Japanese (HOMMA, 1981; IDEMARU, GUION, 2008; but see (HUSSAIN, SHINOHARA, 2019)) and Lebanese Arabic (AL-TAMIMI, KHATTAB, 2018). In the current study, we contribute to the debate on the voicing contrast in geminate obstruents by providing data from Polish, a language where fake geminates prevail but the true ones also occur.

In this study, we aim to realize the following research tasks: (1) compare the proportion of voicing in the closure between voiced geminates and voiced singletons, and (2) compare closure durations between voiced and voiceless geminates and see if they match differences in the closure durations between voiced and voiceless singletons.

## 2. The types and distribution of Polish geminates

Polish is regarded as a true geminating language which allows tautomorphic lexical geminates, as in *lekki* “light” (adj. m.) vs *leki* “medicines”. Generally, apart from plosives, also nasal, fricative, and even affricate consonants may form geminates in intervocalic and, less frequently, initial positions. Although

most geminates in Polish belong to the “fake” category, and even in less obvious cases a diachronic analysis would almost always reveal either a phonological assimilation or morphological concatenation as gemination sources, a native speaker, save a handful of linguists, is not likely to identify *lekki* or *wanna* “bathtub” as derived forms and manifest this awareness in any phonetic modification of such double consonants. Therefore these geminates gravitate towards the “true” category, which is in fact more reliably represented by numerous foreign borrowings, thus confirming the strong geminating potential of the Polish language. This potential is also illustrated, for example, by the word *fontanna* “fountain”, a Latin or Italian borrowing, which features no geminate in the source language.

It is easy to observe that the occurrence of a double consonant letter clearly indicates a geminate in Polish pronunciation. Even if the source language allows no long consonants, a double consonant letter in the original (and borrowed) spelling may trigger gemination, as in the English word *hobby*. This tendency, however, as observed by PORZUCZEK and ROJCZYK (2014), has been attenuated in more recent borrowings, e.g., *mobbing*, which come from non-geminating languages, such as English or German.

The distribution of geminates in Polish follows certain universal tendencies (cf. THURGOOD, 1993; MULLER, 2001; PAJAŁ, 2009). There are fewer sonorant than obstruent geminates, the intervocalic position is the most typical, and coronal geminates (including the nasal) outnumber representatives of the other places of articulation. Except for the coronal nasal, very frequently geminated in Polish, voiced geminates are otherwise rather rare so, in general, voiced obstruent geminates are far less frequent than their voiceless counterparts.

The more marked distributional characteristics (MADDIESON, 1985; LADEFOGED, MADDIESON, 1996; DMITRIEVA, 2009) comprise the existence of word-initial geminates, e.g., *ssak* “mammal”, including affricates, a category highly marked in this position, e.g., *dżdżownica* /dʒdʒɔvˈnitsa/ “earthworm” (see (ROJCZYK, PORZUCZEK, 2019a; 2022) for a detailed presentation of geminate types in Polish). Potential word-final geminates or consonant-adjacent ones (only possible in concatenated fake geminates) are normally degeminated (RUBACH, BOOIJ, 1990), as in the noun-to-adjective derivation *Sewilla* → *sewil* + *ski*. Even in such examples, though, degemination is not always obligatory, e.g., *bez + stronny* /-s+(s)tr-/ “im+partial” (PAJAŁ, 2009), and the reduction may not occur at all or remain incomplete.

## 3. The phonetic realization of Polish geminates

Polish geminates are realized phonetically using a prolonged consonant constriction as the primary cue.

Interestingly, quite a large proportion of affricate geminates are also pronounced with one release, while either the occlusion or the fricative phase is lengthened (THURGOOD, DEMENKO, 2003; PORZUCZEK, ROJCZYK, 2021). The geminate/singleton duration ratio varies across studies and seems to depend on the consonant class, e.g., 1.7 for affricates (THURGOOD, DEMENKO, 2003), 2.4 for stops and 2.1 for fricatives (MALISZ, 2013), up to 2.88 for nasals (ROJCZYK, PORZUCZEK, 2014).

Unlike other geminating languages, Polish allows rearticulation (also in undisputable true geminates in borrowings, e.g., *pizza*), which is more frequent in citation forms, formal speech, or at lower speech rates. Other potential rearticulation triggers may be linked to the etymology of a word and segmental context (KOZYRA, 2008). Naturally, rearticulation rate also differs depending on the consonant type. Figure 1 shows single-articulated (left) and rearticulated (right) productions of the word *Budda* /'budda/. Rearticulation is manifested in the release burst of the first consonant.

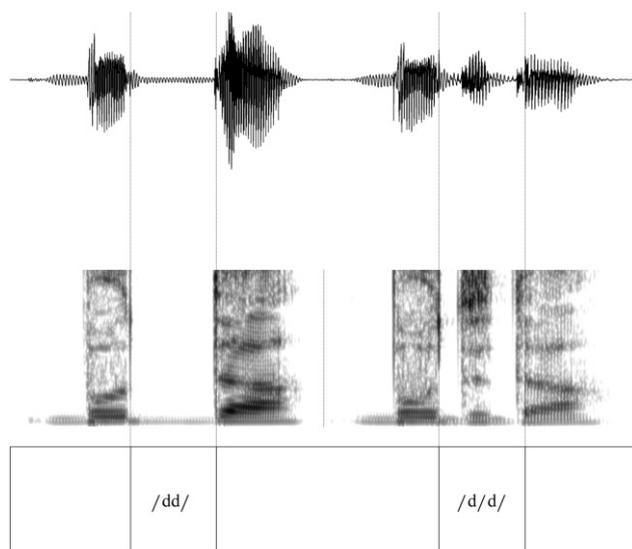


Fig. 1. Single-articulated (left) and rearticulated (right) productions of the word *Budda* /'budda/ by the first author.

ROJCZYK and PORZUCZEK (2019a) found that those with a more salient release phase (affricates and plosives) are rearticulated more often (35 % and 29 % of the cases, respectively), nasals, pronounced with an oral closure and unobstructed nasal air flow only in 18 %, while fricative geminates, as continuants, are hardly ever (2 %) separated by a vocalic insertion. It should be emphasized, however, that these mean figures must be perceived in the context of considerable cross-speaker variation, possibly reflecting the liquid boundary between natural/spontaneous and monitored speech. This boundary may also depend on particular experimental conditions or stimuli. For in-

stance, in ROJCZYK and PORZUCZEK's (2019a) experiment, devoted to a wider range of Polish geminates, the affricate rearticulation proportion (35 %) was much lower than in other studies focused particularly on affricates, which yielded a much larger rearticulation rate: 68 % in (THURGOOD, DEMENKO, 2003) and 76 % in (ROJCZYK, PORZUCZEK, 2019b). All these factors lead to various attitudes and reactions of the speakers, which is illustrated by the general results presented by ROJCZYK and PORZUCZEK (2019a), where out of 48 native Polish participants, 16 produced only single-articulated geminates, while nine rearticulated more than 50 % of the tokens. It is possible that experimental conditions lead to hyperarticulation in a certain proportion of speakers but whether or not rearticulation is marginal in spontaneous speech is a question for a separate study regarding at least the speakers' sociolinguistic background and conversational contexts.

#### 4. The phonological voicing contrast in Polish and its realization

As explained in the previous section, the realization and perception of geminates is strongly based on duration, which is also a parameter involved in cuing the phonological voicing contrast in the world's languages. This phonological contrast is naturally associated with the presence or absence of vocal fold vibration, but in fact, the presence or absence of vocal fold activity as the decisive voicing cue is only characteristic of true voicing languages, such as Catalan, Spanish or French (SOLÉ, 2007), among others. Other languages (putting aside those with more complex phonation contrasts – cf. CHO *et al.*, 2019), the aspirating ones, such as, for instance, English and three varieties of German (CHO *et al.*, 2019), may also employ VOT variation, which results in phonologically significant timing adjustments. Polish, together with other Slavic languages, belongs to the former category of voicing languages, with negligible aspiration variation even in varying prominence positions (MALISZ, ŻYGIS, 2015).

The voicing contrast in obstruents is not always realized phonetically. Word-final obstruent neutralization is a characteristic feature of Polish, as well as numerous other languages. Some languages, e.g., English, also feature word-final devoicing, but they preserve the phonological contrast by the preceding vowel length modifications. The present study focuses on the intervocalic, post-stressed position, the most typical for geminates to occur in (DMITRIEVA, 2017; and references therein). This position is rarely considered in the studies of voicing, and it is regarded as one where true voicing rather than VOT or vowel clipping is the main cue to the contrast. In Polish, the phonetic realization of word-internal intervocalic consonants re-

flects their phonological voicing category (as in *koza* /'kɔza/ “a goat” vs *kosa* /'kɔsa/ “a scythe”), although consonants in obstruent clusters tend to undergo either progressive (more typical of initial and final positions) or regressive assimilation, as in *łódka* /'wutka/ “a boat” or *prośba* /'prɔzba/ “a request”<sup>1</sup>. The post-stressed intervocalic position generally preserves the voiced-voiceless distinction in obstruents (e.g., LISKER, 1957).

As mentioned in Introduction, voiceless consonants tend to be pronounced with longer constriction time, which is also true for a voicing language such as Polish. ROJCZYK and PORZUCZEK (2019a) observed that voiceless plosive singletons were 30 % longer than their voiced counterparts. The difference amounted to 15 % in the case of geminates. The consonant length variation marking the two distinctions: voiced vs voiceless, and singleton vs geminate in Polish may interact with the characteristics of the glottal activity (vocal fold vibration) related to the constriction, and this interaction is examined in the present empirical study.

## 5. The current study

In the current study, we investigate the voicing profile of Polish stop geminates by addressing the following issues:

- 1) Are Polish phonologically voiced geminates devoiced and what is the magnitude of devoicing?
- 2) Is there observable speaker variability in the magnitude of devoicing?
- 3) Is the duration of voicing correlated with the duration of the closure?
- 4) Is there interaction of voicing and the geminate/singleton contrast for release duration and release intensity?
- 5) Are there atypical realizations of voicing in Polish stop geminates?

### 5.1. Materials

The materials included voiced and voiceless stop geminates for the three constriction positions in Polish – bilabial /bb/ vs /pp/, dental /dd/ vs /tt/, velar /gg/ vs /kk/ – and their corresponding singletons /b/ vs /p/, /d/ vs /t/, /g/ vs /k/. The targets were placed in an identical phonological context in potential Polish words represented orthographically as *Upe\_ak* /u'pɛ\_ak/. The Polish orthographic system is transparent for stop consonants by cueing pronunciation in an unambiguous way. Accordingly, each consonant had the same word frame representation as shown in Table 1.

<sup>1</sup>The realization of voicing in sandhi contexts, which are outside the scope of this paper, is more complex, and varies across regional dialects. For a phonological interpretation of the two main variants (see CYRAN, 2011).

Table 1. Structure of the materials for recording.

	Geminate				Singleton			
	Voiced		Voiceless		Voiced		Voiceless	
Bilabial	bb	<i>Upebbak</i>	pp	<i>Upeppak</i>	b	<i>Upebak</i>	p	<i>Upepak</i>
Dental	dd	<i>Upeddak</i>	tt	<i>Upettak</i>	d	<i>Upedak</i>	t	<i>Upetak</i>
Velar	gg	<i>Upeggak</i>	kk	<i>Upekkak</i>	g	<i>Upegak</i>	k	<i>Upekak</i>

The choice for nonwords with this specific phonological structure was motivated by the desire to avoid the effects of unbalanced word frequency or familiarity with token pairs (JURAFSKY *et al.*, 2001; MUNSON, SOLOMON, 2004; RAYMOND *et al.*, 2006), which may distort both durational and spectral relations between geminates and singletons. In other words, besides the desired word structure and measurement feasibility, we were looking for a context where the insertion of any plosive consonant should not form or closely resemble a real Polish word as the speakers may reduce or unnaturally boost the phonetic contrast in minimal pairs with varying lexical familiarity levels, even if associations are only based on word form similarity.

Moreover, previous research has shown that singleton/geminate contrasts are affected by the prosodic position within a word (PORT, O'DELL, 1987; AMANO, HIRATA, 2010; YOSHIDA *et al.*, 2015). In order to reduce the impact of position and prominence level, the targets were placed in a carrier phrase *Pan Adam Upe\_ak, panie prezesie* “Mr Adam Upe\_ak, Mr. President”. By using nonwords functioning as surnames, we provided plausible examples of acceptable Polish utterances with the tested item bearing phrase accent in a non-phrase-final position. Furthermore, it aided the speakers to sustain relatively stable tempo of articulation, which is especially important considering previous reports that overall tempo may influence geminate duration (PIND, 1995; PICKETT *et al.*, 1999; HIRATA, WHITON, 2005; HERMES *et al.*, 2021; YOSHIDA *et al.*, 2015; RIDOUANE, 2022). The target contrasts were interspersed with other contrasts (proportion 1 to 3) that served as distractors and were not used in the current study. They included the same word tokens with the singleton/geminate contrasts but with fricative and nasal consonants.

### 5.2. Participants

A total of 49 native speakers of Polish (39 females and 10 males) ranging in age from 19 to 22 ( $M = 20.08$ ) participated in the study. They were all students at the University of Silesia in Katowice, Poland. All of them had pronunciation features of standard Polish without any dialectal traces. The unbalanced sample reflects the gender proportions in Polish students of humanities. Recruiting university students rather than ran-

dom participants was to ensure more natural speech production and offset the effect of laboratory conditions, where the proportion of disfluencies even in short reading performance dramatically rises. None of the participants had spent more than two months outside Poland. Each speaker was paid 20 PLN (approximately 5 EUR) for their participation. None of the speakers reported any speech or hearing disorders.

### 5.3. Recording and measurements

All recordings took place in a sound-proof booth in the Speech Processing Laboratory, University of Silesia in Katowice. The speakers were seated in front of a 15-inch monitor screen which flashed the test phrases in black 28-point font against white background. The recording was self-paced in that the speakers read a phrase and proceeded to the next one by pressing an arrow key. The targets were randomized for each speaker individually. The signal was captured at 44 100 Hz through a dynamic Shure microphone fed by a Steinberg UR44 (Yamaha) audio interface unit stored as wav. files. Each session, which included other sound contrasts not used in the current study, took approximately 25 min, after which the speakers completed a questionnaire asking for basic personal data.

Measurements were taken from waveform and spectrograms using textgrids in Praat (BOERSMA, WEENINK, 2001; n.d.). As discussed earlier, Polish has both single-articulated and rearticulated geminates. Only single-articulated productions were analysed in this study, because they have uninterrupted closure duration, compared to rearticulated geminates that have two release bursts. We decided to discard rearticulations (29% in our recorded corpus) in order to obtain the results comparable to other languages which all have single-articulated geminates. Rearticulation was most frequently observed in velars (38.5%), followed by dentals (37.2%) and bilabials (24.4%). Phonological voicing had a slight impact on rearticulation with 56.4% of rearticulation in voiceless compared to 43.6% in voiced geminates. Since the constriction duration is relevant for geminate/singleton contrast only in typical, single articulation, we have not decided to analyse the duration of the two constrictions in rearticulated geminates, where the additional release burst and/or vocalic epenthesis is a very clear indication of a doubled consonant. Moreover, the difference in articulatory complexity of the two variants made us refrain from direct comparison of respective temporal patterns. Moreover, the problem with rearticulated productions is that a large part the consonant duration is filled with the release burst of the second consonant. Frequently the closure durations are very short or even incomplete, which makes temporal durations of closure and periodic activity of voicing largely impossible.

The example of rearticulation provided in Fig. 1 is an instance of careful clear speech by the first author for demonstrative purposes. In casual speech the closures (specially the first one) are either very short or even incomplete.

The following segmentation criteria were used (visualization in Fig. 2):

- 1) Closure phase was determined as a time interval between the offset of a preceding vowel indicated by a large drop in intensity and the cessation of formant structure (especially  $F2$ ) and the onset of the following release burst indicated by a sudden rise of spectral energy.
- 2) Release phase was determined as an interval beginning with as the observable rise in (frequently aperiodic) energy to the onset of the following vowel marked by the emergence of visible formant structure.
- 3) Voicing in closure was determined as a period in the closure marked by low-frequency periodic energy (voice bar).

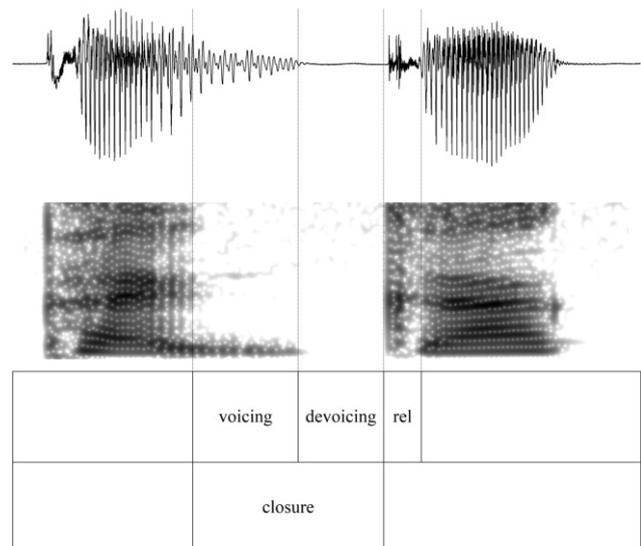


Fig. 2. Segmentation criteria based on the geminate /gg/ produced by Speaker 14 (female, age 20).

Each recorded token was inspected and measured manually by the authors. The following parameters were calculated: closure duration in milliseconds, voicing in closure in milliseconds, voicing/closure duration ratio (VC ratio: the proportion of voicing in closure to total duration of closure ranging from 0 to 1), release duration in milliseconds, and mean release intensity in dB. The release intensity was the mean intensity over the entire release burst. We found no cases of stop productions without a visible release burst. Although they may have differed in duration, there was always a visible plosion prior to the onset of the following vowel.

## 6. Analysis and results

Out of the recorded 588 targets (49 speakers  $\times$  6 geminates  $\times$  6 singletons) 41 (7%) were discarded due to observable disfluencies, resulting in 547 valid tokens (267 geminates and 280 singletons). As mentioned earlier, rearticulated geminates were precluded from further analysis (78 cases, 29%). The final corpus for acoustic measurements included 189 single-articulated geminates (92 voiced and 97 voiceless) and 280 singletons (137 voiced and 143 voiceless). We start our analysis by providing descriptive grouping analysis, which is followed by inferential statistics. We tested statistical significance of differences by using a mixed model analysis of variance in Statistica 13 (TIBCO Software Inc.) with speaker and word as random effects and the singleton/geminate contrast and the voiced/voiceless contrast as fixed effects. This is the model that utilises the method for estimating the variance of random factors which starts with constructing the sums of squares and cross product matrix for the independent variables. The sums of squares and cross products for the random effects are subsequently residualized on the fixed effects, leaving the random effects independent of the fixed effects. Between effects are tested for significance using relevant error terms based on the covariation of random sources of variation in the design using Satterthwaite's method of denominator synthesis (SEARLE *et al.*, 1992; LUKE, 2017; TIBCO Software Inc., 2017). Finally, we provide individual atypical cases of voicing/devoicing in geminates together with their acoustic descriptions.

### 6.1. Voicing profile

The mean voicing ratio in voiced geminates was 0.69 (SE = 0.03). In order to quantify the proportions of voiced, partially devoiced and devoiced productions, we provided the following categorizations: 0.00–0.49 devoiced; 0.50–0.89 partially devoiced; 0.90–1.00 voiced (for different categorizations see (DAVIDSON, 2016; ABRAMSON, WHALEN, 2017)). The analysis revealed that the distribution of the three categories was relatively equal with 33% of voiced, 38% of partially devoiced, and 29% of devoiced productions. There was an observable between-speaker variability in the realization of voicing. Seven speakers (14%) produced only fully voiced geminates, ten speakers (20%) devoiced all geminates, and the remaining thirty-two speakers (66%) produced partially devoiced geminates. The mean voicing ratio in voiced singletons was 0.91 (SE = 0.02). The majority of the productions (80%) were fully voiced, 20% were partially devoiced. No fully devoiced singletons were observed. The singletons had a significantly higher voicing ratio than the corresponding geminates [ $F(1, 39) = 38.81, p < .001, \eta_p^2 = 0.5$ ]. Table 2 presents means

Table 2. Means and standard error for the voicing ratio (ranging from 0 to 1) across the three places of articulation.

	Geminate		Singleton	
	<i>M</i>	SE	<i>M</i>	SE
Bilabial /b/	0.66	0.06	0.92	0.03
Dental /d/	0.75	0.05	0.96	0.01
Velar /g/	0.66	0.05	0.86	0.03

and standard error of the voicing ratio for each place of articulation.

The prediction, based on aerodynamic principles governing sustaining voicing in closure, was that the shorter the duration of closure, the more voiced the consonant should be. Multiple regression analysis revealed that there was no significant regression between closure duration and the voicing ratio in geminates [ $b^* = -0.17, R^2 = 0.29, F(1, 90) = 2.64, p = 0.11$ ], however, as indicated by a negative  $b^*$  value, shorter closure durations tended to predict more voicing. The same regression for singletons yielded a significant result [ $b^* = -0.18, R^2 = 0.32, F(1, 134) = 4.32, p = 0.04$ ], revealing that in the case of geminates shorter closures predicted more voicing. Figure 3 shows the scatterplot regression lines for geminates (left) and singletons (right).

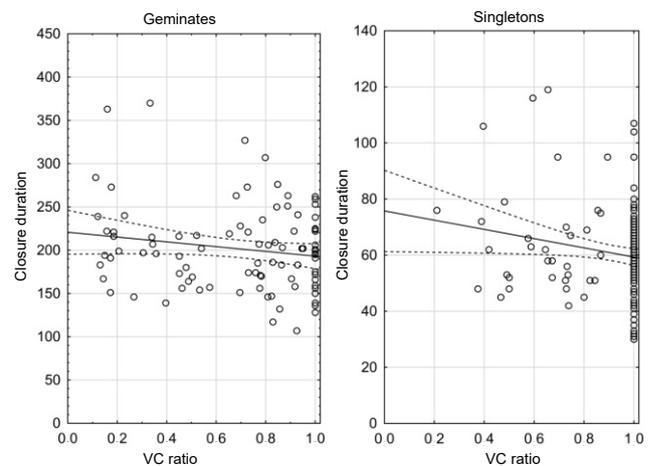
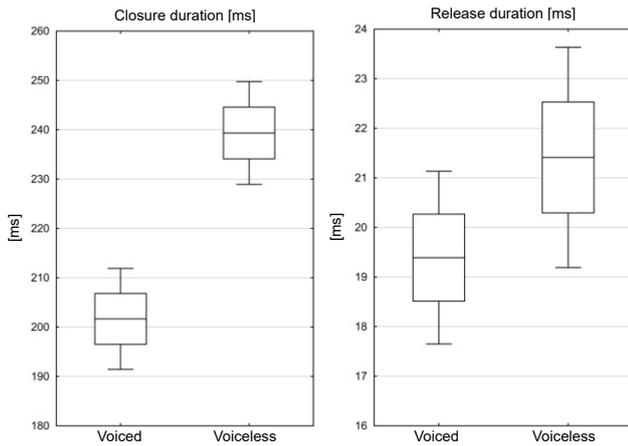


Fig. 3. Scatterplot of interaction between closure duration and voicing-to-closure ratio (VC ratio) in voiced geminates (left) and singletons (right).

Phonologically voiced geminates had a significantly shorter closure duration ( $M = 202, SE = 5.15$ ) than their voiceless counterparts ( $M = 239, SE = 5.25$ ) [ $F(1, 37.7) = 30.63, p < 0.001, \eta_p^2 = 0.45$ ]. This difference was not accompanied by statistically significant differences in duration of the release burst between voiced ( $M = 19.4, SE = 0.88$ ) and voiceless ( $M = 21.4, SE = 1.12$ ) geminates [ $F(1, 49.04) = 1.1, p = 0.3$ ]. Figure 4 shows mean closure and release duration for voiced and voiceless geminates and Table 3 presents numerical values across the three places of articulation.

Table 3. Means in millisecond and standard error for the closure duration and release duration in voiced and voiceless gemimates across the three places of articulation.

	Voiced gemimates				Voiceless gemimates			
	Closure duration		Release duration		Closure duration		Release duration	
	<i>M</i>	SE	<i>M</i>	SE	<i>M</i>	SE	<i>M</i>	SE
Bilabial /b/	220.3	10.4	12.3	0.68	254.1	7.83	14.2	1.08
Dental /d/	205.6	8.78	16.9	0.85	235.6	10.92	18.6	1.22
Velar /g/	183.1	6.65	27.5	1.21	224.3	8.15	33.3	1.65

Fig. 4. Closure duration (left) and release duration (right) in ms in voiced and voiceless gemimates. Lines = *M*, boxes = SE, whiskers = 0.95 CI.

### 6.2. Interaction of voicing and the geminate/singleton contrast

The recorded singletons had a mean closure duration of 61 ms (SE = 1.4) for voiced and 75 ms (SE = 1.38) for voiceless productions. The release burst had a mean duration of 16.1 ms (SE = 0.79) for voiced and 25.9 (SE = 1.05) for voiceless singletons, respectively. Statistical analyses of the observed differences showed that voiced singletons had a significantly shorter closure [ $F(1, 46.57) = 53.16, p < 0.001, \eta_p^2 = 0.53$ ] and release burst [ $F(1, 47.2) = 84.71, p < 0.001, \eta_p^2 = 0.64$ ]. Table 4 presents numerical values for closure and release duration in singletons across the three places of articulation.

The computations of the interaction between the closure duration, the voiced/voiceless contrast, and

the singleton/geminate opposition yielded a significant result [ $F(1, 34.05) = 11.04, p = 0.002, \eta_p^2 = 0.25$ ]. The analysis of the interaction (Fig. 5 top) shows that, even though post-hoc comparison revealed that voiced productions were significantly shorter for both singletons and gemimates [Bonferroni, both  $p < 0.001$ ], the effect of shortening was greater in magnitude for gemimates. This effect clearly results from the fact that gemimates have considerably longer closure durations and thus have more room for durational variability. The analysis of the release burst duration using the same interaction of the voiced/voiceless contrast and the singleton/geminate opposition (Fig. 5 middle) yielded another significant result [ $F(1, 42.86) = 36.18, p < 0.001, \eta_p^2 = 0.46$ ]. In this case, however, post-hoc Bonferroni tests revealed that duration of the release burst was significantly longer for voiceless productions in singletons [ $p < 0.001$ ] but not in gemimates [ $p = 0.14$ ]. It suggests that the durational difference of the release burst found between voiced and voiceless singletons (longer for voiceless and shorter for voiced) is largely cancelled in gemimates (voiceless gemimates do not have a longer release burst than voiced gemimates). The same model of interaction was calculated for mean intensity of the release burst in dB (Fig. 5 bottom). The model yielded a significant result [ $F(1, 36.91) = 17.77, p < 0.001, \eta_p^2 = 0.32$ ]. Post-hoc tests revealed that in the case of singletons, release intensity was significantly lower for voiced ( $M = 37.6, SE = 0.49$ ) than for voiceless stops ( $M = 40.1, SE = 0.49$ ) [ $p < 0.001$ ]. The difference was largely cancelled in the case of gemimates manifesting itself in a smaller and non-significant intensity difference between voiced ( $M = 39.4, SE = 0.62$ ) and voiceless ( $M = 41.5, SE = 0.58$ ) productions [ $p = 0.51$ ].

Table 4. Means in millisecond and standard error for the closure duration and release duration in voiced and voiceless singletons across the three places of articulation.

	Voiced singletons				Voiceless singletons			
	Closure duration		Release duration		Closure duration		Release duration	
	<i>M</i>	SE	<i>M</i>	SE	<i>M</i>	SE	<i>M</i>	SE
Bilabial /b/	71.2	2.55	12.2	1.0	85.9	1.92	17.6	1.06
Dental /d/	53.5	2.12	12.9	0.61	69.4	2.0	21.8	1.02
Velar /g/	58.3	1.81	22.9	1.65	69.7	2.45	38.9	1.62

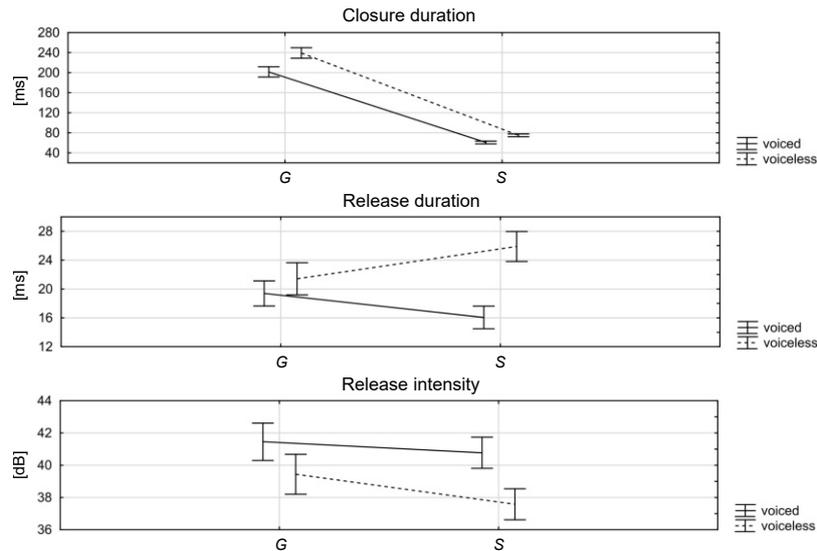


Fig. 5. Interactions between the singleton (S)/geminate (G) opposition and the voicing contrast (voiced – solid line, voiceless – dotted line) for closure duration in ms (top), release duration in ms (middle), and release intensity in dB (bottom). Whiskers = 0.95 CI.

6.3. Voicing in stop geminates: Case analysis

In this section, we discuss individual realizations of voicing that deserve attention in acoustic analysis. We are of opinion that automatic analyses of speech parameters by using scripts frequently lead to overlooking interesting instances of articulatory behavior. Our inspections of each individual production allowed us to isolate four different manifestations of voicing/devoicing in stop geminates other than what is typically reported in the literature. We will briefly discuss them in the following as they may be of interest for future studies investigating voicing in geminates or more generally in consonants.

Figure 6 shows waveform and spectrogram of an alveolar geminate /dd/ in the word *Upeddak* produced

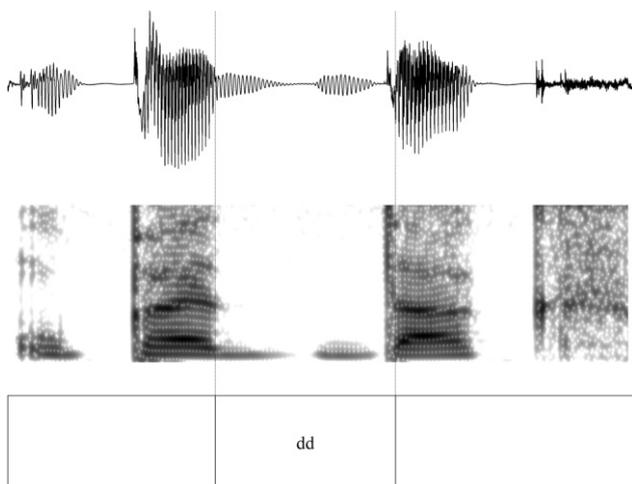


Fig. 6. Geminate /dd/ in the word *Upeddak* by Speaker 34 (female, age 20). Two visible portions of voicing separated by a period of voicelessness.

by Speaker 34 (female, age 20). Voicing in closure has two separate portions of almost equal durations, the first one has a duration of 84 ms and the second one of 77 ms. They are separated by an interval of voicelessness that lasts for 28 ms. A possible interpretation of this, apart from the physiological difficulty with maintaining longer periods of voicing in closure, is that it demonstrates underlying rearticulation of this geminate. Even though rearticulation does not emerge here in its standard manifestation of the release burst of the first consonant, it is signalled by the two separate voicing portions.

Figure 7 shows a bilabial geminate /bb/ in the word *Upebbak* produced by Speaker 7 (female, age 21).

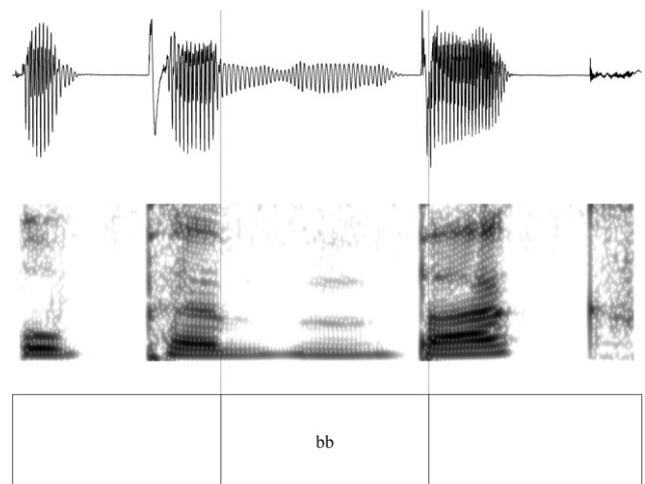


Fig. 7. Geminate /bb/ in the word *Upebbak* by Speaker 7 (female, age 21). Two visible portions of voicing differing in intensity followed by full devoicing prior to the release burst.

Voicing has two portions of intensity. The second portion with the duration of 139 ms is characterized by a rise in intensity compared to the first portion that lasts for 83 ms. There is an interval of full devoicing to the end of the geminate, prior to the release burst, that lasts for 25 ms. Such articulation with two separate portions of periodicity differing in magnitude may also suggest underlying rearticulation.

Figure 8 shows a bilabial geminate /bb/ in the word *Upebbak* produced by Speaker 41 (female, age 20). Closure duration has an observable division into two portions. The first portion with the duration of 114 ms is characterized by full voicing of very strong intensity. In the middle of the closure there is an abrupt cessation of voicing into the second portion that lasts for 100 ms and is completely devoiced. Such a realization may also be indicative of underlying rearticulation, suggesting a nonuniform status of a geminate that may be represented as /bp/. Naturally, as pointed out by a reviewer, such realizations are not uncommon in Japanese voiced geminates and may equally likely result from physiological reasons that we discuss in Introduction.

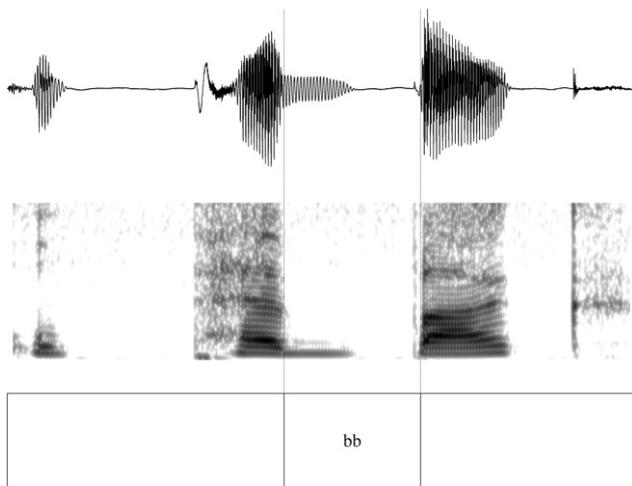


Fig. 8. Geminate /bb/ in the word *Upebbak* by Speaker 41 (female, age 20). The period of high-intensity voicing followed by abrupt complete devoicing.

Figure 9 shows a bilabial geminate /bb/ in the word *Upebbak* produced by Speaker 23 (female, age 20). In this case, devoicing is atypical, because normally devoiced stops have a period of voicing followed by voicelessness. Here, there is a period of long voicelessness in the onset portions of the consonant that lasts 254 ms followed by a sudden emergence of a short interval of voicing prior to the release with the duration of 58 ms. Interestingly, the auditory impression of this geminate is that it is almost fully voiced. One of the reasons for the “voiced” auditory impression may be a very abrupt energy rise at the onset of the vowel.

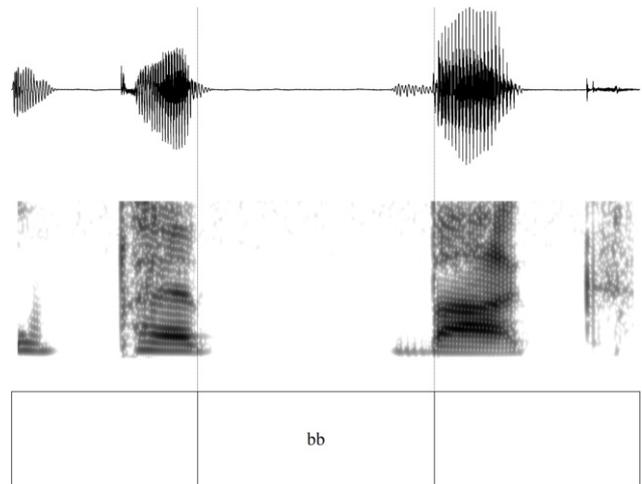


Fig. 9. Geminate /bb/ in the word *Upebbak* by Speaker 23 (female, age 20). Long portion of voicelessness followed by a sudden emergence of short voicing.

## 7. Discussion

The analysis of the collected data allows us to address aforementioned research issues.

*Are Polish phonologically voiced geminates devoiced and what is the magnitude of devoicing?*

Polish stop geminates have a mean voicing ratio of 0.69, which permits a conclusion that globally Polish stop geminates are partially devoiced. This locates Polish in the group of languages that devoice geminates fully or partially such as Japanese (FUJIMOTO, KATAOKA, 2016; HIROSE, ASHBY, 2007; HUSSAIN, SHINOHARA, 2019; KAWAHARA, 2006) or Tashlhiyt Berber (RIDOUANE, 2010), and in contrast to languages that retain voicing throughout the geminate closure such as Buginese (COHN *et al.*, 1999) or Egyptian Arabic (KAWAHARA, 2006). Compared to Japanese, Polish geminates with the ratio of 0.69 may be referred to as mildly devoiced, because the voicing ratios in Japanese range, depending on the study, from 0.3 to 0.4 (KAWAHARA, 2006), 0.53 (HIROSE, ASHBY, 2007), or 0.2 to 0.25 (HUSSAIN, SHINOHARA 2019). The results of our study and previous reports on geminate devoicing may explain why voiced geminates are considered to be typologically marked and appear less frequently in the world’s languages than voiceless geminates (OHALA, 1983; WESTBURY, KEATING, 1986; BLEVINS, 2004; HAYES, STERIADE, 2004). Since global measures of devoicing, defined as the proportion of voicing duration to total closure duration, obscure the analysis of actual individual realizations, we provided quantification of the voicing ratio into three voicing groups: devoiced (0.00 to 0.49), partially devoiced (0.50 to 0.89), voiced (0.90 to 1.00). The distribution in these categories was relatively balanced (33 % of voiced, 38 % of partially devoiced, 29 % of devoiced productions). Polish stop

singletons have a mean voicing ratio of 0.91, which is significantly higher than the one in geminates. This allows us to answer the question formulated in the title of the study concerning the amount of voicing in a geminate. The answer is that Polish phonologically voiced stop geminates are significantly less voiced than their corresponding singletons.

*Is there observable speaker variability in the magnitude of devoicing?*

There was an observable speaker-dependent variability in the realization of voicing. Only voiced geminates were produced by 14 % of the speakers, 20 % of the speakers fully devoiced all geminates, and 66 % of the speakers partially devoiced geminates in their speech. Such variability indicates that speakers differ in their voicing realizations and devoicing is not a uniform and omnipresent speech strategy for all speakers of Polish.

*Is the duration of voicing correlated with the duration of the closure?*

Since longer closure durations are not conducive to maintaining glottal activity necessary for voicing (CATFORD, 1977; JAEGER, 1978), it was hypothesized that shorter closures would predict more voicing, as formulated by the Aerodynamic Voicing Constraint (AVC) (OHALA, 1997), according to which longer closure favors devoicing due to balancing of air pressure in the sub- and supra-glottal cavities. The current results showed a trend in this direction ( $b^* = -0.17$ ), however not at the level of statistical significance. The reason for the observed limited regression power may be the fact that the geminate/singleton contrast strongly relies on durational differences (ESPOSITO, DI BENEDETTO, 1999; KINGSTON *et al.*, 2009; AMANO, HIRATA, 2010; AL-TAMIMI, KHATTAB, 2018; ROJCZYK, PORZUCZEK, 2019a). In more detail, even though shorter closure durations may promote more voicing in closure, geminates are restricted in the shortening of the closure portion at the pressure of remaining distinct from singletons. This hypothesis is supported by the fact that, as shown in this and other studies (HOMMA, 1981; IDEMARU, GUION 2008; AL-TAMIMI, KHATTAB, 2018), voiced geminates tend to be shorter than their voiceless counterparts. This shortening tendency may be motivated by the need to sustain voicing throughout closure to the maximum, however not at the expense of merging durationally with singletons. Another interpretation may be supported by studies showing that speakers do not necessarily have to rely on reducing closure duration to sustain voicing if they resort to other strategies to keep the voicing conditions effective such as lowering the larynx, advancing the tongue root, or raising the tongue and the soft palate with the aim of increasing the size of the pharyngeal cavity (HALLE, STEVENS, 1971;

WESTBURY, 1983; SOLÉ, 2009; AL-TAMIMI, KHATTAB, 2018).

*Is there interaction of voicing and the geminate/singleton contrast for closure duration, release duration and release intensity?*

In terms of closure duration, both voiced geminates and singletons were significantly shorter than their voiceless counterparts, confirming previously cited reports that voicing is reflected in shorter closure durations. The interaction was contributed to by larger shortening effect for geminates (37 ms) than for singletons (14 ms). This difference is predictable on the grounds that longer durations of geminates provide them with more room for durational variability. This result stands in contrast to the results in (BUTCHER, 2004), who found shortening for voiced singletons only, but not for geminates, suggesting that the closure lengthening as a cue to gemination has a primacy over durational differences cuing the voicing contrast.

Duration of the release burst interacted significantly with the voiced/voiceless contrast and the singleton/geminate opposition in that the release duration differentiated voiced and voiceless categories in the case of singletons, but not in the case of geminates. Voiceless singletons had a significantly longer release burst than their voiced counterparts by 9.8 ms. This effect was not observed for geminates, suggesting that the voicing cue of release duration is neutralized in geminates, at least in the case of stop geminates in Polish. A similar interaction was found for intensity of the release burst. In the case of singletons, voiced productions had mean intensity of the release burst that was significantly lower (by 2.6 dB) than in corresponding voiceless productions. This difference was cancelled in geminates in which intensity of the release did not significantly differentiate between voiced and voiceless tokens. In general, these results show that the voiced/voiceless cues in the release burst, such as duration and intensity which are operative in singletons, are neutralized in geminates. It may suggest that the whole burden of cuing the voicing contrast in geminates relies on the closure duration and voicing in closure.

*Are there atypical realizations of voicing in Polish stop geminates?*

In the current results, we have isolated instances of atypical voicing realizations in Polish stop geminates. Although devoicing is commonly defined as gradual cessation of low-energy periodic signal towards the end of closure, some of the discussed realizations show divergent patterns. In the collected sample of Polish geminates, voicing may have two peaks separated by either an interval of weak voicing or complete voicelessness. We interpret these realizations as reflections of underlying rearticulation. Voicing may also emerge in the final portions of the closure preceded by voicelessness.

## 8. Conclusions

The long duration of geminate constriction facilitates a variety of phonetic effects related to the coordination of glottal and supraglottal articulatory gestures. Generally, considering the wide range of devoiced closure portion durations, and large inter-speaker variation, we may suggest little significance of true voicing for phonological voicing identification in Polish geminates or little communicative importance of such identification. The articulatory effort to maintain vocal fold activity throughout the constriction appears superfluous in contexts where no lexical ambiguity may be expected, and where other consonant clusters tend to assume a uniform phonation type. Similarly, the burst intensity level, whether or not actually used as a cue to voicing in singletons, is not modified systematically by speakers to cue the contrast in geminates, where it would also require more articulatory effort. Singletons, on the other hand, are characterized by more salient voicing in closure and burst length distinctions for the same two reasons that were signaled above. First, articulatory control of these features is easier in shorter constrictions, and second, possible voicing neutralization in singletons is more likely to trigger lexical identification problems.

Finally, the separate portions of voicing and abrupt voicing intensity changes within an individual closure presented in case analysis section, as well as the existence of rearticulation may give evidence to the hypothesis that Polish geminates may be regarded as biphonemic combinations. This idea deserves further phonological investigation, because it still needs to accommodate the majority of productions which are single-articulated and thus phonetically they fulfil the requirement of a long consonant similar to, e.g., Italian or Japanese. It is still not clear how geminates should be treated in various languages (the original term itself suggests two elements) and we believe that various empirical studies, including the present one, may help us better understand their nature.

## Acknowledgments

Research supported by the National Science Centre grant Acoustic properties of Polish geminate consonants (grant no. UMO-2017/25/B/HS2/02548) to the first author and co-financed by the funds granted under the Research Excellence Initiative of the University of Silesia in Katowice.

## References

1. ABRAMSON A.S., WHALEN D.H. (2017), Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions, *Journal of Phonetics*, **63**: 75–86, doi: [10.1016/j.wocn.2017.05.002](https://doi.org/10.1016/j.wocn.2017.05.002).
2. AL-TAMIMI J., KHATTAB G. (2018), Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops, *Journal of Phonetics*, **71**: 206–325, doi: [10.1016/j.wocn.2018.09.010](https://doi.org/10.1016/j.wocn.2018.09.010).
3. AMANO S., HIRATA Y. (2010), Perception and production boundaries between single and geminate stops in Japanese, *Journal of the Acoustical Society of America*, **128**(4): 2049–2058, doi: [10.1121/1.3458847](https://doi.org/10.1121/1.3458847).
4. BLEVINS J. (2004), *Evolutionary Phonology: The Emergence of Sound Patterns*, University Press, Cambridge.
5. BOERSMA P., WEENINK D. (2001), Praat: Doing phonetics by computer, *Ear and Hearing*, **32**(2): 266, doi: [10.1097/AUD.0b013e31821473f7](https://doi.org/10.1097/AUD.0b013e31821473f7).
6. BOERSMA P., WEENINK D. (n.d.), Praat programme, Version 6.1.40, <http://www.praat.org/> (access: 15.05.2023).
7. BUTCHER A. (2004), Fortis/lenis revisited one more time: The aerodynamics of some oral stop contrasts in three continents, *Clinical Linguistics and Phonetics*, **18**(6–8): 547–557, doi: [10.1080/02699200410001703565](https://doi.org/10.1080/02699200410001703565).
8. CATFORD J.C. (1977), *Fundamental Problems in Phonetics*, Edinburgh University Press, Edinburgh.
9. CHO T., WHALEN D.H., DOCHERTY G. (2019), Voice onset time and beyond: Exploring laryngeal contrast in 19 languages, *Journal of Phonetics*, **72**: 52–65, doi: [10.1016/j.wocn.2018.11.002](https://doi.org/10.1016/j.wocn.2018.11.002).
10. COHN A.C., HAM W.H., PODESVA R.J. (1999), The phonetic realization of singleton-geminate contrasts in three languages of Indonesia, [in:] *Proceedings of the 14th International Congress of Phonetic Sciences*, pp. 587–590.
11. CYRAN E. (2011), Laryngeal realism and laryngeal relativism: Two voicing systems in Polish?, *Studies in Polish Linguistics*, **6**(1): 45–80.
12. DAVIDSON L. (2016), Variability in the implementation of voicing in American English obstruents, *Journal of Phonetics*, **54**: 35–50, doi: [10.1016/j.wocn.2015.09.003](https://doi.org/10.1016/j.wocn.2015.09.003).
13. DMITRIEVA O. (2009), Geminate typology and perception of consonant length, Paper presented at the 82nd Annual Meeting of the Linguistic Society of America.
14. DMITRIEVA O. (2017), Production of geminate consonants in Russian: Implications for typology, [in:] *The Phonetics and Phonology of Geminate Consonants*, Kubozono H. [Ed.], pp. 34–65, Oxford University Press, Oxford, doi: [10.1093/oso/9780198754930.003.0003](https://doi.org/10.1093/oso/9780198754930.003.0003).
15. ESPOSITO A., DI BENEDETTO M.G. (1999), Acoustical and perceptual study of gemination in Italian stops, *The Journal of the Acoustical Society of America*, **106**(4): 2051–2062, doi: [10.1121/1.428056](https://doi.org/10.1121/1.428056).
16. FUJIMOTO M., KATAOKA R. (2016), Oral/nasal airflow during Japanese stop consonants, *The Journal of the Acoustical Society of America*, **140**(4): 3108, doi: [10.1121/1.4969709](https://doi.org/10.1121/1.4969709).
17. HALLE M., STEVENS K.N. (1971), A note on laryngeal features, *MIT Quarterly Progress Report*, **101**: 198–212.

