



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

ARCHIVES of ACOUSTICS

QUARTERLY

Vol. 48, No. 3, 2023

WARSAW



ARCHIVES of ACOUSTICS

QUARTERLY, Vol. 48, No. 3, 2023

Review Papers

- M. Igras-Cybulska, D. Hemmerling, M. Ziółko, W. Datka, E. Stogowska, M. Kucharski, R. Rzepka, B. Ziółko, *Speech analysis as a tool for detection and monitoring of medical conditions: A review* 289

Research Papers

- T. Mushtaq, A. Kamran, M.Z.A. Qureshi, Z. Iqbal, *A symmetric approach in the three-dimensional digital waveguide modeling of the vocal tract* 317
- P. Lindh, P. Lemenkova, *Ultrasonic P- and S-wave reflection and CPT soundings for measuring shear strength in soil stabilized by deep lime/cement columns in Stockholm Norvik Port* 325
- A. Yori, *Field study on underwater noise emitted by small tourist boats. Comparison between the use of electric and combustion motors* 347
- H. Xiaohui, L. Zhongle, Y. Chao, Y. Zhiyong, *Study on noise attenuation characteristics of hydrofoil with specific cavitation number* 359
- Y. Chen, J. Yang, L. Li, S. Xiao, *Design and experiments of a new internal cone type traveling wave ultrasonic motor* 373
- S.-H. Kim, M.-J. Kim, *Relationship between the sound transmission through the finite double-panel structure with a cylindrical shell array and the vibro-acoustic characteristics of its constituents* 381
- M. Kafil, K. Darabi, S. Ziaei-Rad, *An improved EMD method based on utilizing certain inflection points in the construction of envelope curves* 389
- Q. Lu, M. Li, *VMD and CNN-Based classification model for infrasound signal* 403
- C. Jiang, X. Cao, F. Yang, Z. Liu, *Numerical investigation of the propagation characteristics of surface transverse wave considering various quartz substrate and electrode configurations* 413
- A. Brański, R. Kuras, *PZT asymmetrical shape optimization in active vibration reduction of triangular plates* .. 425
- M. Li, J. Liu, *A microscopic prediction model for traffic noise in adjacent regions to arterial roads* 433

Chronicle

- 51st Winter School on WQA and Lth Winter School on EA&V 451

Editorial Board

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

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

Associate Editors

General linear acoustics and physical acoustics

RDZANEK Wojciech P. (University of Rzeszów, Poland)

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

Architectural acoustics

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

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

Musical acoustics and psychological acoustics

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

PREIS Anna (Adam Mickiewicz University, Poland)

Underwater acoustics and nonlinear acoustics

GRELOWSKA Grażyna (Gdańsk University of Technology, Poland)

MARSZAŁ Jacek (Gdańsk University of Technology, Poland)

Speech, computational acoustics, and signal processing

KOCIŃSKI Jędrzej (Adam Mickiewicz University, Poland)

Ultrasonics, transducers, and instrumentation

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

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

Sonochemistry

DZIDA Marzena (University of Silesia in Katowice, Poland)

Electroacoustics

ŻERA Jan (Warsaw University of Technology, Poland)

Vibroacoustics, noise control and environmental acoustics

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

KLEKOT Grzegorz (Warsaw University of Technology, Poland)

KOMPALA Janusz (Central Mining Institute, Poland)

LENIOWSKA Lucyna (University of Rzeszów, Poland)

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

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

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

Advisory Editorial Board

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

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

BLAUERT Jens (Ruhr University, Germany)

BRADLEY David (The Pennsylvania State University, USA)

CROCKER Malcolm J. (Auburn University, USA)

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

HANSEN Colin (University of Adelaide, Australia)

HESS Wolfgang (University of Bonn, Germany)

KOZACZKA Eugeniusz (Gdańsk University of Technology, Poland)

LEIGHTON Tim G. (University of Southampton, UK)

LEWIN Peter A. (Drexel University, USA)

MAFFEI Luigi (Second University of Naples SUN, Italy)

PUSTELNY Tadeusz (Silesian University of Technology, Poland)

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

SUNDBERG Johan (Royal Institute of Technology, Sweden)

ŚLIWIŃSKI Antoni (University of Gdańsk, Poland)

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

TORTOLI Piero (University of Florence, Italy)

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

Polish Academy of Sciences
Institute of Fundamental Technological Research PAS
Committee on Acoustics 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^AT_EX: JEZIERSKA Katarzyna (Institute of Fundamental Technological Research PAS, Poland)

Review Paper

Speech Analysis as a Tool for Detection and Monitoring
of Medical Conditions: A review

Magdalena IGRAS-CYBULSKA^{(1),(3)*} , Daria HEMMERLING^{(1),(3)}, Mariusz ZIÓŁKO⁽¹⁾,
Wojciech DATKA^{(2),(4)}, Ewa STOGOWSKA⁽²⁾, Michał KUCHARSKI⁽¹⁾,
Rafał RZEPKA⁽⁵⁾, Bartosz ZIÓŁKO^{(1),(5)}

⁽¹⁾ *Techmo sp. z o.o.*
Kraków, Poland

⁽²⁾ *Medical University of Białystok*
Białystok, Poland

⁽³⁾ *AGH University of Science and Technology*
Kraków, Poland

⁽⁴⁾ *Faculty of Medicine, Jagiellonian University*
Kraków, Poland

⁽⁵⁾ *Hokkaido University*
Kita Ward, Sapporo, Hokkaido, Japan

*Corresponding Author e-mail: migras@agh.edu.pl

(received May 22, 2022; accepted March 21, 2023)

The goal of this article is to present and compare recent approaches which use speech and voice analysis as biomarkers for screening tests and monitoring of some diseases. The article takes into account metabolic, respiratory, cardiovascular, endocrine, and nervous system disorders. A selection of articles was performed to identify studies that assess voice features quantitatively in selected disorders by acoustic and linguistic voice analysis. Information was extracted from each paper in order to compare various aspects of datasets, speech parameters, methods of applied analysis and obtained results. 110 research papers were reviewed and 47 databases were summarized. Speech analysis is a promising method for early diagnosis of certain disorders. Advanced computer voice analysis with machine learning algorithms combined with the widespread availability of smartphones allows diagnostic analysis to be conducted during the patient's visit to the doctor or at the patient's home during a telephone conversation. Speech analysis is a simple, low-cost, non-invasive and easy-to-provide method of medical diagnosis. These are remarkable advantages, but there are also disadvantages. The effectiveness of disease diagnoses varies from 65% up to 99%. For that reason it should be treated as a medical screening test and should be an indication of the need for classic medical tests.

Keywords: speech analysis; speech features; acoustic parameters; linguistic analysis; voice biomarkers; screening tests.



Copyright © 2023 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

Acc – accuracy,	AMDF – average magnitude difference function,
AD – Alzheimer's disease,	APQ – amplitude perturbation quotient,
AI – artificial intelligence,	AR – average recall,
ALS – amyotrophic lateral sclerosis,	ASCVD – atherosclerotic cardiovascular disease,
ALSFRS-R – sclerosis functional rating scale,	ASR – automatic speech recognition,
	AUC – area under a curve,
	AUROC – area under the receiver operating characteristic,

- AVEC – audio visual emotion challenges,
 AVQI – acoustic voice quality index,
 BDI – Beck depression inventory,
 BFI – big five inventory,
 BIDR – balanced inventory of desirable responding,
 BMI – body mass index,
 BPRS – brief psychiatric rating scale,
 CAD – coronary artery disease,
 CH – control healthy,
 CHD – coronary heart disease,
 CHR – clinical high-risk,
 CNN – convolutional neural network,
 CSL – computerized speech lab,
 CVR – cockpit voice recorder,
 DAIC – distress assessment interview corpus,
 DCT – discrete cosine transform,
 DDK – diadochokinetic,
 DM – diabetes mellitus,
 DNN – deep neural network,
 DSM – diagnostic and statistical manual of mental disorders,
 EM – expectation–maximization,
 F0 – fundamental frequency,
 FBS – fetal bovine serum,
 FT4 – free thyroxine,
 GC – NN gated convolutional neural network,
 GFI – glottal function index,
 GMM – the Gaussian mixture model,
 GMM-UBM – the Gaussian mixture model-universal background model,
 GPR – the Gaussian processes regression,
 GRBAS – grade-roughness-breathiness-asthenia-strain scale,
 H&Y – the Hoehn and Yahr scale,
 HAMD – the Hamilton depression rating scale,
 HbA1c – glycated hemoglobin A1c,
 HC – healthy controls,
 HLM – hierarchical linear modeling,
 HMM – hidden Markov model,
 HRSD – the Hamilton rating scale for depression,
 HSC – hierarchical spectral clustering,
 IQR – interquartile range,
 IVR – interactive voice response,
 JAD – Just Add Data,
 K-SADS-PL – kiddie schedule for affective disorders and schizophrenia (present and lifetime version),
 KNN – K-nearest neighbours,
 LLD – low-level descriptors,
 LR – logistic regression,
 LSA – latent semantic analysis,
 LSTM – multi-layer long short-term memory,
 LTAS – long term average spectrum,
 MAE – mean absolute error,
 MAP – maximum a posteriori,
 MDS-UPDRS – Movement Disorders Society UPDRS,
 MDVP – Multi-Dimensional Voice Program,
 MFCC – mel-frequency cepstral coefficients,
 MHMC – Multimedia Human-Machine Communication,
 MLP – multilayer perceptron,
 MMSE – mini-mental state examination,
 MPT – maximum phonation time,
 NB – naive Bayes,
 NN – neural networks,
 NN LSTM – neural net multi-layer long short-term memory,
 PANAS – positive and negative affect schedule,
 PANSS – positive and negative syndrome scale,
 PCL-C – post-traumatic stress disorder checklist,
 PCOS – polycystic ovary syndrome,
 PD – Parkinson’s disease,
 PDD – phase distortion deviation,
 PHQ-9 – patient health questionnaire-9,
 PPQ – period perturbation quotient,
 PTP – phonation threshold pressure,
 PTSD – post-traumatic stress disorder,
 PVRQoL – Pediatric Voice-Related Quality-of-Life,
 QIDS – quick inventory of depressive, symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR),
 RAP – relative average perturbation,
 RBF – radial basis function,
 RF – random forest,
 RLR – randomized logistic regression,
 RME – mental and emotional reinforcement,
 RMSE – root mean square error,
 RSI – reflux severity index,
 RVM – relevance vector machines,
 Sens – sensitivity,
 SER – standard error of regression,
 SIPS – semi-structured interview,
 SIT – sentence intelligibility test,
 SNR – signal-to-noise ratio,
 SOPS – scale of prodromal symptoms,
 Spec – specificity,
 SPT – speech pause time,
 STAI – state-trait anxiety inventory,
 SVM – support vector machine,
 T4 – thyroxine, thyroid hormone,
 TSH – thyroid-stimulating hormone,
 TSST-C – trier social stress test for children,
 UAR – unweighted average recall,
 UBM – universal background model,
 UD – unipolar depression,
 UPDRS – Unified Parkinson’s Disease Rating Scale,
 Var – variance,
 VC – vital capacity,
 VHDAIC – virtual human distress assessment interview corpus,
 VHI – voice handicap index,
 YMRS – Young’s rating scale for mania.

1. Introduction

Organs involved in the speech generation process are highly sensitive to both physical and mental ailments, hence the health of the speakers significantly affects their manner of speaking, voice emission, syntax, semantics and specific speech habits. Early detection and treatment of disorders can improve the effectiveness of treatment. In spite of this, speech analysis is currently rarely used in medical diagnostics of disorders other than those directly affecting the organs involved in speech generation and the respiratory system.

There is a high volume of publications on the diagnosis of specific disorders using speech analysis. The

largest number of publications concerns the diagnosis of Parkinson’s disease (PD). This high number of publications is reflected in this paper.

Speech analysis is a simple, low-cost, non-invasive and easy-to-provide preliminary test for disorders which affect speech, even marginally. Publications usually present problems related to the diagnosis of a single disease entity and state a high likelihood of diagnosis.

In this article we summarize different approaches which create systems to detect voice and speech impairments tackled by researchers around the world. Authors take into account different languages, dialects, types of speech (vowels, read text, monologue, etc.), algorithms, and different disorders to be analyzed. The current review aims to summarize the state-of-the-art of voice analysis as a biomarker of diseases. In particular, we want to compare properties of speech corpora and different approaches to speech processing paths.

Although some systematic reviews have been prepared recently, they are usually dedicated to certain groups of disorders (CUMMINS *et al.*, 2015a; DOGAN *et al.*, 2017; LOW *et al.*, 2020; MORO-VELAZQUEZ *et al.*, 2021; STOGOWSKA *et al.*, 2022). To the best of our knowledge, there is a lack of a systematic review of voice analysis in connection with somatic disorders. In this study we did not include disorders which directly affect organs involved in speech generation, such as vocal cords or a vocal tract, as well as pulmonary disorders. The papers present a range of disorders that were analyzed by researchers and the results they obtained. This includes cardiovascular, metabolic, endocrine, COVID-19, schizophrenia, depression, amyotrophic lateral sclerosis (ALS), affective and neurodegenerative (Parkinson’s, Alzheimer’s, dementia) disorders. Creating a system that could monitor and reveal whether a patient has any voice/speech abnormalities and whether further diagnostics are required for a specific disease entity would be extremely useful in the medical environment. Another highly desirable tool would be a system allowing monitoring of treatment through voice analysis. In this article, we present scientific approaches to the problems of detecting disorders through voice analysis and a summary of the results, challenges, and problems.

2. Methods

2.1. *Speech as an objective biomarker*

For the purpose of this work, we have analyzed the PRISMA checklist which includes reports of reviews evaluating randomized trials. It is also a basis for reporting systematic reviews of different types of research. Articles with publication dates between January 2011 and November 2020 were selected from PubMed and ISCA Archive, using keywords ‘voice’ and

‘speech’ and respective disorder names. The records were screened for relevance to the topic of this review in order to identify studies that quantitatively assessed voice quality in the selected disorders by voice analysis. Studies which included only a perceptual assessment of voice were rejected.

For cardiovascular, metabolic and endocrine disorders, schizophrenia and ALS, all records were included as these disorders are less well documented (number of research papers less than 10). In contrast, affective disorders and neurodegenerative disorders (Parkinson’s, Alzheimer’s, dementia) are well investigated and the majority of publications describe automatic recognition with machine learning methods or even applications. Therefore only selected articles were included in our analysis, selected on the basis of publication in the ISCA archive and a high level of advancement. In this review we mentioned the research articles as well as three recent systematic reviews.

Determination of speech features we divided into several categories depending on the source of their origin (prosodic, spectral, voice source, linguistic) without going into further details. In some cases, standard feature vectors were selected. Machine learning approaches are summarized with information on the classifiers, the evaluation method and achieved best scores. If there was no attempt at automatic classification or regression, we report which statistical tests were used.

Finally, reviewed papers were flagged (Tables 10, 11, and 13) with one of the tags describing how advanced it is: 1 – basic research (investigating the statistical significance of acoustic parameters in the context of disorders); 2 – automated (using machine learning to classify regression); 3 – application (usually a smartphone app) for types of diagnostics (not simply collecting recordings).

In Sec. 3, each subsection starts with a medical description of how each disorder affects the voice and contains a summary of different processing approaches; they are supplemented by tables comparing cited studies. We reviewed a total of 110 papers, including disorders: endocrine, cardiovascular, metabolic, neurodegenerative, mental, and COVID-19.

Creating a computerized system of speech recognition-based diagnostic tools includes three steps. Voice recordings of a control group and individuals affected by the disorder in question are required first. Next, specific speech features are selected and calculated for both sets. In the training phase, the computer compares features of both sets of recordings and creates a classifier. This classifies the recordings into one of two sets: non affected individuals, or individuals suspected to be affected by the ailment. The structure of such a system is shown in Fig. 1.

Currently, the most difficult element is obtaining a sufficiently high number of recordings for the

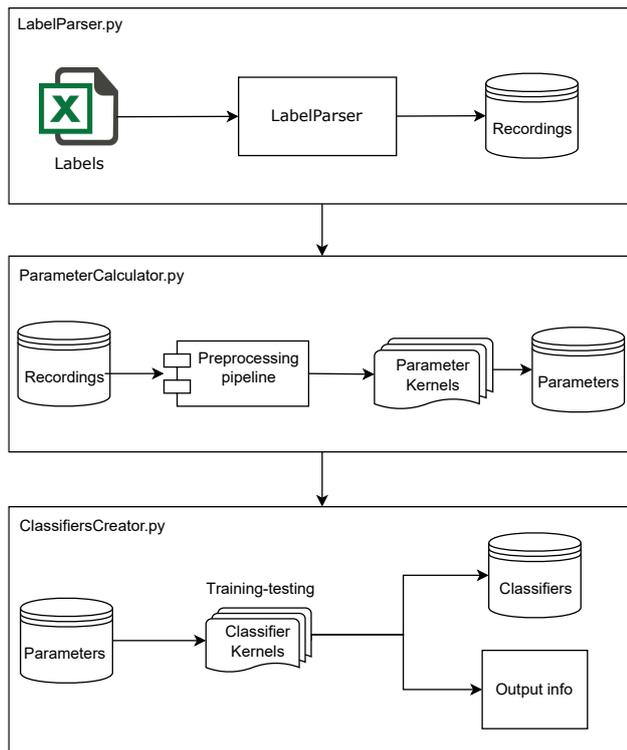


Fig. 1. Elements of a system for determining: speech features, training classifiers, and medical diagnosis.

training, validation, and testing phases. The recording databases are diverse and require a brief introduction. Software for determining speech features and creating classifiers is generally widely available, so the main task is selecting an effective algorithm. This means that the same algorithms are used for a variety of issues in speech technology, e.g., speech and speaker recognition, emotion detection and diagnosis of disease states.

We also performed a summary of available databases mentioned in the literature. Our analysis focused on the language and content of speech (with categories including sustained vowels, read speech, spontaneous monologue, dialogue with a human or virtual interviewer), recording protocol (how many times the individual was recorded, recording procedure, recording duration), number of speakers, their age and gender (in both the Control Healthy (CH) and affected persons), and other modalities (usually video, sometimes motion capture or biometric signals). We also included information on metadata: clinical evaluation of patients and perceptual evaluation of voice. Although the quality of recordings may be crucial for further processing, we did not compare technical details of recording procedures (equipment, acoustic conditions, sampling frequency, file parameters) because they are not usually systematically reported in the articles.

Finally, we reviewed 47 databases: 10 corpora of endocrine diseases (three of diabetes, four of polycystic ovary syndrome (PCOS) and related disorders, three

of thyroid disorders), two of cardiovascular disorders (one of CAD, one of CHD), six of metabolic disorders (obesity), 12 of neurodegenerative disorders (two of ALS, eight of Parkinson's disease, two of Alzheimer's disease), and 17 of mental disorders (three of bipolar, 11 depression and/or anxiety and/or PTSD, three of schizophrenia).

Additional diagrams were prepared to illustrate general tendencies and to compare advancements in the state-of-the-art in the analyzed groups of disorders. The main sources of problems were identified and some recommendations for future research were set.

2.2. Databases

Speech corpus development is generally time and cost consuming; however, good quality recordings are crucial for further processing. Figure 2 indicates that the majority of speech recording databases for the purposes of medical diagnostics emerged in 2016. This was likely in response to publications reporting a satisfactory effectiveness of speech analysis in an initial recognition of disease symptoms.

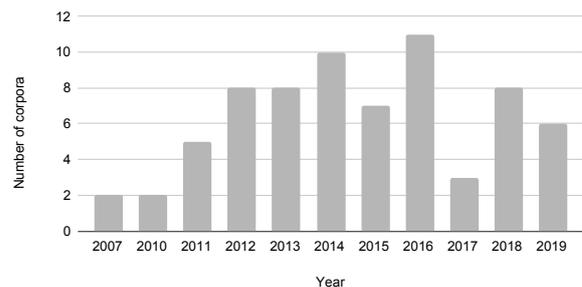


Fig. 2. Number of corpora created in 2007–2019.

There is a wide range of speech protocols, from the shortest (sustained phonation of vowels only) to the longest (interviews with a virtual agent). For read speech, there are standard text passages to be read, usually excerpts from stories or a short natural sentence. Counting one to 10 is also used. Sometimes the patient is asked for a short monologue. Just 21 of the 47 corpora contained more than two categories of speech (Fig. 3).

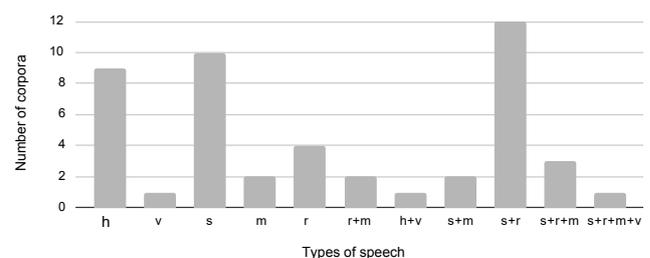


Fig. 3. Types of speech recorded in corpora 2009–2019: h – dialogue with a human; v – dialogue with a virtual agent; s – sustained vowels; m – monologue; r – read speech.

2.3. Features

Deviations of voice features are generally the result of anatomical changes or changes in the functioning of the nervous system. The cited publications do not concern analysis of sources of voice deviations, but only the ability to detect the effects of these changes is considered.

Three sets of speech parameters are usually distinguished: source parameters, vocal tract parameters, and prosody. The first and second groups are usually analyzed in the frequency domain of 20-30 milliseconds frames. Therefore, they are described as low-level descriptors (LLD).

Signal processing algorithms such as filtering or linear prediction allow for the extraction of acoustic features of the source and filter separately. The most popular source parameters include jitter and shimmer, which describe the stability of fundamental frequency production, and voice trembling. Examples of typical vocal tract features include formants and mel-frequency cepstral coefficients (MFCC). Prosodic features, such as syllables, phrases, and sentences, are observed in larger frames. For this reason, prosody parameters are also known as supra segmental or high-level features. Prosody describes intonation (modulation of fundamental frequency (F0) within the utterance), intensity (loudness, energy), and rhythm of speech (pauses, duration of speech segments, speech tempo).

Most of the authors of the reviewed articles used standard software tools to extract acoustic features. The most popular software tools for acoustic feature extraction are openSMILE, Praat, Multi-Dimensional Voice Program (MDVP), Kay Elemetrics-Computer Speech Lab, Dr. Speech, Snack Sound Toolkit and MATLAB (toolboxes such as Voice Sauce).

2.4. Classification models and evaluation metrics

In the reviewed papers focusing on the relationship between voice and the specified disorders, several study design scenarios can be found:

- ill/healthy comparison which leads to binary classification;
- comparison of subclasses of given disorders, which leads to multiclass classification;
- comparison between pre-treatment and post-treatment;
- monitoring the progression of disease severity;
- monitoring of treatment success;
- correlation with prodromal symptoms which leads to the measure of risk of the given disorder.

Depending on the case, classification or regression methods are applied. The most popular models are Gaussian mixture model (GMM) and *i*-vectors. Two type of classifiers are the most frequently used: support vector machine (SVM) and neural networks (NN). Different measures are used according to the classification of the regression model (Tables 2, 4, 6, 8, 10, 11, 13, 15, and 17).

3. Results

3.1. Cardiovascular diseases

Cardiovascular disease is the most common cause of death in both developed and developing countries (KONES, RUMANA, 2017). There is currently little evidence on any association between cardiovascular disease and voice features (Table 1). It has been posited that the disease process may affect anatomical structures associated with voice generation, for example in atherosclerosis; as a systemic inflammatory process, it is associated with multiple pathological processes such as chronic kidney disease, cerebrovascular disease, vascular dementia, retinopathy and peripheral artery disease (MAOR *et al.*, 2018).

PAREEK and SHARMA (2016) studied coronary heart disease (CHD) (Table 2). Their research reveals significant variations in spectrograms, the long-term average spectrum (LTAS) and other voice parameters such as jitter, shimmer and amplitude perturbation quotient (APQ), smoothed APQ, relative average perturbation (RAP), period perturbation quotient (PPQ),

Table 1. Comparison of speech databases in CAD and CHD

(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
Cardiovascular: CHD												
PAREEK, SHARMA, 2016	N/A	80	80	53.4	NI	–	1	0	0	0	0	00:00:04
Cardiovascular: CAD												
MAOR <i>et al.</i> , 2018	English	71	37	63	ASCVD risk score	–	0	1	1	0	0	NI

N/A – not applicable; NI – no information.

Table 2. Comparison of research methods in cardiovascular diseases
(pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Disease	Tool used for features extraction	Features					Classifier/Test
			pr	sp	vs	li	naf	
PAREEK, SHARMA, 2016	CHD	MDVP, CSL	0	1	1	0	18	–
MAOR <i>et al.</i> , 2018	CAD	Beyond verbal communications	0	1	0	0	81	Logistic regression

smoothed PPQ in comparison with the control group ($P < 0.05$).

In research conducted by (MAOR *et al.*, 2018), coronary artery disease (CAD) patients were compared to CH (Table 2). MFCC parameters were extracted. Univariate binary logistic regression analysis identified five voice features that were associated with CAD. Multivariate binary logistic regression with adjustment for atherosclerotic cardiovascular disease (ASCVD) risk scores identified two voice features that were independently associated with CAD (odds ratio OD = 0.37; 95% CI, interquartile range IQR = 0.18–0.79; OD = 4.01; 95% CI, IQR = 1.25–12.84; p -value = 0.009 and p -value = 0.02, respectively). Both features were more strongly associated with CAD when patients were asked to describe an emotionally significant experience.

3.2. COVID-19

Acute respiratory disease caused by the SARS-CoV-2 locates and multiplies mainly in the cytoplasm of lung cells. Thanks to the techniques enabling precise voice analysis combined with artificial intelligence (AI), it is possible to effectively and, above all early, diagnose COVID-19. Diagnostic scenarios can be conducted based on voice samples sent over a telephone.

HAN *et al.* (2020) focused on developing some potential use-cases of intelligent speech analysis for COVID-19 diagnosed patients. By analysing speech recordings they constructed audio-only-based models to automatically categorise the health state of patients from four aspects: severity of illness, sleep quality, fatigue, and anxiety.

The prominent symptoms of COVID-19 include cough and breathing difficulties. SHARMA *et al.* (2020) claim that respiratory sounds (cough, breath, and voice) can provide useful insights, enabling the design of a diagnostic tool. In their research, they determined 9 sound categories describing the voice, breath and cough. The acoustic analysis included the spectral analysis, energy description and zero-crossing rate. The accuracy on test data was 67%.

In order to better evaluate the COVID-19 infection, WEI *et al.* (2020) proposed an end-to-end method for cough detection and classification. It is based on real human-robot conversation data, which processes speech signals to detect cough and classifies it if de-

tected. They find that the weighted sum can generate a 76% top-1 accuracy.

PINKAS *et al.* (2020) studied the harnessed deep machine learning and speech processing to detect the SARS-CoV-2 positives. Their dataset of cellular phone recordings included vocal utterances, speech, and coughs that were self-recorded by the subjects in either hospitals or isolation sites. They achieved the following diagnostic efficiency: a recall of 78% and a probability of false alarm (PFA) of 41%.

The papers of (LECHIEN *et al.*, 2020; STASAK *et al.*, 2021) presented voice analysis to classify the severity of COVID-19, from mild to moderate. They have reported the severity of COVID-19 might have influenced abnormally high rates of vocal dysphonia likely due to glottic (e.g., vocal folds) edema and tissue inflammation. In the study, the scientists used glottal, prosodic and spectral acoustic features from short-duration speech segments and applied them to machine learning algorithms. Experimental results indicate that certain feature-task combinations can produce COVID-19 classification accuracy of up to 80% as compared with using the all-acoustic feature baseline (68%).

DESPOTOVIC *et al.* (2021) presents the experiments with cough patterns using standard acoustic features sets, wavelet scattering features and deep audio embeddings extracted from low-level feature representations. The models achieve accuracy of 89% confirming the applicability of audio signatures to identify the COVID-19 symptoms.

The authors of (HASSAN *et al.*, 2020) applied six speech features from a collected dataset and deep neural network (DNN) to create system for COVID-19 detection. The results show the classification accuracy for breathing sound reaching up to 98%, for cough sounds an accuracy of 97% was attained, while the voice accuracy of the system was only 88%. Their analysis shows that in the first place collecting cough and breathing sounds should make a COVID-19 detection system.

SUBIRANA *et al.* (2020) showed that AI transfer learning algorithms trained on cough phone recordings results in diagnostic tests for COVID-19. They suggest a novel open collective approach to large-scale real-time health care AI. They evaluated the performance of four shallow machine learning classification algorithms: SVM, K-nearest neighbors, random forest,

logistic regression. The presented graphs show that, depending on the methods used, the accuracy of COVID-19 diagnostics ranged from less than 80% (KNN and DenseNet201) to almost 100% (logistic regression and DenseNet201).

LAGUARTA *et al.* (2020) noticed that COVID-19 subjects, especially including asymptomatic, could be accurately discriminated from a forced-cough cell phone recording using AI. When validated with subjects diagnosed using an official test, the model achieved COVID-19 sensitivity of 99% with a specificity of 94%.

DESHPANDE and SCHULLER (2020) summarised efforts taken by the research community towards helping the individuals and the society in the fight against COVID-19 using speech signal processing.

3.3. Obesity and metabolic syndrome

Obesity is a growing health problem in many parts of the world. Excessive body fat is associated with multiple disorders such as diabetes, heart disease, hypertension, and stroke. Obesity itself is characterized by chronic low grade inflammation with permanently increased oxidative stress (KOPP, 2019). Given the potential influence of obesity (body mass index (BMI) of 30 or above) on the size and configuration of upper airway structures, it follows that other structures involved in voice production may be affected by body mass (KOPP, 2019).

The effect of obesity on voice changes is scarcely analyzed in the literature. The results of the conducted research suggest that there is a link between vocal tract morphology and obesity. Research requires vast databases and selecting patients with specific comorbidities. In this case, more effort to obtain data must be made. The summary of databases used so far in the literature is shown in Table 3.

SOLOMON *et al.* (2011) conducted a longitudinal analysis over a period of six months on eight obese and eight non-obese adults who underwent bariatric surgical procedures. No significant differences were detected between the groups during the preoperative assessment for acoustic parameters, maximum phonation time, laryngeal airway resistance and airflow during a sustained vowel. The only minor differences were detected for strain, pitch and loudness perception of voice over time, but not between groups. Phonation threshold pressure (PTP), at comfortable and high pitches (30% and 80% of the F0 range), changed significantly over time, but not between groups. Analysis of individual data revealed a trend for PTP at 30% F0 to decrease as BMI decreased.

DA CUNHA *et al.* (2011) posited that obese individuals' voices are more aperiodic than non-obese individuals' voices, as jitter and shimmer were increased and harmonic-to-noise ratio was decreased in the former group in their study.

Table 3. Comparison of speech databases used for obesity analysis

(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
SOLOMON <i>et al.</i> , 2011	English	8	8	53.1	laryngeal imaging	severity, roughness, breathiness, strain, pitch, loudness	1	1	0	0	0	NI
CELEBI <i>et al.</i> , 2013	Turkish	20	20	38.8	BMI, laryngoscopic en	VHI, GRBAS	1	1	0	0	0	NI
DE SOUZA <i>et al.</i> , 2014	N/A	44	30	42.45	–	vocal complaint	1	0	0	0	0	00:00:10
HAMDAN <i>et al.</i> , 2014	English	9	0	35.56	BMI, laryngeal examination	simplified GRBAS	1	1	0	0	0	NI
BARSTIES <i>et al.</i> , 2013	German	22	7	21.4	BMI, body fat volume	roughness, hoarseness, breathiness, AVQI	1	1	0	0	0	NI
DE SOUZA, SANTOS, 2018	N/A	42	42	26.83	BMI		1	0	0	0	0	NI

N/A – not applicable; NI – no information.

HAMDAN *et al.* (2014) investigated 15 subjects undergoing bariatric surgery. They also found no significant difference in any of the acoustic features or in the laryngeal findings before and after surgery.

In the study conducted by CELEBI *et al.* (2013), 20 obese and 20 non-obese volunteers underwent voice evaluation by laryngoscopy, acoustic analysis, aerodynamic measurement and perceptual analysis, using the grade-roughness-breathiness-asthenia-strain (GRBAS) scale and the 10 scales voice handicap index (VHI). No differences were found in acoustic analysis parameters between the two groups ($P > 0.05$). Maximum phonation time in the obese group (mean \pm standard deviation, 19.6 ± 4.9 seconds) was significantly shorter than in the control group (26.4 ± 4.1 seconds) ($P < 0.001$), although the S/Z ratio was similar between the two groups.

DE SOUZA *et al.* (2014) verified the presence of vocal complaints and a correlation between the auditory-perceptual analysis of voice and vocal self-assessment in a group of women with morbid obesity before and after bariatric surgery. There were no statistically significant differences regarding the mean fundamental frequency of the voice in both groups; however, there was a significant difference between the two groups regarding maximum phonation.

BARSTIES *et al.* (2013) analyzed the impact of body mass and body fat volume on selected parameters of vocal quality, a phonatory range and aerodynamics in women. Significant differences between three weight groups were found across several measures of intensity: vital capacity (VC), maximum phonation time (MPT), and shimmer. As compared to other groups, significantly higher values of maximum and minimum intensity levels, as well as sound pressure level during habitual running speech, were observed for the obese group. In contrast, the underweight group had significantly lower values for VC and the ratio of expected to measured VC. Furthermore, underweight subjects differed significantly as compared to normal weight subjects with lower MPT and the higher lowest F0. Finally, the obese group showed significantly lower shimmer values than normal-range weight subjects.

DE SOUZA and SANTOS (2018) investigated the relationship between BMI and average acoustic voice features. The subjects were grouped according to BMI: 19 underweight, 23 in the normal range, 20 overweight, and 22 obese. Regarding the average F0, there was a statistically significant difference between underweight and overweight and obese groups, and the normal range and overweight and obese groups. The average MPT revealed a statistically significant difference between underweight and obese, the normal range and obese, and overweight and obese individuals. Obese women showed lower MPT.

3.4. PCOS

Polycystic ovary syndrome (PCOS) is the most common cause of hyperandrogenism in women of reproductive age, with the prevalence of 10–15%. The main characteristics of this endocrinopathy are menstrual disorders, clinical and/or laboratory hyperandrogenism and polycystic ovary morphology on ultrasonography. The elevated serum concentration of testosterone can account for symptoms such as hirsutism, acne and androgenic alopecia, as well as a deep, low voice. Moreover, other common conditions in women with PCOS are insulin resistance, disturbances of glucose metabolism, and dyslipidemia. Recent studies have shown that insulin resistance is associated with poorer verbal fluency in women (EKBLAD *et al.* 2015; SIRMANS, PATE, 2014). The significant delay in diagnosing this endocrinopathy still remains a worldwide issue. The use of speech analysis, as an easily accessible, a convenient screening test, could possibly expedite establishing a proper diagnosis and, in consequence, help provide the women with a proper treatment and an early monitoring for metabolic complications of PCOS. The speech databases and methods applied for classification to analyze PCOS diseases by speech are shown in Tables 5 and 6.

HANNOUN *et al.* (2011) found that there was no statistically significant difference in the acoustic parameters except for an increase in the relative average perturbation ($P < 0.035$) and a decrease in the maximum phonation time ($P < 0.001$) in patients with PCOS.

Table 4. Comparison of research methods used for obesity analysis (pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Disease	Tool used for features extraction	Features					Classifier/Test
			pr	sp	vs	li	naf	
SOLOMON <i>et al.</i> (2011)	obesity	MDVP	0	0	1	0	7	ANOVA
BARSTIES <i>et al.</i> (2013)		Voice Profiler 4.2, Speech Tool, Praat	0	1	1	0	20	Mann-Whitney U-test
CELEBI <i>et al.</i> (2013)		Praat	0	0	1	0	7	Mann-Whitney U
DE SOUZA <i>et al.</i> (2014)		ANAGRAF	0	0	1	0	4	–
HAMDAN <i>et al.</i> (2014)		Visi-Pitch IV	0	0	1	0	7	Wilcoxon
DE SOUZA, SANTOS (2018)		Praat	0	0	1	0	2	Mann-Whitne

Table 5. Comparison of speech databases in diabetes, thyroid and PCOS diseases (v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
Endocrine: diabetes												
CHITKARA, SHARMA, 2016	N/A	NI	NI	40–60	–	GRBAS	1	0	0	0	0	00:00:04
PINYOPODJANARD <i>et al.</i> , 2019	N/A	83	70	54	HbA1c, FBS	–	1	0	0	0	0	00:00:05
HAMDAN <i>et al.</i> , 2012	NI	82	29	52.83	HbA1c, FBS	GRBAS	1	1	0	0	0	NI
Endocrine: thyroid												
DASSIE-LEITE <i>et al.</i> , 2018	NI	100	100	3–12	altered T4, FT4, TSH	PVRQoL	1	1	0	0	0	NI
MOHAMMADZADEH <i>et al.</i> , 2011	N/A	120	88	35.9	T4, serum TSH	GRBAS	1	NI	NI	NI	NI	NI
JUNOZOVIĆ-ŽUNIĆ <i>et al.</i> , 2019	N/A	47	0	45	–	GRBAS	1	1	0	0	0	00:00:02
Endocrine: PCOS												
HUANG <i>et al.</i> , 2015	English	48	0	41–62	free T concentrations	VHI	1	1	0	0	0	NI
HANNOUN <i>et al.</i> , 2011	N/A	17	21	26	testosterone level	–	1	0	0	0	0	00:00:02
GUGATSCHKA <i>et al.</i> , 2013	German	24	10	29	endocrinologic	VHI	1	1	0	0	0	NI
AYDIN <i>et al.</i> , 2016	N/A	30	22	23.8	endocrinologic, laryngeal	VHI, GFI, RSI	1	0	0	0	0	NI

N/A – not applicable; NI – no information.

Table 6. Comparison of research methods in endocrine diseases (pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Disease	Tool used for features extraction	Features					Classifier/Test
			pr	sp	vs	li	naf	
DASSIE-LEITE <i>et al.</i> , 2018	thyroid	VOXMETRIA	0	0	1	0	4	Student’s t-test, Mann-Whitney
HAMDAN <i>et al.</i> , 2012	diabetes	VISI-PITCH IV	0	0	1	0	7	Wilcoxon Mann–Whitney rank sum, Pearson’s Chi-square
CHITKARA, SHARMA, 2016		MDVP, CSL	0	0	1	0	22	–
PINYOPODJANARD <i>et al.</i> , 2019		MDVP, CSL	0	0	1	0	7	Logistic regression
MOHAMMADZADEH <i>et al.</i> , 2011		thyroid	Visipitch III, MDVP	0	0	1	0	13
JUNOZOVIĆ-ŽUNIĆ <i>et al.</i> , 2019	thyroid	Speech Training for Windows, Dr. Speech, EZ Voice Plus	0	0	1	0	4	Paired-samples t-test
HUANG <i>et al.</i> , 2015	hysterectomy	CSL	0	0	1	0	5	Linear regression
HANNOUN <i>et al.</i> , 2011	PCOS	VISI Pitch (Model 3300)	0	0	1	0	6	Chi-square, Mann-Whitney tests
GUGATSCHKA <i>et al.</i> , 2013		MDVP	0	0	1	0	9	Student t-test
AYDIN <i>et al.</i> , 2016		Dr. Speech	0	0	1	0	8	Pearson chi-square, Fisher’s exact test, Student t-test, Mann-Whitney U

GUGATSCHKA *et al.* (2013) observed a trend towards a lower mean fundamental frequency, although it was not statistically significant. Elevated serum levels of androgens, as found in women with PCOS, were shown not to have an impact on the subjective and objective voice parameters.

AYDIN *et al.* (2016) claimed that abnormal muscle tension patterns and impaired vocal fold vibration are common in patients with PCOS, although they are not accompanied by increased vocal symptoms or deteriorated acoustic voice parameters.

HUANG *et al.* (2015) showed that testosterone administration in women with low T levels over 24 weeks was associated with dose- and concentration-dependent decreases in average pitch in the higher dose groups. These changes were seen in spite of an absence of self-reported changes in voice. The participants were healthy women, 41–62 years of age, who had undergone hysterectomy with or without partial or total oophorectomy.

3.5. Diabetes

Diabetes mellitus (DM) is a group of metabolic diseases characterized by chronic hyperglycemia which is a result of defective secretion and/or action of insulin. Type 2 DM is the most prevalent type of diabetes and concerns about 90% of diabetic patients worldwide. In the pathogenesis of Type 2 DM both mechanisms – impaired insulin action (insulin resistance) and impaired insulin secretion – play a role. Chronic hyperglycemia leads to development of diabetic complications and affects among others neurological, vascular, and muscular systems, all of which are essential components of the phonatory apparatus (HAMDAN *et al.*, 2012) and methods applied for classification used by the researchers to analyze diabetes diseases are shown in Tables 5 and 6, respectively.

HAMDAN *et al.* (2012) measured fundamental frequency, shimmer, relative average perturbation, harmonic-to-noise ratio and voice turbulence index, and reported no significant differences in any of the acoustic variables between diabetic patients and CH. There was no significant difference in the mean score of any of the perceptual evaluation parameters between diabetic patients and CH, despite the fact that mean scores were all higher in the diabetic group except for roughness. Patients with type 2 DM and poor glycemic control or neuropathy showed a significant difference in the grade GRBAS classification of their voice compared to CH.

In research conducted by CHITKARA and SHARMA (2016), the goal was to distinguish between vocal characteristics of patients with type 2 DM and control group. All the voice parameters that were investigated (jitter, shimmer, smoothed amplitude perturbation quotient, noise to harmonic ratio, relative average

perturbation, amplitude perturbation quotient) show a significant difference in their values for the diabetic group versus CH.

PINYOPODJANARD *et al.* (2019) found that F_0 in female diabetic patients was significantly lower than controls (222.23 ± 27.89 Hz versus 241.08 ± 28.21 Hz, $P < 0.01$). In female diabetic subgroups with disease duration of over 10 years, poor glycemic control or neuropathy, F_0 remained significantly lower. Multivariate analysis showed that F_0 was significantly associated with diabetes after controlling for age, BMI, presence of hypertension, and dyslipidemia. However, F_0 was not able to predict the presence of diabetes as shown by the logistic regression analysis ($P = 0.243$).

3.6. Hypothyroidism and hyperthyroidism

Hypothyroidism is the state of insufficient hormone production by the thyroid gland. Commonly reported symptoms in patients with this condition are hoarseness, deep or weak voice, vocal fatigue and tension while speaking as a result of vagus nerve edema, laryngeal muscle weakness and vocal cord paresis caused by an enlarged thyroid gland. Hyperthyroidism, as a state of increased thyroid hormone secretion, can also significantly reduce voice intensity and deepen its timbre. Hoarseness, roughness and trembling voice are also observed (JUNOZOVIĆ-ŽUNIĆ *et al.*, 2019). The research databases used in literature are shown in Table 5 and methods for further speech analysis are shown in Table 6.

MOHAMMADZADEH *et al.* (2011) found that F_0 , voice turbulence index and soft phonation index were significantly different from control values. There was positive correlation between thyroid-stimulating hormone (TSH) concentration and variation in F_0 and prevalence of voice disorders.

DESSIE-LEITE *et al.* (2018) led an observational, analytical, cross-sectional study including 200 pre-pubertal children, of whom 100 had congenital hypothyroidism. The following parameters were evaluated: 1) history (identification, complaints, and interfering variables); 2) auditory-perceptual and acoustic evaluation; 3) self-assessment scores in the Pediatric Voice-Related Quality-of-Life (PVRQoL) survey; 4) laryngological evaluation; 5) medical records (congenital hypothyroidism etiology, age at treatment initiation, disease severity at diagnosis, treatment quality, and thyroid function tests on the day of the examination). Both groups had mean/median acoustic measurements within normal limits. There was no association between voice/larynx characteristics and endocrinological data.

JUNOZOVIĆ-ŽUNIĆ *et al.* (2019) reported that patients with hypothyroidism displayed significant differences in amplitude perturbation, jitter and noise-to-harmonics ratio between pre-treatment and post-

treatment periods. In the group of patients with hyperthyroidism, significant differences were noted in the aerodynamic parameter maximum phonation time only. There were significant differences in all perceptual parameters in both groups of patients ($P < 0.05$) in pre- and post-treatment, except in the grade and asthenia parameter in the group of patients with hypothyroidism. The parameter grade was a border line insignificant in the group of patients with hyperthyroidism.

3.7. Mental and neurodegenerative disorders

There is a high volume of publications on the diagnosis of neurodegenerative and mental disorders using speech analysis. Most of the works concern the acoustic analysis of speech, but there are publications informing about the high effectiveness of linguistic analysis, an example is (STASAK *et al.*, 2017).

The application described in (KISS *et al.*, 2021) is capable of estimating the probability of three types of voice disorders in English and Hungarian: depression, dysphonia, and Parkinson’s disease.

VILLATORO-TELLO *et al.* (2021) used the available lexicon for a mentally ill and control subjects in a classification process to detect depression and dementia.

3.7.1. Schizophrenia

Schizophrenia is a chronic psychiatric disorder that affects 1% of the world’s adult population. Language, thought and communication dysfunction characterize all its symptoms. They can be broadly divided into two

groups: positive and negative. Positive thought disorder leads to a discourse that is difficult to understand (derailment, contact, neologisms, etc.). The research databases and conducted research are summarized in Tables 7 and 8, respectively.

MOTA *et al.* (2012) compared patients with schizophrenia and mania. Global speech graph measures were not significantly different for the groups; however, patients with schizophrenia produced significantly fewer words per report than patients with mania. The authors observed poor speech, logorrhea and flight of thoughts.

BEDI *et al.* (2015) reported a proof-of-principle study which aimed to test automated speech analysis combined with machine learning to predict later psychosis onset in youths at clinical high-risk (CHR) for psychosis. Thirty-four CHR youths had baseline interviews and were assessed quarterly for up to two and a half years; five transitioned to psychosis. Speech features included a latent semantic analysis (LSA) measure of semantic coherence and two syntactic markers of speech complexity: maximum phrase length and the use of determiners (e.g., which). These speech features predicted later psychosis development with 100% accuracy, outperforming classification from clinical interviews. Speech features were significantly correlated with prodromal symptoms.

GOSZTOLYA *et al.* (2018) matched 10 subjects with schizophrenia and eight HC with age and gender. The speakers performed spontaneous speech about their previous day for 5 min. The automatic speech recognition (ASR) system was trained to recognize the phones

Table 7. Comparison of speech databases used for schizophrenia analysis

(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
MOTA <i>et al.</i> , 2012	Spanish	16	8	35.75	DSM-IV, BPRS, PANSS		0	0	0	1	0	00:20:00 –00:40:00
BEDI <i>et al.</i> , 2015	English	34	0	14–27	SIPS/SOPS	–	0	0	0	1	0	01:00:00
GOSZTOLYA <i>et al.</i> , 2018	Hungarian	10	8	39.9	MMSE	–	0	0	1	0	0	NI

NI – no information.

Table 8. Comparison of research methods in schizophrenia examination

(pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Disease	Tool used for features extraction	Features					Classifier/Test
			pr	sp	vs	li	naf	
Mental								
MOTA <i>et al.</i> , 2012	schizophrenia	Network Analysis Toolkit, MATLAB	0	0	0	1	11	NB, RBF, MLP, SVM, DT
BEDI <i>et al.</i> , 2015		Natural Language Toolkit (NLTK), POS-Tag Pen Tree Bank	1	0	0	1	7	Convexhull classifier
GOSZTOLYA <i>et al.</i> , 2018		ASR with DNN,	1	0	0	0	8	SVM

in the utterances. For acoustic modeling, a standard deep neural network (DNN) with feed-forward topology was applied. A detailed examination revealed that, among the pause-related temporal parameters, those which took into account both the silent and filled pauses were the most useful in distinguishing the two speaker groups.

In studies of links between voice and schizophrenia, more research has been conducted in the area of voice perception (e.g., PINHEIRO, NIZINKIEWICZ, 2019). Their findings support the hypothesis that higher-order operations reflected in amplitude modulations are abnormal in schizophrenia in a valence-dependent manner. The altered detection of vocal changes with a positive quality may lead to deficits in the comprehension of emotional states and intentions of social partners during vocal communication.

3.7.2. Depression and bipolar disorder

Reduced speech activity in patients with depression, especially with psychomotor impairment, is confirmed by many systematic studies. A number of clinical observations suggest that changes in voice features, such as pitch, may be important measures in diagnosing the early-stages of depression. They can also assess the progress of treatment for a depressive episode. Speech pause times (SPT), a silent interval between phonations during automatic speech, can be useful as an objective pathophysiological marker in depression. In clinical remission, depressive patients had comparable SPT values to the CH group. Studies show that in bipolar disorder, increased speech activity can predict a switch to hypomania. Depression is characterized by psychomotor retardation; in speech,

Table 9. Comparison of speech databases in bipolar and depression disorder (v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
Mental: bipolar												
GRÜNERBL <i>et al.</i> , 2014	German	NI	NI	18–65	HAMD, YMRS	–	0	0	0	1	0	NI
GUIDI <i>et al.</i> , 2015	NI	11	18	NI	QID, YMRS	–	0	1	0	0	0	NI
FAURHOLT-JEPSEN <i>et al.</i> , 2016	Danish	28	0	30.3 ± 9.3	HAMD, YMRS	–	0	0	0	1	0	NI
ALGHOWINEM <i>et al.</i> , 2016	English	30	30	21–75	DSM-IV	–	0	0	0	1	0	00:08:33
YANG <i>et al.</i> , 2012	English	57	0	39.7(19–65)	HRS-D	–	0	0	0	1	0	NI
MUNDT <i>et al.</i> , 2012	English	165	0	37.8 ± 12.5	HAM-D, DSM-IV, QIDS-C, QIDS-SR	–	1	1	1	0	0	NI
VALSTAR <i>et al.</i> , 2013	German	NI	NI	31.5 ± 12.3	BDI-II	Valence, Arousal	1	1	1	0	1	00:25:00
HÖNIG <i>et al.</i> , 2014	German	219	0	31.5 ± 12.9	BDI	–	0	1	1	0	0	00:08:08
DEVVAULT <i>et al.</i> , 2014	English	351	0	45.6 ± 12.2	–	–	0	0	0	1	1	00:17:30
AFSHAN <i>et al.</i> , 2018	Mandarin	735	953	NI	DSM-IV	–	0	0	0	1	0	00:01:52
MCGINNIS <i>et al.</i> , 2019	English	71	0	3–8	TSST-C, K-SADS-PL	–	0	0	1	0	0	00:03:00
SCHERER <i>et al.</i> , 2013b	English	43	0	41.2 ± 11.6	PHQ-9, PCL-C	–	0	0	0	0	1	00:60:00
GRATCH <i>et al.</i> , 2014	English	110	NI	18–65	PCL-C, PHQ-9, STAI-T, BIDR, BFI, RME, PANAS	–	0	0	0	1	0	00:38:11

NI – no information.

this shows up in reduction of pitch (variation, range), loudness, tempo and in voice qualities different from those of typical modal speech (MUNDT *et al.*, 2012; FAURHOLT-JEPSEN *et al.*, 2016). The comparison of speech databases and conducted research summary are presented in Tables 9, 10, and 11.

One of the most popular corpora used for depression research is the Mundt database. Thirty five physician-referred patients beginning treatment for depression were assessed weekly, using standard depression severity measures during a six-week observational study. Speech samples were also obtained over the telephone each week using an interactive voice response (IVR) system to automate data collection (FAURHOLT-JEPSEN *et al.*, 2016).

HELPER *et al.* (2013) measured articulatory precision manifested through formant frequency tracking. GMM and SVM were applied using the Mundt database. They showed that a depression state can be classified with only formant frequencies and their dynamics given by the velocity and acceleration.

CUMMINS *et al.* (2015a) also used the Mundt corpus and the GMM model, although they focused on the spectral and energy-based properties of speech. They stated that depression-induced changes are in the

laryngeal coordination and the vocal tract behavior. They show that depression is associated with a decrease in the pitch variability, changes in formant frequencies and decreases in the sub-band energy variability.