18. HAMZAH M.H., FLETCHER J., HAJEK J. (2016), Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay, *Journal of Phonetics*, **58**: 135–151, doi: [10.1016/j.wocn.2016.08.002](https://doi.org/10.1016/j.wocn.2016.08.002).
19. HANKAMER J., LAHIRI A., KOREMAN J. (1989), Perception of consonant length: Voiceless stops in Turkish and Bengali, *Journal of Phonetics*, **17**(4): 283–298, doi: [10.1016/S0095-4470\(19\)30445-0](https://doi.org/10.1016/S0095-4470(19)30445-0).
20. HAYES B., STERIADE D. (2004), The phonetic bases of phonological markedness, [in:] *Phonetically Based Phonology*, Hayes B., Kirchner R., Steriade D. [Eds.], pp. 1–33, Cambridge University Press, Cambridge.
21. HERMES A., TILSEN S., RIDOUANE R. (2021), Cross-linguistic timing contrast in geminates: A rate-dependent perspective, [in:] *Proceedings of the 12th International Seminar on Speech Production*, pp. 52–55.
22. HIRATA Y., WHITON J. (2005), Relational acoustic invariance in the single/geminate stop distinction in Japanese, *The Journal of the Acoustical Society of America*, **117**(4): 2569, doi: [10.1121/1.4788556](https://doi.org/10.1121/1.4788556).
23. HIROSE A., ASHBY M. (2007), An acoustic study of devoicing of the geminates obstruents, [in:] *Proceedings of the 14th International Congress of Phonetic Sciences*, pp. 909–912.
24. HOMMA Y. (1981), Durational relationship between Japanese stops and vowels, *Journal of Phonetics*, **9**(3): 273–281, doi: [10.1016/S0095-4470\(19\)30971-4](https://doi.org/10.1016/S0095-4470(19)30971-4).
25. HUSSAIN Q., SHINOHARA S. (2019), Partial devoicing of voiced geminate stops in Tokyo Japanese, *The Journal of the Acoustical Society of America*, **145**(1): 149–163, doi: [10.1121/1.5078605](https://doi.org/10.1121/1.5078605).
26. IDEMARU K., GUION S. (2008), Acoustic covariants of length contrast in Japanese stops, *Journal of the International Phonetic Association*, **38**(2): 167–186, doi: [10.1017/S0025100308003459](https://doi.org/10.1017/S0025100308003459).
27. JAEGER J.J. (1978), Speech aerodynamics and phonological universals, [in:] *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, **4**: 312–329.
28. JURAFSKY D., BELL A., GREGORY M., RAYMOND W.D. (2001), Probabilistic relations between words: Evidence from reduction in lexical production, [in:] *Frequency and the Emergence of Linguistic Structure*, Bybee J., Hopper P. [Eds.], pp. 229–254, John Benjamins, Amsterdam, doi: [10.1075/tsl.45.13jur](https://doi.org/10.1075/tsl.45.13jur).
29. KAWAHARA S. (2006), A faithfulness ranking projected from the perceptibility scale: The case of [+voice] in Japanese, *Language*, **82**(3): 536–574.
30. KAWAHARA S. (2015), The phonetics of sokuon, or obstruent geminates, [in:] *The Mouton Handbook of Japanese Language and Linguistics*, Kubozono H. [Ed.], pp. 43–77, Mouton de Gruyter, Berlin.
31. KINGSTON J., KAWAHARA S., CHAMBLESS D., MASH D., BRENNER-ALSOP E. (2009), Contextual effects on the perception of duration, *Journal of Phonetics*, **37**(3): 297–320, doi: [10.1016/j.wocn.2009.03.007](https://doi.org/10.1016/j.wocn.2009.03.007).
32. KOTZOR S., WETTERLIN A., ROBERTS A.C., LAHIRI A. (2016), Processing of phonemic consonant length: Semantic and fragment priming evidence from Bengali, *Language and Speech*, **59**(1): 83–112, doi: [10.1177/0023830915580189](https://doi.org/10.1177/0023830915580189).
33. KOZYRA A. (2008), Geminates in Slavic languages [in Polish: Geminaty w językach słowiańskich], *LingVaria*, **1**(5): 257–266.
34. KUBOZONO H. (2017), *The Phonetics and Phonology of Geminate Consonants*, Oxford University Press, Oxford.
35. LADEFOGED P., MADDIESON I. (1996), *The Sounds of the World's Languages*, Blackwell, Oxford UK, Cambridge MA.
36. LAHIRI A., HANKAMER G. (1988), The timing of geminate consonants, *Journal of Phonetics*, **16**(3): 327–338, doi: [10.1016/S0095-4470\(19\)30506-6](https://doi.org/10.1016/S0095-4470(19)30506-6).
37. LISKER L. (1957), Closure duration and the intervocalic voiced-voiceless distinction in English, *Language*, **33**(1): 42–49.
38. LOCAL J., SIMPSON A.P. (1999), Phonetic implementation of geminates in Malayalam nouns, [in:] *Proceedings of the 14th International Congress of Phonetic Sciences*, Ohala J.J., Hasegawa Y., Ohala M., Granville D., Bailey A.C. [Eds.], pp. 595–598, University of California Press, Berkeley.
39. LUKE S.G. (2017), Evaluating significance in linear mixed-effects models in R, *Behavior Research Methods*, **49**: 1494–1502, doi: [10.3758/s13428-016-0809-y](https://doi.org/10.3758/s13428-016-0809-y).
40. MADDIESON I. (1985), Phonetic cues to syllabification, [in:] *Phonetics Linguistics: Essays in Honor of Peter Ladefoged*, Fromkin V.A. [Ed.], pp. 203–221, Academic Press, New York.
41. MALISZ Z. (2013), *Speech rhythm variability in Polish and English: A study of interaction between rhythmic levels*, Ph.D. Thesis, Adam Mickiewicz University, Poznań.
42. MALISZ Z., ŻYGIS M. (2015), Voicing in Polish: Interactions with lexical stress and focus, [in:] *The 18th International Congress of Phonetic Sciences*.
43. MULLER J.S. (2001), *The phonology and phonetics of word-initial geminates*, Ph.D. Thesis, The Ohio State University, Columbus.
44. MUNSON B., SOLOMON N.P. (2004), The effect of phonological neighborhood density on vowel articulation, *Journal of Speech, Language, and Hearing Research*, **47**(5): 1048–1058, doi: [10.1044/1092-4388\(2004\)078](https://doi.org/10.1044/1092-4388(2004)078).
45. OH G.E., REDFORD M.A. (2012), The production and phonetic representation of fake geminates in English, *Journal of Phonetics*, **40**(1): 82–91, doi: [10.1016/j.wocn.2011.08.003](https://doi.org/10.1016/j.wocn.2011.08.003).
46. OHALA J.J. (1983), The origin of sound patterns in vocal tract constraints, [in:] *The Production of Speech*, MacNeilage P.F. [Ed.], pp. 189–216, Springer, New York.

47. OHALA J.J. (1997), Aerodynamics in phonology, [in:] *Proceedings of the Seoul International Conference on Linguistics*, pp. 92–97.
48. OHALA M. (2007), Experimental methods in the study of Hindi geminate consonants, [in:] *Experimental Approaches to Phonology*, Solé M.J., Beddor P., Ohala M. [Eds.], pp. 351–368, Oxford University Press, Oxford.
49. PAJAŁ B. (2009), Contextual constraints on geminates: The case of Polish, [in:] *Proceedings of the 35th Annual Meeting of Berkeley Linguistic Society*.
50. PICKETT E.R., BLUMSTEIN S.E., BURTON M.W. (1999), Effects of speaking rate on the singleton/geminate consonant contrast in Italian, *Phonetica*, **56**: 135–157, doi: [10.1159/000028448](https://doi.org/10.1159/000028448).
51. PIND J. (1995), Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues, *Perception and Psychophysics*, **57**(3): 291–304, doi: [10.3758/BF03213055](https://doi.org/10.3758/BF03213055).
52. PORT R., DALBY J., O'DELL M. (1987), Evidence for mora timing in Japanese, *Journal of the Acoustical Society of America*, **81**: 1574–1585.
53. PORZUCZEK A., ROJCZYK A. (2014), Gemination strategies in L1 and English pronunciation of Polish learners, *Research in Language*, **12**(3): 291–300, doi: [10.2478/rela-2014-0020](https://doi.org/10.2478/rela-2014-0020).
54. PORZUCZEK A., ROJCZYK A. (2021), Complex patterns in L1-to-L2 phonetic transfer: The acquisition of English plosive and affricate fake geminates and non-homorganic clusters by Polish learners, *Research in Language*, **19**(1): 1–13, doi: [10.18778/1731-7533.19.1.01](https://doi.org/10.18778/1731-7533.19.1.01).
55. RAYMOND W.D., DAUTRICOURT R., HUME E. (2006), Word internal /t, d/ deletion in spontaneous speech: Modeling the effects of extralinguistic, lexical, and phonological factors, *Language Variation and Change*, **18**(1): 55–97, doi: [10.1017/S0954394506060042](https://doi.org/10.1017/S0954394506060042).
56. RIDOUANE R. (2010), Geminate at the junction of phonetics and phonology, [in:] *Papers in Laboratory Phonology X*, Fougeron C., Kühnert B., D'Imperio M., Vallée N. [Eds.], pp. 61–90, Mouton de Gruyter, Berlin.
57. RIDOUANE R. (2022), Phonetics and phonology of Tashlhiyt geminates: An overview, *HAL Open Science*, Version halshs-03511107v2, <https://shs.hal.science/halshs-03511107v2>.
58. ROJCZYK A., PORZUCZEK A. (2014), Acoustic properties of nasal geminates in Polish, [in:] *Crossing Phonetics-Phonology Lines*, Cyran E., Szpyra-Kozłowska J. [Eds.], pp. 347–364, Cambridge Scholars Publishing, Cambridge upon Tyne.
59. ROJCZYK A., PORZUCZEK A. (2019a), Durational properties of Polish geminate consonants, *The Journal of the Acoustical Society of America*, **146**(6): 4171–4182, doi: [10.1121/1.5134782](https://doi.org/10.1121/1.5134782).
60. ROJCZYK A., PORZUCZEK A. (2019b), Rearticulated geminates are not sequences of two identical sounds: Evidence from Polish affricate geminates, [in:] *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, Calhoun S., Escudero P., Tabain M., Warren P. [Eds.], pp. 3671–3675, Australasian Speech Science and Technology Association Inc., Canberra.
61. ROJCZYK A., PORZUCZEK A. (2022), POLGEM – the recorded corpus of Polish geminate consonants, *Poznań Studies in Contemporary Linguistics*, **58**(2): 253–262, doi: [10.1515/psicl-2022-0014](https://doi.org/10.1515/psicl-2022-0014).
62. RUBACH J., BOOIJ G.E. (1990), Edge of constituent effects in Polish, *Natural Language and Linguistic Theory*, **8**(3): 427–463, doi: [10.1007/BF00135620](https://doi.org/10.1007/BF00135620).
63. SEARLE S.R., CASELLA G., MCCULLOCH C.E. (1992), *Variance Components*, Wiley, New York.
64. SOLÉ M.-J. (2007), Controlled and mechanical properties in speech: A review of the literature, [in:] *Experimental Approaches to Phonology*, Solé M.-J., Beddor P.S., Ohala M. [Eds.], pp. 302–322, Oxford University Press, Oxford, doi: [10.1093/oso/9780199296675.003.0018](https://doi.org/10.1093/oso/9780199296675.003.0018).
65. SOLÉ M.-J. (2009), Acoustic and aerodynamic factors in the interaction of features, [in:] *Phonetics and Phonology: Interactions and Interrelations*, Vigário M., Freitas M.J. [Eds.], pp. 205–234, John Benjamins, Amsterdam, doi: [10.1075/cilt.306.10sol](https://doi.org/10.1075/cilt.306.10sol).
66. TAGLIAPIETRA L., MCQUEEN J.M. (2010), What and where in speech recognition: Geminates and singletons in spoken Italian, *Journal of Memory and Language*, **63**(3): 306–323, doi: [10.1016/j.jml.2010.05.001](https://doi.org/10.1016/j.jml.2010.05.001).
67. THURGOOD E., DEMENKO G. (2003), Phonetic realizations of Polish geminate affricates, [in:] *Proceedings of the 15th International Congress of Phonetic Sciences*, Solé M.-J., Recasens D., Romero J. [Eds.], pp. 1895–1898.
68. THURGOOD G. (1993), Geminates: A cross-linguistic examination, [in:] *Papers in Honor of Frederick H. Bregelman on the Occasion of the 25th Anniversary of the Department of Linguistics, CSU*, Nevis J.A., McMenamin G., Thurgood G. [Eds.], pp. 129–139, Department of Linguistics, California State University, Fresno CA.
69. TIBCO Software Inc. (2017), Statistica (data analysis software system), Version 13, <http://statistica.io> (access: 15.05.2023).
70. WESTBURY J.R. (1983), Enlargement of the supraglottal cavity and its relation to stop consonant voicing, *The Journal of the Acoustical Society of America*, **74**(4): 1322–1336, doi: [10.1121/1.389236](https://doi.org/10.1121/1.389236).
71. WESTBURY J.R., KEATING P.A. (1986), On the naturalness of stop consonant voicing, *Journal of Linguistics*, **22**(1): 145–166, doi: [10.1017/S0022226700010598](https://doi.org/10.1017/S0022226700010598).
72. YOSHIDA K., DE JONG K.J., KRUSCHKE J.K., PÄIVIÖ P.-M. (2015), Cross-language similarity and difference in quantity categorization of Finnish and Japanese, *Journal of Phonetics*, **50**: 81–98, doi: [10.1016/j.wocn.2014.12.006](https://doi.org/10.1016/j.wocn.2014.12.006).



## Research Paper

## Study on the Impact of Drainage Noise in Residential Bathrooms Based on Finite Element Simulation

Yaping WANG, Xin DENG, Bingyuan CHI, Yanqiu CUI\*

*School of Architecture and Urban Planning, Shandong Jianzhu University  
China*\*Corresponding Author e-mail: [cyq@sdjzu.edu.cn](mailto:cyq@sdjzu.edu.cn)*(received June 8, 2023; accepted March 28, 2024; published online June 18, 2024)*

Residential bathroom drainage noise is a primary source of indoor noise that directly affects quality of life and physical and mental health. Therefore, based on the acoustic theory and the finite element simulation technology, this paper proposes a method to simulate the drainage noise characteristics and its impact range jointly using the flow and acoustic fields. The pressure at the pipe wall caused by the internal flow field of the bathroom drainage pipe is calculated by the Fluent software. Simulations are carried out with the Virtual Lab software to predict the drainage noise characteristics and spatial distribution and to analyse the influence of factors such as the door position, riser position, and the partition wall material on the noise distribution. The results show that drainage noise has prominent high-frequency characteristics, the position of the bathroom drainage pipes and doors affects the spatial noise distribution, and the sound insulation performance of a partition wall with ordinary fired bricks in the bathroom is slightly better than that of ordinary concrete bricks, lightweight aggregate concrete blocks or fly ash blocks. This paper provides a theoretical basis and practical reference for reducing the impact of residential drainage noise and creating a healthy and comfortable indoor acoustic environment.

**Keywords:** residential building; bathroom drainage noise; numerical simulation; sound field distribution.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the rapid development of society, public demand for environmental quality has gradually increased. Noise is one of four major environmental pollutants in the world and has become a key concern. Residential spaces are most closely associated with people's daily lives, and indoor noise can significantly reduce quality of life, work efficiency and health; notably, residential bathroom drainage noise is a primary source of indoor noise (PARK *et al.*, 2018; PUJOL *et al.*, 2014). Previous studies have used field tests, laboratory tests, numerical simulations, subjective evaluations, and virtual reality to explore the objective parameters of residential bathroom drainage noise and residents' subjective evaluations and have found that drainage noise causes significant disturbance to residents (JEON *et al.*, 2019; RYU, SONG, 2019). Therefore, research on residential bathroom drainage noise characteristics and radiated sound fields is essential for cre-

ating a healthy and comfortable acoustic environment in living spaces.

Pipe drainage noise research worldwide has focused mainly on the causes of noise and the factors influencing it, as well as noise reduction control measures, and the prediction, numerical simulation of flow noise, and radiation sound fields.

One study reported that the causes of pipe drainage noise include water flow impact, air pressure fluctuations, and water flow direction mutation. The leading cause was determined to be unstable water and air pressure in a non-full flow pipe (FUCHS, 1983; 1993a). Additionally, factors such as the drainage method, pipe material and fittings, sanitary ware, water flow velocity, and pressure can significantly affect drainage noise (FUCHS, 1993b; VILLOT, 2000). Laboratory tests of UPVC pipe drainage noise revealed that the noise impact of the same-floor drainage pipes was lower than that of different-floor drainage pipes (YANG *et al.*, 2016). Increasing the wall thickness of new UPVC

pipes or increasing the material density of pipes could reduce the impact of drainage noise (JEONG *et al.*, 2017; YEON *et al.*, 2014), and the drainage noise of a vacuum double-bend pipe was approximately 3.5 dB(A) lower than that of an ordinary pipe (JUNG *et al.*, 2012). In addition, the use of new low-noise systems and connections (JEONG *et al.*, 2017; YEON *et al.*, 2014), the use of pipe cladding, the addition of pipe wells, and other measures (XU *et al.*, 2014) were also found to reduce noise.

In recent years, numerical simulation methods, mainly including the sound line method, the finite element method, the boundary element method, and the statistical energy method, have gradually become important research tools in architectural acoustics. Among them, the finite element and boundary element methods are simulation methods based on the wave acoustic theory (MARBURG, NOLTE, 2008) and are widely used in research on pipe noise and enclosed sound fields.

Previous studies have used finite element methods to analyse the coupled vibration noise of pipes and simulate the external radiated sound field of pipes (LIANG *et al.*, 2006; MORI *et al.*, 2017). The vibration noise of gas transmission pipe systems is simulated via ANSYS and SysNoise (LIU *et al.*, 2016; HAN *et al.*, 2020). The acoustic, vibration, and aeroacoustic noise characteristics of *T*-shaped and rectangular cross-section tubes were simulated by experiments with different flow velocities and finite element simulations (MORI *et al.*, 2017). The Virtual Lab software was used to investigate the flow noise propagation mechanism in pipes based on the flow and sound field synergy principle (CAO *et al.*, 2017). CFD was used to analyse the characteristics and main factors influencing flow-induced noise in the variable cross-section pipe. The flow-induced noise and flow field characteristics, including the sound source intensity distribution, pressure distribution, velocity distribution, and spectrum of sound pressure level, were obtained. This provided a theoretical basis for the optimization of variable cross-section piping systems and the investigation of flow-induced noise control techniques (SUN *et al.*, 2021). Additionally, in the context of enclosed space sound fields, the COMSOL Multiphysics software has also been used to evaluate the effect of the addition of sound absorption and the change in absorption position on the sound field in small rooms. This method allows more accurate calculation of low-frequency sound field data in rooms (LAU, POWELL, 2018).

In summary, previous studies have explored the generation principles, characteristics, and factors influencing residential bathroom drainage noise. The flow-induced noise of pipes has a significant impact on the indoor acoustic environment of adjacent floors, and the use of silent pipes, wrapping acoustic insulation materials, and other measures have been proposed to

reduce drainage noise. These works provide an essential theoretical and practical basis for future research. However, in previous studies of pipe drainage noise, most of the flow noise in the pipe was assumed to arise from a line sound source, supplemented by a constant flow rate and flow velocity to the actual noise measurement, which cannot sufficiently reflect the actual state of fluid flow in the pipe, and there is a certain discrepancy with the actual sound field.

In recent years, the use of finite element and boundary element methods with simulation software such as Ansys and COMSOL to simulate fluid flow within pipes and enclosed space sound fields has become a target of active research. However, there has been a relative lack of research on predicting bathroom drainage noise and its radiation distribution using numerical simulation methods. Therefore, this paper proposes a method for predicting pipe drainage noise and its effects using finite element simulation techniques with the help of the numerical simulation software Ansys Fluent (version 2020) and LMS Virtual Lab (v13.6), analyses the factors affecting the noise sound field, and proposes noise reduction design recommendations, providing new ideas and methods for indoor sound environment research.

## 2. Numerical simulation methods

Pipe drainage noise, which is generated by the interaction between the fluid inside the pipe and the pipe wall, includes the flow noise generated by fluid motion inside the pipe and the radiation noise generated by the vibration of the pipe wall. Therefore, based on a theory related to flow and vibration noise, a method of simulating the indoor drainage noise distribution using joint simulation of the flow field and sound fields is proposed, as shown in Fig. 1.

### 2.1. Theoretical analysis of pipe drainage noise

#### 2.1.1. Flow noise

The root cause of flow noise is that the fluid flow produces a distributed sound source throughout the flow field, which is considered aeroacoustics and is generated by the interaction of the fluid medium, the solid wall, and the sound field (LIGHTHILL, 1952; CURLE, 1955; FLOWCS WILLIAMS, HAWKINGS, 1969). The effect of a solid wall on flow-induced noise is considered, and the FW-H equation elucidates three types of flow noise sources in terms of the occurrence mechanism, namely, monopole sources generated by fluid volume pulsations, dipole sources generated by pulsations at solid boundaries, and quadrupole sources generated by free fluid turbulence (LIGHTHILL, 1952; CURLE, 1955; FLOWCS WILLIAMS, HAWKINGS, 1969; ZHANG *et al.*, 2016). The flow in the drainage pipe of

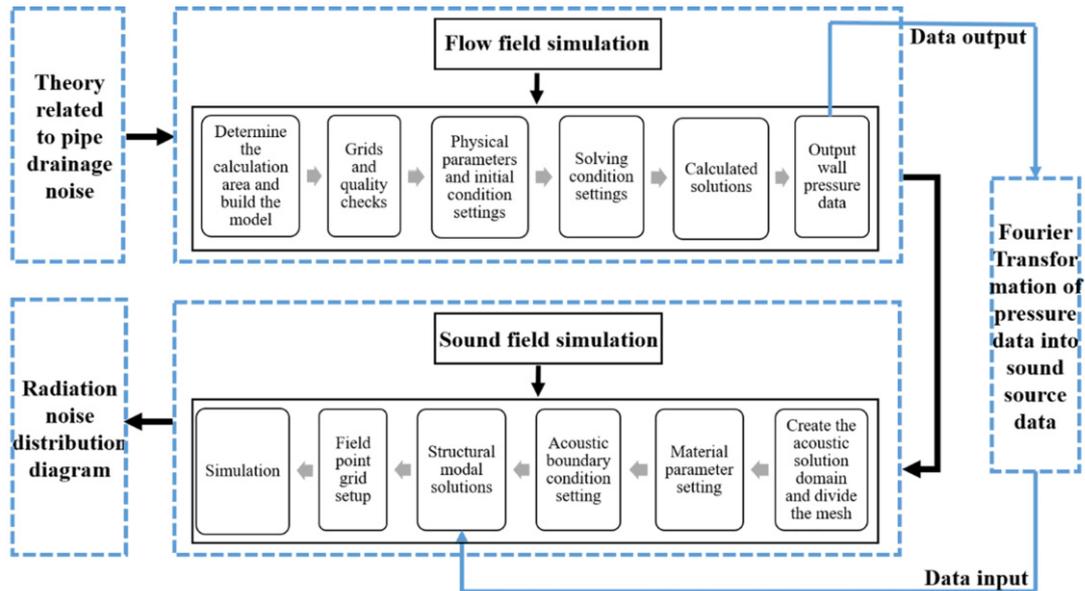


Fig. 1. Numerical simulation process for the radiated sound field of pipe drainage noise.

the residential bathroom is a type of gas-liquid two-phase flow with the low Mach number; the flow velocity is much lower than the sound velocity, the boundary condition is a static solid wall, the compressibility effect of the mixed fluid is negligible, and the fluid acceleration is weak. Therefore, the effect of monopole sources and quadrupole sources on the noise can be ignored, and the noise sources in the flow field are treated as distributed dipole sources caused by pressure pulsations generated by the flow field near the wall surface.

### 2.1.2. Vibration radiation noise

The mixed fluid in the drainage pipe impacts the wall to excite structural vibration, generating structural noise and propagating to the space inside and outside the pipe, producing indoor interference noise. From the perspective of the vibration generation mechanism, the vibration of thin-walled pipes derives from the excitation and transmission of fluid motion, which is influenced to some extent by characteristics such as the intensity of turbulent motion. The vibration is also influenced by the physical properties of the structure itself, with the frequency characteristics of some structural components (NORTON, KARCZUB, 2003a; NORTON, BULL, 1984).

The inherent properties of structural vibrations are described by structural modes, which are related only to the shape and dimensions of the structure and the external excitation frequency (NORTON, KARCZUB, 2003b; MAO, PIETRZKO, 2013). Analysis of the structural modes can reveal the vibration characteristics of a structure, predicting its vibration response under external excitation conditions and hence the distribution of radiated noise.

## 2.2. Simulation method and process

### 2.2.1. Flow field simulation method and process

This paper uses the CFD software Fluent, which is integrated with the Ansys Workbench environment, to simulate the internal flow field characteristics of a bathroom drainage pipe, to obtain the flow state of the mixed media inside the drainage pipe, and to output the pressure data on the inner wall of the pipe as the sound source parameters for acoustic analysis. The calculation models include the VOF multiphase flow model and the realizable  $k-\varepsilon$  turbulence model. The primary process of using the Fluent software to simulate and analyse the flow field information in the pipe is shown in Fig. 1.

### 2.2.2. Sound field simulation method and process

The paper utilizes the Virtual Lab software to simulate the sound field distribution in the interior space using the model-based finite element method. The sound field simulation process is shown in Fig. 1. Firstly, the acoustical solution domains for the inside of the pipe and the room are created in the Virtual Lab software, and the automatic matching layer (AML) is used. Then, the three-dimensional body mesh of the solid air domain is generated. The division of the acoustic mesh should meet the requirement that the minimum unit size is no greater than  $1/6$  of the wavelength corresponding to the maximum frequency, as shown in Eq. (1):

$$L_s \leq \frac{c_{\text{air}}}{6f_{\text{max}}}. \quad (1)$$

The material parameters are set for the acoustical computational domain and the exterior-protection

construction. The wall pressure data obtained from the flow field simulation are used as the excitation conditions for the sound field. The structural modes required for acoustic simulations are calculated based on the Nastran platform. Finally, the field point grid is set up to monitor the simulation results.

### 2.3. Simulation of model settings

#### 2.3.1. Plan forms

The bedroom adjacent to the bathroom in the residence is the area most severely disturbed by drainage noise. Therefore, this paper takes the master bedroom suite with a bathroom as the base model to study the sound field distribution of drainage noise in the bedroom. The base model plane of the master bedroom suite and its dimensions are determined according to the Chinese standard GB 11977-2008 (Standardization Administration of the P.R.C., 2008), as shown in Fig. 2.

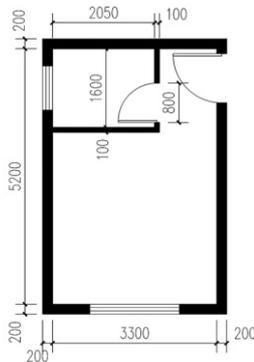


Fig. 2. Basic model plan.

This study focuses on the propagation of pipe drainage noise in a bedroom space. To simplify the model and facilitate calculations, only factors such as the position of the bathroom door, the position of the riser, and the material of the partition wall are considered. The lateral sound transmission effects of the external walls, floor, and other components are disregarded. The positions of the bathroom doors are shown in Fig. 3, and the positions of the riser arrangements are shown in Fig. 3c. To study the most unfavourable

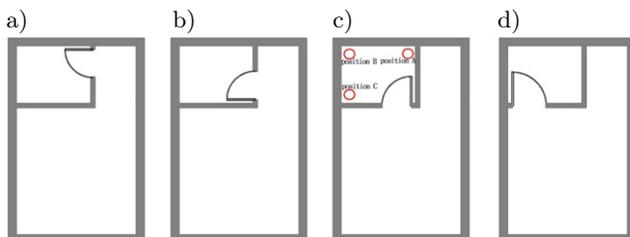


Fig. 3. Schematic diagram of the bathroom and plumbing positions: a) position 1; b) position 2; c) position 3; d) position 4.

case of noise impact, the bathroom door is set to be open. The partition wall between the bathroom and the bedroom is a block infill wall, and the material and thickness of the partition wall are shown in Table 1.

Table 1. Material and wall thickness of the bathroom partition walls.

Masonry materials	Ordinary fired brick	Ordinary concrete bricks	Lightweight aggregate concrete block	Fly ash block
Masonry thickness [mm]	120	100	100	100

#### 2.3.2. Drainage methods

The common drainage methods used in residential bathrooms are the same-floor and different-floor drainage. At present, the most widely used method of drainage in China's completed ordinary houses is different-floor drainage. However, during the drainage of upper-floor occupants, the flow noise and wall vibration noise from flow impacting the pipe section can have a serious impact on the lower floors (YANG *et al.*, 2016). In addition, the toilet is the sanitary ware with the highest incremental sound pressure level of drainage noise (JEONG *et al.*, 2017), so this paper uses a toilet drainage system with different-floor drainage as the simulation model.

#### 2.3.3. Pipe forms

Currently, depending on the combination of drainage and ventilation pipes, common UPVC single-riser and double-riser systems are more widely used in residential bathrooms in China. Compared with single risers, double risers are equipped with ventilation piping, which balances the air pressure and reduces the noise of air plugs. Therefore, the noisier single riser system is chosen as the base model (JIANG, WU, 2019). The pipe material is UPVC. The pipe size was determined according to the Chinese standard GB 50015-2019 (Ministry of Housing and Urban-Rural Development, China, 2019), as shown in Table 2.

Table 2. Drainage pipe simulation model sizes.

Parameter	Symbol	Value
Pipe diameter [mm]	$D$	110
Pipe wall thickness [mm]	–	3.2
Height of riser [mm]	$H_1$	2700
Length of transverse branch pipe [mm]	$L_1$	550
Slope of transverse branch pipe [°]	–	1.49
Height of sanitary ware pipe [mm]	$H_2$	200
Length of sanitary ware pipe [mm]	$L_2$	150

### 3. Drainage pipe sound source simulation

#### 3.1. Wall pressure due to the flow field in the drain

##### 3.1.1. Simulation model building

According to the pipe form model setting in Subsec. 2.3.3, the fluid simulation model of the internal flow field of the single riser system is created in the CATIA platform, as shown in Fig. 4.

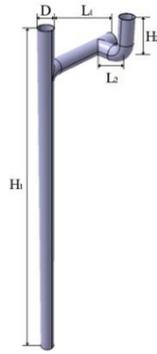


Fig. 4. Fluid simulation model.

The overall model is divided with a tetrahedral mesh, with boundary layers added near the walls. In addition, local mesh encryption is carried out at the intersection of the riser pipe and the transverse branch pipe, the transverse branch pipe and the sanitary ware drainage pipe, as shown in Fig. 5.

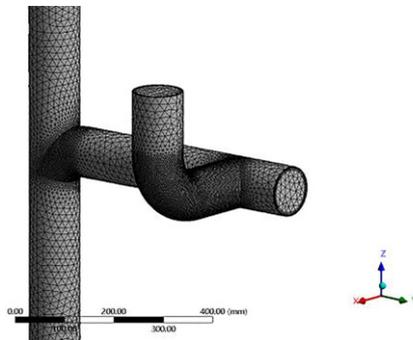


Fig. 5. Illustration of local grid encryption.

Finally, the number of mesh elements in the fluid computational domain reaches 3.5 million, the maximum value of the mesh skewness is  $0.76 < 0.9$ , and the minimum value of the orthogonal quality parameter is  $0.32 > 0.2$ . The overall mesh quality meets the Fluent solution requirements.

##### 3.1.2. Simulation parameter settings

**3.1.2.1. Boundary condition setting.** The boundary conditions are set consistent with the actual drainage process. The inlet of the sanitary ware pipe is defined as the velocity inlet, and the siphon toilet, which is widely used in China, was selected as the upper floor drainage appliance, with the inlet flow velocity set at

$v = 1.0$  m/s. Moreover, the total flow rate of the primary drainage is controlled by writing the equation of the velocity at the inlet and the proportion of each phase, adjusting the velocity inlet to stop feeding water when the total flow reaches the primary drainage of the toilet, that is,  $4.8 L$ . The velocity at the inlet is automatically adjusted to  $v = 0$  m/s. The other inlets of the pipe are pressure inlets, and the inlet pressure is  $p = 0$  Pa. The bottom cross-trunk pipe outlet is set as a pressure outlet with a pressure value of  $0$  Pa. The diameter of each pipe section is  $103.6$  mm, and the turbulence intensity is  $5\%$ . The pipe wall is set as a no-slip interface, and the wall roughness is set according to the wall material. The roughness is  $0.009$  mm when the wall is made of plastic, such as UPVC.

**3.1.2.2. Solution setup.** A double-precision solver was selected for this simulation. The pressure-velocity coupling is based on the pressure-implicit with splitting of the operator (PISO) algorithm. The discrete pressure format is set to body-force-weighted. The pressure and momentum terms are both second-order upwind terms, and the gradient is based on the least squares cell. The number of time steps is set to  $3.2 \times 10^4$ .

The flow field within the height range of the sound field simulation is selected, and the wall pressure data obtained are used as the acoustic boundary conditions for the subsequent simulation. In addition, the time step size of the transient simulation directly affects the frequency range of the acoustic simulation, and the quantitative relationship between the two is shown in Eq. (2):

$$t_s = \frac{1/f_{\max}}{2}. \quad (2)$$

In the sound field simulation, the highest frequency of interest in this study is  $4000$  Hz, and the time step size of the flow field simulation is set to  $1.25 \times 10^{-4}$  s based on the relationship between the step size and frequency.

##### 3.1.3. Field-induced wall pressure data

The vibration noise generation principle is described in Subsec. 2.1.2. For the bathroom pipe drainage process, the water flows to a certain flow rate from the upper end of the sanitary ware drainage pipe under the joint action of gravity and wall friction along the ware drainage pipe into the drainage transverse branch pipe and drainage riser, and the wall of the pipe at each connection produces constant changes in the impact, inducing vibration of thin-walled pipes and driving fluctuations in the external air medium, causing fluctuations in the sound pressure in the space and thus spreading the drainage noise to the indoor space. Therefore, the analysis of the pressure distribution on the wall surface of the pipe is the analysis of the sound source of pipe vibration.

Fluent transient calculations are used to simulate the primary drainage process of the toilet, intercepting part of the flow field in the room and calculating the pipe wall pressure at each time step in that range. As the pipe drainage is constantly changing, the velocity and pressure field distributions in the computational domain vary at each moment. When analysing the change in pressure distribution at the pipe wall during the primary drainage process, as the fluid flows through each section of the pipe, the pressure distribution at each part of the pipe wall within the height range of the interior space shows a specific change rule. To demonstrate the everchanging pipe wall pressure distribution, at 0 s–2 s, representative time points in different periods of fluid motion changes are selected to indicate the distribution of the pipe wall pressure in that period, as shown in Fig. 6.

Fluid starts to flow into the pipe from  $t = 0$  s. During the period from 0 s to 0.2 s, as shown in Fig. 6a, the pressure is positive in all areas except for the inlet pipe section, which is negative, and the pressure increases as the drainage volume increases. During the period from 0.2 s to 0.38 s, the fluid flows through the bend of the drainage apparatus pipe, and the vertically falling fluid has a significant impact on the outer wall of the bend, as shown in Fig. 6c, where the pressure gradually increases and is much greater than the wall pressure at other locations. From 0.38 s to 0.85 s, the fluid moves from the sanitary ware drain pipe to the drain transverse branch pipe, the impact of the fluid on the outside of the bend gradually decreases, the range of the negative pressure zone extends to the upper sidewall of the transverse sanitary ware drain pipe, as shown in Fig. 6c, and the maximum positive pressure value is reached at  $t = 0.422$  s. During the 0.85 s to 1.15 s period, the fluid moves mainly in the drain transverse branch pipe, and the extent of the negative pressure zone at the inlet increases further, while its pressure decreases. After this period, the fluid in the transverse branch pipe starts to move into the riser, the outer wall at the intersection of the riser and transverse branch

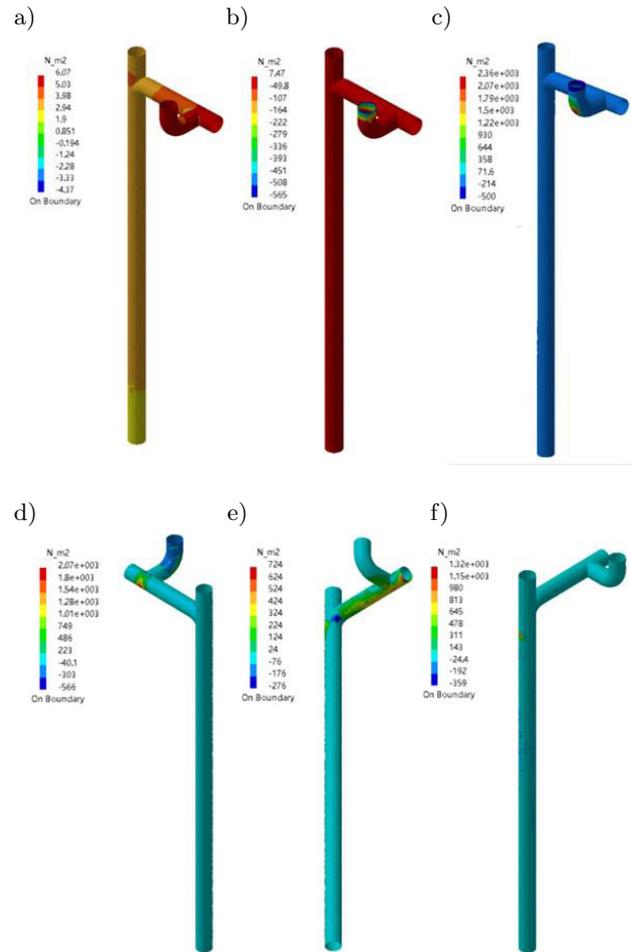


Fig. 6. Pressure distribution at the pipe wall at different moments: a)  $t = 0$  s; b)  $t = 0.2$  s; c)  $t = 0.256$  s; d)  $t = 0.422$  s; e)  $t = 1.142$  s; f)  $t = 1.782$  s.

pipe is subjected to a greater impact of the fluid, and a positive pressure zone starts to appear, but the area and pressure values of the positive pressure zone are smaller than before, as shown in Fig. 6f.

The maximum pressure data and the location of the wall surface at different times of the drainage process are shown in Table 3. A maximum negative pressure

Table 3. Maximum wall pressure data and locations at different times.

Period [s]	Maximum positive pressure value [N/m <sup>2</sup> ]	Maximum positive pressure position	Maximum negative pressure value [N/m <sup>2</sup> ]	Maximum negative pressure position
$t = 0-0.2$	<10	Other areas outside the inlet pipe section	More than -560	Inlet pipe section
$t = 0.2-0.38$	$4.32 \times 10^3$	At the bend	Above -500	Inlet to the area above the bend
$t = 0.38-0.85$	$2.07 \times 10^3$	A lower sidewall of the bend at the junction with the transverse branch pipe	Approximately -300	An upper sidewall of the bend where it meets the transverse branch pipe
$t = 0.85-1.15$	0–800	A lower sidewall of the transverse branch pipe	–	Upper sidewalls of sanitary ware drains and drainage of the transverse branch pipe
$t = 1.15$ later	$1.32 \times 10^3$	Outer pipe wall at the junction of the riser and the transverse branch pipe	–	–

of more than  $-560 \text{ N/m}^2$  occurs during the 0 s to 0.2 s period, and a maximum positive pressure of  $4.32 \times 10^3 \text{ N/m}^2$  occurs during the 0.2 s to 0.38 s period.

The impact of the water flow on the wall at different locations in the drainage process is constantly changing with the fluid movement, and the vibration of the pipe wall generated by the impact also changes at all times, with a larger pipe wall impact in the 0 s–1 s period. Therefore, the pressure data in the 0 s–1 s period are selected as the sound source for the acoustic simulation to simulate the impact of vibration noise caused by the fluid impacting the pipe wall at different locations within the pipe.

### 3.2. Pipe sound field simulation

#### 3.2.1. Simulation model and parameters

In practice, the flow field and the solid wall are coupled, and the change in the fluid flow characteristics inside the pipe has a large impact on the solid wall of the pipe. Therefore, the vibration modal and acoustic finite element methods are used to analyse the acoustic characteristics of noise in the pipe under fluid–solid coupling using the LMS Virtual Lab software. A three-dimensional geometric model of the pipe wall and the sound field inside the pipe was created based on the pipe dimensions described in Subsec. 2.3.3. A tetrahedral mesh of the structural and acoustical computational domains is delineated and extracted. The upper-frequency limit of interest for this simulation is 4000 Hz, and the mesh size of the pipe wall and the internal acoustic field is 14 mm, as shown in Fig. 7.

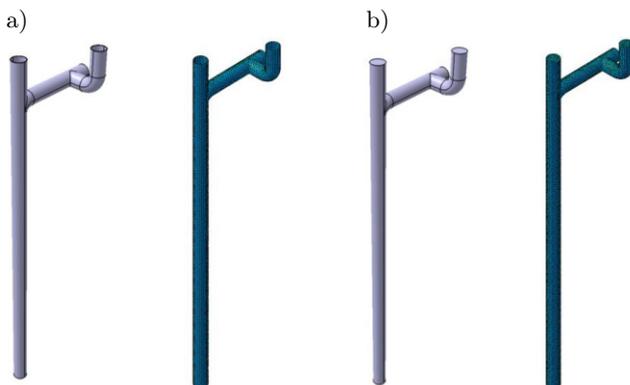


Fig. 7. Geometric modelling and meshing: a) pipe wall models and structural meshes; b) fluid domain and acoustic mesh in pipe.

The structural grid of the pipe wall is made of UPVC material. The material has Young’s modulus of  $3.14 \times 10^9$ , Poisson’s ratio of 0.32, and a density of  $1400 \text{ kg/m}^3$ . The volume proportion of water inside the pipe is so small that its effect is negligible. The material of the fluid domain in the pipe is set to air. The

wall pressure data in the 0 s–1 s time interval are used as the source conditions for the acoustic simulation, and the acoustic–vibrational coupling method is used in the acoustic finite element analysis module to simulate the distribution of the sound field inside the pipe.

#### 3.2.2. Pipe wall vibration modes

The vibration modes of the pipe wall are calculated in the noise and vibration module. Constraints are set in the middle of the riser and the transverse branch pipe to simulate the fixing effect of the pipe clamps and displacement constraints are set in the  $x$ -,  $y$ -, and  $z$ -directions at the upper and lower cross-sections of the riser and the upper cross-section of the sanitary ware pipe. Since the frequency range of concern for acoustic simulation is 100 Hz–4000 Hz, the modal solution range is set slightly larger than the acoustic solution range, which is 90 Hz–4010 Hz. The vibration modes of several frequencies are intercepted, as shown in Fig. 8.

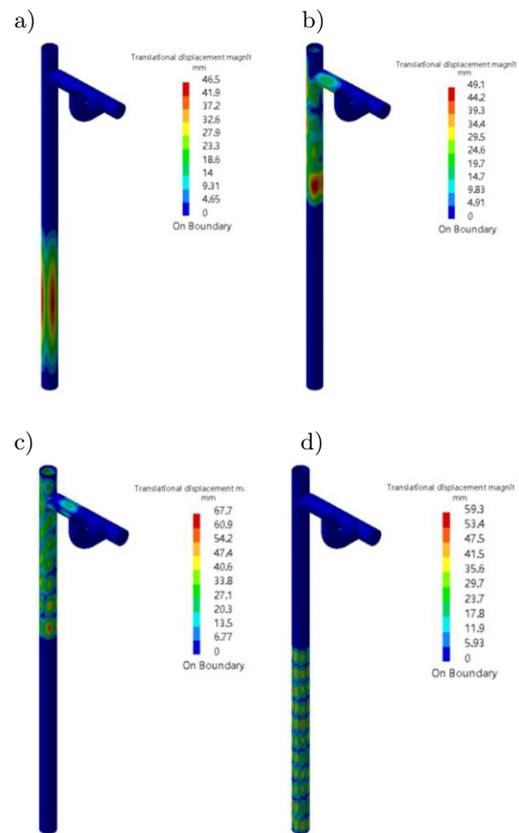


Fig. 8. UPVC pipe vibration modes: a) 250 Hz; b) 500 Hz; c) 1000 Hz; d) 2000 Hz.

The simulation results show that with the gradual increase in frequency, the vibration amplitude of the pipe wall gradually increases, and the maximum value of its vibration displacement gradually increases from 39.1 mm near 125 Hz to 129 mm near 4000 Hz. Due to the constraint effect of pipe clamps, the deformation

degree of different positions of the pipe is very different at the same frequency. For example, in Fig. 8b, the vibration deformation of the pipe at 500 Hz is mainly concentrated in the upper part of the riser, and the maximum vibration displacement in the middle can reach 49.1 mm, while the vibration deformation of the lower part of the riser and the transverse branch is smaller, and the resulting displacement can be ignored.

3.2.3. Sound field distribution inside the pipe

Based on the structural modal simulation results, the sound field distribution inside the pipe under external motivation is calculated. Sound pressure level monitoring points are set up in the acoustic solution domain at the positions shown in Fig. 9 to observe the radiated noise at different positions inside the pipe. A cloud diagram of the sound pressure level distribution in the solution domain is shown in Fig. 10a.

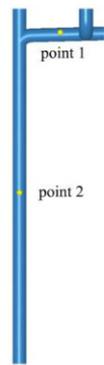


Fig. 9. Distribution of measurement points.

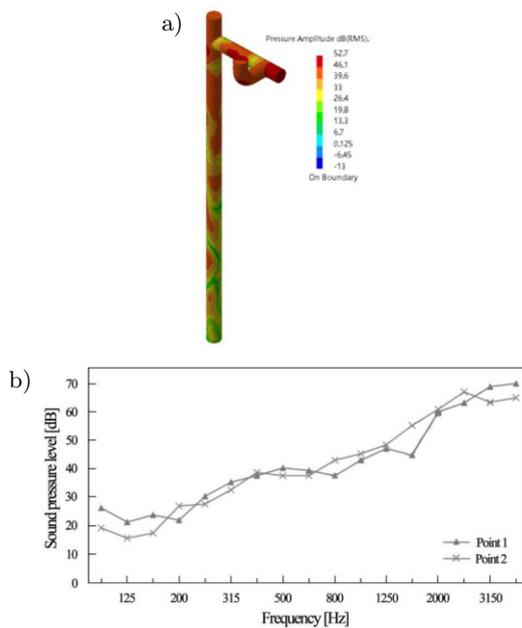


Fig. 10. Sound pressure level distribution cloud and spectral distribution of the measurement points: a) sound field distribution inside the pipe (1000 Hz); b) spectrum of sound pressure levels at monitoring points.

The spectral distribution of each observation point is shown in Fig. 10b.

The results show that the rheological vibration noise inside the pipe exhibits notable high-frequency characteristics, and the sound pressure gradually increases with increasing frequency. The maximum value occurs in the high-frequency range of 2000 Hz–4000 Hz, and the maximum sound pressure can reach more than 70 dB.

4. Results of the radiation sound field distribution of pipe drainage

To investigate the effects of the door position, pipe position, and partition wall material on the distribution characteristics of pipe drainage noise in the bathroom and the adjacent usable space.

4.1. Simulation model setup

4.1.1. Model building

According to the model settings in Subsec. 2.3, structural components such as pipe walls and bathroom partition walls are established in LMS Virtual Lab, and the three-dimensional geometric model of the interior acoustic solution domain is shown in Fig. 11. The positions of the different doors and pipes in the bathroom are shown in Table 4.