HÖNIG *et al.* (2014) observed a similar speech disorder in depression and sleepiness. They employed a small group of acoustic features, modeling prosody and spectrum, enriched with voice quality. The dataset comprises 1122 recordings from 219 German subjects (66 male); mean age 31.5 years, total duration of all files 29.7h. The database consists of read and spontaneous speech. The length of the speech tasks is between 5.8sec and 5.3min (mean 1.6min).

SCHERER *et al.* (2013a; 2013b) examined another group of features – voice quality related, especially in the breathy-tense dimension – and used SVM as a classifier. The authors introduced a new dataset: virtual human distress assessment interview corpus with patients with depression and PTSD.

Experiments presented by CUMMINS *et al.* (2015b) support the hypothesis that reduction in acoustic variation of speech is produced under more severe depression. Their dataset comprises the Mundt corpus and AVEC2013. MFCC features were extracted. Universal

Table 10. Comparison of research methods and results in mental disorders (pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Database	Tool used for features extraction	Features					Classifier/Test	Evaluation metrics	Best score	Tag
			pr	sp	vs	li	naf				
Mental: depression											
ALGHOWINEM <i>et al.</i> , 2012	Black Dog	openSMILE	1	1	1	0	NI	HMM, GMM	VAR, F1	VAR = 71%	2
MUNDT <i>et al.</i> , 2012	Mundt 2	Praat	1	1	1	0	12	logistic regression, t-test	Log Odds Response, <i>p</i> -value	–	1
CUMMINS <i>et al.</i> , 2013	Mundt 1, Black Dog	openSMILE	0	1	0	0	39	GMM-UBM	Acc	~65%	2
SCHERER <i>et al.</i> , 2013a	VHDAIC	NI	1	1	1	0	4	SVM	Acc, F1	75%	2
VALSTAR <i>et al.</i> , 2013	AViD	openSMILE	1	1	1	0	2268	CVR	RMSE, MAE	10.35, 14.12	2
HÖNIG <i>et al.</i> , 2014	AVEC2014	openSMILE	1	1	1	0	3805	RLR	spearman's ρ , pearson's <i>r</i>	-0.46	1
BOZKURT <i>et al.</i> , 2014	Mundt 1	openSMILE, Praat	0	1	0	0	2860	SVM	UAR	69.48%	2
CUMMINS <i>et al.</i> , 2015c	Mundt, AVEC2013	openSMILE	0	1	1	0	2268	RVM	RMSE	10.89	2
ZHAO <i>et al.</i> , 2020	AVEC2013, AVEC2014	openSMILE	0	1	1	0		DCNN	RMSE, MAE	9.57 7.9	2
SENEVIRATNE <i>et al.</i> , 2020	Mundt	Aperiodicity, Periodicity, Pitch (APP) detector	0	1	1	0	20	SVM	Acc	81.77%	2
ALGHOWINEM <i>et al.</i> , 2016	Black Dog, Pitt, AVEC	openSMILE	1	1	1	0	504	SVM	AR	96.90%	2

NI – no information.

Table 11. Comparison of research methods and results in mental disorders
(pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Database	Tool used for features extraction	Features					Classifier/Test	Evaluation metrics	Best score	Tag
			pr	sp	vs	li	naf				
STASAK <i>et al.</i> , 2016	AVEC2014	openSMILE	1	1	1	0	2155	SVM, RVM	RMSE	11.20%	2
SHAU, ESPY-WILSON, 2016	Mundt 1	NI	0	0	1	0	NI	SVM	Acc	62–87%	2
SIMANTIRAKI <i>et al.</i> , 2017	AVEC2014	openSMILE	0	0	1	0	247	JAD	AUC	0.88	2
STASAK <i>et al.</i> , 2017	DAIC	COVAREP speech toolkit	0	1	0	1	74	DT	Acc	82%	2
AFSHAN <i>et al.</i> , 2018	CONVERGE	openSMILE, VoiceSauce	0	1	1	0	>6300	GMM + <i>i</i> -vectors	P, R, F1, A	F1 = 0.95	2
AL HANAI <i>et al.</i> , 2018	DAIC	NI	1	1	1	1	553	NN LSTM	MAE, RMSE	MAE = 4.97 RMSE = 6.27	2
XEZONAKI <i>et al.</i> , 2020	DAIC, General Psychotherapy Corpus (GPC)	NI	1	1	1	1		SVM, hierarchical attention-based Network	UAR, F1-macro	UAR = 0.72 F1-macro = 0.69	2
MCGINNIS <i>et al.</i> , 2019	own	MATLAB	0	1	1	0	NI	LR, SVM	Acc, Sens, Spec	80%, 54%, 93%	2
Mental: bipolar disease											
GRÜNERBL <i>et al.</i> , 2014	own	openSMILE	1	1	1	1	17	NB	Acc, recall, precision	70%	2
GUIDI <i>et al.</i> , 2015	own	–	1	0	1	0	2	non-parametric Friedman test for paired data	$P < 0.05$	–	1
FAURHOLT-JEPSEN <i>et al.</i> , 2016	own	openSMILE	0	1	1	0	6552	RF	Acc, Sens, Spec	0.74, 0.97, 0.52	2
GUIDI <i>et al.</i> , 2017	own	NI	1	0	1	0	12	Friedman's test for paired data, Mann-Whitney U-test	$P < 0.05$	–	1

NI – no information.

background models (UBMs) were trained with the expectation-maximization (EM) algorithm. Speaker-specific GMMs were formed using full adaptation, with five iterations of the maximum a posteriori (MAP) algorithm.

The goal of research conducted by BOZKURT *et al.* (2014) was to find speech features that can distinguish speaking patterns of individuals with a diagnosis of clinical depression on a speaker-independent basis. Speech parameters were obtained from spectral analysis. Experiments for two-class depressed vs. non-depressed subjects were performed using SVM classifiers implemented on free speech recordings. Features advanced recognition rates up to 69% of the arithmetic average of individual class accuracies.

STASAK *et al.* (2016) showed performance boosting for the depression classification by using speech affect ratings in combination with low level features

vs using these descriptors alone. They also showed the importance of setting thresholds for data selection for a specific affect. An automatic emotion-rating system derived from GeMAPS may positively contribute to performance in combination with low level features.

KHORRAM *et al.* (2016) took an approach for depression detection in bipolar patients. They combined two systems. The first was patient-specific and used unlabeled personal calls along with assessment calls to develop a unified background model and *i*-vectors, respectively. The second system was cohort-general and based on rhythm features. Their results showed improvement from the baseline with the unweighted average recall increasing from 0.66 ± 0.11 to 0.73 ± 0.09 and the area under the receiver operating curve increasing from 0.69 ± 0.15 to 0.78 ± 0.12 .

HUANG *et al.* (2016) collected six speech responses from 30 subjects – 15 with unipolar depression (UD)

and 15 with bipolar disorder – and used the hierarchical spectral clustering (HSC) algorithm to adapt the larger Multimedia Human-Machine Communication (MHMC) emotion database to the obtained mood database CHI-MEI. Experimental results show that the proposed method achieved 73%, improving the detection accuracy by 13% compared to the commonly used SVM-based classifiers.

SAHU and ESPY-WILSON (2016) explored features which are important in detecting depression. They used short audio samples of sustained vowels (5–6 s) and longer (30 s – 2 min) of free speech from the Mundt database and explored breathiness, jitter, and shimmer-based features using the average magnitude difference function (AMDF) to quantify them. Using the AMDF based feature they got 62–87% frame-wise accuracy for 5 out of 6 speakers.

ALGHOWINEM *et al.* (2016) investigated the feasibility of cross-cultural UD detection from z -score normalized prosody features. They used German and English datasets (AVEC, BlackDog, and Pitt, respectively). Authors showed that binary (depressed/mild or not depressed) classification trained on SVM performs very well on each individual dataset.

LOPEZ-OTERO *et al.* (2017) presented a study on how speaker de-identification affects the performance of a depression-detection system based on speech transcriptions. For depression detection, it is necessary to know which words are related to depression.

STASAK *et al.* (2017) proposed a novel measure for quantifying articulation effort and demonstrated that depression classification can be achieved by selecting speech with the higher articulation effort, linguistic complexity, or word-based arousal/valence.

SIMANTIRAKI *et al.* (2017) tested the hypothesis that UD could be detected from glottal source signals by using discrete cosine transform (DCT) coefficients of phase distortion deviation (PDD) preceded by the feature selection using Just Add Data (JAD). The models (SVM, ridge logistic regression and random forest) were evaluated on read and spontaneous speech from the AVEC2014 dataset. The researchers concluded that the lack of harmonic clarity is mainly disruptive in the region above F1 where harmonic amplitudes are relatively low.

AFSHAN *et al.* (2018) stressed that depression can be characterized by prosodic abnormalities and/or articulatory and phonetic errors. They verified various aspects of speech signals: cepstral features, difference in harmonic amplitudes, formant amplitudes, and i -vector-based systems using MFCCs and score-level fusion to combine the two systems.

3.7.3. Parkinson's disease

PD is the second-most prevalent neurodegenerative disease in the world. It is caused by a loss of dopaminergic neurons and it causes severe motor and cognitive dysfunctions. It is characterized by hypophonia (reduced voice volume) and dysphonia (breathiness, hoarseness or creakiness in the voice), typically preceding more generalized speech disorders. About 90% of PD patients develop speech impairments such as monopitch, monoloudness, imprecise articulation and other symptoms (OROZCO-ARROYAVE *et al.*, 2014a; 2014b). The summary of speech databases and research methods are presented in Tables 12 and 13.

In comprehensive review MORO-VELAZQUEZ *et al.* (2021) identified the most common features and ma-

Table 12. Comparison of speech databases in PD analysis
(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
PINTO <i>et al.</i> , 2016	French, Portugese	139	65	65–70	NI	–	0	1	1	1	0	14:54:36
SKODDA <i>et al.</i> , 2011	German	138	50	66.74(8.48)	UPDRS III, H&Y	–	1	1	0	0	0	NI
RUSZ <i>et al.</i> , 2018	Czech	78	30	62.3(11.3)	MDS-UPDRS III	MDS-UPDRS III	1	0	1	0	0	NI
OROZCO-ARROYAVE <i>et al.</i> , 2014a	Spanish	50	50	62.2(11.2)	UPDRS-III and H&Y	–	1	1	1	0	0	00:01:34
PETTORINO <i>et al.</i> , 2017	Mandarin	13	12	62.1(52–72)	NI	–	0	1	0	0	0	NI
ZHAN <i>et al.</i> , 2016	English	121	105	57.6(9.1)	UPDRS	–	1	0	0	0	0	NI
HEMMERLING <i>et al.</i> , 2016	Polish	27	0	65(7.9)	UPDRS, H&Y	–	1	1	0	0	0	00:00:59
ANATOLÍK, FOUGERON, 2013	French	79	26	32–89	NI	intelligibility and articulatory imprecision	0	1	0	0	0	NI

NI – no information.

Table 13. Comparison of research methods and results in Parkinson’s disease (pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Database	Tool used for features extraction	Features					Classifier/Test	Evaluation metrics	Best score	Tag
			pr	sp	vs	li	naf				
OROZCO-ARROYAVE <i>et al.</i> , 2014b	German, Spanish, Czech	Praat	0	1	1	0	68	SVM	Acc Sens, Spec, AUC	Acc 97.6%	2
MALLELA <i>et al.</i> , 2020	Indian	NI	0	1	0	0	NI	CNN + Bidirectional Long Short-Term Memory	Acc	97.41%	2
POMPILI <i>et al.</i> , 2020	Portugese	HTK	1	1	1	0	52	DNN TensorFlow	Acc	95.27%	2
VÁSQUEZ-CORREA <i>et al.</i> , 2015	PC-GITA, German, Czech	NI	0	1	0	0	NI	CNN	Acc	89%	2
VILLA-CANAS <i>et al.</i> , 2015	PC-GITA	NI	0	1	1	0	NI	GMM, GMM-UBM, SVM	Acc	77%	2
PETTORINO <i>et al.</i> , 2017	Mandarin, Polish, Italian	NI	1	0	0	0	2	t-test	$P < 0.05$	–	1
ZHAN <i>et al.</i> , 2016	own	NI	1	0	1	0	9	RF	Acc	71.0(± 0.4)%	3
KLUMPP <i>et al.</i> , 2017	NI	NI	1	0	0	0	NI	–	SER	1.34	3
RUSZ <i>et al.</i> , 2018	own	NI	1	0	1	0	NI	binary logistic regression	AUC	0.85	2
WODZINSKI <i>et al.</i> , 2019	PC-GITA	NI	0	1	0	0	NI	LSTM	Acc	0.917	2

NI – no information.

chine learning techniques employed in automatically detecting and assessing the severity of PD using phonatory and articulatory aspects of speech and voice.

OROZCO-ARROYAVE *et al.* (2014b) investigated the detection of PD by analyzing speech recordings in German, Spanish, and Czech. The Spanish database contains speech recordings of 50 PD patients (mean age 61), the set of German patients includes 88 individuals (mean age 66.5), and in the group of Czech native speakers 21 were diagnosed with idiopathic PD (mean age 62.2). For each language a CH group with the same number of individuals was selected. The databases contain recordings of sustained phonation of vowels, diadochokinetic (DDK) evaluation, sentences, and monologues. Three features in the phonation process were taken into account: harmonics-to-noise ratio, normalized noise energy, and glottal-to-noise excitation ratio. Additionally, the first and second formant and 12 MFCC were extracted. Authors suggest that it is possible to detect PD using the same method in different languages. Additionally, OROZCO-ARROYAVE

et al. (2014a) presented further tests using jitter and shimmer.

VILLA-CANAS *et al.* (2015) analyzed low-frequency components of speech signals by using three different time-frequency techniques. Their results showed that the changes in the low frequency components are able to discriminate between people with Parkinson’s and healthy speakers with an accuracy of 77%, using a single sentence.

ZLOTNIK *et al.* (2015) considered four groups of features: phonation, articulation, prosody, and intelligibility. Many techniques were tested for predicting the stage of PD using the patient’s voice exclusively. Finally, an ensemble of classifiers obtained the best results (0.609), combining the output of the best Random Forest, with intelligibility features.

VÁSQUEZ-CORREA *et al.* (2015; 2017; 2018) found that voice impairments appear in about 90% of speech samples as reduced loudness, monopitch, monoloudness, reduced stress, breathy, hoarse voice quality, and imprecise articulation. Speech samples were subjected

to frequency analysis: F0, its variability and MFCC (VÁSQUEZ-CORREA *et al.*, 2015), wavelet transform and short-time Fourier transform (VÁSQUEZ-CORREA *et al.*, 2017; 2018). Physical parameters of voice were used for modeling the SVM and the convolutional neural network (CNN). In the publication from 2015, they declared the highest accuracies with the voiced frames a range from 64% to 86%, while the results with unvoiced frames range from 78% to 99%. In the paper published in 2017, they mentioned accuracies of up to 89% for the classification of Parkinson's patients vs. control group when a convolutional neural network was used to extract features from the time-frequency representations. Their study published in 2018 proposes a multitask learning approach based on CNNs to assess at the same time eleven speech aspects, including difficulties of the patients to move articulators such as lips, palate, tongue, and larynx. The input to the CNNs are time-frequency representations obtained from transitions between voiced and unvoiced segments.

HEMMERLING *et al.* (2016) explored phonatory and articulatory features for modulated vowels in PD detection. They used jitter, shimmer, F0, energy and MFCC statistical values along side instantaneous energy and its range coming from the Hilbert-Huang transformation. The data used contained sustained and modulated vowels. As the result, the sustained vowels covered higher binary accuracy classification.

PETTORINO *et al.* (2017) investigated whether speech of PD patients presents rhythmic abnormalities. Twenty-five Mandarin speakers (13 PD and 12 HC matched on age) and thirty-one Polish speakers (18 PD and 13 HC matched on age) read aloud a passage of story. The vowel percentage and the interval between two consecutive vowel onset points were calculated. They segmented the recorded speech into vocalic and consonantal intervals, and then calculated the vocalic portion in the utterance and the duration of the interval between two consecutive vowels. The effectiveness of the rhythmic metric appears to be language-dependent. For Polish was distinctly higher while for Mandarin there was no significant difference. They concluded that the analyzed method could be used for automatic diagnosis of PD for Polish and Italian, but not for Mandarin.

KLUMPP *et al.* (2017) introduced Apkinson – a smartphone application providing a mobile monitoring solution for PD patients. The severity and progression of PD can be tracked. The patient has to perform a speech exercise in which the person has to constantly produce syllables of subsequent consonant-vowel combinations. The Levenshtein distance evaluates the similarity by computing the required insertions, deletions and substitutions to transform one string to the other. For the recordings of the HC, the recognizer correctly counted the number of keywords in 76% of the cases. For the

group of PD patients, the number of correctly assessed records was slightly lower with 72%.

The most common scale for assessing the severity of PD is the Unified Parkinson's Disease Rating Scale (UPDRS) and Movement Disorders Society UPDRS (MDS-UPDRS), which evaluates non-motor and motor experiences of daily living and motor complications.

The Computational Paralinguistics Challenge Special Session of Interspeech 2015 was dedicated, alongside two other topics, to the degree to which PD can be detected by speech analysis. Recordings of 50 PD patients were provided. The dataset was divided into 42 tasks per speaker, yielding 1470 recordings in the training set (3 speakers), 630 recordings in the development set (15 speakers), and 462 recordings (11 speakers) in the test set. The duration of recordings ranges from 0.24 seconds to 154 seconds and consist of a monologue, a read text and sentence recordings.

GRÓSZ *et al.* (2015) applied two state-of-the-art machine learning methods in the regression Sub-Challenges of the Interspeech 2015 Computational Paralinguistics Challenge. They showed that both DNN and Gaussian process regression (GPR) are competitive with the baseline SVM, and the results can be improved by combining the classifiers. They trained DNNs with five hidden layers and 1000 neurons in each hidden layer. The best results (0.671) they obtained by using a feature selection and by averaging out the scores of multiple recordings clustered to the same person.

SZTAHÓ *et al.* (2015) presented the method of linear regression models on a set of extracted acoustic features from the middle of vowels in words, sentences and continuous speech, and the partitioning of speech samples according to their total length into parts with long, medium and short duration. Jitter, shimmer, articulation rate, intensity and its variation, rate of transients and MFCC were extracted. They notice that in terms of the final results on the test set that was uploaded for the challenge, many conclusions cannot be deduced due to the variation in the data. They emphasized that correlations (and also the baseline results) count as weak. They experienced high intra-variation of the extracted features.

3.7.4. Alzheimer's disease and dementia

Alzheimer's disease (AD) is a progressive neurodegenerative disorder clinically defined as an impairment of certain cognitive and functional abilities. As the result of the aging society, there are growing number of people affected by AD.

GOSZTOLYA *et al.* (2019) took an automatic approach focusing on features describing the number of pauses in spontaneous speech, specifically filled gaps. The authors constructed a large set of descriptors and used correlation and the sequential forward selection algorithm to find the most promising ones. Based on

only the acoustic features, they were able to separate the various groups with accuracy scores between 74–82%. They attained similar accuracy scores using only the linguistic features. With the combination of the two types of features, the accuracy scores between 80–86%.

SADEGHIAN *et al.* (2017) presented empirical evidence that AD patients can be reliably distinguished from HC through a combination of acoustic features from speech and linguistic features extracted from an automatically determined transcription of speech.

WANKERL *et al.* (2017) proposed a purely statistical approach towards automatic diagnosis of AD, solely based on n-gram models with subsequent evaluation of the perplexity. The system works independently in a concrete language. AD patients show emotional prosodic impairment.

The Pitt corpus used by WARNITA *et al.* (2018) consists of speech samples and their transcriptions from 244 HC and 309 dementia patients. They used a gated convolutional neural network (GCNN) for speech data. The presented study is the non-linguistic approach for detecting AD by utilizing only the speech audio data. Since it does not utilize linguistic information, authors can apply it to low resource languages. The proposed method achieved the accuracy of 74%.

WEINER *et al.* (2016; 2018) investigated a way of automatic classification of AD using conversational speech. They derived 10 prosodic and textual features from 98 voice samples from the interdisciplinary longitudinal study on adult development and the aging (ILSE) dataset. They compared two pipelines of fea-

ture extraction for dementia detection: manual transcription and ASR using samples from the ILSE corpus. The acoustic and linguistic features were extracted and several models were built with the Gaussian classifier as the top performer. WEINER *et al.* (2018) stated that early detection of dementia is possible by automatically processing conversational speech. They tested a group of more than 200 subjects. Conversational speech (12 min of interview) was chosen since it is a natural form of communication that can be recorded without causing stress to subjects. The best results were obtained through a combination of acoustic and linguistic features. Finally, a Gaussian classifier was trained to discriminate three cognitive diagnoses. The authors declare that it is possible to detect dementia using speech of duration 2.5 min, although the most reliable results require between 10 and 15 min. In the publication of 2016, the authors declared the F-score of 0.8 for the detection of AD. In the paper from 2018, they stated detected dementia with an UAR of 0.64 using acoustic features extracted from speech segments.

ROHANIYAN *et al.* (2021) present two multimodal fusion-based deep learning models that consume ASR transcribed speech and acoustic data simultaneously to classify whether a speaker in a structured diagnostic task has AD. They achieved an accuracy of 84% using words, word probabilities, disfluency features, pause information, and a variety of acoustic features.

The summary of speech databases and research methods applied for AD and dementia analysis are shown in Tables 14 and 15.

Table 14. Comparison of speech databases in Alzheimer’s disease and dementia

(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
SATTLER <i>et al.</i> , 2017	German	NI	NI	>40	medical, psychological, cognitive, physical, dental examinations, semi-standardized biographic interview	–	0	0	0	1	0	NI
UJIRO <i>et al.</i> , 2018	Japanese	12	12	74.5 ± 4.3	DSM-IV-T	–	0	1	0	0	1	NI

NI – no information.

Table 15. Comparison of research methods and results in Alzheimer’s and dementia disorder

(pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Database	Tool used for features extraction	Features					Classifier/Test	Evaluation metrics	Best score	Tag
			pr	sp	vs	li	naf				
WARNITA <i>et al.</i> , 2018	Pittcorpus	openSMILE	0	1	1	0	12459	GCNN	Acc	73.6%	2
WEINER <i>et al.</i> , 2018	ILSE	NI	1	1	0	0	40	GMM	UAR	0.645	2
UJIRO <i>et al.</i> , 2018	own	Snack Sound Toolkit	0	0	1	1	21	SVM and logistic	AUROC	0.95	2
PAN <i>et al.</i> , 2020	DementiaBank dataset	NI	1	0	1	1		bi-LSTM	F-score	78.34%	2
MIRHEIDARI <i>et al.</i> , 2018	own	Praat	1	0	1	1	12	HMM-GMM	Acc	91%	2

NI – no information.

Table 16. Comparison of speech databases in ALS
(v – sustained vowel; r – read speech; m – monologue; h – dialog with human; a – dialog with virtual agent).

Published	Speaker language	Number of patients	Number of HC	Age	Clinical evaluation	Voice evaluation	Type of speech					Duration per speaker [h:min:s]
							v	r	m	h	a	
HORWITZ-MARTIN <i>et al.</i> , 2016	English	34	0	NI	–	SIT	1	1	0	0	0	NI
WANG <i>et al.</i> , 2016	English	11	11	60	ALSFRS-R	SIT	1	1	0	0	0	NI

NI – no information.

Table 17. Comparison of research methods in ALS
(pr – prosodic; sp – spectral; vs – voice source; li – linguistic; naf – number of acoustic features).

Published	Disease	Tool used for features extraction	Features					Classifier/Test
			pr	sp	vs	li	naf	
Neurodegenerative								
HORWITZ-MARTIN <i>et al.</i> , 2016	ALS	MATLAB	1	1	0	0	36	Spearman correlations, regression
WANG <i>et al.</i> , 2016		openSMILE	1	1	1	0	6373	RLR, <i>i</i> -vectors, SVM, DNN

3.7.5. Amyotrophic lateral sclerosis

ALS is a rapidly progressing disorder that causes the death of neurons controlling voluntary movements. Symptoms include difficulty in speaking or swallowing.

ANTOLÍK and FOUGERON (2013) found that among consonant distortions, the most frequent type of distortion in ALS is an incomplete closure of stops.

WANG *et al.* (2016) explored the option to diagnose ALS from short speech acoustic and articulatory samples. They examined 11 affected patients and 11 HC, and constructed a large dimensionality dataset of acoustic features from audio samples and articulatory features derived from tongue and lip sensors. Furthermore, randomized logistic regression (RLR) was used as a feature selection method, and the *i*-vector was calculated for each speaker and concatenated to feature the vector for speaker normalization. Next, they trained two classes of models: SVM with a radial basis function (RBF) kernel, and DNN achieving maximum performance when all acoustic and sensor features were combined and provided to train the DNN model.

HORWITZ-MARTIN *et al.* (2016) presented research into an objective and automatic assessment of speech loss with features extracted from the first and second formant. They found that acceleration features derived from F2 were the most informative for speech predicting the rate of speech decline and assessing the intelligibility decline.

The speech databases and research methods applied for ALS analysis are shown in Tables 16 and 17.

4. Discussion

4.1. Recruitment process

To create an effective computer system it is necessary to have the right number of recordings of individ-

uals affected by each given disorder. Each patient must have an official diagnosis from a physician. Recordings taken at early disease stages are the most desirable since the main purpose of computer systems is to recognize the onset of the disease. Sometimes it is difficult to convince patients of the desirability of collecting recordings. Patients must consent to participate in research. Thus, the recruitment process is difficult and frequently time-consuming. Collecting recordings requires approval from the Ethics Committee.

4.2. Standardization of database descriptions

We commonly encountered insufficient descriptions of the collected speech databases. As shown in Tables 1, 3, 5, 7, 9, 12, 14, 16 only some of the parameters of recordings are reported systematically across the corpora. The descriptions frequently omit the duration of the recordings or detailed technical information on the conditions of the recording process. We propose to standardize the descriptions as a table or an annex, which would clearly outline the properties of the corpus, especially when the database is not publicly available. Important parameters assessing the usefulness of recording databases are the number of recorded speakers (Fig. 4), the duration of each speech and total recording times representing the size of the database speaker (Tables 1, 3, 5, 7, 9, 12, 14, 16). In addition, it may be useful to divide the recordings by gender. Generally, patients' age ranges are given, although some authors present the average age of patients and standard deviation. Similar information is presented regarding the age of CH. Comparing these data sets shows that frequently the patients are more advanced in age than the CH group. The correlation between the patient age and their health may affect the functionality of the classifiers. An older patient's voice may be incorrectly classified as a voice of an ill

patient. Conversely, a young person's voice may be incorrectly classified as a voice of a healthy individual.

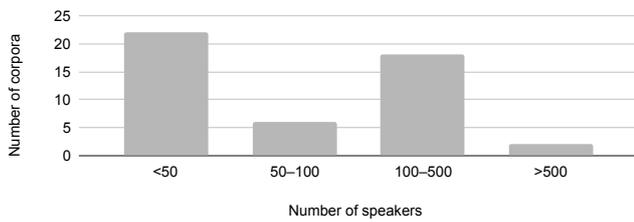


Fig. 4. Corporas from 2009–2019.

The recordings were divided into three groups. The first includes read texts or, far less frequently, individual words or chosen vowels. Such recordings can be used to analyze the same statements for all respondents, which is a valuable advantage when analyzing the acoustic features of voices. In this case, there is no dependency on the content of the statements. The second group includes recordings of spontaneous statements. These enable linguistic analysis, which is especially useful when diagnosing mental illness (STASAK *et al.*, 2017; LOPEZ-OTERO *et al.*, 2017). The third group includes recordings from dialogues between subjects and interviewers (WEINER *et al.*, 2016). Conversational speech is the most natural form of communication that can be recorded without putting subjects under stress. Moreover, acoustic and language information can be obtained from conversations between patients and virtual agents. Generally, subjects receive instruction on how to speak. This applies to duration, avoiding emotions, monologue topics and sometimes loudness. For example, in sporadic cases, patients are asked to give a prolonged statement of the desired vowel.

The most frequently recorded statements are in one language only (sporadically dialect). The main reason is that such recordings are the easiest to obtain. Their effective analysis is also easier, especially linguistic analysis.

Recordings can be obtained using a variety of equipment, such as computers or mobile phones. Specialist recording equipment can be used to obtain a wider frequency range. However, such equipment is less widely available, and recordings must be taken at specific locations. Acoustic conditions for recordings using mobile phones are poorer than those in recording studios. If possible, the signal-to-noise ratio (SNR) parameter should be determined.

There is a general view that speech does not contain frequencies above 8 kHz. Determining the threshold frequency of recorded speech determines the sampling frequency of the digital signals. The size of sampling and the resolution (usually 16 bits per sample) determines the size of the acoustic files.

It is extremely important to label recordings with accurate information on the health of the subjects. In-

formation about any medication taken by the patient is also useful, as it can have a significant impact on the features of the recorded speech.

4.3. Standardization of recording protocols

There is no consistent scenario for the content of recorded speech. Some corpora contain recordings of sustained vowels only, while others use read speech or interviews (WEINER *et al.*, 2018). Each form of speech has its own advantages: sustained vowels provide material which is standard in phoniatric investigations, read speech gives the same content for each speaker, while dialogues are usually the most natural.

4.4. Controlled clinical laryngeal examination

Only a few of the studies included phoniatric/laryngeal imaging (i.e., laryngoscopic, laryngostroboscopic, videokymography, and high-speed digital imaging). In the absence of such evaluation, the condition of the vocal folds and surrounding structures cannot be determined accurately, and potential laryngeal disorders could interfere with the disorder being investigated and have an impact on acoustic parameters of the voice. Such findings would have been bolstered by including additional information regarding the vocal fold structure and physiology. Future studies should include endoscopic methods to place these results in the diagnostic context, especially in case of somatic disorders. This recommendation is especially relevant to disorders for which the state-of-the-art is less advanced or where voice sample sizes are small. The issue is that such examinations are costly.

4.5. Exclusion criteria

There is a lack of consistency concerning excluding voice recording of certain subjects. In some cases, researchers only excluded recordings which did not follow the recording protocol; however, in most cases specific exclusion criteria were chosen, e.g., excluding subjects who were under the influence of alcohol or drugs, smokers (DE SOUZA, SANTOS, 2018) or individuals whose jobs put an increased strain on their voice. In some cases exclusions were made on the basis of medical records (e.g., the disorder was too advanced). In general, the criteria should be standardized. Additionally, the excluded recordings can be valuable for case studies investigating factors influencing the human voice.

4.6. Selection criteria for the HC group

In the majority of studies, control groups were matched by age and gender only. In some cases, additional factors were also considered, e.g., BMI. In general, there is no standardized approach.

4.7. Single disease

While some of the disorders may be intercorrelated (e.g., obesity and CAD, depression, and diabetes), the reviewed articles rarely took into account more than one or two disorders at a time.

4.8. Sample size

The number of speakers in the corpora varies from around 10 to several thousand. If the number is low, their medical records are generally more detailed. In order to apply certain modern machine learning methods, e.g., DNN, large datasets are required. This could be achieved through crowdsourcing, which has been effective in other fields of speech technology.

4.9. Legal issues of voice recording

While state-of-the-art algorithms are able to identify individuals using their voice prints, new problems are emerging. When speech is recorded on smartphones (DOGAN *et al.*, 2017), user data should be treated as sensitive and protected with appropriate privacy policy security procedures for transfer and storage.

4.10. Numbers of recordings

In most cases, just one recording is taken for each individual and there are no reference points from the onset of disease or from its progression. There is a need for more frequent monitoring of the patient voice status.

4.11. Availability of corpora

The majority of datasets were prepared for specific studies and were not available publicly to other researchers (LOW *et al.*, 2020). This makes results less comparable and reproducible while new methods are being developed. However, several corpora have become international benchmark standards on which novel methods can be validated. This trend should also be adopted for other disorders.

4.12. Feature sets

Tables 2 and 4 clearly show that there are no standards in speech parameterization or, if they exist, that they are rarely used. Only a few results were obtained using standard feature vectors, such as those offered by openSMILE. The number of extracted features varies from several to several thousand features (e.g., openSMILE vectors such as AVEC or ComPare) (Fig. 5). In most cases, researchers choose one or two parameters from four categories (prosodic, spectral, voice source, linguistic), on the basis of their previous experience or their tools. Technically, there are no limitations on verifying features from all the categories,

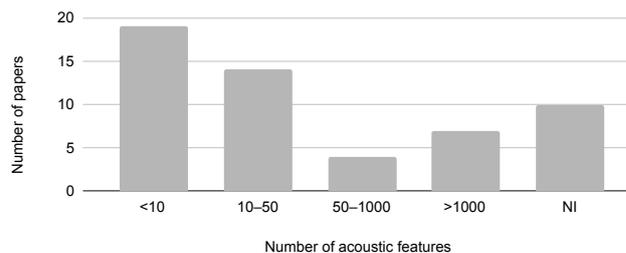


Fig. 5. Acoustic features investigated in the articles 2009–2019.

which could improve our understanding of the correlations of voice features with disorders (SADEGHIAN *et al.*, 2017). When novel parameterizations are introduced, their implementations are usually not made public.

4.13. Lack of cross-cultural and cross-language comparisons

The majority of articles focus on a single corpus or language. Gathering recordings to create databases is time consuming, and requires a high level of cooperation between engineers and medical professionals from different specializations. It is necessary to identify patients who meet the criteria set out in the study (e.g., no specific comorbidities, addictions, etc.). Being able to share databases (including speech and voice signals) among researchers usually requires formal consent and the willingness to cooperate. In the literature, there are still few studies that take into account more than one language (OROZCO-ARROYAVE *et al.*, 2014b), even though there is a need for systematic comparisons of findings between languages (MAOR *et al.*, 2018; VÁSQUEZ-CORREA *et al.*, 2017; PETTORINO *et al.*, 2017). The importance of heterogeneous training sets in machine learning should also be emphasized. Acoustic features of speech are affected by the diagnosed disorder, as well as by the language spoken by the subject. In many cases the same algorithms can be used effectively regardless of the language, and only the acoustic feature values will change. The situation is significantly more complicated in linguistic testing of speech. Transferring research methods from one language to another may not be sufficiently effective (WARNITA *et al.*, 2018).

5. Conclusions

Current disease diagnostic methods are adapted from those which have been verified as useful in speech or speaker recognition and investigating emotions. However, it should be remembered that speech analysis for medical purposes should follow different rules than speech recognition analysis. As a result of evolution, the human voice has adapted to the perceptual

capabilities of the hearing system. This has greatly improved the efficiency of communication by speech. The human ear is a frequency analyzer with nonlinear characteristics. Imitating these properties in speech technology has been shown to be highly effective, e.g., MFCC. We suppose that in the case of speech analysis for medical purposes there are no indications to take the perceptual properties of the hearing system into account. Authors of the studies included in our paper attempt to detect speech features which can be indicative of specific disorders. To achieve this, they mainly use acoustic analysis and sometimes linguistic analysis.

Determining the reasons (e.g., anatomical or neurological) for voice changes in specific disease states is a separate issue and not the subject of the cited publications. Characteristic features of speech result from the anatomical structure of the vocal tract and the way it is stimulated by the nervous system. Therefore, medical speech diagnostics are aimed at neurodegenerative and mental disorders as well as disorders affecting the physical structure of the vocal tract. The majority of publications refer to speech deviations caused by disorders of the nervous system (e.g., depression, dementia, Parkinson's and Alzheimer's diseases).

Both acoustic and linguistic properties of speech are taken into account. The methods of acoustic analysis are more useful and were tested for all the diseases presented above. Linguistic methods have been tested for mental and neurodegenerative diseases. The conducted experiments show that the combination of both methods improves the efficiency of diagnosis. The relatively large number of publications in this field is testimony to the influence of the nervous system on the generation of speech. This is particularly important for psychiatry, which usually lacks objective clinical measurements used in other specializations. While nervous system disorders result in both acoustic and linguistic features of speech, somatic disorders are diagnosed by analyzing acoustic parameters of the voice.

Speech analysis systems provide a promising approach for creating low-cost, non-invasive and remote diagnostic tools for automatic assessment or monitoring of certain disorders. Early and sensitive disease detection can support medical intervention and treatment. Speech corpora can be collected conveniently in a clinical environment or at home using smartphones.

Various methods of assessing the effectiveness of diagnosing diseases through acoustic voice analysis and linguistic speech analysis were used. This diversity makes it difficult to answer the question for which diseases the analyzed methods are by far the most effective and for which the least reliable. Effectiveness of disease diagnosis varies from 65% up to 99%. From a medical point of view, such results should be treated as a screening tests only and should be an indication of the need for standard medical tests. According

to literature reports, the highest effectiveness was obtained for: COVID-19, schizophrenia, and Parkinson's disease. Worse results were obtained for depression, bipolar disorder and Alzheimer's disease. Relatively weaker results were obtained for PCOS, diabetes, hypothyroidism, hyperthyroidism, and amyotrophic lateral sclerosis. The weakest accuracies were obtained for the diagnosis of obesity and metabolic syndrome.

Acknowledgments

This paper was supported by the National Centre for Research and Development [grant VAMP on <http://voicebiometrics.pl/> number POIR.04.01.04-00-0052/18].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. AFSHAN A., GUO J., PARK S.J., RAVI V., FLINT J., ALWAN A. (2018), Effectiveness of voice quality features in detecting depression, [in:] *Interspeech*, pp. 1676–1680, doi: [10.21437/Interspeech.2018-1399](https://doi.org/10.21437/Interspeech.2018-1399).
2. AL HANAI T., GHASSEMI M.M., GLASS J.R. (2018), Detecting depression with audio/text sequence modeling of interviews, [in:] *Interspeech*, pp. 1716–1720, doi: [10.21437/Interspeech.2018-2522](https://doi.org/10.21437/Interspeech.2018-2522).
3. ALGHOWINEM S., GOECKE R., EPPS J., WAGNER M., COHN J.F. (2016), Cross-cultural depression recognition from vocal biomarkers, [in:] *INTERSPEECH*, pp. 1943–1947, doi: [10.21437/Interspeech.2016-1339](https://doi.org/10.21437/Interspeech.2016-1339).
4. ALGHOWINEM S., GOECKE R., WAGNER M., EPPS J., BREAKSPEAR M., PARKER G. (2012), From joyous to clinically depressed: Mood detection using spontaneous speech, [in:] *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference*, Youngblood G.M., McCarthy P.M. [Eds.], pp. 141–146, <https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS12/paper/view/4478/4782>.
5. ANTOLÍK T.K., FOUGERON C. (2013), Consonant distortions in dysarthria due to Parkinson's disease, amyotrophic lateral sclerosis and cerebellar ataxia, [in:] *Interspeech*, pp. 2152–2156, doi: [10.21437/Interspeech.2013-509](https://doi.org/10.21437/Interspeech.2013-509).
6. AYDIN K. *et al.* (2016), Voice characteristics associated with polycystic ovary syndrome, *The Laryngoscope*, **126**(9): 2067–2072, doi: [10.1002/lary.25818](https://doi.org/10.1002/lary.25818).
7. BARSTIES B., VERFAILLIE R., ROY N., MARYN Y. (2013), Do body mass index and fat volume influence vocal quality, phonatory range, and aerodynamics in females?, *CoDAS*, **25**(4): 310–318, doi: [10.1590/s2317-17822013000400003](https://doi.org/10.1590/s2317-17822013000400003).

8. BEDI G. *et al.* (2015), Automated analysis of free speech predicts psychosis onset in high-risk youths, *npj Schizophrenia*, **1**(1): 15030, doi: [10.1038/npjischz.2015.30](https://doi.org/10.1038/npjischz.2015.30).
9. BOZKURT E., TOLEDO-RONEN O., SORIN A., HOORY R. (2014), Exploring modulation spectrum features for speech-based depression level classification, [in:] *Interspeech*, doi: [10.21437/Interspeech.2014-312](https://doi.org/10.21437/Interspeech.2014-312).
10. CELEBI S. *et al.* (2013). Acoustic, perceptual and aerodynamic voice evaluation in an obese population, *The Journal of Laryngology and Otology*, **127**(10): 987–990, doi: [10.1017/s0022215113001916](https://doi.org/10.1017/s0022215113001916).
11. CHITKARA D., SHARMA R.K. (2016), Voice based detection of type 2 diabetes mellitus, [in:] *2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, pp. 83–87, doi: [10.1109/AEEICB.2016.7538402](https://doi.org/10.1109/AEEICB.2016.7538402).
12. CUMMINS N., EPPS J., SETHU V., BREAKSPEAR M., GOECKE R. (2013), Modeling spectral variability for the classification of depressed speech, [in:] *Interspeech*, pp. 857–861, doi: [10.21437/Interspeech.2013-242](https://doi.org/10.21437/Interspeech.2013-242).
13. CUMMINS N., SCHERER S., KRAJEWSKI J., SCHNIEDER S., EPPS J., QUATIERI T.F. (2015a), A review of depression and suicide risk assessment using speech analysis, *Speech Communication*, **71**: 10–49, doi: [10.1016/j.specom.2015.03.004](https://doi.org/10.1016/j.specom.2015.03.004).
14. CUMMINS N., SETHU V., EPPS J., KRAJEWSKI J. (2015b), Relevance vector machine for depression prediction, [in:] *Interspeech*, pp. 110–114, doi: [10.21437/Interspeech.2015-37](https://doi.org/10.21437/Interspeech.2015-37).
15. CUMMINS N., SETHU V., EPPS J., SCHNIEDER S., KRAJEWSKI J. (2015c), Analysis of acoustic space variability in speech affected by depression, *Speech Communication*, **75**: 27–49, doi: [10.1016/j.specom.2015.09.003](https://doi.org/10.1016/j.specom.2015.09.003).
16. DA CUNHA M.G.B., PASSEROTTI G.H., WEBER R., ZILBERSTEIN B., CECCONELLO I. (2011), Voice feature characteristic in morbid obese population, *Obesity Surgery*, **21**(3): 340–344, doi: [10.1007/s11695-009-9959-7](https://doi.org/10.1007/s11695-009-9959-7).
17. DASSIE-LEITE A.P., BEHLAU M., NESI-FRANÇA S., LIMA M.N., DE LACERDA L. (2018), Vocal evaluation of children with congenital hypothyroidism, *Journal of Voice*, **32**(6): 11–19, doi: [10.1016/j.jvoice.2017.08.006](https://doi.org/10.1016/j.jvoice.2017.08.006).
18. DESHPANDE G., SCHULLER B. (2020), An overview on audio, signal, speech, & language processing for COVID-19, *arXiv preprint*, doi: [10.48550/arXiv.2005.08579](https://doi.org/10.48550/arXiv.2005.08579).
19. DESPOTOVIC V., ISMAEL M., CORNIL M., MC CALL R., FAGHERAZZI G. (2021), Detection of COVID-19 from voice, cough and breathing patterns: Dataset and preliminary results, *Computers in Biology and Medicine*, **138**: 104944, doi: [10.1016/j.compbiomed.2021.104944](https://doi.org/10.1016/j.compbiomed.2021.104944).
20. DEVULT D. *et al.* (2014), SimSensei kiosk: A virtual human interviewer for healthcare decision support, [in:] *AAMAS '14: Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 1061–1068.
21. DOGAN E., SANDER C., WAGNER X., HEGERL U., KOHLS E. (2017), Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? Systematic review, *Journal of Medical Internet Research*, **19**(7): e262, doi: [10.2196/jmir.7006](https://doi.org/10.2196/jmir.7006).
22. EKBLAD L.L. *et al.* (2015), Insulin resistance is associated with poorer verbal fluency performance in women, *Diabetologia*, **58**(11): 2545–2553, doi: [10.1007/s00125-015-3715-4](https://doi.org/10.1007/s00125-015-3715-4).
23. FAURHOLT-JEPSEN M. *et al.* (2016), Voice analysis as an objective state marker in bipolar disorder, *Translational psychiatry*, **6**(7): e856–e856, doi: [10.1038/tp.2016.123](https://doi.org/10.1038/tp.2016.123).
24. GOSZTOLYA G., BAGI A., SZALÓKI S., SZENDI I., HOFFMANN I. (2018), Identifying schizophrenia based on temporal parameters in spontaneous speech, doi: [10.13140/RG.2.2.10884.78721](https://doi.org/10.13140/RG.2.2.10884.78721).
25. GOSZTOLYA G., VINCZE V., TÓTH L., PÁKÁSKI M., KÁLMÁN J., HOFFMANN I. (2019), Identifying mild cognitive impairment and mild alzheimer's disease based on spontaneous speech using ASR and linguistic features, *Computer Speech & Language*, **53**: 181–197, doi: [10.1016/j.csl.2018.07.007](https://doi.org/10.1016/j.csl.2018.07.007).
26. GRATCH J. *et al.* (2014), The distress analysis interview corpus of human and computer interviews, [in:] *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pp. 3123–3128, http://www.lrec-conf.org/proceedings/lrec2014/pdf/508_Paper.pdf.
27. GRÓSZ T., BUSA-FEKETE R., GOSZTOLYA G., TÓTH L. (2015), Assessing the degree of nativeness and Parkinson's condition using Gaussian processes and deep rectifier neural networks, [in:] *Interspeech*, pp. 919–923, doi: [10.21437/Interspeech.2015-195](https://doi.org/10.21437/Interspeech.2015-195).
28. GRÜNERBL A. *et al.* (2014), Smartphone-based recognition of states and state changes in bipolar disorder patients, *IEEE Journal of Biomedical and Health Informatics*, **19**(1): 140–148, doi: [10.1109/jbhi.2014.2343154](https://doi.org/10.1109/jbhi.2014.2343154).
29. GUGATSCHKA M. *et al.* (2013), Subjective and objective vocal parameters in women with polycystic ovary syndrome, *Journal of Voice*, **27**(1): 98–100, doi: [10.1016/j.jvoice.2012.07.007](https://doi.org/10.1016/j.jvoice.2012.07.007).
30. GUIDI A., SCHOENTGEN J., BERTSCHY G., GENTILI C., SCILINGO E.P., VANELLO N. (2017), Features of vocal frequency contour and speech rhythm in 1250 bipolar disorder, *Biomedical Signal Processing and Control*, **37**: 23–31, doi: [10.1016/j.bspc.2017.01.017](https://doi.org/10.1016/j.bspc.2017.01.017).
31. GUIDI A., SCILINGO E. P., GENTILI C., BERTSCHY G., LANDINI L., VANELLO N. (2015), Analysis of running speech for the characterization of mood state in bipolar patients, [in:] *2015 AEIT International Annual Conference (AEIT)*, pp. 1–6, doi: [10.1109/AEIT.2015.7415275](https://doi.org/10.1109/AEIT.2015.7415275).

32. HAMDAN A.-I., JABBOUR J., NASSAR J., DAHOUK I., AZAR S.T. (2012), Vocal characteristics in patients with type 2 diabetes mellitus, *European Archives of Oto-Rhino-Laryngology*, **269**(5): 1489–1495, doi: [10.1016/j.amjoto.2012.03.008](https://doi.org/10.1016/j.amjoto.2012.03.008).
33. HAMDAN A.-L., SAFADI B., CHAMSEDDINE G., KASTY M., TURFE Z.A., ZIADE G. (2014), Effect of weight loss on voice after bariatric surgery, *Journal of Voice*, **28**(5): 618–623, doi: [10.1016/j.jvoice.2014.03.004](https://doi.org/10.1016/j.jvoice.2014.03.004).
34. HAN J. *et al.* (2020), An early study on intelligent analysis of speech under COVID-19: Severity, sleep quality, fatigue, and anxiety, *arXiv preprint*, doi: [10.48550/arXiv.2005.00096](https://doi.org/10.48550/arXiv.2005.00096).
35. HANNOUN A., ZREIK T., HUSSEINI S.T., MAHFOUD L., SIBAI A., HAMDAN A.-I. (2011), Vocal changes in patients with polycystic ovary syndrome, *Journal of Voice*, **25**(4): 501–504, doi: [10.1016/j.jvoice.2009.12.005](https://doi.org/10.1016/j.jvoice.2009.12.005).
36. HASSAN A., SHAHIN I., ALSABEK M.B. (2020), COVID-19 detection system using recurrent neural networks, [in:] *2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, pp. 1–5, doi: [10.1109/CCCI49893.2020.9256562](https://doi.org/10.1109/CCCI49893.2020.9256562).
37. HELFER B.S., QUATIERI T.F., WILLIAMSON J.R., MEHTA D.D., HORWITZ R., YU B. (2013), Classification of depression state based on articulatory precision, [in:] *Interspeech*, pp. 2172–2176, doi: [10.21437/Interspeech.2013-513](https://doi.org/10.21437/Interspeech.2013-513).
38. HEMMERLING, D., OROZCO-ARROYAVE J.R., SKALSKI A., GAJDA J., NÖTH E. (2016), Automatic detection of Parkinson’s disease based on modulated vowels, [in:] *Interspeech*, pp. 1190–1194, doi: [10.21437/Interspeech.2016-1062](https://doi.org/10.21437/Interspeech.2016-1062).
39. HÖNIG F., BATLINER A., NÖTH E., SCHNIEDER S., KRAJEWSKI J. (2014), Automatic modelling of depressed speech: relevant features and relevance of gender, [in:] *Interspeech*, pp. 1248–1252, doi: [10.21437/Interspeech.2014-313](https://doi.org/10.21437/Interspeech.2014-313).
40. HORWITZ-MARTIN R.L. *et al.* (2016), Relation of automatically extracted formant trajectories with intelligibility loss and speaking rate decline in amyotrophic lateral sclerosis, [in:] *Interspeech*, pp. 1205–1209, doi: [10.21437/Interspeech.2016-403](https://doi.org/10.21437/Interspeech.2016-403).
41. HUANG G., PENCINA K.M., COADY J.A., BEL-EVA Y.M., BHASIN S., BASARIA S. (2015), Functional voice testing detects early changes in vocal pitch in women during testosterone administration, *The Journal of Clinical Endocrinology Metabolism*, **100**(6): 2254–2260, doi: [10.1210/jc.2015-1669](https://doi.org/10.1210/jc.2015-1669).
42. HUANG K.-Y., WU C.-H., KUO Y.-T., JANG F.-L. (2016), Unipolar depression vs. bipolar disorder: An elicitation-based approach to short-term detection of mood disorder, [in:] *Interspeech*, pp. 1452–1456, doi: [10.21437/Interspeech.2016-620](https://doi.org/10.21437/Interspeech.2016-620).
43. JUNUZOVIĆ-ŽUNIĆ L., IBRAHIMAGIĆ A., ALTUMBABIĆ S. (2019), Voice characteristics in patients with thyroid disorders, *The Eurasian Journal of Medicine*, **51**(2): 101–105, doi: [10.5152/eurasianjmed.2018.18331](https://doi.org/10.5152/eurasianjmed.2018.18331).
44. KHORRAM S., GIDEON J., MCINNIS M.G., PROVOST E.M. (2016), Recognition of depression in bipolar disorder: Leveraging cohort and person specific knowledge, [in:] *Interspeech*, pp. 1215–1219, doi: [10.21437/Interspeech.2016-837](https://doi.org/10.21437/Interspeech.2016-837).
45. KISS G., SZTAHÓ D., TULICS M.G. (2021), Application for detecting depression, Parkinson’s disease and dysphonic speech, [in:] *Interspeech*, pp. 956–957.
46. KLUMPP P., JANU T., ARIAS-VERGARA T., VÁSQUEZ-CORREA J.C., OROZCO-ARROYAVE J.R., NÖTH E. (2017), Apkinson – A mobile monitoring solution for Parkinson’s disease, [in:] *Interspeech*, pp. 1839–1843, doi: [10.21437/Interspeech.2017-416](https://doi.org/10.21437/Interspeech.2017-416).
47. KONES R., RUMANA U. (2017), Cardiometabolic diseases of civilization: History and maturation of an evolving global threat. An update and call to action, *Annals of Medicine*, **49**(3): 260–274, doi: [10.1080/07853890.2016.1271957](https://doi.org/10.1080/07853890.2016.1271957).
48. KOPP W. (2019), How western diet and lifestyle drive the pandemic of obesity and civilization diseases, *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, **12**: 2221–2236, doi: [10.2147/DMSO.S216791](https://doi.org/10.2147/DMSO.S216791).
49. LAGUARTA J., HUETO F., SUBIRANA B. (2020), COVID-19 artificial intelligence diagnosis using only cough recordings, *IEEE Open Journal of Engineering in Medicine and Biology*, **1**: 275–281, doi: [10.1109/OJEMB.2020.3026928](https://doi.org/10.1109/OJEMB.2020.3026928).
50. LECHIEN J. *et al.* (2020), Features of mild-to-moderate COVID-19 patients with dysphonia, *Journal of Voice*, doi: [10.1016/j.jvoice.2020.05.012](https://doi.org/10.1016/j.jvoice.2020.05.012).
51. LOPEZ-OTERO P., DOCIO-FERNANDEZ L.D., ABAD A., GARCIA-MATEO C. (2017), Depression detection using automatic transcriptions of de-identified speech, [in:] *Interspeech*, pp. 3157–3161, doi: [10.21437/Interspeech.2017-1201](https://doi.org/10.21437/Interspeech.2017-1201).
52. LOW D.M., BENTLEY K.H., GHOSH S.S. (2020), Automated assessment of psychiatric disorders using speech: A systematic review, *Laryngoscope Investigative Otolaryngology*, **5**(1): 96–116, doi: [10.1002/lio2.354](https://doi.org/10.1002/lio2.354).
53. MALLELA J. *et al.* (2020), Raw speech waveform based classification of patients with ALS, Parkinson’s disease and healthy controls using CNN-BLSTM, [in:] *Interspeech*, pp. 4586–4590, doi: [10.21437/Interspeech.2020-2221](https://doi.org/10.21437/Interspeech.2020-2221).
54. MAOR E., SARA J.D., ORBELO D.M., LERMAN L.O., LEVANON Y., LERMAN A. (2018), Voice signal characteristics are independently associated with coronary artery disease, *Mayo Clinic Proceedings*, pp. 840–847, doi: [10.1016/j.mayocp.2017.12.025](https://doi.org/10.1016/j.mayocp.2017.12.025).
55. MCGINNIS E.W. *et al.* (2019), Giving voice to vulnerable children: Machine learning analysis of speech detects anxiety and depression in early childhood, *IEEE Journal of Biomedical and Health Informatics*, **23**(6): 2294–2301, doi: [10.1109/JBHI.2019.2913590](https://doi.org/10.1109/JBHI.2019.2913590).