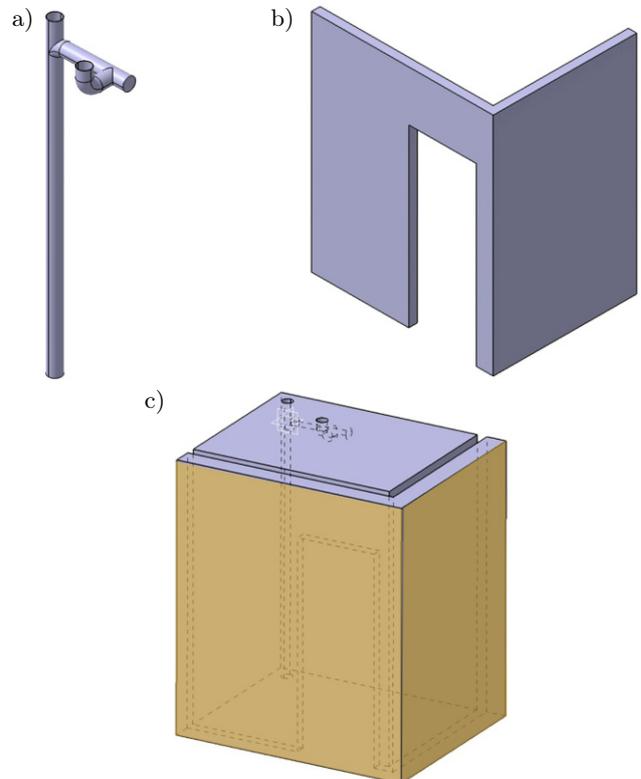
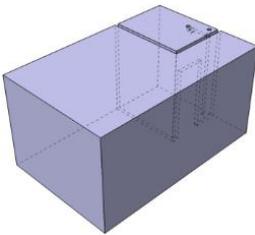
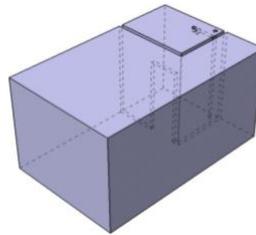
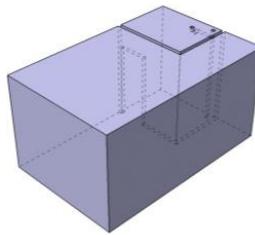
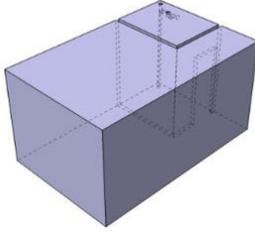
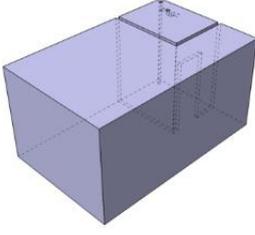
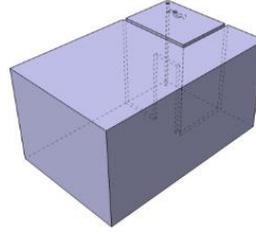
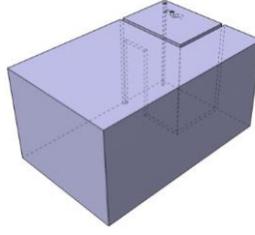
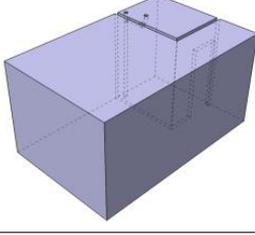
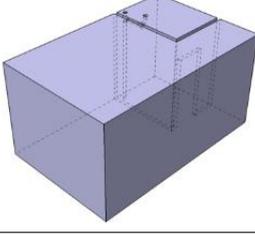
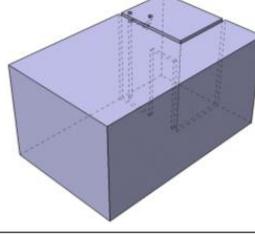


Fig. 11. Simplified simulation model of the residential bathroom: a) pipe wall; b) bathroom partition wall; c) AML boundary layer.

Table 4. Acoustic model settings.

Pipe positions	Door positions			
	1	2	3	4
A	(positions overlap)			
B				
C			(positions overlap)	

A, B, and C represent different pipe positions, with 1, 2, 3, and 4 representing different door positions in the bathroom, excluding the A1 and C3 working conditions where the pipe and door positions overlap, forming a total of 10 different working conditions.

4.1.2. Parameter setting

The mesh size of the pipe wall structure and the bathroom partition wall is set to 20 mm, and the mesh size of the indoor acoustic domain is set to 14 mm considering the upper limit of the simulation frequency and the computer performance. A field point grid and noise monitoring points were set up 1.2 m above the indoor floor to observe the distribution of drainage noise in the interior space. The structural mesh is imported into the noise and vibration module, and the contact surface between the partition wall and other surrounding enclosure components is set as a fixed constraint to determine the structural mode of the partition wall in the acoustic calculation frequency range.

4.2. Analysis of the sound field simulation results

4.2.1. Influence of the door position

To analyse the influence of the door position on the noise distribution of the pipe drainage in the external space, the partition wall material is set as ordinary concrete brick, and the pipe position in the model B is

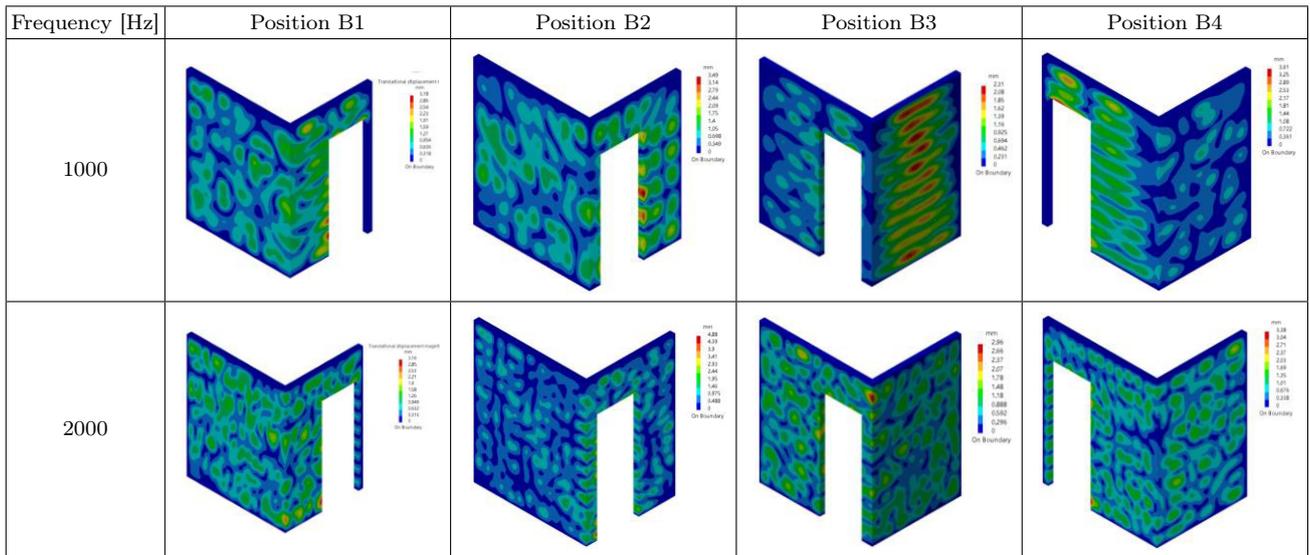
used. For different door positions, the drainage noise distribution is calculated under working conditions B1, B2, B3, and B4. The structural modes of the partition wall caused by pipe drainage in the acoustic calculation frequency range for medium frequencies of 1000 Hz and 2000 Hz are shown in Table 5.

The spatial sound field distribution at different door positions in the bathroom is simulated, and the results are shown in Fig. 12. The noise spectrum at the indoor monitoring points is shown in Fig. 13.

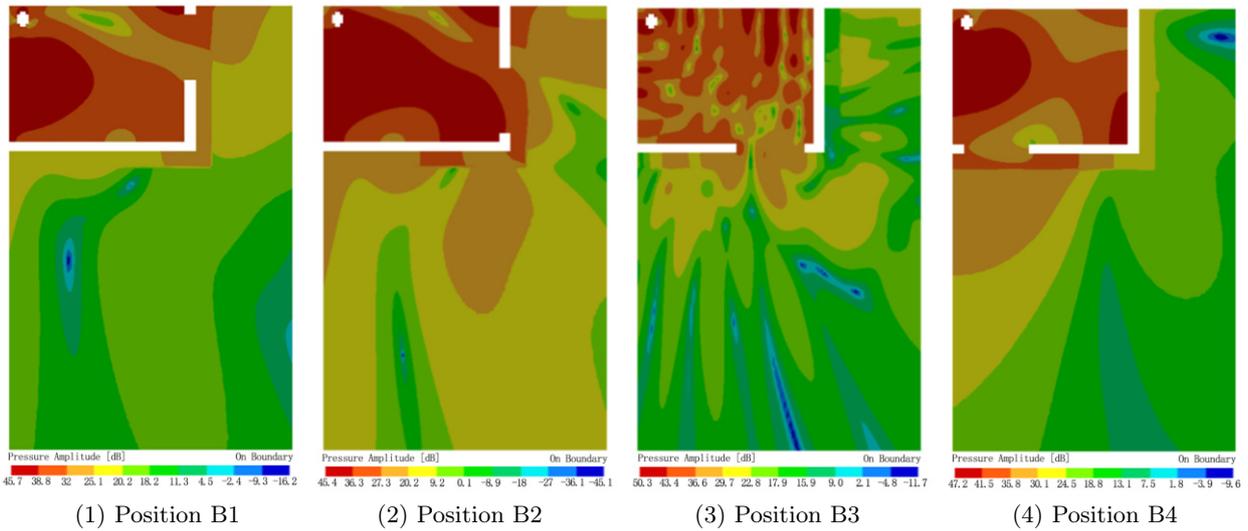
Analysis of the simulation cloud diagram and data shows that the pipe drainage noise has prominent high-frequency characteristics, and the sound pressure level is higher in the frequency range above 1000 Hz. The maximum noise level outside the pipe is distributed near the pipe, with the noise level gradually decreasing with increasing distance from the pipe. Analysis of the distribution cloud diagram of the planar sound pressure level indicates that the sound pressure level is significantly greater than that in the surrounding area of the connecting line between the riser and the door, and the difference between the middle- and high-frequency ranges can reach more than 10 dB.

The noise values of the monitoring points in the bedroom space for the B1, B2, B3, and B4 working conditions are 41.3 dB(A), 43.5 dB(A), 46.1 dB(A), and 43.2 dB(A), respectively, among which the noise value of the B3 working condition facing the bedroom space is greater than that of the other working conditions.

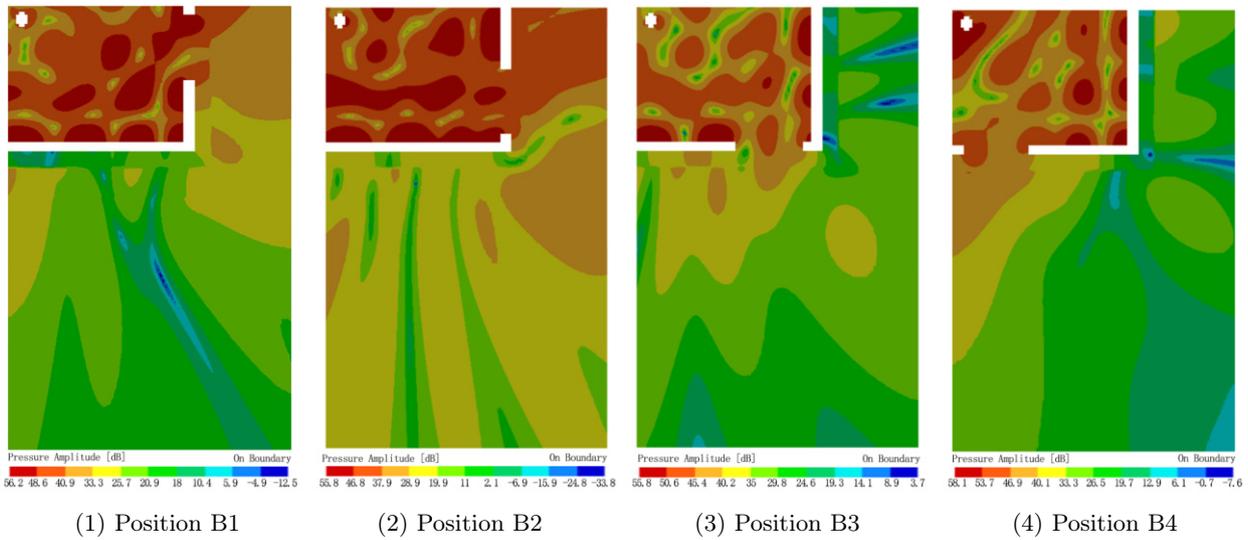
Table 5. Schematic representation of the structural modes of the partition wall.



a) 250 Hz

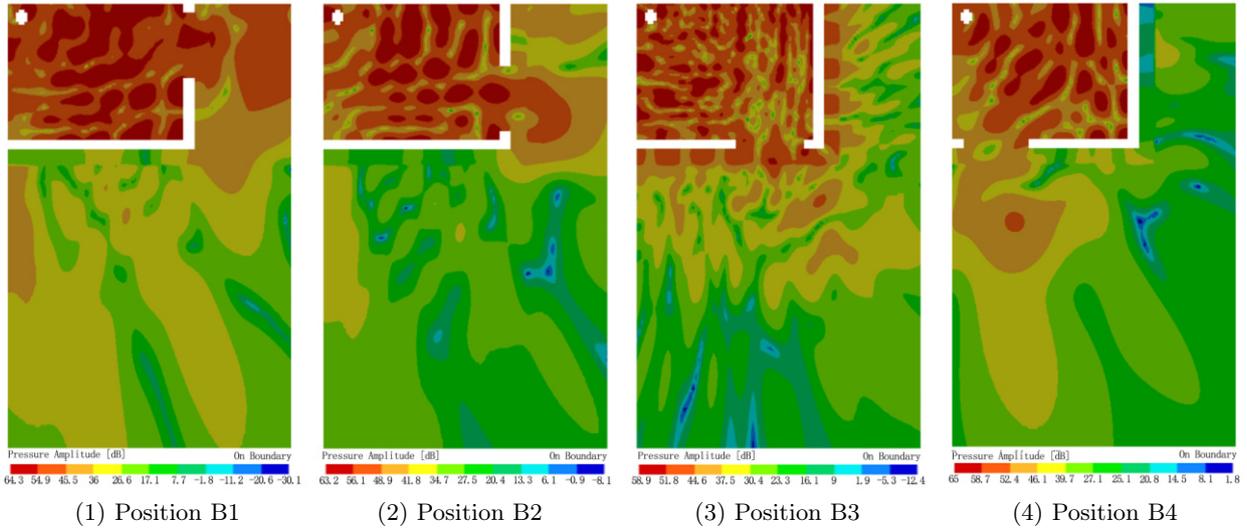


b) 500 Hz

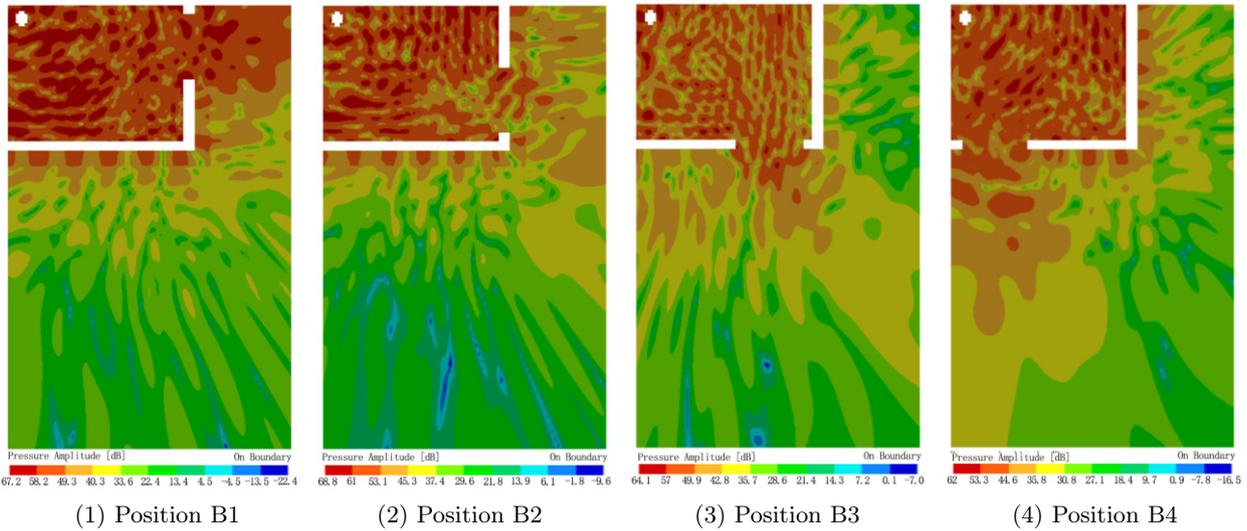


[Fig. 12ab.]

c) 1000 Hz



d) 2000 Hz



e) 4000 Hz

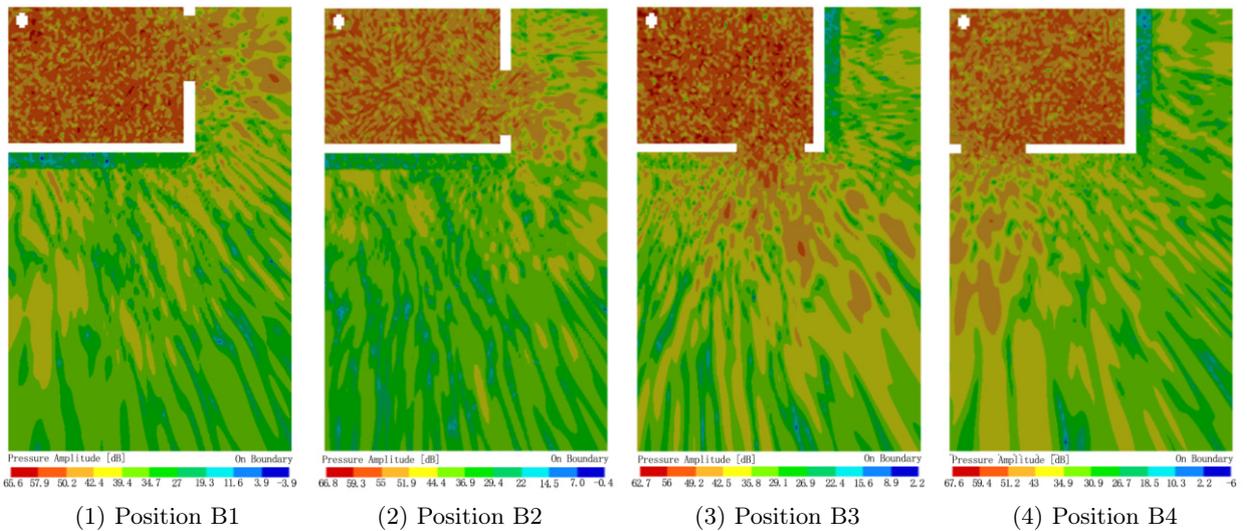


Fig. 12. Distribution of sound pressure levels in the external space at different door positions.

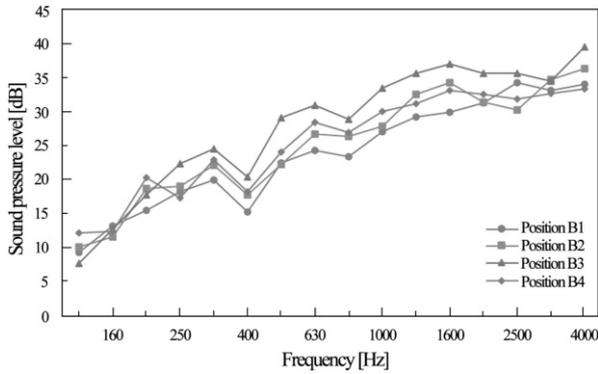


Fig. 13. Sound pressure levels at monitoring points.

When the door position corresponds to the B1 and B2 working conditions, the area with higher noise is mainly distributed at the entrance of the bedroom suite, and the impact on the noise in the bedroom space

is relatively small; under the B3 and B4 working conditions, as the relative angle between the door and the riser increases, the noise value of the indoor monitoring points increases.

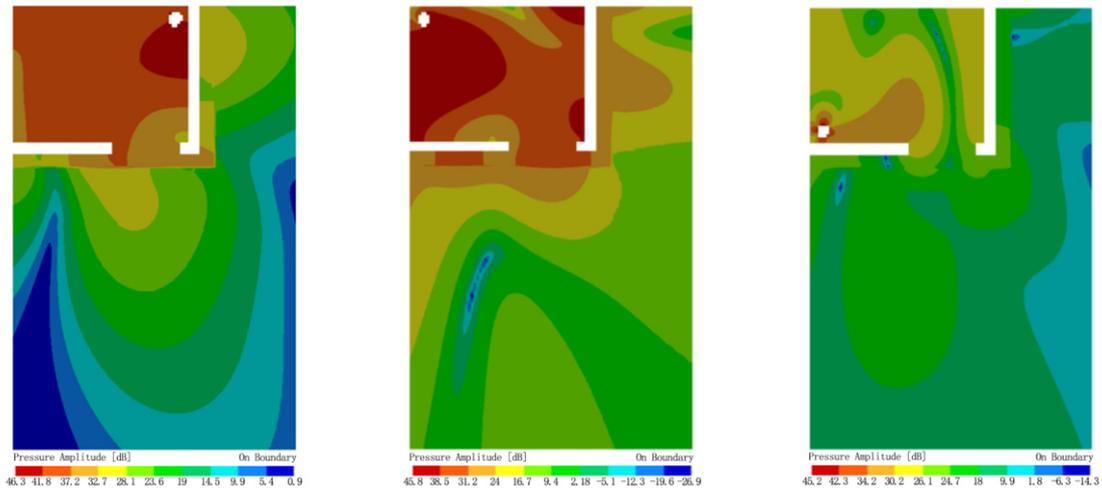
4.2.2. Influence of the riser position

To analyse the influence of the drainage riser position on the distribution of pipe drainage noise in the external space, the door position is configured for the Model 3, and the drainage noise distributions of the A3, B3, and C3 working conditions are calculated for different pipe positions.

The external spatial sound field distribution of the riser at different positions is shown in Fig. 14, and the noise spectrum of the indoor monitoring points is shown in Fig. 15.

In a comparison of the calculations for different riser positions, when the riser positions are in the A3,

a) 250 Hz

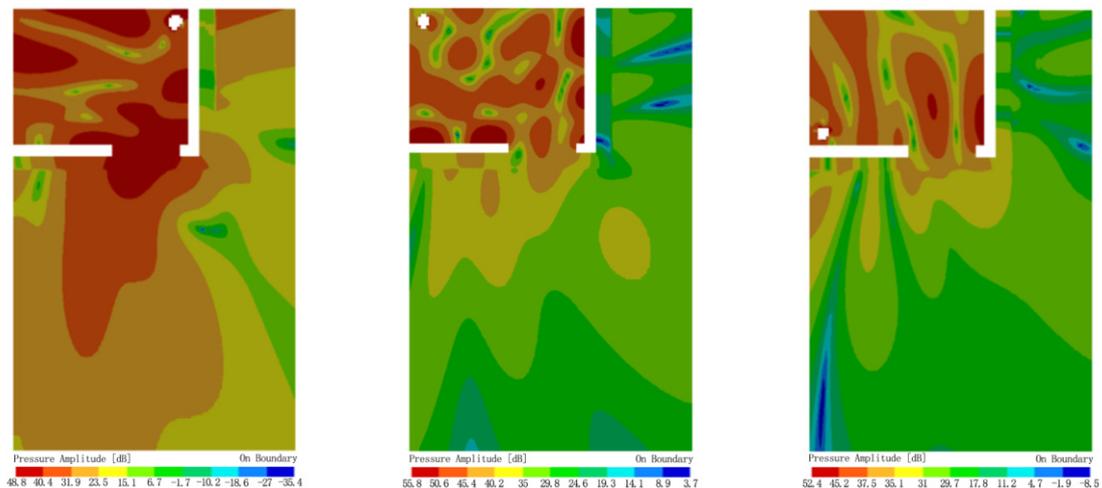


(1) Riser position A3

(2) Riser position B3

(3) Riser position C3

b) 500 Hz



(1) Riser position A3

(2) Riser position B3

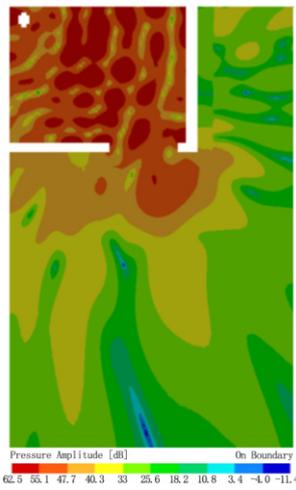
(3) Riser position C3

[Fig. 14ab.]

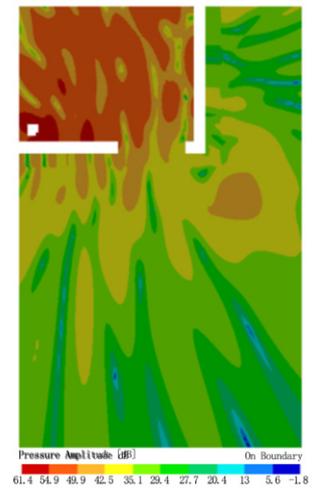
c) 1000 Hz



(1) Riser position A3

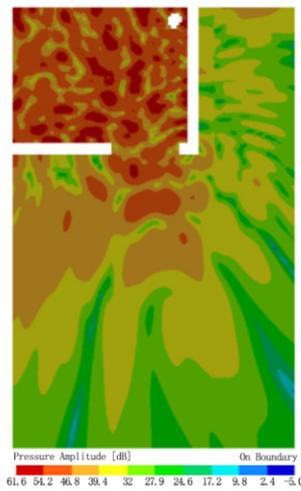


(2) Riser position B3

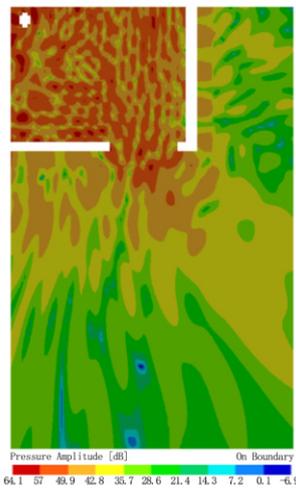


(3) Riser position C3

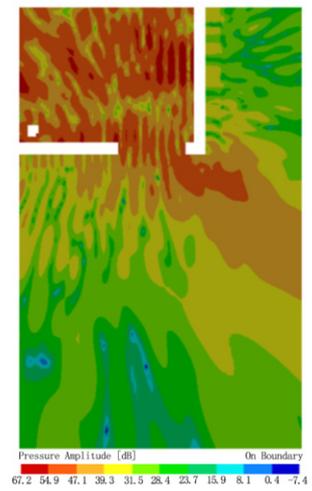
d) 2000 Hz



(1) Riser position A3

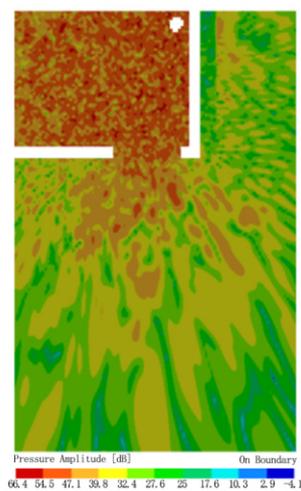


(2) Riser position B3

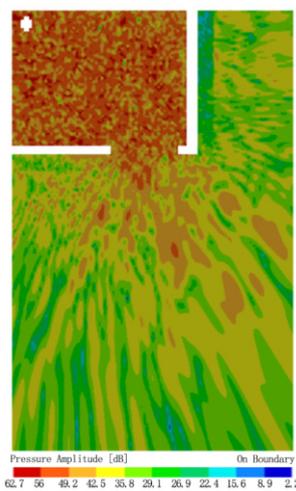


(3) Riser position C3

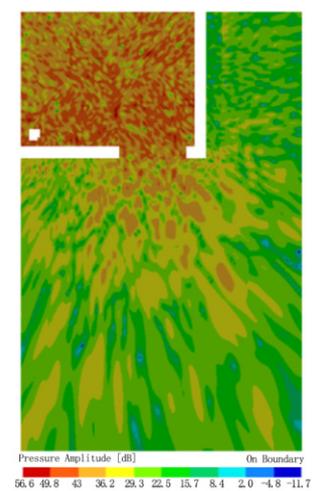
e) 4000 Hz



(1) Riser position A3



(2) Riser position B3



(3) Riser position C3

Fig. 14. Distribution of sound pressure levels in the external space at different riser positions.

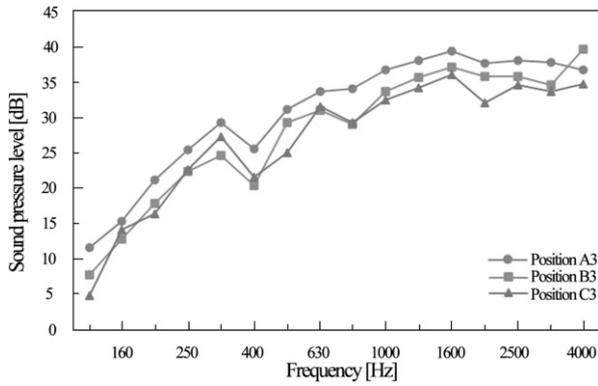


Fig. 15. Sound pressure levels at monitoring points.

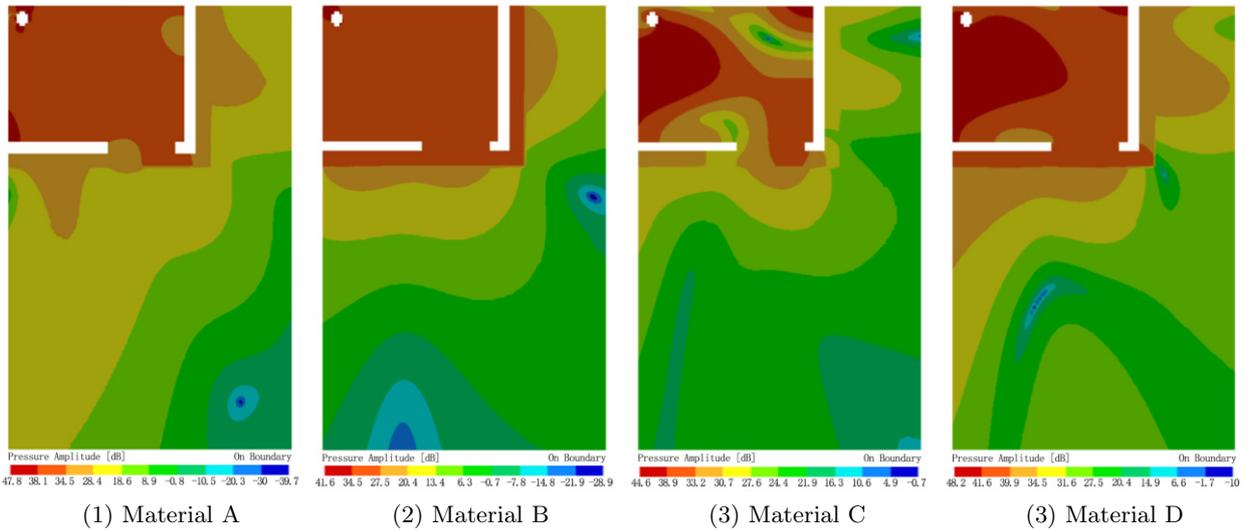
B3, and C3 working conditions, the pipe drainage noise values at the monitoring points in the bedroom are 48.2 dB(A), 46.1 dB(A), and 43.9 dB(A), respectively.

The A3 working condition has the highest noise impact because there is no direct shielding between the riser and the external space, so the drainage noise is directly transmitted to the outside.

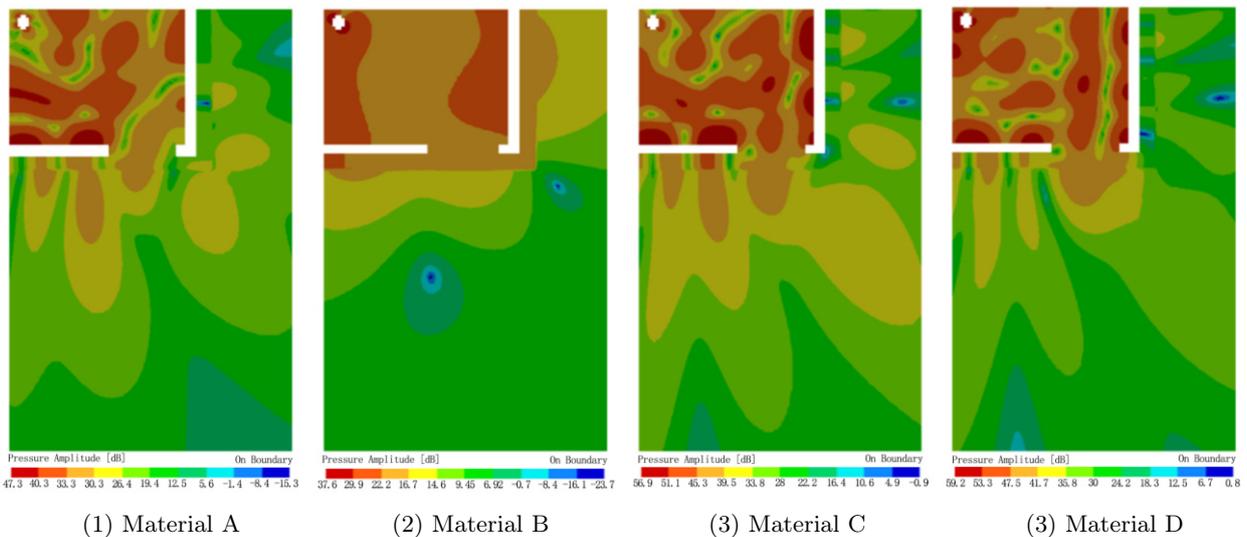
4.2.3. Influence of the partition wall material

To analyse the influence of the bathroom partition wall material on the noise distribution in the external space, the partition wall material parameter is used as the dependent variable, and the position of the pipe and door is set to the B3 working condition. Common masonry materials are selected for the partition wall material, including ordinary fired brick masonry, ordinary concrete brick masonry, a lightweight aggregate concrete block wall, and a fly ash block wall, and the corresponding property setting parameters are shown in Table 6.

a) 250 Hz

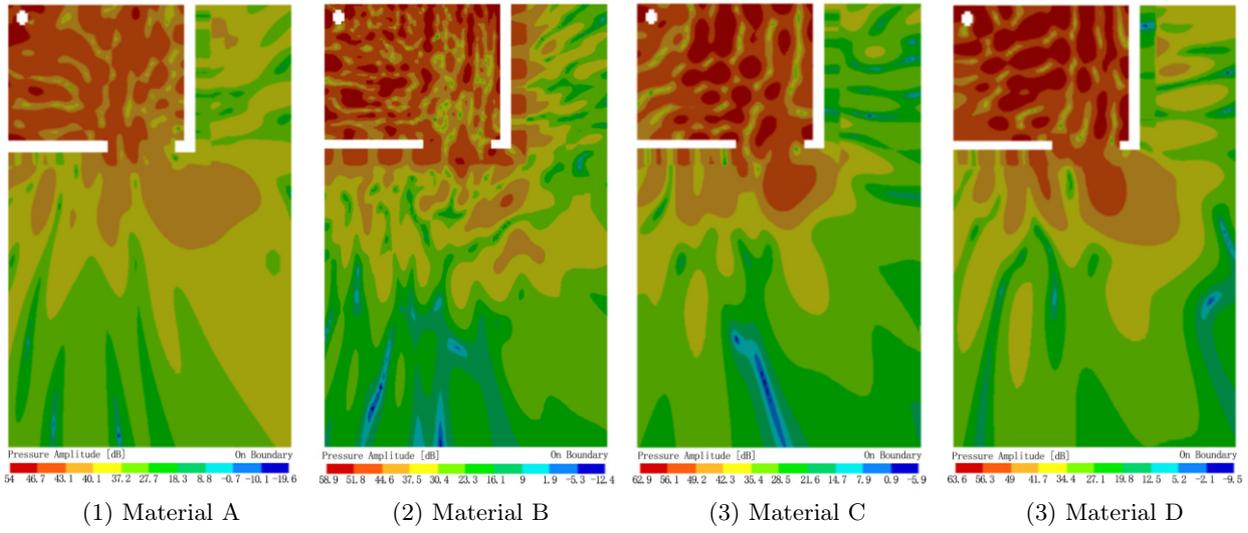


b) 500 Hz

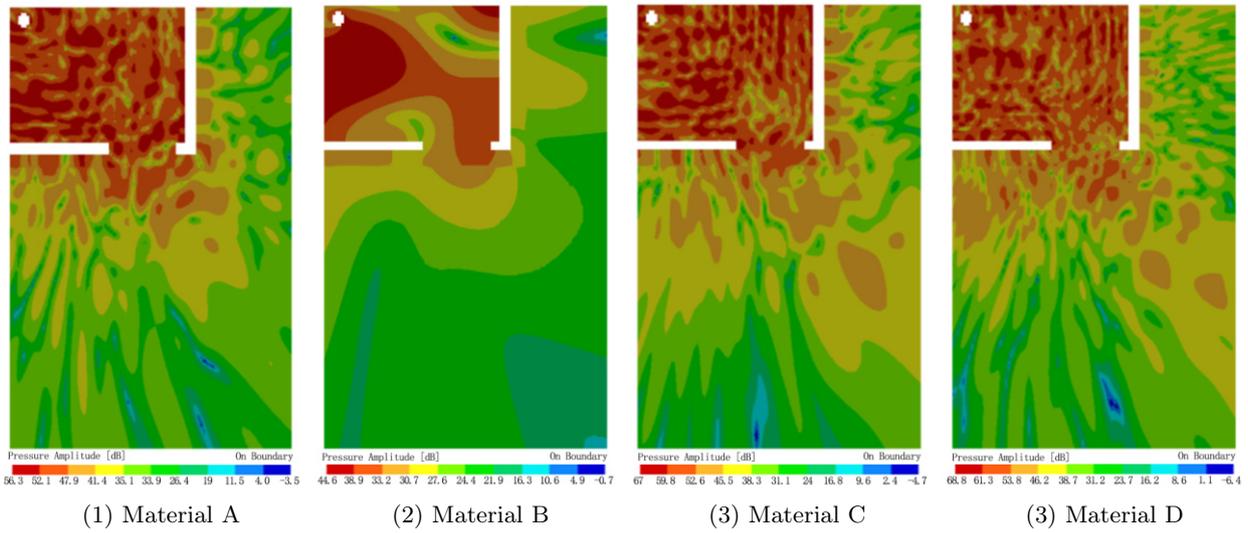


[Fig. 16ab.]

c) 1000 Hz



d) 2000 Hz



e) 4000 Hz

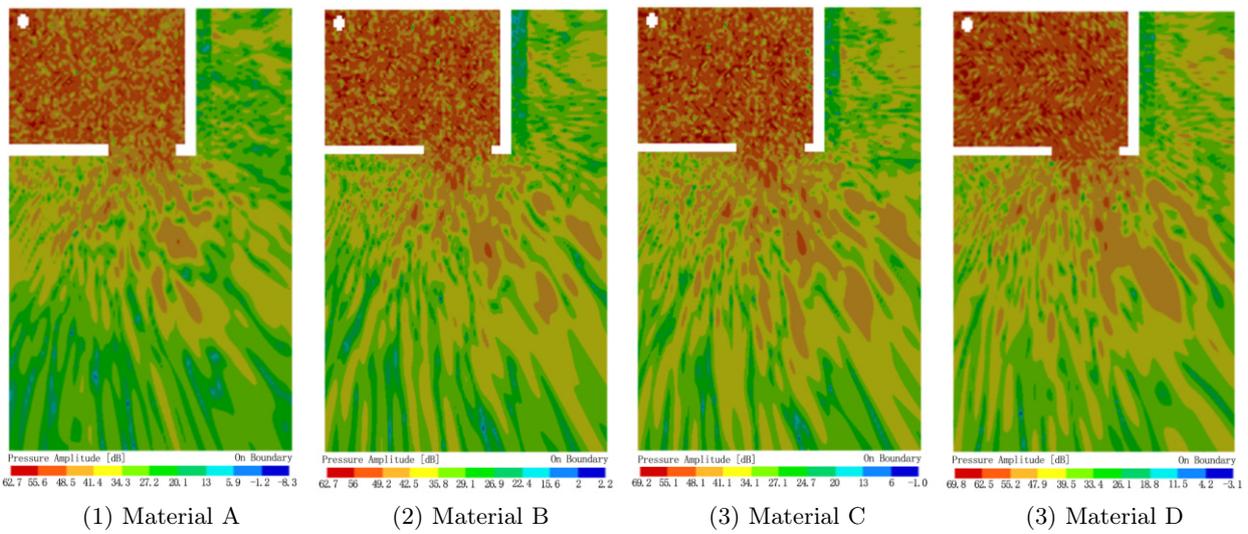


Fig. 16. Distribution of sound pressure levels in the external space for different partition materials.

Table 6. Materials for bathroom partitions.

Materials	Young's modulus [N/m <sup>2</sup> ]	Poisson's ratio	Density [kg/m <sup>3</sup> ]
Ordinary fired brick	$4.4 \times 10^9$	0.15	1800
Ordinary concrete bricks	$4.4 \times 10^9$	0.2	2000
Lightweight aggregate concrete block	$3.4 \times 10^9$	0.2	1000
Fly ash block	$2.83 \times 10^9$	0.2	1400

The sound field distribution in the external space with different partition wall materials. The results are shown in Fig. 16. The four materials in the diagram are represented by materials A, B, C, and D. The noise spectrum at the indoor monitoring point is shown in Fig. 17.

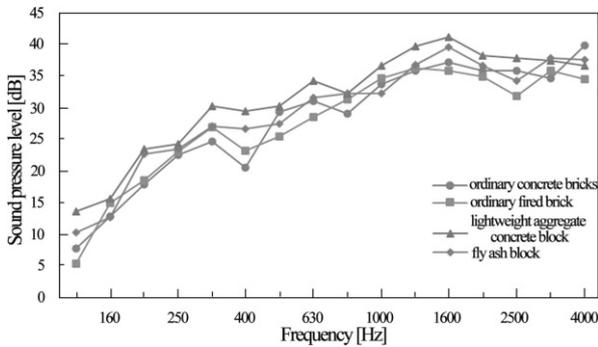


Fig. 17. Sound pressure levels at monitoring points.

A comparison of the noise data from the monitoring points in the bedroom shows that the noise gradually increases with increasing frequency, showing an increasing trend, with a few frequencies, such as 400 Hz, 800 Hz, and 2500 Hz, showing a slight decrease in the noise. The noise values of the monitoring points in the bedroom space under the four working conditions of ordinary concrete bricks, ordinary fired bricks, lightweight aggregate concrete blocks, and fly ash blocks are 46.1 dB(A), 44.6 dB(A), 48.1 dB(A) and 48.6 dB(A), respectively. When wall materials with higher masses, densities, and Young's moduli, such as ordinary concrete bricks and ordinary fired bricks, are used, the external space noise is slightly lower than that under working conditions with lightweight concrete and fly ash blocks. However, the difference between the noise values simulated under the four working conditions is not prominent.

### 5. Test of drainage noise

To verify the accuracy of the simulation results, the B3 working condition with the most simulated cases in Sec. 4 was chosen for the drainage noise field test to compare the results of the simulation and the test of drainage noise.

### 5.1. Test programme

The plan layout of the spaces within the test suite is shown in Fig. 18a. The partition wall between

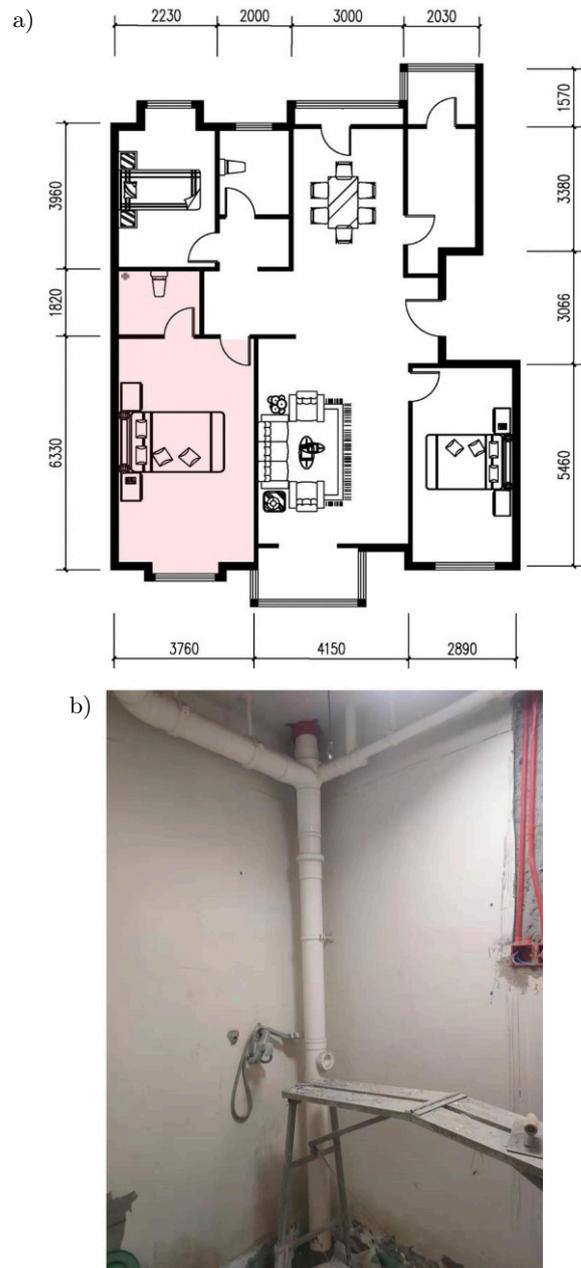


Fig. 18. Test condition master bedroom suite: a) plan layout of test condition; b) master bathroom riser.

the bathroom and the adjacent space is a 100 mm lightweight block wall. The sanitary ware is a toilet. A common UPVC single-riser drainage system with different floor drainage methods is used in the master bathroom, as shown in Fig. 18b. The riser in the master bathroom is located at the intersection of the second bedroom partition wall and the splitting wall, with the door opening towards the master bedroom. The layout is the same as that of the B3 working condition.

The field test consists of the indoor background noise and drainage noise. The background noise was measured with the doors and windows closed. During pipe drainage in the master bathroom, the noise in the adjacent space was measured under the most unfavourable conditions, that is, with the door open.

The test period was from 2 p.m.–4 p.m., and the instrument was a BK2260 precision noise analyser with a range of 0.8 dB–80 dB. The sound pressure level was measured at  $1/3$  octave centre frequencies in the range of 10 Hz–20 kHz. The sound pressure level at the centre of the room is used to represent the noise level of the whole space according to the relevant specifications for sound pressure level testing, with the height of the measurement point being 1.2 m above the indoor floor.

### 5.2. Comparison of simulation and test results

A comparison of the simulated and tested sound pressure levels for the B3 working condition is shown in Fig. 19. When the doors and windows are closed, the background noise level during the day is 33.7 dB(A), as shown in Fig. 20. Analysis of the data shows that

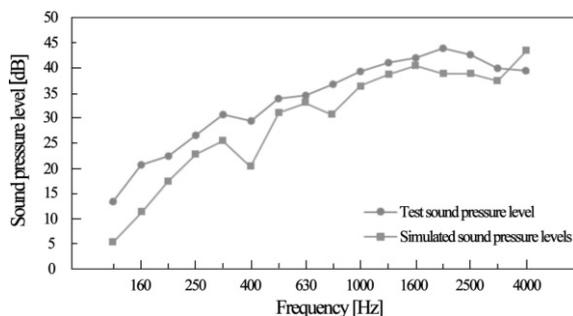


Fig. 19. Comparison of the sound pressure levels of simulations and tests.

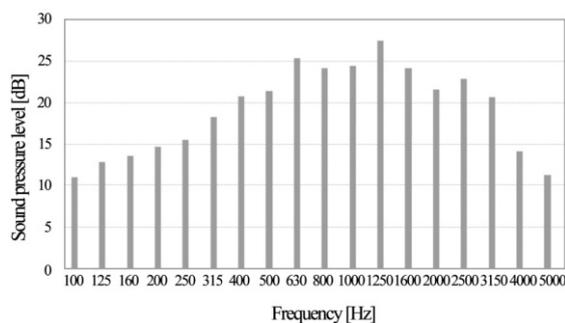


Fig. 20. Indoor daytime background noise.

the tested sound pressure level for drainage noise was 48.5 dB(A), which is 2.4 dB different from the simulated sound pressure level of 46.1 dB(A) for the B3 working conditions in Subsecs. 4.2.1 and 4.2.2, mainly due to the influence of background noise at the test site. Analysis of the sound pressure level frequency curves clearly shows that in the range of 125 Hz–3150 Hz, the test sound pressure level is slightly greater than the simulated sound pressure level, and in the range of 3150 Hz–4000 Hz, the simulated sound pressure level is slightly greater than the test sound pressure level, but the overall trend of the simulated and test sound pressure level frequency curves is similar. At the same time, both the simulated and tested sound pressure levels exhibit high-frequency characteristics, with the maximum noise occurring at approximately 1600 Hz–2000 Hz. A comparison of the conclusions shows that the results obtained by the numerical calculation method are consistent with the test results.