56. MIRHEIDARI B., BLACKBURN D., WALKER T., VENNARI A., REUBER M., CHRISTENSEN H. (2018), Detecting signs of dementia using word vector representations, [in:] *Interspeech*, pp. 1893–1897, doi: [10.21437/Interspeech.2018-1764](https://doi.org/10.21437/Interspeech.2018-1764).
57. MOHAMMADZADEH A., HEYDARI E., AZIZI F. (2011), Speech impairment in primary hypothyroidism, *Journal of Endocrinological Investigation*, **34**(6): 431–433, doi: [10.1007/BF03346708](https://doi.org/10.1007/BF03346708).
58. MORO-VELAZQUEZ L., GOMEZ-GARCIA J.A., ARIAS-LONDOÑO J.D., DEHAK N., GODINO-LLORENTE J.I. (2021), Advances in Parkinson’s disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects, *Biomedical Signal Processing and Control*, **66**: 102418, doi: [10.1016/j.bspc.2021.102418](https://doi.org/10.1016/j.bspc.2021.102418).
59. MOTA N.B. *et al.* (2012), Speech graphs provide a quantitative measure of thought disorder in psychosis, *PLOS ONE*, **7**(4): e34928. doi: [10.1371/journal.pone.0034928](https://doi.org/10.1371/journal.pone.0034928).
60. MUNDT J.C., VOGEL A.P., FELTNER D.E., LENDERKING W.R. (2012), Vocal acoustic biomarkers of depression severity and treatment response, *Biological psychiatry*, **72**(7): 580–587, doi: [10.1016/j.biopsych.2012.03.015](https://doi.org/10.1016/j.biopsych.2012.03.015).
61. OROZCO-ARROYAVE J.R., ARIAS-LONDOÑO J.F., VARGAS-BONILLA J.F., GONZALEZ-RATIVA M.C., NÖTH E. (2014a), New spanish speech corpus database for the analysis of people suffering from Parkinson’s disease, [in:] *LREC*, pp. 342–347.
62. OROZCO-ARROYAVE J.R. *et al.* (2014b), Automatic detection of Parkinson’s disease from words uttered in three different languages, [in:] *Interspeech*, doi: [10.21437/Interspeech.2014-375](https://doi.org/10.21437/Interspeech.2014-375).
63. PAN Y., MIRHEIDARI B., REUBER M., VENNARI A., BLACKBURN D., CHRISTENSEN H. (2020), Improving detection of Alzheimer’s disease using automatic speech recognition to identify high-quality segments for more robust feature extraction, [in:] *Interspeech*, pp. 4961–4965, doi: [10.21437/Interspeech.2020-2698](https://doi.org/10.21437/Interspeech.2020-2698).
64. PAREEK V., SHARMA R.K. (2016), Coronary heart disease detection from voice analysis, [in:] *2016 IEEE Students’ Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1–6, doi: [10.1109/SCEECS.2016.7509344](https://doi.org/10.1109/SCEECS.2016.7509344).
65. PETTORINO M., GU W., PÓLROLA P., FAN P. (2017), Rhythmic characteristics of Parkinsonian speech: A study on Mandarin and Polish, [in:] *Interspeech*, pp. 3172–3176, doi: [10.21437/Interspeech.2017-850](https://doi.org/10.21437/Interspeech.2017-850).
66. PINHEIRO A.P., NIZNIKIEWICZ M. (2019), Altered attentional processing 1395 of happy prosody in schizophrenia, *Schizophrenia Research*, **206**: 217–224, doi: [10.1016/j.schres.2018.11.024](https://doi.org/10.1016/j.schres.2018.11.024).
67. PINKAS G., KARNY Y., MALACHI A., BARKAI G., BACHAR G., AHARONSON V. (2020), SARS-CoV-2 detection from voice, *IEEE Open Journal of Engineering in Medicine and Biology*, **1**: 268–274, doi: [10.1109/ojemb.2020.3026468](https://doi.org/10.1109/ojemb.2020.3026468).
68. PINTO S. *et al.* (2016), Dysarthria in individuals with Parkinson’s disease: A protocol for a binational, cross-sectional, case-controlled study in French and European Portuguese (FraLusoPark), *BMJ Open*, **6**(11): doi: [10.1136/bmjopen-2016-012885](https://doi.org/10.1136/bmjopen-2016-012885).
69. PINYOPODJANARD S., SUPPAKITJANUSANT P., LOMPREW P., KASEMKOSIN N., CHAILURKIT L., ONGPHIPHADHANAKUL B. (2019), Instrumental acoustic voice characteristics in adults with type 2 diabetes, *Journal of Voice*, **35**(1): 116–121, doi: [10.1016/j.jvoice.2019.07.003](https://doi.org/10.1016/j.jvoice.2019.07.003).
70. POMPILI A. *et al.* (2020), Assessment of Parkinson’s disease medication state through automatic speech analysis, *arXiv preprint*, doi: [10.48550/arXiv.2005.14647](https://doi.org/10.48550/arXiv.2005.14647).
71. ROHANIAN M., HOUGH J., PURVER M. (2021), Alzheimer’s dementia recognition using acoustic, lexical, disfluency and speech pause features robust to noisy inputs, [in:] *Interspeech*, pp. 3820–3824, doi: [10.21437/Interspeech.2021-1633](https://doi.org/10.21437/Interspeech.2021-1633).
72. RUSZ J. *et al.* (2018), Smartphone allows capture of speech abnormalities associated with high risk of developing Parkinson’s disease, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **26**(8): 1495–1507, doi: [10.1109/TNSRE.2018.2851787](https://doi.org/10.1109/TNSRE.2018.2851787).
73. SADEGHIAN R., SCHAFFER J.D., ZAHORIAN S.A. (2017), Speech processing approach for diagnosing dementia in an early stage, [in:] *Interspeech*, pp. 2705–2709, doi: [10.21437/Interspeech.2017-1712](https://doi.org/10.21437/Interspeech.2017-1712).
74. SAHU S., ESPY-WILSON C.Y. (2016), Speech features for depression detection, [in:] *Interspeech*, pp. 1928–1932, doi: [10.21437/Interspeech.2016-1566](https://doi.org/10.21437/Interspeech.2016-1566).
75. SATTLER C. *et al.* (2017), Interdisciplinary longitudinal study on adult development and aging (ILSE), [in:] *Encyclopedia of Geropsychology*, Pachana N.A. [Ed.], pp. 1213–1222, Springer, doi: [10.1007/978-981-287-082-7_238](https://doi.org/10.1007/978-981-287-082-7_238).
76. SCHERER S., STRATOU G., GRATCH J., MORENCY L.-P. (2013a), Investigating 1435 voice quality as a speaker-independent indicator of depression and PTSD, [in:] *Interspeech*, pp. 847–851, doi: [10.21437/Interspeech.2013-240](https://doi.org/10.21437/Interspeech.2013-240).
77. SCHERER S. *et al.* (2013b), Automatic behavior descriptors for psychological disorder analysis, [in:] *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pp. 1–8, doi: [10.1109/FG.2013.6553789](https://doi.org/10.1109/FG.2013.6553789).
78. SENEVIRATNE N., WILLIAMSON J.R., LAMMERT A.C., QUATIERI T.F., ESPY-WILSON C. (2020), Extended study on the use of vocal tract variables to quantify neuromotor coordination in depression, [in:] *Interspeech*, pp. 4551–4555, doi: [10.21437/Interspeech.2020-2758](https://doi.org/10.21437/Interspeech.2020-2758).
79. SHARMA N. *et al.* (2020), Coswara – A database of breathing, cough, and voice sounds for COVID-19 diagnosis, *arXiv preprint*, pp. 4811–4815, doi: [10.21437/Interspeech.2020-2768](https://doi.org/10.21437/Interspeech.2020-2768).

80. SIMANTIRAKI O., CHARONYKTAKIS P., PAMPOUCHIDOU A., TSIKNAKIS M., COOKE M. (2017), Glottal source features for automatic speech-based depression assessment, [in:] *Interspeech*, pp. 2700–2704, doi: [10.21437/Interspeech.2017-1251](https://doi.org/10.21437/Interspeech.2017-1251).
81. SIRMANS S.M., PATE K.A. (2014), Epidemiology, diagnosis, and management of polycystic ovary syndrome, *Clinical Epidemiology*, **6**: 1–13, doi: [10.2147/cep.s37559](https://doi.org/10.2147/cep.s37559).
82. SKODDA S., GRÖNHEIT W., SCHLEGEL U. (2011), Intonation and speech rate in Parkinson's disease: General and dynamic aspects and responsiveness to levodopa admission, *Journal of Voice*, **25**(4): e199–e205, doi: [10.1016/j.jvoice.2010.04.007](https://doi.org/10.1016/j.jvoice.2010.04.007).
83. SOLOMON N.P., HELOU L.B., DIETRICH-BURNS K., STOJADINOVIC A. (2011), Do obesity and weight loss affect vocal function?, [in:] *Seminars in Speech and Language*, **31**(1): 31–42, doi: [10.1055/s-0031-1271973](https://doi.org/10.1055/s-0031-1271973).
84. DE SOUZA L.B.R., PEREIRA R.M., DOS SANTOS M.M., GODOY C.M.A. (2014), Fundamental frequency, phonation maximum time and vocal complaints in morbidly obese women, *ABCD. Arquivos Brasileiros de Cirurgia Digestiva*, **27**(1): 43–46. doi: [10.1590/s0102-67202014000100011](https://doi.org/10.1590/s0102-67202014000100011).
85. DE SOUZA L.B.R., DOS SANTOS M.M. (2018), Body mass index and acoustic voice parameters: Is there a relationship?, *Brazilian Journal of Otorhinolaryngology*, **84**(4): 410–415, doi: [10.1016/j.bjorl.2017.04.003](https://doi.org/10.1016/j.bjorl.2017.04.003).
86. STASAK B., EPPS J., CUMMINS N., GOECKE R. (2016), An investigation of emotional speech in depression classification, [in:] *Interspeech*, pp. 485–489, doi: [10.21437/Interspeech.2016-867](https://doi.org/10.21437/Interspeech.2016-867).
87. STASAK B., EPPS J., GOECKE R. (2017), Elicitation design for acoustic depression classification: An investigation of articulation effort, linguistic complexity, and word affect, [in:] *Interspeech*, pp. 834–838, doi: [10.21437/Interspeech.2017-1223](https://doi.org/10.21437/Interspeech.2017-1223).
88. STASAK B., HUANG Z., RAZAVI S., JOACHIM D., EPPS J. (2021), Automatic detection of COVID-19 based on short-duration acoustic smartphone speech analysis, *Journal of Healthcare Informatics Research*, **5**(2): 201–217, doi: [10.1007/s41666-020-00090-4](https://doi.org/10.1007/s41666-020-00090-4).
89. STOGOWSKA E., KAMŃSKI K.A., ZIÓŁKO B., KOWALSKA I. (2022), Voice changes in reproductive disorders, thyroid disorders and diabetes: A review, *Endocrine Connections*, **11**(3): e201505, doi: [10.1530/EC-21-0505](https://doi.org/10.1530/EC-21-0505).
90. SUBIRANA B. *et al.* (2020), Hi sigma, do I have the Coronavirus?: Call for a new artificial intelligence approach to support health care professionals dealing with the COVID-19 pandemic, *arXiv preprint*, doi: [10.48550/arXiv.2004.06510](https://doi.org/10.48550/arXiv.2004.06510).
91. SZTAHÓ D., KISS G., VICSI K. (2015), Estimating the severity of Parkinson's disease from speech using linear regression and database partitioning, [in:] *Interspeech*, pp. 498–502, doi: [10.21437/Interspeech.2015-183](https://doi.org/10.21437/Interspeech.2015-183).
92. UJIRO T. *et al.* (2018), Detection of dementia from responses to atypical questions asked by embodied conversational agents, [in:] *Interspeech*, pp. 1691–1695, doi: [10.21437/Interspeech.2018-1514](https://doi.org/10.21437/Interspeech.2018-1514).
93. VALSTAR M. *et al.* (2013), Avec 2013: The continuous audio/visual emotion and depression recognition challenge, [in:] *AVEC'13 Proceedings of the 3rd ACM International Workshop on Audio/Visual Emotion Challenge*, pp. 3–10, doi: [10.1145/2512530.2512533](https://doi.org/10.1145/2512530.2512533).
94. VÁSQUEZ-CORREA J.C., ARIAS-VERGARA T., OROZCO-ARROYAVE J.R., NÖTH E. (2018), A multitask learning approach to assess the dysarthria severity in patients with Parkinson's disease, [in:] *Interspeech*, pp. 456–460, doi: [10.21437/interspeech.2018-1988](https://doi.org/10.21437/interspeech.2018-1988).
95. VÁSQUEZ-CORREA J.C., ARIAS-VERGARA T., OROZCO-ARROYAVE J.R., VARGAS-BONILLA J.F., ARIAS-LONDOÑO J.D., NÖTH E. (2015), Automatic detection of Parkinson's disease from continuous speech recorded in non-controlled noise conditions, [in:] *Interspeech*, pp. 105–109, doi: [10.21437/Interspeech.2015-36](https://doi.org/10.21437/Interspeech.2015-36).
96. VÁSQUEZ-CORREA J.C., OROZCO-ARROYAVE J.R., NÖTH E. (2017), Convolutional neural network to model articulation impairments in patients with Parkinson's disease, [in:] *Interspeech*, pp. 314–318, doi: [10.21437/Interspeech.2017-1078](https://doi.org/10.21437/Interspeech.2017-1078).
97. VILLA-CAÑAS T., ARIAS-LONDOÑO J.D., OROZCO-ARROYAVE J.R., VARGAS-BONILLA J.F., NÖTH E. (2015), Low-frequency components analysis in running speech for the automatic detection of Parkinson's disease, [in:] *Interspeech*, pp. 100–104, doi: [10.21437/Interspeech.2015-35](https://doi.org/10.21437/Interspeech.2015-35).
98. VILLATORO-TELLO E., DUBAGUNTA P., FRITSCH J., RAMÍREZ-DE-LA ROSA G., MOTLICEK P., MAGIMAI-DOSS M. (2021), Late fusion of the available lexicon and raw waveform-based acoustic modeling for depression and dementia recognition, [in:] *Interspeech*, pp. 1927–1931, doi: [10.21437/Interspeech.2021-1288](https://doi.org/10.21437/Interspeech.2021-1288).
99. WANG J., KOTHALKAR P.V., CAO B., HEITZMAN D. (2016), Towards automatic detection of amyotrophic lateral sclerosis from speech acoustic and articulatory samples, [in:] *Interspeech*, pp. 1195–1199, doi: [10.21437/Interspeech.2016-1542](https://doi.org/10.21437/Interspeech.2016-1542).
100. WANKERL S., NÖTH E., EVERT S. (2017), An n-gram based approach to the automatic diagnosis of Alzheimer's disease from spoken language, [in:] *Interspeech*, pp. 3162–3166, doi: [10.21437/Interspeech.2017-1572](https://doi.org/10.21437/Interspeech.2017-1572).
101. WARNITA T., INOUE N., SHINODA K. (2018), Detecting Alzheimer's disease using gated convolutional neural network from audio data, [in:] *Interspeech*, pp. 1706–1710, doi: [10.21437/Interspeech.2018-1713](https://doi.org/10.21437/Interspeech.2018-1713).
102. WEI W., WANG J., MA J., CHENG N., XIAO J. (2020), A real-time robot-based auxiliary system for risk evaluation of COVID-19 infection, *arXiv preprint*, doi: [10.48550/arXiv.2008.07695](https://doi.org/10.48550/arXiv.2008.07695).
103. WEINER J., ANGRICK M., UMESH S., SCHULTZ T. (2018), Investigating the effect of audio duration on dementia detection using acoustic features, [in:] *Interspeech*, pp. 2324–2328, doi: [10.21437/Interspeech.2018-57](https://doi.org/10.21437/Interspeech.2018-57).

104. WEINER J., HERFF C., SCHULTZ T. (2016), Speech-based detection of Alzheimer's disease in conversational German, [in:] *Interspeech*, pp. 1938–1942, doi: [10.21437/Interspeech.2016-100](https://doi.org/10.21437/Interspeech.2016-100).
105. WODZINSKI M., SKALSKI A., HEMMERLING D., OROZCO-ARROYAVE J.R., NÖTH E. (2019), Deep learning approach to Parkinson's disease detection using voice recordings and convolutional neural network dedicated to image classification, [in:] *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 717–720, doi: [10.1109/embc.2019.8856972](https://doi.org/10.1109/embc.2019.8856972).
106. XEZONAKI D., PARASKEVOPOULOS G., POTAMIANOS A., NARAYANAN S. (2020), Affective conditioning on hierarchical attention networks applied to depression detection from transcribed clinical interviews, [in:] *Interspeech*, pp. 4556–4560, doi: [10.21437/Interspeech.2020-2819](https://doi.org/10.21437/Interspeech.2020-2819).
107. YANG Y., FAIRBAIRN C., COHN J.F. (2012), Detecting depression severity from vocal prosody, *IEEE Transactions on Affective Computing*, **4**(2): 142–150, doi: [10.1109/T-AFFC.2012.38](https://doi.org/10.1109/T-AFFC.2012.38).
108. ZHAN A. *et al.* (2016), High frequency remote monitoring of Parkinson's disease via smartphone: Platform overview and medication response detection, *arXiv preprint*, doi: [10.48550/arXiv.1601.00960](https://doi.org/10.48550/arXiv.1601.00960).
109. ZHAO Z. *et al.* (2020), Hybrid network feature extraction for depression assessment from speech, [in:] *Interspeech*, pp. 4956–4960, doi: [10.21437/Interspeech.2020-2396](https://doi.org/10.21437/Interspeech.2020-2396).
110. ZLOTNIK A., MONTERO J.M., SAN-SEGUNDO R., GALLARDO-ANTOLÍN A. (2015), Random forest-based prediction of Parkinson's disease progression using acoustic, ASR and intelligibility features, [in:] *Interspeech*, pp. 503–507, doi: [10.21437/Interspeech.2015-184](https://doi.org/10.21437/Interspeech.2015-184).

Research Paper

A Symmetric Approach in the Three-Dimensional Digital Waveguide Modeling of the Vocal Tract

Tahir MUSHTAQ^{(1)*}, Ahmad KAMRAN⁽¹⁾,
Muhammad Zubair Akbar QURESHI⁽²⁾, Zafar IQBAL⁽³⁾

⁽¹⁾ *Department of Mathematics, COMSATS University Islamabad
Vehari Campus, Vehari, Pakistan*

⁽²⁾ *Department of Mathematics, Air University
Islamabad, Pakistan*

⁽³⁾ *Department of Mathematics, Government Graduate College of Science
Multan, Pakistan*

*Corresponding Author e-mail: tahirmushtaq@ciitvehari.edu.pk

(received November 4, 2022; accepted April 4, 2023)

Simulation of wave propagation in the three-dimensional (3D) modeling of the vocal tract has shown significant promise for enhancing the accuracy of speech production. Recent 3D waveguide models of the vocal tract have been designed for better accuracy but require a lot of computational tasks. A high computational cost in these models leads to novel work in reducing the computational cost while retaining accuracy and performance. In the current work, we divide the geometry of the vocal tract into four equal symmetric parts with the introduction of two axial perpendicular planes, and the simulation is performed on only one part. A novel strategy is defined to implement symmetric conditions in the mesh. The complete standard 3D digital waveguide model is assumed as a benchmark model. The proposed model is compared with the benchmark model in terms of formant frequencies and efficiency. For the demonstration, the vowels /ɔ/, /i/, /ε/, /ɑ/, and /u/ have been selected for the simulations. According to the results, the benchmark and current models are nearly identical in terms of frequency profiles and formant frequencies. Still the current model is three times more effective than the benchmark model.

Keywords: symmetric; digital waveguide; vocal tract; delay lines; rectilinear uniform grid.



Copyright © 2023 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

Three-dimensional (3D) acoustic simulations of the vocal tract show tremendous promise for the research of speech acoustics (GULLY, TUCKER, 2019). Many studies have focused on simulating the vocal tract because of its significance in the speech production system (LIM *et al.*, 2019; QURESHI *et al.*, 2020; QURESHI, ISHAQ, 2019; QURESHI, SYED, 2015; 2019; SCHICKHOFER, MIHAESCU, 2020; VAMPOLA *et al.*, 2015). Many researchers have used computational two- and three-dimensional techniques for voice generation based on wave propagation in the vocal tract using the finite element method (FEM) (ARNELA *et al.*, 2016a;

2019; BLANDIN *et al.*, 2021; MOHAPATRA *et al.*, 2019). The most common method for analyzing vocal tract acoustics is FEM. However, the existing technology limits its usage in the real-time application of wave propagation due to the long computation time.

The use of waveguide modeling in computational vocal tract modeling is quite common. Computational models are used in digital waveguide modeling for wave propagation in the vocal tract. It is assumed that each node in the vocal tract grid acts as a scattering junction connected by a unit of waveguide element, which is known as the delay line in the simulation (SPEED *et al.*, 2013a; TREYSSCDE, 2021). High-quality sound generation in real-time is made possible by the

use of digital waveguide models, which are efficient and realistic computational models of the waveguide model (MULLEN *et al.*, 2007; SPEED *et al.*, 2013a). A few attempts have been made to extend the Kelly-Lochbaum model by incorporating fractional delays for the vocal tract's elongation (MATHUR *et al.*, 2006; QURESHI, SYED, 2019). The Kelly-Lochbaum model was based on the same length of each cylindrical segment.

In the literature, the transformation of the vocal tract into concatenated segments with various cross-sectional areas has proven successful for one-dimensional waveguide modeling (MULLEN *et al.*, 2003; 2006), whereas the vocal tract has been converted into better approximated one using conical segmentations (MAKAROV, 2009; STRUBE, 2003). The multi-dimensional digital waveguide model offers more precision than the one-dimensional waveguide model but has a higher computing cost (ARNELA *et al.*, 2016b; BEESON, MURPHY, 2004; BLANDIN *et al.*, 2015; MOHAPATRA *et al.*, 2022; MURPHY, BEESON, 2007; MURPHY, HOWARD, 2000). The restriction of uniform structured mesh is a major limitation in the digital waveguide models (SPEED *et al.*, 2013a; VAN DUYN, SMITH, 1995). The complexity and computational cost of the three-dimensional digital waveguide model for vocal tract modeling limits the interest of the researchers. When accuracy is a concern, the three-dimensional model becomes a suitable candidate for modern computers (SPEED *et al.*, 2013a; 2013b). In recent works, three-dimensional digital waveguide modeling for the vocal tract has been adopted for the uniform linear grid as well as the non-uniform rectilinear grid proposed for voiced stop consonants (GULLY *et al.*, 2017; GULLY, TUCKER, 2019; QURESHI, ISHAQ, 2019; SPEED *et al.*, 2013b).

For the modeling of the vocal tract, the current work is focused on implementing a symmetric 3D digital waveguide model. For this purpose, a 3D geometry of the vocal tract is divided into four equal symmetric parts with the introduction of two axial perpendicular planes. The modeling of the single symmetric part is referred to as a 3D symmetric digital waveguide model of the vocal tract. A novel strategy has been devised to implement symmetric conditions on the two perpendicular axial sides of the meshing.

The proposed approach will reduce the high computation cost in the simulation while keeping the model's accuracy. For the validation of our approach, we choose a complete standard three-dimensional waveguide as a benchmark model in the proposed work. For the demonstrations of our model, we consider cross-sectional areas of the different vowels /ɔ/, /i/, /ε/, /α/, and /u/ from (STORY *et al.*, 1996). In the proposed work, the developed model is compared with the benchmarked model in terms of formant frequencies and efficiency. As a result of the simulations, the symmetric model is found to match the benchmark model

in terms of accuracy and is three times more efficient than the benchmark model.

2. Three-dimensional symmetric waveguide modeling of the vocal tract

For the realistic modeling of the vocal tract, the accuracy of the wave propagation in the vocal tract using 3D waveguide modeling makes it popular in the field of speech production. However, a significant restriction in the application of the three-dimensional waveguide model is the computational cost. Not many works have been reported in the literature to reduce the computational cost using the three-dimensional waveguide model of the vocal tract. Furthermore, the current work is focused on reducing computational cost in the vocal tract modeling. The computational cost is greatly reduced with the help of implementing symmetric three-dimensional waveguide modeling of the vocal tract. Using two axial perpendicular planes, a 3D geometry of the vocal tract is divided into four equal symmetric parts and the modeling of the single symmetric part is referred to as a symmetric three-dimensional digital waveguide model of the vocal tract. The symmetric conditions are implemented on the symmetric sides of the meshing. The visualization of the current approach in the proposed work is demonstrated in Figs. 1 and 2.

Figure 1 shows the three-dimensional geometry of the vocal tract orientation in the case of the vowel /ɔ/, where the length of the vocal tract is along the x -axis. Figure 2 presents the symmetric part with the green color for the vocal tract, as shown in Fig. 1. The re-

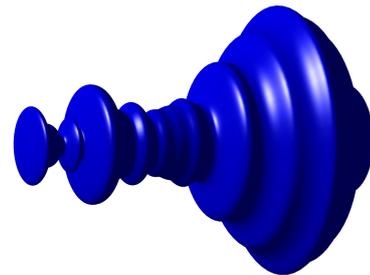


Fig. 1. Three-dimensional geometrical view of the vocal tract.

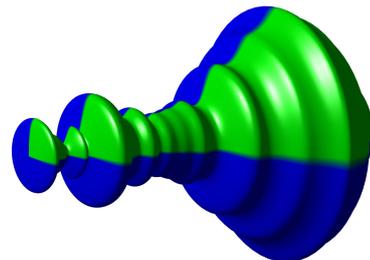


Fig. 2. Three-dimensional symmetric part of the vocal tract.

maining three symmetric parts are shown with the same blue color in the figure, which is not considered in the simulation of wave propagation in the vocal tract.

A vocal tract is a cylindrical tube of varying cross-sectional area in a one-dimensional waveguide model. The total vocal tract length is divided into multiple cylindrical tube segments with different cross-sectional areas. The wave equation is used to formulate the relationship between velocity and pressure in this tube (MARKEL, GRAY, 1976; RABINER, SCHAFER, 1978). D'Alembert's process helps to solve one-dimensional wave equations as a sum of left and right wave components. In the continuity and momentum equation at the intersection of two cylinders k -th and $(k + 1)$ -th (KARJALAINEN, ERKUT, 2004), R_k is the reflection coefficient, A_k is the cross-sectional area of the k -th cylinder, and A_{k+1} is the cross-sectional area of the $(k + 1)$ -th cylinder:

$$R_k = \frac{A_k - A_{k+1}}{A_k + A_{k+1}}. \quad (1)$$

The accuracy is increased by getting the three-dimensional structure in the waveguide model, which improves spatio-temporal connectivity in the sample grid. In this case, the mesh of the waveguide model increases with the expansion of the spatio-temporal connection in the sample grid of the three-dimensional (BEESON, MURPHY, 2004; MULLEN *et al.*, 2006; VAN DUYNÉ, SMITH, 1993). In the standard three-dimensional waveguide model, every node is called a scattering junction, where each wave is modified and moves in many directions. Every scattering junction is situated at the same delay lines in the grid and each scattering junction has six more neighboring scattering junctions at a point of 90 degrees, as illustrated in Fig. 3. except at the walls of the vocal tract. The interconnection of multiple junctions is shown in Fig. 4 with a three-dimensional view.

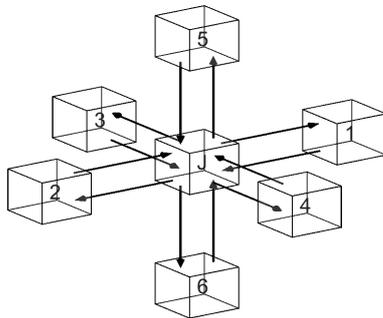


Fig. 3. Single junction with six-neighboring junctions.

Multiple grid topologies are utilized in the literature for the three-dimensional digital waveguide modeling of the vocal tract. However, we selected a three-dimensional uniform rectilinear grid topology in the work because of its simplicity.

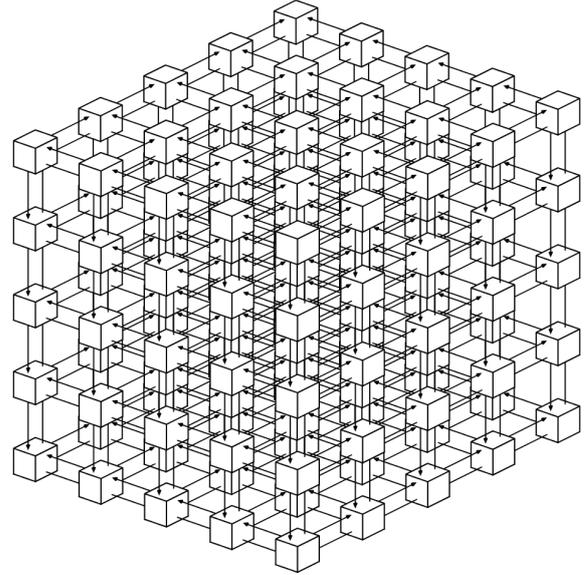


Fig. 4. Interconnection of multiple junctions.

At each junction J , the incoming wave pressures are symbolized as $p_{J,1}^+$, $p_{J,2}^+$, $p_{J,3}^+$, $p_{J,4}^+$, $p_{J,5}^+$, and $p_{J,6}^+$, respectively, and the outgoing pressures are denoted by $p_{J,1}^-$, $p_{J,2}^-$, $p_{J,3}^-$, $p_{J,4}^-$, $p_{J,5}^-$, and $p_{J,6}^-$, respectively. Finally, the pressure $p_{J,i}$ on each scattering junction J can be formulated as (MULLEN, 2006; VAN DUYNÉ, SMITH, 1993):

$$p_{J,i} = p_{J,i}^+ + p_{J,i}^-. \quad (2)$$

For N -port neighboring junctions of the junction J , the total pressure p at the junction J can be written as (MULLEN, 2006; VAN DUYNÉ, SMITH, 1993):

$$p_J = 2 \frac{\sum_{i=1}^N Y_i p_{J,i}^+}{\sum_{i=1}^N Y_i}. \quad (3)$$

For a three-dimensional case with uniform impedance, the aforementioned equation can be rewritten as:

$$p_J = \frac{\sum_{i=1}^6 p_{J,i}^+}{3}. \quad (4)$$

The outgoing pressure components can be considered as:

$$p_{J,i}^- = p_J - p_{J,i}^+. \quad (5)$$

The current work is focused on the symmetric three-dimensional waveguide mesh. To increase the efficiency of the three-dimensional waveguide model, we impose the symmetric conditions on the symmetric sides of the waveguide mesh. In the configuration of the rectilinear mesh in Fig. 3, we consider that each junction has six neighboring junctions at an angle of 90 degrees from one another. Junctions 1 and 2 are

along the x -axis, junctions 3 and 4 are along the y -axis, while junctions 5 and 6 are along the z -axis. Consider that the length of the three-dimensional vocal tract is along the x -direction while the cross-sectional area is along the yz -plane. In our work, we consider the part of the vocal tract that rests in the first quadrant in the coordinate axes. To make a symmetric model, we impose symmetric boundary conditions on the xy - and xz -planes.

For the symmetry on the xy -plane, the neighboring nodes 5 and 6 are between the line of symmetry along the y -axis. The outgoing wave from junction J to port 6 is unimportant and is omitted, while the incoming wave from port 6 to junction J is important, and we consider it. To employ symmetric, it must be the same as an outgoing wave from J to port 5. In Eq. (4), the pressure at the junction J for the symmetric line can be derived as:

$$p_J = \frac{\sum_{\substack{i=1 \\ i \neq 6}}^6 p_{J,i}^+ + p_{J,5}^+}{3}. \quad (6)$$

Output traveling-wave components at the node are calculated as:

$$p_{J,i}^- = p_J - p_{J,i}^+, \quad i \neq 6. \quad (7)$$

For the symmetry on the xz -plane, the neighboring nodes 3 and 4 are between the line of symmetry along the z -axis. The outgoing wave from junction J to port 4 is unimportant and is omitted, while the incoming wave from port 4 to junction J is important, and we consider it. To employ symmetric, it must be the same as an outgoing wave from J to port 3. In Eq. (4), the pressure at the junction J for the symmetric line can be derived as:

$$p_J = \frac{\sum_{\substack{i=1 \\ i \neq 4}}^6 p_{J,i}^+ + p_{J,3}^+}{3}. \quad (8)$$

Output traveling-wave components at the node are calculated as:

$$p_{J,i}^- = p_J - p_{J,i}^+, \quad i \neq 4. \quad (9)$$

3. Numerical simulation

A series of cross-sectional areas are acquired from the particular geometric configurations of the vocal tract. In the proposed work, we consider the different configurations of the vocal tract for the vowels /ɔ/, /i/, /ε/, /a/, and /u/. The cross-sectional areas of these vowels are taken from (STORY *et al.*, 1996) with varying vocal tract lengths of 17.46, 15.88, 16.67, 17.46, and 18.25 cm, respectively. For the production of the three-dimensional mesh, the cross-sectional areas are interpolated into the smooth shape of the three-dimensional vocal tract by using a piecewise third-order spline.

For the best resolution of the graphs, the length of the impulse response is chosen as 25 000 input samples for every vowel considered in this work. All input samples have zero values except the first input sample, which has a value of 1. The numerical solution of the current vocal model comprises two basic steps for each iteration. The first step is used to find the scattering of the waves at each junction of the mesh, while the wave delay of each junction is passed to the neighboring junctions in the second step. In the first step, we apply the glottal boundary condition (SPEED *et al.*, 2013a) on each junction that lies on the inlet of the vocal tract. For the scattering of the wave, Eq. (4) is used to calculate the total pressure on each interior junction of the mesh and we also measure the outgoing pressure in all six directions with the help of Eq. (5). The scattering of the waves at outlet junctions (SPEED *et al.*, 2013a) is also computed. For the junctions that lie on the symmetric xy -plane, Eqs. (6) and (7) are employed to calculate total and outgoing pressures, respectively. Similarly, for the case of a symmetric xz -plane, Eqs. (8) and (9) are used. In the second step, all components of the wave delay of each junction are passed to its neighboring junctions. We iterate this process according to the total number of input samples to get output samples. The frequency response is then obtained from the transfer function by applying fast Fourier transformation (FFT) and natural logarithms on the output samples. The simulations of the current work are compiled in the high-level computer programming language MATLAB 2020. MATLAB computer language also offers the use of C++-coded files to increase the efficiency of the simulations. We also developed a C++-coded file used in MATLAB to increase the efficiency for the high-cost computational work in the 3D symmetric meshing. The reflection coefficient of the glottis r_G , the reflection coefficient of the wall r_W and the reflection coefficient lip r_L are assigned with values of 0.97, 1.0, and -0.9, respectively.

4. Results and discussion

In this section, we demonstrate our current approach in the execution of the symmetric three-dimensional digital waveguide model of the vocal tract. Our benchmark model is used to compare and validate our symmetric model. Based on accuracy and efficiency, the comparison between the symmetric and the benchmark models is established. The first three lowest formant frequencies are known to be sufficient for every vowel to be recognized. In the current work, we take six formant frequencies of the benchmark and symmetric models and compare them with each other so that we can get a better idea about the relative error, and check the efficiency of our symmetric model. The six formant frequencies are denoted by f_1 , f_2 , f_3 ,

f_4, f_5, f_6 , which are mentioned in the first column of all tables.

For the comparison process, the sample delay d_s is taken the same for both models. In the current work, vowels /ɔ/, /i/, /ε/, /α/, and /u/ are taken with different vocal tract lengths. The efficiency and accuracy of the current model are compared with the benchmark model in the collapsed time, frequency profiles and formant frequencies, respectively. At least three formant frequencies are required for each vowel. In our work, we consider the first six formant frequencies produced by the benchmark model and our proposed model (ARNELA *et al.*, 2016b).

Figure 5 shows the frequency profiles generated by a transfer function of a vowel /ɔ/ of the vocal tract up to 6000 Hz. We observe that the frequency profiles of benchmark and symmetric models overlap. There is no difference between the profiles of the benchmark and the symmetric models. In other words, the frequency profiles of both models are identical to each other. Numerically calculated values are given in Table 1. In this table, column 4 shows the relative error between the benchmark and symmetric models. This column presents the zero relative errors in all six formant frequencies in the current case of vowel /ɔ/. Column 7 shows that the proposed model is about 3 times more efficient than the benchmark model due to the high efficiency of the symmetric model.

The frequency profiles of the vowel /i/ of the vocal tract obtained by its transfer function are shown in Fig. 6. It is also noted that the frequency profile of the symmetric model matches the frequency profile of the benchmark model. Measured values of formant frequencies are presented in Table 2. Column 4 of this table shows the zero relative errors in all six formants frequencies of vowel /i/. Column 7 shows that

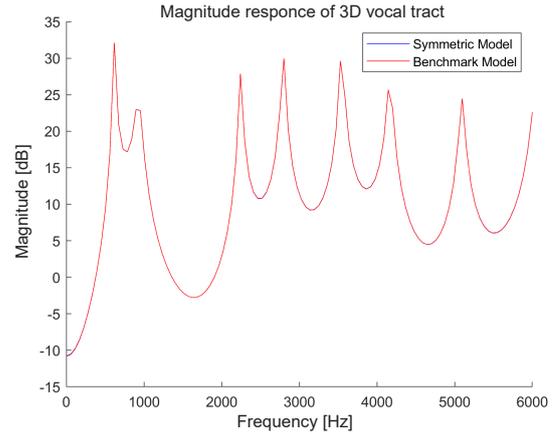


Fig. 5. Comparison between the frequency profiles of the benchmark and symmetric models in the vowel /ɔ/ case.

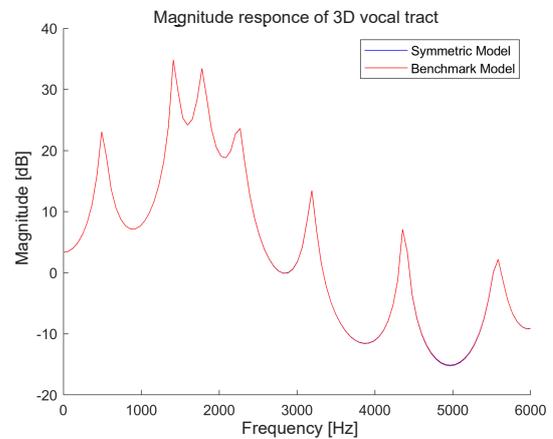


Fig. 6. Comparison between the frequency profiles of the benchmark and symmetric models in the vowel /i/ case.

the symmetric model, in this case, is about 3 times more efficient than the benchmark model.

Table 1. Comparison between the formant frequencies of the benchmark and symmetric models in the vowel /ɔ/ case.

Format frequency	Benchmark model [Hz]	Symmetric model [Hz]	Error [%]	Elapsed time (average) [sec]		Efficiency
				Benchmark model	Symmetric model	
f_1	597	597	0	627.323	209.391	2.995
f_2	894	894	0			
f_3	2234	2234	0			
f_4	2755	2755	0			
f_5	3573	3573	0			
f_6	4243	4243	0			

Table 2. Comparison between the formant frequencies of the benchmark and symmetric models in the vowel /i/ case.

Format frequency	Benchmark model [Hz]	Symmetric model [Hz]	Error [%]	Elapsed time (average) [sec]		Efficiency
				Benchmark model	Symmetric model	
f_1	492	492	0	550.188	178.868	3.076
f_2	1557	1557	0			
f_3	1885	1885	0			
f_4	2294	2294	0			
f_5	3195	3195	0			
f_6	4423	4423	0			

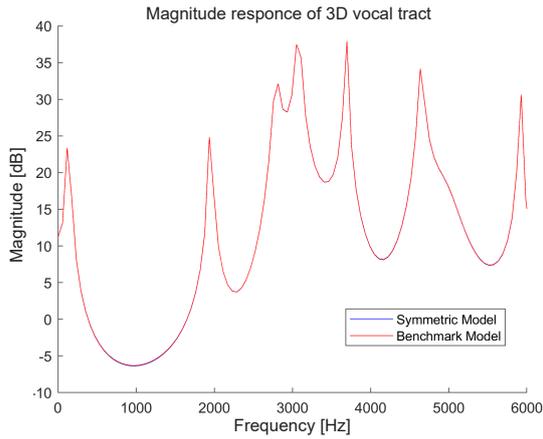


Fig. 7. Comparison between the frequency profiles of the benchmark and symmetric models in the vowel /ε/ case.

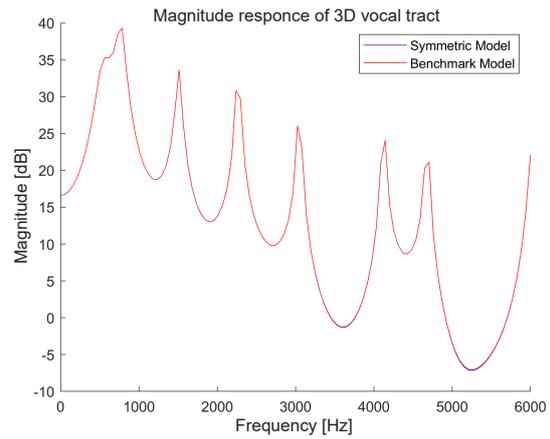


Fig. 8. Comparison between the frequency profiles of the benchmark and symmetric models in the vowel /α/ case.

Table 3. Comparison between the formant frequencies of the benchmark and symmetric models in the vowel /ε/ case.

Format frequency	Benchmark model [Hz]	Symmetric model [Hz]	Error [%]	Elapsed time (average) [sec]		Efficiency
				Benchmark model	Symmetric model	
f_1	157	157	0	513.045	155.675	3.296
f_2	1955	1955	0			
f_3	2815	2815	0			
f_4	3127	3127	0			
f_5	3752	3752	0			
f_6	4690	4690	0			

Table 4. Comparison between the formant frequencies of the non-symmetric and symmetric models in the vowel /α/ case.

Format frequency	Benchmark model [Hz]	Symmetric model [Hz]	Error [%]	Elapsed time (average) [sec]		Efficiency
				Benchmark model	Symmetric model	
f_1	819	819	0	674.704	210.66	3.202
f_2	1638	1638	0			
f_3	2307	2307	0			
f_4	3052	3052	0			
f_5	4168	4168	0			
f_6	4764	4764	0			

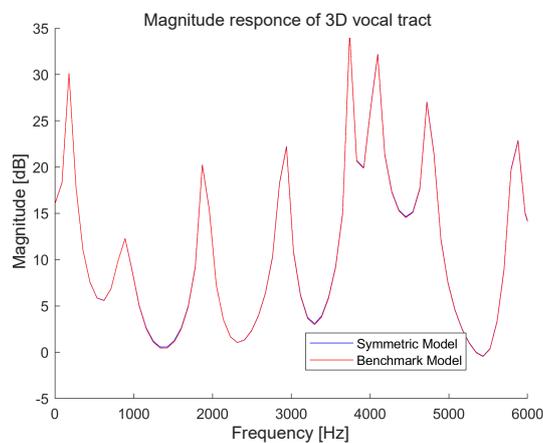


Fig. 7. Comparison between the frequency profiles of the benchmark and symmetric models in the vowel /u/ case.

Table 5. Comparison between the formant frequencies of the benchmark and symmetric models in the vowel /u/ case.

Format frequency	Benchmark model [Hz]	Symmetric model [Hz]	Error [%]	Elapsed time (average) [sec]		Efficiency
				Benchmark model	Symmetric model	
f_1	144	144	0	559.679	185.035	3.024
f_2	855	855	0			
f_3	1924	1924	0			
f_4	2921	2921	0			
f_5	3704	3704	0			
f_6	3989	3989	0			

Figures 7, 8, and 9 show the frequency profiles of vowels / ε /, / α /, and /u/ of the vocal tract, respectively. In all these cases, the frequency profiles of the benchmark and symmetric models overlap. Numerically measured values in Tables 3, 4, and 5 of column 4 present zero relative errors in all 6 formant frequencies in all cases. The efficiency of our model over the benchmark model, as shown in Figs. 7, 8, and 9, is about 3.3, 3.2, and 3 times better, respectively.

5. Conclusions

In the proposed work, the symmetric model was used for modeling three-dimensional waveguide model of the vocal tract. The computational cost was greatly reduced by employing a symmetric approach in this work. The simulation was performed on the vowels / υ /, /i/, / ε /, / α /, and /u/. By studying tables and figures, we draw the following conclusions:

- successful implementation of the symmetric approach in the three-dimensional waveguide modeling of the vocal tract was conducted;
- the formant frequencies of the symmetric model are the same as the benchmark model;
- the frequency profiles of the current model overlap with that of the benchmark model;
- the symmetric model is more efficient than the benchmark model. In all the cases, the proposed model is about 3 times more efficient than the benchmark model.

We conclude that the symmetric model presents itself as a highly efficient and accurate three-dimensional waveguide model of the vocal tract. The proposed model provides an opportunity to efficiently investigate the vocal tract's frequency response for speech production. We believe the symmetric model may serve as a useful vocal tract model in speech synthesizers and provide a new dimension for further investigations.

References

1. ARNELA M. *et al.* (2016a), Influence of lips on the production of vowels based on finite element simulations and experiments, *The Journal of the Acoustical Society of America*, **139**(5): 2852–2859, doi: [10.1121/1.4950698](https://doi.org/10.1121/1.4950698).
2. ARNELA M. *et al.* (2016b), Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds, *The Journal of the Acoustical Society of America*, **140**(3): 1707–1718, doi: [10.1121/1.4962488](https://doi.org/10.1121/1.4962488).
3. ARNELA M., DABBAGHCHIAN S., GUASCH O., ENGWALL O. (2019), MRI-based vocal tract representations for the three-dimensional finite element synthesis of diphthongs, *IEEE/ACM Transactions on Audio, Speech, Language Processing*, **27**(12): 2173–2182, doi: [10.1109/TASLP.2019.2942439](https://doi.org/10.1109/TASLP.2019.2942439).
4. BEESON M.J., MURPHY D.T. (2004), RoomWeaver: A digital waveguide mesh based room acoustics research tool, [in:] *Proceedings of the Seventh International Conference on Digital Audio Effects*, pp. 268–273.
5. BLANDIN R. *et al.* (2015), Effects of higher order propagation modes in vocal tract like geometries, *The Journal of the Acoustical Society of America*, **137**(2): 832–843, doi: [10.1121/1.4906166](https://doi.org/10.1121/1.4906166).
6. BLANDIN R., FÉLIX S., DOC J.-B., BIRKHOLZ P. (2021), Combining multimodal method and 2D finite elements for the efficient simulation of vocal tract acoustics, [in:] *Proceedings of the 27th International Congress on Sound and Vibration*.
7. GULLY A.J., DAFFERN H., MURPHY D.T. (2017), Diphthong synthesis using the dynamic 3D digital waveguide mesh, *IEEE/ACM Transactions on Audio, Speech, Language Processing*, **26**(2): 243–255, doi: [10.1109/TASLP.2017.2774921](https://doi.org/10.1109/TASLP.2017.2774921).
8. GULLY A.J., TUCKER B. (2019), Modeling voiced stop consonants using the 3D dynamic digital waveguide mesh vocal tract model, [in:] *Proceedings of the International Congress of Phonetic Sciences 2019*, Australasian Speech Science and Technology Association Inc.
9. KARJALAINEN M., ERKUT C. (2004), Digital waveguides versus finite difference structures: Equivalence and mixed modeling, *EURASIP Journal on Applied Signal Processing*, **2004**(7): 978–989, doi: [10.1155/S1110865704401176](https://doi.org/10.1155/S1110865704401176).
10. LIM Y., ZHU Y., LINGALA S.G., BYRD D., NARAYANAN S., NAYAK K.S. (2019), 3D dynamic MRI of the vocal tract during natural speech, *Magnetic Resonance in Medicine*, **81**(3): 1511–1520, doi: [10.1002/mrm.27570](https://doi.org/10.1002/mrm.27570).
11. MAKAROV I.S. (2009), Approximating the vocal tract by conical horns, *Acoustical Physics*, **55**(2): 261–269, doi: [10.1134/S106377100902016X](https://doi.org/10.1134/S106377100902016X).