## 6. Conclusions

This paper analyses the causes and factors influencing bathroom drainage noise and uses the finite element numerical simulation technology with the help of the Fluent software and the LMS Virtual Lab to present a method to simulate the indoor drainage noise distribution using a joint simulation of the flow and sound fields. A three-dimensional model of the drainage noise computational domain, pipe wall, and partition wall was established to simulate the distribution of pipe drainage noise in the space under different working conditions, such as different door opening positions, riser arrangement positions, and partition wall materials. The conclusions are as follows:

- 1) The noise generated by drainage pipes increases with increasing frequency and has prominent high-frequency characteristics. Larger sound pressure levels often appear between high frequencies of 1600 Hz–2000 Hz.
- 2) The noise outside the pipe decreases gradually with an increasing distance from the pipe. The distribution of noise in the bedroom plane exhibits notable directionality due to the shielding and obstruction effect of the enclosure components. The sound pressure level near the connection line between the drainage riser and the door is greater than that of the surrounding area, and the difference increases with increasing frequency.
- 3) Among the four working conditions of the bathroom door position, the noise simulation value is 46.1 dB(A) when the bathroom door is arranged facing the used space, which is significantly greater than that under other working conditions. When the drainage riser is set towards the door, the simulated noise can reach 48.2 dB(A), which

is slightly greater than that under other working conditions. When the bathroom partition wall is made of four common materials – ordinary concrete bricks, ordinary fired bricks, lightweight aggregate concrete blocks, and fly ash blocks – the simulated noise at the bedroom monitoring point ranges from 44.6 dB(A) to 48.1 dB(A), and ordinary fired bricks have better sound insulation performance. Therefore, to reduce the interference of drainage noise from residential bathrooms, the position of bathroom doors, the position of pipes, and the construction of partition walls should be reasonably considered when designing building plans.

However, due to the limitations of the current level of the numerical simulation technology and the capabilities of the finite element simulation software, the three-dimensional model used in this paper has been simplified, and more ideal parameter settings have been adopted. However, it is hoped that the actual research and related conclusions in this paper can provide a reference for further research in the future.

### Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors. The authors would like to sincerely thank the anonymous reviewers for their constructive suggestions, which have substantially improved the quality of this study.

### References

- CAO Y., KE H., LIN Y., ZENG M., WANG Q. (2017), Investigation on the flow noise propagation mechanism in pipes of shell-and-tube heat exchangers based on synergy principle of flow and sound fields, *Applied Thermal Engineering*, **122**: 339–349, doi: [10.1016/j.applthermaleng.2017.04.057](https://doi.org/10.1016/j.applthermaleng.2017.04.057).
- CURLE N. (1955), The influence of solid boundaries upon aerodynamic sound, *Proceedings of the Royal Society of London, Series A. Mathematical and Physical Sciences*, **231**(1187): 505–514, doi: [10.1098/rspa.1955.0191](https://doi.org/10.1098/rspa.1955.0191).
- FFOWCS WILLIAMS J.E., HAWKINGS D.L. (1969), Sound generation by turbulence and surfaces in arbitrary motion, *Philosophical Transactions of the Royal Society of London, Series A, Mathematical and Physical Sciences*, **264**(1151): 321–342, doi: [10.1098/rsta.1969.0031](https://doi.org/10.1098/rsta.1969.0031).
- FUCHS H.V. (1983), Generation and control of noise in water supply installations: Part 1: Fundamental aspects, *Applied Acoustics*, **16**(5): 325–346, doi: [10.1016/0003-682x\(83\)90025-7](https://doi.org/10.1016/0003-682x(83)90025-7).
- FUCHS H.V. (1993a), Generation and control of noise in water supply installations. Part 2: Sound source mechanisms, *Applied Acoustics*, **38**(1): 59–85, doi: [10.1016/0003-682x\(93\)90041-4](https://doi.org/10.1016/0003-682x(93)90041-4).
- FUCHS H.V. (1993b), Generation and control of noise in water supply installations. Part 3: Rating and abating procedures, *Applied Acoustics*, **39**(3): 165–190, doi: [10.1016/0003-682x\(93\)90002-N](https://doi.org/10.1016/0003-682x(93)90002-N).
- HAN T. *et al.* (2020), Flow-induced noise analysis for natural gas manifolds using LES and FW-H hybrid method, *Applied Acoustics*, **159**: 107101, doi: [10.1016/j.apacoust.2019.107101](https://doi.org/10.1016/j.apacoust.2019.107101).
- JEON J.Y., JO H.I., KIM S.M., YANG H.S. (2019), Subjective and objective evaluation of water-supply and drainage noises in apartment buildings by using a head-mounted display, *Applied Acoustics*, **148**: 289–299, doi: [10.1016/j.apacoust.2018.12.037](https://doi.org/10.1016/j.apacoust.2018.12.037).
- JEONG A., KIM K., SHIN H., YANG K. (2017), Characteristics of reducing the water-drainage noise of toilet-bowl according to the composition of water drainage piping materials of the bathrooms of apartment housing [in Korean], *Transactions of the Korean Society for Noise and Vibration Engineering*, **27**(1): 114–120, doi: [10.5050/ksnve.2017.27.1.114](https://doi.org/10.5050/ksnve.2017.27.1.114).
- JIANG X., WU B. (2019), Design and exploration of drainage system based on pressure pipe in buildings, [in:] *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, **300**(2): 022016, doi: [10.1088/1755-1315/300/2/022016](https://doi.org/10.1088/1755-1315/300/2/022016).
- JUNG H.Y., SUNG S.H., JUNG H.J. (2012), Drainage noise reduction system using a vacuum double pipe elbow, [in:] *Proceedings of the Korean Society for Noise and Vibration Engineering Conference*, pp. 817–818.
- LAU S.K., POWELL E.A. (2018), Effects of absorption placement on sound field of a rectangular room: A statistical approach, *Journal of Low Frequency Noise, Vibration and Active Control*, **37**(2): 394–406, doi: [10.1177/1461348418780027](https://doi.org/10.1177/1461348418780027).
- LIANG Z., ZHOU G., ZHANG Y., LI Z., LIN S. (2006), Vibration analysis and sound field characteristics of a tubular ultrasonic radiator, *Ultrasonics*, **45**(1–4): 146–151, doi: [10.1016/j.ultras.2006.07.022](https://doi.org/10.1016/j.ultras.2006.07.022).
- LIGHTHILL M.J. (1952), On sound generated aerodynamically. I. General theory, *Proceedings of the Royal Society of London, Series A. Mathematical and Physical Sciences*, **211**(1107): 564–587, doi: [10.1098/rspa.1952.0060](https://doi.org/10.1098/rspa.1952.0060).
- LIU E., YAN S., PENG S., HUANG L., JIANG Y. (2016), Noise silencing technology for manifold flow noise based on ANSYS fluent, *Journal of Natural Gas Science and Engineering*, **29**: 322–328, doi: [10.1016/j.jngse.2016.01.021](https://doi.org/10.1016/j.jngse.2016.01.021).
- MAO Q., PIETRZKO S. (2013), *Control of noise and structural vibration*, Springer, London, pp. 3–4, doi: [10.1007/978-1-4471-5091-6](https://doi.org/10.1007/978-1-4471-5091-6).
- MARBURG S., NOLTE B. [Eds.] (2008), *Computational Acoustics of Noise Propagation in Fluids: Finite and Boundary Element Methods*, Springer, Berlin, doi: [10.1007/978-3-540-77448-8](https://doi.org/10.1007/978-3-540-77448-8).

18. Ministry of Housing and Urban-Rural Development of China (2019), *Standard for design of building water supply and drainage (GB 50015-2019)* [in Chinese], China Planning Press, Beijing, China.
19. MORI M., MASUMOTO T., ISHIHARA K. (2017), Study on acoustic, vibration and flow induced noise characteristics of T-shaped pipe with a square cross-section, *Applied Acoustics*, **120**: 137–147, doi: [10.1016/j.apacoust.2017.01.022](https://doi.org/10.1016/j.apacoust.2017.01.022).
20. NORTON M.P., BULL M.K. (1984), Mechanisms of the generation of external acoustic radiation from pipes due to internal flow disturbances, *Journal of Sound and Vibration*, **94**(1): 105–146, doi: [10.1016/s0022-460x\(84\)80008-5](https://doi.org/10.1016/s0022-460x(84)80008-5).
21. NORTON M.P., KARZUB D.G. (2003a), Pipe flow noise and vibration: a case study, [in:] *Fundamentals of Noise and Vibration Analysis for Engineers*, Cambridge University Press, Cambridge, pp. 441–487, doi: [10.1017/cbo9781139163927.009](https://doi.org/10.1017/cbo9781139163927.009).
22. NORTON M.P., KARZUB D.G. (2003b), Mechanical vibrations: a review of some fundamentals, [in:] *Fundamentals of Noise and Vibration Analysis for Engineers*, Cambridge University Press, Cambridge, pp. 1–127, doi: [10.1017/cbo9781139163927.003](https://doi.org/10.1017/cbo9781139163927.003).
23. PARK S.H., LEE P.J., JEONG J.H. (2018), Effects of noise sensitivity on psychophysiological responses to building noise, *Building and Environment*, **136**: 302–311, doi: [10.1016/j.buildenv.2018.03.061](https://doi.org/10.1016/j.buildenv.2018.03.061).
24. PUJOL S. *et al.* (2014), Indoor noise exposure at home: a field study in the family of urban schoolchildren, *Indoor air*, **24**(5): 511–520, doi: [10.1111/ina.12094](https://doi.org/10.1111/ina.12094).
25. RYU J., SONG H. (2019), Comparison between single-number quantities for rating noises from sanitary installations in residential buildings by objective and subjective methods, *Building and Environment*, **164**: 106378, doi: [10.1016/j.buildenv.2019.106378](https://doi.org/10.1016/j.buildenv.2019.106378).
26. SUN L., ZHE C., GUO C., CHENG S., HE S., GAO M. (2021), Numerical simulation regarding flow-induced noise in variable cross-section pipes based on large eddy simulations and Ffowcs Williams–Hawkings methods, *AIP Advances*, **11**(6): 065118, doi: [10.1063/5.0052148](https://doi.org/10.1063/5.0052148).
27. Standardization Administration of the P.R.C. (2008), *Functions and dimensions series of bathrooms in housing (GB 11977-2008)* [in Chinese], China Quality and Standards Publishing and Media Co., Ltd., China.
28. VILLOT M. (2000), Characterization of building equipment, *Applied Acoustics*, **61**(3): 273–283, doi: [10.1016/S0003-682X\(00\)00034-7](https://doi.org/10.1016/S0003-682X(00)00034-7).
29. XU N., ZHANG C., MA L. (2014), Residential building sound pollution and control measures, *Advanced Materials Research*, **945**: 711–716, doi: [10.4028/www.scientific.net/amr.945-949.711](https://doi.org/10.4028/www.scientific.net/amr.945-949.711).
30. YANG H.S., CHO H.M., KIM M.J. (2016), Field measurements of water supply and drainage noise in the bathrooms of Korea’s multi-residential buildings, *Applied Acoustics*, **6**(11): 372, doi: [10.3390/app6110372](https://doi.org/10.3390/app6110372).
31. YEON J.O., KIM K.W., YANG K.S., LEE B.K. (2014), Evaluation of noise characteristics of drainage and water supply systems in apartment bathrooms, *Advanced Materials Research*, **1025**: 987–990, doi: [10.4028/www.scientific.net/amr.1025-1026.987](https://doi.org/10.4028/www.scientific.net/amr.1025-1026.987).
32. ZHANG N., XIE H., XING W., WU B. (2016), Computation of vortical flow and flow induced noise by large eddy simulation with FW-H acoustic analogy and Powell vortex sound theory, *Journal of Hydrodynamics*, **28**(2): 255–266, doi: [10.1016/s1001-6058\(16\)60627-3](https://doi.org/10.1016/s1001-6058(16)60627-3).



## Research Paper

## Effects of Ultrasonic Power and Intensity of Mechanical Agitation on Pretreatment of a Gold-Bearing Arsenopyrite

Won Chol HONG\*, Ye Yong KIM, Chang Dok KWON, Kwang Chol SO

Faculty of Mining Engineering  
Kim Chaek University of Technology  
Democratic People's Republic of Korea

\*Corresponding Author e-mail: [hwc89217@star-co.net.kp](mailto:hwc89217@star-co.net.kp)

(received November 30, 2023; accepted March 5, 2024; published online August 1, 2024)

In this paper, the effects of an ultrasonic power and the intensity of mechanical agitation for pulp on alkaline pretreatment of gold-bearing arsenopyrite were investigated. The effect of pulp temperature on leaching efficiency in alkaline pretreatment of arsenopyrite was investigated under ultrasound and non-ultrasound conditions. Pre-treatment was followed by gold leaching tests with a cyanide solution. Compared with the non-ultrasound condition at the temperature of 60 °C, arsenic extraction and gold extraction was increased 20 %, 14.4 %, respectively, in the presence of ultrasound at ambient temperature. The characteristics of the ultrasonic power level as a function of the intensity of mechanical agitation were evaluated by a numerical simulation with CFD software – Ansys Fluent. The simulation results demonstrated that the stronger intensity of mechanical agitation, the lower ultrasonic power level. These results were proved through leaching experiments at different rotation speeds of impeller and ultrasonic powers.

The study results demonstrate that the ultrasound is an effective factor for pretreatment of gold bearing arsenopyrite and gold extraction is related to an ultrasonic power and the intensity of mechanical agitation.

**Keywords:** gold; arsenopyrite; alkaline pretreatment; ultrasound; computational fluid dynamics (CFD); Ansys Fluent.



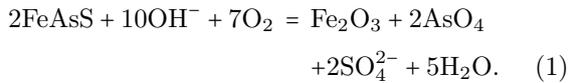
Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The refractory gold ore means the mineral of which gold extractions are less than 80 % after fine grinding using a traditional cyanide leaching (LA BROOY *et al.*, 1994). Commonly, arsenic-bearing gold ore is very refractory. Gold exists in the form of ultrafine particles or microscopic lattice within arsenic mineral in these ores, so it cannot attach to cyanide during the cyanide leaching process (MESA ESPITIA, LAPIDUS, 2015). Therefore, it is necessary to break these ores up to the gold particle size or decompose to make gold exposed before leaching (CHRYSOULIS, McMULLEN, 2005; DENG, GU, 2018; HASHEMZADEHFINI *et al.*, 2011). The process such as decomposition and crushing prior to leaching is called the pretreatment of refractory gold ore. The pretreatment is significant in the leaching of gold bearing arsenopy-

rite (CORKHILL, VAUGHAN, 2009). Recently, many researchers have studied the pretreatment of a gold bearing arsenopyrite to expose gold and remove arsenic by changing mineral composition, electrochemical and physicochemical properties of the gold ore (NAN *et al.*, 2014). There are many pretreatment methods of refractory gold ores including roasting oxidation, bacterial oxidation, pressure oxidation and chemical pretreatment. Particularly, the chemical treatment is widely used due to its low cost (DANG *et al.*, 2016). The typical methods for the chemical pretreatment include alkaline pretreatment, acid treatment, wet chlorination, and HNO<sub>3</sub> catalyzing oxidation decomposition. However, the alkaline pretreatment is considered as the most economical and eco-friendly (MENG *et al.*, 2003; MIKHLIN *et al.*, 2006). BHAKTA *et al.* (1989) developed the alkaline pretreatment of gold bearing arsenopyrite on the basis of previous literatures. This process needs

low potential for oxidizing gold bearing arsenopyrite on alkaline medium:



They found that the iron oxide film was hydrated and transferred into the porous state when sodium hydroxide was used. They also found that temperature was very significant. AWE and SANDSTRÖM (2010) pretreated a tetrahedrite-rich complex ore using selective dissolution of arsenic and antimony. The extraction ratio was related to the concentration of sulphur and hydroxide, temperature and reaction time. MENG *et al.* (2003) investigated to increase gold extraction in the leaching of arsenic-bearing gold ores. They confirmed that gold was separated from sulphur and arsenic in the concentrate when the alkaline pretreatment was applied to it in the conditions of ambient temperature and pressure after ultra-grinding. This method is not widely used due to the dependence of the alkaline pretreatment on alkaline concentration, temperature and pressure.

Recently, the ultrasound has been widely used to accelerate the speed of chemical reaction (CONTAMINE *et al.*, 1994). Unlike microwave, electrochemical and photochemical processing, all it needs is the medium for its transmission. Ultrasound can break down cavitation bubbles near the heterogeneous solid-liquid interface prior to generation of asymmetric flow and liquid jet. It can increase mass and heat transformation due to the breakdown of boundary layers. Collision between particles can occur corrosion and washing of a solid surface and decrease wettability. Ultrasound can cause a weak flow of the solid-liquid boundary surface and reduce the diffusion thickness. It can accelerate the diffusion speed to increase extraction. Many researchers have already studied the gold extraction using ultrasound (ZHU *et al.*, 2012), the effect of ultrasound on zinc extraction (SLACZKA, 1986), the effect of ultrasound on copper extraction from copper ore (BESE, 2007; WANG *et al.*, 2017), the effect of ultrasound on leaching of phosphate using hydrochloric acid (TEKIN, 2002), and the effect of ultrasound on sulfuric acid leaching of colemanite (TAYLAN *et al.*, 2007).

As the application range of ultrasound broadens, many studies have been made for analyzing fluid dynamic characteristics as a function of ultrasound parameters during leaching combined ultrasound such as size distributions and energy levels of cavitation bubbles as functions of power and frequency of ultrasound (MEROUANI *et al.*, 2014; MÜLLER *et al.*, 2014), change of the characteristics of fluid flow due to ultrasound and mechanical agitation (SAJJADI *et al.*, 2015; KOJIMA *et al.*, 2010). However, the effects of ultrasonic power and intensity of mechanical agitation on leaching of minerals are rarely involved.

Therefore, an investigation concerning the effect of ultrasound and the relation between ultrasonic power and intensity of mechanical agitation was conducted by computational fluid dynamics simulation and experiments.

## 2. Experimental

### 2.1. Materials

The refractory gold bearing arsenopyrite was obtained from DokSong Mine, Democratic People's Republic of Korea. The arsenopyrite samples with a particle size – 0.074 mm were used in all experiments. The chemical composition of material is given in Table 1. The samples were crushed, ground, and then sieved using ASTM standard sieves.

Table 1. Composition analysis result.

Elements	Amount [%]
Au ( $\times 10^{-6}$ )	19.87
Ag ( $\times 10^{-6}$ )	1.44
As	41.53
Fe	35.51
S	20.21
SiO <sub>2</sub>	2.49
Al <sub>2</sub> O <sub>3</sub>	0.13
CaO	0.05
MgO	0.08

### 2.2. Experimental methods

The experiments were conducted in three stages.

In the first stage, gold bearing arsenopyrite was pre-leached using alkali. This experiment was carried out in 1000 ml beaker placed in the thermostatic bath within a precision of  $\pm 0.1$  °C, using a mechanical agitator (LR500A, Yamato Scientific Co., Ltd., Japan) and hydraulic ultrasonic generator (GuangZhou Hengda Ultrasonic Electric Technological Ltd., China) with the power adjustable in the range of 0 W–150 W and 20 kHz frequency, 200 g arsenopyrite sample, and 800 ml distilled water were introduced into the experimental beaker equipped with the ultrasonic probe and thermometer. After that, 20 g sodium hydroxide (analytically pure) was put into the solution and mechanically stirred for 6 h at 350 rpm at different temperatures (20 °C, 40 °C, 60 °C, and 80 °C) without ultrasound. Then, it was stirred for 6 h at 20 °C in the presence of ultrasound. While these experiments were carried out, leaching pulps were taken and filtered at the time intervals of 1 h to survey arsenic extraction using atomic absorption spectrometer (model Varian SpectrAA 220FS). After the alkaline pretreatment, the pulp was filtered, washed with distilled water three

times and then the residue was dried at 120 °C. Dried residues were used in the next stage of experiments.

In the second stage, the pretreated arsenopyrite samples were leached using sodium cyanide; 150 g pretreated sample and 600 ml distilled water were poured into the beaker. The pulp was mechanically stirred at 600 rpm, at room temperature. The pH of solution was adjusted to 10.5~11 using sodium hydroxide. Then, 1.2 g chemically pure sodium cyanide was added to the pulp. In other words, the consumption of sodium cyanide is 8 kg/t on plant scale. The leaching was conducted for 8 h. At the time intervals of 1 h, pulps were taken and filtered to survey gold extraction. The concentration of gold in pulp was analyzed using an atomic absorption spectrometer (model Varian SpectraAA 220FS).

In the third stage, the relation between ultrasonic power and intensity of mechanical agitation was analyzed by computational fluid dynamic simulation and experiments. For analyzing transmission characteristics of ultrasonic power, unsteady LES (large eddy simulation) model and FW-H (Ffows-Williams & Hawkings) model were used. Based on the simulation result, the arsenic and gold extractions were compared as a function of time by varying the ultrasonic power (150 W, 400 W, 1000 W) and agitation intensity (150 rpm, 350 rpm, 550 rpm, 750 rpm).

The experimental set-up is shown in Fig. 1.

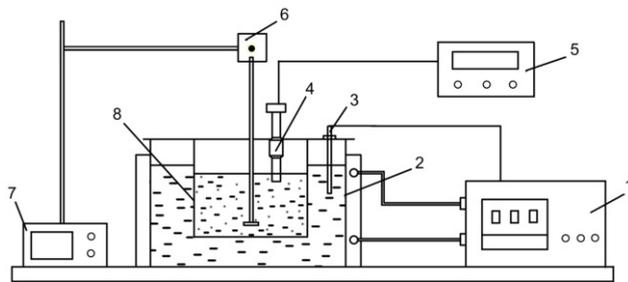


Fig. 1. Schematic experimental set-up; 1) thermostat; 2) thermostatic bath; 3) thermometer; 4) ultrasonic generator; 5) hydraulic pump; 6) mechanical agitator; 7) agitator controller; 8) beaker.

### 3. Results and discussions

#### 3.1. Effect of ultrasound on alkaline pretreatment

The effect of ultrasound on the alkaline pretreatment of gold-bearing arsenopyrite was shown in Figs. 2 and 3.

As shown in Fig. 2, the temperature of pulp had a great influence on leaching dynamics under no ultrasound. The arsenic extraction increased as the temperature increased. However, the initial extraction was significantly high at 80 °C, after 5 h it became lower than the one at 60 °C. Under ultrasound condition, arsenic extraction was a little lower than the one without ultra-

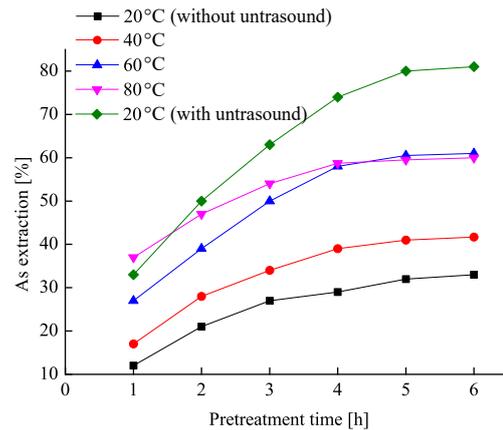


Fig. 2. Arsenic extractions with and without ultrasound on alkaline pretreatment.

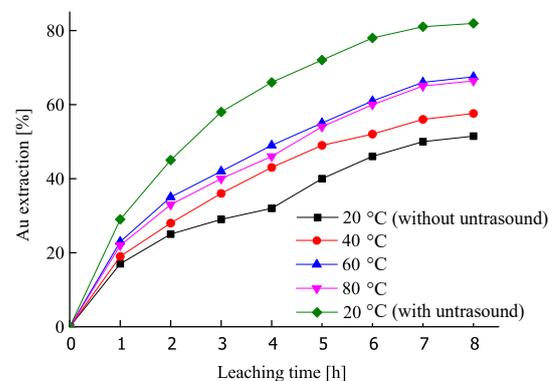


Fig. 3. Gold extractions with and without ultrasound on alkaline pretreatment.

sound at 1 h and it increased 20 % more than the one without ultrasound at the end of leaching. This indicates that ultrasound has a great influence on leaching dynamics. Due to the cavitation, the extremely high temperature and pressure are provided on the interface between solution and solid matrix. PENN *et al.* (1959) found that ultrasound wave reduced the thickness of boundary layer for the mass transfer, thus facilitating the mass transfer.

Figure 3 shows that the gold extraction in cyanidation was in proportion to arsenic extraction.

#### 3.2. Relation between ultrasonic power and intensity of mechanical agitation on alkaline pretreatment

##### 3.2.1. Computational fluid dynamics simulation

##### 3.2.1.1. Modeling for the simulation. Acronyms:

- $r$  – radial coordinates of the receiver location,
- $I(r, \theta; \mathbf{y})$  – directional ultrasonic intensity per unit volume of a jet,
- $C$  – modified convection factor,

- $u$  – turbulence velocity,  
 $M$  – dimensionless parameter,  
 $N$  – dimensionless parameter,  
 $k$  – turbulence kinetic energy,  
 $P_A$  – total ultrasonic power,  
 $P_{\text{ref}}$  – reference ultrasonic power,  
 $\theta$  – angular coordinates of the receiver location,  
 $\epsilon$  – turbulence dissipation rate.

The Jet Noise Source model was proposed by Goldstein and Ribner who modified the model originally defined by Ribner. It is considered for anisotropy of turbulence in axisymmetric turbulent jets.

Goldstein proposed that the total ultrasonic power emitted by the unit volume of a turbulent jet is calculated from:

$$\begin{aligned}
 P_A(\mathbf{y}) &= \int_{2\pi}^0 \int_{\pi}^0 I(r, \theta; \mathbf{y}) r^2 \sin \theta \, d\theta \, d\psi \\
 &= 2\pi r^2 \int_{\pi}^0 I(r, \theta; \mathbf{y}) \sin \theta \, d\theta, \quad (2)
 \end{aligned}$$

where  $r$  and  $\theta$  are the radial and angular coordinates of the receiver location, and  $I(r, \theta; \mathbf{y})$  is the directional ultrasonic intensity per unit volume of a jet defined by:

$$\begin{aligned}
 I(r, \theta; \mathbf{y}) &= \frac{12\rho_0\omega_f^4 L_1 L_2^2 \overline{u_{t1}^2}}{5\pi a_0^5 r^2} \frac{D_{\text{self}}}{C^5} \\
 &+ \frac{24\rho_0\omega_f^4 L_1 L_2^4 \overline{u_{t1}^2}}{\pi a_0^5 r^2} \left( \frac{\partial U}{\partial r} \right)^2 \frac{D_{\text{shear}}}{C^5}. \quad (3)
 \end{aligned}$$

$C$  is the modified convection factor defined as:

$$C = 1 - M_c \cos \theta \quad (4)$$

and

$$\begin{aligned}
 D_{\text{self}} &= 1 + 2 \left( \frac{M}{9} - N \right) \cos^2 \theta \sin^2 \theta \\
 &+ \frac{1}{3} \left[ \frac{M^2}{7} + M - 1.5N \left( 3 - 3N + \frac{1.5}{\Delta^2} - \frac{\Delta^2}{2} \right) \right] \sin^4 \theta, \quad (5)
 \end{aligned}$$

$$D_{\text{shear}} = \cos^2 \theta \left[ \cos^2 \theta + \frac{1}{2} \left( \frac{1}{\Delta^2} - 2N \right) \sin^2 \theta \right]. \quad (6)$$

The other parameters are defined by:

$$\Delta = \frac{L_2}{L_1}, \quad (7)$$

$$M = \left[ \frac{3}{2} \left( \Delta - \frac{1}{\Delta} \right) \right]^2, \quad (8)$$

$$N = 1 - \frac{\overline{u_{t2}^2}}{\overline{u_{t1}^2}}, \quad (9)$$

$$L_1 = \frac{\overline{u_{t1}^2}^{1.5}}{\epsilon}, \quad (10)$$

$$L_2 = \frac{\overline{u_{t2}^2}^{1.5}}{\epsilon}, \quad (11)$$

$$\omega_f = 2\pi \frac{\epsilon}{k}, \quad (12)$$

where  $\overline{u_{t1}^2}$  and  $\overline{u_{t2}^2}$  are computed, respectively, according to the turbulence model chosen for computing. If the RSM is selected, they are computed from the corresponding normal stresses. For all other two-equation turbulence models, they are calculated from:

$$\overline{u_{t1}^2} = \frac{8}{9}k, \quad (13)$$

$$\overline{u_{t2}^2} = \frac{4}{9}k. \quad (14)$$

In *Ansys Fluent* (2016), the ultrasonic power both in the dimensional units and dB is computed from:

$$L_P = 10 \log \left( \frac{P_A}{P_{\text{ref}}} \right), \quad (15)$$

where  $P_{\text{ref}}$  is the reference ultrasonic power ( $P_{\text{ref}} = 10^{-12}$  W by default).

*3.2.1.2. Model structure.* Model of mechanical agitator equipped ultrasonic generator is as shown in Fig. 4.

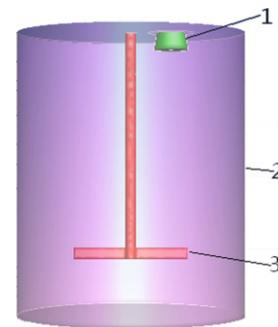


Fig. 4. Model of mechanical agitator equipped ultrasonic generator: 1) ultrasonic generator; 2) agitator box; 3) agitator impeller.

The unstructured mesh of a mechanical agitator equipped with an ultrasonic generator was created in Ansys ICEM. The model domain was discretized into 583.742 hexahedral elements and tetrahedral elements, 964.532 nodes.

Figure 5 shows the mesh model structure of mechanical agitator equipped with an ultrasonic generator. The model was meshed in minimum size 0.5 mm (around ultrasonic probe) and the stability of the mesh is performed by check of mesh.

Base data to create calculation structure was determined such as experimental conditions for simulation analysis (Table 2).

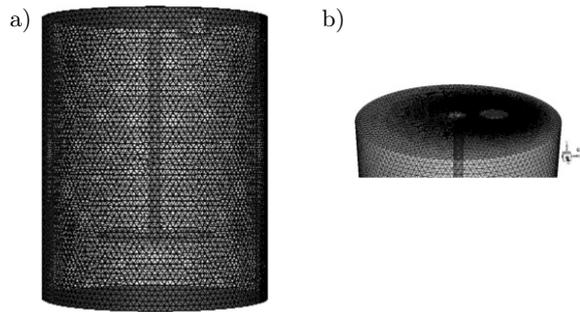


Fig. 5. Computational domain of model: a) computational domain of mechanical agitator modeling; b) computational domain of regions around ultrasonic generator.

Table 2. Data to create calculation structure.

Data	Value
Dimension of calculation region	3
Dimension of ultrasonic generation region	1
Dimension of agitation region	1
Size of agitator's impeller [mm]	200
Maximum mesh scale [mm]	10
Minimum mesh scale [mm]	0.5
Dimension of calculation region [mm × mm]	∅400 × 500
Number of phase	1

3.2.1.3. Calculation condition. Boundary conditions for simulation analysis were shown in Table 3.

Table 3. Calculation condition.

Classification	Data	Unit	Value
Boundary condition	Ultrasonic power	W	150, 400, 1000
	Frequency	Hz	20000
	Temperature	K	293
	Agitation intensity	rpm	150, 350, 550, 750
	Chemical composition	–	Water
Calculation model	RSM model		
	Abnormal problem		

3.2.1.4. Simulation results. The effect of agitation intensity on an ultrasonic power was analyzed using the CFD simulation while varying agitation intensity.

Firstly, the ultrasonic power level as a function of agitation intensity was investigated under ultrasonic power of 150 W. That result was shown in Figs. 6 and 7.

Based on the simulation result, the attenuation ratio which reflects the change of ultrasonic power, was investigated at 25 cm apart from the bottom of beaker on the axis of oscillator. That result was shown in Fig. 8.

As shown in Fig. 8, the ultrasonic power level decreases as agitation intensity increases. For the agitation speeds of 150 rpm, 350 rpm, 550 rpm, 750 rpm, the attenuation ratios were 55.1 %, 60.2 %, 68.0 %, 75.8 %, respectively.

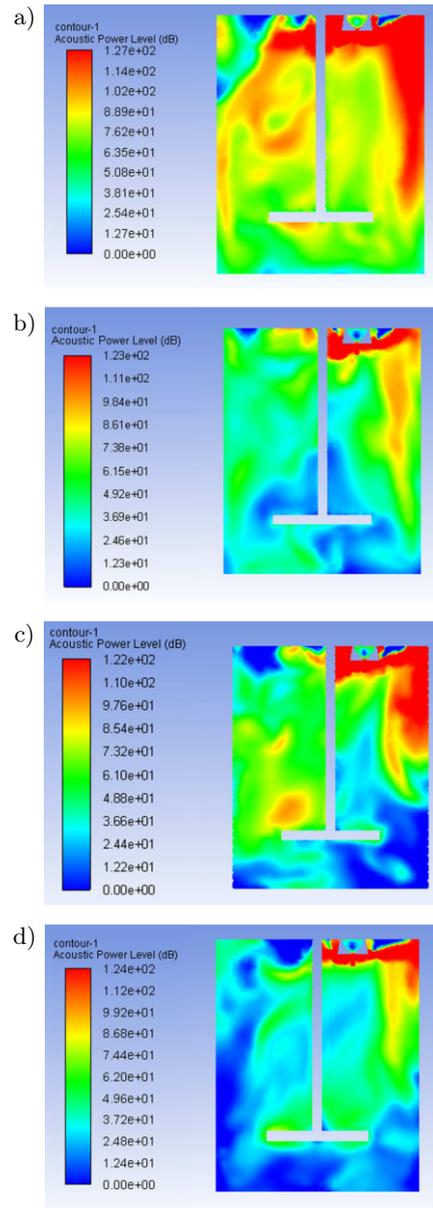


Fig. 6. Distribution character of acoustic power level according to the different agitation intensity under ultrasonic power of 150 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

Next, the ultrasonic power level as a function of agitation intensity was also investigated under an ultrasonic power of 400 W. That result was shown in Figs. 9 and 10.

Similarly, the attenuation ratio of an ultrasonic power as a function of agitation speed was shown in Fig. 11.

As shown in Fig. 11, at the agitation speeds of 150 rpm, 350 rpm, 550 rpm, 750 rpm, the attenuation ratios were 41.8 %, 47.3 %, 56.3 %, 63.5 %, respectively.

Finally, the ultrasonic power level as a function of agitation intensity was also investigated under an ul-

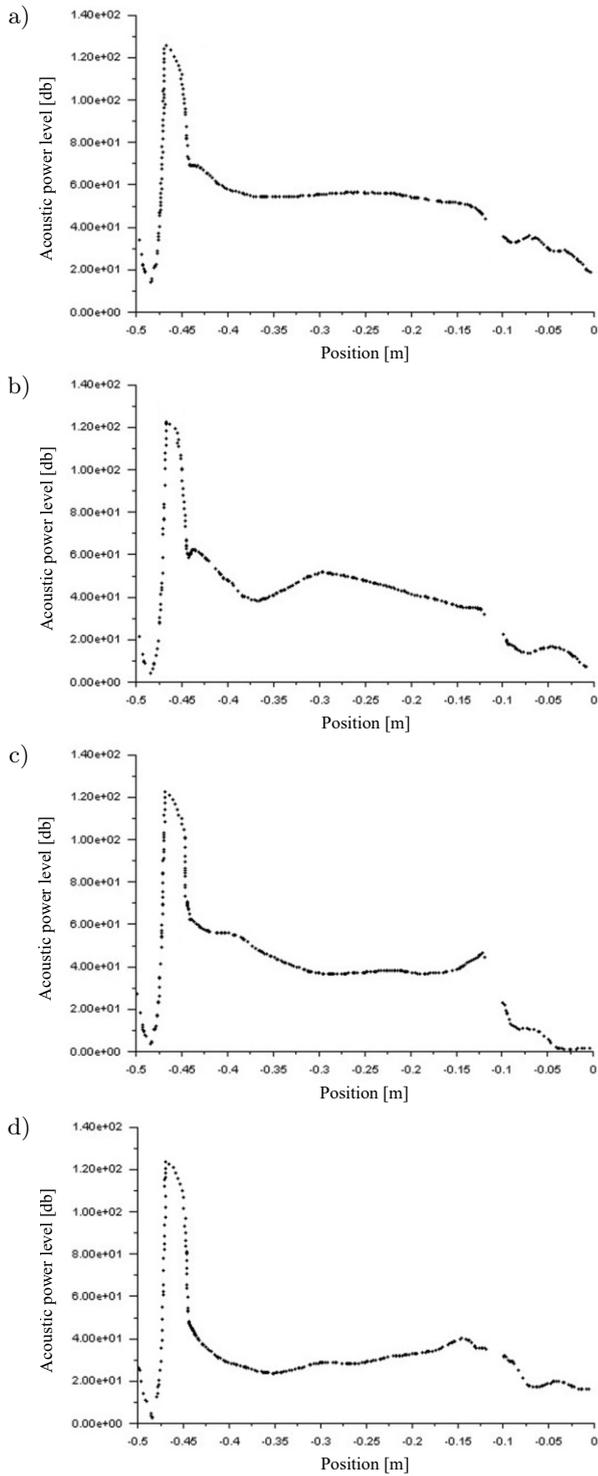


Fig. 7. Characteristic curve of acoustic power level according to the distance from agitator's bottom under ultrasonic power of 150 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

trasonic power of 1000 W. That result was shown in Figs. 12 and 13.

Similarly, the attenuation ratio of the ultrasonic power level as a function of agitation speed was shown in Fig. 14.

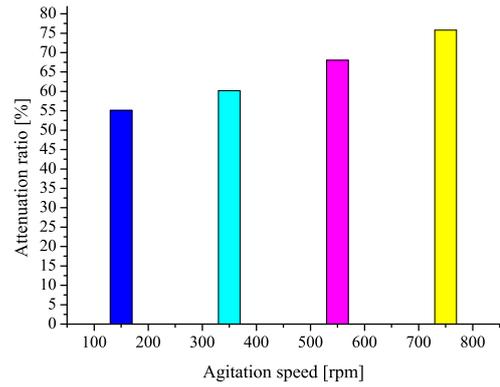


Fig. 8. Attenuation ratio of acoustic power level according to the different agitation intensity under ultrasonic power of 150 W.

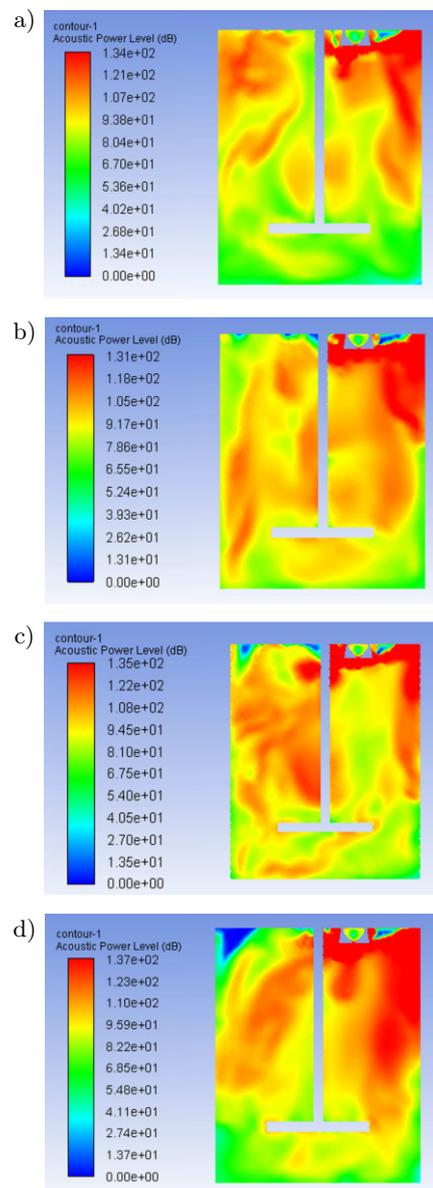


Fig. 9. Distribution character of acoustic power level according to the different agitation intensity under an ultrasonic power of 400 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

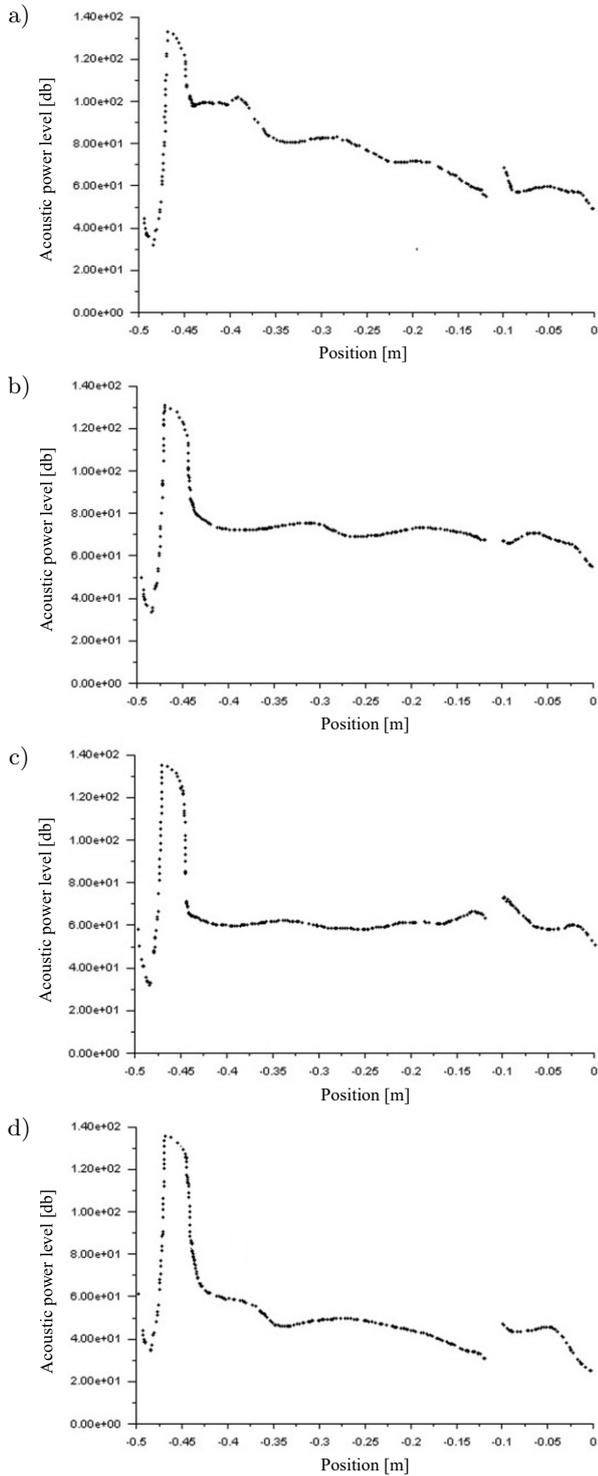


Fig. 10. Characteristic curve of acoustic power level according to the distance from agitator's bottom under ultrasonic power of 400 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

As shown in Fig. 14, for the agitation speeds of 150 rpm, 350 rpm, 550 rpm, 750 rpm, the attenuation ratios were 29.6 %, 32.0 %, 35.3 %, 35.1 %, respectively.

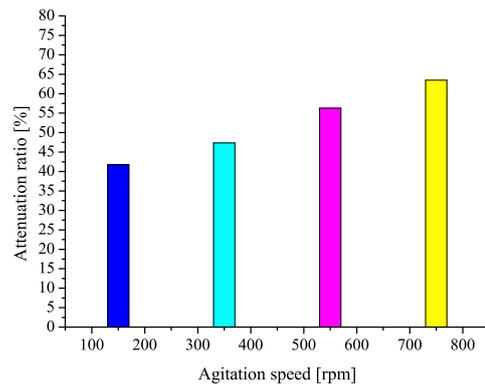


Fig. 11. Attenuation ratio of acoustic power level according to the different agitation intensity under ultrasonic power of 400 W.

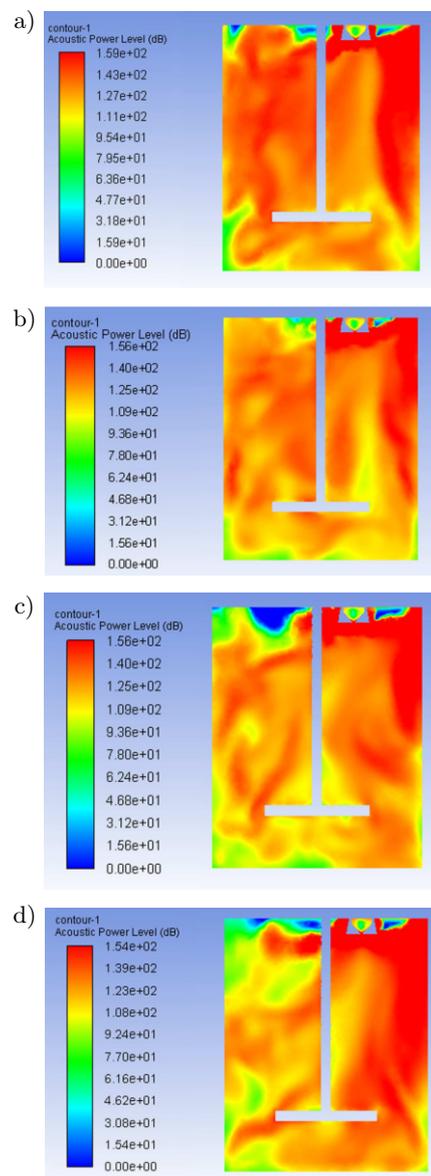


Fig. 12. Distribution character of acoustic power level according to the different agitation intensity under ultrasonic power of 1000 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

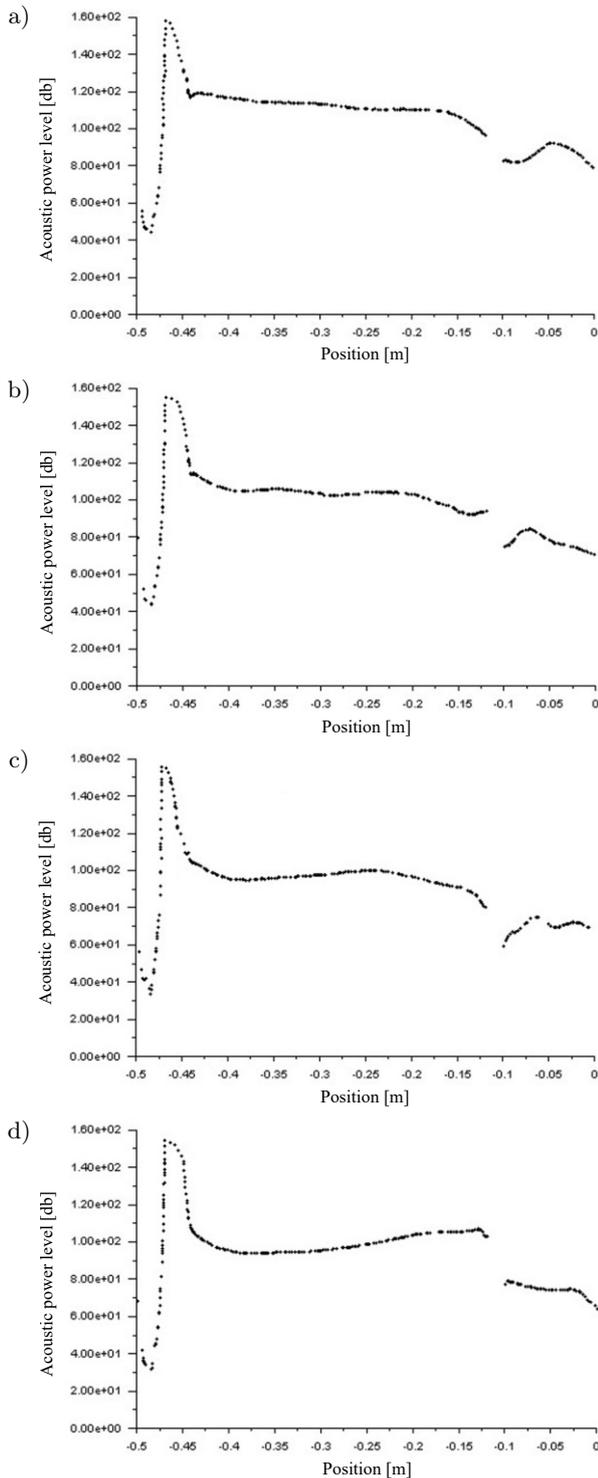


Fig. 13. Characteristic curve of acoustic power level according to the distance from agitator's bottom under ultrasonic power of 1000 W: a) 150 rpm; b) 350 rpm; c) 550 rpm; d) 750 rpm.

The simulation results show that an ultrasonic power decreases as agitation intensity increases. Besides, it also demonstrated that the detrimental effect of intensity of mechanical agitation decreased as an ultrasonic power increased.