12. MARKEL J.E., GRAY A.H. (1976), *Linear Prediction of Speech*, Springer.
13. MATHUR S., STORY B.H., RODRÍGUEZ J.J. (2006), Vocal-tract modeling: Fractional elongation of segment lengths in a waveguide model with half-sample delays, *IEEE Transactions on Audio, Speech, and Language Processing*, **14**(5): 1754–1762, doi: [10.1109/TSA.2005.858550](https://doi.org/10.1109/TSA.2005.858550).
14. MOHAPATRA D.R., FLEISCHER M., ZAPPI V., BIRK-HOLZ P., FELLS S. (2022), Three-dimensional finite-difference time-domain acoustic analysis of simplified vocal tract shapes, [in:] *Proceedings of Interspeech*, pp. 764–768, doi: [10.21437/Interspeech.2022-10649](https://doi.org/10.21437/Interspeech.2022-10649).
15. MOHAPATRA D.R., ZAPPI V., FELLS S. (2019), An extended two-dimensional vocal tract model for fast acoustic simulation of single-axis symmetric three-dimensional tubes, [in:] *Proceedings of Interspeech 2019*, pp. 3760–3764, doi: [10.21437/Interspeech.2019-1764](https://doi.org/10.21437/Interspeech.2019-1764).
16. MULLEN J. (2006), *Physical modelling of the vocal tract with the 2D digital waveguide mesh*, Ph.D. Thesis, The University of York.
17. MULLEN J., HOWARD D.M., MURPHY D.T. (2003), Digital waveguide mesh modeling of the vocal tract acoustics, [in:] *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 119–122, doi: [10.1109/ASPAA.2003.1285834](https://doi.org/10.1109/ASPAA.2003.1285834).
18. MULLEN J., HOWARD D.M., MURPHY D.T. (2006), Waveguide physical modeling of vocal tract acoustics: Flexible formant bandwidth control from increased model dimensionality, *IEEE Transactions on Audio, Speech, and Language Processing*, **14**(3): 964–971, doi: [10.1109/TSA.2005.858052](https://doi.org/10.1109/TSA.2005.858052).
19. MULLEN J., HOWARD D.M., MURPHY D.T. (2007), Real-time dynamic articulations in the 2-D waveguide mesh vocal tract model, *IEEE Transactions on Audio, Speech, and Language Processing*, **15**(2): 577–585, doi: [10.1109/TASL.2006.876751](https://doi.org/10.1109/TASL.2006.876751).
20. MURPHY D.T., BEESON M. (2007), The KW-boundary hybrid digital waveguide mesh for room acoustics applications, *IEEE Transactions on Audio, Speech, and Language Processing*, **15**(2): 552–564, doi: [10.1109/TASL.2006.881681](https://doi.org/10.1109/TASL.2006.881681).
21. MURPHY D.T., HOWARD D.M. (2000), 2-D digital waveguide mesh topologies in room acoustics modelling, [in:] *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFx)*, pp. 211–216.
22. QURESHI T.M., ISHAQ M. (2019), Real-time vocal tract model for elongation of segment lengths in a waveguide model, *Archives of Acoustics*, **44**(2): 287–300, doi: [10.24425/aoa.2019.128492](https://doi.org/10.24425/aoa.2019.128492).
23. QURESHI T.M., SYED K.S. (2015), Two dimensional featured one dimensional digital waveguide model for the vocal tract, *Computer Speech & Language*, **33**(1): 47–66, doi: [10.1016/j.csl.2014.12.004](https://doi.org/10.1016/j.csl.2014.12.004).
24. QURESHI T.M., SYED K.S. (2019), Improved vocal tract model for the elongation of segment lengths in a real time, *Computer Speech & Language*, **57**(4): 41–58, doi: [10.1016/j.csl.2019.02.001](https://doi.org/10.1016/j.csl.2019.02.001).
25. QURESHI T.M., SYED K.S., ZAFAR A. (2020), Non-uniform rectilinear grid in the waveguide modeling of the vocal tract, *Archives of Acoustics*, **45**(4): 585–600, doi: [10.24425/aoa.2020.135247](https://doi.org/10.24425/aoa.2020.135247).
26. RABINER L.R., SCHAFFER R.W. (1978), *Digital Processing of Speech Signals*, Prentice-Hall.
27. SCHICKHOFER L., MIHAESCU M. (2020), Analysis of the aerodynamic sound of speech through static vocal tract models of various glottal shapes, *Journal of Biomechanics*, **99**: 109484, doi: [10.1016/j.jbiomech.2019.109484](https://doi.org/10.1016/j.jbiomech.2019.109484).
28. SPEED M., MURPHY D., HOWARD D. (2013a), Modeling the vocal tract transfer function using a 3D digital waveguide mesh, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **22**(2): 453–464, doi: [10.1109/TASLP.2013.2294579](https://doi.org/10.1109/TASLP.2013.2294579).
29. SPEED M., MURPHY D., HOWARD D. (2013b), Three-dimensional digital waveguide mesh simulation of cylindrical vocal tract analogs, *IEEE Transaction on Audio, Speech, and Language Processing*, **21**(2): 449–455, doi: [10.1109/TASL.2012.2224342](https://doi.org/10.1109/TASL.2012.2224342).
30. STORY B.H., TITZE I.R., HOFFMAN E.A. (1996), Vocal tract area functions from magnetic resonance imaging, *The Journal of the Acoustical Society of America*, **100**(1): 537–554, doi: [10.1121/1.415960](https://doi.org/10.1121/1.415960).
31. STRUBE H.W. (2003), Are conical segments useful for vocal-tract simulation? (L), *The Journal of the Acoustical Society of America*, **114**(6): 3028–3031, doi: [10.1121/1.1623789](https://doi.org/10.1121/1.1623789).
32. TREYSSCÖDE F. (2021), A model reduction method for fast finite element analysis of continuously symmetric waveguides, *Journal of Sound and Vibration*, **508**: 116204, doi: [10.1016/j.jsv.2021.116204](https://doi.org/10.1016/j.jsv.2021.116204).
33. VAMPOLA T., HORÁČEK J., LAUKKANEN A.-M., ŠVEC J.G. (2015), Human vocal tract resonances and the corresponding mode shapes investigated by three-dimensional finite-element modelling based on CT measurement, *Logopedics Phoniatrics Vocology*, **40**(1): 14–23, doi: [10.3109/14015439.2013.775333](https://doi.org/10.3109/14015439.2013.775333).
34. VAN DUYN S.A., SMITH J.O. (1993), Physical modeling with the 2-D digital waveguide mesh, [in:] *Proceedings of the International Computer Music Conference*.
35. VAN DUYN S.A., SMITH J.O. (1995), The tetrahedral digital waveguide mesh, [in:] *Proceedings of 1995 Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 234–237, doi: [10.1109/ASPAA.1995.482998](https://doi.org/10.1109/ASPAA.1995.482998).

Research Paper

Ultrasonic P- and S-Wave Reflection and CPT Soundings for Measuring Shear Strength in Soil Stabilized by Deep Lime/Cement Columns in Stockholm Norvik Port

Per LINDH^{(1),(2)}, Polina LEMENKOVA⁽³⁾

⁽¹⁾ *Department of Investments, Technology and Environment, Swedish Transport Administration*
Malmö, Sweden, e-mail: per.lindh@trafikverket.se

⁽²⁾ *Faculty of Engineering, Department of Building and Environmental Technology*
Division of Building Materials, Lund University
Lund, Sweden, e-mail: per.lindh@byggtek.lth.se

⁽³⁾ *École Polytechnique de Bruxelles, Laboratory of Image Synthesis and Analysis (LISA)*
Université Libre de Bruxelles (ULB)
Brussels, Belgium

*Corresponding Author e-mail: polina.lemenkova@ulb.be

(received December 22, 2022; accepted March 20, 2023)

In this research project, the measurements of the ultrasonic P- and S-waves and seismic cone penetration testing (CPT) were applied to identify subsurface conditions and properties of clayey soil stabilized with lime/cement columns in the Stockholm Norvik Port, Sweden. Applied geophysical methods enabled to identify a connection between the resistance of soil and strength in the stabilized columns. The records of the seismic tests were obtained in the laboratory of Swedish Geotechnical Institute (SGI) through estimated P- and S-wave velocities using techniques of resonance frequency measurement of the stabilized specimens. The CPT profiles were used to evaluate the quality of the lime/cement columns of the reinforced soil by the interpretation of signals. The relationship between the P- and S-waves demonstrated a gain in strength during soil hardening. The quality of soil was evaluated by seismic measurements with aim to achieve sufficient strength of foundations prior to the construction of the infrastructure objects and industrial works. Seismic CPT is an effective method essential to evaluate the correct placement of the CPT inside the column. This work demonstrated the alternative seismic methods supporting the up-hole technology of drilling techniques for practical purpose in civil engineering and geotechnical works.

Keywords: civil engineering; soil stabilization; compressive strength; cement; lime; seismic waves.

PACS: 81.40.Cd; 81.40.Ef; 62.20.Qp; 83.50.Xa; 45.70.Mg; 92.40.Lg; 81.40.Lm; 62.20.M-

2000 MSC: 76Axx; 74Exx; 74Fxx.



Copyright © 2023 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

ASCII – American Standard Code for Information Interchange,
CPT – cone penetration testing,
DSM – deep soil mixing,
FEM – finite element method,
LKD – lime kiln dust,
OPC – ordinary Portland cement,
OCR – overconsolidation ratio,
QL – quick lime,
SPT – standard penetration test,

SGI – Swedish Geotechnical Institute,
UCS – uniaxial compressive strength.

1. Introduction

1.1. Background

Stabilization of soil is a fundamental issue in civil engineering. Prior to engineering and construction works, weak soils should be stabilized using binders

such as cement (BACHE *et al.*, 2022; CHEN *et al.*, 2021; LINDH, LEMENKOVA, 2022f), lime (KICHOU *et al.*, 2022; CHAND, SUBBARAO, 2007), slag (LINDH, LEMENKOVA, 2021a) or others (MIRZABABAEI *et al.*, 2018). Compared to native soil, stabilized mixture exhibits the improved properties such as increased strength, enhanced elastic modulus and stiffness (MADHYANNAPU *et al.*, 2010; SUNDARY *et al.*, 2022), improved liquefaction resistance (ITO *et al.*, 1994). Other advantages include the reduced porosity, permeability and shrinkage (TONINI DE ARAÚJO *et al.*, 2023; EL-RAWI, AWAD, 1981; ÅHNBERG, 2003), decreased swell potential in soils prone to freeze-thaw cycles (SHIHATA, BAGHDADI, 2001; BIN-SHAFIQUE *et al.*, 2011; ORAKOGLU *et al.*, 2017), increased resistance against the impacts from moisture and temperature (ZHANG *et al.*, 2020; KIM *et al.*, 2012). Such improved properties make soil suitable to earthworks and ensure the safety of constructions and civil infrastructure systems (roads, communications, tunnels, highways). This especially concerns the high plasticity expansive clay soils of Sweden.

Existing methods of soil stabilization can be divided into two broad categories: traditional methods of soil stabilization and evaluation of strength (TRHLÍKOVÁ *et al.*, 2012; LINDH, LEMENKOVA, 2022b; HEIDARIZADEH *et al.*, 2021) and advanced seismic methods using non-destructive techniques of acoustic soundings (VARMA *et al.*, 2022; FOTI *et al.*, 2002; LINDH, LEMENKOVA, 2021b; GARCIA-SUAREZ *et al.*, 2021). The traditional methods of evaluation of the improvements in the stabilized soil include the uniaxial tests (AVCI, MOLLAMAHMUTOĞLU, 2016; LAPOINTE *et al.*, 2012; XU, YI, 2021) or triaxial (ALVARADO, COOP, 2012) using laboratory equipment which are robust and widely applied approaches. However, they have certain limitations and disadvantages. Firstly, these methods are destructive, i.e., tested specimens are crashed after the experiments. Secondly, these tests are not applicable to be carried out on the fissured clay. Moreover, the uniaxial compressive strength (UCS) may be not precise for soils with the angle of shearing resistance not equal to zero where the shear strength is not equal to half the compressive strength. On that basis, the use of the applied methods of data analysis based on mathematical modelling (JEFFERIES, 2022) is better applicable for evaluation of soil characteristics.

The applied geophysical methods have been widely used in the context of seismic and acoustical tests. These are based on the evaluation of the velocities of P- and S-waves penetrating the soil (SAFAEE *et al.*, 2022; LINDH, LEMENKOVA, 2022d, FOTI, LANCELLOTTA, 2004). Other examples include the Rayleigh–Ritz method which evaluates the free vibration characteristics in the ground (SONKAR, MITTAL, 2022; JONES, 1958) or attenuation of waves through the

evaluation of vibration over time (COLOMBERO *et al.*, 2015). Several experimental papers have been published on the application of wave velocity to evaluate soil parameters. Among these, SANTAMARINA and CASCANTE (1998) report on the existing relationship between the wave propagation and the inherent properties of the materials, such as fabric, mineralogical structure, surface roughness, size of particles and angularity, which control soil strength and stiffness.

Ultrasonic and bender element tests used to measure the elastic modulus and evaluate the rigidity of soil as the shear modulus from compression waves are reported in (AMARAL *et al.*, 2011) with a case of sand samples stabilized by cement. The use of the free-free resonant column tests in combination with the UCS tests has also proved to be applicable to the sand-cement mixtures for measurement of the electrical resistivity and mechanical properties of the stabilized soil (RUSATI *et al.*, 2020). Other examples in this category include measuring the shear wave velocity during the penetration testing (HEPTON, 1989), and determining the dynamic shear modulus of the ground, independent of Poisson's ratio (ABBISS, 1981).

The cone penetration testing (CPT) is an effective method of soil investigation and identification of the subsurface conditions. Originally developed in the Netherlands in the 1930s and initially named as the Dutch cone test (BROUWER, 2007; MCCALLUM, 2014), it has nowadays numerous applications in geotechnical engineering, such as prediction of liquefaction resistance (OLSEN, 2018; FITZGERALD, ELSWORTH, 2012; SAYE *et al.*, 2021), evaluation of strength (PRICE *et al.*, 2016; JAMIOLKOWSKI *et al.*, 2003), and measuring excess pore pressure (SULLY, CAMPANELLA, 1991; ELSWORTH *et al.*, 2006). Nowadays, CPT is one of the most used methods for measurements of the geotechnical properties of soil. For instance, the applications of CPT for estimation of deformation modulus of soil (SHAHIEN, FAROUK, 2013; BENZ NAVARRETE *et al.*, 2022) and characterising soil liquefaction potential by GUAN *et al.* (2022) are worth mentioning. The related methods include, for instance, estimating the cone penetration resistance for analysis of the soil compressibility (BISHT *et al.*, 2021).

The deep soil mixing (DSM) method is an in situ stabilization of soil in which soil is mixed with binders, typically cementitious materials, lime, slag or similar stabilizing agents. Over the last decades, there has been a rise in research into the DSM methods, as a response to the appearance of novel construction practices and equipment in addition to the existing general techniques of drilling in cored boreholes (HEPTON, 2015). These are the techniques of stabilizing the unsaturated expansive subsoils at the moderate active depths (MADHYANNAPU, PUPPALA, 2014; MADHYANNAPU *et al.*, 2009), or settlement control in the highway embankment (ARCHEWA *et al.*, 2011).

The cases of the reinforced embankments supported by the DSM cement columns are presented by LAMBRECHTS *et al.* (2012) for stability and the settlement control, BERGADO *et al.* (2008) with computed permeability and the compressibility ratio and the consolidation of the columns and clay, and YAPAGE *et al.* (2014) with evaluated settlements, excess pore-water pressures and lateral deformations in a strain-softening behavior of the deep cement columns.

Cohesive soil with high moisture content and fine-grained structure, such as clays, silts or loams, are best stabilized by lime and cementitious binders. As a result, soil has enhanced parameters which result in the increased strength, reduced permeability and compressibility. However, the engineering properties of soil stabilized by the DSM method depend on various factors including the following ones: the original characteristics of the native soil, the types and the amount of binders, technical and operational parameters defined by the construction types, curing time, depth, external parameters such as temperature and moisture, and loading conditions.

In this paper, we propose a framework of novel applications of the DSM and seismic methods to address the problem of the evaluation of strength of the expansive clay stabilized with lime/cement columns. The methodology includes the in-situ fieldwork and laboratory based measurements and modelling. The CPT soundings were performed in the area of the Stockholm Norvik Port. The resistance and strength of soil in the stabilized columns was evaluated by seismic tests. A method of resonant frequency measurements of the compressional wave velocities was embedded into these tests. We applied the alternative seismic methods supporting the up-hole technology of the drilling techniques. This enabled to evaluate the pressure by the S- and P-waves for lime/cement columns. We report the results of seven columns with performed CPT probing (ACS155, ACW151, ADA147, ADC148, AEG35, AEE356, and AEA358). We evaluated the relationship between the P- and S-waves which showed the gain in strength.

1.2. Objectives and goals

The background to the project is the need to correctly evaluate the quality of the lime/cement columns in terms of strength and homogeneity in the region of the Stockholm Norvik Port – a new deep sea port of Sweden in the Baltic Sea. Norvik consisted of a bay until the beginning of the 1980s, surrounded by a hilly strip of land and a mountainous island. Nowadays, this area is largely reconstructed for the container terminal of the Stockholm Norvik Port. The Stockholm Norvik Port is an important port transporting cargo in the capital with a direct connection to other regions of Sweden in the Baltic Sea and northern Europe. The

sustainable and efficient operation of the infrastructure in the port modalities requires safe constructions on the stabilized ground. Therefore, the objectives of this work are to determine the strength properties of soil stabilized by lime/cement columns using the CPT. The CPT was evaluated as a test method for the stabilization of soil prior to construction works. The practical goal of this work is to connect the laboratory tests and field measurements to obtain a better and safer optimization of the binders in terms of types and proportions for soil stabilization. In the response to these needs, the aim is to evaluate seismic methods for evaluating the strength of the soil, stabilized by various binders.

2. Methods

2.1. Fieldwork

In this project, the field investigations were carried out during the period of 7–10 November 2016 in Norvik. The technical equipment included the multi-purpose drilling rig GEORIG – Model 607 used for soundings. The GEORIG – 607, developed by Geotech AG, was selected as a proper instrument for drilling, since it is equipped with all the necessary devices for soil rock drilling, Swedish weight sounding, dynamic sounding, standard penetration test (SPT) and CPT. The GEORIG 607 was anchored with a screw to achieve the sufficient holding force, Fig. 1.



Fig. 1. Process of CPT sounding in Norvik area.
Photo source: Per Lindh.

During the construction of a gas storage facility in Nynäshamn, the bay was filled with the explosives from the construction. The filling was carried out by the tipping masses from several fronts, which meant that the large volumes of clay were enclosed in the blast-stone filling. The two large areas, northern and southern ones, were identified and used as study areas, filled with clay up to 30 m. Previously, several geotech-

nical field investigations in this area were carried out in 1982, 2008, 2010, and 2011 with technical details reported clay properties and stabilization (ERIKSSON, 2015). The requirements for the lime/cement columns in this project were set on the achieving of a minimum in situ shear strength of 150 kPa after 28 days of curing period. To ensure this, the requirement for the laboratory packed samples is set to at least 200 kPa. Several mixtures of the stabilized soil that contained cement, quicklime and the lime kiln dust (LKD) met the requirements of at least 200 kPa after 28 days of storage. Since the in situ temperature in the columns is partly dependent on the type and the amount of binder, the laboratory tests differed from the fieldwork results, especially when the laboratory cured samples were stored at 7°C in a climate room where the cooling is significantly greater than in in-situ conditions.

2.2. Laboratory tests

2.2.1. Test chamber

Prior to the laboratory tests, the test chamber was well examined and documented in terms of technical conditions for natural soil collected in the Stockholm Norvik port. The laboratory measurements on the stabilized soil were carried out using seismic tests of P-wave and flexwave (shear wave) velocities that evaluated the strength of the stabilized samples. Seismic tests were carried out as a reference to the measurements in the field.

2.2.2. Determination of the material parameters

The soil stabilized in this project included natural clay with parameters reported in Figs. 2 and 3. The overall density values vary in the range of 1.40 to 2.00 t/m³ in the depth up to 25 m (Fig. 2, right). The values of both the natural water ratio (W_N) of soil and the yield strength (W_L) vary from 40 to 100% (Fig. 2). The density mostly varies between 1.5 to 1.7 tons per m³. The corrected shear strength, $c_u\text{-korr}$, is mainly between 5 and 15 kPa and the clay is low to medium, in terms of sensitivity.

2.2.3. Binders

Three different binders were used in this study: 1) ordinary Portland cement (OPC) of type CEM II/A; 2) burnt lime, or quick lime (QL); 3) LKD. These binders were selected due to their effectiveness and applicability, as tested in previous studies (LINDH, LEMENKOVA, 2022c). Using these three general types of binders, nine combinations of the blended mixes were fabricated and tested. The mixing quantities corresponded to 80, 100, and 120 kg of binder per m³ of clay. To evaluate binder combinations for stabilization using the deep mixing method, the experimental setup using the simplex centroid design was used following the existing methodology (LINDH, LEMENKOVA, 2022e). Hence, various percentage of the three binders was tested in an experiment which consists of a triangle with pure binders in the corners and a mixture of two binders along the triangle's stripes.

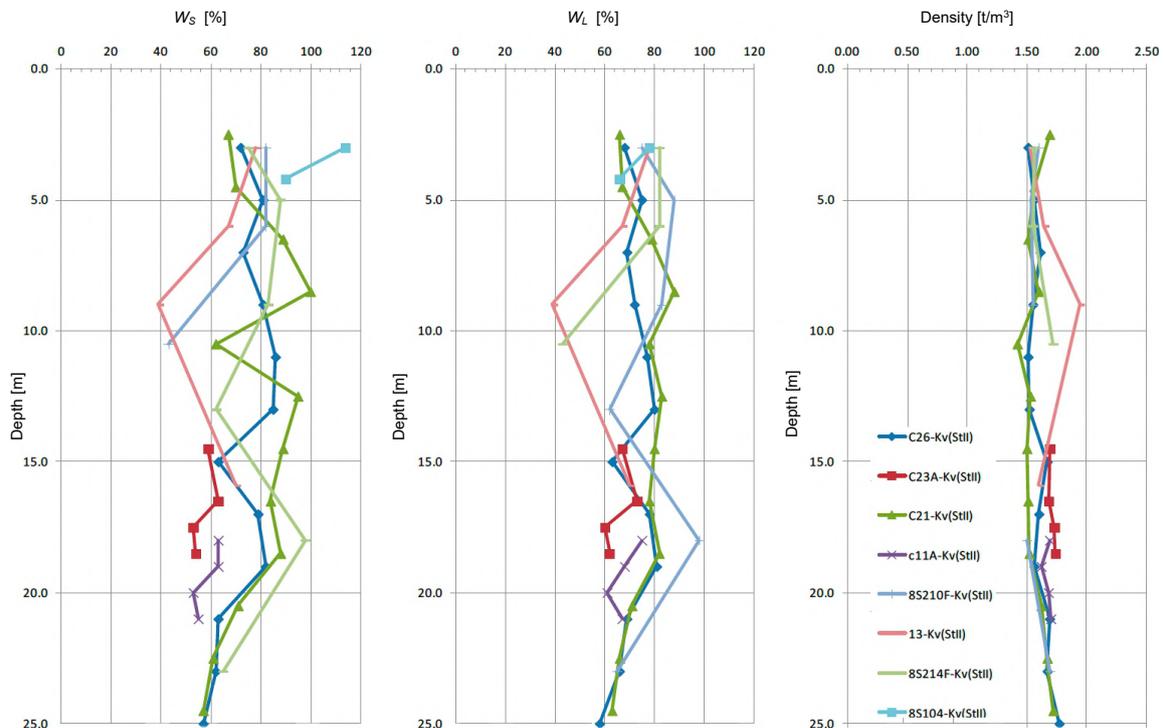


Fig. 2. Variations in water ratio, yield strength and density in natural clay over depth.

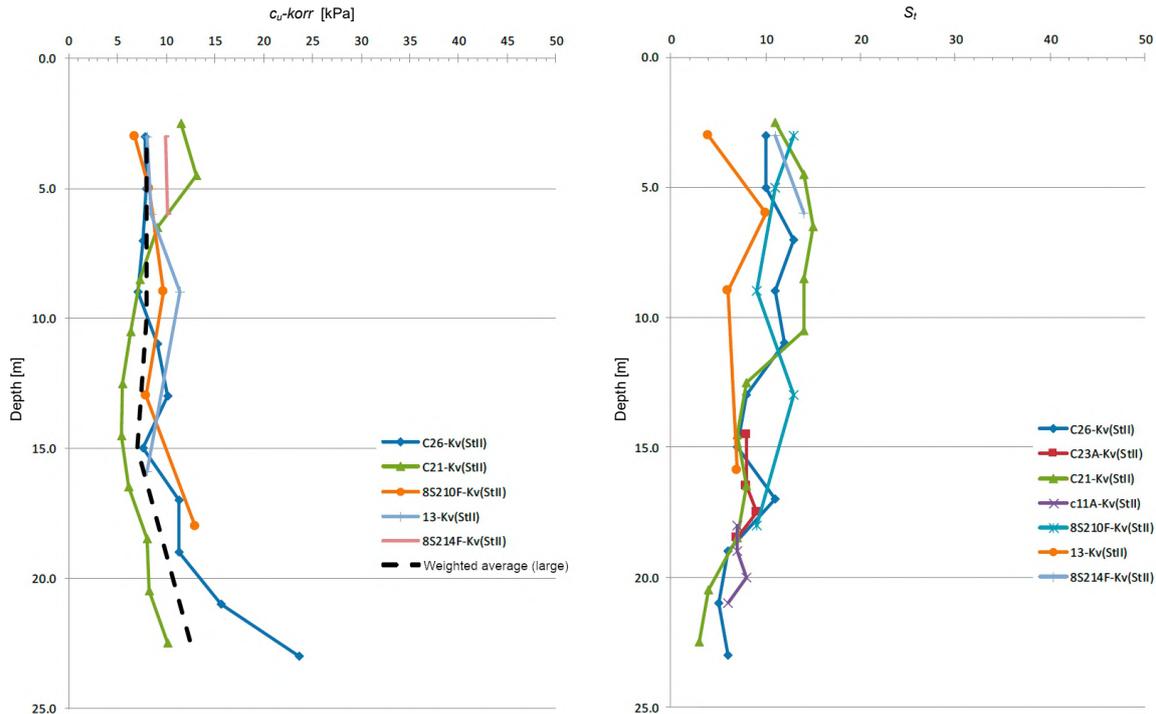


Fig. 3. Shear strength and sensitivity changed over depth for the large clay area.

A combination of all the three binders was tested experimentally inside the triangle with a total binder content as a constant. The trial surface in a three-factor simplex centroid is structured as a triangle with the corners of the triangle representing 100% of one binder and the centre of the triangle corresponds to 33% of the three different binders. The mutual ratio between the binders varies depending on where in the triangle the test point is located. These tests were carried out according to the laboratory quality handbook of the Swedish Geotechnical Institute (SGI) (document 29a, rev c). Each trial point was performed as a double trial. The points on the lower edge of the triangle correspond to either a pure binder (cf., CEM II/A or QL), or, alternatively, as a mixture of 50% CEM II/A and 50% QL. Testing the proportions and the amount of binders was carried out on clay collected from the test excavation sites carried out at Norvik.

The curing times for the samples was determined as 7, 14, 21, and 28 days with measurements of strength performed on the reference days. Some extra specimens were evaluated on the 80th day of curing. In addition, extra reference tests were carried out for calibration of seismic measurements in relation to the compressive strength at 7 and 14 days. The experimental setup was performed as a three factor simplex centroid with a restriction limitation of the LKD content which was increased up to 50%. The purpose of the statistical trial planning is to minimise the number of trials in view of the large amount of materials that should be tested in real conditions (dozens tons of soil) and to statis-

tically optimise the binder blends. At the same time, the experimental planning maintains the statistical significance and ensures that both positive and negative interactions between the binder components are detected.

2.3. Seismic tests

The natural resonance frequency (f_n) is the number of oscillations per second (Hz) in a test body that is allowed to oscillate and swing freely without damping. The lowest resonant frequency is called the fundamental mode. All natural resonance frequencies can be physically related to the elastic constants, E-modulus (E) and transverse contraction number or seismic velocities such as primary wave (P-wave) and secondary wave (S-wave or shear wave). For specimens with a length twice the diameter ($L/D \geq 2$), the one-dimensional (1D) wave propagation velocity was calculated using Eq. (1):

$$V_{P1D} = 2Lf_d, \tag{1}$$

where f_d is a damped resonance frequency.

Seismic measurements were performed as a resonance frequency measurement of the P-wave and S-wave, Fig. 4. The advantage of the P-wave measurements is that these can be carried out on the specimens while they harden in the sleeve so that the sleeve does not significantly affect the measurements. The disadvantage of measuring the P-wave is that it is strongly affected when soil becomes saturated with water, when

S_r goes towards one. When soil is saturated with water, the measurement of P-wave velocity of soil sample is affected by the included water for which the P-wave velocity is about 1500 m/s. Necessarily, this may lead to the incorrect and biased results. However, this bias is not significant in the laboratory-made specimens, because these cannot be saturated with water without a very high pressure.

Measuring the shear wave instead enables to solve this problem in cases when it is not affected by high water saturation levels. However, the measurement of the true shear wave is more difficult with a free-free resonant column setup. A common way to evaluate a shear wave, although not entirely accurate, is to measure the bending mode, i.e., flex mode or transverse mode (VERÁSTEGUI-FLORES *et al.*, 2015). According to this approach, the shear wave velocity is underestimated by approximately 5% for a specimen with a slenderness factor of 2 (where slenderness is a ratio of length/diameter) and $a = 0.2$. The setup for measuring the flex mode is shown in Fig. 4.

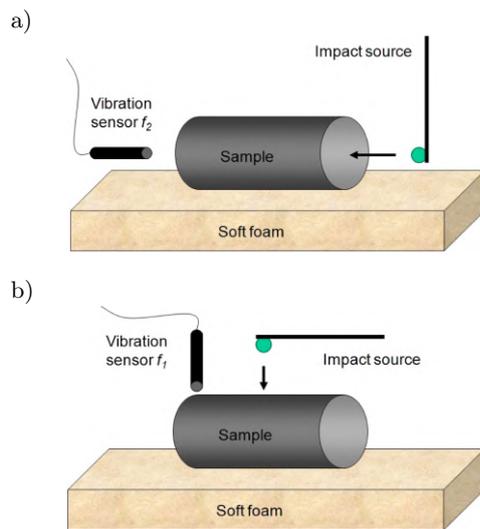


Fig. 4. Setup of test measurements: a) P-wave frequency of a sample body by a free-resonance column approach; b) flex wave frequency (f_f) of a sample body using a free resonance frequency of column.

The evaluation of the shear wave velocity is calculated according to Eq. (2):

$$V_s = 2Lf_f. \quad (2)$$

To measure true shear waves in a laboratory, it is necessary to use bending elements. However, this methodology is cumbersome and requires a long time. To solve this problem, a new methodology has been tried at the SGI to evaluate the torsional mode aimed at a better calculation of the shear wave, Fig. 5. However, this approach is not yet sufficiently tested and should be verified with more measurements and analysis using the finite element method (FEM) to check that the correct mode is evaluated.

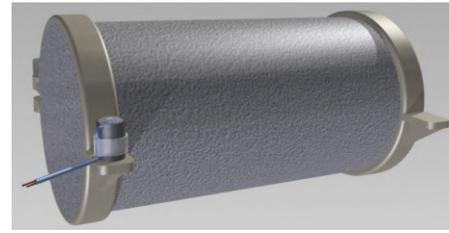


Fig. 5. Sketch of the experimental setup for measuring torsion mode according to the free-free resonance column principle.

To evaluate how both the P-wave velocity and shear strength develop with time, reference soil samples were fabricated. The reference samples consisted of a standard recipe consisting of 50% OPC and 50% QL. The advantage of the use of P-wave velocity in the laboratory tests is that it can also be measured on soil specimens stored in sleeves. However, in such case, the length of the sleeves must be equal to the length of the sample. Otherwise, the resonance frequency of the sleeve is measured additionally, which biases the results. The binder quantities were chosen corresponding to 80, 100, and 120 kg/m³. After 7 days of curing, the P-wave velocity and shear strength were evaluated on the two samples with each binder quantity. This was repeated on days 14, 21, 28, and 80, respectively. This type of correlation has previously been used for the surface-stabilized specimens and indicated robust results (LINDH, LEMENKOVA, 2022a). The choice of 80 days was based on the degree days to compare with the 28 day strength of samples stored at 20°C. A better procedure would have been to use maturity numbers (M_T) instead of the degree days.

2.4. Seismic CPT

The CPT probe is performed according to the standard SS-EN ISO 22476-1:2012 for subsurface exploration. The reason for using seismic CPT is that in this method, a peak pressure and both seismic values obtained from the P-wave (compression wave) and S-wave (shear wave) are evaluated. During the CPT probing, a cylindrical probe with a cross-sectional area of 1000 mm² is driven, where the probe has a tip angle of 60°. The pushing of the instrumented cone into the ground was performed at a continuous rate of 2.0 cm/s. During the CPT, the force required to drive the probe down was measured by the mechanical measurements to evaluate the penetration resistance of soil when pushing a cone with a conical tip into the soil. The casing friction was measured through a backlash coupling, to distinguish it from the tip pressure. The pore pressure generated during the pushing was measured using a filter system located behind the probe tip, and the friction force gauge was used to measure the friction ratio between the sleeve friction and the

tip resistance, which was measured as a percentage, see the scheme in Fig. 6.

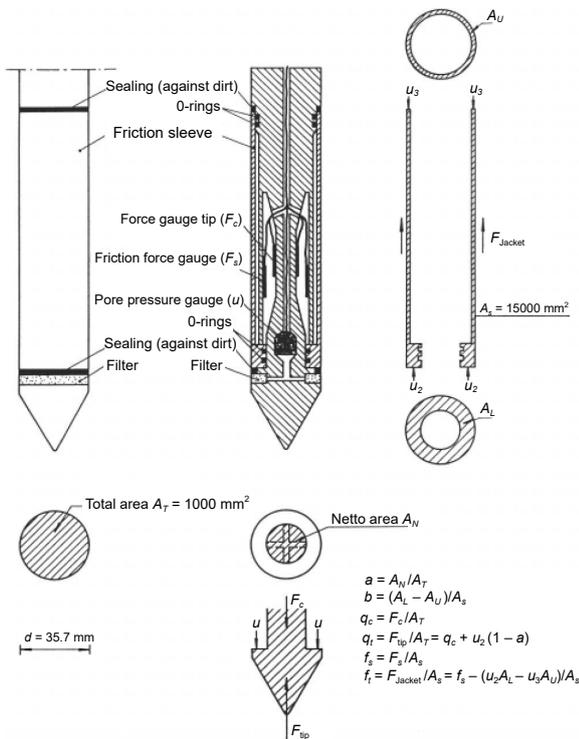


Fig. 6. Schematic diagram of the composition of a CPT probe. Modified after: SGI Information 15.

During the tests, the tip resistance (MPa) is recorded as a force required to push the tip of the cone through soil. It is determined as the force per unit area obtained by dividing the measured force by the cross-sectional area of the tip. This peak pressure is denoted as q_c if the pore pressure is not taken into account and q_t if the value is corrected taking into account error sources from the effects of the soil pores. In a special case, when $u \approx 0$ or is negligible, q_c becomes q_t . The CPT probe is reported and registered in standard protocols using Conrad software, as shown in Fig. 16. The sounding was carried out following the traditional methods of the CPT sounding. The specifics of the fieldwork is that at the depth where seismic measurements were carried out, the CPT sounding was paused, while the P- and S-wave measurements started.

The measurements used a sledgehammer which was connected with an earth cable to the measuring equipment to generate shear and compression waves. At the start of seismic measurements, the problems arose with the signal, which could be traced to motor vibrations and transmission between the drill chuck and drill steel. This was resolved by releasing the drill chuck and shutting down the motor of the drill track carriage. The CPT soundings were not performed at the centre of the columns, to avoid the disturbances at the centre arising from the column installation itself.

The Seismic CPT included the measurements of shear and compression waves, in addition to the regular CPT soundings. The measurements were performed using a series-connected device that was placed between the CPT probe and the probing rods. The measurements were performed by stopping the CPT probe at the required level of 25 m. Alternatively, the systems with continuous measurements can also be used. In such cases, measurements should be stopped after splicing the bars.

There are two different principles of seismic measurements.

The first principle is based on the two rounds of accelerometers (A1 and A2) with a fixed distance between them, Fig. 7a. The accelerometers measure the x -, y -, and z -directions, where y -direction identifies the pitch. This principle calculates the penetration speed based on the time measurement between the first

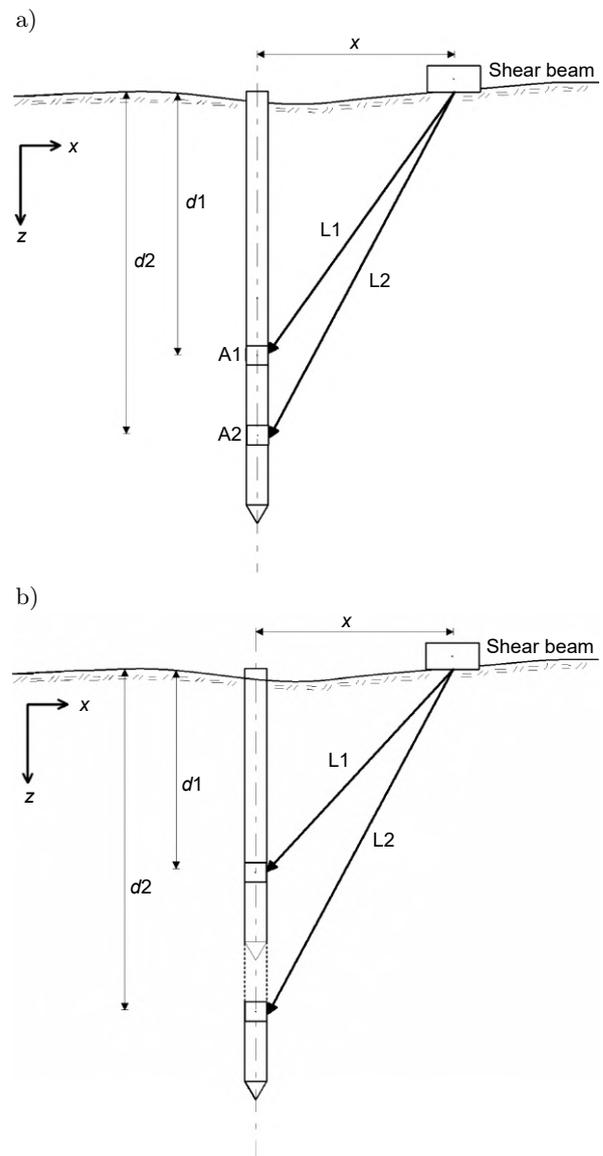


Fig. 7. Schematic diagram of a seismic CPT probe with: a) dual accelerometers; b) one accelerometer.

wave propagation to accelerometers 1 and 2. Therefore, this method is sometimes referred to as “true-time”, since the fixed distance between the accelerometers provides a safer determination of the time difference and thus a precise determination of the ground running speed, Fig. 7a. As the distance between the accelerometers is constant, the speed can be calculated directly.

The second principle is based on carrying out the measurements with only one set of the accelerometers with registration in three directions (x , y , and z). The measurement is first performed at one level, after which the CPT probing continues and a new measurement is performed at the next level, Fig. 7b. In this case, the accelerometer measures signals in the x -, y -, and z -directions, where y -direction identifies the pitch. This method is often called “pseudo-time”, as the distance is based on the two different measurements where the difference in depth is taken from the CPT logging. First, a depth measurement $d1$ is performed, after which the probe is pressed down to the depth $d2$ to continue with the next measurement, Fig. 7b.

An alternative to the above is to combine the limestone probe with the seismic CPT approach. In cases where a more reliable determination of the soil properties is needed, we recommend the application of these alternatives. An alternative to using the seismic CPT is the up hole technique for quality control of the lime/cement columns. This methodology involves the installation of the pipe in or near the centre of the column which means that the pipe is installed using lime/cement column machine. Such technology was developed for installation of the measuring pipes and reported in previous investigations of the SGI in SD Technical Report 35. This can be a good solution to quick and easy installation of the pipes for drilling.

After the initial hardening of soil specimens, the measurements of soil strength were performed using the following techniques. On the top of the lime/cement column, four accelerometers were installed, positioned with a 90° angle between the accelerometers. The test included the evaluation of the P- and S-wave sources which were lowered into the pipe to the target depth, after which the probe was clamped as a “packer”, the waves were excited and their velocity was measured, Fig. 8. The experiment was repeated in the directions where the accelerometers are placed. After the measurements are finished on one level, the probe is moved up to the next level, where the measurements continued in an iterative manner.

The advantage of this method is that it ensures a good contact between the columns and the accelerometers. Besides, the excitation wave and the location of the seismic source are very well defined. Moreover, an additional advantage is that the measurements can be carried out by one person without costs for drilling rigs, which significantly decreases the financial

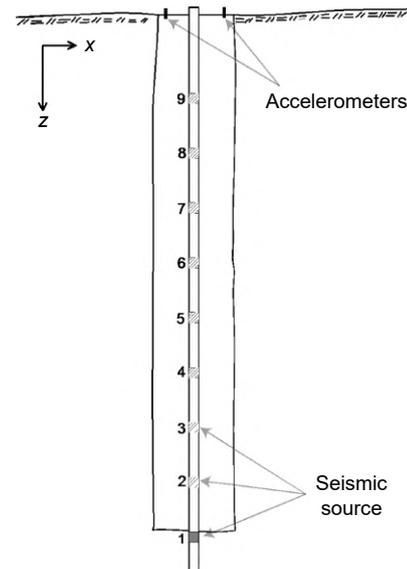


Fig. 8. Schematic sketch of the “up hole” measurements of the lime/cement column.

costs in the project. However, there are also some disadvantages of this method which should be mentioned. The quality of the columns is usually the worst at the top. This can be solved either by excavating the top or by attaching the accelerometers at a deeper level. Another disadvantage compared to the seismic CPT is that it is necessary to decide in advance which columns are to be measured. Therefore, in order to obtain a direct comparison with the limestone probe, a slotted pipe can be used for the up hole measurements, after which a normal test with the limestone probe is carried out.

2.5. Shear strength evaluation from CPT

The evaluation of the CPT probe has been performed using Conrad software, version 3.1. In a fine-grained soil, shear strength is primarily evaluated as the net peak pressure. An alternative method for clays is to use the generated pore overpressure. The empirical relationships developed for the conditions of Swedish soil are based on the evaluated dataset containing field measurements and laboratory data from a variety of soil types collected in Sweden. For natural clay soil, the relationship between the net peak pressure and shear strength is sensitive to the yield strength of soil (W_L) and to some extent also to the degree of the overconsolidation ratio (OCR), which is described according to Eq. (3):

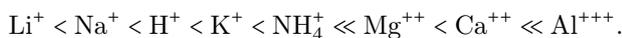
$$c_u = \frac{q_t - \sigma_{v0}}{13.4 + 6.65W_L} \left(\frac{\text{OCR}}{1.3} \right)^{-0.2} \quad (3)$$

In cases where values of the yield strength of clay are missing, Eq. (4) is used instead:

$$c_u \approx \frac{q_t - \sigma_{v0}}{N_{kt}}, \quad (4)$$

where N_{kt} (empirical confactor) is set to 16.3 for clay. In the Conrad evaluation software, N_{kt} is denoted by N_{11} . For a binder-stabilized soil, there is not the same amount of the empirical evidence and therefore, not much literature about which value of N_{kt} should be used. There are recommendations that N_{kt} values should be set between 17–20 for a stabilized soil (LARSSON, 2017; MAKUSA *et al.*, 2014). Furthermore, there are limitations in evaluation of the firmness in different layers of highly layered soil, because the tip pressure is affected from the above as the underlying layer at a distance of 5 to 20 times tip diameters. To evaluate a firmer layer in a loose soil using this method, the thickness of the layer should be between 4.4–4.7 m. In the opposite case, with a looser layer in a firmer soil, the thickness of soil layer should be between 0.2 to 0.4 m. This means that the diapason of weakness in a lime/cement column can easily be missed in the CPT evaluation. This becomes the most important issue for the evaluation of a lime/cement column reinforcement in the shear zone.

In a binder-stabilized soil, the pozzolanic reactions take place and continue for many months, therefore the properties of the stabilized soil (texture, structure, strength, etc.) partly change over time. For instance, change in the texture of clay is based on the exchange of ions in the clay and change in the water ratio. The ion exchange is described by the lyotropic series:



The change in the structure of the clay mineral is shown in Fig. 9.

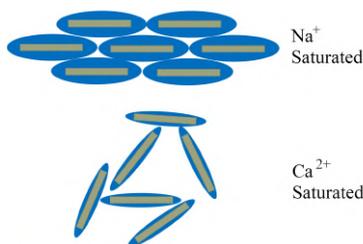


Fig. 9. Change in texture and water-holding of clay as it moves from a sodium saturated system to a calcium-saturated system.

Thus, the stabilized clay has a change in plasticity, yield strength, and other parameters. The degree of change depends both on the type of clay mineral and on the type of binder. The changes in the structure of the stabilized clay mean that the empirical relationships that exist between the tip pressure and shear strength of the original non-stabilised clay differ for a stabilized clay. To evaluate samples cured at different temperatures, soil specimens can be compared between laboratory-processed samples cured in climate rooms at 7°C and those fabricated in situ under the differ-

ent temperature conditions. In this case, the maturity number M_T is defined according to Eq. (5):

$$M_T = [20 + (T - 20) \times K]^2 \times \sqrt{t}, \quad (5)$$

where T – temperature (°C), t – time (days), K – material-dependent empirical constant. The material-dependent empirical constant K is calculated by curing samples at different temperatures, e.g., 7°C and 20°C. The samples were then measured with the resonance frequency measurement at different time intervals. The results from the seismic measurements were used to fit the curves so that they coincide, i.e., different values of K are used, after which the curves are compared. A common value of K accepted in this study is 0.5, although literature has also shown $K \approx 1$ (ÅHNBERG, HOLM, 1987).

3. Results and discussion

3.1. Results of the lime/cement column recipes

3.1.1. Strength determination

In the testing of binder combinations, a reference recipe of the OPC CEM II/A and QL were used in a mutual ratio of 50/50%. The evaluated shear strength for reference samples at different curing times is reported in Fig. 10. Here the quantities of binders represented 80, 100, and 120 kg of binders per m³ of clay. The samples were stored in a climate room with a temperature of 7°C. The results show the higher strength in the samples stabilized with 120 kg of binders per m³ of clay, while the lowest values for the 80 kg of binder. Thus, on the 80th day of stabilization, the highest values of shear strength reached 430, 445, and 465 kPa for soil mixtures stabilized with 80, 100, and 120 kg of binders, respectively. This indicates that the amount of binders directly affects the strength of soil in the final output.

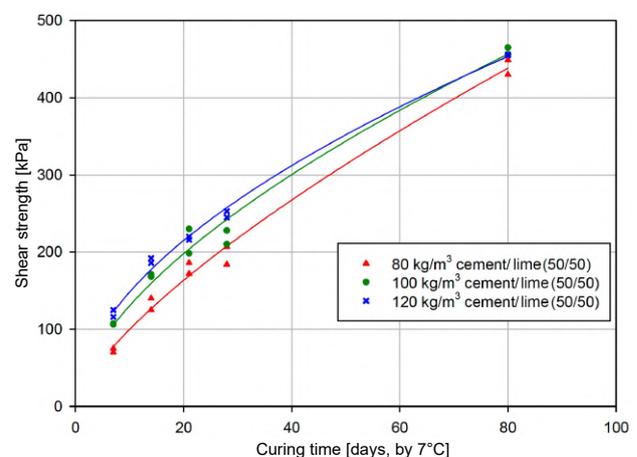


Fig. 10. Shear strength as a function of curing time for test specimens stabilized by a combination of CEM II/A and burned lime (50/50).

In order to evaluate the effects of different combinations of binders on shear strength of the stabilized soil material, a test setup was used supported by Pareto diagrams, Fig. 11. The Pareto chart shows the effect of the different components of binders on the stabilization results with the achieved significance value $p = 0.05$. For a binder amount at 80 kg/m^3 , the results show a clear positive interaction between various components (OPC, QL, and LKD) in a binder combination, Fig. 11. The direct and positive interaction means that binder components generate a higher strength than can be expected from a linear regression between the effects of the individual binders. The model has an explanation rate of 95%, i.e., it explains 95% of the variations in shear strength. The setup of test for binder amount of 100 kg/m^3 clay contains an extended number of tests. For comparison with other amounts of binders, the analysis was performed with the same number of trial points as for 80 and 120 kg of binders, respectively. The results of this analysis show a lower positive interaction between the binders. However, this should be compared with the result from the analysis of the extended trials. In a trial setup with more internal trial points, the resolution increases, and the degree of the explanation rate of the model increases as well, from 95% to 97.5%. The extended model shows a greater interaction between various binder components for a mix amount corresponding to 120 kg of binder per m^3 of clay. At this mix amount, the positive interaction of the binder components was insignificant. The degree of explanation of the response surface, according to the Response Surface Methodology (RSM) (MYERS, MONTGOMERY, 1995; MONTGOMERY, 1996) was 98%.

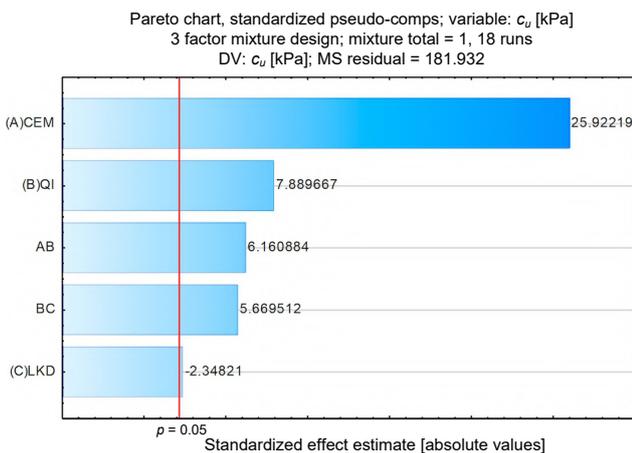


Fig. 11. Pareto chart showing the effects of binders on soil stabilization with the significance level 0.05 for a mixture of 80 kg of binders per m^3 of soil.

3.1.2. Seismic measurements

Seismic measurements included the evaluation of the velocity of P-waves, according to the resonance frequency method. The P-wave is an axial wave pass-

ing through the cylindrical specimen. In this case, the P-wave was correlated against the shear strength of the stabilized soil specimens. For the dimensioning of the lime/cement columns, the values of shear strength of the material were used. Various binder recipes were assessed for shear strength of soil specimens obtained after a certain curing time. The connection between the P-wave velocity and the compressive strength is indirect and material-dependent, i.e., based on the composition of the soil material. A correlation between the S-wave velocity and shear strength of the stabilized soil is reported in Fig. 12a.

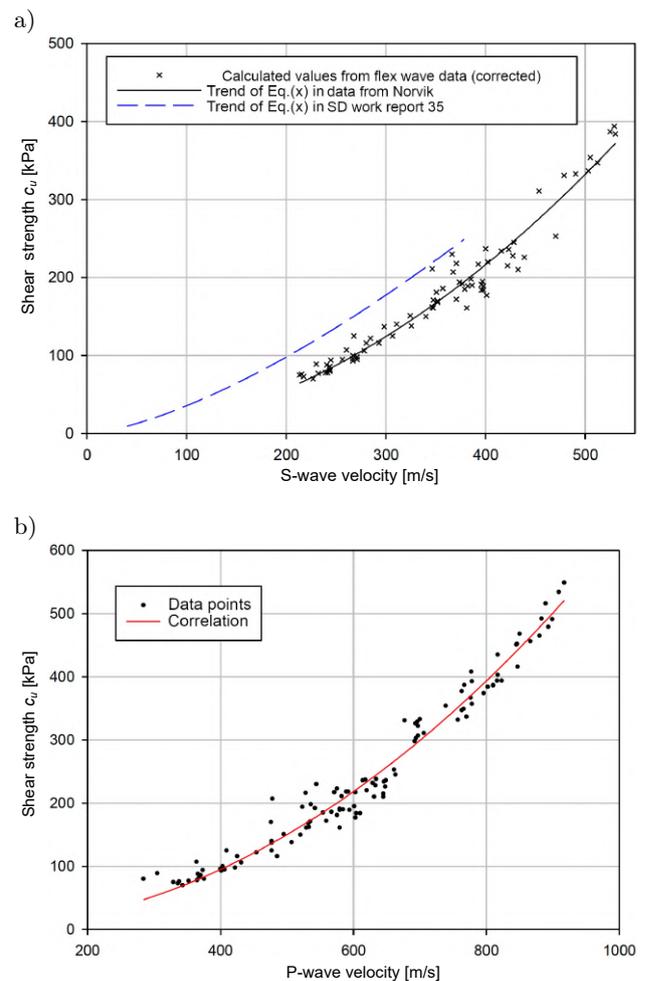


Fig. 12. Shear strength as a function of: a) S-wave speed for stabilized clay from Norvik; b) P-wave rate for stabilized Norvik samples.

Here, the shear wave velocity is adjusted according to Eqs. (6) and (7) (VERÁSTEGUI-FLORES *et al.*, 2015):

$$\tau_{fu} = 0.0424 \times V_s^{1.462}, \quad (6)$$

$$\tau_{fu} = 0.0021 \times V_s^{1.9244}. \quad (7)$$

Here Eq. (6) comes from a reference data using Report 35, and Eq. (7) comes from the data investigation

using stabilized clay from Norvik. The results from the above measurements can be used for the shear strength prediction based on measured P- and S-wave velocities from the in situ measurements, e.g., with down hole measurements. For the Norvik project, the major measurements were the P-waves where a correlation between the P-wave velocity and the shear strength of the soil samples was evaluated and presented in Fig. 12b. The relationship between the P-wave and shear strength corresponds to Eq. (8):

$$\tau_{fu} = 0.0004 \times V_p^{2.0497}. \quad (8)$$

To make a correct comparison between the different stabilized soil samples solidified with varying curing ages, the ambient temperature was taken into account. Thus, in the laboratory tests, the specimens were stored in the climate rooms with a constant temperature maintained at 7°C, where exothermic energy from specimens is cooled away, which affected strength development in soil. In contrast, the temperature in the field is significantly higher due to the less effects from cooling.

The temperature in the field also depends on the amount and type of binder. The comparison between the samples hardened at different temperatures is most appropriately done using the maturity numbers. In Fig. 13, the P-path measurements are reported as a function of curing time and recalculated to the maturity numbers on soil samples cured in 7 and 20°C. In the calculation of the maturity rate (Eq. (5)), the K factor was set to 0.55 by the empirical fitting. Figure 13 shows the differences in the evaluated strength between the samples cured in the laboratory conditions with those from the in-situ field measurements before the final strength is achieved. In the Norvik project, no temperature measurements were carried out in the field. The degree days were used as measurement parameters of curing progress of soil hardening to replace the maturity numbers.

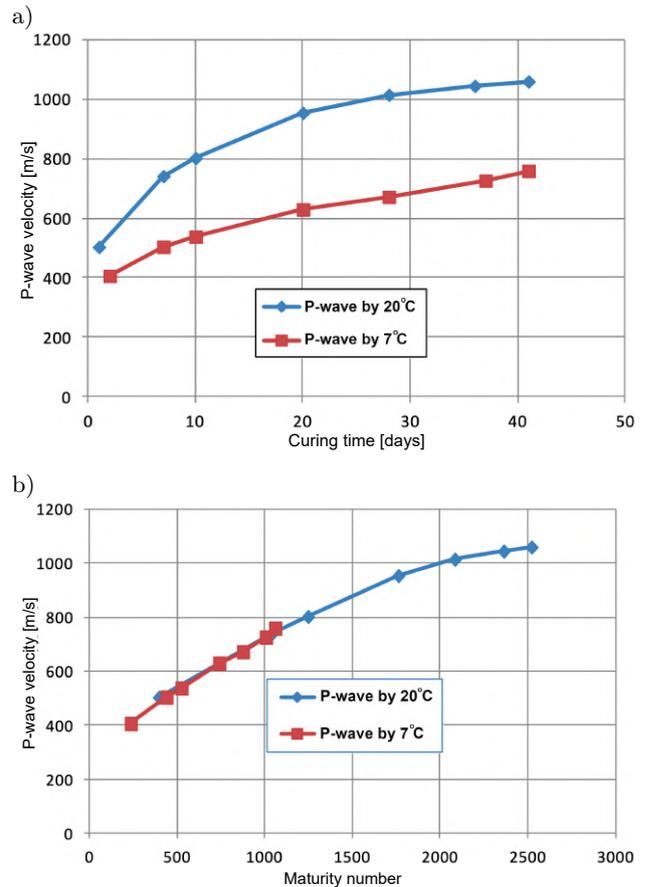


Fig. 13. P-wave rate as a function of: a) the curing time; b) the maturity number (M_T).

3.2. CPT soundings

The results from the CPT soundings are reported in Figs. 17–22. In total, 7 different lime/cement columns with a varying curing age between 9 days and 49 days were examined. The data for the investigated columns are summarised in Table 1. Here the drilling below the column refers to the penetration level of the tool, which is 0.8 meters below the stabilised part. This is

Table 1. Compilation of data from probed columns.

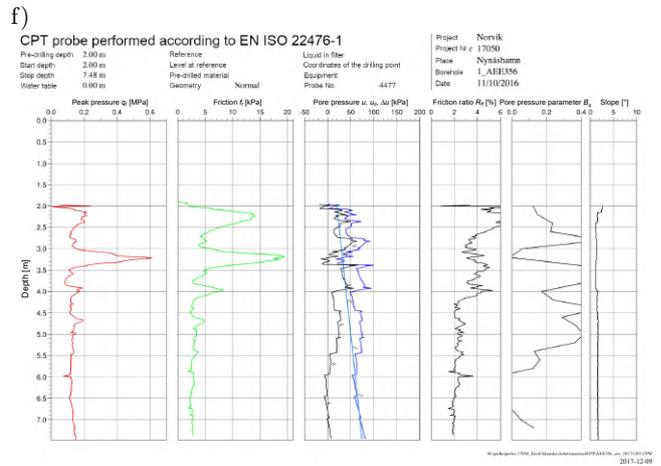
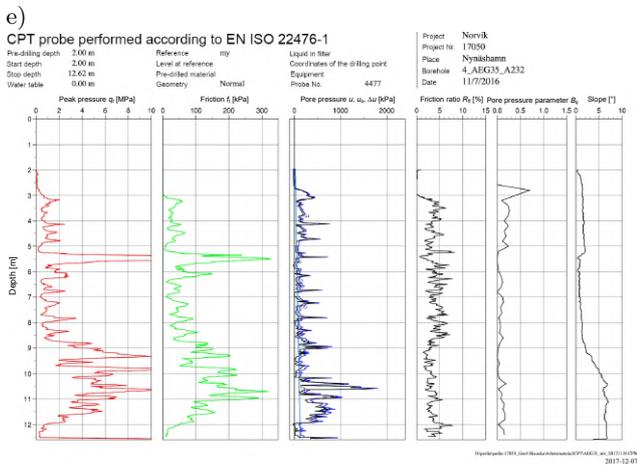
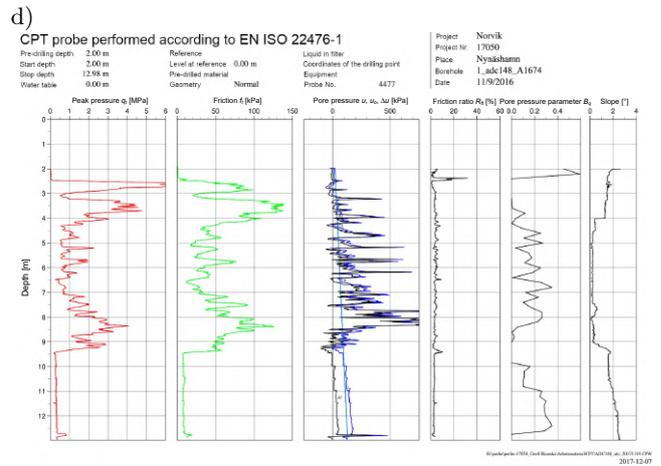
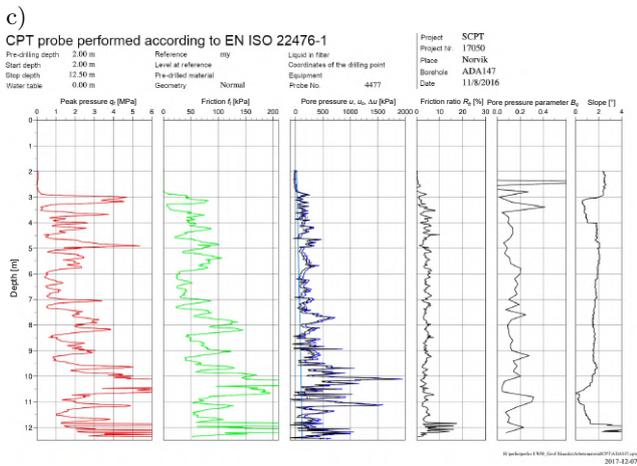
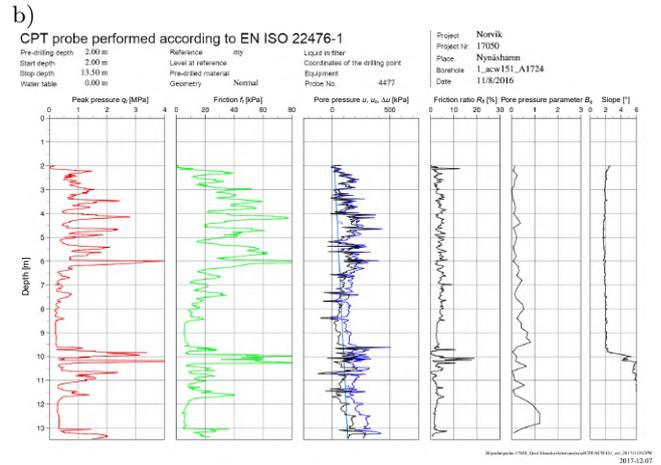
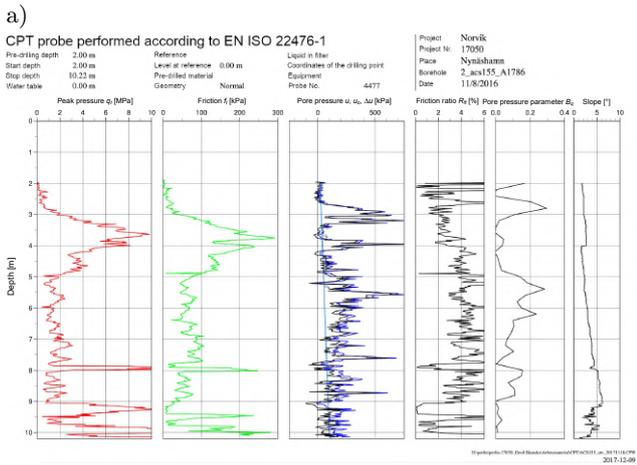
Column ID	Top level of column	Bottom level of column	Stabilized column length for binder LC/SC 50 kg	Drilling below column [m]	Average binder content [kg/m]	Lifting speed [mm/revolution]	Production date	SCPT date	Curing time [days]
ACS155	-0.9	-25.1	24.2	0.8	51.0	20	2016.10.12	2016.11.08	26
ACW151	-0.7	-24.2	23.5	0.8	50.3	20	2016.10.11	2016.11.08	28
ADA147	-0.7	-24.2	23.5	0.8	50.2	20	2016.10.10	2016.11.08	28
ADC148	-0.6	-24.2	23.6	0.8	50.0	20	2016.10.10	2016.11.09	29
AEA358	0.1	-12.0	12.1	0.8	50.5	20	2016.10.31	2016.11.09	9
AEG35	-0.5	-11.5	11.0	0.8	51.1	20	2016.09.19	2016.11.07	49
AEE356	-	-	-	-	-	-	-	2016.11.10	-

due to the fact that the mixing tool releases the binder above the bottom of the tool and normally it is not as much as 0.8 meters. The rise indicates the lifting, or the ascend speed of the tool for each revolution while drilling cycle.

Since there are significant uncertainties related to the evaluation of shear strength of the stabilized soil based on the peak pressure or the pore pressure evaluated from the CPT sounding, this study includes the comparison of the standardised values between the

peak pressure q_c and standardised results from seismic measurements. The normalisation was carried out by dividing the results by the largest value. Seismic measurements were recorded in the three lines, of which x - and y -lines represent shear wave measurements and z -line represent the compression wave measurements. The evaluations are reported in Figs. 17–22.

The measured parameters for the lime/cement columns are reported in Fig. 14. The results show a large variation and a very low tip resistance at the depth



[Fig. 14a-f]

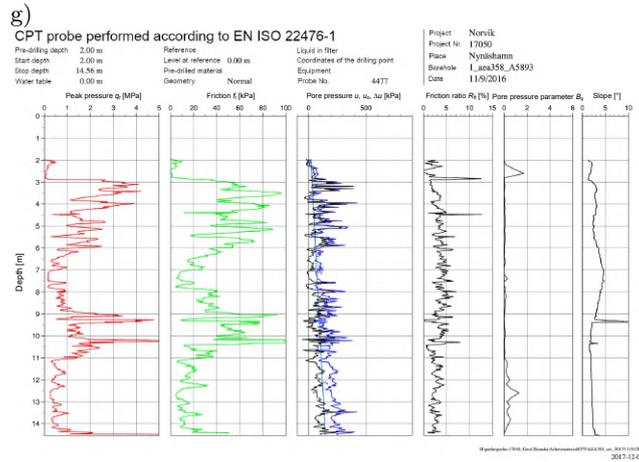


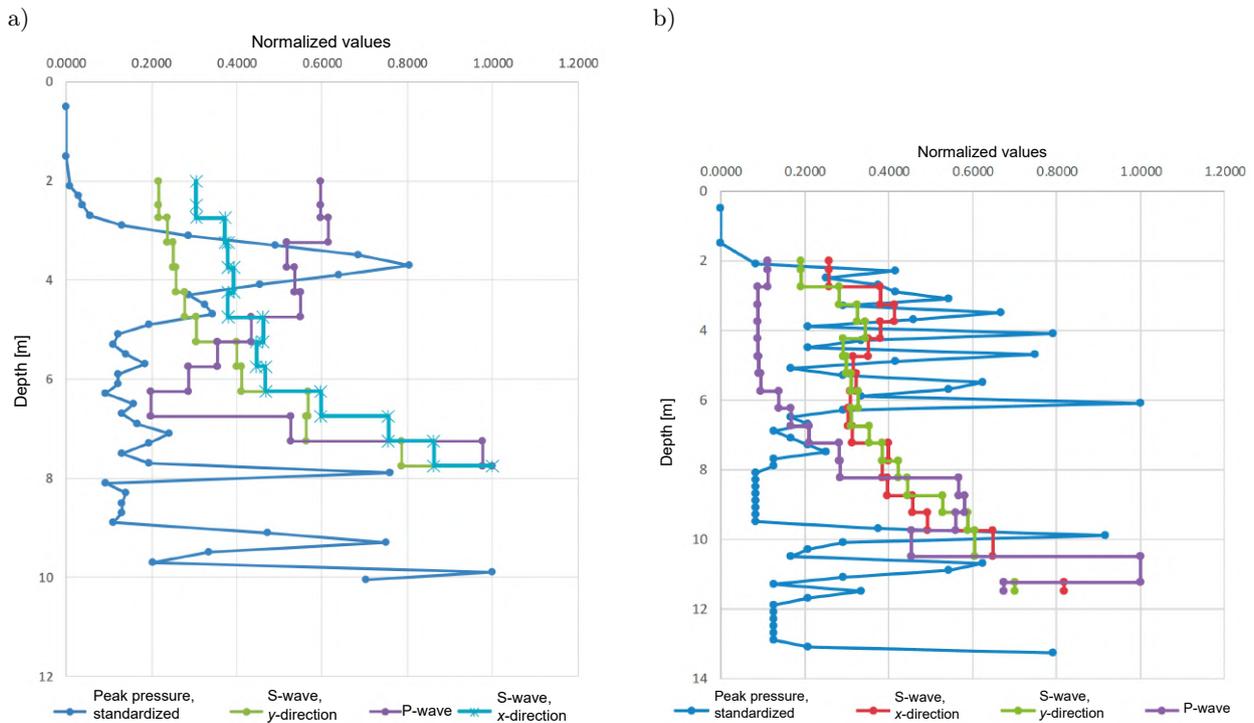
Fig. 14. Evaluated results from the CPT probing of the lime/cement columns (start and end depths of the seven test columns): a) ACS155 (2.00–10.22 m); b) ACW151 (2.00–13.50 m); c) ADA147 (2.00–12.50 m); d) ADC148 (2.00–12.98 m); e) AEG35 (2.00–12.62 m); f) AEE356 (2.00–7.48 m); g) AEA358 (2.00–14.56 m).

between 2 and 2.5 m below the ground surface, after which the tip pressure increases up to about 10 MPa at the level of 3.5 m. Between 5 and 8 m, the peak pressure reaches values of 1–3 MPa; at 8 m of depth it rises steeply to over 10 MPa. The sounding was interrupted at the depth of 10 m due to the stop caused by a high tip pressure. The measured friction shows the consistent results, although with some delay at the end of the probe. The results from seismic measurements show a good agreement between the shear waves in the *x*- and *y*-directions.

However, the high peak pressure around the level of 3.5 m is not achieved. The compression wave shows

a better agreement with the peak pressure and the evaluated P-wave velocities for various columns, see Fig. 15. The P-wave velocity down to the level of 7 m is within the compression wave velocities measured in the laboratory, see Fig. 17a and Fig. 12b. The values are, however, quite high, especially considering that the curing age of the column at the time of probing was only 26 days.

The measured shear wave speed for the columns is reported in Figs. 17 to 22 for various directions. The graph shows the results from the evaluated shear wave excited from both right and left in *x*-direction, the means of signals. The results show generally low



[Fig. 15ab]

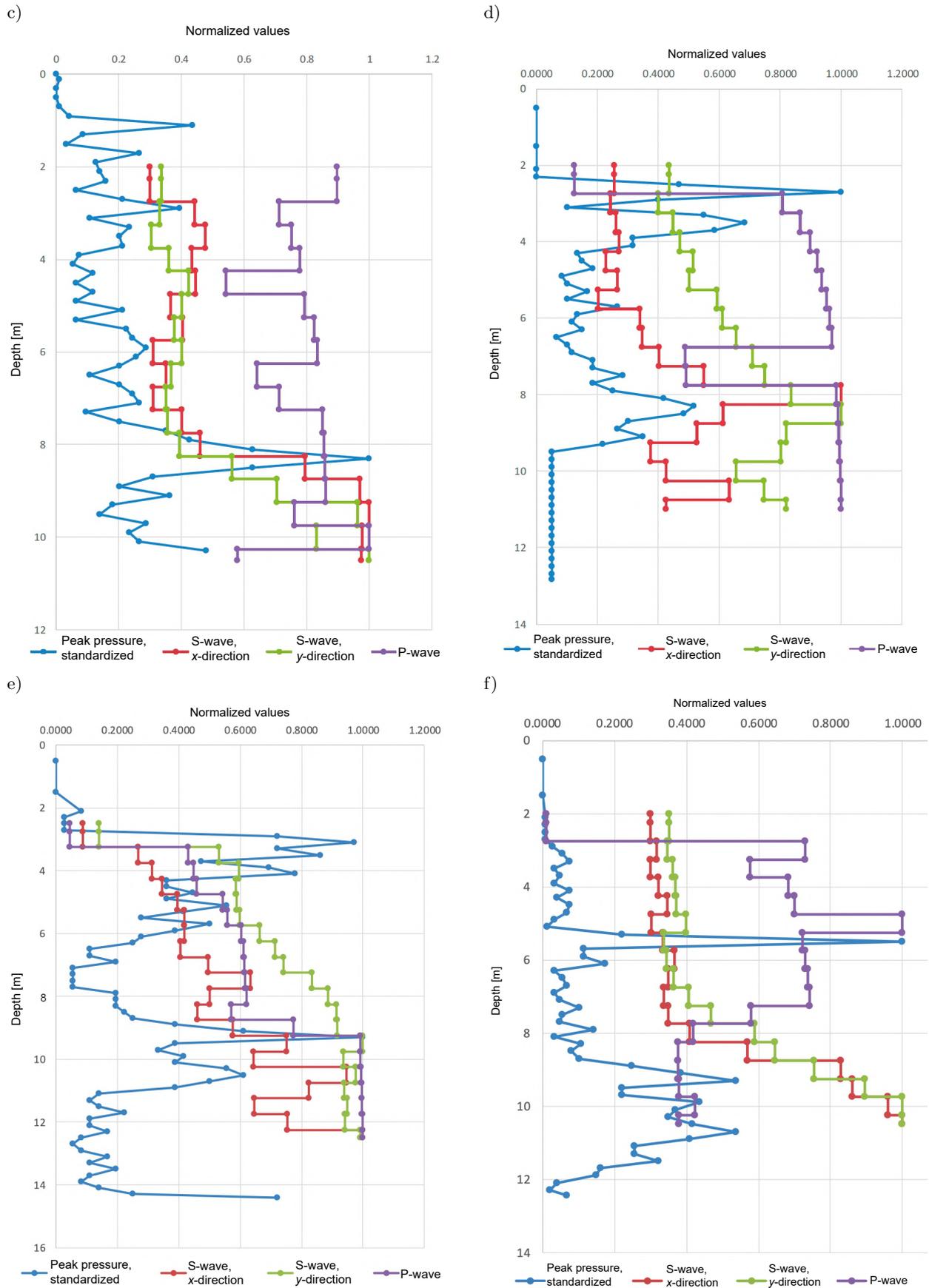


Fig. 15. Normalized peak pressure, S-wave and P-wave for the lime/cement columns: a) ACS155; b) ACW151; c) ADA147; d) ADC148; e) AEA358; f) AEG35

Table 2. Table showing evaluated results from seismic CPT for column ACS155.

Depth [m]	P-wave [m/s]	Ext. c_u from P-wave [kPa]	S-wave, x -dir. [m/s]	S-wave, y -dir. [m/s]	S-wave mean [m/s]	Ext. c_u from S-wave [kPa]
2.25	852	406	98.94	105.27	102.11	15
2.75	877	431	121.06	115.52	118.29	20
3.25	739	303	123.13	122.65	122.89	22
3.75	766	326	127.60	125.87	126.74	23
4.25	785	343	123.17	135.54	129.36	24
4.75	620	212	149.97	158.20	149.09	32
5.25	507	140	144.83	194.82	169.83	41
5.75	410	91	152.01	200.14	176.08	44
6.25	284	43	194.30	276.78	235.54	77
6.75	750	313	245.61	274.35	259.98	93
7.25	1395	1115	280.15	383.49	331.82	149
7.75	1427	1169	325.17	488.08	406.63	220

shear wave velocities. Table 2 shows P- and S-wave velocities for column ACS155. The table also contains calculated shear strength values based on the correlation determined in the laboratory. It is clear that the soil strength evaluated using P-wave measurements is relatively high with values above the actual strength, while strength evaluated using S-wave measurements shows values clearly below the expected results. Down to a depth of 6 m, the evaluated strength values are only slightly above the initial strength values of natural clay. Figure 14 shows the results of testing various columns. These results in the column ACS155 (Fig. 14a) indicate a significant difference compared to a probing in natural clay, cf., Fig. 14f in the column AEE356. The results from the probing of column ACS155 indicate that based on the seismic CPT sounding, the soil has been reinforced with a very solid part around 3.5 m below the ground surface.

The evaluated CPT sounding for the column ACW151 is reported in Fig. 14b. The peak pressure varies from <0.5 MPa to >4 MPa. This may appear to be low for a column that was cured during 28 days at the time of probing. For the column ACW151, the standardised values between the peak pressure, P- and S-wave show rather poor agreement, see Fig. 15b. The evaluated S-wave here gives a fairly good agreement between the x - and y -directions, but does not follow the results from the peak pressure. The highest evaluated P-wave velocity is almost 5000 m/s, which is in parity with the velocities in steel columns. The probing results from the CPT of the column ADA147 are reported in Fig. 14c. Here the tip pressure varies from about 0.5 MPa up to over 6 MPa. This shows a very large variation in strength and indicates difficulties in assessing the strength of a column based on the CPT. Figure 15c shows the standardised values for

the column ADA147. The shear wave velocities show low strength values, Fig. 21. Seismic testing gives bulk modulus values which do not have large variations in the peak pressure but rather an average value. The correlation between the peak pressure and shear wave velocities is low. The evaluated compression wave speed shows values above 2000 m/s.

The column ADC148 also shows large variations in the measured tip pressure. The CPT appears to have exited the column about 9 m below the ground surface, Fig. 14d. This column shows an abnormally large difference in the shear wave between the values measured in x -direction and y -direction. Neither the S- nor P-wave velocities show a good agreement with the tip resistance, Fig. 15d. The P-wave velocities here vary between 1000 and 2000 m/s which does not reflect the expected strength. The sounding result from the column AEA358 shows a similar pattern to the other columns, Fig. 14g. The peak pressure here shows very low values at 7.5 m of depth. The evaluated shear wave velocities here give the credible values of shear strength, c.m.f., Fig. 20, but does not reflect the variation in the peak pressure, Fig. 15e. The test results for column AEG35 are reported in Fig. 14e. This sounding also shows large variations in the tip pressure, which makes it difficult to find a connection between the CPT sounding results and the shear wave velocities, Fig. 15f.

4. Conclusion

The laboratory testing showed a high repeatability between the trials of seismic measurements. The variation in values of the compressive strength and shear strength slightly increased during the period of soil stabilization and increased strength of soil. This

is explained by minor defects and inaccuracies in the test specimens which have a greater effect on strength and result in higher resistance. The results from the field tests show large variations in the tip resistance, but also large variations in shear and compression wave velocities. These results were partly expected but the hypothesis was to find the same trends in the peak pressure and in seismic results. At the seismic CPT, there are several uncertainty factors regarding the seismic measurements which will be investigated in further steps of research. These include the question of how the generated S- and P-waves are transmitted to the column and whether there are cross signals originating from other sources of vibration. Further, the issues include the investigation of how well shear and compression waves are generated at greater depths, and to precise the position of probe within the column.

Seismic CPT has many advantages over the traditional CPT. To ensure that the CPT probe is in the column, a resistivity CPT can be connected there because the stabilized soil has a much lower resistivity due to free ions. One of the great advantages of seismic CPT is that it enables to measure the compressive strength and to apply seismic measurements as determination of the P- and S-wave velocities. Another important parameter for seismic CPT is the financial costs of the works and technical equipment. Here, both a drilling rig and a special expensive equipment are required. This results in the fact that this method is currently quite time-consuming and financially costly. By using a better signal source in seismic methods, some of the disadvantages of the methods can be eliminated. Therefore, this methodology can be a way to evaluate the effects of the lime/cement columns but cannot compete in terms of cost with the limestone column probes.

Acknowledgments

We are grateful to Helen Åhnberg and Rolf Larsson for project idea development, Roger Wisén (Ramböll) and Prof. Dr. Nils Rydén (LTH/Peab) for their support with data modelling and simulations, and Torbjörn Edstam, Skanska for the review and valuable comments. We would like to thank the anonymous reviewers and a technical editor who provided careful reading and constructive commenting of this work.

Funding

The Port of Stockholm Management, Nordic Construction Company (NCC) and GeoMind KB Engineering Consultancy are acknowledged for the assistance with data collection and the survey support. The Swedish Transport Administration, Skanska Group AB and the Development Fund of the Swedish Construction Industry (SBUF) provided a financial contribution

to this project. The authors express gratitude to the Swedish Maritime Administration (SMA) and Cemex AB for provided adhesives and technical support of this research project.

References

1. ABBISS C.P. (1981), Shear wave measurements of the elasticity of the ground, *Géotechnique*, **31**(1): 91–104, doi: [10.1680/geot.1981.31.1.91](https://doi.org/10.1680/geot.1981.31.1.91).
2. ÅHNBERG H. (2003), Measured permeabilities in stabilized Swedish soils, [in:] *Third International Conference on Grouting and Ground Treatment*, pp. 622–633, doi: [10.1061/40663\(2003\)34](https://doi.org/10.1061/40663(2003)34).
3. ÅHNBERG H., HOLM G. (1987), On the influence of the curing temperature on the shear strength of lime- and cement-stabilized soil [in Swedish: Om inverkan av härdningstemperaturen påskjuvhållfastheten hos kalk- och cementstabiliserad jord], *Swedish Geotechnical Institute*.
4. ALVARADO G., COOP M. (2012), On the performance of bender elements in triaxial tests, *Géotechnique*, **62**(1): 1–17, doi: [10.1680/geot.7.00086](https://doi.org/10.1680/geot.7.00086).
5. AMARAL M.F., DA FONSECA A.V., ARROYO M., CASCANTE G., CARVALHO J. (2011), Compression and shear wave propagation in cemented-sand specimens, *Géotechnique Letters*, **1**(3): 79–84, doi: [10.1680/geolett.11.00032](https://doi.org/10.1680/geolett.11.00032).
6. ARCHEEWA E., PUPPALA A.J., SARIDE S., KALLA S. (2011), Numerical model studies of deep soil mixing (DSM) column to mitigate bridge approach settlements, [in:] *Geo-Frontiers Congress 2011*, pp. 3286–3295.
7. AVCI E., MOLLAMAHMUTOĞLU M. (2016), UCS Properties of superfine cement-grouted sand, *Journal of Materials in Civil Engineering*, **28**(12): 06016015, doi: [10.1061/\(ASCE\)MT.1943-5533.0001659](https://doi.org/10.1061/(ASCE)MT.1943-5533.0001659).
8. BACHE B.K.F., WIERSHOLM P., PANIAGUA P., EMDAL A. (2022), Effect of temperature on the strength of lime-cement stabilized Norwegian clays, *Journal of Geotechnical and Geoenvironmental Engineering*, **148**(3): 04021198, doi: [10.1061/\(ASCE\)GT.1943-5606.0002699](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002699).
9. BENZ NAVARRETE M.A., BREUL P., GOURVÈS R. (2022), Application of wave equation theory to improve dynamic cone penetration test for shallow soil characterisation, *Journal of Rock Mechanics and Geotechnical Engineering*, **14**(1): 289–302, doi: [10.1016/j.jrmge.2021.07.004](https://doi.org/10.1016/j.jrmge.2021.07.004).
10. BERGADO D.T., JAMSAWANG P., TANCHAISAWAT T., LAI Y.P., LORENZO G.A. (2008), Performance of reinforced load transfer platforms for embankments supported by deep cement mixing piles, [in:] *GeoCongress 2008: Geosustainability and Geohazard Mitigation*, pp. 628–637.

11. BIN-SHAFIQUE S., RAHMAN K., AZFAR I. (2011), The effect of freezing-thawing cycles on performance of fly ash stabilized expansive soil subbases, [in:] *Advances in Geotechnical Engineering*, pp. 697–706.
12. BISHT V., SALGADO R., PREZZI M. (2021), Material point method for cone penetration in clays, *Journal of Geotechnical and Geoenvironmental Engineering*, **147**(12): 04021158, doi: [10.1061/\(ASCE\)GT.1943-5606.0002687](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002687).
13. BROUWER J.J.M. (2007), *In-situ Soil Testing*, IHS BRE Press.
14. CHAND S.K., SUBBARAO C. (2007), In-place stabilization of pond ash deposits by hydrated lime columns, *Journal of Geotechnical and Geoenvironmental Engineering*, **133**(12): 1609–1616, doi: [10.1061/\(ASCE\)1090-0241\(2007\)133:12\(1609\)](https://doi.org/10.1061/(ASCE)1090-0241(2007)133:12(1609)).
15. CHEN Y., ZHOU Y., KONG G., CHEN L., CHEN G. (2021), Settlement and stress analysis of soil-cement column-reinforced foundation under an in situ stabilized layer, *International Journal of Geomechanics*, **21**(12): 06021032, doi: [10.1061/\(ASCE\)GM.1943-5622.0002205](https://doi.org/10.1061/(ASCE)GM.1943-5622.0002205).
16. COLOMBERO R., KONTOS S., FOTI S., POTTS D.M. (2015), Numerical modelling of wave attenuation through soil, [in:] *Geotechnical Engineering for Infrastructure and Development*, pp. 3959–3964.
17. EL-RAWI N.M., AWAD A.A. (1981), Permeability of lime stabilized soils, *Transportation Engineering Journal of ASCE*, **107**(1): 25–35, doi: [10.1061/TPEJAN.0000907](https://doi.org/10.1061/TPEJAN.0000907).
18. ELSWORTH D., LEE D.S., HRYCIW R., SHIN S. (2006), Pore pressure response following undrained uCPT sounding in a dilating soil, *Journal of Geotechnical and Geoenvironmental Engineering*, **132**(11): 1485–1495, doi: [10.1061/\(ASCE\)1090-0241\(2006\)132:11\(1485\)](https://doi.org/10.1061/(ASCE)1090-0241(2006)132:11(1485)).
19. ERIKSSON H. (2015), *Personal and written communication with Håkan Eriksson*, Unpublished discussions, Geomind.
20. FITZGERALD M., ELSWORTH D. (2012), Liquefaction resistance recovered from “on-the-fly” CPTu-measured pore pressures, [in:] *Processing of the Geotechnical Earthquake Engineering and Soil Dynamics Congress IV*, doi: [10.1061/40975\(318\)69](https://doi.org/10.1061/40975(318)69).
21. FOTI S., LAI C.G., LANCELLOTTA R. (2002), Porosity of fluid-saturated porous media from measured seismic wave velocities, *Géotechnique*, **52**(5): 359–373, doi: [10.1680/geot.2002.52.5.359](https://doi.org/10.1680/geot.2002.52.5.359).
22. FOTI S., LANCELLOTTA R. (2004), Soil porosity from seismic velocities, *Géotechnique*, **54**(8): 551–554, doi: [10.1680/geot.2004.54.8.551](https://doi.org/10.1680/geot.2004.54.8.551).
23. GARCIA-SUAREZ J., SEYLABI E., ASIMAKI D. (2021), Seismic harmonic response of inhomogeneous soil: Scaling analysis, *Géotechnique*, **71**(5): 392–405, doi: [10.1680/jgeot.19.P.042](https://doi.org/10.1680/jgeot.19.P.042).
24. GUAN Z., WANG Y., ZHAO T. (2022), Adaptive sampling strategy for characterizing spatial distribution of soil liquefaction potential using cone penetration test, *Journal of Rock Mechanics and Geotechnical Engineering*, **14**(4): 1221–1231, doi: [10.1016/j.jrmge.2022.01.011](https://doi.org/10.1016/j.jrmge.2022.01.011).
25. HEIDARIZADEH Y., LAJEVARDI S.H., SHARIFIPOUR M. (2021), Correlation between small-strain shear stiffness and compressive strength of clayey soils stabilized with cement and Nano-SiO₂, *International Journal of Geosynthetics and Ground Engineering*, **7**: 12, doi: [10.1007/s40891-021-00258-x](https://doi.org/10.1007/s40891-021-00258-x).
26. HEPTON P. (1989), Shear wave velocity measurements during penetration testing, [in:] *Penetration testing in the UK*, pp. 275–278, The Institution of Civil Engineers.
27. HEPTON P. (2015), Deep rotary cored boreholes in soils using wireline drilling, [in:] *Advances in Site Investigation Practice*, pp. 269–280, The Institution of Civil Engineers.
28. ITO T., MORI Y., ASADA A. (1994), Evaluation of resistance to liquefaction caused by earthquakes in sandy soil stabilized with quick-lime consolidated briquette piles, *Soils and Foundations*, **34**(1): 33–40, doi: [10.3208/sandf1972.34.33](https://doi.org/10.3208/sandf1972.34.33).
29. JAMIOLKOWSKI M., PRESTI D.C.F.L., MANASSERO M. (2003), Evaluation of relative density and shear strength of sands from CPT and DMT, [in:] *Symposium on Soil Behavior and Soft Ground Construction Honoring Charles C. “Chuck” Ladd*, pp. 201–238, doi: [10.1061/40659\(2003\)7](https://doi.org/10.1061/40659(2003)7).
30. JEFFERIES M. (2022), On the fundamental nature of the state parameter, *Géotechnique*, **72**(12): 1082–1091, doi: [10.1680/jgeot.20.P.228](https://doi.org/10.1680/jgeot.20.P.228).
31. JONES R. (1958), In-situ measurement of the dynamic properties of soil by vibration methods, *Géotechnique*, **8**(1): 1–21, doi: [10.1680/geot.1958.8.1.1](https://doi.org/10.1680/geot.1958.8.1.1).
32. KICHOU Z., MAVROULIDOU M., GUNN M.J. (2022), Investigation of the strength evolution of lime-treated London clay soil, [in:] *Proceedings of the Institution of Civil Engineers – Ground Improvement*, doi: [10.1680/jgrim.21.00053](https://doi.org/10.1680/jgrim.21.00053).
33. KIM S., GOPALAKRISHNAN K., CEYLAN H. (2012), Moisture susceptibility of subgrade soils stabilized by lignin-based renewable energy coproduct, *Journal of Transportation Engineering*, **138**(11): 1283–1290, doi: [10.1061/\(ASCE\)TE.1943-5436.0000097](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000097).
34. LAMBRECHTS J.R., GANSE M.A., LAYHEE C.A. (2012), Soil mixing to stabilize organic clay for I-95 widening, [in:] *Third International Conference on Grouting and Ground Treatment*, pp. 575–585.
35. LAPOINTE E., FANNIN J., WILSON B.W. (2012), Cement-treated soil: Variation of UCS with soil type, [in:] *Proceedings of the Fourth 502 International Confer-*

- ence on Grouting and Deep Mixing, pp. 512–521, doi: [10.1061/9780784412350.0037](https://doi.org/10.1061/9780784412350.0037).
36. LARSSON S. (2017), *Personal communication with Stefan Larsson*, Unpublished discussions, KTH Royal Institute of Technology.
 37. LINDH P., LEMENKOVA P. (2021a), Evaluation of different binder combinations of cement, slag and CKD for S/S treatment of TBT contaminated sediments, *Acta Mechanica et Automatica*, **15**: 236–248, doi: [10.2478/ama-2021-0030](https://doi.org/10.2478/ama-2021-0030).
 38. LINDH P., LEMENKOVA P. (2021b), Resonant frequency ultrasonic P-Waves for evaluating uniaxial compressive strength of the stabilized slag–cement sediments, *Nordic Concrete Research*, **65**: 39–62, doi: [10.2478/ncr-2021-0012](https://doi.org/10.2478/ncr-2021-0012).
 39. LINDH P., LEMENKOVA P. (2022a), Dynamics of strength gain in sandy soil stabilised with mixed binders evaluated by elastic P-Waves during compressive loading, *Materials*, **15**(21): 7798, doi: [10.3390/ma15217798](https://doi.org/10.3390/ma15217798).
 40. LINDH P., LEMENKOVA P. (2022b), Geochemical tests to study the effects of cement ratio on potassium and TBT leaching and the pH of the marine sediments from the Kattegat Strait, Port of Gothenburg, *Baltica*, **35**: 47–59, doi: [10.5200/baltica.2022.1.4](https://doi.org/10.5200/baltica.2022.1.4).
 41. LINDH P., LEMENKOVA P. (2022c), Impact of strength-enhancing admixtures on stabilization of expansive soil by addition of alternative binders, *Civil and Environmental Engineering*, **18**(2): 726–735, doi: [10.2478/cee-2022-0067](https://doi.org/10.2478/cee-2022-0067).
 42. LINDH P., LEMENKOVA P. (2022d), Seismic velocity of P-waves to evaluate strength of stabilized soil for Svenska Cellulosa Aktiebolaget Biorefinery Östrand AB, Timrå, *Bulletin of the Polish Academy of Sciences: Technical Sciences*, **70**(4): e141593, doi: [10.24425/bpasts.2022.141593](https://doi.org/10.24425/bpasts.2022.141593).
 43. LINDH P., LEMENKOVA P. (2022e), Simplex lattice and X-ray diffraction for analysis of soil structure: A case of cement-stabilised compacted tills reinforced with steel slag and slaked lime, *Electronics*, **11**(22): 3726, doi: [10.3390/electronics11223726](https://doi.org/10.3390/electronics11223726).
 44. LINDH P., LEMENKOVA P. (2022f), Soil contamination from heavy metals and persistent organic pollutants (PAH, PCB and HCB) in the coastal area of Västernorrland, Sweden, *Gospodarka Surowcami Mineralnymi – Mineral Resources Management*, **38**(2): 147–168, doi: [10.24425/gsm.2022.141662](https://doi.org/10.24425/gsm.2022.141662).
 45. MADHYANNAPU R.S., PUPPALA A.J. (2014), Design and construction guidelines for deep soil mixing to stabilize expansive soils, *Journal of Geotechnical and Geoenvironmental Engineering*, **140**(9): 04014051, doi: [10.1061/\(ASCE\)GT.1943-5606.0001149](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001149).
 46. MADHYANNAPU R.S., PUPPALA A.J., BHADRIRAJU V., NAZARIAN S. (2009), Deep soil mixing (DSM) treatment of expansive soils, [in:] *Advances in Ground Improvement: Research to Practice in the United States and China*, pp. 130–139, doi: [10.1061/41025\(338\)14](https://doi.org/10.1061/41025(338)14).
 47. MADHYANNAPU R.S., PUPPALA A.J., NAZARIAN S., YUAN D. (2010), Quality assessment and quality control of deep soil mixing construction for stabilizing expansive subsoils, *Journal of Geotechnical and Geoenvironmental Engineering*, **136**(1): 119–128, doi: [10.1061/\(ASCE\)GT.1943-5606.0000188](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000188).
 48. MAKUSA G.P., MATSSON H., KNUTSSON S. (2014), Shear strength evaluation of preloaded stabilized dredged sediments using CPT, [in:] *Proceedings of the 3rd International Symposium on Cone Penetration Testing*, pp. 753–760, <http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-35248> (access: 20.12.2022).
 49. MCCALLUM A. (2014), A brief introduction to cone penetration testing (CPT) in frozen geomaterials, *Annals of Glaciology*, **55**(68): 7–14, doi: [10.3189/2014AoG68A005](https://doi.org/10.3189/2014AoG68A005).
 50. MIRZABABAEI M., ARULRAJAH A., HAQUE A., NIMBALKAR S., MOHAJERANI A. (2018), Effect of fiber reinforcement on shear strength and void ratio of soft clay, *Geosynthetics International*, **25**(4): 471–480, doi: [10.1680/jgein.18.00023](https://doi.org/10.1680/jgein.18.00023).
 51. MONTGOMERY D.C. (1996), *Design and Analysis of Experiments*, 7th ed., Wiley.
 52. MYERS R.H., MONTGOMERY D.C. (1995), *Response Surface Methodology*, Wiley.
 53. OLSEN R.S. (2018), CPT prediction of liquefaction resistance using the CPT soil characterization curve lookup technique, [in:] *Geotechnical Earthquake Engineering and Soil Dynamics V*, pp. 170–181.
 54. ORAKOGLU M.E., LIU J., LIN R., TIAN Y. (2017), Performance of clay soil reinforced with fly ash and lignin fiber subjected to freeze-thaw cycles, *Journal of Cold Regions Engineering*, **31**(4): 04017013, doi: [10.1061/\(ASCE\)CR.1943-5495.0000139](https://doi.org/10.1061/(ASCE)CR.1943-5495.0000139).
 55. PRICE A.B., DEJONG J.T., BOULANGER R.W., PARRA BASTIDAS A.M., MOUG D. (2016), Effect of prior strain history on the cyclic strength and CPT penetration resistance of silica silt, [in:] *Proceedings of the Geotechnical and Structural Engineering Congress 2016*, pp. 1664–1674.
 56. RUSATI P.K., SONG K.-I., YOON Y.-W., HWANG W., LIU L. (2020), Electrical resistivity and elastic wave velocity of sand–cement–inorganic binder mixture, *Environmental Geotechnics*, **7**(5): 318–329, doi: [10.1680/jenge.17.00082](https://doi.org/10.1680/jenge.17.00082).
 57. SAFAEE A.M., MAHBOUBI A., NOORZAD A. (2022), Seismic behavior of tiered geogrid reinforced soil (GRS) using treated backfill soil, *Geosynthetics International*, **30**(2): 200–224, doi: [10.1680/jgein.21.00060](https://doi.org/10.1680/jgein.21.00060).
 58. SANTAMARINA C., CASCANTE G. (1998), Effect of surface roughness on wave propagation parameters, *Géo-*

- technique, **48**(1): 129–136, doi: [10.1680/geot.1998.48.1.129](https://doi.org/10.1680/geot.1998.48.1.129).
59. SAYE S.R., OLSON S.M., FRANKE K.W. (2021), Common-origin approach to assess level-ground liquefaction susceptibility and triggering in CPT-compatible soils using Q , *Journal of Geotechnical and Geoenvironmental Engineering*, **147**(7): 04021046, doi: [10.1061/\(ASCE\)GT.1943-5606.0002515](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002515).
60. SHAHIEN M.M., FAROUK A. (2013), Estimation of deformation modulus of gravelly soils using dynamic cone penetration tests, *Ain Shams Engineering Journal*, **4**(4): 633–640, doi: [10.1016/j.asej.2013.01.008](https://doi.org/10.1016/j.asej.2013.01.008).
61. SHIHATA S.A., BAGHDADI Z.A. (2001), Simplified method to assess freeze-thaw durability of soil cement, *Journal of Materials in Civil Engineering*, **13**(4): 243–247, doi: [10.1061/\(ASCE\)0899-1561\(2001\)13:4\(243\)](https://doi.org/10.1061/(ASCE)0899-1561(2001)13:4(243)).
62. SONKAR C., MITTAL A. (2022), Evaluation of axial strength of sheathed cold-formed steel wall panels using Rayleigh-Ritz method and direct strength method: Comparative study, *Practice Periodical on Structural Design and Construction*, **27**(1): 04021064, doi: [10.1061/\(ASCE\)SC.1943-5576.0000623](https://doi.org/10.1061/(ASCE)SC.1943-5576.0000623).
63. SULLY J.P., CAMPANELLA R.G. (1991), Effect of lateral stress on CPT penetration pore pressures, *Journal of Geotechnical Engineering*, **117**(7): 1082–1088, doi: [10.1061/\(ASCE\)0733-9410\(1991\)117:7\(1082\)](https://doi.org/10.1061/(ASCE)0733-9410(1991)117:7(1082)).
64. SUNDARY D., MUNIRWAN R.P., AL-HUDA N., MUNIRWANSYAH, SUNGKAR M., JAYA R.P. (2022), Shear strength performance of dredged sediment soil stabilized with lime, *Physics and Chemistry of the Earth, Parts A/B/C*, **128**: 103299, doi: [10.1016/j.pce.2022.103299](https://doi.org/10.1016/j.pce.2022.103299).
65. TONINI DE ARAÚJO M., TONATTO FERRAZZO S., MANSUR CHAVES H., GRAVINA DA ROCHA C., CESAR CONSOLI N. (2023), Mechanical behavior, mineralogy, and microstructure of alkali-activated wastes-based binder for a clayey soil stabilization, *Construction and Building Materials*, **362**: 129757, doi: [10.1016/j.conbuildmat.2022.129757](https://doi.org/10.1016/j.conbuildmat.2022.129757).
66. TRHLÍKOVÁ J., MAŠÍN D., BOHÁČ J. (2012), Small-strain behaviour of cemented soils, *Géotechnique*, **62**(10): 943–947, doi: [10.1680/geot.9.P.100](https://doi.org/10.1680/geot.9.P.100).
67. VARMA M., MAJI V.B., BOOMINATHAN A. (2022), Seismic assessment of shotcrete support in jointed rock tunnels, *International Journal of Geosynthetics and Ground Engineering*, **8**: 51, doi: [10.1007/s40891-022-00392-0](https://doi.org/10.1007/s40891-022-00392-0).
68. VERÁSTEGUI-FLORES R.D., DI EMIDIO G., BEZUIJEN A., VANWALLEGHEM J., KERSEMANS M. (2015), Evaluation of the free-free resonant frequency method to determine stiffness moduli of cement-treated soil, *Soils and Foundations*, **55**(5): 943–950, doi: [10.1016/j.sandf.2015.09.001](https://doi.org/10.1016/j.sandf.2015.09.001).
69. XU B., YI Y. (2021), Soft clay stabilization using three industry byproducts, *Journal of Materials in Civil Engineering*, **33**(5): 06021002, doi: [10.1061/\(ASCE\)MT.1943-5533.0003710](https://doi.org/10.1061/(ASCE)MT.1943-5533.0003710).
70. YAPAGE N.N.S., LIYANAPATHIRANA D.S., KELLY R.B., POULOS H.G., LEO C.J. (2014), Numerical modeling of an embankment over soft ground improved with deep cement mixed columns: Case history, *Journal of Geotechnical and Geoenvironmental Engineering*, **140**(11): 04014062, doi: [10.1061/\(ASCE\)GT.1943-5606.0001165](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001165).
71. ZHANG Y., DANIELS J.L., CETIN B., BAUCOM I.K. (2020), Effect of temperature on pH, conductivity, and strength of lime-stabilized soil, *Journal of Materials in Civil Engineering*, **32**(3): 04019380, doi: [10.1061/\(ASCE\)MT.1943-5533.0003062](https://doi.org/10.1061/(ASCE)MT.1943-5533.0003062).

Appendix A. An example of the protocol evaluated using Conrad v.3.1

CPT probing carried out according to EN ISO 22476-1

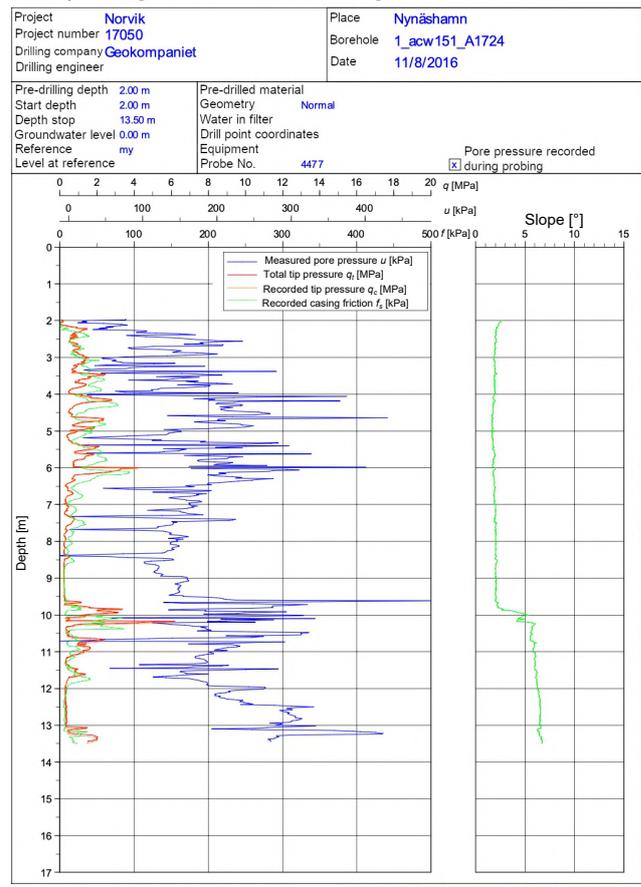
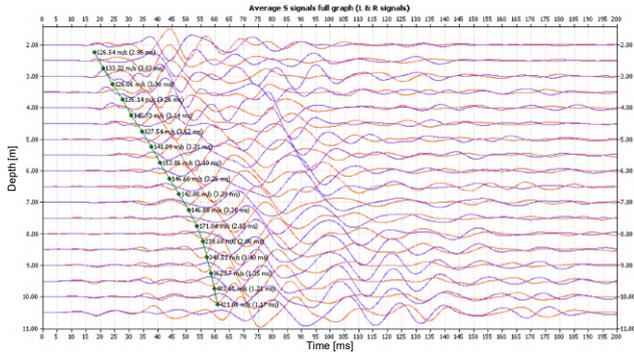


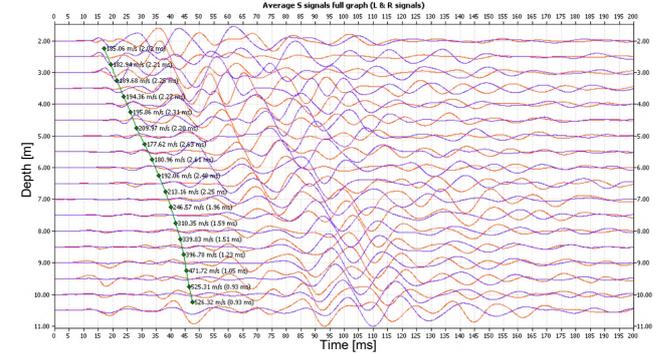
Fig. 16. Protocol with standard scales evaluated using Conrad software, version 3.1.