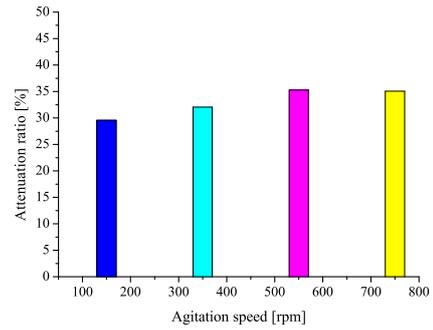


Fig. 14. Attenuation ratio of acoustic power level according to the different agitation intensity under ultrasonic power of 1000 W.

### 3.2.2. Experimental results

Based on the simulation results, experiments were carried out for proving them.

Relations between an ultrasonic power and agitation intensity on the alkaline pretreatment were shown in Figs. 15–17.

Figure 15 shows the effect of agitation intensity under an ultrasonic power of 150 W.

As shown in Fig. 15, when agitation intensity increased from 150 rpm to 350 rpm under an ultrasonic power of 150 W, arsenic extraction increased more and more. Also, the initial arsenic extraction rate was high at 550 rpm during 3 h after starting leaching under ultrasound, but it decreased again afterwards. The extraction of arsenic was higher at 350 rpm than that at

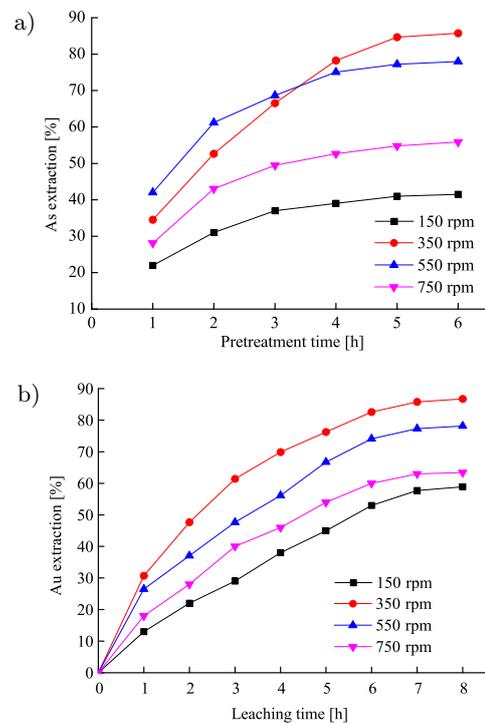


Fig. 15. Effect of agitation intensity under ultrasonic power of 150 W: a) arsenic extraction; b) gold extraction.

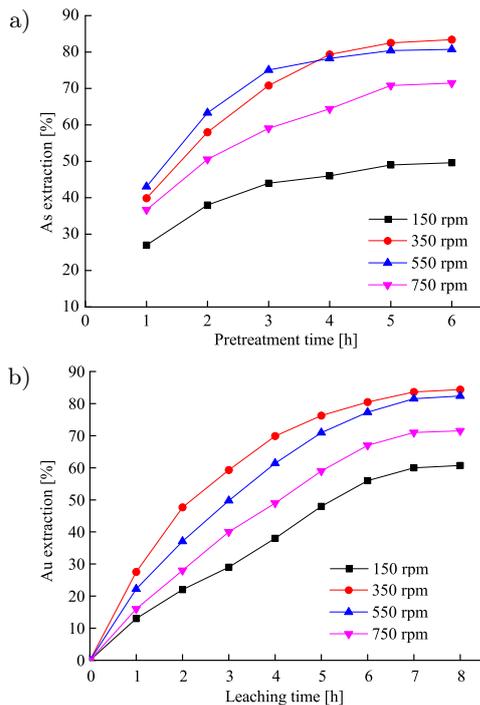


Fig. 16. Effect of agitation intensity under ultrasonic power of 400 W: a) arsenic extraction; b) gold extraction.

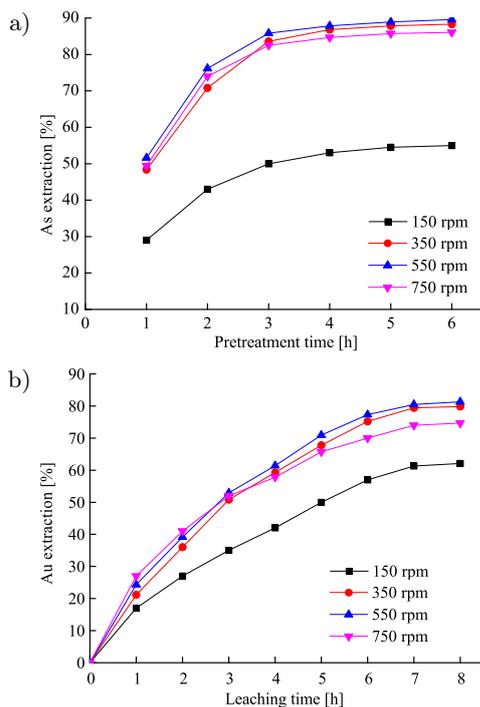


Fig. 17. Effect of agitation intensity under ultrasonic power of 1000 W: a) arsenic extraction; b) gold extraction.

750 rpm. This indicates that the maximum of extraction is available at 350 rpm under an ultrasonic power of 150 W. The extraction of gold in the pretreated sample might be proportional to the arsenic extraction.

As shown in Fig. 16, the extraction of arsenic increased as the agitation intensity increased from

150 rpm to 550 rpm under an ultrasonic power of 400 W. When the pretreated sample was leached for 5 h under ultrasound, the extraction of arsenic was 1 % lower at 550 rpm than the one at 350 rpm. This shows that the agitation intensity had influence on the extraction of arsenic unlike an ultrasonic power of 150 W.

The arsenic extraction was 67 % at agitation intensity of 750 rpm at leaching time of 6 h. It was higher than that at 150 W. But the extraction of arsenic and gold were also the highest at 350 rpm.

Figure 17 shows the effect of agitation intensity under ultrasonic power of 1000 W. As shown in Fig. 17, when agitation intensity changed from 150 rpm to 550 rpm, the arsenic extraction increased. The extraction of arsenic at 750 rpm was 28.3 %, 13.7 % higher than that at 150 W and 400 rpm. Therefore, it can be thought that the detrimental effect of agitation intensity decreased with the increase of an ultrasonic power. In this case, the extraction of gold was not proportional to the arsenic extraction. It could be attributed to the acceleration of a formation of  $\text{SiO}_2$  gel and adsorption of gold in pulp to it due to a high ultrasonic power (ZHANG *et al.*, 2016).

At the agitation speed of 150 rpm, an ultrasonic power was attenuated least from the simulation, whereas the extractions of arsenic and gold were lowest. It attributes to the settling of mineral particles onto the bottom due to the low agitation speed. But both results from experiment and simulation were consistent at different agitation speeds.

#### 4. Conclusions

Arsenic and gold extractions at 60 °C were 61 % and 67.5 %, respectively, without ultrasound, whereas 81 % and 81.9 % at 20 °C with ultrasound. This shows that ultrasound might be favourable for an alkaline pretreatment of gold bearing arsenopyrite.

Through CFD simulation, it was demonstrated that an ultrasonic power decreased as intensity of mechanical agitation increased, and the detrimental effect of mechanical stirring decreased as an ultrasonic power increased.

The experiment results showed that the gold extraction was the highest at 150 W and 350 rpm. These results showed that the proper ultrasonic power and agitation intensity should be selected on the pretreatment of gold-bearing arsenopyrite and the ultrasonic power could be interfered with agitation intensity. But, the interference decreased with the increase of an ultrasonic power.

#### Acknowledgments

The authors are grateful for the support of Professor Yong Chol Son from Kim Il Sung University.

## References

1. Ansys Fluent (2016), User's guide manual fluent Inc., Ansys Fluent 16.0, Pittsburgh, USA.
2. AWE S., SANDSTRÖM Å. (2010), Selective leaching of arsenic and antimony from a tetrahedrite rich complex sulphide concentrate using alkaline sulphide solution, *Minerals Engineering*, **23**(15): 1227–1236, doi: [10.1016/j.mineng.2010.08.018](https://doi.org/10.1016/j.mineng.2010.08.018).
3. BESE A.V. (2007), Effect of ultrasound on the dissolution of copper from copper converter slag by acid leaching, *Ultrasonics Sonochemistry*, **14**(6): 790–796, doi: [10.1016/j.ultsonch.2007.01.007](https://doi.org/10.1016/j.ultsonch.2007.01.007).
4. BHAKTA P., LANGHANS J., LEI K. (1989), *Alkaline oxidative leaching of gold-bearing arsenopyrite ores*, Report of Investigations 9258, Bureau of Mines, United States Department of the Interior, p. 1–11.
5. CHRYSOULIS S.L., McMULLEN J. (2005), Mineralogical investigation of gold ores, *Developments in Mineral Processing*, **15**: 21–71, doi: [10.1016/S0167-4528\(05\)15002-9](https://doi.org/10.1016/S0167-4528(05)15002-9).
6. CONTAMINE F., FAID F., WILHELM A.M., BERLAN J., DELMAS H. (1994), Chemical reactions under ultrasound: discrimination of chemical and physical effects, *Chemical Engineering Science*, **49**(24, Part 2): 5865–5873, doi: [10.1016/0009-2509\(94\)00297-5](https://doi.org/10.1016/0009-2509(94)00297-5).
7. CORKHILL C.L., VAUGHAN D.J. (2009), Arsenopyrite oxidation – A review, *Applied Geochemistry*, **24**(12): 2342–2361, doi: [10.1016/j.apgeochem.2009.09.008](https://doi.org/10.1016/j.apgeochem.2009.09.008).
8. DANG X. KE W., TANG C., LV J., ZHOU X., LIU C. (2016), Increasing leaching rate of gold cyanide of two-stage calcination generated from refractory ore containing arsenopyrite and pyrrhotite, *Rare Metals*, **35**(10): 804–810, doi: [10.1007/s12598-015-0470-0](https://doi.org/10.1007/s12598-015-0470-0).
9. DENG S., GU G. (2018), An electrochemical impedance spectroscopy study of arsenopyrite oxidation in the presence of *Sulfobacillus thermosulfidooxidans*, *Electrochimica Acta*, **287**: 106–114, doi: [10.1016/j.electacta.2018.08.051](https://doi.org/10.1016/j.electacta.2018.08.051).
10. HASHEMZADEHFINI M., FICERIOVÁ J., ABKHOSHK E., SHAHRAKI B. (2011), Effect of mechanical activation on thiosulphate leaching of gold from complex sulfide concentrate, *Transactions of Nonferrous Metals Society of China*, **21**(12): 2744–2751, doi: [10.1016/S1003-6326\(11\)61118-7](https://doi.org/10.1016/S1003-6326(11)61118-7).
11. KOJIMA Y., ASAKURA Y., SUGIYAMA G., KODA S. (2010), The effects of ultrasonic flow and mechanical flow on the sonochemical efficiency in a rectangular sonochemical reactor, *Ultrasonics Sonochemistry*, **17**(6): 978–984, doi: [10.1016/j.ultsonch.2009.11.020](https://doi.org/10.1016/j.ultsonch.2009.11.020).
12. LA BROOY S.R., LINGE H.G., WALKER G.S. (1994), Review of gold extraction from ores, *Minerals Engineering*, **7**(10): 1213–1241, doi: [10.1016/0892-6875\(94\)90114-7](https://doi.org/10.1016/0892-6875(94)90114-7).
13. MENG Y., WU M., SU S., WANG L. (2003), Intensified alkaline leaching pretreatment of refractory gold concentrate at common temperature and pressure [in Chinese], *Transactions of Nonferrous Metals Society of China*, **13**(2): 426–430.
14. MEROUANI S., HAMDAR O., REZGUI Y., GUEMINI M. (2014), Energy analysis during ultrasonic bubble oscillations: Relationship between bubble energy and sonochemical parameters, *Ultrasonics*, **54**(1): 227–232, doi: [10.1016/j.ultras.2013.04.014](https://doi.org/10.1016/j.ultras.2013.04.014).
15. MESA ESPITIA S.L., LAPIDUS G.T. (2015), Pretreatment of a refractory arsenopyritic gold ore using hydroxyl ion, *Hydrometallurgy*, **153**: 106–113, doi: [10.1016/j.hydromet.2015.02.013](https://doi.org/10.1016/j.hydromet.2015.02.013).
16. MIKHLIN Y.L., ROMANCHENKO A.S., ASANOV I.P. (2006), Oxidation of arsenopyrite and deposition of gold on the oxidized surfaces: A scanning probe microscopy, tunneling spectroscopy and XPS study, *Geochimica et Cosmochimica Acta*, **70**(19): 4874–4888, doi: [10.1016/j.gca.2006.07.021](https://doi.org/10.1016/j.gca.2006.07.021).
17. MÜLLER S., FISCHPER M., MOTTYLL S., SKODA R., HUSSONG J. (2014), Analysis of the cavitating flow induced by an ultrasonic horn – Experimental investigation on the influence of actuation phase, amplitude and geometrical boundary conditions, *EPJ Web of Conferences*, **67**: 02079, doi: [10.1051/epjconf/20146702079](https://doi.org/10.1051/epjconf/20146702079).
18. NAN X. CAI X., KONG J. (2014) Pretreatment process on refractory gold ores with As, *ISIJ International*, **54**(3): 543–547, doi: [10.2355/isijinternational.54.543](https://doi.org/10.2355/isijinternational.54.543).
19. PENN R., YEAGER E., HOVORKA F. (1959), The effect of ultrasonic waves on the dissolution of Nickel, *Journal of the Acoustical Society of America*, **31**: 1372–1376.
20. SAJJADI B., RAMAN A.A.A., IBRAHIM S. (2015), A comparative fluid flow characterisation in a low frequency/high power sonoreactor and mechanical stirred vessel, *Ultrasonics Sonochemistry*, **27**: 359–373, doi: [10.1016/j.ultsonch.2015.04.034](https://doi.org/10.1016/j.ultsonch.2015.04.034).
21. ŚLACZKA A.St. (1986), Effect of ultrasound on ammonium leaching of zinc from galmei ore, *Ultrasonics*, **24**(1): 53–55, doi: [10.1016/0041-624X\(86\)90075-2](https://doi.org/10.1016/0041-624X(86)90075-2).
22. TAYLAN N., GÜRBÜZ H., BULUTCU A.N. (2007), Effects of ultrasound on the reaction step of boric acid production process from colemanite, *Ultrasonics Sonochemistry*, **14**(5): 633–638, doi: [10.1016/j.ultsonch.2006.11.001](https://doi.org/10.1016/j.ultsonch.2006.11.001).
23. TEKIN T. (2002), Use of ultrasound in the dissolution kinetics of phosphate rock in HCl, *Hydrometallurgy*, **64**(3): 187–192, doi: [10.1016/S0304-386X\(02\)00040-3](https://doi.org/10.1016/S0304-386X(02)00040-3).
24. WANG S., CUI W., ZHANG G., ZHANG L., PENG J. (2017), Ultra fast ultrasound-assisted decopperization from copper anode slime, *Ultrasonics Sonochemistry*, **36**: 20–26, doi: [10.1016/j.ultsonch.2016.11.013](https://doi.org/10.1016/j.ultsonch.2016.11.013).
25. ZHANG G., WANG S., ZHANG L., PENG J. (2016), Ultrasound-intensified leaching of gold from a refractory ore, *ISIJ International*, **56**(4): 714–718, doi: [10.2355/isijinternational.ISIJINT-2015-476](https://doi.org/10.2355/isijinternational.ISIJINT-2015-476).
26. ZHU P. ZHANG X., LI K., QIAN G., ZHOU M. (2012), Kinetics of leaching refractory gold ores by ultrasonic-assisted electro-chlorination, *International Journal of Minerals, Metallurgy, and Materials*, **19**(6): 473–477, doi: [10.1007/S12613-012-0582-6](https://doi.org/10.1007/S12613-012-0582-6).

## Research Paper

Mutually Orthogonal Complementary Golay Coded Sequences:  
An In-vivo StudyIhor TROTS\*, Jurij TASINKIEWICZ<sup>ID</sup>, Andrzej NOWICKI<sup>ID</sup>*Department of Ultrasound, Institute of Fundamental Technological Research  
Polish Academy of Sciences  
Warsaw, Poland*\*Corresponding Author e-mail: [igortr@ippt.pan.pl](mailto:igortr@ippt.pan.pl)*(received April 22, 2024; accepted July 26, 2024; published online August 19, 2024)*

Fast and high-quality ultrasound imaging allows to increase the effectiveness of detecting tissue changes at the initial stage of disease. The aim of the study was to assess the quality of ultrasound imaging using mutually orthogonal, complementary Golay coded sequences (MOCGCS). Two 16-bits MOCGCS sets were implemented in the Verasonics Vantage™ scanner. Echoes from a perfect reflector, a custom-made nylon wire phantom, a tissue-mimicking phantom, and in-vivo scans of abdominal aorta and common carotid artery were recorded. Three parameters of the detected MOCGCS echoes: signal-to-noise ratio (SNR), side-lobe level (SLL), and axial resolution were evaluated and compared to the same parameters of the echoes recorded using standard complementary Golay sequences (CGS) and a short, one sine cycle pulse. The results revealed that MOCGCS transmission maintained comparable echo quality metrics (SNR, SLL, and axial resolution) compared to CGS and short pulses. Notably, both MOCGCS and CGS offered similar SNR improvements (5 dB–9 dB) in comparison to the short pulse for wires placed at depths up to 8 cm. Analysis of axial resolution, estimated at the full width at half maximum level, revealed near-identical values for all transmitted signals (0.17  $\mu$ s for MOCGCS, 0.16  $\mu$ s for CGS, and 0.18  $\mu$ s for short pulse). MOCGCS implementation in ultrasound imaging offers the potential to significantly reduce image reconstruction time while maintaining image quality comparable to CGS sequences. In the experimental study we have shown that MOCGCS offers advantages over conventional CGS by enabling two times faster data acquisition and image reconstruction without compromising image quality.

**Keywords:** coded excitation; Golay codes; synthetic aperture; ultrasound imaging.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Modern ultrasound diagnostics require exceptional image quality, characterized by deep penetration depth, high signal-to-noise ratio (SNR), and excellent resolution. However, achieving this trade-off between penetration and resolution necessitates alternative approaches. Wide-band coded sequences, combined with echo compression techniques, offer a promising solution (NOWICKI *et al.*, 2003). These sequences allow for increased SNR and deeper visualization without exceeding power limitations, even enabling the use of higher frequencies. Among the various coded sequences explored in ultrasound, Golay codes have emerged as frontrunners (GOLAY, 1961). Their unique ability to suppress side lobes in the transmitted sig-

nal makes them particularly attractive. This property minimizes interference and improves overall signal quality.

While computational capabilities play a role, a crucial factor limiting the frame rate (images per second) in ultrasound imaging is the speed of data acquisition. This speed depends on the sound wave velocity in the target tissue, imaging depth, and the number of scanned lines.

A promising approach to overcome this limitation involves transmitting multiple scanning signals simultaneously. However, separating the resulting echoes requires the transmitted signals to be orthogonal, minimizing interference. Fortunately, specific pairs of Golay coded sequences, known as MOCGCS, possess this desired property (BAE, 2003).

MOCGCS have attracted significant research interest in recent years due to their potential applications in various fields, including ultrasound (GRAN, JENSEN, 2006; DEMI *et al.*, 2013; TROTS *et al.*, 2004; 2011; WU *et al.*, 2020; TIAN *et al.*, 2021). Their key advantage lies in the orthogonality property, enabling significant increases in data acquisition rate and consequently, faster image reconstruction (frame rate). While prior studies explored using MOCGCS for 2D B-mode image reconstruction by comparing simultaneous transmission with short pulse approaches (CHIAO, THOMAS, 2000; BAE *et al.*, 2002; KIM, SONG, 2003; MISARIDIS, JENSEN, 2005; PENG *et al.*, 2006; RAMALLI *et al.*, 2015; ZHAO, LUO, 2018; KUMRU, KOYMEN, 2018), a critical gap remains. No research has conducted an in-depth analysis of how MOCGCS echoes behave in real-world scenarios, particularly regarding the effectiveness of matched filtering for echo separation.

This paper addresses this gap by experimentally evaluating the feasibility of MOCGCS in ultrasound diagnostics. We analyze and compare echoes acquired using MOCGCS with those obtained from conventional complementary Golay sequences (CGS) and short pulses. Our analysis focuses on echoes from various targets, including a brass plate, a custom-made nylon wire phantom, and a tissue-mimicking phantom. Specifically, we compare signal parameters such as pulse duration, side-lobe level (SLL), and SNR for each excitation signal. Matched filtering is employed for MOCGCS and CGS emissions to separate echoes and suppress side-lobes. These parameters include pulse duration at the  $-6$  dB and  $-20$  dB levels, signal SLL, and SNR. In the case of MOCGCS and CGS emissions, matched filters were used to separate the echoes and suppress the signal side-lobes. The experimental setup utilizes a commercially available ultrasound system (Verasonics Vantage™) equipped with a 128-element linear array transducer.

## 2. Materials and methods

The experiment utilized a Verasonics Vantage™ research ultrasound system equipped with a 128-element linear array transducer (L7-4), operating at a center frequency of 5.2 MHz with a bandwidth of 60 %. Three types of signals were generated for transmission:

- a conventional 16-bit complementary Golay sequence (CGS) pair,
- two 16-bit, mutually orthogonal complementary Golay code sets (MOCGCS),
- a one sine cycle pulse (short pulse) with a nominal frequency of 5.2 MHz.

Ultrasound echoes were collected using three different targets immersed in a water tank:

- perfect reflector: this initial measurement with a perfect reflector allowed us to estimate the

compressed pulse duration for each signal type (MOCGCS, CGS, and short pulse),

- custom-made wire phantom: a phantom consisting of fine nylon wires (diameter 0.25 mm) arranged vertically with 20 mm axial spacing was used to compare the side-lobe level (SLL) for different excitation signals,
- tissue-mimicking phantom: this phantom with an attenuation of 0.5 dB/(MHz·cm) facilitated the comparison of signal-to-noise ratio (SNR) across different signals.

The basic algorithm of MOCGCS compression is given below. Consider two  $L$ -bit long codes  $G_{1i}$  and  $G_{2i}$ ,  $i = 1, 2, \dots, M$ , that are complementary pairs ( $G_{11}$  and  $G_{12}$  is a complementary pair #1, and  $G_{21}$  and  $G_{22}$  is a complementary pair #2) and obey the following condition (TROTS, 2015; TROTS *et al.*, 2015):

$$R_{G_{11}}(n) + R_{G_{12}}(n) = \begin{cases} 2L, & n = 0, \\ 0, & n \neq 0, \end{cases} \quad (1)$$

$$R_{G_{21}}(n) + R_{G_{22}}(n) = \begin{cases} 2L, & n = 0, \\ 0, & n \neq 0, \end{cases}$$

where the autocorrelation function  $R$  of the coded sequence in Eq. (1) is defined as follows (MISARIDIS, 2001):

$$R(n) = \begin{cases} \sum_{k=0}^{L-1-n} C(k)C(k-n), & n = 0, \dots, L-1, \\ R(-n), & n = -(L-1), \dots, -1, \end{cases} \quad (2)$$

where  $C(k)$ ,  $k = 0, \dots, L-1$  denote the coded sequence  $G_{1i}$ ,  $G_{2i}$ ,  $i = 1, 2$  of the length  $L$ .

Two pairs of Golay codes CGC,  $G_{1i}$  and  $G_{2i}$ , are said to be mutually orthogonal, or MOCGCS pairs, if the sum of their cross-correlation functions is zero:

$$\sum_{i,j=1}^2 R_{G_{1i}G_{2j}} = 0. \quad (3)$$

The cross-correlation functions in Eq. (3) are defined as follows (TSENG, LIU, 1972):

$$R_{G_{1i}G_{2j}}(n) = \begin{cases} \sum_{k=0}^{L-1-n} G_{1i}(k)G_{2j}(k-n), & n = 0, \dots, L-1, \\ \sum_{k=0}^{L-1-n} G_{1i}(k-n)G_{2j}(k), & n = -(L-1), \dots, -1. \end{cases} \quad (4)$$

Thus, the CGC pairs  $G_{1i}$  and  $G_{2i}$ ,  $i = 1, 2$ , form a MOCGCS set if they satisfy Eqs. (1) and (3) simultaneously. The key benefit of MOCGCS, as defined by Eq. (3), lies in their ability to enable the simultaneous transmission of two CGS pairs without interference during reception, allowing for the separation of individual echoes.

In this work, two sets of MOCGCS – orthogonal sets of  $\{G_1, G_2\}$  – are used:

$$G_{11} = [-1 \ -1 \ -11 \ -1 \ -11 \ -1 \ -1 \ -1 \ -1 \ 1 \ 1 \ 1 \ -1 \ 1],$$

$$G_{12} = [-1 \ 1 \ -1 \ -1 \ -1 \ 1 \ 1 \ 1 \ -1 \ 1 \ -1 \ -1 \ -1 \ -1],$$

$$G_{21} = [-1 \ -1 \ -11 \ -1 \ -11 \ -1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ 1 \ -1],$$

$$G_{22} = [-1 \ 1 \ -1 \ -1 \ -1 \ 1 \ 1 \ 1 \ 1 \ -1 \ 1 \ 1 \ 1 \ 1].$$

Assuming orthogonal Golay sequences of length  $L = 16$  bits, the two code sequences  $\{G_{1i}, G_{2i}\}$  are transmitted by two transducer elements (namely, #1 and #2). First, the orthogonal signals  $\{G_{11}, G_{21}\}$  are transmitted. Then, the corresponding echoes are detected and stored for further processing. Next, the process is repeated for the second pair of orthogonal signals  $\{G_{12}, G_{22}\}$ . The detected signals, which are a superposition of echoes corresponding to  $\{G_{1i}, G_{2i}\}$ , can be compressed by summing the correlation functions of each received sequence with the sequence transmitted by the same transducer element. For example, to recover the echo for transducer element #1 (i.e., the pair  $G_{1i}$  that was transmitted by transducer element #1), for the first and second transmissions, one should compute the sum of cross-correlation functions  $P_{G_1}$  of the received signals  $R_{S_i G_{1i}}$  with the corresponding transmitted codes  $G_{1i}$ ,  $i = 1, 2$ , using Eq. (4) as follows:

$$P_{G_1} = \sum_{i=1}^2 R_{S_i G_{1i}}. \quad (5)$$

Figure 1 illustrates the data acquisition scheme for MOCGCS transmission. Two sets of MOCGCS codes, denoted as  $G_{1i}$  and  $G_{2i}$ , were transmitted simultaneously using two adjacent elements of the linear array transducer. During the first transmission, element #1

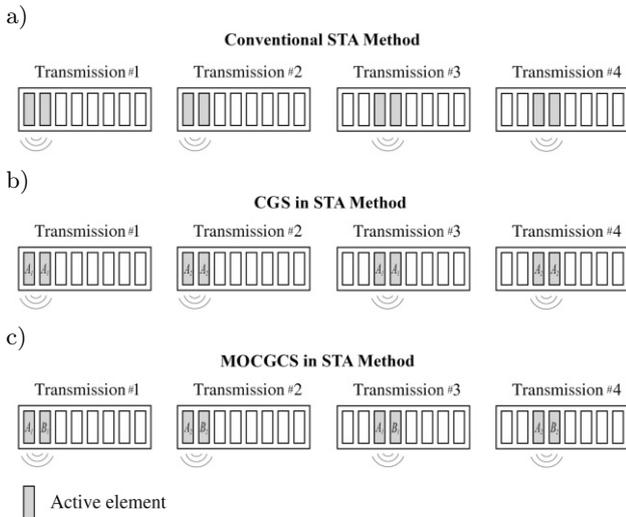


Fig. 1. Data acquisition scheme for: a) short pulse; b) CGS; c) MOCGCS transmission in a 2-element synthetic transmit aperture method with a 2-elements stride. In all three configurations, all elements of the array are active.

transmitted code  $G_{11}$  while element #2 transmitted code  $G_{21}$ . This pattern was reversed for the second transmission, with element #1 transmitting  $G_{12}$  and element #2 transmitting  $G_{22}$ . In both transmissions, echoes were acquired by all transducer elements simultaneously. For comparison, CGS transmission involved transmitting only one complementary pair ( $G_{11}$  and  $G_{12}$ ) by the same elements (#1 and #2) during two consecutive transmissions. Similarly, the short pulse was transmitted and received by elements #1 and #2, implementing a synthetic transmit aperture (STA) with a 2-element sub-aperture in transmit mode and a 2-element stride. This STA scheme was used throughout the experimental measurements (detailed in Sec. 3).

Raw data were collected at a sampling rate of 20.8 MHz and stored for further processing. The processing methods differed based on the transmitted signal type:

- MOCGCS echoes: processing followed the method described in (TROTS *et al.*, 2022),
- CGS echoes: a conventional matched filtering technique (MISARIDIS, 2001) was employed to compress the signals and suppress side-lobes.

All signal processing algorithms were implemented in MATLAB<sup>®</sup>.

### 3. Results

Figure 2 presents the radio frequency (RF) signals detected from a perfect reflector (brass plate) for different transmitted signals at a 5.2 MHz center frequency:

- Fig. 2a: echo for one cycle short pulse,
- Figs. 2b and 2c: MOCGCS echoes, representing the superposition of 16-bit codes  $G_{11}$  and  $G_{21}$ , and  $G_{12}$  and  $G_{22}$ , respectively, transmitted simultaneously by elements #1 and #2,
- Figs. 2d and 2e: CGS echoes resulting from single transmissions of codes  $G_{11}$  and  $G_{12}$  by elements #1 and #2. Elements #1 and #2 were used for transmission, and element #1 received the echoes. All coded sequences were 16 bits long, and the overall bandwidth was about 60 %.

Figure 3 depicts the envelopes of the signals reflected from a brass plate (corresponding RF echoes shown in Fig. 2). For MOCGCS and CGS signals, matched filtering was applied to compress the RF echoes before envelope detection. Figure 3 shows only the  $G_1$  code envelope from the MOCGCS transmission. The same code  $G_1 = \{G_{11}, G_{12}\}$  was used as the CGS example in Figs. 2d and 2e.

The half-maximum durations (measured at  $-6$  dB) of the envelopes for both MOCGCS and CGS were nearly identical, at approximately  $0.17 \mu\text{s}$  and  $0.16 \mu\text{s}$ , respectively. The corresponding value for the short

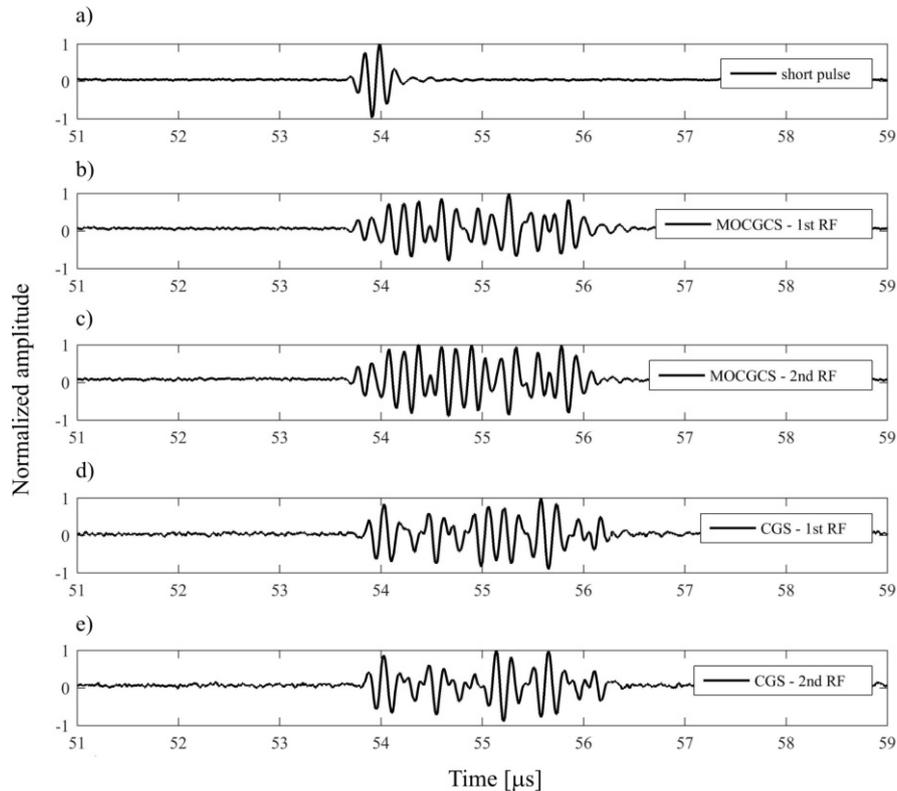


Fig. 2. RF echo signal from a perfect reflector for: a) short pulse transmission; b) MOCGCS echo being a superposition of codes  $G_{11}$  and  $G_{21}$ ; c) MOCGCS echo being a superposition of codes  $G_{12}$  and  $G_{22}$ ; d) CGS echo resulting from  $G_{11}$  sequence transmitted; e) CGS echo resulting from  $G_{12}$  sequence transmitted. Elements #1 and #2 were used in TX mode; the echoes were detected by element #1.

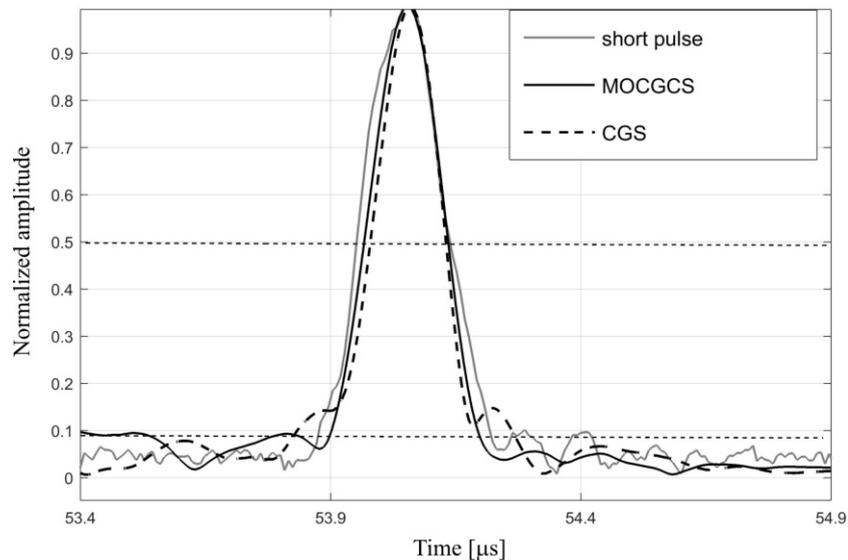


Fig. 3. Envelopes of the processed RF echoes obtained from the perfect reflector using the short pulse, the MOCGCS and the CGS transmitted signals.

pulse was slightly higher at  $0.18 \mu\text{s}$ . Translating these time durations to spatial distance using the ultrasound speed provides axial resolutions of approximately  $0.52 \text{ mm}$  and  $0.49 \text{ mm}$  for MOCGCS and CGS, respectively. The short pulse resulted in a slightly lower resolution of  $0.55 \text{ mm}$ .

The envelope durations at the  $-20 \text{ dB}$  level were  $0.30 \mu\text{s}$  for MOCGCS,  $0.36 \mu\text{s}$  for CGS, and  $0.35 \mu\text{s}$  for the short pulse transmission. These durations translate to spatial resolutions of  $0.92 \text{ mm}$  for MOCGCS,  $1.11 \text{ mm}$  for CGS, and  $1.08 \text{ mm}$  for the short pulse.

In Fig. 4, the envelopes of the RF signals collected from the nylon wire phantom are shown as the functions of depth.

In the case of MOCGCS and CGS, the RF signals were compressed using a match filtering technique prior the envelope detection.

In Fig. 5, the SLL versus depth is shown for different excitation signals. A comparison of the SLL for MOCGCS and CGS shows the efficiency of matched filtering (i.e., the extraction of echoes corresponding to different MOCGCS pairs from the received signal) at various depths. The SLL for the short pulse was also estimated for comparison.

As shown in Fig. 5, the SLL for all types of transmitted signals increases with depth.

In Fig. 6, the SNR determined from the beamformed RF signal along scanline #65, which coincides with the central column of point scatterers of the tissue-mimicking phantom (see Fig. 7), is shown for the short pulse (Fig. 6a), the MOCGCS (Fig. 6b), and the conventional CGS (Fig. 6c) transmission. The SNR was determined as the ratio of the average signal power to the RMS noise power at different depths. The average signal power was estimated for echo samples in 1.5 mm windows corresponding to 5 wavelengths (the wavelength is approximately 0.3 mm). The center of

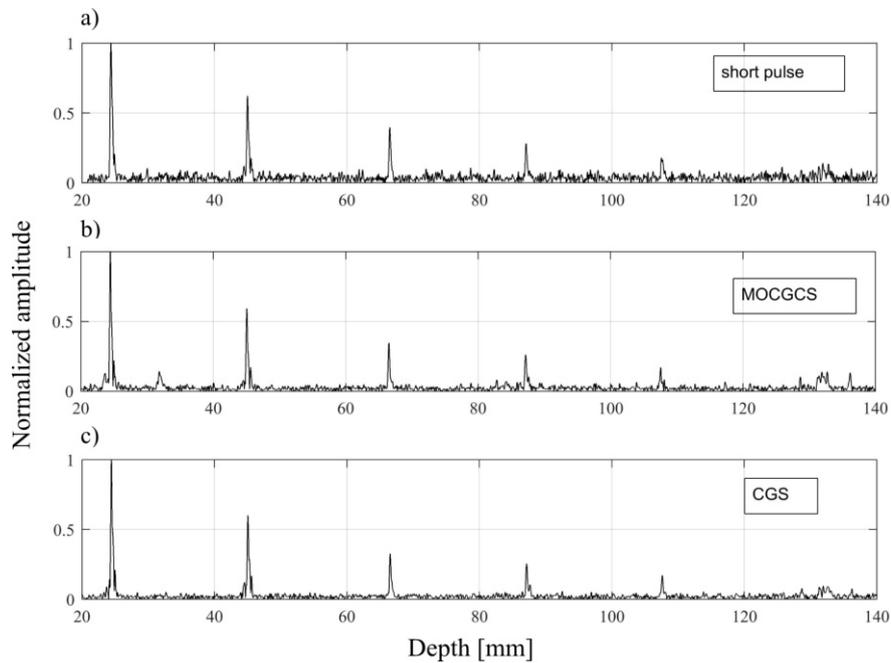


Fig. 4. Envelopes of the processed RF echoes from the nylon wire phantom obtained using: a) the short pulse; b) MOCGCS; c) CGS transmitted signals.

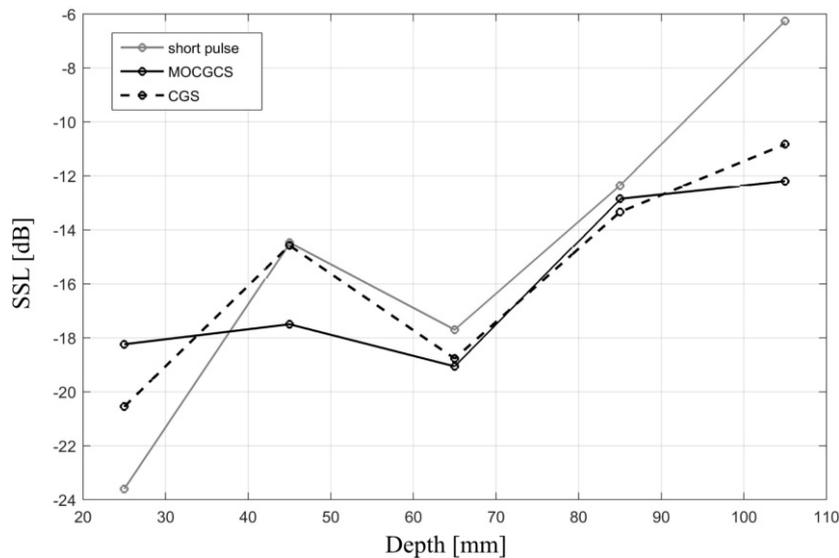


Fig. 5. The SLL vs. depth for the processed signals obtained from the wire phantom (see Fig. 4).

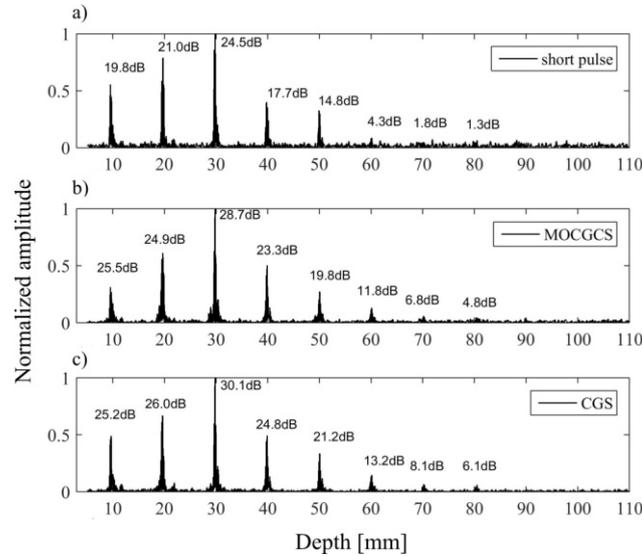


Fig. 6. The SNR for the beamformed RF signal along scanline #65 obtained from the tissue-mimicking phantom using: a) the short pulse; b) MOCGCS; c) conventional CGS transmission. The transmitted voltage was 1.6 V. The absolute values of the beamformed signals are shown.

each window coincided with the peak of echo amplitude at a given depth (corresponding to the wire's location). The RMS noise power was computed for echo samples in a 10 mm (about 34 wavelengths) window starting from 100 mm where only noise is present and the signal component can be neglected (see Fig. 6).

To demonstrate the advantage of the encoded transmission in detecting deeper located wires due to better SNR, the peak-to-peak transmitted voltage was chosen to be 1.6 V, and the measurements were performed without amplifying the detected echoes. As shown in Fig. 6, there is only a slight decrease in the SNR values (not exceeding 1.5 dB across the entire depth range) for MOCGCS compared to CGS transmitted signals. In contrast, the corresponding decrease in the SNR value for the short pulse signal varied from about 5 dB to 8.9 dB across the entire depth range.

In Fig. 7, B-mode images of the tissue-mimicking phantom obtained using the STA image reconstruction method with MOCGCS excitation signals, the conventional STA method with CGS, and short pulse excitation are shown.

The extension of the point spread function for MOCGCS and CGS signals in the axial direction can be observed in Fig. 6. This is due to the SLL increase resulting from the compression of corresponding echo signals. The increase in SLL is primarily influenced by the transducer bandwidth and the non-ideal shape of the generated transmitter sequences (see Fig. 2b through Fig. 2e) as well as the scattering properties of the phantom material, which are similar to those of the real biological tissue.

Finally, the B-mode images of the tissue-mimicking phantom obtained using the STA image reconstruc-

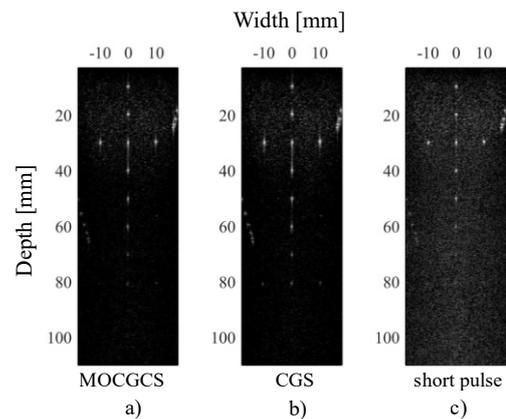


Fig. 7. B-mode images of the tissue-mimicking phantom obtained using the STA image reconstruction method with: a) MOCGCS excitation signals; b) the conventional STA method with CGS; c) short pulse excitation. The transmitted voltage was 1.6 V. All images are displayed on a logarithmic scale with a 40 dB dynamic range.

tion method with MOCGCS excitation signals, as discussed in (TROTS *et al.*, 2022), and the conventional STA method with CGS and short pulse excitation signals, shown in Fig. 7.

Figures 8 and 9 show the 2D B-mode reconstructed in-vivo images of the abdominal aorta and common carotid artery from one of the report's authors obtained using MOCGCS, CGS, and short pulse for comparison.

The results of tests for improving sensitivity, increasing the amplitude of received echoes for the three tested transmitter systems, are illustrated in Fig. 10. The received RF echoes were recorded by a single element of the linear array for each transmitted sequence.

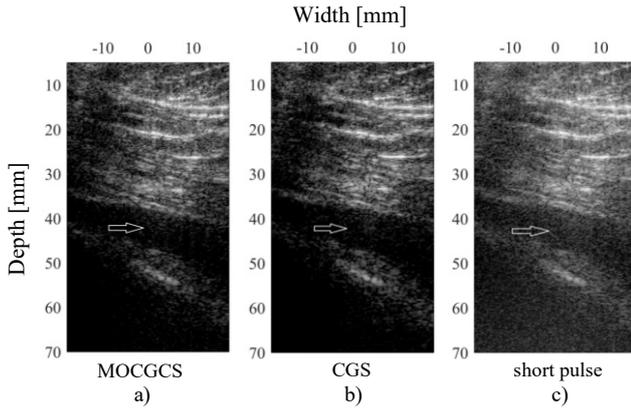


Fig. 8. Comparison of the 2D B-mode images of the abdominal aorta (arrows point to the center of the aorta) obtained in-vivo using: a) MOCGCS; b) CGS; c) short pulse transmission. All images are displayed on a logarithmic scale with a 40 dB dynamic range. The transmitted voltage was 10 V.

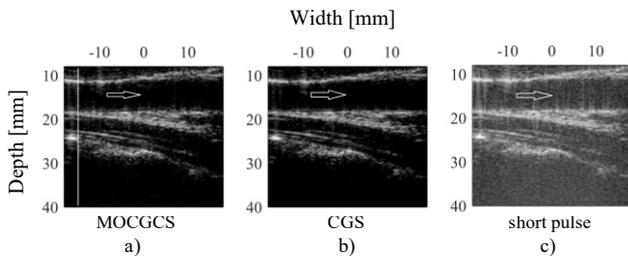


Fig. 9. Comparison of the 2D B-mode images of the common carotid artery (arrows point to the center of the artery) obtained in-vivo: a) using MOCGCS; b) CGS; c) short pulse transmission. All images are displayed on a logarithmic scale with a 40 dB dynamic range. The transmitted voltage was 1.6 V. The vertical line in the leftmost panel indicates the position of the RF echoes shown in Fig. 10.

The recorded echoes were normalized to the maximum value of the CGS signal to show the range of the spanned amplitudes.

The first peaks of the RF signals in Fig. 10 correspond to the signal reflection from the upper wall of the carotid artery at the depth of 10 mm. The amplitude of echoes obtained using the short pulse were at the level of  $-14.3$  dB compared to the coded CGS transmission. In the case of MOCGCS transmission, the decrease in amplitude was about  $-1.6$  dB.

The 2D ultrasound images obtained clearly show that the image resolution is nearly the same in all cases. The reconstruction time for MOCGCS is half that of CGS and the same as that of the short pulse. Using MOCGCS of 4th order or higher can further reduce the reconstruction time, which allows for a proportional increase in frame rate (HUANG, 2005). Specifically, in the case of 4th order MOCGCS, a set of four coded sequences is transmitted by four transducers during one transmission, and as a result, four image lines can be constructed after two transmissions.

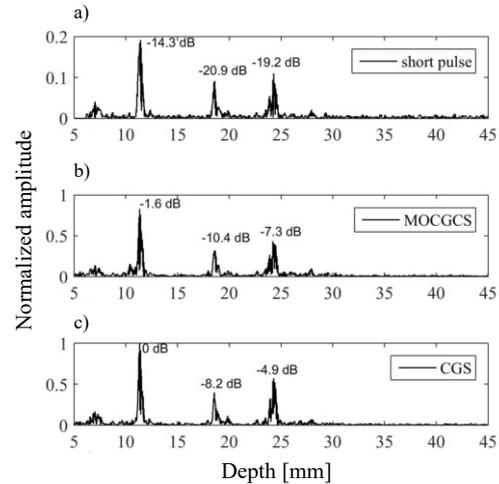


Fig. 10. Normalized magnitudes of the processed RF signals from the human common carotid artery depicted in Fig. 9 for: a) short pulse; b) MOCGCS; c) CGS transmitted signals. The RFs were recorded with element #16 (denoted by the vertical line in the leftmost panel in Fig. 9) during transmission #8 of the STA data acquisition.