Appendix B. Supplementary material

a) Borehole: AEG35, *x*-direction – mean signal



b) Borehole: AEG35, *y*-direction – mean signal



c) Borehole: AEG35, *z*-direction – *p*-signal

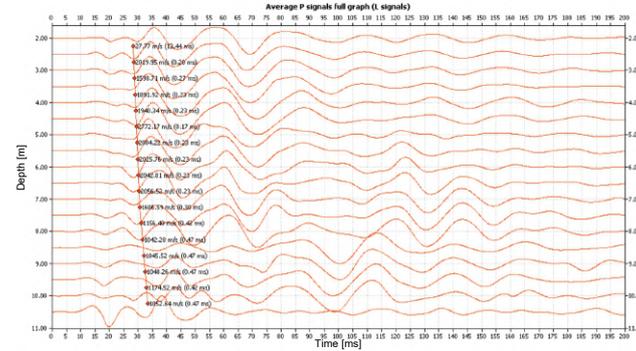
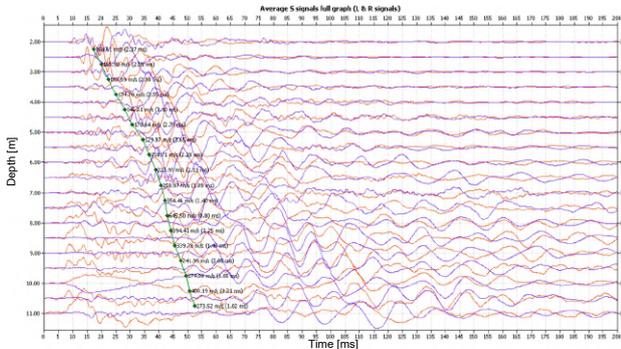
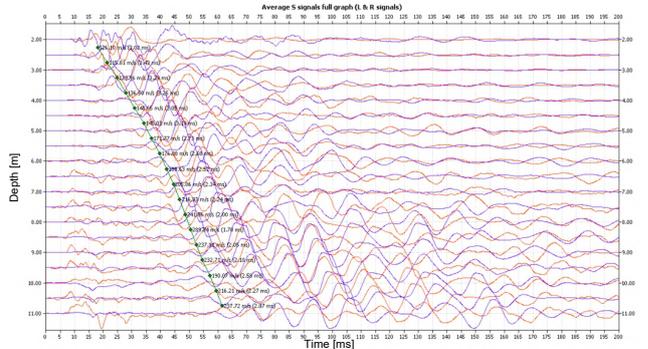


Fig. 17. Signals for AEG35 column.

a) Borehole: ADC148, *x*-direction – mean signal



b) Borehole: ADC148, *y*-direction – mean signal



c) Borehole: ADC148, *z*-direction – *p*-signal

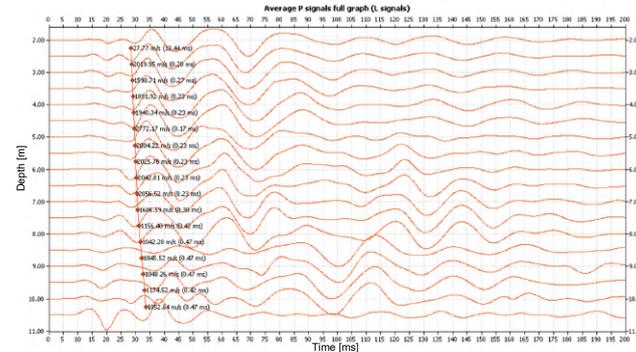
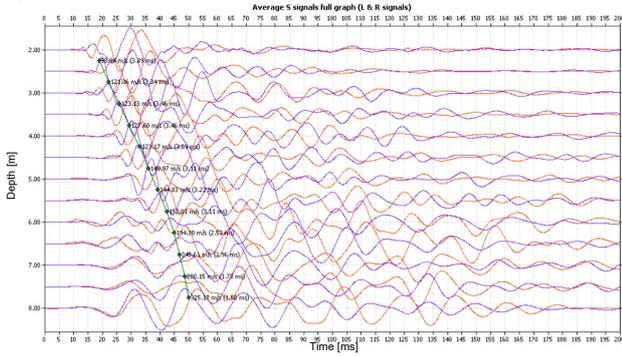
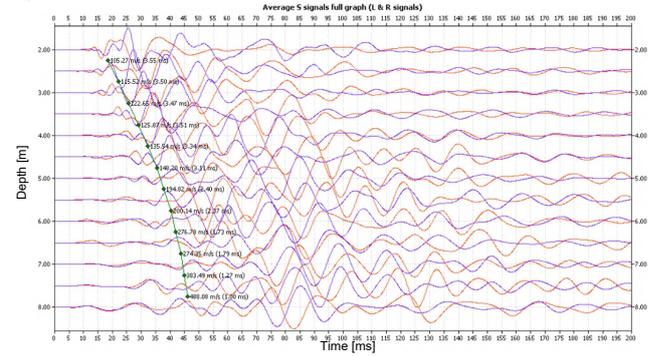


Fig. 18. Signals for ADC148 column.

a) Borehole: ACS155, *x*-direction – mean signal



b) Borehole: ACS155, *y*-direction – mean signal



c) Borehole: ACS155, *z*-direction – *p*-signal

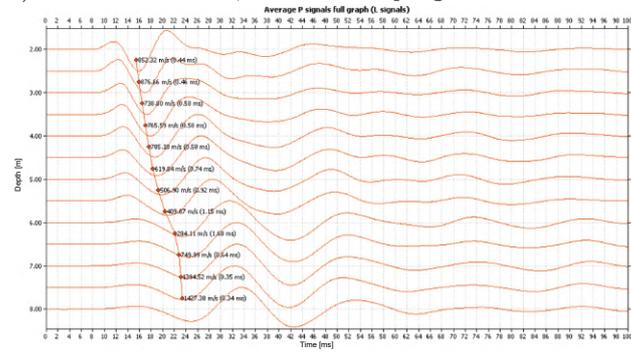
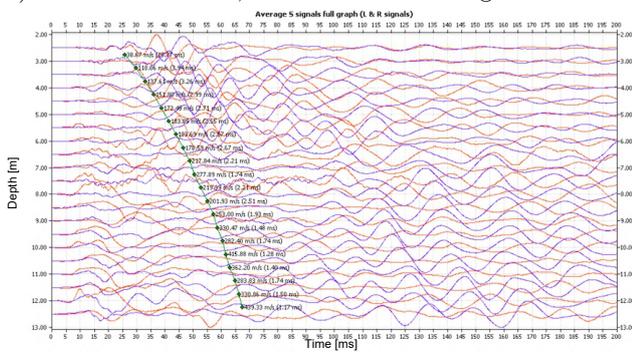
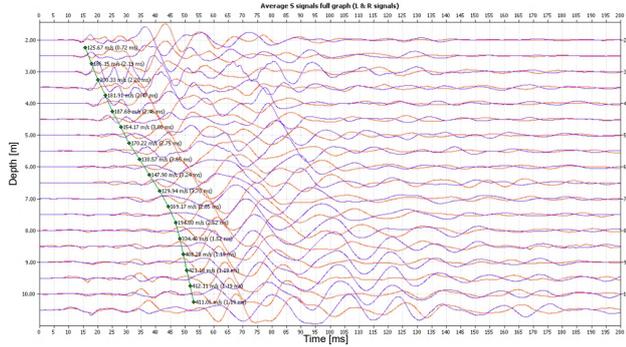


Fig. 19. Signals for ACS155 column.

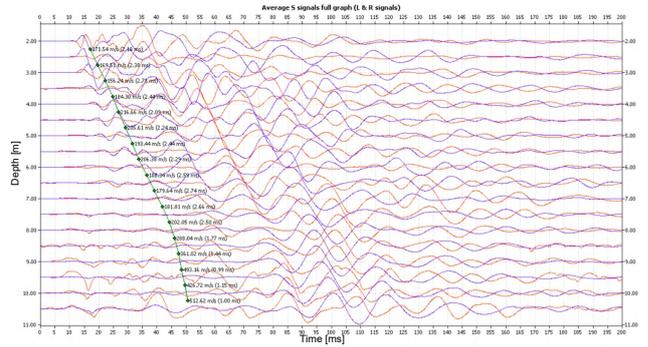
a) Borehole: AEA358, *x*-direction – mean signal



a) Borehole: ADA147, *x*-direction – mean signal



b) Borehole: ADA147, *y*-direction – mean signal



c) Borehole: ADA147, *z*-direction – *p*-signal

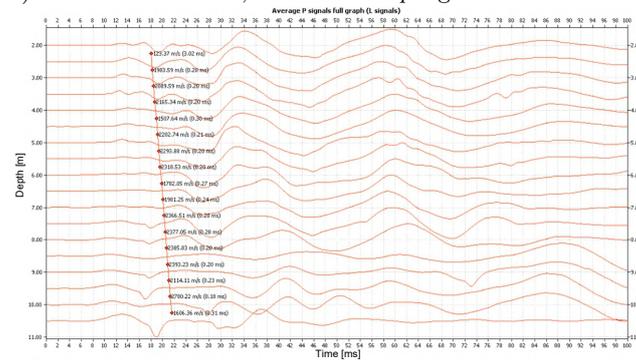
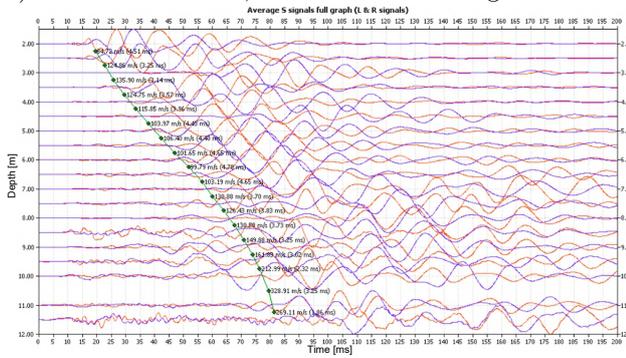
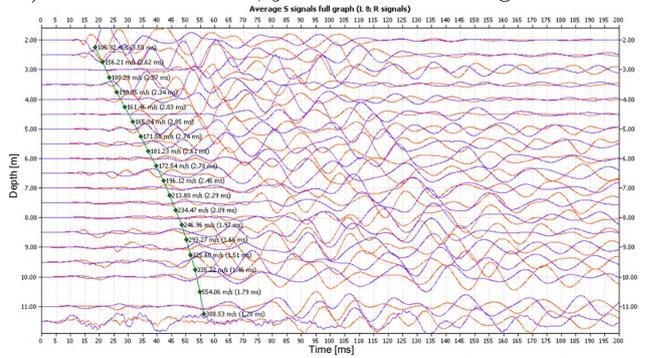


Fig. 21. Signals for ADA147 column.

a) Borehole: ACW151, *x*-direction – mean signal



b) Borehole: ACW151, *y*-direction – mean signal



c) Borehole: ACW151, *z*-direction – *p*-signal

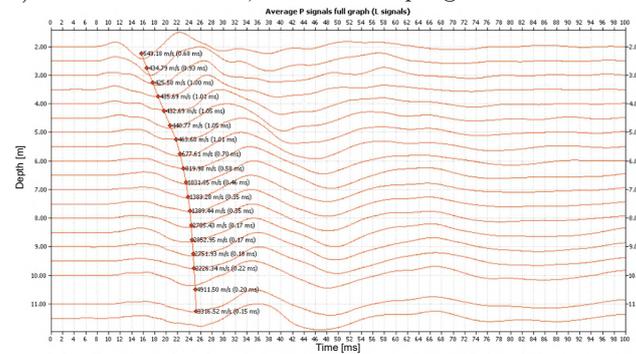


Fig. 22. Signals for ACW151 column.

Research Paper

Field Study on Underwater Noise Emitted by Small Tourist Boats. Comparison Between the Use of Electric and Combustion Motors

Alfio YORI

*Facultad de Ciencias de la Ingeniería
Instituto de Acústica
Universidad Austral de Chile
Valdivia, Chile; e-mail: ayori@uach.cl*

(received November 25, 2021; accepted January 26, 2023)

Small boats, possessing outboard engines, are widely used in tourism and mammal watching within marine protected areas. Noise generated by this type of vessels has the capacity to negatively affect marine fauna, especially marine mammals, which use sound throughout all the phases of their lives. These tourism boats used in mammal watching may use different propulsion systems, such as gas, diesel or electric engines. To characterize underwater noise emitted by this type of vessels becomes relevant not only when assessing the acoustic impact produced by these different propulsion systems over the marine fauna living inside these protected marine areas, but also when determining which one produces the least impact. A comparative study of underwater noise emissions coming from small touristic boats was made in this study. Boats were similar in capacity and functions, although possessing different propulsion systems. Measurements were made on two boats with a 50 Hp internal combustion engine and one 5 Hp electric boat. These boats were selected to be studied because they have practically the same size, possess the same passenger-capacity and are used to make similar jobs and routes inside the protected area where they are operated. The electric boat showed a considerable decrease in underwater noise emissions, especially in low frequencies. This boat will produce a lower accumulated exposition of the fauna to the noise or will allow a closer approach to the observed species. Measurements were made between September 2018 and January 2020.

Keywords: small boat underwater noise; anthropogenic underwater noise; electric motor boat; internal combustion engine boat; underwater noise produced by tourist boats.



Copyright © 2023 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

From all the different types of existing anthropogenic underwater noise, the most commonly found worldwide is that produced by vessels, which has the potential to negatively affect marine organisms. In Europe, efforts to monitor and manage acoustic contamination by vessels are included in the Marine Strategy Framework Directive, MSFD, and its descriptor 11.2 (Commission of the European Communities, 2008), which forces member states to guarantee that underwater noise levels will not exceed those thresholds that endanger the Good Environmental Status (GES) of EU waters (European Commission, 2017).

Underwater noise emitted by a vessel mainly depends on its size, on the engine power and on its speed

(GRELOWSKA *et al.*, 2013; SANTOS-DOMÍNGUEZ *et al.*, 2016; KLAUSON, MUSTONEN, 2017). Vessels possess different types of noise sources, where the relative acoustic intensity emitted by these noise sources depends on the type of vessel and their velocity (ABRAHAMSEN, 2012).

In Valdivia, a city located in the south of Chile, in South America, there is an estuary and a marine protected area inside this estuary, very much visited by tourists, mainly through small boats powered by internal combustion outboard engines. This type of boat is the most commonly used inside Chilean marine protected areas, where touristic activities of marine mammal watching are carried out. Boats powered by electric motors were included in this activity. The goal of introducing this type of motors was to use touristic

boats more friendly towards the environment, thus diminishing water contamination by wastes left by internal combustion engines and airborne noise emissions, which have almost completely disappeared with this type of motor. Regarding underwater noise emissions, the answer whether a boat with an electric motor is more silent than another similar boat with an internal combustion engine is not simple.

As small boats, with passenger capacity between 10 and 20 persons, are widely used in tourism and mammal watching within Chilean marine protected areas, to characterize the underwater noise produced by this type of boats is very important for the evaluation of the impact that the different types of vessels may produce over the marine fauna existing inside these marine protected areas.

There is little available information on the underwater noise level emitted by small boats as well as on the difference in the noise level emitted by the different types of engines used by small boats, such as electric motors and internal combustion engines.

Due to the above, it was established as an objective for this work to carry out a comparative study of underwater noise emissions, generated by small tourist boats with different propulsion systems. Two boats with an internal combustion engine and one with an electric motor were subjected to studies. Through the results of this noise emissions study, decisions on the type of vessel that should be used in marine mammal watching tourism in marine protected areas can be made.

These three boats were chosen for this study since they are considered as similar and appropriate to be compared. This is because they have almost the same size, bearing the same passengers capacity and are used to perform similar routes and duties inside the estuary. The Chilean Navy Port Authorities classify these three vessels as “small”. Worth remembering is that the efficiency of internal combustion engines is under 50%, whereas that of an electric engine is about 90% (EKDAHL ESPINOZA, 2014).

2. Underwater noise impact

Anthropogenic underwater noise has increased its presence in the oceans and coasts due to intensification in human activity (ANDREW *et al.*, 2022; McDONALD *et al.*, 2006; ROSS, 2005), thus increasing existing concern about its impact over marine life (TYACK, 2008; KUNC *et al.*, 2016).

Marine mammals have evolved exploiting water capacity to transmit sound waves through long distances and with low absorption. This efficiency is not the same for electromagnetic waves such as light, which are rapidly attenuated (GRELOWSKA, KOZACZKA, 2014; URICK, 1996; KINSLER *et al.*, 1999). This is why marine mammals, especially cetaceans, use mainly sound,

more than any other signal to perceive their surrounding and interact with it (TAVOLGA, 1965; 1971; HAWKINS, MYRBERG, 1983).

Marine mammals use principally their auditory system for sailing, feeding, distance calculation, communicating, hunting, socializing, etc., since this is their sense possessing a broader range (RICHARDSON *et al.*, 1995; TYACK, 2000).

This is why noise generated by human beings under the water has the capacity to negatively affect marine fauna, especially cetaceans, which use sound throughout their life phases. These negative effects can range from physiological stress, behavior changes, masking, temporal hearing loss, permanent hearing loss and death (MCCAULEY *et al.*, 2003; WYSOCKI *et al.*, 2006; WRIGHT *et al.*, 2007). A little longer than a decade ago, the first scientific guide about maximum limits of underwater noise exposition was published; it was targeted at avoiding the occurrence of hearing impairment, both temporal (TTS) and permanent (PTS) as well, in marine animals exposed to it (SOUTHALL *et al.*, 2007). Other new guides have been updating these exposition limits according to advances in research about noise impact on marine mammals (National Marine Fisheries Service [NMFS], 2016; 2018; ERBE *et al.*, 2016).

3. Noise sources in vessels

Noise sources in ships may be grouped into three categories: machinery noise, propeller noise and hydrodynamic noise (URICK, 1996).

Machinery noise is originated as a mechanical vibration of the different parts of the propulsion system and auxiliary machinery of a ship, which is transmitted to water through the hull. This type of noise is present in ships with rigid mount diesel engines and are originated in hydraulic systems, compressors, generators, shafts and gears, which transmit not only their vibration, but also their airborne noise transmitted to the environment through the ship hull. This type of rotary motion machine with a determined number of revolutions per minutes (rpm) produces a noise spectrum dominated by tonal components of the fundamental frequency and harmonics corresponding to rpm of the rotary movements of machines.

The main source of propeller noise is the cavitation noise induced by the propeller rotation. When the propeller rotates in the water, areas of low pressure are generated, both at the tip and on the surface of the propeller blade. If this negative pressure is low enough, a change in the water phase is produced, thus generating hundreds of tiny air bubbles, which rapidly implode or collapse after moving away and due to high increase in pressure, hence generating a strong impulsive noise.

Consisting of a big number of random implosions produced by the collapse of gas bubbles, the cavita-

tion noise possesses a uniform frequency spectrum. The spectrum (due to the propeller cavitation) has a negative slope in high frequencies of -6 dB/octave and a positive slope of 6 dB/octave in low frequencies. There is a peak in the spectrum located between 100 Hz and 1000 Hz for big vessels. The peak amplitude and its position in the frequency axis change with the vessel velocity. The spectrum amplitude increases as velocity increases; whereas the peak frequency decreases as velocity increases. There is a velocity from which the phenomenon of cavitation starts. The level of cavitation noise increases from this critical velocity.

Moreover, the passage of water through the propeller produces a propeller noise composed by tonal components of low frequency, which add to the uniform spectrum of the cavitation noise. These components coincide with the propeller turning ratio and are originated by the propeller resonance, excited by the vortices shedding from it.

Finally, hydrodynamic noise is produced by the irregularity and fluctuation of the flow along the hull as a result of the movement of the ship. Pressure fluctuations produced by this irregular flow are radiated as noise. This noise is also known as flow noise. Under normal circumstances of operation and design, hydrodynamic noise has little contribution to the total noise irradiated by vessels, since the most probable fact is that this noise is masked by the machinery noise and propeller noise.

The vessels in this study are small boats used for tourism and mammal watching. These boats have thin, narrow hulls with no protruding appendages from the hull, so their main source of underwater noise is the noise produced by the propeller.

4. Method

4.1. Measuring system

The system used to measure the noise emitted by the boats consists of the hydrophone Cetacean Research model C55/736 with a lineal response from 0.15 Hz to 44 kHz and the digital recorder Tascam DR680MKII possessing quantization rates, Q: $16/24$ bit and sampling frequencies, fs: $48/96/192$ kHz, which are used to record noise emissions. Afterwards, sound files are loaded to the software SpectraPLUS-SC 5.1D, where the required noise descriptors are obtained. A sampling frequency of 48 kHz was used in this study.

Due to the lack of facilities to carry out the calibration of the system under water, the system was calibrated inside an anechoic chamber following the steps indicated in the standard EN 60565 (2007) which allows calibration in air and in free field conditions, replacing the reference hydrophone with a calibrated microphone. A sensitivity of -167 dB re.1V/ μ Pa for the hydrophone and a sensitivity equal to -154.9 dB re.1V/ μ Pa for the

system hydrophone plus recorder were obtained with the standard. The software SpectraPLUS-SC5.1D was calibrated using this last sensitivity and a pure tone of 400 Hz.

The process was made inside the anechoic room of the Acoustic Department from Universidad Austral de Chile, whose dimensions are 2.45 m wide, 4.45 m long, and 3.8 m high, with a chamber cut-off frequency of 120 Hz and a wedge cutoff frequency equal to 170 Hz.

The complete system consists of a hydrophone, a buoy, a recorder and a kayak as means of transport. The hydrophone and its buoy are deployed from the kayak.

4.2. Vessels evaluated in the study

To carry out comparisons of underwater noise emissions produced by small touristic boats with different types of propulsion, two boats with internal combustion outboard engine and one boat with solar-powered electric inboard motor were used.

These types of boats were chosen because in the marine protected areas, where touristic watching activities are made, internal combustion outboard engine boats are the most widely used. The motivation to include boats with electric engines was to know how much more friendly they are with the environment, when contrasted with those using internal combustion engines.

4.2.1. Internal combustion outboard engine boats

Boats belonging to this type were identified as “combustion motor boat 1” and “combustion motor boat 2”. Figures 1 and 2 show these boats, respectively. Table 1 shows the characteristics of these vessels.



Fig. 1. Combustion motor boat 1.



Fig. 2. Combustion motor boat 2.

Table 1. Characteristics of evaluated boats.

Name	Length [m]	Motor	Power [Hp]	Passengers	Size*
Combustion motor boat 1	9.5	Internal combustion/Outboard	50	16	small
Combustion motor boat 2	8.2	Internal combustion/Outboard	50	12	small
Electric motor boat	9.5	Electric/Inboard	5	16	small

*According to classifications by the Chilean army.

These two boats are used for touristic rides inside the estuary of the Valdivia River. They are also used for touristic routes inside a marine protected area known as “Nature Sanctuary of Cruces River” (Ministerio del Medio Ambiente, 2015; Ministerio de Educación Pública, 1981). These two boats represent all touristic vessels of the estuary, listed as small by the Port Authority of Valdivia, belonging to the Chilean Navy.

4.2.2. Solar-powered electric motor boat

This type of vessel was introduced in 2012; the goal was to implement a sustainable river transport system in the rivers of the estuary. Contrasted with their competitors, these boats are recognized as more friendly with the environment, since they do not produce contamination by wastes, as do boats powered by internal combustion engines. The assessed boat of this type, identified as electric motor boat, has the same touristic route as that of combustion motor boats 1 and 2. Figure 3 shows the assessed electric boat and Table 1 shows its characteristics. We must remember that the efficiency of an internal combustion engine is between 35% and 45%, and that of an electric motor is between 85% and 95% (SALAS *et al.*, 2013).



Fig. 3. Evaluated electric motor boat.

The reason to include this electric motor boat in the study was to answer the question whether they are also more friendly with the environment, contrasted with those powered by internal combustion motors, when emitted underwater noise is considered.

4.3. Measurement procedures

Measurements of underwater emitted noise are carried out while the boat under evaluation passes in

front of the hydrophone at a defined distance. Measurements were made for three boat velocities. For the internal combustion engine boats, velocities were 1000, 3000, and 6000 rpm. For the electric motor boat, velocities were 1100, 3000, and 5500 rpm.

Figure 4 shows the area of the Valdivia River where evaluations are carried out. Figure 5 shows the boat track during measurement and the hydrophone position. This is shown only as a reference, since position and relative distance changed between measurements because the hydrophone is allowed to drift during the

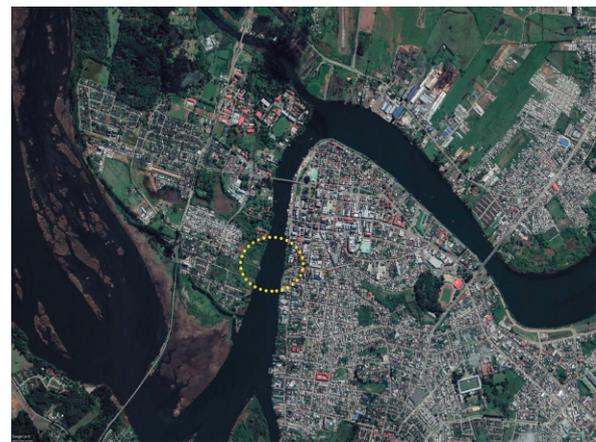


Fig. 4. Area of Valdivia River where the study was made.



Fig. 5. The orange line shows the boat track in front of the hydrophone and the yellow circle indicates the hydrophone position.

measurements, thus avoiding hydrodynamic noise of low frequency that may be produced by currents or tides if the microphone is stationary. Measurements were made during high tide, with atmospheric conditions without rain or wind and Beaufort wind force scale between 0 and 1 (Hydrographic and Oceanographic Service of the Chilean Navy, 2002). The river depth is from 8 to 10 meters in the area where measurements were made.

The hydrophone and its floating buoy are displayed from the kayak over which the recorder and the operator stay. The hydrophone is submerged as deep as 4 meters. Through a communication radio, instructions are given to the vessel to start the measurement. The distance at which the boat passes in front of the hydrophone is measured through an optical rangefinder. The boat follows a track of approximately 250 m long in front of the hydrophone (see Fig. 5). The 4 meters correspond to half of the depth to which the marine bottom is located. The underwater sound pressure depends on the depth due to the Lloyd's Mirror effect (CAREY, 2009). For a depth of 4 m this effect affects frequencies lower than 93 Hz (URICK, 1996; National Physical Laboratory, 2014). However, this is a comparative study, where all sound emissions evaluated will be affected by this effect.

From the audio file recorded during the passage of the boat, a time interval of 15 seconds centered on the time where the boat reached the shortest distance to the hydrophone is selected. This distance is called closest point of approach or CPA, and corresponds to the distance measured with the optical rangefinder (see Fig. 5) (ISO 17208, 2012; ANSI S12.64, 2009). In this work, the time where the CPA occurs is determined through the interference pattern produced due to the sound reflection over the marine bottom (see Fig. 6) (BJØRNØ, 2017; CAREY, 2009). From this audio file, three equivalent noise levels $L_{eq,T}$ of 5 seconds are obtained (see Eq. (1)). Central $L_{eq,5s}$ is chosen as the

underwater noise emitted by the boat during its passage in front of the hydrophone (see Figs. 5 and 6). Thus, this level will represent the worst condition on the receiver, since it only considers the highest emission interval. This method differs from what is indicated for standardized methods, where an rms level is measured for a time interval called data window period or DWP (ISO 17208, 2012, ANSI S12.64, 2009).

The descriptor equivalent noise level $L_{eq,T}$ of a fluctuating noise of the time interval T gives the continuous noise level possessing the same amount of energy as that of the fluctuating noise, in the same time interval T . The formula for $L_{eq,T}$ is:

$$L_{eq,T} = 10 \log \left[\frac{1}{T} \int_0^T \left(\frac{p(t)}{p_0} \right)^2 dt \right] \text{ [dB]}, \quad (1)$$

where $p(t)$ is the instant sound pressure in the considered time interval T , and p_0 is the reference pressure equal to 1 μPa . Likewise, noise level may be expressed as an RMS value through the expression:

$$L_p = 20 \log \left(\frac{p_{\text{RMS}}}{p_0} \right) \text{ [dB]}, \quad (2)$$

with

$$p_{\text{RMS}} = \sqrt{\frac{1}{T} \int_0^T (p(t))^2} \text{ [dB]}, \quad (3)$$

where Eq. (2) gives the same value as that given by Eq. (1) for an equal time T .

Another noise descriptor used in underwater acoustics is the sound exposure level SEL. The formula for SEL is:

$$\text{SEL} = 10 \log \left(\int_0^T \left(\frac{p(t)}{p_0} \right)^2 dt \right) \text{ [dB]}. \quad (4)$$

The SEL descriptor gives the total energy contained in the sound, where the energy of a sound is proportional to the time integral of the pressure squared. The descriptors L_{eq} and SEL are related through the following equation:

$$\text{SEL} = L_{eq,T} - 10 \log \left(\frac{1}{T} \right) \text{ [dB]}, \quad (5)$$

equation that considers the time interval T with which the L_{eq} level was obtained.

Similarly, if one has a sound with a constant amplitude level L_p over time, one can obtain the total energy level for an exposure time of duration t in seconds, through the equation:

$$\text{SEL} = L_p + 10 \log(t) \text{ [dB]}. \quad (6)$$

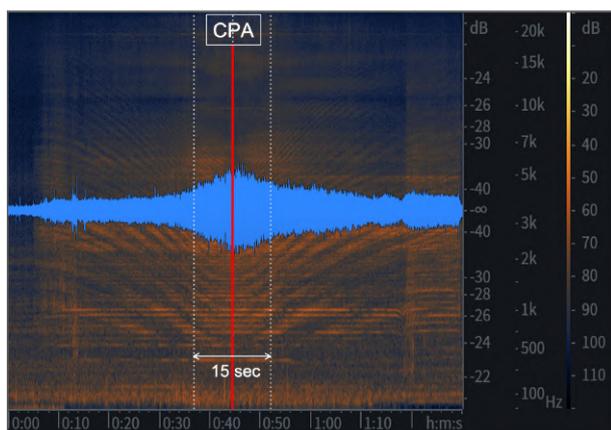


Fig. 6. Waveform and spectrogram. Example of a 15-second time interval centered on the point of the shortest distance from the boat to the hydrophone or CPA.

4.4. Correction by distance and source level

To obtain the source levels $L_{p,1m}$ of the evaluated boats, which is the noise level emitted by the source at a distance of 1 m, noise levels obtained during measurements must be corrected based on the distance they were made. There are different theoretical models for the propagation of underwater noise, where spherical propagation is assumed for large depths and cylindrical propagation is assumed in the presence of sound channels or certain conditions of shallow waters (RICHARDSON *et al.*, 1995). In shallow waters, the type of marine bottom may widely change the loss of sound propagation, where for example, if the bottom is too absorbent, even a behavior closer to spherical may be found (RICHARDSON *et al.*, 1995).

There are many environmental factors that influence the transmission of sound in shallow water, so developing adequate theoretical models is very complicated. Theory and empirical data are commonly combined to obtain reliable propagation predictions. In very shallow water, with sound wavelengths λ comparable to water depths, sound propagation can be analyzed using the mode theory. This theory indicates that if the effective depth of the water is less than $\lambda/4$, the wave will experience very large propagation losses. To accommodate the large variability observed in sound propagation in shallow water, semi-empirical propagation models have been designed for shallow water applications (RICHARDSON *et al.*, 1995).

The place of the river where this study was carried out presents a depth between 8 and 10 meters, with a mud bottom. According to the mode theory, this location has a low cut-off frequency for the sound propagation equal to 46 Hz. However, for cases of shallow water, it is common to use semi-empirical models of sound propagation. In field studies with comparative purposes, as is the case in this study, it is usual to use a sound propagation between spherical and cylindrical, as the attenuation shown by Eq. (7) (National Oceanic and Atmospheric Administration, 2020; California Department of Transportation, 2020; MCGARRY *et al.*, 2020; Servicio de Evaluación Ambiental, 2022). This is a recommended practical model, since it provides a pragmatic estimation of transmission losses (National Physical Laboratory, 2014; Marine Management Organisation, 2015). Due to the above, this model was used to obtain the source level in this work:

$$L_{p,1m} = L_{eq,T} - 15 \log\left(\frac{1}{d}\right) \text{ [dB]}, \quad (7)$$

where $L_{p,1m}$ is the source level and d is the distance to the source from where $L_{eq,T}$ was measured (CPA distance). The sound propagation model used in this study is correctly adjusted to the empirical results (see Figs. 7–10). The period time T used in this work was 5 seconds.

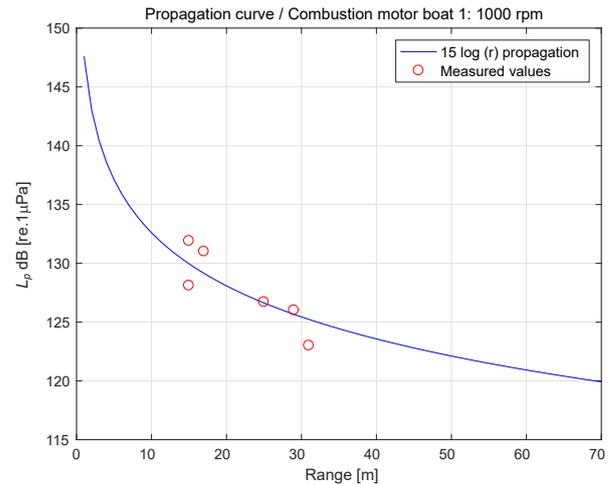


Fig. 7. Comparison between calculated transmission loss, distance and those levels obtained during measurement for the combustion boat 1, at velocity 1000 rpm.

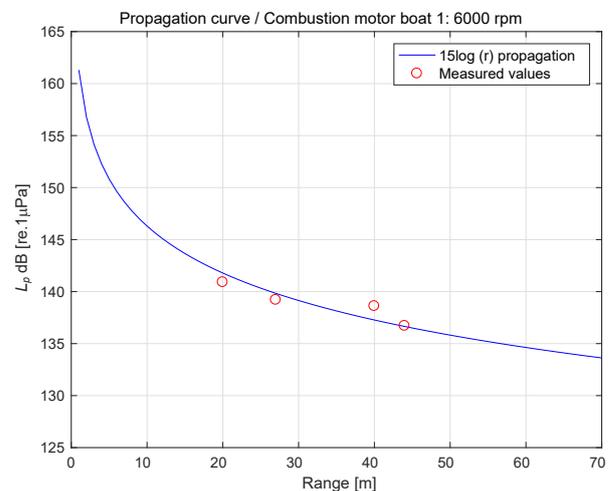


Fig. 8. Comparison between calculated transmission loss, distance and those levels obtained during the measurement for combustion boat 1, at velocity 6000 rpm.

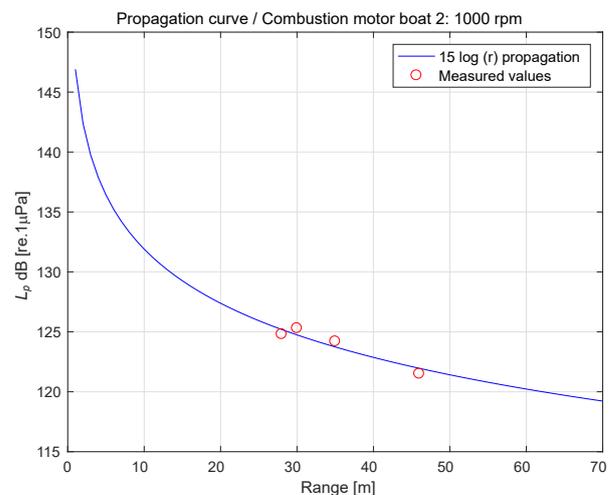


Fig. 9. Comparison between calculated transmission loss, distance and those levels obtained during the measurement for combustion boat 2, at velocity 1000 rpm.

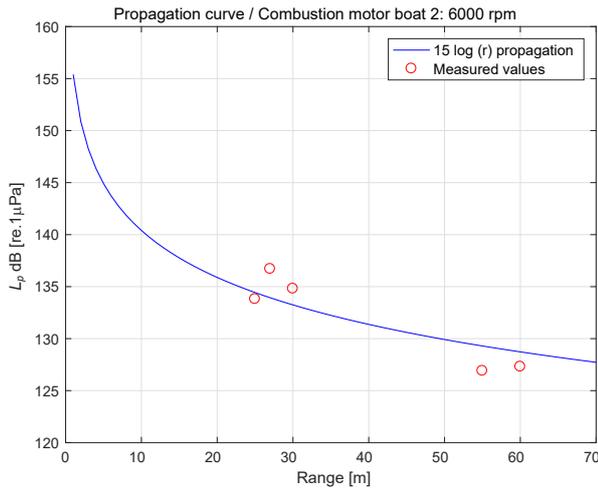


Fig. 10. Comparison between calculated transmission loss, distance and those levels obtained during the measurement for combustion boat 2, at velocity 6000 rpm.

In shallow waters and estuaries, where continuous tide changes occur, a dependence between sound velocity and depth is little probable, since a column of mixed and isothermal water is produced (Marine Management Organisation, 2015). Sound absorption was not considered in this study, since it is negligible for distances under 100 m (RICHARDSON *et al.*, 1995; URICK 1996; Marine Management Organisation, 2015).

5. Results

5.1. Source levels

For the three boats evaluated in this study, 44 measurements were made, with a minimum of 13 measurements per boat. Measurements with three different velocities were carried out for each boat: highest, mid and lowest velocities, which corresponded to 6000, 3000, and 1000 rpm, respectively, for internal combustion motor boats; and to 1100, 3000, and 5500 rpm, respectively, for the electric motor boat. A minimal of 4 measurements were carried out for each one of these velocities.

Measurements took place between September 2018 and January 2020, always under weather conditions of no rain or wind and a Beaufort wind scale between 0 and 1 (Hydrographic and Oceanographic Service of the Chilean Navy, 2002). This allowed working with very low levels of natural background noise, with values between 99 dB and 101 dB [re.1μPa]. The hydrophone depth was constant at 4 meter. The distance between the hydrophone and boats (distance *d*) constantly varied due to the field conditions under which the study was carried out.

A 15-second time interval centered on the closest point between the boat and the hydrophone (CPA), was selected from audio files. Three levels $L_{eq,T}$ of

5 seconds were taken from this range. Central $L_{eq,5s}$ showed the highest level and was chosen in this work as the level emitted from the ship. Thus, by using this emission level, the worst condition is being considered when evaluating environmental impact. Tables 2–4 show levels $L_{eq,5s}$ obtained and those distances they were measured at.

Table 2. Sound pressure levels measured for the electric motor boat; *d* is the measurement distance.

No.	Electric motor boat					
	$L_{eq,5s}$ dB [re.1μPa], <i>d</i> [m]					
	1100 rpm		3000 rpm		5500 rpm	
	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>
1	106.4	7	113	7	123.5	7
2	110.9	7	113.9	7	125.9	7
3	110.8	7	112.8	7	122.9	7
4	111.4	7	114.6	7	122.1	7
5	111.5	7	115.1	7	122.4	7
6					121.5	7
7					122.7	7

Table 3. Sound pressure levels measured for the combustion engine boat 1; *d* is the measurement distance.

No.	Combustion motor boat 1					
	$L_{eq,5s}$ dB [re.1μPa], <i>d</i> [m]					
	1000 rpm		3000 rpm		6000 rpm	
	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>
1	123	31	131.6	19	140.9	20
2	128.1	15	127.7	35	138.6	40
3	131	17	133.2	20	139.2	27
4	126	29	129.4	35	136.7	44
5	131.9	15				
6	126.7	25				

Table 4. Sound pressure levels measured for the combustion engine boat 2; *d* is the measurement distance.

No.	Combustion motor boat 2					
	$L_{eq,5s}$ dB [re.1μPa], <i>d</i> [m]					
	1000 rpm		3000 rpm		6000 rpm	
	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>	$L_{eq,5s}$	<i>d</i>
1	121.5	46	123.6	40	126.9	55
2	124.2	35	124.8	41	133.8	25
3	125.3	30	125.3	42	134.8	30
4	124.8	28	128.4	27	136.7	27
5					127.3	60

To obtain the source levels $L_{p,1m}$ Eq. (7) was used. This transmission loss by distance was the one that best fitted those field values obtained. This fitting may be appreciated in Figs. 7–10, which show some of the obtained results.

Source levels obtained for the assessed boats, estimated from values presented in Tables 2–4, and Eq. (7) are shown in Figs. 11–13.

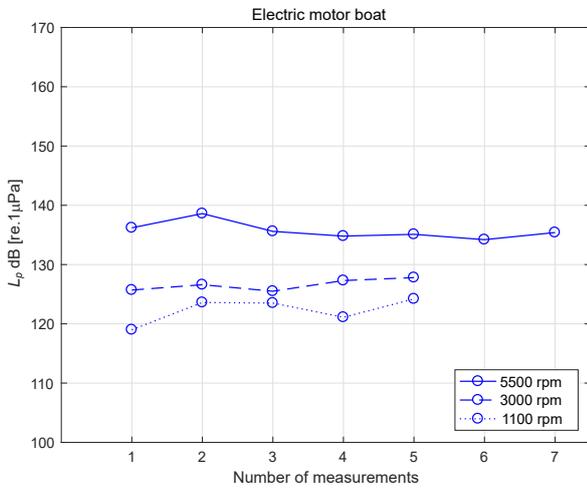


Fig. 11. Source levels $L_{p,1m}$, obtained for the electric motor boat.

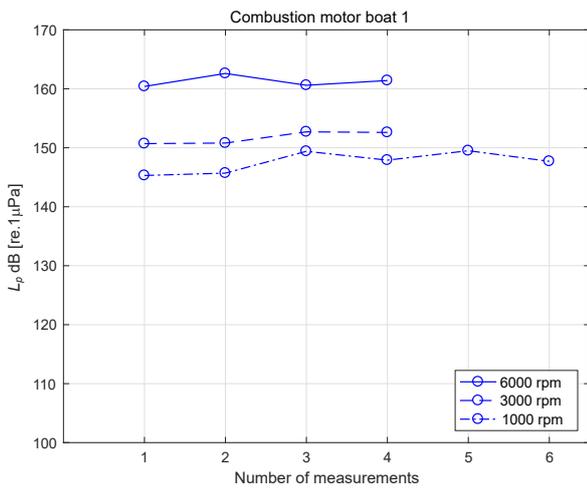


Fig. 12. Source levels $L_{p,1m}$, obtained for the combustion motor boat 1.

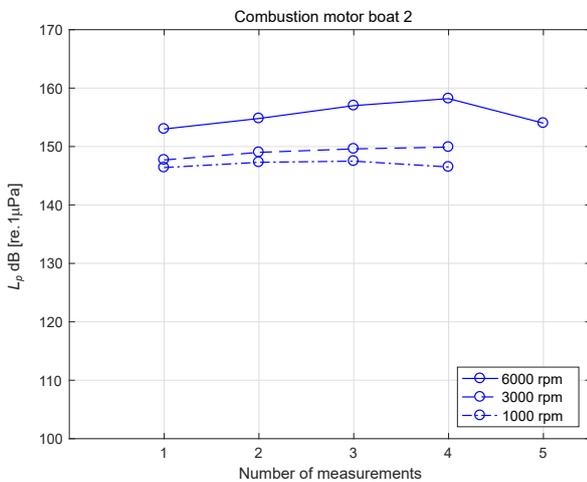


Fig. 13. Source levels $L_{p,1m}$, obtained for the combustion motor boat 2.

Now, the average source levels, together with their standard deviation s , uncertainty and highest and lowest levels are shown in Table 5 and Figs. 14–16. The uncertainty of the measurements was estimated through a confidence interval CI, calculated with a confidence level of 95% and using t-Student’s criterion (VELASCO LUNA, 2002).

Table 5. Average values of the measured source levels $L_{p,1m}$ of the evaluated boats, together with their standard deviation s , confidence interval CI and lowest and highest values.

Electric motor boat, L_p dB [re.1μPa]					
Velocity	$L_{p,1m}$	s	CI	$L_{p,max}$	$L_{p,min}$
1100 rpm	122.3	2.17	122.3 ± 3	124.2	119
3000 rpm	126.6	0.99	126.6 ± 1	127.8	125.5
5500 rpm	135.7	1.43	135.7 ± 1	138.6	134.2
Combustion motor boat 1, L_p dB [re.1μPa]					
Velocity	$L_{p,1m}$	s	CI	$L_{p,max}$	$L_{p,min}$
1000 rpm	147.6	1.8	147.6 ± 2	149.5	145.3
3000 rpm	151.7	1.07	151.7 ± 2	152.7	150.7
6000 rpm	161.3	1	161.3 ± 2	162.6	160.4
Combustion motor boat 2, L_p dB [re.1μPa]					
Velocity	$L_{p,1m}$	s	CI	$L_{p,max}$	$L_{p,min}$
1000 rpm	146.9	0.56	146.9 ± 1	147.5	146.4
3000 rpm	149	0.99	149 ± 2	149.9	147.7
6000 rpm	155.4	2.15	155.4 ± 3	158.2	153

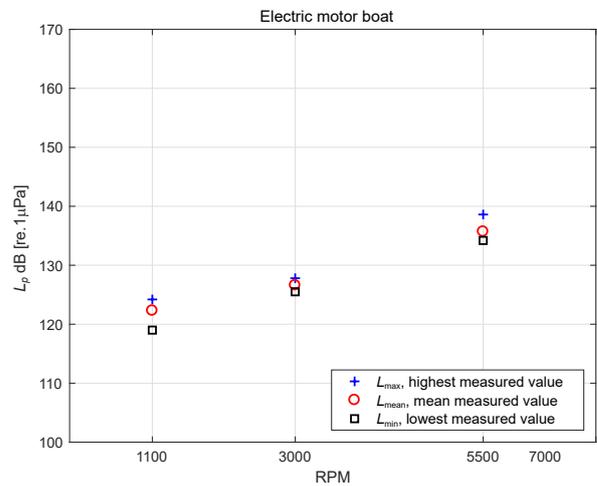


Fig. 14. Average source levels $L_{p,1m}$ of the electric motor boat, together with highest and lowest levels for the three evaluated velocities.

5.2. Noise emission spectra

The 1/3 octave band noise spectra for the three assessed boats were obtained, which are shown in Figs. 17, 18, and 19. Spectra are from the sound emissions corresponding to the moment when boats pass in front of the hydrophone and at the closest distance or

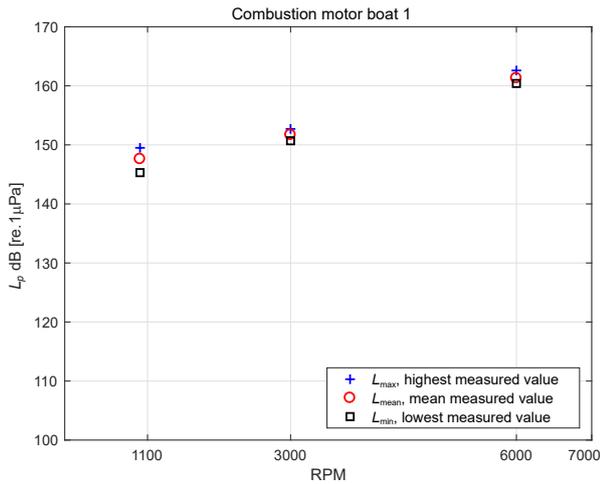


Fig. 15. Average source levels $L_{p,1m}$ of the combustion motor boat 1, together with the highest and lowest levels, for the three evaluated velocities.

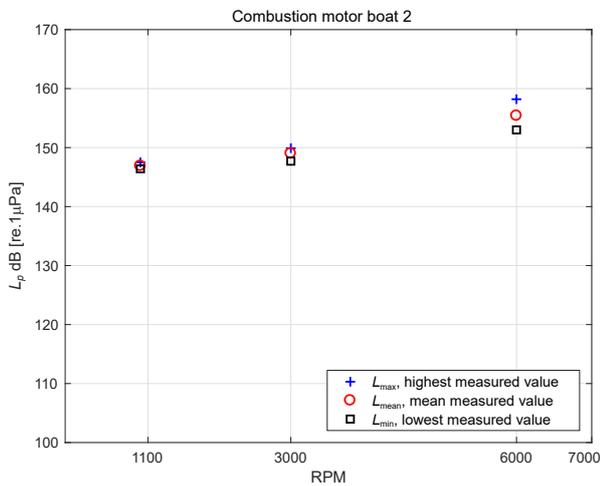


Fig. 16. Average source levels $L_{p,1m}$ of the combustion motor boat 2, together with the highest and lowest levels, for the three evaluated velocities.

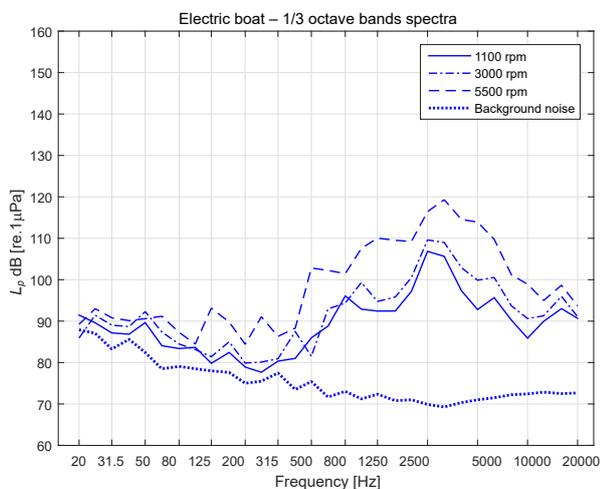


Fig. 17. 1/3 octave bands spectra of the electric motor boat, for the three velocities evaluated. Likewise, the background noise registered in the area is shown.

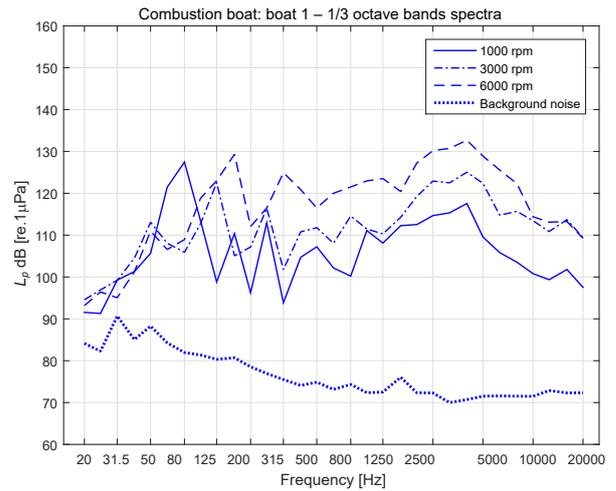


Fig. 18. 1/3 octave bands spectra combustion motor boat 1, for the three velocities evaluated. Background noise measured in the area is also shown.

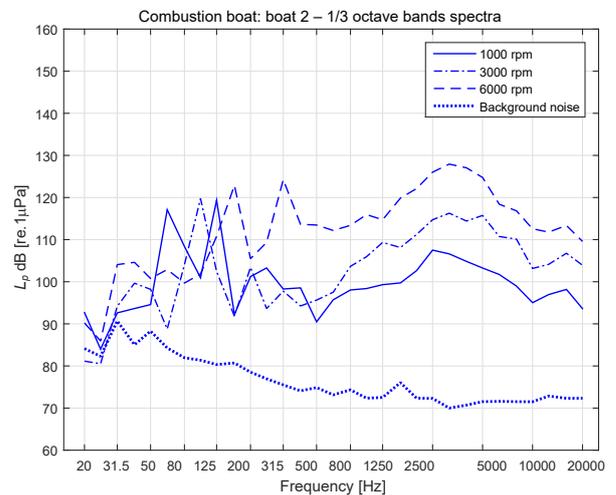


Fig. 19. 1/3 octave bands spectra of the combustion motor boat 2, for the three velocities evaluated. Background noise measured in the place is also shown.

CPA. Spectra correspond to only one passage of the boat in front of the hydrophone and for each one of the three velocities. Together with the boats spectra, figures show the spectra of the background noise evaluated in the area.

6. Discussion

This work comprises a comparative study of the underwater noise levels emitted by boats with similar dimensions and passenger's capacity, used in identical touristic duties, though using different propelling systems. Results show that source levels $L_{p,1m}$ of the assessed boats propelled by internal combustion outboard motor are between 145.3 dB and 162.6 dB [re.1μPa], for the combustion motor boat 1, and between 146.4 dB and 158.2 dB [re.1μPa], for the com-

bustion motor boat 2. Source levels $L_{p,1m}$ obtained for the electric motor boat fluctuated between 119 dB and 138.6 dB [re.1 μ Pa]. All these emission levels were registered for the boats velocities between 1000 rpm and 6000 rpm.

The background noise level with which measurements were carried out was between 99 dB and 101 dB [re.1 μ Pa]. Wide band levels measured and observed, emitted by the evaluated sources, were always superior to the background noise observed over 10 dB, even for the lowest velocities. Regarding the lowest 1/3 octave frequency bands due to differences inferior to 10 dB, some level corrections were made. Nevertheless, these corrections did not influence the broadband level obtained.

Source levels $L_{p,1m}$ obtained in this work, corresponding to boats propelled by internal combustion engines, coincide with those levels shown by other studies for the same type of boats (RICHARDSON *et al.*, 1995; Marine Management Organisation, 2015; ERBE, 2002; WLADICHUK *et al.*, 2019). Regarding levels emitted by the electric motor vessel, it was not possible to find literature showing values of emission levels for this type of boat.

According to the results obtained, the boat propelled by electric motor emits underwater noise levels much lower than those emitted by the internal combustion boats. For the low velocity (1000–1100 rpm), the underwater noise level emitted by the electric motor boat is approximately 25 dB lower than the level emitted by the combustion motor boat 1, and 24.6 dB lower than the level emitted by the combustion motor boat 2. For the mid velocity (3000 rpm), the electric motor boat emits approximately 25.1 dB less than the level emitted by the combustion motor boat 1, and 22.4 dB less than the emissions of the combustion motor boat 2. For the highest velocity (5500–6000 rpm), the level of underwater noise emitted by the electric motor boat is approximately 25.6 dB lower than the level emitted by the combustion motor boat 1, and 19.7 dB lower than the level emitted by the combustion motor boat 2.

This difference between underwater noise levels emitted by the electric motor boat and those levels emitted by combustion motor engines arises as highly important when we consider their use inside marine protected areas, in duties such as marine mammal watching and also regarding the fulfillment of the highest recommended noise levels to avoid any type of disturbance, temporal damage TTS or permanent damage PTS, to the observed species (NMFS, 2016; 2018). Letting the source level $L_{p,1m}$ in Eq. (7) be the dependent variable, one can see that a difference of 25 dB between the level of underwater noise emitted by two boats means that there will be an approximate difference of 46 times in the closest distance that the less noisy boat could reach when approaching a certain marine mam-

mal, compared to the closest distance that the loudest boat could reach, to generate the same level of noise on the animal exposed to noise.

Similarly, Eq. (6) shows that this 25 dB difference means a 316 times lower accumulated energy for an exposure to noise emitted by the electric boat compared to the exposure to noise emitted by the boat with an internal combustion engine.

The 1/3 octave band frequency spectra agree with the obtained source levels $L_{p,1m}$. It can be clearly seen how the level of frequency components increases as velocity of boats increases. Combustion engine boats present similar spectra regarding shape and behavior (see Figs. 18 and 19). In them, low-frequency tonal components are observed, which are produced by the mechanical vibrations of the rotating parts of the motor. The frequency of the tonal components increases with increasing propeller rpm. A continuous spectrum is observed over 500 Hz, which is mainly due to the cavitation noise produced by the propeller, which is of impulsive origin and broadband. Spectra show how the amplitude relation between tonal components and continuous components goes changing. Thus, as the boat velocity increases, cavitation noise produced by the propeller becomes dominant in noise emissions over the mechanical noise produced by the motor, which is predominant at low velocities.

Regarding the emission spectra of the electric motor boat (see Fig. 17), the cavitation noise produced by the propeller increases as the boat velocity increases. The cavitation noise produced by this vessel is lower than that produced by combustion motor boats for all evaluated velocities. What is interesting in these spectra is that low frequency tonal components, corresponding to periodic vibrations of the motor and propeller, show extremely low amplitudes. This may be due to different reasons. As it is an electric engine, the high periodic vibration produced by combustion does not exist; the rotating parts of the motor – as it is an inboard motor – are inside the hull; and finally, the modern design of the hull allows for a less turbulent water flow around the propeller, which reduces the excitation of resonances in the propeller (URICK, 1996).

7. Conclusion

The results reached in this study show that for the type of evaluated vessels, using electric motor boats instead of those commonly used boats – propelled by internal combustion motors – underwater noise levels emitted by an amount between 20 dB and 25 dB are reduced. This reduction in the level of emitted noise allows the electric boat, when necessary, a closer approach to the animal under observation. Furthermore, it will allow larger time of observation or a much lower accumulated exposition of the animal to the noise.

A comparative study to evaluate the acoustic impact that each one of these vessels would cause over a determined species would be very interesting and necessary.

References

1. ABRAHAMSEN K. (2012), The ship as an underwater noise source, [in:] *Proceedings of the 11th European Conference on Underwater Acoustics*, **17**(1): 070058, doi: [10.1121/1.4772953](https://doi.org/10.1121/1.4772953).
2. ANDREW R.K., HOWE B.M., MERCER J.A., DZIECIUCH M.A. (2002), Ocean ambient sound: Comparing the 1960s with the 1990s for a receiver off the California coast, *Acoustics Research Letters Online*, **3**(2): 65–70, doi: [10.1121/1.1461915](https://doi.org/10.1121/1.1461915).
3. ANSI/ASA S12.64 (2009), *Quantities and procedures for description and measurement of underwater sound from ships – Part 1: General requirements*.
4. BJØRNØ L. (2017), *Applied Underwater Acoustics*, Elsevier, eBook ISBN: 9780128112472.
5. California Department of Transportation (2020), *Technical guidance for the assessment of hydroacoustic effects of pile driving on fish*, <https://dot.ca.gov/-/media/dot-media/programs/environmental-analysis/documents/env/hydroacoustic-manual-a11y.pdf>.
6. CAREY W.M. (2009), Lloyd's mirror – Image interference effects, *Acoustics Today*, **5**(2): 14–20, doi: [10.1121/1.3182842](https://doi.org/10.1121/1.3182842).
7. Commission of the European Communities, (2008), *Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive)*, <http://data.europa.eu/eli/dir/2008/56/oj>.
8. EKDAHL ESPINOZA H. (2014), Solar powered boats [in Spanish: Embarcaciones propulsadas por energía solar], *Revista de Marina*.
9. EN 60565 (2007), *Underwater acoustics – Hydrophones – Calibration in the frequency range 0.01 Hz to 1 MHz*, European Standard, pp. 29–30.
10. ERBE C. (2002), Underwater noise of whale-watching boats and potential effects on killer whales (*Orcinus orca*), based on an acoustic impact model, *Marine Mammal Science*, **18**(2): 394–418, doi: [10.1111/j.1748-7692.2002.tb01045.x](https://doi.org/10.1111/j.1748-7692.2002.tb01045.x).
11. ERBE C., REICHMUTH C., CUNNINGHAM K., LUCKE K., DOOLING R. (2016), Communication masking in marine mammals: A review and research strategy, *Marine Pollution Bulletin*, **103**(1–2): 15–38, doi: [10.1016/j.marpolbul.2015.12.007](https://doi.org/10.1016/j.marpolbul.2015.12.007).
12. European Commission (2017), *Commission decision (EU) 2017/848 of 17 May 2017*.
13. GRELOWSKA G., KOZACZKA E. (2014), Underwater acoustic imaging of the sea, *Archives of Acoustics*, **39**(4): 439–452, doi: [10.2478/aoa-2014-0048](https://doi.org/10.2478/aoa-2014-0048).
14. GRELOWSKA G., KOZACZKA E., KOZACZKA S., SZYMCAK W. (2013), Underwater noise generated by a small ship in the shallow sea, *Archives of Acoustics*, **38**(3): 351–356, doi: [10.2478/aoa-2013-0041](https://doi.org/10.2478/aoa-2013-0041).
15. HAWKINS A.D., MYRBERG A.A.Jr., (1983), Hearing and sound communication underwater, [in:] *Bioacoustics, a Comparative Approach*, Lewis B. [Ed.], pp. 347–405, Academic Press, London.
16. Hydrographic and Oceanographic Service of the Chilean Navy (2002), *Tide and Current Glossary* [in Spanish: *Glosario de Marea y Corrientes*], 3rd ed., Pub. SHOA 3013.
17. ISO/PAS 17208-1 (2012), *Acoustics – Quantities and procedures for description and measurement of underwater sound from ships – Part 1: General requirements for measurements in deep water*, ISO Standard.
18. KINSLER L.E., FREY A.R., COPPENS A.B., SANDERS J.V. (1999), *Fundamentals of Acoustics*, 4th ed, Wiley, Germany.
19. KLAUSON A., MUSTONEN M. (2017), Ship source strength estimation in shallow water, *Proceedings of Meetings on Acoustics*, **142**(4): 2686, doi: [10.1121/1.5014794](https://doi.org/10.1121/1.5014794).
20. KUNC H.P., McLAUGHLIN K.E., SCHMIDT R. (2016), Aquatic noise pollution: Implications for individuals, populations, and ecosystems, *Proceedings of the Royal Society B: Biological Sciences*, **283**(1836), doi: [10.1098/rspb.2016.0839](https://doi.org/10.1098/rspb.2016.0839).
21. Marine Management Organisation [MMO] (2015), *Modelled mapping of continuous underwater noise generated by activities*, MMO Project No: 1097, ISBN: 978-1-909452-87-9.
22. McCAULEY R.D., FEWTRELL J., POPPER A.N. (2003), High intensity anthropogenic sound damages fish ears, *The Journal of the Acoustical Society of America*, **113**(1): 638–642, doi: [10.1121/1.1527962](https://doi.org/10.1121/1.1527962).
23. McDONALD M.A., HILDEBRAND J.A., WIGGINS S.M. (2006), Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California, *The Journal of the Acoustical Society of America*, **120**(2): 711–718, doi: [10.1121/1.2216565](https://doi.org/10.1121/1.2216565).
24. MCGARRY *et al.* (2020), *Evidence base for application of Acoustic Deterrent Devices (ADDs) as marine mammal mitigation (Version 3)*, Joint Nature Conservation Committee, Report No. 615.
25. Ministerio del Medio Ambiente (2015), *Decreto no. 41*, Chile.
26. Ministerio de Educación Pública (1981), *Decreto no. 2734*, Chile.
27. National Marine Fisheries Service [NMFS] (2016), *Technical guidance for assessing the effects of anthropogenic sound on marine mammal hearing: Underwater acoustic thresholds for onset of permanent and temporary threshold shifts*, U.S. Department of Commerce, NOAA, NOAA Technical Memorandum NMFS-OPR-55.
28. National Marine Fisheries Service [NMFS] (2018), *2018 Revisions to: Technical guidance for assessing the effects of anthropogenic sound on marine mammal hearing (Version 2.0): Underwater thresholds for onset of*

- permanent and temporary threshold shifts*, U.S. Department of Commerce, NOAA, NOAA Technical Memorandum NMFS-OPR-59.
29. National Oceanic and Atmospheric Administration [NOAA] (2020), *Guidelines for safely deterring marine mammals*, <https://www.federalregister.gov/documents/2020/08/31/2020-18718/guidelines-for-safely-detering-marine-mammals>.
 30. National Physical Laboratory [NPL] (2014), *Good practice guide for underwater noise measurement*, National Measurement Office, Marine Scotland, The Crown Estate, NPL Good Practice Guide No. 133, ISSN: 1368-6550, <https://www.npl.co.uk/special-pages/guides/gpg133underwater>.
 31. RICHARDSON W.J., GREENE C.R., MALME C.I., THOMSON D.H. (1995), *Marine Mammals and Noise*, Academic Press, San Diego.
 32. ROSS D. (2005), Ship sources of ambient noise, *IEEE Journal of Oceanic Engineering*, **30**(2): 257–261, doi: [10.1109/JOE.2005.850879](https://doi.org/10.1109/JOE.2005.850879).
 33. SALAS M., LUCO R., EKDAHL H. (2013), Solar energy for propulsion of river vessels [in Spanish: Energía solar para propulsión de embarcaciones fluviales], [in:] *International Conference IDS2013 – Amazonia, Iquitos, Peru*.
 34. SANTOS-DOMÍNGUEZ D., TORRES-GUIJARRO S., CARDENAL-LÓPEZ A., PENA-GIMENEZ A. (2016), ShipsEar: An underwater vessel noise database, *Applied Acoustics*, **113**: 64–69, doi: [10.1016/j.apacoust.2016.06.008](https://doi.org/10.1016/j.apacoust.2016.06.008).
 35. Servicio de Evaluación Ambiental [SEA] (2022), *Evaluation criteria in the SEIA: Prediction and evaluation of impacts by submarine noise* [in Spanish: Criterio de evaluación en el SEIA: Predicción y evaluación de impactos por ruido submarino].
 36. SOUTHALL B.L. *et al.* (2007), Marine mammal noise exposure criteria, *Aquatic Mammals*, **33**(4): 411–522, doi: [10.1578/AM.33.4.2007.411](https://doi.org/10.1578/AM.33.4.2007.411).
 37. TAVOLGA W.N. (1965), *Review of marine bio-acoustics, Technical report: NAVTRADEVCEEN 1212-1*, U.S. Naval Training Device Center.
 38. TAVOLGA W.N. (1971), Sound production and detection, [in:] *Fish Physiology*, HOAR W.S., RANDALL D.J. [Eds.], **5**: 135–205, Academic Press, New York, doi: [10.1016/S1546-5098\(08\)60047-3](https://doi.org/10.1016/S1546-5098(08)60047-3).
 39. TYACK P.L., (2000), Functional aspects of cetacean communication, [in:] *Cetacean Societies: Field Studies of Dolphins and Whales*, Mann J., Connor R.C., Tyack P.L., Whitehead H. [Eds.], Chicago, The University of Chicago Press, pp. 270–307, doi: [10.1017/S0376892901290183](https://doi.org/10.1017/S0376892901290183).
 40. TYACK P.L. (2008), Implication for marine mammals of large-scale changes in the marine acoustic environment, *Journal of Mammalogy*, **89**(3): 549–558, doi: [10.1644/07-MAMM-S-307R.1](https://doi.org/10.1644/07-MAMM-S-307R.1).
 41. URICK R.J. (1996), *Principles of Underwater Sound*, 3rd ed., Peninsula Publishing.
 42. VELASCO LUNA F. (2002), Introduction and bibliography of the student t-distribution and its uses in statistical inference [in Spanish: Introducción y bibliografía de la distribución t de student y sus usos en inferencia estadística], *Universidad Veracruzana, Facultad de Estadística e Informática*.
 43. WLADICHUK J.L., HANNAY D.E., MACGILLIVRAY A.O., LI Z., THORNTON S.J. (2019), Systematic source level measurements of whale watching vessels and other small boats, *The Journal of Ocean Technology*, **14**(3).
 44. WRIGHT A.J. *et al.* (2007), Do marine mammals experience stress related to anthropogenic noise?, *International Journal of Comparative Psychology*, **20**(2): 274–316, doi: [10.46867/ijcp.2007.20.02.01](https://doi.org/10.46867/ijcp.2007.20.02.01).
 45. WYSOCKI L.E., DITTAMI J.P., LADICH F. (2006), Ship noise and cortisol secretion in European freshwater fishes, *Biological Conservation*, **128**(4): 501–508, doi: [10.1016/j.biocon.2005.10.020](https://doi.org/10.1016/j.biocon.2005.10.020).

Research Paper

Study on Noise Attenuation Characteristics of Hydrofoil with Specific Cavitation Number

He XIAOHUI^{(1)*}, Liu ZHONGLE⁽²⁾, Yang CHAO⁽¹⁾, Yuan ZHIYONG⁽²⁾

⁽¹⁾ Jiangnan Industry Group Co., Ltd.
Wuyi Village, China

⁽²⁾ Naval University of Engineering
Wuhan, China

*Corresponding Author e-mail: xiaohui.he16@gmail.com

(received September 23, 2022; accepted January 26, 2023)

In this study, the modified Sauer cavitation model and Kirchhoff-Ffowcs Williams and Hawkings (K-FWH) acoustic model were adopted to numerically simulate the unsteady cavitation flow field and the noise of a three-dimensional NACA66 hydrofoil at a constant cavitation number. The aim of the study is to conduct and analyze the noise performance of a hydrofoil and also determine the characteristics of the sound pressure spectrum, sound power spectrum, and noise changes at different monitoring points. The noise change, sound pressure spectrum, and power spectrum characteristics were estimated at different monitoring points, such as the suction side, pressure side, and tail of the hydrofoil. The noise characteristics and change law of the NACA66 hydrofoil under a constant cavitation number are presented. The results show that hydrofoil cavitation takes on a certain degree of pulsation and periodicity. Under the condition of a constant cavitation number, as the attack angle increases, the cavitation area of the hydrofoil becomes longer and thicker, and the initial position of cavitation moves forward. When the inflow velocity increases, the cavitation noise and the cavitation area change more drastically and have a superposition tendency toward the downstream. The novelty is that the study presents important calculations and analyses regarding the noise performance of a hydrofoil, characteristics of the sound pressure spectrum, and sound power spectrum and noise changes at different monitoring points. The article may be useful for specialists in the field of engineering and physics.

Keywords: sound pressure spectrum; noise; sound power spectrum; numerical prediction.



Copyright © 2023 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

Cavitation, the formation of bubbles in a liquid, is a phenomenon that generally occurs at the interface between a fluid and a solid with relative motion (CAO *et al.*, 2014). In other words, when the local pressure in the flow field drops to the saturated vapor pressure at its proper temperature, the liquid medium will explosively vaporize and form bubbles. When cavitation reaches a certain threshold, it will be accompanied by the burst and detachment of bubble groups, which causes strong noise, vibration and cavitation erosion (SULTANOV *et al.*, 2020). The spatial differences of the closing emerge throughout the cavity's development, and under specific circumstances, the pocket becomes unstable and violently implodes. The volume of the va-

por cavity oscillates between a minimum and a maximum during this operating regime. The destabilizing process results in the emission of biphasic and vortex structures known as cavitation clouds, which are highly erosive and known to produce large overpressures.

Cavitation flow encompasses almost all complex flow phenomena: turbulence, multiphase flow, phase transition, compressible, and unsteady characteristics, etc., (HUANG *et al.*, 2018; PROKOPOV *et al.*, 1993). Cavitation can be divided into different forms: primary cavitation, sheet cavitation, cavitation cloud, eddy cavitation, and super-cavitation (WANG *et al.*, 2001). Cavitation often has adverse effects on underwater and water conservancy equipment, for example, by reducing thrust efficiency, severely corroding the structure of equipment, and affecting normal performance

(CHANG, 2011; SHIN *et al.*, 2021). On the other hand, scholars and engineers have also explored the potential benefits of cavitation. For example, it is used for sterilizing and crushing stones in the medical field, industrial sewage treatment, oil exploitation, underwater cutting, shock absorption, noise reduction for underwater weapons, etc.

Many scholars have conducted extensive theoretical and experimental research on hydrofoils and acoustic radiation. From a theoretical perspective, Euler proposed the cavitation phenomenon for the first time in 1753. In 1839, Reynolds and Besant (LI, SHI, 1997) studied cavitation in the laboratory. LORD RAYLEIGH (1917) formulated the mathematical equation of cavitation bubbles in an incompressible fluid, and PLESSET (1949) made a correction to the theory of Rayleigh. NOLTINGK, NEPPIRAS (1950) and NEPPIRAS, NOLTINGK (1951) added an additional pressure correction to the surface tension in the Rayleigh model, and in 1952, Poritsky added a liquid viscosity correction (KORZHYK *et al.*, 2017). Since then, many scholars have continuously developed this theory.

An important work is presented in (WANG *et al.*, 2021). The research examines the effect of water injection on broadband noise and hydrodynamic performance for NACA66 (MOD) hydrofoils under cloud cavitation conditions. The influence of water injection on the hydrodynamic performance and noise sources for a NACA66 (MOD) hydrofoil under cloud cavitation is computed in this work ($\sigma = 0.83$, $\text{Re} = 5.1 \times 10^5$). The results of the analysis show that the water injection may effectively stop the growth of cloud cavitation and significantly reduce the severe pressure fluctuation. As a result, the flow field's dipole/quadrupole noise can be reduced. KUBOTA *et al.* (1992) proposed a cavitation model based on the transport equation on the basis of the Rayleigh–Plesset equation. Based on the Rayleigh–Plesset equation, ZWART (2004), SCHNERR, SAUER (2001), and SINGHAL *et al.* (2002) established their respective cavitation transport equations representing the relationship between mass transport and pressure change. KIELDSSEN *et al.* (2000) conducted experimental research on a NACA0015 hydrofoil. From an experimental perspective, LEROUX *et al.* (2003; 2004) carried out a study on the fracture and detachment phenomenon of cavitation generated by the unsteady cavitation of a single NACA66 hydrofoil. FUJI *et al.* (2007) studied the influence of the geometric shapes of hydrofoils on cavitation dynamics and summarized NACA0015 hydrofoil decrease in different water holes. Through FFT analysis, the influence of instability on cavitation dynamics was obtained. HONG *et al.* (2017) studied the Clark-Y hydrofoil characteristics with different cavitation numbers.

WANG *et al.* (2009) and ZHANG *et al.* (2009) performed a time-frequency analysis on the unsteady dynamic characteristics of the cavitation around the hy-

drofoil. DUCOIN *et al.* (2009; 2012), WU *et al.* (2005), WANG, OSTOJA-STARZEWSKI (2007), JI *et al.* (2010), and others studied fluid-solid coupling and the cavitation of two-dimensional and three-dimensional hydrofoils by combining numerical simulation and experiments. FAN (2015) performed research on the vibration and acoustic radiation of a hydrofoil. The purpose of the article is to conduct and analyze numerically the unsteady cavitation flow field and the noise of a three-dimensional NACA66 hydrofoil under a constant cavitation number.

2. Materials and methods

2.1. Mathematical, noise, and cavitation models

In the calculation process, the homogeneous equilibrium flow model is adopted. Assuming that there is no velocity slip between the gas and liquid, the mass conservation equation, the momentum equation, and the density equation of a three-phase mixture are, respectively:

Mass equation:

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m u) = 0. \quad (1)$$

Momentum equation:

$$\frac{\partial \rho_m u}{\partial t} + \nabla \cdot (\rho_m uu) = -\nabla p + \nabla \cdot \tau + S_M, \quad (2)$$

where $\tau = \mu \left(\nabla u + (\nabla u)^T - \frac{2}{3} \delta \nabla \cdot u \right)$.

The density of the mixture is defined by the equation:

$$\rho_m = a_l \rho_l + a_v \rho_v + a_g \rho_g, \quad (3)$$

where a_l , a_v , and a_g represent the volume fraction of gas in the liquid phase, vapor phase and non-condensation state, respectively, and ρ_l , ρ_v , and ρ_g refer to the densities of gas in these three states, respectively.

Mass fraction:

$$y_i = \frac{a_i \rho_i}{\rho_m}. \quad (4)$$

In the numerical calculation, the turbulence motions are roughly divided into three types: firstly, direct simulation (DNS), secondly, large eddy simulation (LES), and thirdly, Reynolds-averaging averaging simulation.

In this study, the large eddy simulation model (LES model) was adopted to solve the transient Navier–Stokes equation, which can directly simulate large eddies in turbulence, but not small eddies. As a result, a similar model was established to simulate the influence of small eddies on large ones. That is, the Navier–Stokes equation is filtered in the wavenumber space or physical space. The filtering process removes small eddies, the width of which is less than the filtering width

or the given physical width, in order to obtain the control equation of large eddies (ALKISHRIWI *et al.*, 2008):

$$\bar{\Phi}(x) = \int_D \Phi(x') G(x; x') \lim_{x \rightarrow \infty} dx', \quad (5)$$

where D is the fluid domain, G is the filter function that determines the scale of the resolved eddies, and $\bar{\Phi}$ is defined by $\Phi' = \Phi - \bar{\Phi}$ and $\bar{\Phi}' \neq 0$.

The discretization of the spatial domain into finite control volumes implicitly provides the filtering operation:

$$\bar{\Phi}(x) = \frac{1}{V} \int_V \Phi(x') dx', \quad x' \in V, \quad (6)$$

where V is the control volume. The filter function $G(x; x')$ implied here is then:

$$G(x; x') = \begin{cases} 1/V, & x' \in V, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Filtering the Navier–Stokes equations leads to additional unknown quantities. The filtered momentum equation can be written in the following way:

$$\begin{aligned} \frac{\partial}{\partial t} (\bar{\rho} \bar{U}_i) + \frac{\partial}{\partial x_j} (\bar{\rho} \bar{U}_i \bar{U}_j) \\ = -\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left[\mu \left(\frac{\partial \bar{U}_i}{\partial x_j} + \frac{\partial \bar{U}_j}{\partial x_i} \right) \right] + \frac{\partial \tau_{ij}}{\partial x_j}, \end{aligned} \quad (8)$$

where τ_{ij} denotes the subgrid-scale stress:

$$\tau_{ij} = \rho \bar{U}_i \bar{U}_j - \bar{\rho} \bar{U}_i \bar{U}_j. \quad (9)$$

Flow-induced noise is generated by the disturbance propagation in the flow process. The disturbance propagation generates pressure fluctuations and propagates outward as a sound source. The boundary layer of the hydrofoil and the detachment of eddies radiate high-frequency noise, which is equivalent to the quadrupole source. The non-uniform flow field around the hydrofoil and the unsteady pulsation force on the hydrofoil surface induced by the pulsation turbulence field radiate low-frequency noise, which is equivalent to the dipole source, and the noise caused by the burst of bubbles is equivalent to the monopole.

The right side of the equation hereinafter refers to two surface source items (monopole and dipole) and one volume source item (quadrupole). The formula consists of a volume integral polynomial and a surface integral polynomial; the surface integral describes the contribution of the monopole source, dipole source, and part of the quadrupole source to the noise, while the volume integral defines the quadrupole contribution outside the control surface. If the quadrupole item is ignored, the equation of the pressure field is defined as:

$$p'(x, t) = p'_T(x, t) + p'_L(x, t). \quad (10)$$

The corresponding monopole, $p'_T(x, t)$, is the sound pressure due to the thickness. The corresponding dipole, $p'_L(x, t)$, defines the sound pressure due to the load. The formula is shown in Eqs. (11) and (12):

$$\begin{aligned} 4\pi p'_T(x, t) = \int_{f=0} \left[\frac{\rho_0 \dot{v}_n}{r(1-M_r)^2} \right] dS \\ + \int_{f=0} \left[\frac{\rho_0 v_n [r \dot{M}_r \hat{r}_i + c_0 (M_r - M^2)]}{r^2 (1-M_r)^3} \right] dS, \end{aligned} \quad (11)$$

$$\begin{aligned} 4\pi p'_L(x, t) = \frac{1}{c_0} \int_{f=0} \left[\frac{l_i \hat{r}_i}{r(1-M_r)^2} \right] dS \\ + \int_{f=0} \left[\frac{l_r - l_i M_i}{r(1-M_r)^2} \right] dS \\ + \frac{1}{c_0} \int_{f=0} \left[\frac{l_r [r \dot{M}_r \hat{r}_i + c_0 (M_r - M^2)]}{r^2 (1-M_r)^3} \right] dS, \end{aligned} \quad (12)$$

where M refers to the Mach number, M_r represents the radial Mach number, and l_i is the local force on the unit area at the direction of i , which can also be defined by Eqs. (13) and (14) (BELJATYNSKIJ *et al.*, 2010; CURLE, 1955; NURTAS *et al.*, 2020; PRENTKOVSKIS *et al.*, 2012; Su *et al.*, 2013; YANG *et al.*, 2014):

$$p'_T(x, t) = \int_{-T}^T \int_{A(\tau)} \rho v_n \frac{DG}{D\tau} dA(y) d\tau, \quad (13)$$

$$p'_L(x, t) = \int_{-T}^T \int_{A(\tau)} F_i \frac{DG}{D\tau} dA(y) d\tau. \quad (14)$$

A cavitation model mathematically describes the mutual transformation between water and vapor, which can be characterized by the modified Sauer cavitation model proposed by YANG *et al.* (2011; 2012):

$$\dot{m}^+ = \frac{C_{\text{prod}} 3a_g (1-a_v) \rho_v}{R_B} \sqrt{\frac{2|P_v - P|}{3\rho_l}} \text{sign}(P_v - P), \quad (15)$$

$$\dot{m}^- = C_{\text{dest}} \frac{3a_v \rho_v}{R_B} \sqrt{\frac{2|P_v - P|}{3\rho_l}} \text{sign}(P_v - P), \quad (16)$$

where \dot{m}^+ and \dot{m}^- represent evaporation and condensation of vapor, the mass fraction is $a_g = 7.8 \times 10^{-4}$, the volume fraction is $a_v = 1 \times 10^{-6}$, R_B is the initial value of the bubble radius, $R_B = 1.0 \times 10^{-6} m$, the evaporation coefficient is $C_{\text{prod}} = 50$, the condensation coefficient is $C_{\text{dest}} = 0.01$, and $P_v = P_{\text{sat}} + 0.5P_{\text{turb}}$, and $P_{\text{turb}} = 2\rho k/3$.

2.2. Divisions of geometric models and computational domain

A NACA66 hydrofoil was used as the research geometric model. The geometric parameters were obtained from PLESSET (1949), and the unsteady calculation was conducted in order to study the acoustic radiation law of the hydrofoil with different inflow velocities and different attack angles for a specific cavitation number (AL-OBAIDI, 2019, 2020; AL-OBAIDI, MISHRA, 2020):

$$\text{cavitation number: } \sigma_n = \frac{P - P_V}{\frac{1}{2}\rho V^2}, \quad (17)$$

$$\text{pressure coefficient: } C_p = \frac{P - P_V}{\frac{1}{2}\rho V^2}, \quad (18)$$

$$\text{Strouhal number: } St = \frac{fc}{V}, \quad (19)$$

$$\text{Reynolds number: } Re = \frac{Vc}{\nu}, \quad (20)$$

where P is the environmental pressure, P_V is the saturated vapor pressure, V is the inflow velocity, f is the falling-off period of the cavity, c is the chord length of the hydrofoil, and ν is the viscosity coefficient.

The computational domain setting is shown in Figs. 1 and 2 (geometric model establishment and grid division).



Fig. 1. The whole computational domain.

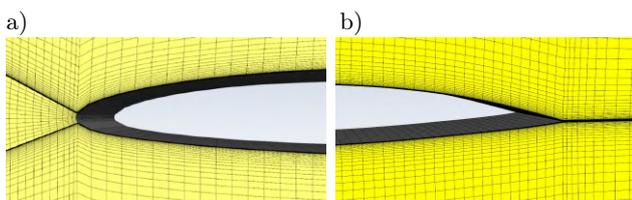


Fig. 2. Close-up view of mesh near the leading edge and the trailing edge of the hydrofoil: a) leading edge near the foil (LE); b) trailing edge near the foil (TE).

To reduce computational resources, the distance between the leading edge of the foil and the incoming flow is set to $2c$, the distance between the leading edge of the foil and the outlet of pressure is $6c$, the foil span is $0.3c$, and the height of the three-dimensional computational domain is $1.28c$. Cavitation number σ_n : 1.25; total working conditions of attack angles: 11, namely, 0° , $\pm 3^\circ$, $\pm 6^\circ$, $\pm 9^\circ$, $\pm 12^\circ$, and $\pm 15^\circ$; inflow velocity: 5.33 m/s, 10.288 m/s, and 20.577 m/s; environmental pressure: $21263.6 P_a$, $71279.5 P_a$, and $267697 P_a$.

The coordinates of the noise monitoring points are set. There are 31 noise monitoring points in total, whose coordinates are shown in Table 1.

Table 1. Settings of monitoring points (Z=0).

Monitoring points	X [mm]	Y [mm]
R1	0	0
R2	12.1195	4.9922
R3	37.5488	7.6615
R4	48.601	7.8085
R5	68.403	6.5358
R6	81.82	4.413
R7	100	-0.1178
R8	100	50
R9	100	100
R10	150	0
R11	150	50
R12	150	100
R13	200	0
R14	200	50
R15	200	100
R16	300	0
R17	300	50
R18	300	100
R19	13.2317	-2.9933
R20	36.6214	-3.9267
R21	47.689	-3.9042
R22	73.21	-2.6703
R23	91.23	-1.2005
R24	100	-50
R25	100	0
R26	150	-50
R27	150	-100
R28	200	-50
R29	200	-100
R30	300	-50
R31	300	-100

3. Results and discussion

When the inflow velocity is 5.33 m/s and the cavitation number is 1.25 ($\sigma_n = 1.25$), the simulation result is in good agreement with the experimental result of LEROUX *et al.* (2004). In addition, typically unstable cloud cavitation occurs, and it can be observed that a large number of vortexes detach from the surface of the foil (LEROUX *et al.*, 2004). The surface load of the foil also changes, and the lift and drag coefficients undergo a certain periodic change. When the length of

the cavity is less than $0.5c$, the bubble groups fluctuate less on the surface of the foil: such cavitation is called quasi-stable cavitation. However, when the length of the cavity is greater than $0.5c$, the cavity structure becomes quite unstable, and the bubbles burst. Cavitation pressure pulsation and the length of the cavity change regularly from the top of the blades. When the length of the cavity L/C reaches the maximum value of $0.7\sim 0.8$, it is called unstable cavitation. After verifying that the simulation is correct, the cavitation number is fixed at $\sigma_n = 1.25$, and then the second working

condition (inflow velocity of 10.288 m/s) and the third working condition (inflow velocity of 20.577 m/s) are selected.

Noise law of cavity in the unsteady growth process:

- working condition 1 – 5.33 m/s , Figs. 3–8;
- working condition 2 – 10.288 m/s , Figs. 9–14;
- working condition 3 – 20.577 m/s , Figs. 15–20.

Overall sound pressure level analysis:

- working condition 1 – 5.33 m/s , Figs. 21–22;
- working condition 2 – 10.288 m/s , Fig. 23;
- working condition 3 – 20.577 m/s , Fig. 24.

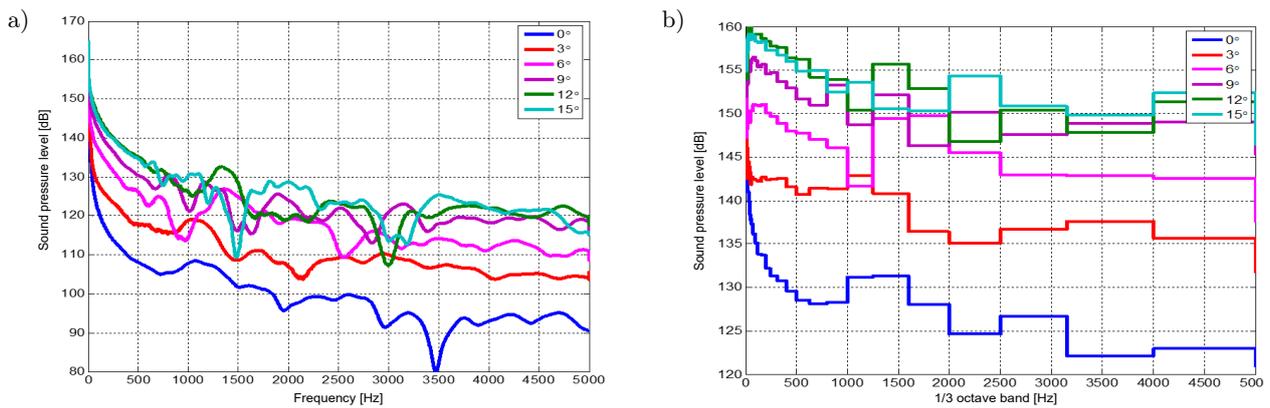


Fig. 3. Sound pressure frequency spectrum curves of R1 at $0^\circ, 3^\circ, 6^\circ, 9^\circ, 12^\circ, 15^\circ$ (cavitation number is fixed at $\sigma_n = 1.25$).

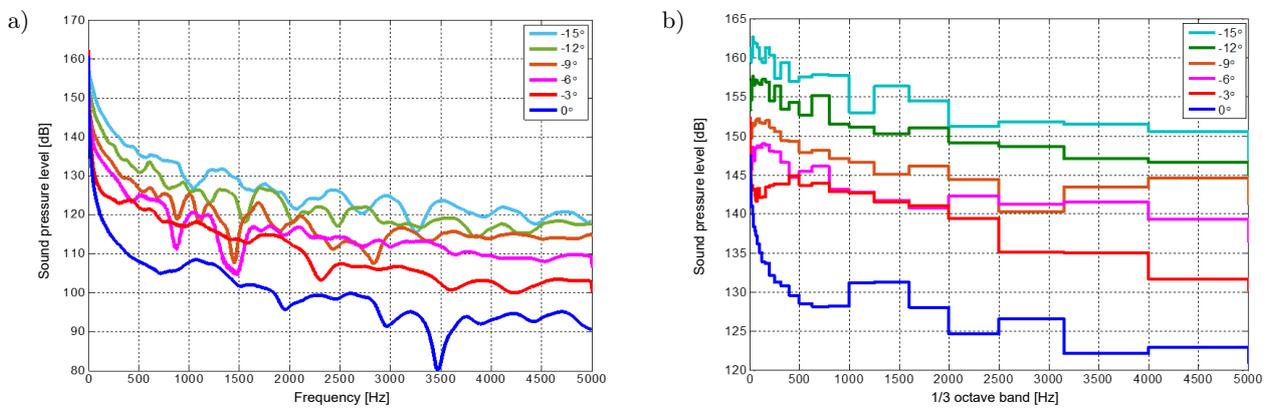


Fig. 4. Sound pressure frequency spectrum curves of R1 at $0^\circ, -3^\circ, -6^\circ, -9^\circ, -12^\circ, -15^\circ$.

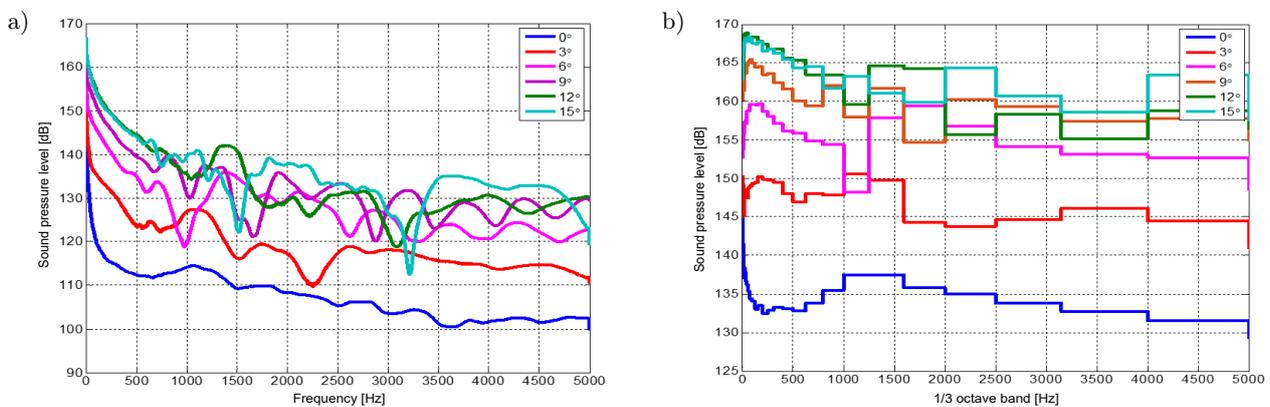


Fig. 5. Sound pressure frequency spectrum curves of R6 at $0^\circ, 3^\circ, 6^\circ, 9^\circ, 12^\circ, 15^\circ$.

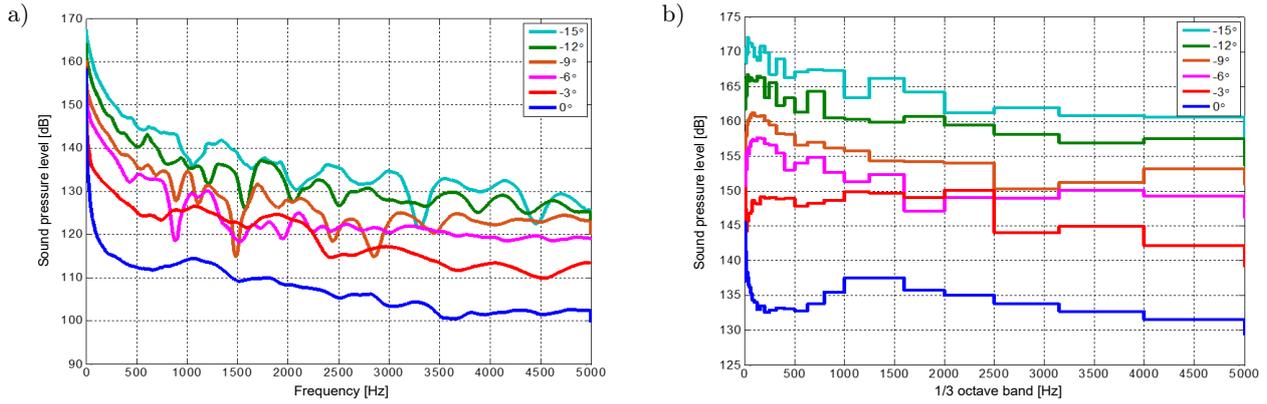


Fig. 6. Sound pressure frequency spectrum curves of R6 at 0°, -3°, -6°, -9°, -12°, -15°.

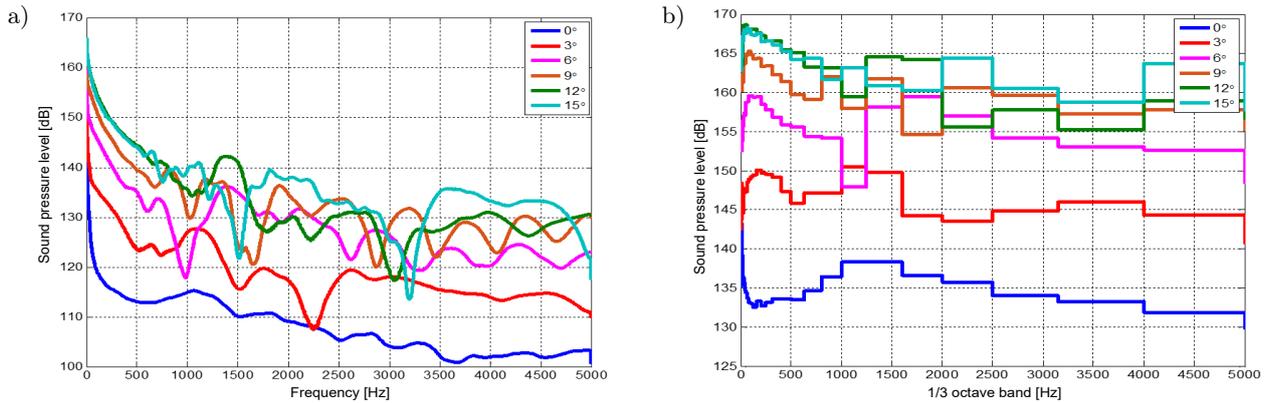


Fig. 7. Sound pressure frequency spectrum curves of R7 at 0°, 3°, 6°, 9°, 12°, 15°.

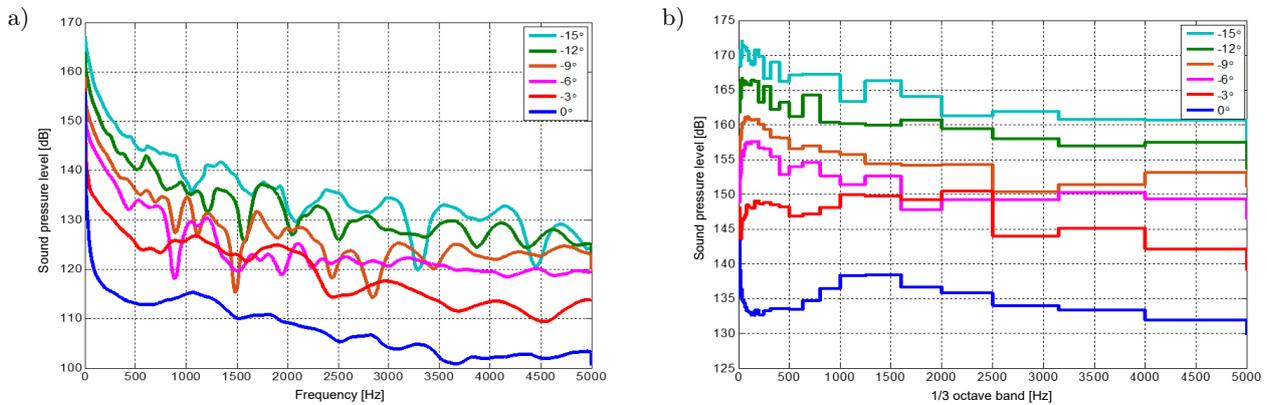


Fig. 8. Sound pressure frequency spectrum curves of R7 at 0°, -3°, -6°, -9°, -12°, -15°.

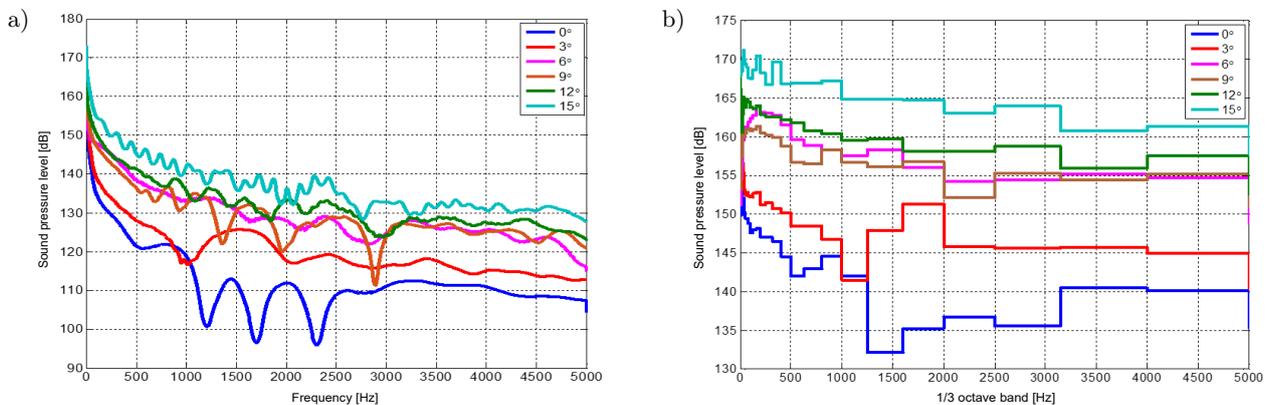


Fig. 9. Sound pressure frequency spectrum curves of R1 at 0°, 3°, 6°, 9°, 12°, 15°.

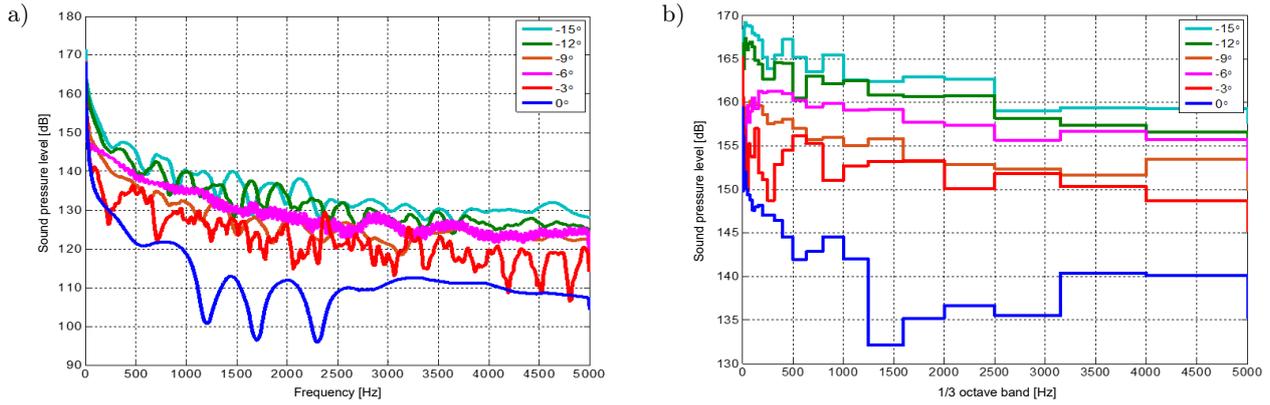


Fig. 10. Sound pressure frequency spectrum curves of R1 at 0° , -3° , -6° , -9° , -12° , -15° .

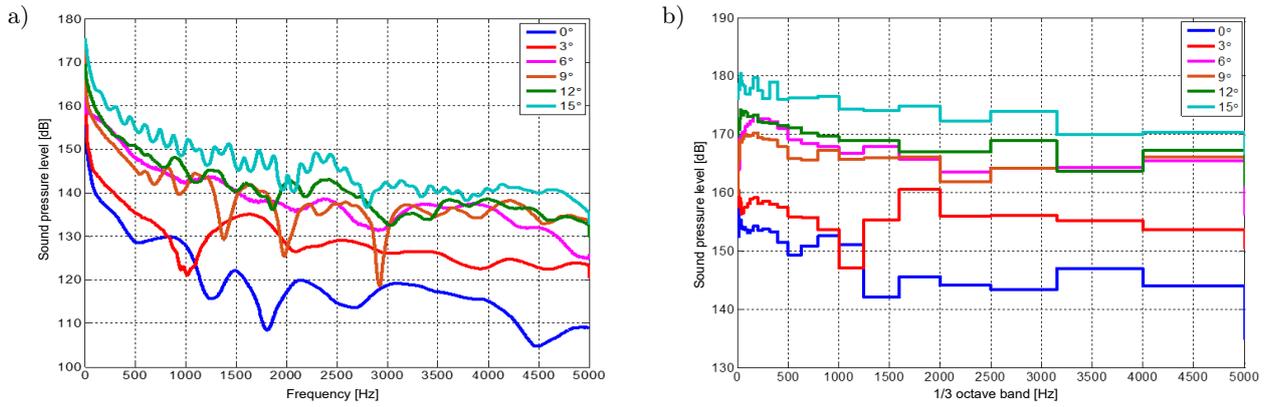


Fig. 11. Sound pressure frequency spectrum curves of R6 at 0° , 3° , 6° , 9° , 12° , 15° .

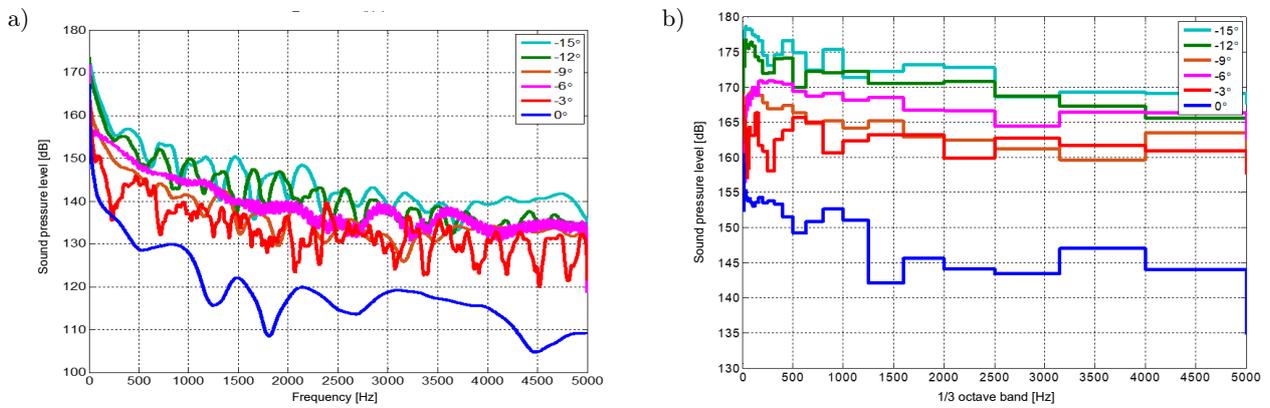


Fig. 12. Sound pressure frequency spectrum curves of R6 at 0° , -3° , -6° , -9° , -12° , -15° .

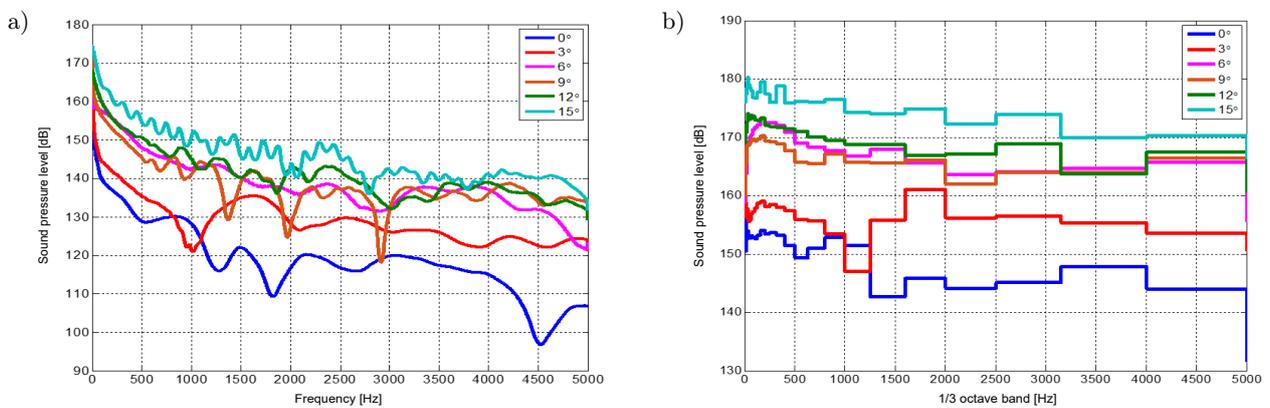


Fig. 13. Sound pressure frequency spectrum curves of R7 at 0° , 3° , 6° , 9° , 12° , 15° .