## 4. Conclusions

This study investigated the feasibility of using a set of two complementary Golay sequences (MOCGCS) in ultrasound diagnostics. MOCGCS offers advantages over conventional CGS by enabling faster data acquisition and image reconstruction without compromising image quality. To evaluate MOCGCS performance, the parameters of recorded RF signals acquired from an ideal reflector using MOCGCS, CGS, and short pulse excitation were compared. These parameters included axial resolution that was estimated using the half-maximum duration ( $-6$  dB level) of the signal envelopes (Fig. 4). MOCGCS exhibited minimal (5.8 %) degradation in axial resolution compared to CGS. Despite a slight increase in spatial duration (0.52 mm vs. 0.49 mm for CGS), MOCGCS maintained performance close to CGS. Short pulse excitation displayed a more significant (11 %) decrease in axial resolution compared to CGS, with a spatial duration of 0.55 mm.

The SLL of echoes acquired from a custom-made wire phantom was estimated from their envelopes (Fig. 4) and illustrated against depth (Fig. 5). As expected, a consistent trend of increasing SLL with depth is observed for all signals. The observed increase in SLL with depth can be attributed to growing attenuation of the signal as the two-way propagation path increases and frequency-dependent attenuation acting as a low-pass filter, suppressing higher frequency components of the temporal spectrum. This leads to broadening of the signal in time domain, and the SLL increases. Figure 5 shows that MOCGCS and CGS received signals had very similar SLL values except at a depth of 45 mm, where a difference of about 3 dB was observed

between MOCGCS and CGS SSL values. On the other hand, the SLL of the short pulse signal was about 5 dB higher at a depth of 25 mm and about 4.5 dB lower at a depth of 105 mm compared to the MOCGCS SLL values at specified depths.

SNR values for MOCGCS, CGS and short pulse transmitted signals were compared using RF echoes from the tissue-mimicking phantom. Specifically, SNR at different depths were determined from the beamformed RF signal along scanline #65 of the synthesized B-mode images of the (see Fig. 7). It can be seen from Fig. 6 that SNR values obtained using MOCGCS and CGS signals did not differ significantly. The maximum deterioration of SNR for MOCGCS was observed at depths of 40 mm, close to 1.5 dB. Moreover, coded transmission yielded an increase in SNR ranging from 4.8 dB to 8.9 dB over the short pulse signal across entire range of depths, with this gain becoming especially pronounced for deeper wires located at depths above 40 mm.

Finally, the experimental data from the tissue-mimicking phantom indicated the benefits of MOCGCS. Specifically, B-mode images obtained using the STA method with CGS, MOCGCS, and the short pulse excitation signals are shown in the Fig. 7. It can be seen (by visual assessment) that the images of comparable quality can be obtained using conventional CGS signal and MOCGCS, with the latter yielding a twofold frame increase in comparison to the conventional CGS, while maintaining better visualization depth in comparison to the short pulse transmit signal at the same time.

In-vivo recorded scans from the abdominal aorta and common carotid artery of a volunteer seem to be identical regarding minor fragments in backscattered echoes. However, the signal gain was clearly obtained and was proportional to the ratio of the code length to the short pulse length, being close to eight time greater.

The results confirmed that simultaneous MOCGCS transmission and separation of combined RF echoes on the receiver side does not degrade signal parameters compared to conventional CGS. Additionally, transmitting two MOCGCS sets maintains the axial resolution and data acquisition speed of short pulse transmission. The frame rate can potentially be further increased by using more than two MOCGCS sets. The results demonstrate the advantage of MOCGCS in modern ultrasonography compared to conventional short pulse excitation and CGS.

## References

1. BAE M.H. (2003), *Ultrasound imaging method and apparatus using orthogonal Golay codes*, U.S. Patent 6638227B2.
2. BAE M.-H., LEE W.-Y., JEONG M.-K., KWON S.-J. (2002), Orthogonal Golay code based ultrasonic imaging without reducing frame rate, [in:] *2002 IEEE Ultrasonics Symposium, 2002. Proceedings*, pp. 1705–1708, doi: [10.1109/ULTSYM.2002.1192625](https://doi.org/10.1109/ULTSYM.2002.1192625).
3. CHIAO R.Y., THOMAS L.J. (2000), Synthetic transmit aperture imaging using orthogonal Golay coded excitation, [in:] *2000 IEEE Ultrasonics Symposium. Proceedings. An International Symposium*, pp. 1677–1680, doi: [10.1109/ULTSYM.2000.921644](https://doi.org/10.1109/ULTSYM.2000.921644).
4. DEMI L., VITI J., KUSTERS L., GUIDI F., TORTOLI P., MISCHI M. (2013), Implementation of parallel transmit beamforming using orthogonal frequency division multiplexing-achievable resolution and interbeam interference, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **60**(11): 2310–2320, doi: [10.1109/TUFFC.2013.6644735](https://doi.org/10.1109/TUFFC.2013.6644735).
5. GOLAY M.J.E. (1961), Complementary series, *IRE Transactions on Information Theory*, **7**(2): 82–87.
6. GRAN F., JENSEN J.A. (2006), Frequency division transmission imaging and synthetic aperture reconstruction, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **53**(5): 900–911, doi: [10.1109/tuffc.2006.1632681](https://doi.org/10.1109/tuffc.2006.1632681).
7. HUANG X. (2005), Simple implementation of mutually orthogonal complementary sets of sequences, [in:] *Proceedings. International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 369–372.
8. KIM B.-H., SONG T.-K. (2003), Multiple transmit focusing using modified orthogonal Golay codes for small scale systems, [in:] *IEEE Symposium on Ultrasonics*, pp. 1574–1577, doi: [10.1109/ULTSYM.2003.1293208](https://doi.org/10.1109/ULTSYM.2003.1293208).
9. KUMRU Y., KOYMEN H. (2018), Beam coding with orthogonal complementary Golay codes for signal-to-noise ratio improvement in ultrasound mammography, *The Journal of the Acoustical Society of America*, **144**(3): 1888–1888, doi: [10.1121/1.5068277](https://doi.org/10.1121/1.5068277).
10. MISARIDIS T. (2001), *Ultrasound imaging using coded signals*, Ph.D. Thesis, Center for Fast Ultrasound Imaging Technical University of Denmark.
11. MISARIDIS T., JENSEN J.A. (2005), Use of modulated excitation signals in medical ultrasound. Part I: basic concepts and expected benefits, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **52**(2): 177–191, doi: [10.1109/TUFFC.2005.1406545](https://doi.org/10.1109/TUFFC.2005.1406545).
12. NOWICKI A., SECOMSKI W., LITNIEWSKI J., TROTS I. (2003), On the application of signal compression using Golay's codes sequences in ultrasound diagnostic, *Archives of Acoustics*, **28**(4): 313–324.
13. PENG H., HAN X., LU J. (2006), Study on application of complementary Golay code into high frame rate ultrasonic imaging system, *Ultrasonics*, **44**: e93–e96, doi: [10.1016/j.ultras.2006.06.030](https://doi.org/10.1016/j.ultras.2006.06.030).
14. RAMALLI A., GUIDI F., BONI E., TORTOLI P. (2015), A real-time chirp-coded imaging system with tissue attenuation compensation, *Ultrasonics*, **60**: 65–75, doi: [10.1016/j.ultras.2015.02.013](https://doi.org/10.1016/j.ultras.2015.02.013).
15. TIAN L., LU X., XU C., LI Y. (2021), New mutually orthogonal complementary sets with non-power-of-two

- lengths, *IEEE Signal Processing Letters*, **28**: 359–363, doi: [10.1109/LSP.2021.3054565](https://doi.org/10.1109/LSP.2021.3054565).
16. TSENG C.C., LIU C.L. (1972), Complementary sets of sequences, *IEEE Transactions on Information Theory*, **18**(5): 644–652, doi: [10.1109/TIT.1972.1054860](https://doi.org/10.1109/TIT.1972.1054860).
  17. TROTS I. (2015), Mutually orthogonal Golay complementary sequences in synthetic aperture imaging systems, *Archives of Acoustics*, **40**(2): 283–289, doi: [10.1515/aoa-2015-0031](https://doi.org/10.1515/aoa-2015-0031).
  18. TROTS I., NOWICKI A., SECOMSKI W., LITNIEWSKI J. (2004), Golay sequences – side-lobe – canceling codes for ultrasonography, *Archives of Acoustics*, **29**(1): 87–97.
  19. TROTS I., TASINKEVYCH Y., NOWICKI A. (2015), Orthogonal Golay codes with local beam pattern correction in ultrasonic imaging, *IEEE Signal Processing Letters*, **22**(10): 1681–1684, doi: [10.1109/LSP.2015.2423619](https://doi.org/10.1109/LSP.2015.2423619).
  20. TROTS I., TASINKEVYCH Y., NOWICKI A., LEWANDOWSKI M. (2011), Golay coded sequences in synthetic aperture imaging systems, *Archives of Acoustics*, **36**(4): 913–926, doi: [10.2478/v10168-011-0061-5](https://doi.org/10.2478/v10168-011-0061-5).
  21. TROTS I., ŻOLEK N., TASINKEVYCH J., WÓJCIK J. (2022), Mutually Orthogonal Golay Complementary Sequences in Medical Ultrasound Diagnostics. Experimental Study, *Archives of Acoustics*, **47**(3): 399–405, doi: [10.24425/aoa.2022.142013](https://doi.org/10.24425/aoa.2022.142013).
  22. WU SW., CHEN CY., LIU Z. (2020), How to construct mutually orthogonal complementary sets with non-power-of two lengths?, *IEEE Transactions on Information Theory*, **67**(6): 3464–3472, doi: [10.1109/TIT.2020.2980818](https://doi.org/10.1109/TIT.2020.2980818).
  23. ZHAO F., LUO J. (2018), Diverging wave compounding with spatio-temporal encoding using orthogonal Golay pairs for high frame rate imaging, *Ultrasonics*, **89**: 155–165, doi: [10.1016/j.ultras.2018.05.009](https://doi.org/10.1016/j.ultras.2018.05.009).



## Research Paper

Influence of Ultrasonic Cavitation on *Botryococcus Braunii* Growth

Asleena SALAEH

Division of Physics, School of Science, Walailak University  
Thailand; e-mail: [as.salaeh@gmail.com](mailto:as.salaeh@gmail.com)

(received November 9, 2023; accepted April 3, 2024; published online June 25, 2024)

This study investigates ultrasonic energy's impact on enhancing the growth of *Botryococcus braunii* (*B. braunii*) microalgae. Microalgae, known for their advantages in greenhouse gas mitigation and biomass conversion, were subjected to various stressors, including ultrasonic waves, to optimize productivity. Ultrasonic waves induce acoustic cavitation, increasing membrane permeability and substrate conversion. The study examined the impact of energy and maximum pressure resulting from bubble collapse on the relative specific growth rate of *B. braunii* microalgae. It was observed that reproduction showed a promotive trend until the energy surpassed 30 kJ. However, when ultrasonic energy reached 18.2 kJ, reproduction was inhibited due to the maximum pressure generated during bubble bursting, which reached  $5.7 \mu\text{N}/\mu\text{m}^2$ , leading to the suppression of reproduction upon encountering bubble collapse events. Under specific ultrasonic conditions (15.1 kJ energy, maximum pressure of  $45.5 \times 10^5 \text{ Pa}$ ), a maximum specific growth rate of  $0.329 \pm 0.020 \text{ day}^{-1}$  in a two-day interval boosted *B. braunii* microalgae biomass productivity. These findings advance our understanding of ultrasonic wave effects on microalgae reproduction and underscore the potential for optimizing ultrasonic parameters to enhance biomass production.

**Keywords:** *Botryococcus braunii*; ultrasonic wave; cavitation; specific growth rate; bubble size.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Microalgae offer numerous benefits over conventional crops, such as rapid growth, ease of cultivation, and space efficiency (ENMAK, 2010). Notably, microalgae can mitigate the greenhouse effect by converting carbon dioxide into biomass via photosynthesis, thereby lowering greenhouse gas levels in the atmosphere, and supporting carbon sequestration (ONYE-AKA *et al.*, 2021).

Despite the challenges of microalgae cultivation, their unique characteristics make them attractive for industrial applications. To fully exploit their potential, researchers have explored various environmental stressors, including nutrient scarcity, high temperatures, intense light, and elevated pH levels (SHAKIROV *et al.*, 2021; FU *et al.*, 2019). Innovative technologies such as ultrasonic waves have been used to apply controlled stress conditions, promoting cell proliferation and optimizing metabolic activity.

Ultrasonic waves, generated through cavitation, employ low-energy waves to stimulate cells. The im-

plosion of cavities near the cell surface enhances membrane permeability, facilitating nutrient and molecule exchange. Research confirms that low-energy ultrasonic waves significantly increase *Botryococcus braunii* (*B. braunii*) growth rates compared to conditions without them (WANG *et al.*, 2014; XU *et al.*, 2014). This study primarily aims to develop ultrasonic wave stimulation for enhanced *B. braunii* growth and biomass production.

## 2. Theory of acoustic cavitation

The Rayleigh–Plesset equation is a second-order ordinary differential equation that governs the dynamics of a spherical bubble within an infinite fluid. It plays a crucial role in understanding cavitation phenomena and predicting the behavior of such a bubble. The Rayleigh–Plesset equation is expressed as follows:

$$\rho_0 \left[ R\ddot{R} + \frac{3}{2}\dot{R}^2 \right] = (p_v - p_\infty(t)) + p_{g0} \left( \frac{R_0}{R} \right)^{3k} - \frac{2S}{R} - 4\mu \frac{\dot{R}}{R}, \quad (1)$$

where  $\dot{R}$  and  $\ddot{R}$  are the first- and second-order derivatives of the bubble radius with respect to time,  $R_0$  is the initial bubble radius in meters [m],  $\rho_0$  is the density of fluid in kilograms per cubic meter [kg/m<sup>3</sup>],  $S$  is the surface tension in Newtons per meter [N/m],  $\mu$  is the dynamic viscosity of fluid in Pascal-second [Pa·s],  $p_v$  is the vapor pressure of fluid,  $p_{g0}$  is the gas pressure in the bubble at its ambient state ( $p_{g0} = p_0 + \frac{2S}{R_0} - p_v$ ), and  $p_\infty(t)$  is the variation in bulk pressure as function of time, which is given by Eq. (2):

$$p_\infty(t) = p_0 - p_A \sin(2\pi ft), \quad (2)$$

where  $p_0$  is the ambient pressure (1 atm, 101.325 kPa),  $f$  is the frequency [Hz],  $p_A$  is the driving pressure [Pa] that is correlated with the acoustic intensity ( $I_{ac}$ ) and acoustic power (WANG, YUAN, 2016):

$$p_A = \sqrt{2I_{ac}\rho_0c_0}, \quad (3)$$

$$I_{ac} = \frac{\text{Power}}{\text{Volume or area}}, \quad (4)$$

where  $c_0$  is the speed of sound in fluid in meters per second [m/s]. Ultrasonic waves impact microalgae cells through cavitation, a phenomenon that occurs when ultrasonic waves pass through a substance. These waves create alternating regions of compression and rarefaction as they propagate. During the rarefaction phase, vapor bubbles can form within the liquid medium, leading to cavitation. As the ultrasonic wave continues to exert pressure on the medium, these bubbles grow until they reach a critical size. When this happens, the bubbles collapse rapidly during the compression phase.

When the cavitation bubble grows to a significant size, the effects of non-condensable gas, surface tension, and viscosity become negligible. As a result, the complex Rayleigh–Plesset equation in Eq. (1) can be simplified to:

$$\rho_0 \left[ R\ddot{R} + \frac{3}{2}\dot{R}^2 \right] = (p_v - p_\infty(t)). \quad (5)$$

During the collapse phase, when the applied pressure exceeds the vapor pressure ( $p_\infty > p_v$ ) and the bubble radius decreases ( $R < R_0$ ), the interface velocity can be described as follows:

$$\dot{R} \cong -\sqrt{\frac{2(p_\infty - p_v)}{3\rho_0} \left( \left( \frac{R_0}{R} \right)^2 - 1 \right)}. \quad (6)$$

The force generated by the collapse of cavitation bubbles plays a significant role in both promoting and disrupting microalgae cells. This force is influenced by the initial size of the cavitation bubbles, which, in turn, depends on the frequency of the ultrasonic waves. This relationship can be expressed through Eq. (7) (BRENNEN, 2005):

$$f_n = \frac{1}{2\pi} \sqrt{\frac{1}{\rho_0 R_e^2} \left( 3k(p_0 - p_v) + \frac{2S}{R_e} (3k - 1) \right)}, \quad (7)$$

where  $k$  is approximately constant,  $S$  is the surface tension,  $R_e$  is the equilibrium radius at pressure ( $p_0$ ), and  $(p_0 - p_v)$  represents tension of the fluid.

When a bubble collapses near a cell wall or cell membrane, it can cause disruption or alteration in the shape of a microalgae cell (LIU *et al.*, 2022). This effect occurs when the bubble experiences a stress exceeding its yield strength, resulting in actual deformation. To comprehend how a collapsing bubble can generate such high pressure, the maximum pressure ( $p_{\max}$ ) in Pascals [Pa] can be approximated using Eq. (8):

$$p_{\max} \cong 0.157(p_\infty - p_v) \left( \frac{R_0}{R} \right)^3 + p_\infty. \quad (8)$$

Furthermore, this maximum pressure occurs at a distance ( $r_{\max}$ ) in meters [m] from the bubble center, given by Eq. (9):

$$r_{\max} \cong 1.59 R. \quad (9)$$

### 3. Material and methods

#### 3.1. Microalgae and culture conditions

Plankton samples were collected from the reservoir at Walailak University in Nakhon Si Thammarat, Thailand (latitude 8°38'32.25"N, longitude 99°54'26.52"E) using a plankton net with a mesh size of 67 μm. This reservoir stores brackish water with a slow flow, a pH of 7.2, and a dissolved oxygen level of 6.47 mg/L (RUBSAI, 2012). Colonies of *B. braunii* were isolated from the collected samples and cultured individually in 5 ml of BG-11 liquid medium for one month. They were then transferred to 250 ml flasks with 100 ml of BG-11 medium for a week. Subsequently, 10 ml of *B. braunii* algae were divided and cultivated in separate 1000 ml flasks with 800 ml of BG-11 medium, resulting in 13 distinct experimental conditions. The cultures were maintained in a sterile chamber at a constant temperature of 25 °C, with 3000 lux illumination during a 16:8-hour light-dark period. The structure of *B. braunii* microalgae is illustrated in Fig. 1. In *B. braunii* microalgae, ovoid cells form tetrad patterns at the colony center. As colonies develop, these cells enlarge and transform into conical

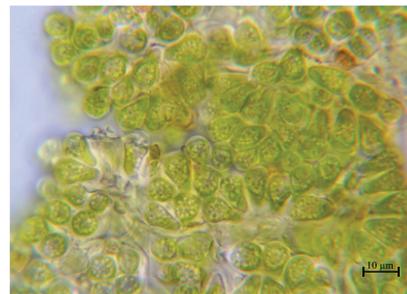


Fig. 1. *Botryococcus braunii* microalgae.

cells, playing a crucial role in defining the outer boundaries (GUPTE, 2012).

### 3.2. Ultrasonic treatment and experiments

The investigation utilized an ultrasonic bath (Elmasonic P60H), operating at frequencies of 37 kHz and 80 kHz, with a 5.75-liter tank. Microalgae were subjected to ultrasonic excitation at specific power levels: 57.2 W, 83.6 W, and 132.3 W for 37 kHz, and 46.3 W, 67.2 W, and 101.2 W for 80 kHz (refer to Table 1). Exposure durations were precisely 3 min and 5 min, with 2-day intervals. Optimal excitation parameters were determined based on the study by SALAEH *et al.* (2017). Each stimulation with ultrasonic wave for all condition was conducted at 9.00 a.m. Thailand time according to these specified conditions. After ultrasonic treatment, microalgae samples were cultured under consistent flask conditions. Each experimental condition was replicated three times to ensure reliability. Biomass measurements were recorded every third day throughout the study.

The effective ultrasonic power of the system was assessed by measuring the power of the ultrasonic bath at different temperatures. This measurement was essential as the application of power to the system induces energy transfer to the liquid, resulting in molecular movement and the generation of thermal energy. The power of the ultrasonic bath ( $P$ , in W) can be calculated using Eq. (10) (LIU *et al.*, 2022):

$$P = mC_p \frac{\Delta T}{\Delta t}, \quad (10)$$

where  $m$  is the mass of fluid,  $C_p$  is the specific heat capacity of fluid,  $\Delta T$  is the temperature difference between the initial temperature and the final temperature after a specific reaction time  $\Delta t$ .

### 3.3. Microalgae growth and biomass measurement

A 1 ml sample of *B. braunii* cells was transferred to a tube for counting cell density using a hemocytometer slide. The sample was carefully placed on the slide, covered, and observed under an optical microscope (Olympus CH20i microscope). Daily counting of the microalgae with a microscope allowed for precise determination of the specific growth rate of *B. braunii* microalgae. Cell density calculations were conducted by cap-

turing images of microalgae samples within a hemocytometer under a microscope. The cell count was determined through image processing, followed by calculating the cell density using Eq. (11) (DILIA *et al.*, 2018):

$$\text{Cell density} = \frac{\text{Number of counted cells} \times 25 \times 10^4}{\text{Number of boxes}}. \quad (11)$$

The growth data from cultivating *B. braunii* under different ultrasonic parameters validated the presented logistic model. The logistic model of population growth is written as follows (ZHANG *et al.*, 2016):

$$N(t) = \frac{KN_0}{N_0 + (K - N_0)e^{-rt}}, \quad (12)$$

where  $N$  is the number of cells as a function of time  $t$ ,  $K$  is the maximum number of cells or the carrying capacity, and  $r$  is the specific growth rate.

Biomass content was determined by measuring the dry weight-to-volume ratio (mg/ml). A 12 ml algae aliquot was centrifuged, filtered onto a weighed 47 mm glass fiber filter, dried at 103 °C for 13 hours, and then weighed. Biomass productivity ( $g \text{ (L day)}^{-1}$ ) is defined as follows (CHEN *et al.*, 2016):

$$\text{Biomass productivity} = \frac{B_2 - B_1}{t_2 - t_1}, \quad (13)$$

where  $B_1$  and  $B_2$  represent the biomass concentration at times  $t_1$  and  $t_2$ , respectively, representing the initial and final points.

## 4. Results and discussion

### 4.1. The behaviors of acoustic cavitation within ultrasonic bath

Acoustic cavitation occurs due to pressure fluctuations induced by ultrasonic waves in a fluid. These waves compress and rarefy the fluid, forming cavitation bubbles that explosively collapse, enlarging until reaching resonance velocity. This implosion generates pores near cell membranes, enhancing permeability. Ultrasonic treatment holds potential for improved substrate conversion through enhanced diffusion (REN *et al.*, 2019; PEREIRA *et al.*, 2023). Multiple factors in Table 1, including the maximum cavitation radius ( $R_{\max}$ ), maximum pressure ( $p_{\max}$ ) during collapse, and

Table 1. Numerical results of cavitation bubble dynamics.

Frequency [kHz]	Power [W]	Intensity [mW/cm <sup>2</sup> ]	$R_0$ [μm]	$R_{\max}$ [μm]	$R_{\min}$ [μm]	$\frac{R_{\max}}{R_{\min}}$	$\dot{R}$	$p_{\max}$ ( $\times 10^5$ Pa)	$r_{\max}$ [μm]
37	57.2	4.9	74	141.5	41.7	3.4	49.77	5.0	66.4
	83.6	7.2		145.5	21.7	6.71	140.2	45.5	34.5
	132.3	11.4		150.4	14.5	10.41	271.5	173.1	23.0
80	46.3	4.0	35	61.4	16.9	3.63	55.2	6.3	26.9
	67.2	5.8		63.5	14.7	4.3	72.29	11.5	23.3
	101.2	8.7		65.9	9.1	7.2	156.49	56.9	14.5

the distance ( $r_{\max}$ ) from the bubble center where the maximum pressure occurs, are influenced by ultrasonic parameters and liquid characteristics.

The numerical results on the dynamic behavior of cavitation bubbles for stimulated microalgae in an ultrasonic bath are summarized in Table 1, presenting the following characteristics. According to the experiment, the medium's density ( $\rho_m$ ) is 1001.68 kg/m<sup>3</sup>, and its sound speed ( $c_m$ ) is 1353.22 m/s, which is nearly equal to that of water. Consequently, the dynamic viscosity and applied surface tension were 0.00891 Pa·s and 0.0072 N/m, respectively. The relationship between ultrasonic power and intensity is described by Eq. (4), wherein increasing ultrasonic power enhances the stretching effect of the ultrasonic wave on cavitation formation in the positive pressure region (HAO *et al.*, 2021).

The maximum size of the cavitation bubble was calculated by solving Eq. (1) using the numerical method. During the expansion phase, the bubble grows until it reaches its maximum size ( $R_{\max}$ ) before collapsing. Importantly, the  $R_{\max}$  is inversely related to the emitted ultrasonic frequency. According to Eq. (7), the initial bubble size is 74  $\mu\text{m}$  and 35  $\mu\text{m}$  at ultrasonic frequencies of 37 kHz and 80 kHz, respectively. The ratio of maximum bubble size to initial bubble size increases with higher ultrasonic power. Specifically, at 37 kHz, the ratio ranges from 1.91 to 2.03, while at 80 kHz, it ranges from 1.75 to 1.88. It is important to note that the ratio of bubble size at 80 kHz is relatively lower due to reduced acoustic pressure in the ultrasonic bath and the smaller wavelength ( $\lambda$ ) compared to the 37 kHz ultrasonic frequency. During the collapse phase (from  $R_{\max}$  to  $R_{\min}$ ), the  $p_{\max}$  increases with ultrasonic power for each frequency. For instance, at 37 kHz, with power levels of 57.2 W, 83.6 W, and 132.3 W,  $R_{\max}$  is 141.5  $\mu\text{m}$ , 145.5  $\mu\text{m}$ , and 150.4  $\mu\text{m}$ , respectively. The maximum pressure points are situated at distances of 66.4  $\mu\text{m}$ , 34.5  $\mu\text{m}$ , and 23.0  $\mu\text{m}$  away from the bubble as it collapses. Similarly, at an ultrasonic frequency of 80 kHz, with power levels of 46.3 W, 67.2 W, and 101.2 W, the  $R_{\max}$  values are 61.4  $\mu\text{m}$ , 63.5  $\mu\text{m}$ , and 65.9  $\mu\text{m}$ , and the maximum pressure points are located 26.9  $\mu\text{m}$ , 23.3  $\mu\text{m}$ , and 14.5  $\mu\text{m}$  away from the collapsing bubble.

Table 1 provides the  $R_{\max}/R_{\min}$  ratios, offering valuable insights into the volume change experienced by the bubble during collapse, serving as an indicator of its compression ratio. This ratio reflects the severity of the collapse (KANTHALE *et al.*, 2008). For a 37 kHz ultrasonic frequency,  $R_{\max}/R_{\min}$  ratios are 3.39, 6.71, and 10.41 at ultrasonic powers of 57.2 W, 83.6 W, and 132.3 W, respectively. Meanwhile, for an 80 kHz ultrasonic frequency,  $R_{\max}/R_{\min}$  ratios are 3.63, 4.33, and 7.22 at ultrasonic powers of 46.3 W, 67.2 W, and 101.2 W, respectively. As the applied acoustic power increases, the bubbles expe-

rience higher negative pressure during the rarefaction cycle and higher positive pressure during the compression cycle, leading to an elevated compression ratio of the bubble cavity. This ratio increases alongside the collapsed interface velocity ( $\dot{R}$ ), resulting in higher maximum pressures during the collapsed phase. Consequently, the force generated by the collapsing process increases. At an ultrasonic frequency of 37 kHz, this force is estimated to be 498.43 nN/ $\mu\text{m}^2$ , rising to 17 306.6 nN/ $\mu\text{m}^2$  with increasing ultrasonic power. Similarly, at 80 kHz, the force increases from 633.04 nN/ $\mu\text{m}^2$  to 5690.69 nN/ $\mu\text{m}^2$  with increasing ultrasonic power.

#### 4.2. Effect of energy and maximum pressure during collapse phase on the growth rate of microalgae

The total energy provided to the microalgae sample can be determined by taking into account the treatment period and ultrasonic power. The interaction between the total energy delivered to the microalgae and the relative specific growth rate of *B. braunii* microalgae during the two-day interval is shown in Fig. 2. *Botryococcus braunii* microalgae experienced an appropriate growth rate when ultrasonic energy was applied. The growth rate, however, changed from positive to negative when the microalgae received up to 30 kJ of total energy. In addition to specific growth rate, Table 2 shows biomass productivity.

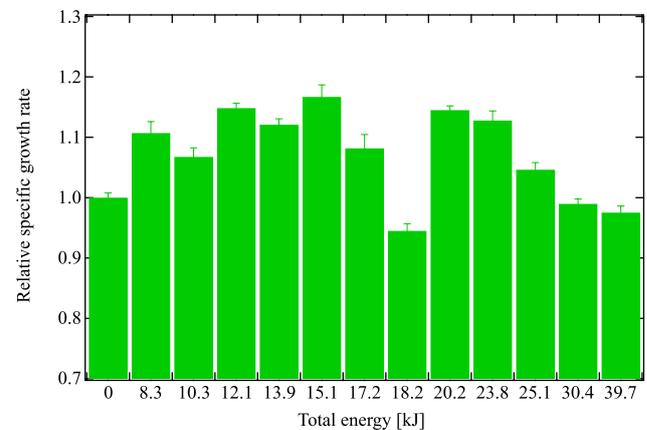


Fig. 2. Effect of total energy per excitation on the relative specific growth rate of *B. braunii* microalgae for a two-day interval. Error bars were used to represent the standard deviations derived from triplicate measurements of the data.

This shows that the number of microalgae cells decreased when subjected to a high level of energy from the ultrasonic wave compared to microalgae that were not exposed to ultrasonic energy. However, a negative growth rate is observed when the microalgae are exposed to an energy level of 18.2 kJ. Therefore, it is vital to consider other aspects of the ultrasonic wave.

Figure 3 displays the relationship between the maximum pressure generated during bubble collapse

Table 2. Specific growth rate of *B. braunii* under the influence of ultrasonic excitation from the logistic model and the biomass productivity.

Ultrasonic power [W]	Time/Energy [s/kJ]	Specific growth rate ( $r$ , day <sup>-1</sup> )	Biomass productivity ( $g (L_{\text{day}})^{-1}$ )
Frequency 37 kHz			
Control	0	0.282 ± 0.008	0.0361
57.2	180/10.3	0.301 ± 0.015	0.0421
83.6	180/15.1	0.329 ± 0.020	0.0486
132.3	180/23.8	0.318 ± 0.016	0.0435
57.8	300/17.2	0.305 ± 0.023	0.0403
83.6	300/25.1	0.295 ± 0.012	0.0454
132.3	300/39.7	0.275 ± 0.011	0.0287
Frequency 80 kHz			
Control	0	0.290 ± 0.007	0.0117
46.3	180/8.3	0.321 ± 0.019	0.0144
67.2	180/12.1	0.333 ± 0.008	0.0206
101.2	180/18.2	0.274 ± 0.012	0.0189
46.3	300/13.9	0.325 ± 0.010	0.0183
67.2	300/20.2	0.325 ± 0.007	0.0194
101.2	300/30.4	0.287 ± 0.008	0.0156

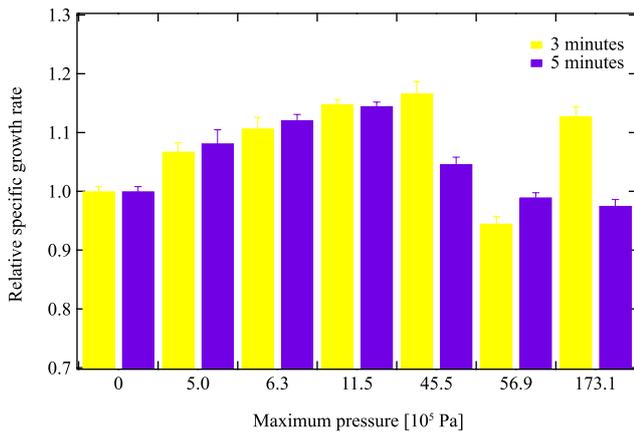


Fig. 3. Effect of maximum pressure during collapsing phase from  $R_{\text{max}}$  to  $R_{\text{min}}$  on the relative specific growth rate of *B. braunii* microalgae for a two-day interval. The yellow bar represents the exposure time of 3 min, and the purple color represents the exposure time of 5 min. Error bars were used to represent the standard deviations derived from triplicate measurements of the data.

and the relative specific growth rate of *B. braunii* microalgae. When the maximum pressure increases, up to a maximum pressure of  $45.5 \times 10^5$  Pa, the growth rate of *B. braunii* microalgae shows an increase. However, the specific growth rate of the microalgae becomes negative when the collapse pressure surpasses  $56.9 \times 10^5$  Pa. This shows that the impact of the maximum pressure during bubble collapse has a negative impact on the growth of the *B. braunii* microalgae above this threshold.

In a previous scenario, exposing microalgae to ultrasonic energy of 18.2 kJ led to a decline in their growth rate. Examination of the maximum pressure

during the collapse phase revealed potential contact between microalgae and collapsed cavity bubbles, reaching a maximum collapse pressure of  $56.9 \times 10^5$  Pa. The force generated by the collapse was estimated based on this maximum pressure. Remarkably, the force exerted by the collapse of the bubble reached up to  $5.7 \mu\text{N}/\mu\text{m}^2$  when considering a typical spherical microalgal cell (*B. braunii*) with a diameter of approximately  $9 \mu\text{m}$  (TASIĆ *et al.*, 2016). Consequently, the force applied to the microalgae was around  $362.6 \mu\text{N}$ . In the report of LEE *et al.* (2012), the force required to rupture the cell wall of algae is about  $11.33 \mu\text{N}$ , surpassing the threshold for growth inhibition in microalgae. In scenarios where maximum pressure exceeds  $45.5 \times 10^5$  Pa for the 3-minute exposure, the relative specific growth rate initially declines, followed by an increase. The study by ANTONY (1963) demonstrated that cavities occur at half-wavelength within the ultrasonic bath for each frequency. Therefore, ultrasonic frequencies of 37 kHz and 80 kHz have wavelengths of 40.41 mm and 18.69 mm, respectively. The distances between bubble collapse locations are approximately 20.21 mm and 9.35 mm for frequencies of 37 kHz and 80 kHz, respectively, with the latter being shorter than the former. Consequently, in the ultrasonic bath, the collapse positions for the 80 kHz ultrasonic wave exceeded those of the 37 kHz. This observation suggests that under the maximum pressure condition of  $56.9 \times 10^5$  Pa, the microalgae were exposed to greater bubble collapse than in the  $173.1 \times 10^5$  Pa maximum pressure condition, leading to inhibited specific growth rate.

The research findings demonstrate that the growth of *B. braunii* microalgae can be stimulated by various parameters. Among these conditions, the maximum

specific growth rate observed was  $0.329 \pm 0.020 \text{ day}^{-1}$ , achieved when the microalgae cells were exposed to ultrasonic energy of 15.1 kJ and a maximum pressure during collapse of approximately  $45.5 \times 10^5 \text{ Pa}$ . These conditions involved the application of 37 kHz ultrasonic frequency, with a three-minute exposure at a two-day interval. The stimulation of *B. braunii* microalgae by ultrasonic waves resulted in higher biomass productivity compared to microalgae not exposed to ultrasonic waves (XU *et al.*, 2014). Moreover, cavitation phenomena generated by ultrasonic waves also contribute to nutrient and oxygen transport through membrane permeability, promoting the growth of *Echinacea purpurea*, *Staphylococcus epidermidis*, *Pseudomonas aeruginosa*, *Escherichia coli*, and *B. braunii* (PITT, ROSS, 2003; XU *et al.*, 2014).

Ultrasound-induced cavitation in a medium involves the creation and oscillation of gas bubbles, which subsequently collapse near cells. This phenomenon converts potential energy into chemical, thermal, and mechanical energies in the form of reactive oxygen species (TOPAZ *et al.*, 2005). The interaction of cavitation near the cell involves a process known as sonoporation (KUDO *et al.*, 2009). Sonoporation refers to the transient and dynamic increase in cell membrane permeability, resulting from complex processes involving bubble physics and bubble-cell interactions (FAN *et al.*, 2014). When exposed to cavitation, microalgae cells with thin cell walls (e.g., *Chlamydomonas concordia*) or without cell walls (e.g., *Dunaliella salina*) are destroyed. However, cells with thick cell walls or those existing in colonies (e.g., *Nannochloropsis oculata*) tend to increase in population due to the ultrasound-induced separation of colonies, resulting in a higher number of individual cells or smaller colonies (JOYCE *et al.*, 2014). It is possible that ultrasonic waves have caused fissures in the cell wall or separation of colonies, facilitating nutrient transport into the cells.

## 5. Conclusion

The study explored the impact of ultrasonic energy and maximum pressure during bubble collapse on the growth of *B. braunii* microalgae. Acoustic cavitation induced by ultrasonic waves generates cavitation bubbles that collapse explosively, creating pores near cell membranes and enhancing substrate conversion. Varied ultrasonic parameters, including frequency, power, exposure time, and intervals, were tested to stimulate microalgae growth. The influence of energy and the maximum pressure generated during bubble collapse was observed under all conditions when the exposure was extended to 2-day intervals. The reproduction of microalgae showed an enhancing trend until the energy exceeded 30 kJ. However, once the energy reached 18.2 kJ, the reproduction of microalgae was inhibited. This inhibitory effect can be attributed to the max-

imum pressure generated during the bubble bursting, which was found to reach  $5.7 \mu\text{N}/\mu\text{m}^2$  under these conditions. Considering a microalga with an approximate diameter of  $9 \mu\text{m}$ , the resulting force is approximately  $362.6 \mu\text{N}$ . Consequently, microalgae can restrain their reproduction when encountering bubble collapse events. The highest observed specific growth rate was  $0.329 \pm 0.020 \text{ day}^{-1}$ , achieved with 15.1 kJ of ultrasonic energy, a maximum pressure of approximately  $45.5 \times 10^5 \text{ Pa}$ , and 37 kHz ultrasonic frequency with a three-minute exposure at a two-day interval. This stimulation resulted in increased biomass productivity compared to untreated microalgae.

## Acknowledgments

The author expresses gratitude for the financial support provided by the Royal Golden Jubilee (RGJ) Ph.D. Programme (grant no. PHD/0221/2560), through the National Research Council of Thailand (NRCT), and the Thailand Research Fund (TRF). The guidance of lecturers from the Division of Physics, Walailak University is also acknowledged.

## References

- ANTONY O.A. (1963), Technical aspects of ultrasonic cleaning, *Ultrasonics*, **1**(4): 194–198, doi: [10.1016/0041-624X\(63\)90167-7](https://doi.org/10.1016/0041-624X(63)90167-7).
- BRENNEN C. (2005), *Fundamentals of Multiphase Flow*, Cambridge University Press, pp. 345.
- CHEN J.T. *et al.* (2016), Preliminary assessment of Malaysian microalgae strains for the production of bio jet fuel, [in:] *IOP Conference Series Materials Science and Engineering*, **152**: 012042, doi: [10.1088/1757-899X/152/1/012042](https://doi.org/10.1088/1757-899X/152/1/012042).
- DILIA P., KALSUM L., RUSDIANASARI R. (2018), Fatty acids from microalgae *Botryococcus braunii* for raw material of biodiesel, *Journal of Physics: Conference Series*, **1095**: 012010, doi: [10.1088/1742-6596/1095/1/012010](https://doi.org/10.1088/1742-6596/1095/1/012010).
- ENMAK P. (2010), Influences of CO<sub>2</sub> concentrations and salinity on acceleration of microalgal oil as raw material for biodiesel production, *Journal of Biotechnology*, **150**: 19, doi: [10.1016/j.jbiotec.2010.08.063](https://doi.org/10.1016/j.jbiotec.2010.08.063).
- FAN Z., KUMON R.E., DENG C.X. (2014), Mechanisms of microbubble-facilitated sonoporation for drug and gene delivery, *Therapeutic Delivery*, **5**(4): 467–486, doi: [10.4155/tde.14.10](https://doi.org/10.4155/tde.14.10).
- FU L., LI Q., YAN G., ZHOU D., CRITTENDEN J.C. (2019), Hormesis effects of phosphorus on the viability of *Chlorella regularis* cells under nitrogen limitation, *Biotechnology for Biofuels*, **12**: 121, doi: [10.1186/s13068-019-1458-z](https://doi.org/10.1186/s13068-019-1458-z).

8. GUPTE Y. (2012), Morpho-physiological characteristics of *Botryococcus braunii* (Kützinger, 1849) & its oil production from the species isolated from Thane, Maharashtra, India, *Asian Journal of Microbiology, Biotechnology and Environmental Sciences*, **14**(4): 523–526.
9. HAO X., SUO H., PENG H., XU P., GAO X., DU S. (2021), Simulation and exploration of cavitation process during microalgae oil extracting with ultrasonic-assisted for hydrogen production, *International Journal of Hydrogen Energy*, **46**(3): 2890–2898, doi: [10.1016/j.ijhydene.2020.06.045](https://doi.org/10.1016/j.ijhydene.2020.06.045).
10. JOYCE E., KING P., MASON T. (2014), The effect of ultrasound on the growth and viability of microalgae cells, *Environmental Biology of Fishes*, **26**: 1741–1748, doi: [10.1007/s10811-013-0202-5](https://doi.org/10.1007/s10811-013-0202-5).
11. KANTHALE P., ASHOKKUMAR M., GRIESER F. (2008), Sonoluminescence, sonochemistry (H<sub>2</sub>O<sub>2</sub> yield) and bubble dynamics: Frequency and power effects, *Ultrasonics Sonochemistry*, **15**(2): 143–150, doi: [10.1016/j.ultsonch.2007.03.003](https://doi.org/10.1016/j.ultsonch.2007.03.003).
12. KUDO N., OKADA K., YAMAMOTO K. (2009), Sonoporation by single-shot pulsed ultrasound with microbubbles adjacent to cells, *Biophysical Journal*, **96**(12): 4866–4876, doi: [10.1016/j.bpj.2009.02.072](https://doi.org/10.1016/j.bpj.2009.02.072).
13. LEE A.K., LEWIS D.M., ASHMAN P.J. (2012), Disruption of microalgal cells for the extraction of lipids for biofuels: Processes and specific energy requirements, *Biomass and Bioenergy*, **46**: 89–101, doi: [10.1016/j.biombioe.2012.06.034](https://doi.org/10.1016/j.biombioe.2012.06.034).
14. LIU Y., LIU X., CUI Y., YUAN W. (2022), Ultrasound for microalgal cell disruption and product extraction: A review, *Ultrasonics Sonochemistry*, **87**: 106054, doi: [10.1016/j.ultsonch.2022.106054](https://doi.org/10.1016/j.ultsonch.2022.106054).
15. ONYEAKA H., MIRI T., OBILEKE K., HART A., ANUMUDU C., AL-SHARIFY Z.T. (2021), Minimizing carbon footprint via microalgae as a biological capture, *Carbon Capture Science Technology*, **1**: 100007, doi: [10.1016/j.ccst.2021.100007](https://doi.org/10.1016/j.ccst.2021.100007).
16. PEREIRA R.N., JAESCHKE D.P., MERCALI G.D., RECH R., MARCZAK L.D.F. (2023), Impact of ultrasound and electric fields on microalgae growth: A comprehensive review, *Brazilian Journal of Chemical Engineering*, **40**: 607–622, doi: [10.1007/s43153-022-00281-z](https://doi.org/10.1007/s43153-022-00281-z).
17. PITT W.G., ROSS S.A. (2003), Ultrasound increases the rate of bacterial cell growth, *Biotechnology Progress*, **19**(3): 1038–1044, doi: [10.1021/bp0340685](https://doi.org/10.1021/bp0340685).
18. REN H.Y. *et al.* (2019), Ultrasonic enhanced simultaneous algal lipid production and nutrients removal from non-sterile domestic wastewater, *Energy Conversion and Management*, **180**: 680–688, doi: [10.1016/j.enconman.2018.11.028](https://doi.org/10.1016/j.enconman.2018.11.028).
19. RUBSAI A. (2012), *Isolation, growth characters and oil accumulation of green colonial microalgae, botryococcus braunii*, MS Degree, Walailak University.
20. SALAEH A., BOONPHASUK S., DANWORAPHONG S. (2017), Expediting growth rate of *Botryococcus braunii* using 37- and 80-kHz ultrasonic waves, [in:] *Proceedings of 24th International Congress on Sound and Vibration 2017 (ICSV 24)*, pp. 1782.
21. SHAKIROV Z.S., KHALILOV I.M., KHUJAMSHUKUROV N.A. (2021), Stress factors' effects on the induction of lipid synthesis of microalgae, *Journal of Applied Biology and Biotechnology*, **9**(06): 149–153, doi: [10.7324/JABB.2021.96019-1](https://doi.org/10.7324/JABB.2021.96019-1).
22. TASIĆ M.B., PINTO L.F.R., KLEIN B.C., VELJKOVIĆ V.B., FILHO R.M. (2016), *Botryococcus braunii* for biodiesel production, *Renewable and Sustainable Energy Reviews*, **64**: 260–270, doi: [10.1016/j.rser.2016.06.009](https://doi.org/10.1016/j.rser.2016.06.009).
23. TOPAZ M. *et al.* (2005), Acoustic cavitation in phacoemulsification and the role of antioxidants, *Ultrasound in Medicine Biology*, **31**(8): 1123–1129, doi: [10.1016/j.ultrasmedbio.2005.02.016](https://doi.org/10.1016/j.ultrasmedbio.2005.02.016).
24. WANG M., YUAN W. (2016), Modeling bubble dynamics and radical kinetics in ultrasound induced microalgal cell disruption, *Ultrasonics Sonochemistry*, **28**: 7–14, doi: [10.1016/j.ultsonch.2015.06.025](https://doi.org/10.1016/j.ultsonch.2015.06.025).
25. WANG S.K., WANG F., STILES A.R., GUO C., LIU C.Z. (2014), *Botryococcus braunii* cells: Ultrasound-intensified outdoor cultivation integrated with in situ magnetic separation, *Bioresource Technology*, **167**: 376–382, doi: [10.1016/j.biortech.2014.06.028](https://doi.org/10.1016/j.biortech.2014.06.028).
26. XU L., WANG S., WANG, F., GUO C., LIU C.Z. (2014), Improved biomass and hydrocarbon productivity of *Botryococcus braunii* by periodic ultrasound stimulation, *BioEnergy Research*, **7**: 986–992, doi: [10.1007/s12155-014-9441-9](https://doi.org/10.1007/s12155-014-9441-9).
27. ZHANG L., LI B., WU Z., GU L., YANG Z. (2016), Changes in growth and photosynthesis of *Mixotrophic Ochromonas sp.* in response to different concentrations of glucose, *Journal of Applied Phycology*, **28**: 2671–2678, doi: [10.1007/s10811-016-0832-5](https://doi.org/10.1007/s10811-016-0832-5).



## Research Paper

## Ultrasonic Haptic Devices: Ultrasonic Noise Assessment

Leszek MORZYŃSKI<sup>ID</sup>, Marlena PODLEŚNA<sup>ID</sup>,  
Grzegorz SZCZEPAŃSKI<sup>ID</sup>, Anna WŁUDARCZYK<sup>ID</sup>

Central Institute For Labour Protection – National Research Institute  
Warsaw, Poland

\*Corresponding Author e-mail: [lmorzyns@ciop.pl](mailto:lmorzyns@ciop.pl)

(received March 23, 2023; accepted March 1, 2024; published online May 27, 2024)

Ultrasonic haptic technology is one of the more interesting novel technologies being intensively developed in recent years. Such technology has a number of undoubted advantages and potential applications, but it can also be a source of ultrasonic noise. Pursuant to the provisions of the labor law, ultrasonic noise at a high sound pressure level can be a harmful factor for human health. The article presents the results of the assessment of ultrasonic noise emitted by an ultrasonic haptic device and the assessment of exposure to noise of a person using the device. The tests were carried out using one of the haptic devices readily available on the market. Ultrasonic noise emission tests were carried out around the device, at selected points placed on the surface of a hemisphere of a radius of 0.5 m, for various haptic objects. The analyzed parameter was the equivalent sound pressure level in the  $1/3$  octave band with a center frequency of 40 kHz. Variable sound pressure levels ranged from 96 dB to 137 dB. Noise exposure tests were carried out both using the KEMAR measurement dummy and with test participants of different heights. In most cases, the sound pressure level exceeded 110 dB, and in the worst case it exceeded 131 dB. Comparison of the results of ultrasonic noise assessments with the permissible values of this noise in the working environment shows that in the case of prolonged or improper use of the device, the permissible values may be exceeded.

**Keywords:** ultrasonic haptic devices; ultrasonic noise.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The effective engagement of the sense of touch in the daily interactions with various interfaces has inspired many, leading to the incorporation of buzzers and various motors into popular devices, such as smartphones or wristwatches. The next step, that researchers have been trying to reach for years (IWAMOTO *et al.*, 2008), is the development of touchless haptic interfaces. Various solutions have been presented and one of the most promising ones is the use of ultrasound for the induction of a feeling of touch on human skin. The research on that topic has been ongoing since 1977 and begun with the presentation of the detection thresholds for human skin (GAVRILOV *et al.*, 1977). This knowledge was later used as a basis for excitation of the touch receptors in human skin using acoustic radiation pressure, first in water (DALECKI *et al.*, 1995), and later, after further dis-

coveries (CARTER *et al.*, 2013; HOSHI *et al.*, 2010), in air (RAKKOLAINEN *et al.*, 2021). In order to create detectable shapes in mid-air, the sound signal used should be of a high frequency, usually in the range of 40 kHz to 70 kHz (FRIER *et al.*, 2019). The signal should be modulated with the frequency of 0.4 Hz to 500 Hz, to allow for a detectable deflection of the skin by the touch receptors (GESHEIDER *et al.*, 2002). At the same time, it should be noted that the ultrasonic haptic technology is based on the use of ultrasound with high sound pressure levels (up to 150 dB at the focus point) and thus can be a source of ultrasonic noise (RADOSZ, PLEBAN, 2018; ŚLIWIŃSKI, 2016). Pursuant to the provisions of labor law in force, in Poland (Internet System of Legal Acts, 2018), ultrasonic noise is a harmful factor, because at sufficiently high sound pressure levels it has an adverse effect on the human body (SMAGOWSKA, PAWLACZYK-ŁUSZCZYŃSKA, 2013). Despite the safety concerns the research usually

focuses not on the exploration of threats, but rather possibilities that ultrasonic haptic technology can create. Since the inventions of new ways to use the technology influences the way in which it operates, it can also influence the specific needs for risk assessment. Research on the characteristics of generated signals can indicate which signal parameters or which generation types can be most commonly used in an industrial or commercial setting. A review of the literature shows current trends in the development and research topics concerning the ultrasonic haptic technology.

The latest studies and publications on the ultrasonic haptic technology can be mostly assigned to one of two categories: applications of the ultrasonic haptic technology or exploring issues related to the generation and perception of haptic signals. An interesting example of the use of ultrasonic haptic technology is presented in (ROMANUS *et al.*, 2019). The authors describe their device, which integrates three technologies: virtual reality goggles, ultrasonic haptic devices, and wearable devices. The haptic device was used to create a holographic, animated image of a heart, which can be felt with the sense of touch and which movement (beating) is synchronized with the heartbeat of the person operating the device. Another interesting proposal was to use the ultrasonic haptic technology in the process of hand-guided programming of collaborative industrial robots, as presented in (RIVERA PINTO *et al.*, 2020). This type of programming consists of manually guiding the robot arm by a human in the way, which later allows the robot to mimic the sequence of actions performed by the human. Such programming could minimize production line downtime and can be implemented with the use of virtual reality technologies. However, for a person performing such a task, the lack of feedback in the form of sensory stimuli is a major obstacle. In this work, virtual reality was enriched with sensory stimuli using the ultrasonic haptic technology and the use of virtual reality goggles can shield the user from potential adverse effects of ultrasonic noise or become a convenient mounting point for hearing protection.

Discussion of the most important of the anticipated applications of the ultrasonic haptic technology is also presented in the review by RAKKOLAINEN *et al.* (2021). The proposed applications are divided into groups covering: sterile medical interfaces, applications in the automotive industry, advertising and sales, and augmented virtual reality and mixed reality. The paper also reviews issues related to the creation of tactile objects with the use of ultrasonic haptic technology, their precision and perception. Issues related to the safety of the technology, resulting from the presence of ultrasonic waves with a high level of sound pressure (ultrasonic noise) were also discussed. The authors point out that even at a great distance from the focus of the ultrasound, its level may exceed 110 dB,

and further research into the impact of ultrasound on the hearing organ is necessary to fully assess this issue, although some recent studies (CARCAGNO *et al.*, 2019) have not demonstrated the impact of ultrasound with a frequency of 40 kHz to shift the threshold of hearing.

DI BATTISTA *et al.* (2022) focused on the burdensome, non-auditory impact of ultrasonic noise, which may occur when ultrasonic haptic technology is used in consumer devices. The conducted research concerned the impact of ultrasound with a frequency of 40 kHz and high sound pressure levels on the cognitive functions of the exposed persons. The conducted research showed no adverse effects of exposure to ultrasonic noise on the test subjects, assessed both by the number of correct answers given in the conducted tests as well as the reaction time. The authors state that ultrasounds with a frequency of 40 kHz and a level of 120 dB have no effect on human cognitive functions.

Taking into account the current progress in the development of ultrasonic haptic technology, ultrasonic haptic devices have a real chance to become tools used in everyday work. Because ultrasonic haptic technology can provide tactile sensations that provide feedback to actions taken, its potential future applications include workplaces related to control, design or diagnostics in virtual reality or augmented reality environments. Ultrasonic haptic technology can be especially valuable for people with visual impairments, helping to accommodate their needs in the workplace. In such cases, ultrasonic haptic devices would turn from technological gadgets into tools intended for long hours of work. This makes it all the more important to assess the technologies introduced to the market in terms of potential hazards to employees caused by the generated ultrasonic noise. The main focus of the studies presented in this article is the possible impact the ultrasonic noise can have on the persons using it on a daily basis. It presents the results of measurement and assessment of ultrasonic noise emitted by an ultrasonic haptic transducer, taking into account the criteria adopted for the assessment of ultrasonic noise in the work environment in Poland.

## 2. Method and experimental setup

Estimating exposure to ultrasonic noise generated by an ultrasonic haptic device is a difficult issue due to the multitude of factors that may affect the value of this exposure, most of which depend on how the device is used. Ultrasonic noise is significantly reduced by propagation in the air. Moreover, ultrasonic noise sources, in particular haptic transducers in which an ultrasonic beam with appropriate parameters is intentionally generated, are directional sources. This means that the exposure of a given person to ultrasonic noise will be influenced by the position of the person's head (especially their ears) in relation to the haptic

device, in terms of both the distance from the device and the angular position in relation to it. The position of the head of a person using an ultrasonic haptic device in relation to this device will be influenced by factors such as: the method of using the device (sitting or standing), the person's height, arm length, the height of the device in relation to the human body, the way the upper limb is positioned (straight, bent). Depending on these factors, the distance of the person's head from the haptic device will most often range from 50 cm to 90 cm. The sound pressure level of the ultrasonic noise produced by the device, as well as the directional characteristics of the noise radiation, will be influenced by factors such as: the type of the generated haptic object, its position in relation to the device and its intensity, and the parameters of the modulation used. It should be noted here that haptic devices may enable adjustment of the intensity of the generated object by adjusting the amplitude of the generated signals, thus reducing the sound pressure level of noise, but reduced intensity of the generated object deteriorates the tactile sensations felt by the user. Another factor that will influence the sound pressure level of noise reaching the ears of a person using a haptic device is the presence of a person's hand touching the generated haptic object and acting as a kind of acoustic screen for the ultrasonic wave. Obviously, due to the directional nature of the spread of ultrasonic noise, this is most important when a person's hand is aligned with both the person's ear and the haptic device. This situation will occur when the person is standing and the haptic device is placed low relative to the user. For a seated person and a haptic device placed higher, the signal shielding effect of the hand may be much smaller or negligible. For safety reasons, the optimal solution would be for the device to generate a haptic object when it detects the user's hand in the space above the ultrasonic matrix of the haptic transducer, which requires the use of a hand position or presence sensor in the device. However, device manufacturers are not

obliged to use this type of sensors, so it is possible that a person is close to a working haptic transducer but does not touch the object, thereby disabling the acoustic shielding effect. For example, as an office worker may use the keyboard as the main tool in their work, an ultrasonic haptic device could also be used in the same manner, yet neither of those tools would be operated constantly throughout the whole working day. Because of this it cannot be assumed that the hands of the user would shield them from potential harm at all times.

In our research, we attempted to assess how the type and parameters of the generated haptic object affect the sound pressure level of ultrasonic noise to which the user of the haptic device may be exposed, and the directionality of noise radiation (in particular in the direction in which the user of the device is located). These studies also allowed to assess the impact of the user's height and hand position on exposure to ultrasonic noise. For this purpose, three experiments described below were carried out, including measurements of sound pressure levels at selected points of a hemisphere with a radius of 50 cm, and measurements of ultrasonic noise near the ears of the user of the haptic device first by carrying out tests using a measuring dummy and then with participants.

Tests of ultrasonic noise emitted by an ultrasonic haptic device at selected points of the hemisphere were carried out in an acoustic test chamber characterized by a short reverberation time (semi-anechoic properties). The tests were carried out using the STRATOS Inspire haptic device, performing, for selected haptic objects, measurements of the equivalent sound pressure level in the  $1/3$  octave band with a center frequency of 40 kHz (which, according to previous studies, poses the greatest hazard to the users) at the selected points of the hemisphere with a radius of 50 cm, which is the closest assumed distance of the head of the user from the haptic device. The diagram of the measuring system is shown in Fig. 1.

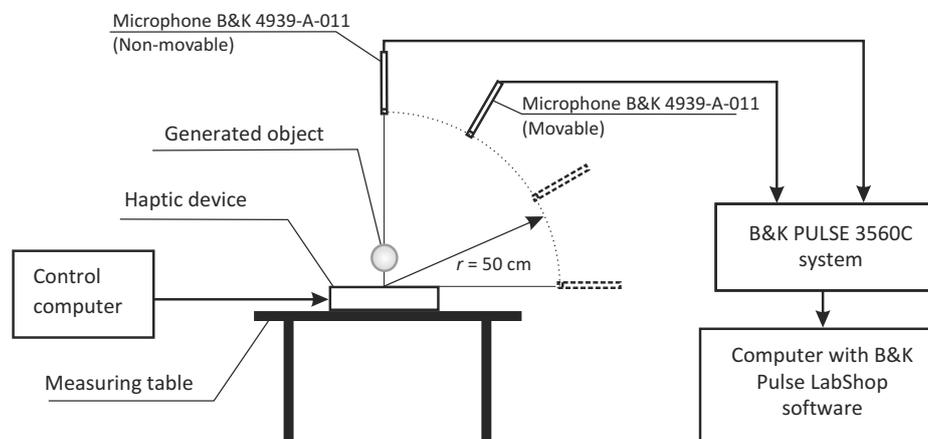


Fig. 1. Scheme of the measurement system for testing ultrasonic noise emitted by an ultrasonic haptic device.

The tested device was connected to the control computer and placed in the center of the test table, located at a distance of not less than 1.5 m from each of the walls of the room. The measurement system consisted of a Brüel & Kjær Pulse type 3560C measuring cassette with a type 3110 input/output module and two measuring microphones type 4939-A-11. The microphone marked in Fig. 1 as “movable” was the actual measurement microphone and was set up at individual measurement points of the hemisphere during the tests. The microphone marked in Fig. 1 as “non-movable” was permanently placed in the upper part of the hemisphere, perpendicular to the surface of the haptic device and performed a control function, allowing to assess the variability of the acoustic signal generated by the haptic device after each repositioning of the movable microphone at the next measurement point and restarting the device. Measurements made with a non-movable microphone made it possible to check whether each switching on and off of the device (e.g., when setting up a moving microphone) does or does not cause significant differences in the sound pressure levels of the ultrasonic noise produced and whether the operation of the device is stable (i.e., whether there are any unpredictable changes in the generated signal resulting, for example, from the applied control of the matrix of ultrasonic transducers). Thus, the fixed microphone made it possible to verify whether the measurements made with the moving microphone are not measurements of random values depending on the successive activations of the haptic device.

A computer with the Pulse LabShop software and the analysis module in frequency bands was used to control the measurements and record the measurement results. During the tests, the values of the equivalent sound pressure level in the frequency band with a center frequency of 40 kHz (averaging time 20 s) were recorded at measurement points in the space around the device, located on a hemisphere. The research focused mainly on the front of the device (angular mark  $0^\circ$ ), since this should be the position assumed by the person operating the device. In the frontal part of the hemisphere, the tests were carried out with a horizontal angular resolution of  $15^\circ$  in the range of  $\pm 45^\circ$ . In the horizontal plane, tests were also carried out for angles of  $\pm 90^\circ$  and  $\pm 180^\circ$ . In the vertical plane, the tests were performed with an angular resolution of  $15^\circ$ . The diagram of the distribution of measurement points in the conducted research is shown in Figs. 2 and 3. For each angular position of the microphones in the horizontal plane (Fig. 2), measurements were made in all angular positions in the vertical plane (Fig. 3), while in the vertical  $90^\circ$  position, measurements were made only with a fixed microphone.

During tests haptic objects were generated in the form of two flat shapes: a point and a circle. The radius of the generated circle was 2 cm. The modulation fre-

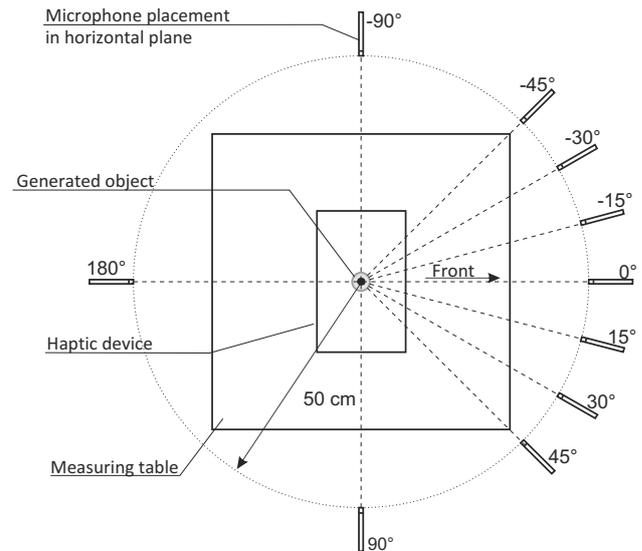


Fig. 2. Scheme of the arrangement of measurement points in the horizontal plane during testing of ultrasonic noise emitted by an ultrasonic haptic device.

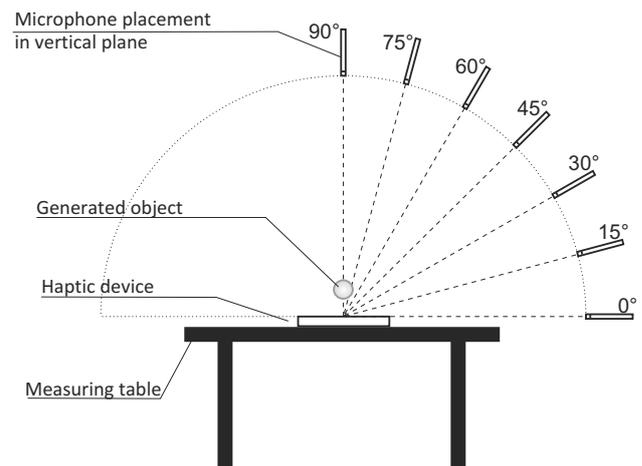


Fig. 3. Scheme of the distribution of measurement points in the vertical plane during measuring of ultrasonic noise emitted by an ultrasonic haptic device.

quency was 50 Hz or 200 Hz. The height of the generated objects above the haptic device (shape generation height,  $h_{sg}$ ) was 10 cm, 20 cm, and 30 cm.

Ultrasonic noise tests using a test dummy were carried out for the STRATOS Inspire haptic device generating selected haptic objects. Measurements of the equivalent sound pressure level were made in the  $1/3$  octave band with a center frequency of 40 kHz. The tested haptic device was connected to the control computer and placed in the center of the test table, located at a distance of not less than 1.5 m from each of the walls of the room. A GRAS KEMAR measuring dummy was placed in front of the test table, where a person using a haptic device would be positioned (Fig. 4). The height of the dummy setting was

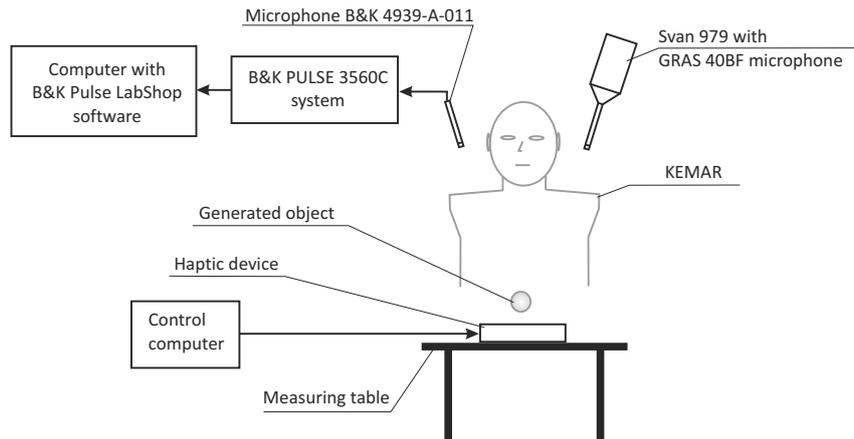


Fig. 4. Scheme of the measurement stand for testing ultrasonic noise with the use of the measuring dummy.

set to 168.5 cm (the average height of a person). Ultrasonic noise tests were carried out using two measurement systems, independently for each ear of the measuring dummy, which made it possible to assess the differences in exposure to ultrasonic noise between the right and left ear.

The first measurement system (measurement at the dummy's right ear) consisted of a Brüel & Kjær Pulse type 3560C measuring cassette with a type 3110 input/output module and a type 4939-A-11 measuring microphone. A computer with the Pulse LabShop software and the analysis module in frequency bands was used to control the measurements and record the measurement results. During the tests, the values of the equivalent sound pressure level in  $1/3$  octave bands (averaging time 20 s) were recorded. The second measurement system (measurement at the dummy's left ear) consisted of an integrating sound level meter SVAN 979

with a GRAS 40BF microphone and an SV17 preamplifier. During tests, the meter recorded the values of the equivalent sound pressure level in  $1/3$  octave bands (averaging time 20 s). In accordance with the ultrasonic noise measurement methodology (RADOSZ, 2012; 2020; RADOSZ, PLEBAN, 2018) and the provisions of the PN-Z-01339:2020-12 standard, the microphones of the measurement systems were placed 10 cm from the entrance to the outer ear canal of the dummy's appropriate ear and were directed towards the ultrasound source. The view of the test stand during the tests is shown in Fig. 5.

During tests, haptic objects were generated in the form of three flat shapes: a point, a circle and a square. The radius of the generated circle was 2 cm, and the side length of the generated square was 4 cm. The modulation frequency for all objects was 200 Hz and the generation height  $h_{sg}$  was 20 cm. The objects differed

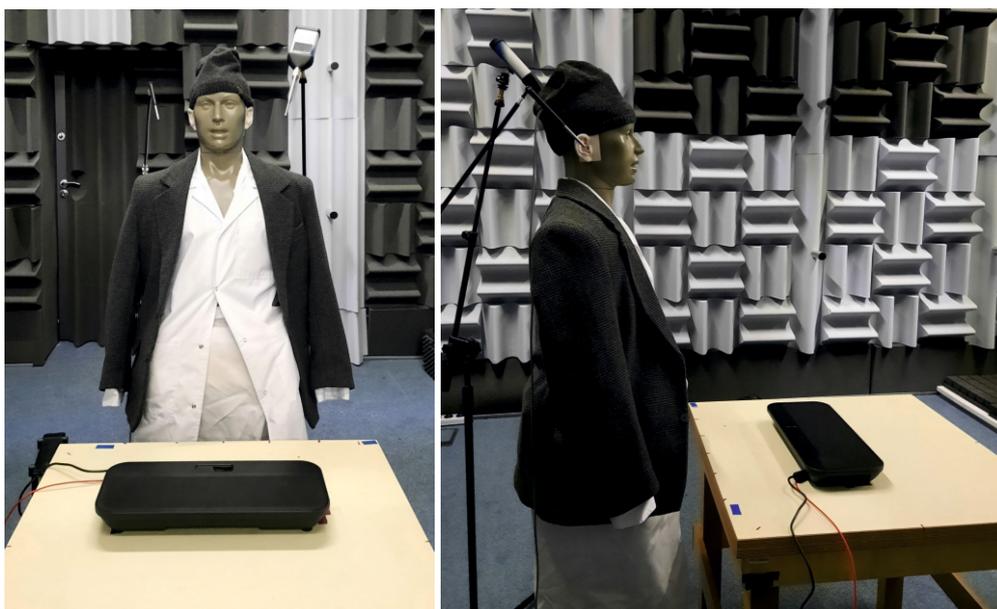


Fig. 5. View (front and side) of the ultrasonic noise testing stand.

in the type of modulation (amplitude – AM, spatio-temporal – SP).

Ultrasonic noise tests with participants for the STRATOS Inspire haptic device generating selected haptic objects were conducted with six participants of different heights, from 159 cm to 177 cm, which corresponds to the height of the measuring microphone in the range from 144 cm to 162 cm (PODLEŠNA *et al.*, 2022). These tests allow to assess the impact of the haptic device user’s height and the presence of his hand on the noise exposure.

The layout of the measuring stand (Fig. 6) was similar to that used for measurements using a measuring dummy, except that only one of the measuring

systems produced by Brüel & Kjær was used. The microphone of the measurement system was placed 10 cm from the entrance to the outer ear canal of the participant’s right ear and was directed towards the ultrasound source. Measurements of the equivalent sound pressure level (averaging time 20 s) were made in the 1/3 octave band with a center frequency of 40 kHz. The users were allowed to position themselves comfortably in front of the device. Measurements were made for two positions of the test participant’s right hand: along the body and while touching a haptic object.

During tests, haptic objects were generated in the form of three flat shapes: a point, a circle, and a square. The radius of the generated circle was 2 cm, and the side length of the generated square was 4 cm. The modulation frequency for all objects was 200 Hz. The objects were generated at different generation heights,  $h_{sg}$ : 10 cm, 20 cm, and 30 cm. All objects were generated in a spatio-temporal manner.

### 3. Results

The test results for a haptic object in the form of a point are presented in Tables 1 and 2, and in Figs. 7 and 8. The tests were carried out for a point located at a height of 20 cm and for two different modulation frequencies: 200 Hz (Table 1 and Fig. 7) and 50 Hz (Table 2 and Fig. 8).

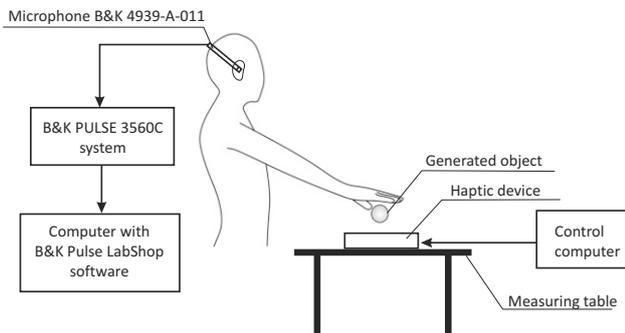


Fig. 6. Scheme of the measurement stand for testing ultrasonic noise with participants.

Table 1. Results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a point at a height of 20 cm (modulation frequency 200 Hz).

Vertical angular position [°]	Equivalent 1/3 octave band sound pressure level $L_{eq,40\text{ kHz}}$ [dB]									
	Horizontal angular position [°]									
	-90	-45	-30	-15	0	15	30	45	90	180
0	96	100	105	108	97	100	102	93	101	105
15	111	108	111	112	115	111	111	104	109	111
30	110	115	111	115	119	112	113	111	120	117
45	113	118	120	124	116	113	124	125	122	114
60	118	121	118	110	119	113	111	118	109	121
75	131	116	119	116	105	120	115	123	131	119
90	131									

Table 2. Results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a point at a height of 20 cm (modulation frequency 50 Hz).

Vertical angular position [°]	Equivalent 1/3 octave band sound pressure level $L_{eq,40\text{ kHz}}$ [dB]									
	Horizontal angular position [°]									
	-90	-45	-30	-15	0	15	30	45	90	180
0	97	102	105	107	97	100	102	87	101	105
15	114	106	111	111	115	110	111	104	108	111
30	110	115	110	114	120	113	114	111	120	117
45	113	119	121	124	116	112	124	126	122	114
60	118	120	119	111	119	113	113	118	108	121
75	131	115	119	116	108	120	114	123	131	119
90	131									

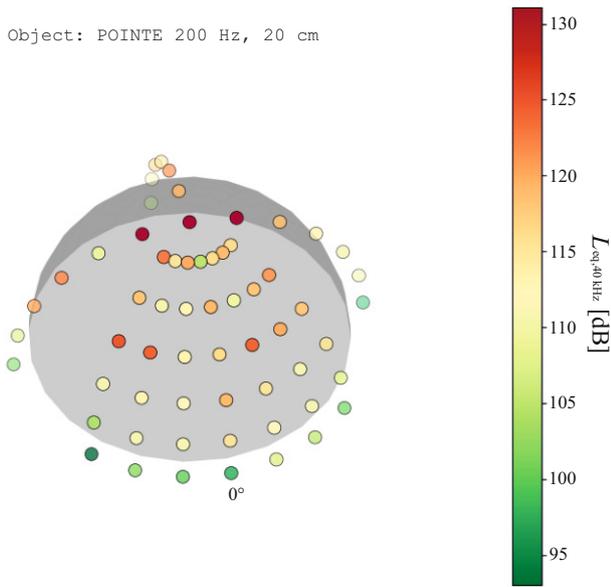


Fig. 7. Visualization of the results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a point at a height of 20 cm (modulation frequency 200 Hz).

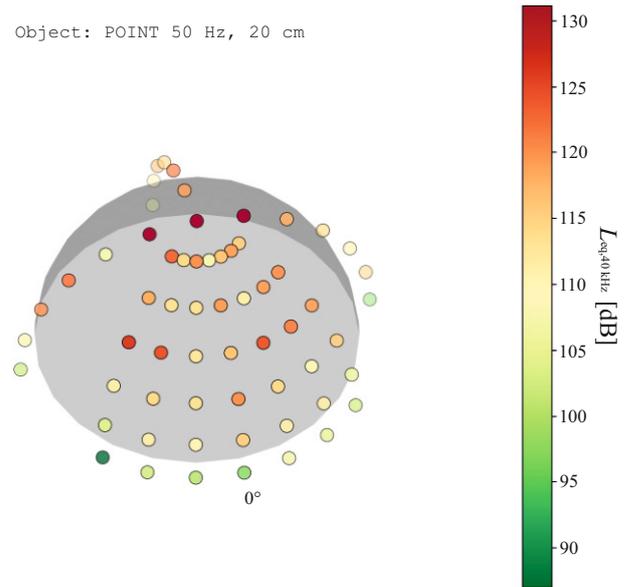


Fig. 8. Visualization of the results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a point at a height of 20 cm (modulation frequency 50 Hz).

The test results for a haptic object in the form of a circle are presented in Tables 3, 4, and 5 as well as in Figs. 9, 10, and 11. The tests were carried out for a circle with a radius of 2 cm and a modulation

frequency of 200 Hz and for three different heights of the circle above the surface of the haptic device: 10 cm (Table 3 and Fig. 9), 20 cm (Table 4 and Fig. 10), and 30 cm (Table 5 and Fig. 11).

Table 3. Results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 10 cm ( $r = 2$  cm, modulation frequency 200 Hz).

Vertical angular position [°]	Equivalent $1/3$ octave band sound pressure level $L_{eq,40\text{ kHz}}$ [dB]									
	Horizontal angular position [°]									
	-90	-45	-30	-15	0	15	30	45	90	180
0	101	111	107	107	111	105	109	110	106	108
15	114	115	117	118	115	116	118	115	120	118
30	121	118	122	122	122	124	124	120	122	121
45	125	122	125	122	124	129	125	126	124	131
60	128	130	131	125	123	127	124	130	126	126
75	126	129	129	128	127	131	130	130	127	126
90	124									

Table 4. Results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 20 cm ( $r = 2$  cm, modulation frequency 200 Hz).

Vertical angular position [°]	Equivalent $1/3$ octave band sound pressure level $L_{eq,40\text{ kHz}}$ [dB]									
	Horizontal angular position [°]									
	-90	-45	-30	-15	0	15	30	45	90	180
0	103	107	110	109	111	104	110	104	100	108
15	116	116	117	118	122	114	119	111	118	118
30	115	119	119	123	123	118	120	118	123	123
45	121	124	126	126	123	121	128	128	128	120
60	121	121	120	121	118	116	117	120	121	122
75	134	122	121	121	119	123	123	127	135	123
90	134									

Table 5. Results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 30 cm ( $r = 2$  cm, modulation frequency 200 Hz).

Vertical angular position [°]	Equivalent $1/3$ octave band sound pressure level $L_{eq,40\text{ kHz}}$ [dB]									
	Horizontal angular position [°]									
	-90	-45	-30	-15	0	15	30	45	90	180
0	104	104	104	114	109	108	102	104	103	107
15	120	116	111	125	117	118	117	110	121	114
30	115	120	123	128	124	115	117	118	121	123
45	121	125	124	119	119	116	127	119	122	114
60	116	118	115	118	113	116	119	112	122	116
75	125	116	120	124	115	122	122	123	126	125
90	137									

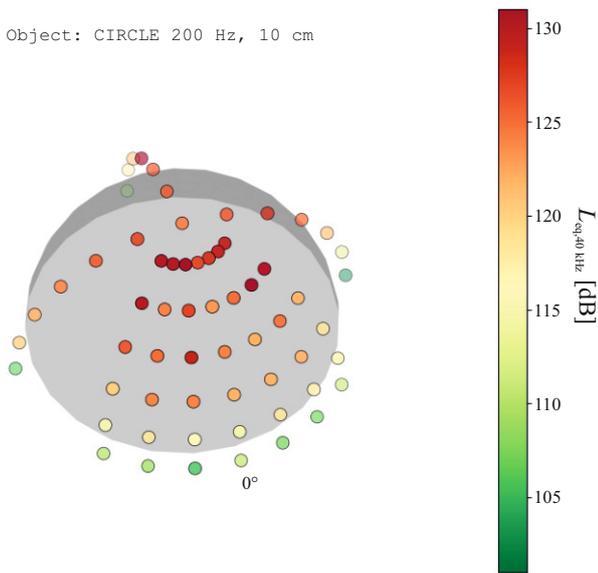


Fig. 9. Visualization of the results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 10 cm (modulation frequency 200 Hz).

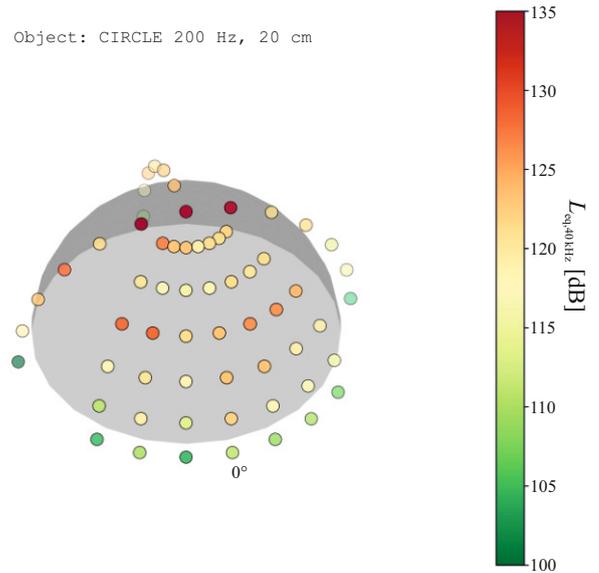


Fig. 10. Visualization of the results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 20 cm (modulation frequency 200 Hz).

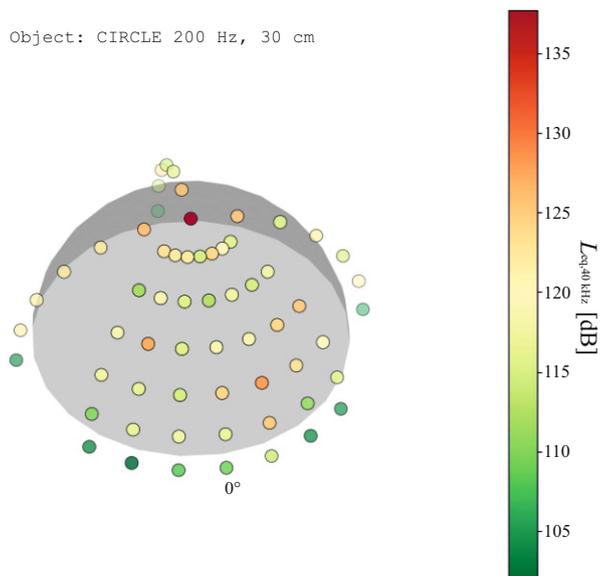


Fig. 11. Visualization of the results of ultrasonic noise tests at selected points of the hemisphere for an object in the form of a circle at a height of 30 cm ( $r = 2$  cm, modulation frequency 200 Hz).

The test results for a haptic object in the form of a point show that the highest equivalent sound pressure level of 131 dB was measured in the upper part of the hemisphere, where the fixed microphone was placed and at measurement points located in the upper part of the hemisphere for the horizontal angular positions  $-90^\circ$  and  $90^\circ$  and the vertical angular position  $75^\circ$ . High sound pressure levels (up to 126 dB) were also recorded in the front part of the hemisphere for a vertical angular position of  $45^\circ$ . The measurements results for a haptic object in the form of a circle show that the change in the generation height of the haptic object changes the angular position of the measurement points in which the highest sound pressure levels were recorded. In the case of a haptic object (circle) located at a height of 10 cm above the matrix of the haptic device, the highest sound pressure levels, reaching 131 dB, are observed for vertical angular positions of  $60^\circ$  and  $75^\circ$ . For a vertical angular position of  $90^\circ$ , the sound pressure level is lower and amounts to 124 dB. For a haptic object generated at a height of 30 cm, the highest sound pressure level, 137 dB, was recorded for a point at a vertical angle of  $90^\circ$ .

The results of the ultrasonic noise measurement with using KEMAR measuring dummy are included in Table 6.

Table 6. Results of ultrasonic noise tests for microphones placed at KEMAR's ear.

Haptic object			Svan	B&K
Shape	Modulation type	Location height [cm]	$L_{eq,40\text{ kHz}}$ [dB]	$L_{eq,40\text{ kHz}}$ [dB]
Point	AM	20	114.1	114
	ST	20	112.3	112
	AM	10	123.6	117
	AM	30	113.4	101
Square	AM	10	117.5	117
	AM	20	113.8	112
	AM	30	114.5	112
Circle	ST	10	126.4	122
	ST	20	117.7	116
	ST	30	117.5	111

The measurements results presented in Table 6 show that the equivalent sound pressure level in the frequency band with a center frequency of 40 kHz near the dummy's ears exceeded 110 dB in each case, with the highest value recorded being 126.4 dB. Differences between sound pressure levels for the right and left ear can be several dB.

The results of the ultrasonic noise measurements with participants of different heights were presented in detail in a previously published work (PODLEŚNA *et al.*, 2022). This article presents only additional synthesis and analysis of the measurement results obtained for the STRATOS Inspire device and the con-

clusions drawn from them. The results of measurements of the equivalent sound pressure level of ultrasonic noise in relation to the height of the measurement microphone, resulting from the height of the research participant are presented in Fig. 12. In the graphs presenting the test results in relation to the height of the measuring microphone, trend lines have been added for the measurement results obtained for the object generation height of 20 cm.

Test results show an impact of the presence of the user's hand above the device. The equivalent sound pressure level has never exceeded 130 dB while the user's hand was extended. That was the case however for 3 measurements while the user's hand was withdrawn. The highest equivalent sound pressure levels were measured in cases where the object generation height was 10 cm. This phenomenon can be explained based on the results of tests carried out on the hemisphere. When the haptic object is generated at a higher height, much of the acoustic energy is emitted upwards, above the haptic device (Fig. 8). When generating an acoustic object at a lower height (Fig. 9), a large part of the acoustic energy is emitted at smaller vertical angles, towards the user of the device. Out of 54 measurement cases the equivalent sound pressure level exceeded 110 dB in 48 cases (89 %) while the user's hand was touching haptic object and in all cases (100 %) while the user's hand was positioned along the body. In most cases, a hand placed on a haptic object reduced the noise by several dB (approx. 4 dB on average), but in extreme cases it was more than 20 dB. Such a large reduction concerned haptic objects generated at low heights. The height of the user has a significant impact on the test results, but this effect varied depending on the position of the research participant's hand. Considering the trend lines in Fig. 12, when a participant's hand touches a haptic object, the equivalent sound pressure level at the participant's ear was lower for taller participants. The reverse relationship can be observed when the research participant holds their hand along the body. These results indicate that when the ultrasound emitted from the haptic device is not shielded by the user's hand, the angular position of the user's head in relation to the transducer is more important than the distance of the user's head from the device.

#### 4. Assessment of ultrasonic noise exposure and conclusions

The assessment of the exposure of the user of an ultrasonic haptic device to ultrasonic noise and the resulting risks should be carried out using appropriate criteria. In order to assess the risk posed by ultrasonic noise generated by an ultrasonic haptic device, the results of the conducted study were compared with

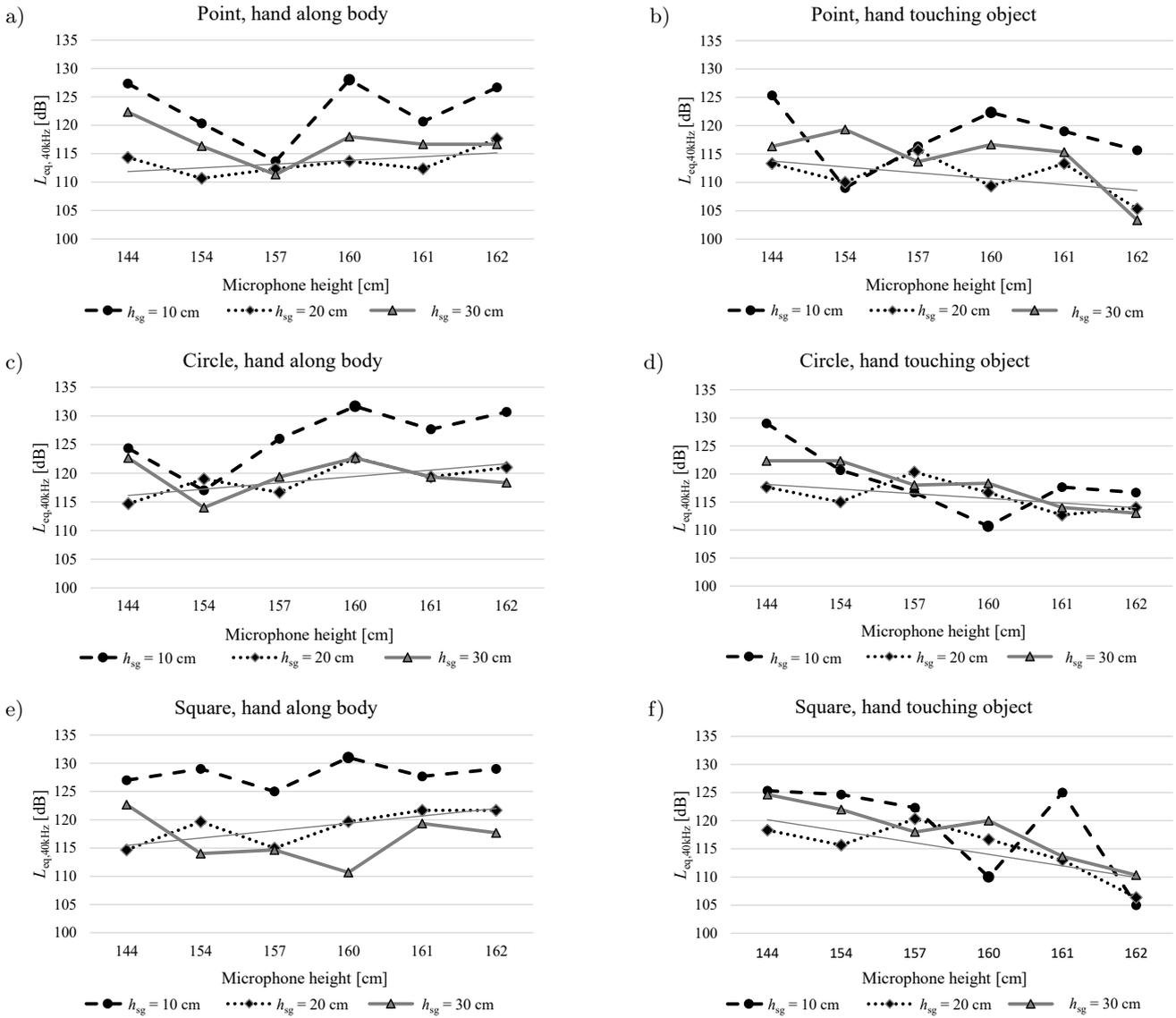


Fig. 12. Results of ultrasonic noise measurements for different haptic objects, object generation heights, participants heights, and their hands placed along body (on the left) and touching the haptic object (on the right). The solid gray line is the trend line for the test results for an object generation height of 20 cm.

the ultrasonic noise limit values applicable in Poland (Internet System of Legal Acts, 2018). The quantities characterizing ultrasonic noise in the working environment are:

- equivalent sound pressure levels in 1/3 octave bands with central frequencies ranging from 10 kHz to 40 kHz related to the 8-hour daily

or weekly average working time specified in the Labour Code;

- maximum sound pressure levels in 1/3 octave bands with center frequencies from 10 kHz to 40 kHz.

The permissible values of ultrasonic noise are presented in Table 7.

Table 7. Permissible values of ultrasonic noise in Poland.

The center frequency of the 1/3 octave band [kHz]	Permissible equivalent sound pressure level related to the 8-hour daily or average weekly working time specified in the Labour Code [dB]	Permissible maximum sound pressure level [dB]
10, 12.5, 16	80	100
20	90	110
25	105	125
31.5, 40	110	130

The equivalent sound pressure level in the  $i$ -th  $1/3$  octave band, related to an 8-hour daily working time,  $L_{fi,eq,8h}$  is determined based on the equation:

$$L_{fi,eq,8h} = L_{fi,eq,T_e} + 10 \log \frac{T_e}{T_0}, \quad (1)$$

where  $L_{fi,eq,T_e}$  is the equivalent sound pressure level in the  $i$ -th  $1/3$  octave band determined for the total exposure time  $T_e$ , and  $T_0$  is the reference time of 8 h.

The results of the research of ultrasonic noise affecting the user of the haptic transducer showed that the values of the equivalent sound pressure level in the  $1/3$  octave band with a frequency of 40 kHz, recorded at the ear of the person using the haptic device, exceeded 110 dB in most of the conducted tests, approaching 131 dB in the worst cases. Taking into account Eq. (1), the test results indicate that with long-term, daily use of the haptic device, the permissible value of ultrasonic noise may be exceeded. If the equivalent sound pressure level exceeds 110 dB, the duration of daily use of the haptic device should be less than 8 hours. Each 3 dB increase in the equivalent sound pressure level above 110 dB means that the operating time of the haptic device must be halved in order to ensure that the permissible value of ultrasonic noise is not exceeded. Obviously, as shown in the presented research, factors influencing the amount of exposure include, among others, the type of generated object and its location, as well as the method of using the device, including the position of the user's hand. For this reason, without precisely defining the application of the transducer and how it is used, estimating the exposure to ultrasonic noise and the duration of use that will not result in exceeding the permissible value of ultrasonic noise is very difficult.

Ultrasonic noise emission tests carried out in selected points of the hemisphere with a radius of 50 cm around the transducer show that, depending on the type of the generated touch object, in some points of this hemisphere the values of the sound pressure level exceed 130 dB, sometimes reaching 137 dB. Equivalent sound pressure levels above 130 dB were also recorded in three cases of measurements with participants. This means that if the user's head is too close to the haptic transducer, e.g., as a result of the user leaning over the transducer they are using, the ultrasonic noise limit values specified for the maximum sound pressure level in the  $1/3$  octave band with a center frequency of 40 kHz will be exceeded.

Test results indicate that for low-lying tactile objects the acoustic energy is radiated at wider angles around the transducer. The noise exposure of the test subjects was significantly higher for objects generated at a height of 10 cm than for objects generated above that height. According to the obtained results, for low-lying tactile objects, the acoustic energy will be emitted towards the head of the person standing next to the transducer. From the point of view of protecting

the employee against ultrasonic noise, it may be advantageous to use the transducer in a sitting position, so that the user's head is at a low vertical angular height relative to the haptic device, as well as avoiding generating low-lying objects.

To conclude, ultrasonic haptic technology piques the interest of many hoping to enhance the quality of human-machine interfaces and bring the dreams of the future to present day. However, many are also wary of the technology, as it employs ultrasound of high pressure levels. Appropriate studies are being conducted in order to properly assess the risk that such technology may pose to its users, including the presented paper. Presented results suggest that in order to safely use ultrasonic haptic technology, especially in work environment, specific guidelines should be created and implemented. Furthermore, additional studies should be conducted, including a wider range of devices, and a wider range of types of usage and environments, where such devices could be implemented. Considering the high sound pressure levels of ultrasonic noise emitted by haptic devices, research into the development of noise reduction measures to enable safer use of this new technology also appears necessary.

### Acknowledgments

This paper is published and based on the results of a research task carried out within the scope of the fifth stage of the National Programme "Improvement of safety and working conditions" supported from the resources of the National Centre for Research and Development – task no. II.PB.13 entitled "Development and testing of the properties of ultrasonic haptic technology with special taking into account the possibility of its application for the needs of disabled people". The Central Institute for Labour Protection – National Research Institute is the Programme's main coordinator.

### References

1. CARCAGNO S., DI BATTISTA A., PLACK C.J. (2019), Effects of high-intensity airborne ultrasound exposure on behavioural and electrophysiological measures of auditory function, *Acta Acustica United with Acustica*, **105**(6): 1183–1197, doi: [10.3813/AAA.919395](https://doi.org/10.3813/AAA.919395).
2. CARTER T., SEAH S. A., LONG B., DRINKWATER B., SUBRAMANIAN S. (2013), Ultrahaptics: Multi-point mid-airhaptic feedback for touch surfaces, [in:] *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, pp. 505–514, doi: [10.1145/2501988.2502018](https://doi.org/10.1145/2501988.2502018).
3. DALECKI D., CHILD S.Z., RAEMAN C.H., CARSTENSEN E.L. (1995), Tactile perception of ultrasound, *The Journal of the Acoustical Society of America*, **97**(5): 3165–3170, doi: [10.1121/1.411877](https://doi.org/10.1121/1.411877).

4. DI BATTISTA A., PRICE A., MALKIN R., DRINKWATER B.W., KUBERKA P., JARROLD C. (2022), The effects of high-intensity 40 kHz ultrasound on cognitive function, *Applied Acoustics*, **200**: 109051, doi: [10.1016/j.apacoust.2022.109051](https://doi.org/10.1016/j.apacoust.2022.109051).
5. FRIER W., PITTEA D., ABLART D., OBRIST M., SUBRAMANIAN S. (2019), Sampling strategy for ultrasonic mid-air haptics, [in:] *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, doi: [10.1145/3290605.3300351](https://doi.org/10.1145/3290605.3300351).
6. GAVRILOV L.R., GERSUNI G.V., ILYINSKI O.B., TSIRULNIKOV E.M., SHCHEKANOV E.E. (1977), A study of reception with the use of focused ultrasound. I. Effects on the skin and deep receptor structures in man, *Brain Research*, **135**(2): 265–277, doi: [10.1016/0006-8993\(77\)91030-7](https://doi.org/10.1016/0006-8993(77)91030-7).
7. GESHEIDER G.A., BOLANOWSKI S.J., POPE J.V., VERRILLO R.T. (2002), A four-channel analysis of the tactile sensitivity of the fingertip: Frequency selectivity, spatial summation, and temporal summation, *Somatosensory & Motor Research*, **19**(2): 114–124, doi: [10.1080/08990220220131505](https://doi.org/10.1080/08990220220131505).
8. HOSHI T., TAKAHASHI M., IWAMOTO T., ANDSHINODA H. (2010), Noncontact tactile display based on radiation pressure of airborne ultrasound, *IEEE Transactions on Haptics*, **3**(3): 155–165, doi: [10.1109/TOH.2010.4](https://doi.org/10.1109/TOH.2010.4).
9. Internet System of Legal Acts [in Polish: Internetowy System Aktów Prawnych] (2018), *Regulation of the Minister of Family, Labour and Social Policy of 12th June 2018 on the Maximum Admissible Concentrations and Intensities of Harmful to Health Agents in the Working Environment*, Dz. U. No 1286, Sejm RP.
10. IWAMOTO T., TATEZONO M., SHINODA H. (2008), Non-contact method for producing tactile sensation using airborne ultrasound, *Haptics: Perception, Devices and Scenarios. EuroHaptics 2008. Lecture Notes in Computer Science*, **5024**: 504–513, doi: [10.1007/978-3-540-69057-3\\_64](https://doi.org/10.1007/978-3-540-69057-3_64).
11. PODLEŚNA M., MORZYŃSKI L., ŁADA K., WŁUDARCZYK A., KWARTNIK M. (2022), Ultrasonic haptic devices – exploration of possibilities and threats while applied in workspace environment, [in:] *Proceedings of the 28th International Congress on Sound and Vibration, ICSV28*.
12. RADOSZ J. (2012), Methodology issues of ultrasonic noise exposure assessment, *Noise Control Engineering Journal*, **60**(6): 645–654, doi: [10.3397/1.3701038](https://doi.org/10.3397/1.3701038).
13. RADOSZ J. (2020), Ultrasonic noise measurements in the work environment, [in:] *Occupational Noise and Workplace Acoustics – Advances in Measurement and Assessment Techniques*, Pleban D. [Ed.], CRC Press, Taylor & Francis Group.
14. RADOSZ J., PLEBAN D. (2018), Ultrasonic noise measurements in the work environment, *The Journal of the Acoustical Society of America*, **144**(2): 2532–2538, doi: [10.1121/1.5063812](https://doi.org/10.1121/1.5063812).
15. RAKKOLAINEN I., FREEMAN E., SAND A., RAISAMO R., BREWSTER S. (2021), A survey of mid-air ultrasound haptics and its applications, *IEEE Transactions on Haptics*, **14**(1): 2–19, doi: [10.1109/TOH.2020.3018754](https://doi.org/10.1109/TOH.2020.3018754).
16. RIVERA PINTO A., KIDAL J., LAZKANO E. (2020), Multimodal mixed reality impact on a hand guiding task with a holographic cobot, *Multimodal Technologies and Interaction*, **4**: 78, doi: [10.3390/mti4040078](https://doi.org/10.3390/mti4040078).
17. ROMANUS T., MAKSYMENKO M., FRIER W., CORENTHY L., GEORGIU O. (2019), Mid-air haptic bioholograms in mixed reality, [in:] *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, doi: [10.1109/ISMAR-Adjunct.2019.00-14](https://doi.org/10.1109/ISMAR-Adjunct.2019.00-14).
18. ŚLIWIŃSKI A. (2016), On the noise hazard assessment within the intermediate range of the high audible and the low ultrasonic frequencies, *Archives of Acoustics*, **41**(2): 331–338, doi: [10.1515/aoa-2016-0034](https://doi.org/10.1515/aoa-2016-0034).
19. SMAGOWSKA B., PAWLACZYK-ŁUSZCZYŃSKA M. (2013), Effects of action of ultrasonic noise on the human body – A bibliographic review, *International Journal of Occupational Safety and Ergonomics*, **19**(2): 195–202, doi: [10.1080/10803548.2013.11076978](https://doi.org/10.1080/10803548.2013.11076978).

## Research Paper

Drone Flight Detection at an Entrance  
to a Beehive Based on Audio SignalsUrszula LIBAL\*<sup>ID</sup>, Pawel BIERNACKI<sup>ID</sup>

Faculty of Electronics, Photonics and Microsystems  
Department of Acoustics, Multimedia and Signal Processing  
Wrocław University of Science and Technology  
Wrocław, Poland; e-mail: pawel.biernacki@pwr.edu.pl

\*Corresponding Author e-mail: [urszula.libal@pwr.edu.pl](mailto:urszula.libal@pwr.edu.pl)

(received June 28, 2023; accepted March 5, 2024; published online May 28, 2024)

Spotting a significant number of drones flying near the entrance of a beehive during late Spring could indicate the occurrence of swarming mood, as the surge in drone presence is related to an overcrowded hive. Swarming refers to a natural reproductive process witnessed in honey bees, wherein half of the bee colony departs from their hive alongside the aging queen. In the paper, we propose an early swarming detection mechanism that relies on the behavior of the drones. The proposed method is based on audio signals registered in a close proximity to the beehive entrance. A comparative study was performed to find the most effective preprocessing method for the audio signals for the detection problem. We have compared the results for three different power spectrum density coefficients estimation methods, which are used as an input of an autoencoder neural network to discriminate drones from worker bees. Through simulations employing real-life signals, it has been demonstrated that drone detection based solely on audio signals is indeed feasible. The attained level of detection accuracy enables the creation of an efficient alarm system for beekeepers.

**Keywords:** signal processing; machine learning; neural networks; anomaly detection; autoencoder; honey bee swarming; drone detection.



Copyright © 2024 The Author(s).  
This work is licensed under the Creative Commons Attribution 4.0 International CC BY 4.0  
(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Swarming is a natural phenomenon that occurs when a honey bee colony reproduces and divides into multiple colonies. Swarming typically occurs during the late spring and early summer months (WRIGHT, 1913; OSTROWSKA, 1980). Early detection of swarming (ZGANK, 2011; HADJUR *et al.*, 2022) in honey bees is essential for swarming prevention, colony health monitoring, queen management, swarm capture, and effective population management. It allows beekeepers to take timely actions to maintain healthy colonies, prevent the loss of bees, and optimize honey production.

There are several ways to predict when honey bees will swarm. Most of them require interference in the hive. Here are a few:

- queen cells present,
- the old age of the queen,

- a hive becomes too crowded,
- increased foraging activity, more drones coming and going.

Thus, the detection of swarming without interfering in the hive can be based on the detection of drones' activity around the hive. This can be accomplished by analyzing the sounds around the hive and identifying drones.

Honey bees use sound as means of communication, both within the hive and with other bees outside the hive. The sound produced by bees is a form of vibration created by the rapid beating of their wings and is used to convey information about the location of food, the presence of danger, and other important information. One of the most well-known sounds produced by bees is the buzzing sound that is heard when they are in flight. This sound is created by the rapid beating of their wings, which can occur at a rate

of up to 200 beats per second. The frequency of the buzzing sound can vary depending on the size and species of the bee (KAWAKITA, ICHIKAWA, 2019). The typical range of frequencies generated by different bees is piping signal in the range: 100 Hz–500 Hz (SEELEY, TAUTZ, 2001), with a fundamental frequency of 384 Hz (SARMA *et al.*, 2002), tooting: 400 Hz–500 Hz, and quacking: around 350 Hz.

Distinguishing between bees and drones, based on the sounds they make, can be done using the fact that the drones are bigger, have bigger wings. The results showed that body shape or wing size can be correlated with fundamental frequency (GRADIŠEK *et al.*, 2017), and the duration of the buzzes has also been shown to differ with body size (larger bees producing shorter buzzes). Moreover, using the amplitude, frequency, and duration of flight, one can distinguish between bees and drones by analyzing the frequency spectrum of their sounds (RIBEIRO *et al.*, 2021).

This paper is divided into the following sections. Section 1 briefly provides information why distinguishing between worker bees and drones is important to control beehive environment and its health. Section 2 presents related works on the bee sound detection and classification. In Sec. 3 the methodology of the proposed solution is discussed. We focus on feature extraction and the implementation of machine learning techniques. Section 4 discusses the identification results based on the collected datasets. The final conclusions are in Sec. 5.

## 2. Related work

### 2.1. Audio analysis methods

Mel-frequency cepstral coefficients (MFCCs) (DAVIS, MERMELSTEIN, 1980; SOARES *et al.*, 2022; LIBAL, BIERNACKI, 2024) are the most common set of features used in studies that exploited a machine learning framework. Many studies analyzed MFCCs to extract information for bee detection, queen absence and swarming detection, and bee species identification, as well as environmental effects, with the three first coefficients showing the greatest discrimination. PENG *et al.* (2020) used the so-called improved MFCC (IMFCC) proposed in (YEGNANARAYANA *et al.*, 2005) to capture additional information from the higher-frequency part of the spectrum that is typically ignored by traditional MFCC. This has been shown to improve the classification accuracy for queenless hive detection tasks. In (ZLATKOVA *et al.*, 2020), the short-time Fourier transform (STFT) calculated with filter banks and the overlapping method was used to detect swarming events. The STFT has been calculated using 128, 256, 512, and 1024 bins to investigate the impact of window width. In the study (GOURISARIA *et al.*, 2024), the MFCC approach was compared with the STFT.

### 2.2. Machine learning algorithms

A typical machine learning framework encompasses signal measurement, preprocessing, feature extraction, and finally classification. In the area of acoustic analysis of bee sounds, many different classifiers have been explored. The most common classifier is a support vector machine (SVM) (CEJROWSKI *et al.*, 2020; NOLASCO *et al.*, 2018), a kernel-based method that projects data into higher dimensions in which a hyperplane can separate the classes.

More recent deep learning neural network-based methods are being introduced. In (RUVINGA *et al.*, 2021; KULYUKIN *et al.*, 2018) the use of so-called long short-term memory (LSTM) recurrent neural networks (RNN) for the queen bee presence detection is proposed. A comparison between LSTM, a multi-layer perceptron (MLP) neural network, and logistic regression was made, and it showed the power of the LSTM for the task at hand. Recently, convolutional neural networks (CNN) (NOLASCO *et al.*, 2018; KIM *et al.*, 2021), have gained popularity, especially within computer vision tasks. To make them directly applicable to the bee acoustics analysis, researchers have relied on image-like inputs, such as spectrograms, mel-scaled spectrograms, or other two-dimensional time-frequency representations of the audio signals.

## 3. Methods

Presented in the article results based on a selected set of audio recordings acquired in the context of the beehive monitoring system capable of identifying and predicting certain events and states of the hive that are of interest to the beekeeper. All recordings are sampled with frequency of 44100 Hz and saved in WAV format without any compression. For signal processing the recordings were divided into 1 second long samples. The data set used for simulations consists of 10000 samples of bee flight sound and around 1700 samples of a flying drone sound. To record the audio samples, we used a directional microphone mounted on top of the hive and aimed at its entrance.

The whole detection process is divided into two parts: feature extraction and classification by autoencoder neural network. The signal processing flow chart is shown in Fig. 1.

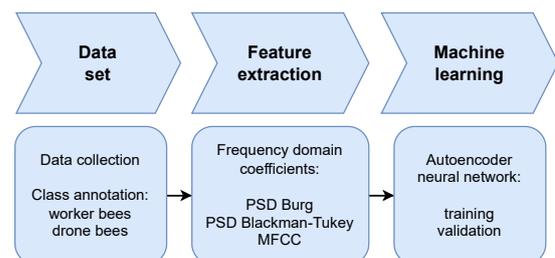


Fig. 1. Signal processing flow chart.

### 3.1. Power spectrum density coefficients

Power spectrum coefficients are the type of features that can be extracted from an audio signal to characterize its spectral content. They are calculated by taking the squared magnitude of the Fourier transform of the signal, which represents the power or energy content of the signal at each frequency component. The power spectrum coefficients can then be used as a feature vector to identify or classify different types of audio signals. We decided to employ power spectrum density (PSD) coefficients as features used in the detection phase of the identification.

Power spectral density estimation techniques can be divided into parametric and nonparametric methods. The non-parametric methods estimate PSD explicitly from signal samples, without any assumptions about a particular process structure. Parametric approaches assume that the signal can be described as the stationary process (MA – moving average, AR – autoregressive, or ARMA – autoregressive moving average) of the order  $m$ . The power spectral density is then calculated using estimated model parameters. This paper presents PSD estimated with the parametric approaches (the Burg method) and nonparametric methods (the Blackman-Tukey method).

#### 3.1.1. Burg algorithm

The Burg algorithm (KAY, 1988; ORFANIDIS, 1995) assumes that a signal can be described as an autoregressive (AR) process of the order  $m$ :

$$\hat{x} = - \sum_{k=1}^m a_m(k)x(n-k). \quad (1)$$

The Burg algorithm solves the ordinary least squares problem. AR parameters  $a_m$  are estimated by minimizing the prediction forward and backward errors which are referred to as the error between the actual value signal and the corresponding estimators in forward and backward:

$$\text{PSD}_x^{\text{Burg}}(f) = \frac{E_m}{\left| 1 + \sum_{k=1}^m a_m(k)e^{-j2\pi fk} \right|^2}. \quad (2)$$

#### 3.1.2. Blackman-Tukey method

The Blackman-Tukey (BLACKMAN, TUKEY, 1958; COOLEY, 1997) power spectrum estimate is calculated with the use of the fast Fourier transform (FFT) in the following way:

$$\text{PSD}_x^{B-T}(f) = |\text{FFT}\{w(n) * R(n)\}|, \quad (3)$$

where  $w(n)$  is a window,  $R(n)$  is an autocorrelation of the input signal  $x(n)$ .

The signal processing scheme for the Blackman-Tukey estimation method of PSD is shown in Fig. 2.

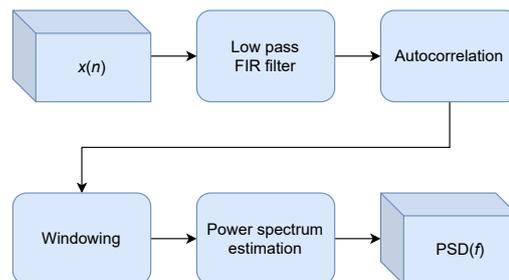


Fig. 2. Schema of Blackman-Tukey method of power spectrum estimation.

The lowpass FIR filter is used to adjust the bandwidth of the signal to investigate its influence on identification. Filter coefficients were changed to obtain the desired filter characteristic. Power spectrum estimation requires the Fourier transform calculation. To minimize leakage of spectrum a windowing procedure was implemented. Different windows were investigated (Hanning, Hamming, Kaiser).

During trait extraction, we noticed some differences in PSD coefficients to distinguish bees from drones. It can be observed in Fig. 3.

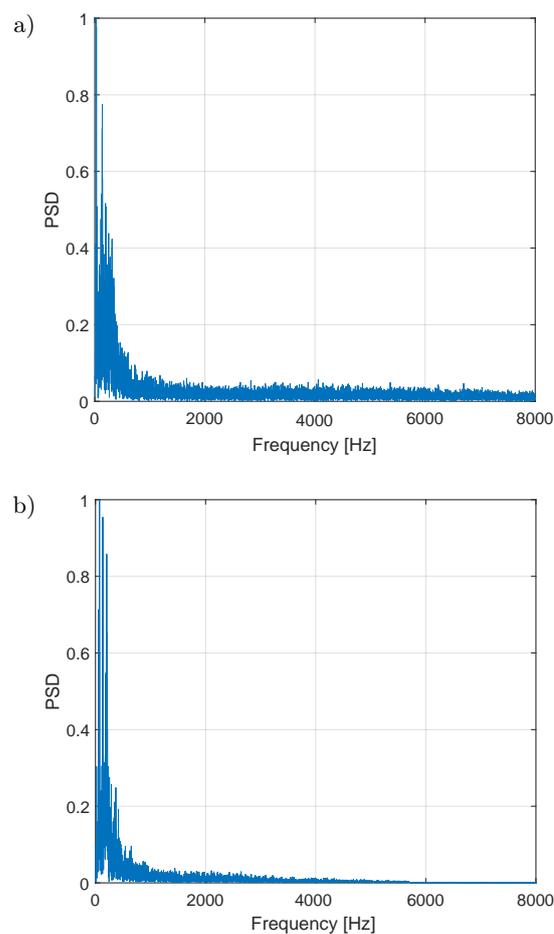


Fig. 3. PSD: a) worker bees; b) drones.

### 3.2. Mel-frequency cepstral coefficient

The motivating idea of MFCCs is to compress information about the vocal tract (smoothed spectrum) into a small number of coefficients based on an understanding of the cochlea in the ear. The basic steps to calculate MFCC are shown in Fig. 4.

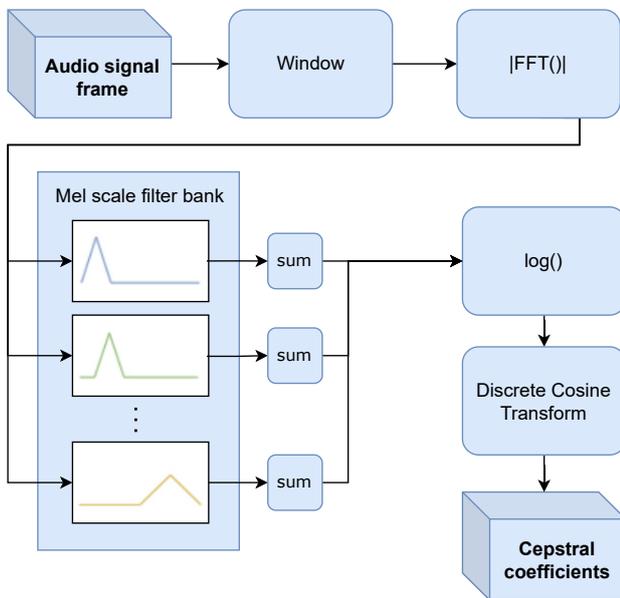


Fig. 4. MFCC calculation diagram.

The extraction of cepstral coefficients allowed for noticing some differences in MFCCs distinguishing worker bees from drones. It can be observed in Fig. 5 and in the calculated difference in Fig. 6.

### 3.3. Autoencoder

An autoencoder neural network is a type of artificial neural network that is used for unsupervised learning of efficient data representations. The network consists of an encoder and a decoder, where the encoder maps the input data to a compressed representation (HINTON, SALAKHUTDINOV, 2006), and the decoder maps the compressed representation back to the original data. The objective of the autoencoder is to minimize the difference between the input and output data, while also enforcing a constraint on the dimensionality of the compressed representation.

In audio signal identification, autoencoder neural networks can be used to learn compact representations of audio signals that capture their essential features. This can be useful for tasks such as pattern recognition, classification, identification, anomaly detection or noise reduction. Autoencoder neural networks have been used in a variety of applications in speech recognition, speaker identification or music genre classification. By training an autoencoder on a large dataset of audio signals, the network can learn to extract features

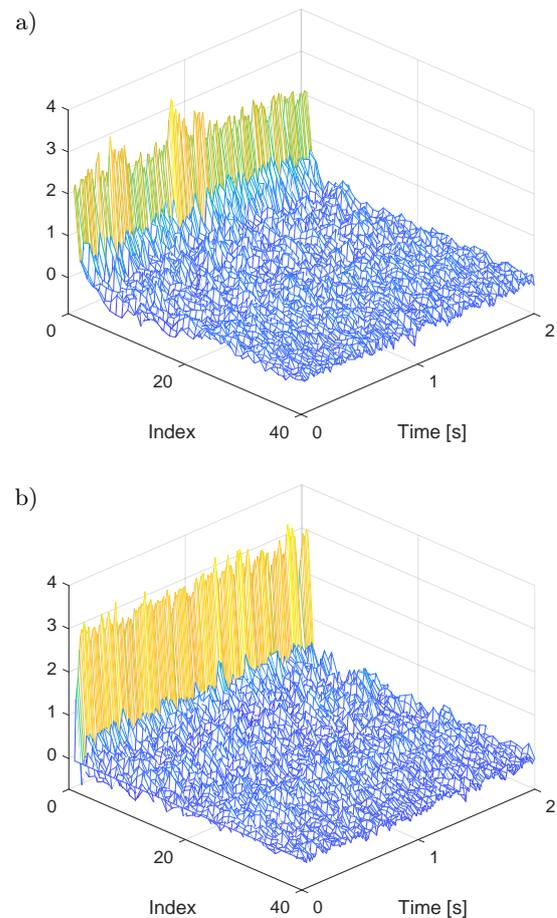


Fig. 5. MFCCs: a) worker bees; b) drones.

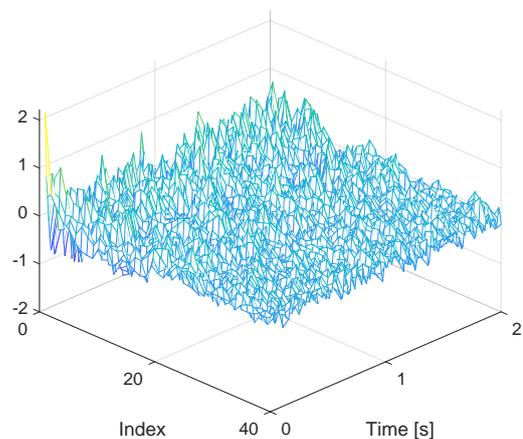


Fig. 6. Difference between MFCCs for worker bees and drones.

that are relevant to the task at hand, while also discarding noise and irrelevant information. They have also been combined with other types of neural networks, such as CNNs or RNNs, to improve performance on more complex tasks.

One common approach for using autoencoders in audio signal identification is to train the network on a reconstruction task, where the input is an audio

signal and the output is the reconstructed audio signal. The loss function used during training is typically a measure of the difference between the input and output signals, such as mean squared error (MSE) or mean absolute error. The mean square error is a reconstruction loss of the output produced by the network, obtained for a particular input vector after encoding and decoding stages. Once the network is trained, the compressed representation learned by the encoder can be used as a feature vector for identifying or classifying different types of audio signals. The scheme of such a network is shown in Fig. 7.

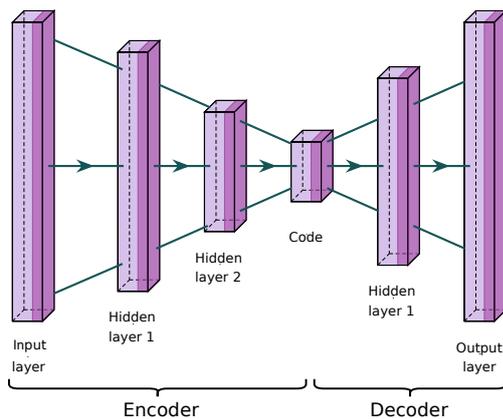


Fig. 7. General schema of autoencoder neural network.

In our simulations we have used neural networks with 1, 2, and 3 activation (ReLU) layers for the encoder. The decoder had always 2 layers: one with ReLU and one with a sigmoid activation function. In the case of power spectrum estimation based on the Burg nad Blackman-Tukey method, we decided for the following neural network settings. For the neural network with 3 activation layers, the number of features for encoder in layer 1 was 64, layer 2 – 32, layer 3 – 16. For the neural network with 2 activation layers, the number of features in layer 1 was 64, and layer 2 – 32. And for the neural network with only 1 activation layer, the number of features in layer 1 was 64. In the case of the usage of the MFCC, we decided to apply only the autoencoder neural network with only 1 activation layer, because the number of cepstral coefficients was relatively small, from 10 to 35 only.

Worker bees are present in the beehive the whole year, while no drones survive the Winter. This is the reason why the detection of a drone can be treated as an anomaly, which occurs most often in the time of the year preceding swarming. The process of anomaly detection using an autoencoder is divided into the following main steps:

- Step 1. Training: in the first step, the autoencoder is trained on the flight sounds of worker bees only.
- Step 2. Testing: in the second step, it is used for a test reconstruction of recordings from both

classes: worker bees and drones. Our hypothesis is that the abnormal signals (sounds of drones) will have a higher reconstruction error.

- Step 3. Classification: the last step is the detection of drone signals as anomalies if the reconstruction errors surpass a fixed threshold.

It is worth mentioning that the system takes into account that there are many more worker bees flying in and out the hive in the spring time than drones. The worker bees are extremely busy collecting nectar and pollen, and they generate huge traffic when flying to the hive. Fortunately, a special property of the autoencoder neural network prevents undetectability of less numerous drones, that could be the case for other classifiers. Autoencoder trained on the set containing only worker bees flight recordings generates much larger MSE for recordings of drones. The autoencoder neural network is a type of generative networks, and the reconstruction loss (after encoding and decoding stages) informs of the quality of recreation of the input by the network. An anomaly given as input to such a network, produces a higher loss value. The standard input gives a minimal loss value, related to a deviation between the input audio frames in the training set.

In Fig. 8, we present an exemplary histogram of the MSE in a series of numerical experiments returned by the autoencoder neural network for training (worker bees) and testing (drones) data sets. The vectors of coefficients representing recordings of drone flights produce a higher MSE because they deviate significantly from the signatures of worker bees, in the sense of the frequency components. The threshold is marked by the red dotted line.

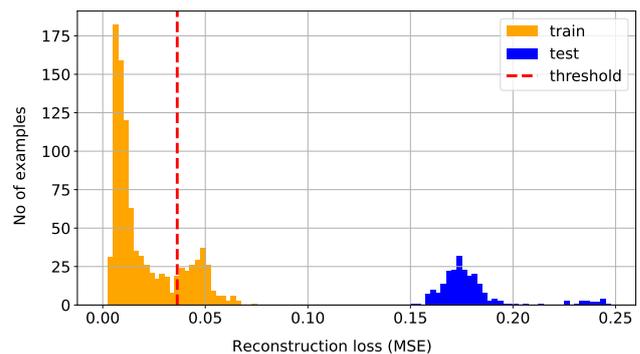


Fig. 8. Exemplary histogram of the MSE loss produced by autoencoder neural network for training and testing data sets.

As we see in Fig. 8, the MSE generated by drones is, on average, significantly greater than the MSE generated for worker bees. In the presented case, both histograms of the MSE are well separated in terms of separability of the corresponding probability densities. The proposed system of detecting drones is in fact detecting only a higher frequency of drones’ flights since

the system does not count particular individuals but only classifies audio frames of 1 second length. In this paper, we present results of classification recordings of bees into two classes based on a cleared and tagged (unquestionably by worker bees or drone labels) data set. It is absolutely possible that in practice we will encounter numerous situations of huge traffic of bees next to beehive entrance when many worker bees and drones can be present while recording an audio. This problem should be investigated in future research, and we hope that the special property of the autoencoder (giving the reconstruction loss as output) can also make it possible to detect a drone in the presence of many worker bees.

#### 4. Results

The experiment was performed for an autoencoder neural networks with 1, 2, and 3 activation layers, and for three preprocessing methods resulting in three different signal representations in the frequency domain by the following estimates:

- the Burg parametric method of power spectral estimate,
- the Blackman-Tukey nonparametric method of power spectral estimate,
- MFCC calculation.

Two first two methods, Burg and Blackman-Tukey methods, operate on a linear scale in the frequency domain. The third method, MFCC, uses a logarithmic scale. The goal was to investigate whether any approach will present a higher recognition results, considering that the learning process is done by neural networks and the calculated spectral coefficients are not analysed directly, but consist an input of the autoencoder neural networks.

For linear frequency scale, we cut the frequency bandwidth with the step 100 Hz in the range from 100 Hz to 3000 Hz. For the Mel-frequency spectrum scale, we have chosen 10, 15, 20, 25, 30 or 35 cepstral coefficients. The comparison study should determine the most effective method for estimating spectral coefficients, but also the number of cepstral coefficients or frequency bandwidth, depending on the chosen method.

##### 4.1. Statistical evaluation

Our database contained significantly less drone recordings in comparison to the huge number of worker bee recordings. Due to unbalanced data, it is important not only to focus on classification accuracy, but also on other result parameters such as recall and  $F1$ -score.

The accuracy for the binary classification problem is the proportion of correct predictions, both true positives (TP – number of correct detections) and true

negatives (TN – number of correct rejections), to the total number of cases examined:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (4)$$

where FP is false positive, also called false alarm, and FN is false negative.

In addition to accuracy, a recall value was calculated, meaning a sensitivity of the detection test:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (5)$$

We analyzed also a  $F1$ -score, which is defined as:

$$F1 = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \quad (6)$$

The  $F1$ -score takes values in the range  $[0, 1]$ . The highest possible value of  $F1$ -score (equal 1) indicates perfect precision and recall of the detection method.

In the series of experiments for different sets of settings, the above parameters: accuracy, recall and  $F1$ -score are going to indicate the best method for the considered drone detection problem. The ideal method, with zero incorrect classifications for both worker bees and drones, would have all three parameters equal 1. This is, of course, an unrealistic expectation, but the method which obtains the results closest to the value 1, will be considered the best.

For the better understanding of the obtained results, we also apply a weighted confusion matrix in the form presented in Fig. 9. In standard confusion matrix, there are simple counts: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), showing numbers of cases classified correctly or wrongly to both classes. Due to highly unbalanced data set, with many more worker bee audio recordings (10 000 samples) compared to only 1700 drone samples, the presented values were accordingly divided and presented in percent points. That way of presenting results, shows the actual percentage of correctly classified or misclassified worker bees and drones.

True labels	drone	TP/(TP+FN)	FN/(TP+FN)
	worker bee	FP/(TN+FP)	TN/(TN+FP)
		drone	worker bee
		Predicted labels	

Fig. 9. Weighted confusion matrix.

#### 4.2. Drone detection results

The classification of honey bees into two classes: worker bees and drones, based on audio recordings of the sound generated during their flight was performed for a data set of 10 000 audio samples for workers and 1700 for drones. The training of the autoencoder was performed with the use of 80 % of the honey bee flight recordings. The rest (20 %, 2000 samples) was used for testing, together with all records of drone flights (1700 samples). The experiment was carried out in 100 iterations. In each iteration, the autoencoder was trained and tested. The training stage gives full calculation of all parameters of the neural network. At this stage also a threshold dividing standard class (worker bees) and anomaly (drones) was derived from the formula:

$$\text{threshold} = \text{mean}(\text{MSE}_{\text{train}}) + \text{std}(\text{MSE}_{\text{train}}), \quad (7)$$

where mean is a mean value and std is a standard deviation of mean square errors, which are estimated using only outputs of the worker bee recordings during a training stage. In that way, the autoencoder was trained and ready for testing. The cases for which the mean square error was smaller than the calculated threshold were classified to the worker bees class and the cases for which was bigger to the drone class:

$$\text{class}(x) = \begin{cases} \text{worker class,} & \text{if } \text{MSE}_{\text{test}}(x) \leq \text{threshold,} \\ \text{drone class,} & \text{if } \text{MSE}_{\text{test}}(x) > \text{threshold.} \end{cases} \quad (8)$$

In the end all cases were compared with the true labels, which led to obtaining statistical indicators presented in this section.

The resulting accuracy for the Burg and Blackman-Tukey methods is shown in Fig. 10. The recall values for the methods are placed in Fig. 11. For five of the six methods: NN\_2\_BURG, NN\_3\_BURG, NN\_1\_B\_T, NN\_2\_B\_T, and NN\_3\_B\_T, the obtained values of accuracy and recall, after reaching a certain bandwidth, remain in almost constant value ranges, different for each specific method. Surprisingly, the only exception is a neural network with one activation layer for input in the form of Burg estimates

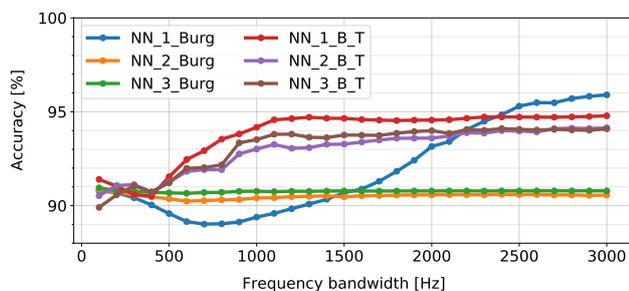


Fig. 10. Accuracy for neural networks with 1, 2, and 3 activation layers, and power spectrum estimation by Burg and Blackman-Tukey method.

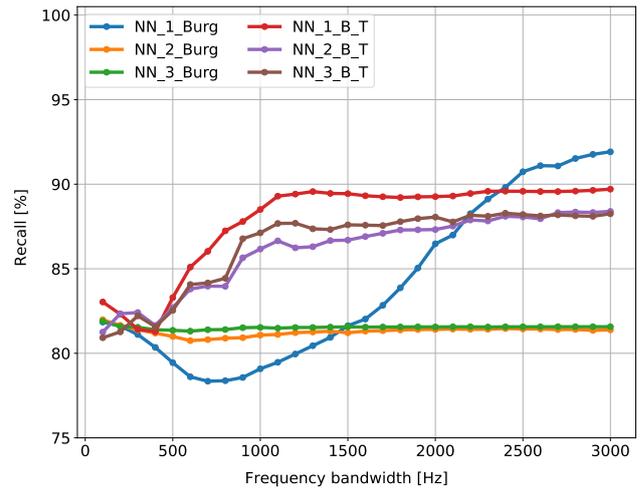


Fig. 11. Recall for neural networks with 1, 2, and 3 activation layers, and power spectrum estimation by Burg and Blackman-Tukey method.

(NN\_1\_BURG), for which the accuracy and recall values increase significantly as the frequency band is extended. The method has reached the highest evaluation factor values from all analyzed methods for the frequency bandwidth 3000 Hz.

The results for the MFCC are presented in Fig. 12. Both accuracy and recall values are presented as a function of the number of cepstral coefficients on the Mel-frequency scale. The experiment was carried out only for a neural network with 1 activation layer due to the small number of input coefficients, it is 10, 15, 20, 25, 30 or 35 MFCCs. It turns out that for this method (NN\_1\_MEL), the highest accuracy and recall are reached for only 10 cepstral coefficients and for the 15 and more coefficients they fall slightly.

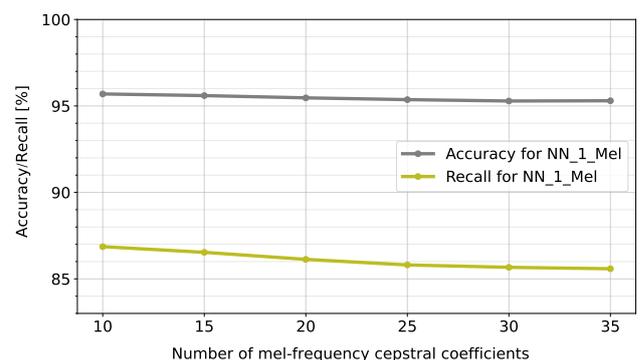


Fig. 12. Accuracy and recall for MFCC as input of neural network with 1 activation layer.

In Figs. 13a–c, we present the confusion matrices of the best results for all the three preprocessing methods: MFCC (NN\_1\_MEL 10 MFCC), Blackman-Tukey (NN\_1\_B\_T 3000 Hz), and Burg (NN\_1\_BURG 3000 Hz), respectively. The results were normalized by dividing by the size of the drone class in the up-

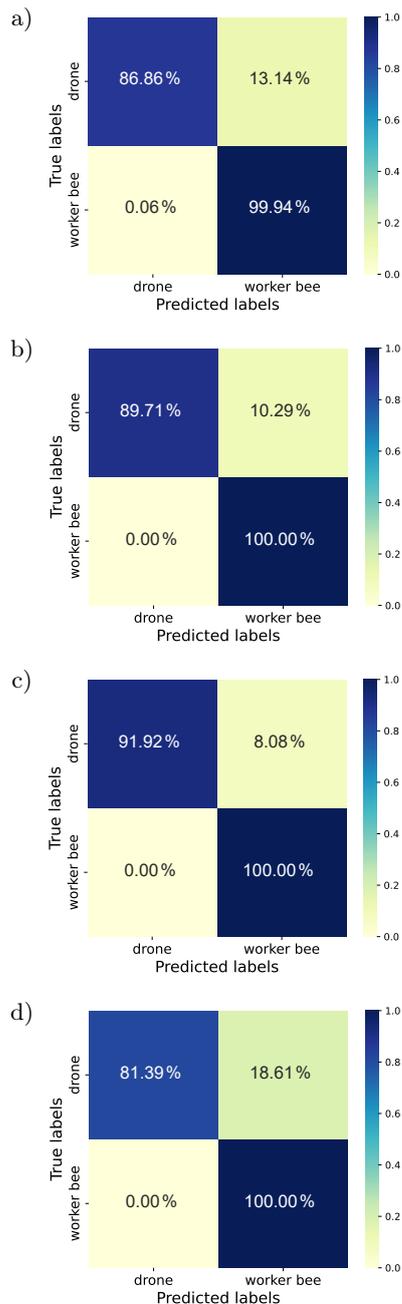


Fig. 13. Weighted confusion matrices, presented in percent points: a) NN\_1\_Mel 10 MFCC; b) NN\_1\_B\_T 3000 Hz; c) NN\_1\_Burg 3000 Hz; d) NN\_2\_Burg 3000 Hz.

per row of the confusion matrix and by the size of the worker bee class in the lower row. The worst result for the same 3000 Hz bandwidth was obtained for the Burg method and the neural network with 2 activation layers (NN\_2\_BURG 3000 Hz). We place the confusion matrix for that case in Fig. 13d.

In all cases with preprocessing based on Blackman-Tukey or Burg methods, the true negative rate (TNR) is 100 % and the false positive rate (FPR) is 0 %, which means that all signals from the worker bee class were correctly labeled as worker bees by neural networks.

On the other hand, for all cases using MFCC, the false positive rate has values higher than zero, meaning that the probability of a false alarm is also non-zero. Our method allows for an early detection of the swarming mood of honey bees, thanks to an analysis of the number of occurrences of drones flying in close proximity to the entrance of a beehive. Drones represent at most around 15 % of the population in beehives at the top, that is, during the late spring, and the worker bees are significantly more numerous at the same time – around 85 % of the population. That is why a higher false alarm rate can lead to many more false classifications of worker bees to the drone class. And as a consequence, the system would start the swarming alarm the whole year, except in winter, when worker bees are not active and stay inside the beehive.

#### 4.3. The best results

The best results, taking into account the highest values of the three statistical factors: the accuracy of drone recognition, the recall value (sensitivity of the method) and the  $F1$ -score, were obtained for the Burg power spectrum estimate method with the widest considered frequency band of 3000 Hz and neural network with one activation layer (marked as NN\_1\_BURG 3000 Hz – see Table 1). Accuracy reached 95.90 %, recall 91.92 %, and  $F1$ -score 96.11 %. The same method with a slightly narrower frequency band (NN\_1\_BURG 2900 Hz and NN\_1\_BURG 2800 Hz) achieved comparatively excellent results. Next in order, methods based on MFCC and Mel-frequency scale (such as, e.g., NN\_1\_MEL 10 MFCC) have given the accuracy at a very similar level, but recall recorded a decrease of around 5 % in all cases. Similarly,  $F1$ -score for MFCC method drops for more than 2 % compared to the Burg method.

However, it is worth noting that the unrivaled Burg method requires recordings with a bandwidth of 2500 Hz–3000 Hz to achieve the high classification results – see Table 1. If this would be a hardware limitation for some reasons, then it is better to use the

Table 1. Ten the best accuracy results and the corresponding recall and  $F1$ -score values.

	Method	Accuracy	Recall	$F1$ -score
1	NN_1_BURG 3000 Hz	0.959010	0.919165	0.961152
2	NN_1_BURG 2900 Hz	0.958210	0.917626	0.960442
3	NN_1_BURG 2800 Hz	0.957031	0.915206	0.959327
4	NN_1_MEL 10 MFCC	0.956934	0.868622	0.938068
5	NN_1_MEL 15 MFCC	0.955990	0.865345	0.936714
6	NN_1_BURG 2600 Hz	0.954863	0.910953	0.957374
7	NN_1_BURG 2700 Hz	0.954778	0.910762	0.957287
8	NN_1_MEL 20 MFCC	0.954680	0.861279	0.934943
9	NN_1_MEL 25 MFCC	0.953653	0.858056	0.933549
10	NN_1_BURG 2500 Hz	0.953010	0.907349	0.955725

MFCC-based method, at the cost of lowering the recall parameter and  $F1$ -score, and thus making more incorrect drone detections.

## 5. Conclusion

In this article, we have investigated the possibility of building an early swarming alert system for beekeepers, based on the detection of a larger number of drones flying at the entrance to a beehive. The system applies neural networks of autoencoder type, which must be previously trained on the basis of a signal database, containing worker bees and drones flight sound recordings, preferably registered in the environment where the system will be installed.

The preliminary study focused on finding the best signal processing methods and settings for the assumed task. We have compared three signal preprocessing methods, producing frequency-domain coefficients representing the recorded signals. They are: the Burg parametric method, the Blackman-Tukey nonparametric method, and the MFCC method. The power spectral or cepstral coefficients were the input of the autoencoder neural network. The detection was performed by the three settings of the encoder-decoder neural network pairs: with various (1, 2 or 3) number of activation layers for the encoder and with fixed 2 activation layers for the decoder.

The results obtained show that the configuration of the autoencoder neural network with only 1 activation layer has given the highest accuracy and recall values for all preprocessing methods. The best results have been received for the Burg parametric method of power spectrum estimation in a linear frequency scale and the frequency bandwidth of 3000 Hz (NN\_1\_BURG 3000 Hz). The accuracy of drone detection is 95.90 %, the recall (sensitivity) – 91.92 %,  $F1$ -score – 96.12 %, and false alarm rate equals 0.00 %. Cutting the bandwidth (to 2900 Hz, 2800 Hz, ..., and so on) has gradually decreased the accuracy of the drone detection.

The method using MFCC and the mel-frequency scale was found to give slightly worse results than the Burg preprocessing method with accuracy 95.69 %, the recall (sensitivity) – 86.86 %,  $F1$ -score – 93.81 %, and false alarm rate equals 0.06 % (for the case NN\_1\_MEL 10 MFCC). The accuracy level stays close to the accuracy for the best Burg method, but recall drops for more than 5 %,  $F1$ -score for more than 2 %, and the false alarm rate increases. The probability of a false alarm for the MFCC method is non-zero in all investigated cases, meaning that worker bees can be classified by the neural network as drones. Contrary to the Burg and Blackman-Tukey preprocessing methods, for which the probability of a false alarm is always zero.

Considering that the aim of the proposed method is an early detection of swarming mood of honey bees, based on more frequent observations of drones close

to a beehive entrance, it is important that the worker bees, which are more numerous in a swarm (around 85 % in late spring) than the drones (around 15 % in late spring), are not mistaken with the drones. This would increase the drone detection rate and falsely alarm beekeepers of a possible start of the swarming mood. In the future, a further study on the behavior of drone bees should be conducted. In particular, the correlation between the frequency of drone observation in relation to other bees and the state of the swarm should be investigated.

## Acknowledgments

The authors would like to express their great appreciation to Dr. Bartłomiej Golenko from the Wrocław University of Science and Technology for his valuable and constructive suggestions during the planning and development of this research work.

## References

1. BLACKMAN R.B., TUKEY J.W. (1958), The measurement of power spectra from the point of view of communications engineering – Part I, *Bell System Technical Journal*, **37**(1): 185–282, doi: [10.1002/j.1538-7305.1958.tb03874.x](https://doi.org/10.1002/j.1538-7305.1958.tb03874.x).
2. CEJROWSKI T., SZYMAŃSKI J., LOGOFAĆU D. (2020), Buzz-based recognition of the honeybee colony circadian rhythm, *Computers and Electronics in Agriculture*, **175**: 105586, doi: [10.1016/j.compag.2020.105586](https://doi.org/10.1016/j.compag.2020.105586).
3. COOLEY J.W. (1997), The re-discovery of the fast Fourier transform algorithm, *Microchimica Acta*, **3**: 33–45, doi: [10.1007/BF01201681](https://doi.org/10.1007/BF01201681).
4. DAVIS S., MERMELSTEIN P. (1980), Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, [in:] *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **28**(4): 357–366, doi: [10.1109/TASSP.1980.1163420](https://doi.org/10.1109/TASSP.1980.1163420).
5. GOURISARIA M.K., AGRAWAL R., SAHNI M., SINGH P.K. (2024), Comparative analysis of audio classification with MFCC and STFT features using machine learning techniques, *Discover Internet of Things*, **4**: 1, doi: [10.1007/s43926-023-00049-y](https://doi.org/10.1007/s43926-023-00049-y).
6. GRADIŠEK A., SLAPNIČAR G., SORN J., LUŠTREK M., GAMS M., GRAD J. (2017), Predicting species identity of bumblebees through analysis of flight buzzing sounds, *Bioacoustics*, **26**(1): 63–76, doi: [10.1080/09524622.2016.1190946](https://doi.org/10.1080/09524622.2016.1190946).
7. HADJUR H., AMMAR D., LEFÈVRE L. (2022), Toward an intelligent and efficient beehive: A survey of precision beekeeping systems and services, *Computers and Electronics in Agriculture*, **192**: 106604, doi: [10.1016/j.compag.2021.106604](https://doi.org/10.1016/j.compag.2021.106604).
8. HINTON G.E., SALAKHUTDINOV R.R. (2006), Reducing the dimensionality of data with neural networks,

- Science*, **313**(5786): 504–507, doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647).
9. KAWAKITA S., ICHIKAWA K. (2019), Automated classification of bees and hornet using acoustic analysis of their flight sounds, *Apidologie*, **50**: 71–79, doi: [10.1007/s13592-018-0619-6](https://doi.org/10.1007/s13592-018-0619-6).
  10. KAY S.M. (1988), *Modern Spectral Estimation: Theory and Application*, Englewood Cliffs, Prentice Hall.
  11. KIM J., OH J., HEO T.-Y. (2021), Acoustic scene classification and visualization of beehive sounds using machine learning algorithms and Grad-CAM, *Mathematical Problems in Engineering* **2021**: 5594498, doi: [10.1155/2021/5594498](https://doi.org/10.1155/2021/5594498).
  12. KULYUKIN V., MUKHERJEE S., AMLATHE P. (2018), Toward audio beehive monitoring: Deep learning vs. standard machine learning in classifying beehive audio samples, *Applied Sciences*, **8**(9): 1573, doi: [10.3390/app8091573](https://doi.org/10.3390/app8091573).
  13. LIBAL U., BIERNACKI P. (2024), MFCC selection by LASSO for honey bee classification, *Applied Sciences*, **14**(2): 913, doi: [10.3390/app14020913](https://doi.org/10.3390/app14020913).
  14. NOLASCO I., TERENCEZ A., CECCHI S., ORCIONI S., BEAR H.L., BENETOS E. (2018), Audio-based identification of beehive states, [in:] *ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8256–8260, doi: [10.1109/ICASSP.2019.8682981](https://doi.org/10.1109/ICASSP.2019.8682981).
  15. ORFANIDIS S.J. (1995), *Introduction to Signal Processing*, Prentice-Hall.
  16. OSTROWSKA W. (1980), *Beekeeping* [in Polish: *Gospodarka Pasiieczna*], Państwowe Wydawnictwo Rolnicze i Leśne, Warszawa.
  17. PENG R., ARDEKANI I., SHARIFZADEH H. (2020), An acoustic signal processing system for identification of queen-less beehives, [in:] *Proceedings of the 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 57–63.
  18. RIBEIRO A.P., DA SILVA N.F.F., MESQUITA F.N., ARAÚJO P.D.C.S., ROSA T.C., MESQUITA-NETO J.N. (2021), Machine learning approach for automatic recognition of tomato-pollinating bees based on their buzzing-sounds, *PLOS Computational Biology*, **17**(9): e1009426, doi: [10.1371/journal.pcbi.1009426](https://doi.org/10.1371/journal.pcbi.1009426).
  19. RUVINGA S., HUNTER G.J., DURAN O., NEBEL J. (2021), Use of LSTM networks to identify “queenlessness” in honeybee hives from audio signals, [in:] *2021 17th International Conference on Intelligent Environments (IE)*, pp. 1–4, doi: [10.1109/IE51775.2021.9486575](https://doi.org/10.1109/IE51775.2021.9486575).
  20. SARMA M.S., FUCHS S., WERBER C., TAUTZ J. (2002), Worker piping triggers hissing for coordinated colony defence in the dwarf honeybee *Apis florea*, *Zoology*, **105**(3): 215–223, doi: [10.1078/0944-2006-00064](https://doi.org/10.1078/0944-2006-00064).
  21. SEELEY T.D., TAUTZ J. (2001), Worker piping in honey bee swarms and its role in preparing for liftoff, *Journal of Comparative Physiology A*, **187**: 667–676, doi: [10.1007/s00359-001-0243-0](https://doi.org/10.1007/s00359-001-0243-0).
  22. SOARES B.S., LUZ J.S., DE MACÊDO V.F., SILVA R.R.V., DE ARAÚJO F.H.D., MAGALHÃES D.M.V. (2022), MFCC-based descriptor for bee queen presence detection, *Expert Systems with Applications*, **201**: 117104, doi: [10.1016/j.eswa.2022.117104](https://doi.org/10.1016/j.eswa.2022.117104).
  23. WRIGHT W.D. (1913), *The Honey Bee*, Albany, J.B. Lyon Company.
  24. YEGNANARAYANA B., PRASANNA S.M., ZACHARIAH J.M., GUPTA C.S. (2005), Combining evidence from source, suprasegmental and spectral features for a fixed-text speaker verification system, *IEEE Transactions on Speech and Audio Processing*, **13**(4): 575–582, doi: [10.1109/TSA.2005.848892](https://doi.org/10.1109/TSA.2005.848892).
  25. ZGANK A. (2021), IoT-based bee swarm activity acoustic classification using deep neural networks, *Sensors*, **21**(3): 676, doi: [10.3390/s21030676](https://doi.org/10.3390/s21030676).
  26. ZLATKOVA A., GERAZOV B., TASHKOVSKI D., KOKOLANSKI Z. (2020), Analysis of parameters in algorithms for signal processing for swarming of honeybees, [in:] *2020 28th Telecommunications Forum (TELFOR)*, pp. 1–4, doi: [10.1109/TELFOR51502.2020.9306562](https://doi.org/10.1109/TELFOR51502.2020.9306562).