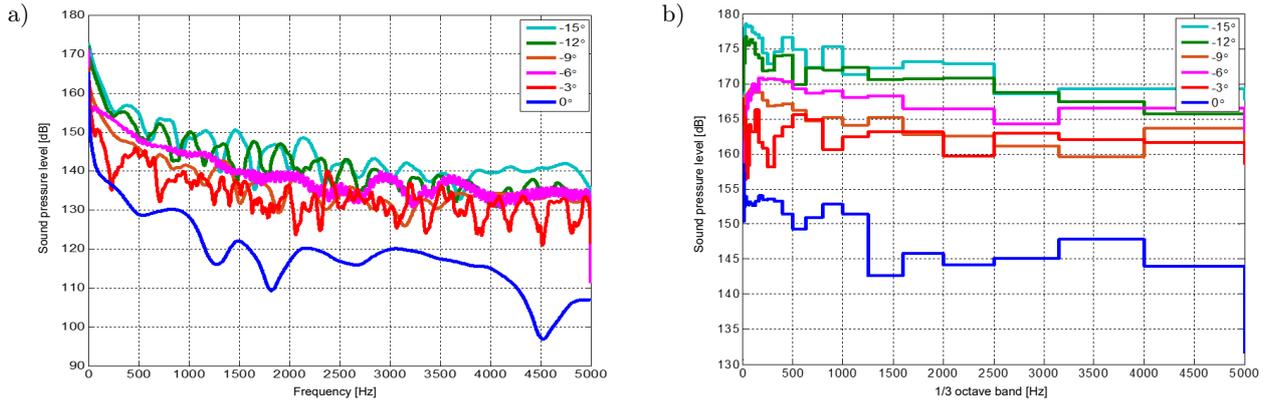


Fig. 14. Sound pressure frequency spectrum curves of R7 at 0°, -3°, -6°, -9°, -12°, -15°.

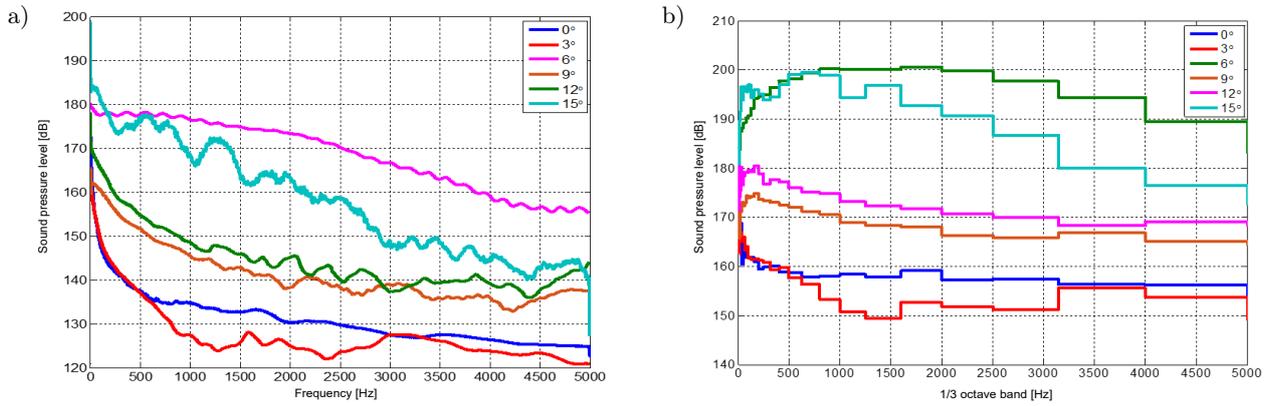


Fig. 15. Sound pressure frequency spectrum curves of R1 at 0°, 3°, 6°, 9°, 12°, 15°.

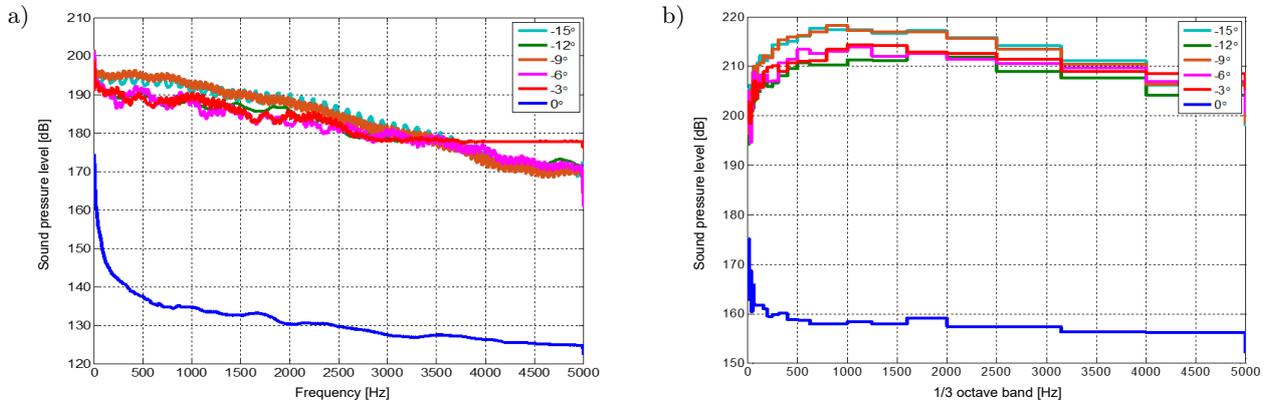


Fig. 16. Sound pressure frequency spectrum curves of R1 at 0°, -3°, -6°, -9°, -12°, -15°.

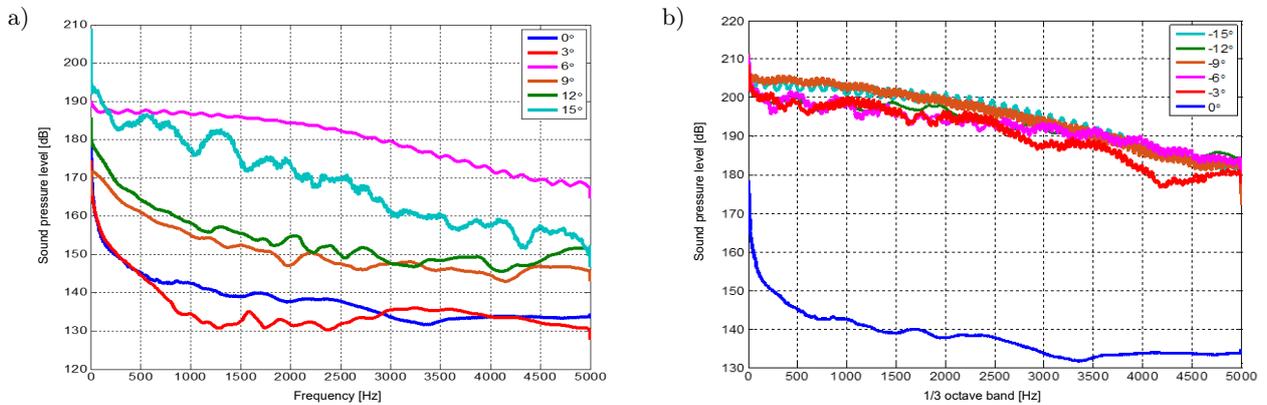


Fig. 17. Sound pressure frequency spectrum curves of R6 at 0°, 3°, 6°, 9°, 12°, 15°.

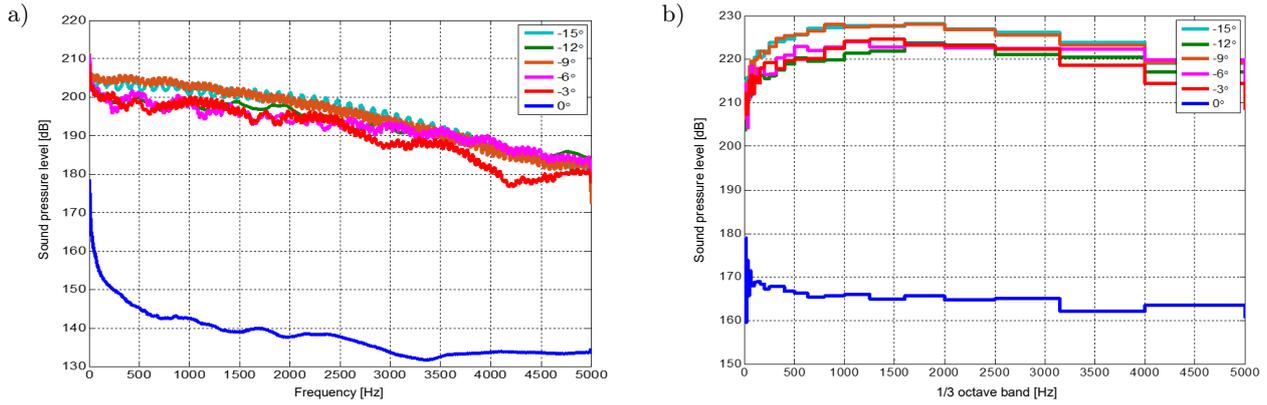


Fig. 18. Sound pressure frequency spectrum curves of R6 at 0°, -3°, -6°, -9°, -12°, -15°.

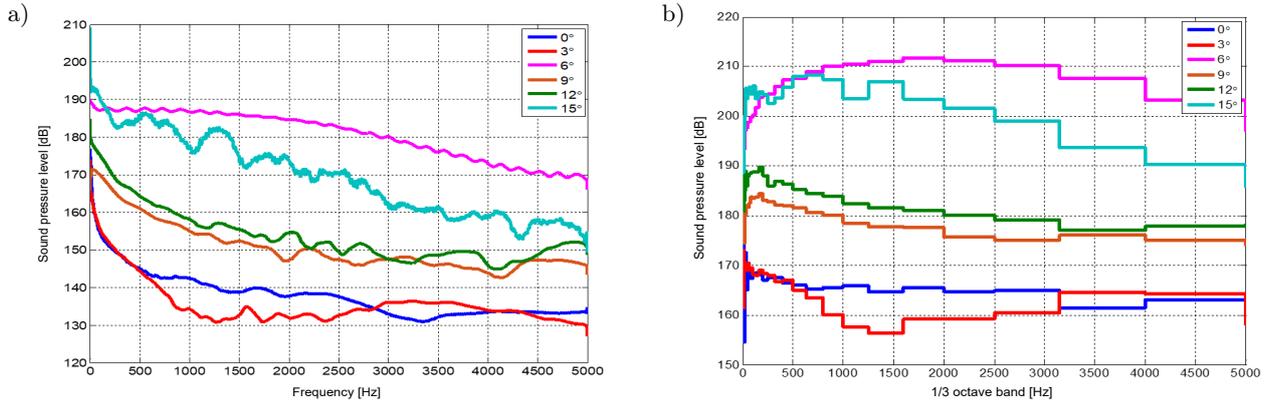


Fig. 19. Sound pressure frequency spectrum curves of R7 at 0°, 3°, 6°, 9°, 12°, 15°.

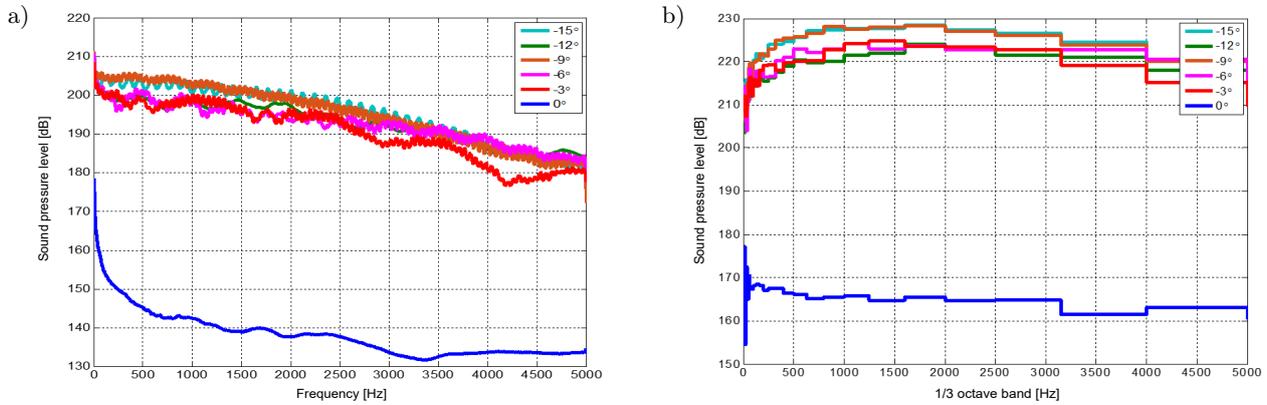


Fig. 20. Sound pressure frequency spectrum curves of R7 at 0°, -3°, -6°, -9°, -12°, -15°.

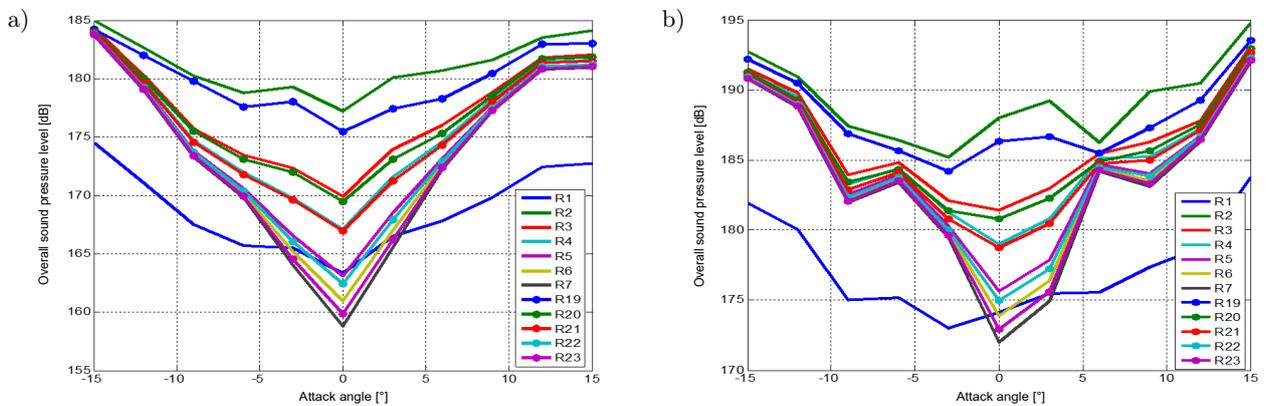


Fig. 21. The change law of overall sound pressure level with the angle at the monitoring points of the surface of the hydrofoil: a) 39.533 m/s; b) 10.288 m/s.

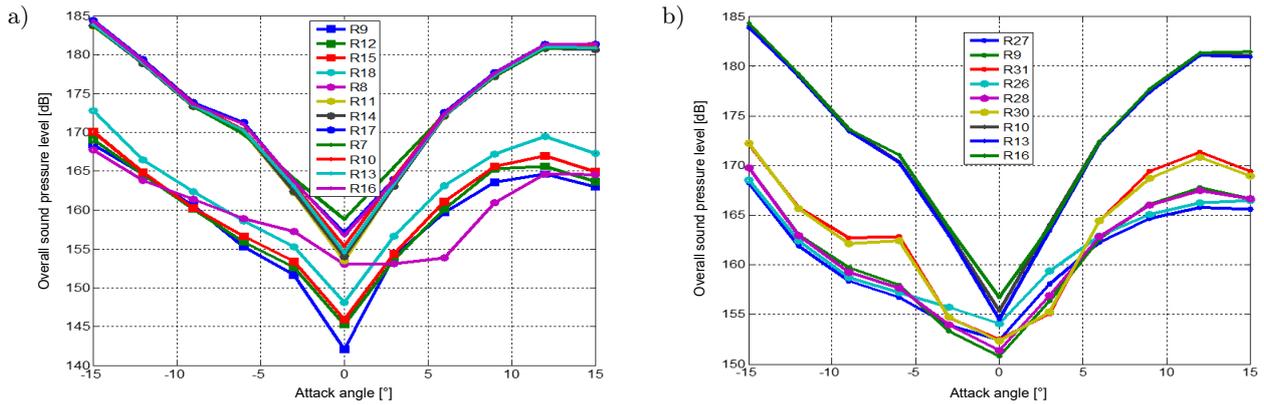


Fig. 22. The change law of overall sound pressure level with the angle at the monitoring points of the horizontal axis and below the tail of the hydrofoil (5.33 m/s).

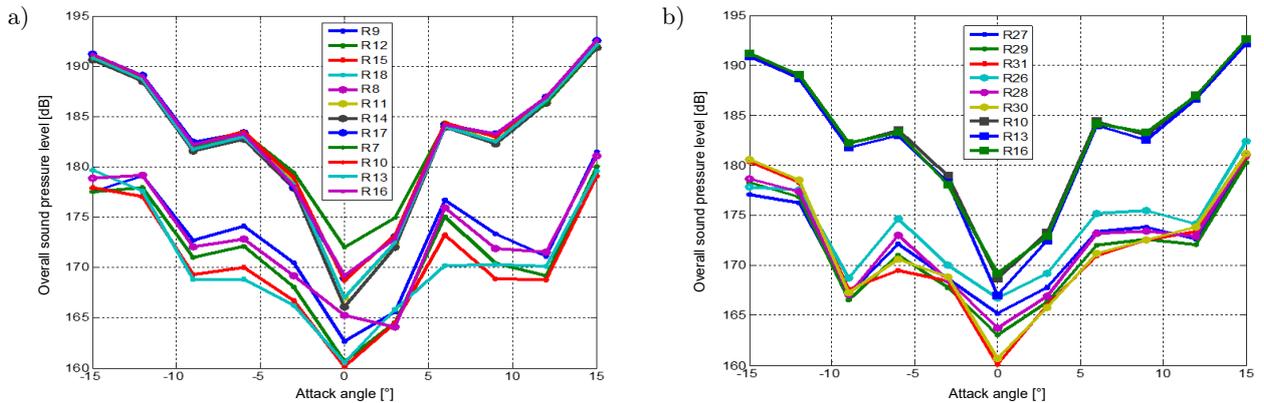


Fig. 23. The change law of overall sound pressure level with the angle at the monitoring points of the horizontal axis and below the tail of the hydrofoil (10.288 m/s).

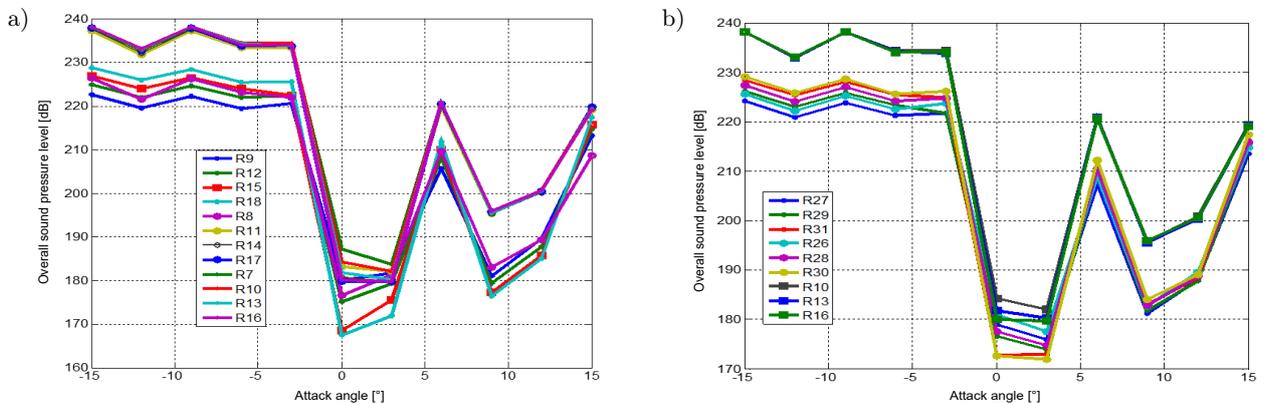


Fig. 24. The change law of overall sound pressure level with the angle at the monitoring points of the horizontal axis and below the tail of the hydrofoil (20.577 m/s).

The curve above shows that the points above the horizontal axis increase with the distance of the vertical axis, and the sound pressure level decreases gradually: $SPL(R7 > R8 > R9)$, $SPL(R10 > R11 > R12)$, $SPL(R13 > R14 > R15)$, $SPL(R16 > R17 > R18)$. At the points on the suction side of the hydrofoil, the sound pressure levels at monitoring points R2~R7 increase with the increase in the absolute value of the angle, and they reach the minimum value when the attack angle becomes 0°. The sound pressure levels at R2~R7 grad-

ually decrease with the increase in the distance from the monitoring point at R1. The condition is the same as the pressure side.

The results in working conditions 1, 2, and 3 show that, for a specific cavitation number, the sound pressure level at each monitoring point increases with the increase in inflow velocity. In working conditions 1 and 2, the sound pressure levels at the monitoring points on the surface of the hydrofoil and the monitoring points at the tail of the hydrofoil are basically

equivalent at the positive and negative angles of attack, and 0° is the axis of symmetry. In working condition 1, the sound pressure level is attenuated by $-6\text{ dB}/3^\circ$ at the negative angle of attack, while at the positive angle of attack, the sound pressure level increases by $6\text{ dB}/3^\circ$ at monitoring points above the horizontal axis of the tail of the hydrofoil and increases by $3\text{ dB}/3^\circ$ at the points below the horizontal axis of the tail. Moreover, the sound pressure values clearly fluctuate at monitoring points R7, R9, R17, R18, R30, and R31. In working condition 2, the sound pressure level is attenuated by $-6\text{ dB}/3^\circ$ at the negative angle of attack, while at the positive angle, the sound pressure level increases by $4\text{ dB}/3^\circ$, at the monitoring points R7, R8, R9, and R10 of 6° to 12° , the sound pressure level is attenuated by $2.5\text{ dB}/3^\circ$. In working condition 3, with the increase in velocity, the sound pressure level at each monitoring point at the tail of the hydrofoil remains relatively consistent at negative attack angles of -15° to -3° , while there is a sharp attenuation of about 50 dB at -3° to 0° . At the positive angle of attack, the overall sound pressure level changes relatively drastically, especially at $3^\circ\sim 6^\circ$, and the overall sound pressure level rapidly increases by about 35 dB ; how-

ever, it rapidly decreases by about 30 dB at $6^\circ\sim 9^\circ$ and increases by about 20 dB at $12^\circ\sim 15^\circ$.

The computational figures of the power spectral density at monitoring points R2, R7, R18, and R31 in working condition 2 are shown in Fig. 25.

From the analysis of the power spectral density at each monitoring point at the surface or the tail of the hydrofoil, it can be observed that the power spectral densities near the surface of the hydrofoil at points R1~R6 are distinctively larger than those above at points R7~R18 for a positive attack angle and also greater than those at points R25~R31 below the horizontal axis of the hydrofoil tail. However, the power spectral densities below the horizontal axis of the hydrofoil tail are greater than those above the axis at equidistant monitoring points, which indicates that the energy near the hydrofoil surface is relatively high. With the increase in the attack angle, the power spectral density increases correspondingly.

At the negative angle of attack, the power spectral density near the surface of the hydrofoil at points R19~R23 is much higher than that at R7~R16 (above the horizontal axis of the hydrofoil tail) and greater than that at R24~R31 (below the axis of the hydrofoil

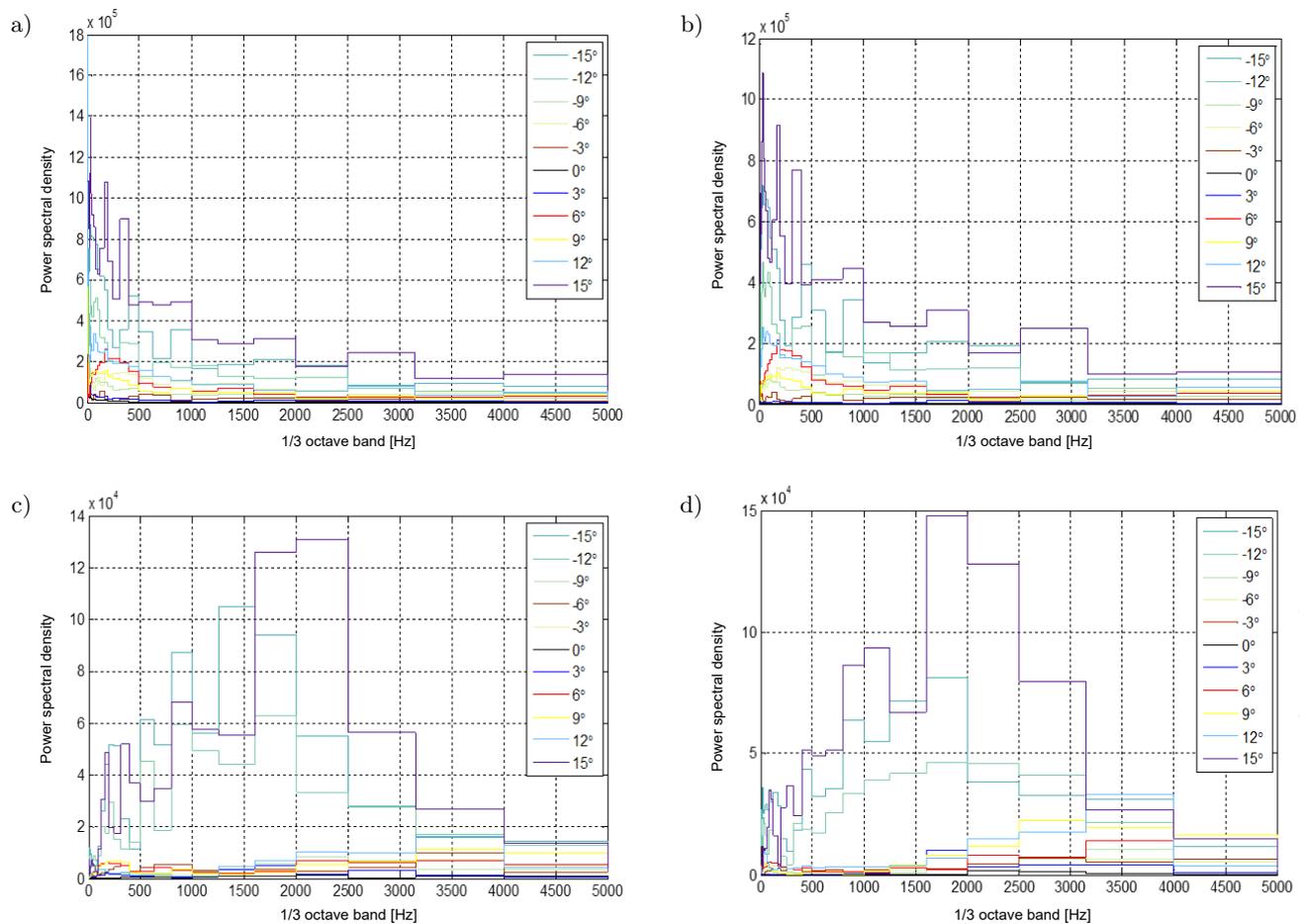


Fig. 25. Power spectral density: a) R2 point; b) R7 point; c) R17 point; d) R31 point.

tail). The power spectral density above the horizontal axis of the hydrofoil tail is greater than that below its horizontal axis at equidistant monitoring points, which indicates that the energy near the hydrofoil surface is relatively high, and the power spectral density increases with the increase in the absolute value of the attack angle.

4. Conclusions

The noise performance of a hydrofoil was numerically predicted and analyzed, and the characteristics of the sound pressure spectrum, sound power spectrum and noise changes at different monitoring points were determined. The noise characteristics and change law of the NACA66 hydrofoil with a specific cavitation number were analyzed. Cavitation bubbles experienced a periodic pulsating process of inception, development, fracture, falling-off, and bursting. With a constant cavitation number, the cavitation area of the foil becomes longer and thicker with the increase in the attack angle, and the initial position of cavitation inception moves forward. As the inflow velocity increases, changes in the cavitation noise and region become more drastic. The results in working conditions show that, for a specific cavitation number, the sound pressure level at each monitoring point increases with the increase in inflow velocity.

The change law of noise was analyzed at each monitoring point of the surface and tail of the NACA66 hydrofoil at different inflow velocities and positive and negative attack angles. The shape of cavitation bubbles has a great influence on the acoustic signal signature of the hydrofoil, particularly at the tail of the hydrofoil, and has a significant effect on the noise of the flow field at the tail. Moreover, because the monitoring points may stack, counteract or interfere with each other, the acoustic signal signatures are weakened or locally reinforced. The novelty is that the study has important calculations and analyses regarding the noise performance of a hydrofoil, characteristics of the sound pressure spectrum, and sound power spectrum and noise changes at different monitoring points. The article may be useful for specialists in the field of engineering and physics. This paper can be of interest both as introductory material and as a basis for further study.

References

1. ALKISHRIWI N., MEINKE M., SCHRÖDER W. (2008), Large eddy simulation of steam wise-rotation turbulent channel flow, *Computer & Fluids*, **37**(7): 786–792, doi: [10.1016/j.compfluid.2007.03.018](https://doi.org/10.1016/j.compfluid.2007.03.018).
2. AL-OBAIDI A.R. (2019), Numerical investigation of flow field behaviour and pressure fluctuations within an axial flow pump under transient flow pattern based on CFD analysis method, *Journal of Physics: Conference Series*, **1279**(1): 012069, doi: [10.1088/1742-6596/1279/1/012069](https://doi.org/10.1088/1742-6596/1279/1/012069).
3. AL-OBAIDI A.R. (2020), Influence of guide vanes on the flow fields and performance of axial pump under unsteady flow conditions: Numerical study, *Journal of Mechanical Engineering and Sciences*, **14**(2): 6570–6593, doi: [10.15282/jmes.14.2.2020.04.0516](https://doi.org/10.15282/jmes.14.2.2020.04.0516).
4. AL-OBAIDI A.R., MISHRA R. (2020), Experimental investigation of the effect of air injection on performance and detection of cavitation in the centrifugal pump based on vibration technique, *Arabian Journal for Science and Engineering*, **45**(7): 5657–5671, doi: [10.1007/s13369-020-04509-3](https://doi.org/10.1007/s13369-020-04509-3).
5. BELJATYNSKIJ A. PRENTKOVSKIS O., KRIVENKO J. (2010), The experimental study of shallow flows of liquid on the airport runways and automobile roads, *Transport*, **25**(4): 394–402, doi: [10.3846/transport.2010.49](https://doi.org/10.3846/transport.2010.49).
6. CAO H., FANG S., ZHU Z. (2014), Numerical prediction and characteristic analysis of slow varying components of cavitation noise of propeller in non-uniform inflow, *High-tech Communications*, **24**(3): 296–304, doi: [10.3772/j.issn.1002-0470.2014.03.012](https://doi.org/10.3772/j.issn.1002-0470.2014.03.012).
7. CHANG F.N. (2011), The prediction of the viscous cavitation flow around a hydrofoil section, [in:] *Proceedings of the Twenty First International Off-shore and Polar Engineering Conference*, **1**: 989–994, <https://onepetro.org/ISOPEIOPEC/proceedings-abstract/ISOPE11/All-ISOPE11/ISOPE-I-11-423/13400> (access: 21.09.2022).
8. CURLE N. (1955), The influence of solid boundaries upon aerodynamic sound, *Proceedings of the Royal Society of London, Series A*, **231**(1187): 505–514, doi: [10.1098/rspa.1955.0191](https://doi.org/10.1098/rspa.1955.0191).
9. DUCOIN A., ASTOLFI J.A., DENISET F., SIGRIST J.F., (2009), Computational and experimental investigation of flow over a transient pitching hydrofoil, *European Journal of Mechanics – B/Fluids*, **28**(6): 728–743, doi: [10.1016/j.euromechflu.2009.06.001](https://doi.org/10.1016/j.euromechflu.2009.06.001).
10. DUCOIN A., ASTOLFI J.A., GOBERT M.L. (2012), An experimental study of boundary-layer induced vibrations on a hydrofoil, *Journal of Fluids and Structures*, **32**(3): 37–51, doi: [10.1016/j.jfluidstructs.2011.04.002](https://doi.org/10.1016/j.jfluidstructs.2011.04.002).
11. FAN G. (2015), *Research on Vibration and Acoustic Radiation Characteristics of Hydrofoil Flow*, Shanghai, School of Mechanics and Power, Shanghai Jiao Tong University.
12. FUJI A., KAWAKAMI D.T., TSUJIMOTO Y., ARNDT R.E.A. (2007), Effect of hydrofoil shapes on partial and transitional cavity oscillations, *Journal of Fluids Engineering*, **129**(6): 669–673, doi: [10.1115/1.2734183](https://doi.org/10.1115/1.2734183).
13. HONG F., YUAN J.P., ZHOU B.L. (2017), Application of a new cavitation model for computations of unsteady turbulent cavitating flows around a hydrofoil, *Journal of Mechanical Science and Technology*, **31**(1): 249–260, doi: [10.1007/s12206-016-1227-x](https://doi.org/10.1007/s12206-016-1227-x).

14. HUANG B., WU Q., WANG G.-Y. (2018), Progress and prospects of investigation into unsteady cavitating flows [in Chinese], *Drainage and Irrigation Mechanical Engineering*, **36**(1): 1–14, doi: [10.3969/j.issn.1674-8530.17.0094](https://doi.org/10.3969/j.issn.1674-8530.17.0094).
15. JI B., LOU X.-W., WU Y.-L., LIU S.-H., XU H.-Y., OSHIMA A. (2010), Numerical investigation of unsteady cavitating turbulent flow around a full-scale marine propeller, *Journal of Hydrodynamics*, **22**(5): 747–752, doi: [10.1016/S1001-6058\(10\)60025-X](https://doi.org/10.1016/S1001-6058(10)60025-X).
16. KIELDSSEN M., ARNDT R.E.A., EFFERTZ M. (2000), Spectral characteristics of sheet/cloud cavitation, *Journal of Fluids Engineering*, **122**(3): 481–487, doi: [10.1115/1.1287854](https://doi.org/10.1115/1.1287854).
17. KORZHYK V., BUSHMA O., KHASKIN V., DONG C., SYDORETS V. (2017), Analysis of the current state of the processes of hybrid laser-plasma welding, [in:] *Proceedings of the 2nd International Conference on Mechanics, Materials and Structural Engineering (ICMMSE 2017) “AER-Advances in Engineering Research”*, pp. 80–90, doi: [10.2991/icmmse-17.2017.14](https://doi.org/10.2991/icmmse-17.2017.14).
18. KUBOTA A., KATO H., YAMAGUCHI H. (1992), A new modeling of cavitating flows: a numerical study of unsteady cavitation on a hydrofoil section, *Journal of Fluid Mechanics*, **240**(1): 59–96, doi: [10.1017/S00221209200003X](https://doi.org/10.1017/S00221209200003X).
19. LEROUX J.B., ASTOLFI J.A., BILLARD J.Y. (2004), An experimental study of unsteady partial cavitation, *Journal of Fluids Engineering*, **126**(1): 94–101, doi: [10.1115/1.1627835](https://doi.org/10.1115/1.1627835).
20. LEROUX J.B., OLIVIER C.D., ASTOLFI J.A. (2003), A joint experimental and numerical study of mechanisms associated to instability of partial cavitation on two-dimensional hydrofoil, *Physics of Fluids*, **17**(5): 052101, doi: [10.1063/1.1865692](https://doi.org/10.1063/1.1865692).
21. LI G., SHI L. (1997), Cavitation and cavitation erosion and their influencing factors [in Chinese], *Journal of Petroleum University (Natural Science Edition)*, **21**(1): 97–101, http://www.alljournals.cn/view_abstract.aspx?pcid=3C5CA5E51F7D0F8A&cid=3C5CA5E51F7D0F8A&jid=AFEA5AA653DB650F7CB44EE42E99DE56&aid=4D3C8AFBB3796195B9EA80098C763CE8&yid=5370399DC954B911&vid=&iid=&sid=&eid=&from_abstract=1 (access: 11.09.2022).
22. LORD RAYLEIGH O.M. F.R.S. (1917), VIII. On the pressure developed in a liquid during the collapse of a spherical cavity, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **34**(200): 94–98, doi: [10.1080/14786440808635681](https://doi.org/10.1080/14786440808635681).
23. NEPPIRAS E.A., NOLTINGK B.E. (1951), Cavitation produced by ultrasonics: Theoretical conditions for the onset of cavitation, *Proceedings of The Physical Science Society, Section B*, **64**(12): 1032–1038, doi: [10.1088/0370-1301/64/12/302](https://doi.org/10.1088/0370-1301/64/12/302).
24. NOLTINGK B.E., NEPPIRAS E.A. (1950), Cavitation produced by ultrasonics, *Proceedings of the Physical Society, Section B*, **63**(9): 674–685, doi: [10.1088/0370-1301/63/9/305](https://doi.org/10.1088/0370-1301/63/9/305).
25. NURTAS M., BAISHEMIROV Z., TSAY V., TASTANOV M., ZHANABEKOV Z. (2020), Convolutional neural networks as a method to solve estimation problem of acoustic wave propagation in poroelastic media, *News of the National Academy of Sciences of the Republic of Kazakhstan – Physico-Mathematical Series*, **4**(332): 52–60, doi: [10.32014/2020.2518-1726.65](https://doi.org/10.32014/2020.2518-1726.65).
26. PLESSET M. (1949), The dynamics of cavitation bubbles, *Journal of Applied Mechanics*, **16**: 277–282.
27. PRENTKOVSKIS O., TRETJAKOVAS J., ŠVEDAS A., BIELIATYNSKYI A., DANIŪNAS A., KRAYUSHKINA K. (2012), The analysis of the deformation state of the double-wave guardrail mounted on bridges and viaducts of the motor roads in Lithuania and Ukraine, *Journal of Civil Engineering and Management*, **18**(5): 761–771, doi: [10.3846/13923730.2012.731252](https://doi.org/10.3846/13923730.2012.731252).
28. PROKOPOV V.G. et al. (1993), Effect of coating porosity on the process of heat-transfer with gas-thermal deposition, *Powder Metallurgy and Metal Ceramics*, **32**(2): 118–121, doi: [10.1007/BF00560034](https://doi.org/10.1007/BF00560034).
29. SCHNERR G.H., SAUER J. (2001), Physical and numerical modeling of unsteady cavitation dynamics, [in:] *4th International Conference on Multiphase Flow ICMF-2001*, <https://www.semanticscholar.org/paper/Physical-and-numerical-modeling-of-unsteady-Schnerr-Sauer/ed8c0ec44ddc18456eb4f32fb7aaf76605bd13aa> (access: 17.09.2022).
30. SHIN S., HONG J.-W., NAGARATHINAM D., AHN B.-K., PARK S.-G. (2021), Tip vortex cavitation and induced noise characteristics of hydrofoils, *Applied Sciences (Switzerland)*, **11**(13): 5906, doi: [10.3390/app1135906](https://doi.org/10.3390/app1135906).
31. SINGHAL A.K., ATHAVALE M.M., LI H., JIANG Y. (2002), Mathematical basis and validation of the full cavitation model, *Journal of Fluids Engineering*, **124**(3): 617–624, doi: [10.1115/1.1486223](https://doi.org/10.1115/1.1486223).
32. SU Y., DOU F., LIU Y., CUI T. (2013), Study of non-cavitating propeller noise, *Journal of Wuhan University of Technology (Transportation Science and Engineering)*, **37**(5): 895–899, doi: [10.3963/j.issn.2095-3844.2013.05.001](https://doi.org/10.3963/j.issn.2095-3844.2013.05.001).
33. SULTANOV K.S., KHUSANOV B.E., RIKHSIEVA B.B. (2020), Longitudinal waves in a cylinder with active external friction in a limited area, [in:] *IV International Scientific and Technical Conference Mechanical Science and Technology Update (MSTU-2020)*, **1546**: 012140, doi: [10.1088/1742-6596/1546/1/012140](https://doi.org/10.1088/1742-6596/1546/1/012140).
34. WANG G., SENOCAK I., SHYY W., IKOHAGI T., CAO S. (2001), Dynamics of attached turbulent cavitating flow, *Progress in Aerospace Science*, **37**(6): 551–581, doi: [10.1016/S0376-0421\(01\)00014-8](https://doi.org/10.1016/S0376-0421(01)00014-8).
35. WANG G., OSTOJA-STARZEWSKI M. (2007), Large eddy simulation of a sheet/cloud cavitation on a NACA0015 hydrofoil, *Applied Mathematical Modelling*, **31**(3): 417–447, doi: [10.1016/j.apm.2005.11.019](https://doi.org/10.1016/j.apm.2005.11.019).
36. WANG G.-Y., ZHANG B., HUANG B., ZHANG M. (2009), Unsteady dynamics of cloud cavitating flows around

- a hydrofoil, *Journal of Experiments in Fluid Mechanics*, **23**(3): 44–49, <https://www.semanticscholar.org/paper/Unsteady-Dynamics-of-Cloud-Cavitating-Flows-around-Wang-Zhang/0db160f3faecc68c656b97b6453252fba40eeded> (access: 19.09.2022).
37. WANG W., LI Z., LIU M., JI X. (2021), Influence of water injection on broadband noise and hydrodynamic performance for a NACA66 (MOD) hydrofoil under cloud cavitation condition, *Applied Ocean Research*, **115**: 102858, doi: [10.1016/j.apor.2021.102858](https://doi.org/10.1016/j.apor.2021.102858).
38. WU J.Y., WANG G.Y., SHYY W. (2005), Time-dependent turbulent cavitating flow computations with interfacial transport and filter based models, *International Journal for Numerical Methods in Fluids*, **49**(7): 739–761, doi: [10.1002/fld.1047](https://doi.org/10.1002/fld.1047).
39. YANG Q., WANG Y., ZHANG Z. (2012), Assessment of the improved cavitation model and modified turbulence model for ship propeller cavitation simulation, *Journal of Mechanical Engineering*, **48**(9): 178–785, doi: [10.3901/JME.2012.09.178](https://doi.org/10.3901/JME.2012.09.178).
40. YANG Q., WANG Y., ZHANG Z., HOU G. (2014), Numerical prediction of cavitation inception radiated noise of contra-rotating propeller with non-uniform in flow, *Shenghixue Xuebao/Acta Acustica*, **39**(5): 589–604.
41. YANG Q.-F., WANG Y.-S., ZHANG Z.-H. (2011), Improvement and evaluation of numerical model for cavitation flow viscous simulation around propeller blade section, *Transactions of Beijing Institute of Technology*, **31**(12): 1401–1407.
42. ZHANG B., WANG G., HUANG B., ZHIYI Y. (2009), Numerical and experimental studies on unsteady shedding mechanisms of cloud cavitation, *Chinese Journal of Theoretical and Applied Mechanics*, **41**(5): 651–659, doi: [10.6052/0459-1879-2009-5-2008-152](https://doi.org/10.6052/0459-1879-2009-5-2008-152).
43. ZWART P., GERBER A.G., BELAMRI T. (2004), A two-phase flow model for prediction cavitation dynamics, [in:] *5th ICMF 2004 International Conference on Multiphase Flow*.

Research Paper

Design and Experiments of a New Internal Cone Type Traveling Wave Ultrasonic Motor

Ye CHEN^{(1)*}, Junlin YANG⁽¹⁾, Liang LI⁽²⁾, Shihao XIAO⁽¹⁾⁽¹⁾ *Institute of Vibration Engineering, Liaoning University of Technology*
Shiying Street, Guta District, Jinzhou, Liaoning Province, China⁽²⁾ *College of Science, Liaoning University of Technology*
Shiying Street, Guta District, Jinzhou, Liaoning Province, China

*Corresponding Author e-mail: jxchenye@lnut.edu.cn

(received November 25, 2022; accepted March 12, 2023)

In order to simplify the motor structure, to reduce the difficulty of rotor pre-pressure application and to obtain better output performance, a new internal cone type rotating traveling wave ultrasonic motor is proposed. The parametric model of the internal cone type ultrasonic motor was established by the ANSYS finite element software. The ultrasonic motor consists of an internal cone type vibrator and a tapered rotor. The dynamic analysis of the motor vibrator is carried out, and two in-plane third-order bending modes with the same frequency and orthogonality are selected as the working modes. The other advantages of this motor are that pre-pressure can be imposed by the weight of the rotor. The prototype was trial-manufactured and experimentally tested for its vibration characteristics and output performance. When the excitation frequency is 22260.0 Hz, the pre-pressure is 0.1 N and the peak-to-peak excitation voltage is 300 V, the maximum output torque of the prototype is 1.06 N·mm, and the maximum no-load speed can reach 441.2 rpm. The optimal pre-pressure force under different loads is studied, and the influence of the pre-pressure force on the mechanical properties of the ultrasonic motor is analyzed. It is instructive in the practical application of this ultrasonic motor.

Keywords: ultrasonic motors; traveling wave; internal cone type; bending vibration.



Copyright © 2023 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 ultrasonic motor is a new type of micro-technical motor that uses the inverse piezoelectric effect of piezoelectric materials to produce ultrasonic frequency vibration in the vibrator and uses the friction between the vibrator and the rotor to achieve the rotor rotation, linear or multi-degree-of-freedom motion. Ultrasonic motors have the advantages of a simple structure, high power density, fast response, no electromagnetic radiation, and high positioning accuracy (ZHAO, 2011). Therefore, more and more scholars have explored and researched them from the aspects of structure design, drive control principle and friction materials, and have achieved certain results (TIAN *et al.*, 2020; MAKAREM *et al.*, 2021; LI *et al.*, 2019; PUOZA, SAKTHIVELSAMY, 2021; WANG *et al.*, 2020). Now ultrasonic motors have been successfully applied

in precision positioning systems, micro-robotic systems, aerospace engineering, biological engineering and other high-tech industrial settings (LIU *et al.*, 2022; OLSSON *et al.*, 2016; MISHRA *et al.*, 2018; WANG, XU, 2017; OH *et al.*, 2009).

Ultrasonic motors can be divided into standing wave ultrasonic motors and traveling wave ultrasonic motors from the point of view of vibration characteristics. In commercial applications, the latter are widely used because of their high efficiency and simple drive control. From the viewpoint of a motion output, they can be divided into rotary, linear and multi-degree-of-freedom ultrasonic motors (RYNDZIONEK, SIENKIEWICZ, 2021). Among them, rotary ultrasonic motors are more well developed and the technology is more mature. Among various types of ultrasonic motors, squiggle and in-plane bending travelling wave ultrasonic motors are often suitable for miniaturiza-

tion and integration (XU *et al.*, 2021; LU *et al.*, 2020; MASHIMO, OBA, 2022; LI *et al.*, 2021).

An important factor affecting the application of the ultrasonic motor is the overall structural size. Therefore a millimeter scale thick film rotating traveling wave ultrasonic motor based on the chemical mechanical thinning and polishing process is proposed (ZHANG *et al.* 2022). The vibration mode of the motor is the B02 mode under the resonant frequency of 26.2 kHz. The motor can achieve stable bidirectional rotation under the excitation of four sinusoidal voltages. Moreover, when the excitation voltage is 50 V_{p-p}, the maximum speed can reach 766 rpm under the preload force of 0.686 mN. A miniature flat cross-shaped rotating ultrasonic motor was designed and manufactured (ČEPONIS *et al.* 2020). The motor rotates the rotor by exciting the first-order in-plane bending vibration of the cross-shaped vibrator. The results of the experimental study show that the motor has a maximum speed of 972.62 rpm at a peak-to-peak of 200 V when a preload force of 22.65 mN is applied. The miniature cross-shaped motor can be mounted directly to a printed circuit board or integrated into other systems with a limited installation space.

The oblate-type ultrasonic motor, extensively desired in small-scale robotics, fuzing, and biomedical technology, however, has not obtained abundant development. A flat ultrasonic micro-motor with multilayer piezoelectric ceramics and a chamfered driving tip is proposed in order to realize a low-voltage drive for ultrasonic motors (ZHAO *et al.* 2016). The vibrator is fabricated with a multilayer piezoelectric ceramic glued to a copper ring with a thickness of 0.5 mm. There are six driving tips on the copper ring as a whole. The driving tips are chamfered in the proper direction and their height is 1 mm. The motor can work smoothly and reach a rotation speed of about 2000 r/min at a voltage amplitude of 20 V_{p-p}. It shows the characteristics of high speed and low load capacity.

As can be seen from the above-mentioned articles, many authors have paid attention to motor miniaturization and structural innovations. Therefore, this paper proposes an internal cone type rotating traveling wave ultrasonic motor, which consists of an internal cone type vibrator and a tapered rotor, and uses friction to drive the rotor in a rotational motion. The internal cone type vibrator and the tapered rotor are in trapezoidal teeth contact with each other, which facilitates the smooth operation of the motor while having a large output speed and an output torque.

2. Ultrasonic motor structure and working principle

2.1. Ultrasonic motor structure

The structure of the internal cone type ultrasonic motor vibrator is shown in Fig. 1. The internal cone

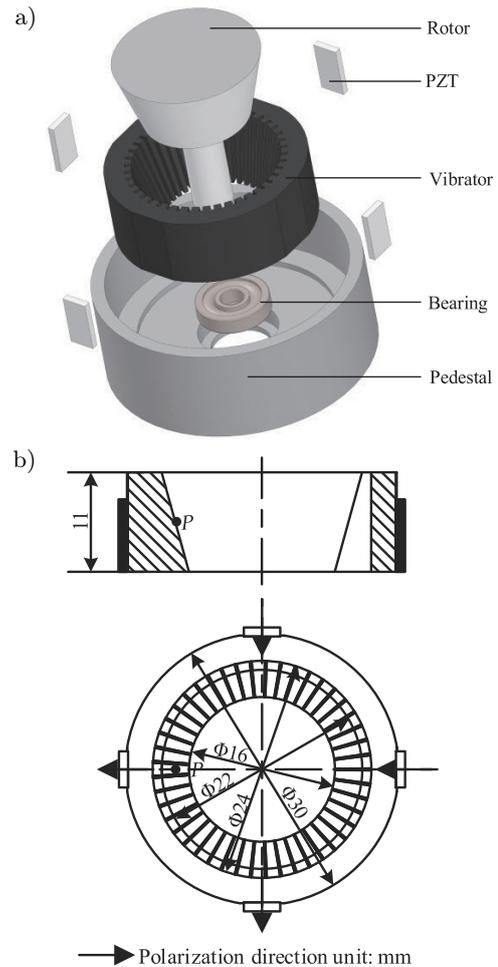


Fig. 1. Structure of ultrasonic motor: a) 3D model of the motor; b) main geometrical parameters of the vibrator.

type vibrator is based on a cylindrical structure with a tapered hole inside. Several uniform inner trapezoidal teeth are designed inside the cylinder, which is conducive to enlarging the amplitude of the inner surface in the circumferential direction. The number of teeth in the vibrator is 45, and the width of the tooth slot is 0.2 mm. Four rectangular piezoelectric ceramic sheets of $8 \times 4 \times 1$ mm are pasted on the outer surface of the internal cone type vibrator. The diameter of the outer cylindrical surface of the internal cone type vibrator of the rotating ultrasonic motor is set to 30 mm.

In Fig. 1, the tapered rotor of the motor and the internal cone type vibrator are in contact with the bevel tooth surface, which is very different from the point contact structure in the contact process between the vibrator and the rotor of the previous motor, which can ensure the stable contact between the vibrator and the rotor and reduce energy loss. And it dissipates heat well, as well as it avoids the problems of unstable operation and small driving torque of the ultrasonic motor in the past. The polarization directions of the two groups of piezoelectric ceramic sheets are shown in Fig. 1.

2.2. Bending vibration of cylindrical shells

The piezoelectric oscillator described in this paper is a thin-walled structure, and its vibration modes can be analyzed by using the cylindrical shell vibration theory. The coordinate system of the cylindrical shell is shown in Fig. 2, which is the radial coordinate, the angular coordinate and the axial coordinate. It is assumed that the vibration displacement is tangential and radial. The displacement distribution of the in-plane vibration mode of the cylindrical shell is a constant along the axial direction (axis), and the displacement distribution along the radial direction (axis) is also considered as a constant due to the thin-walled structure, so each displacement component is a function of the angular coordinate. SOEDEL (2004) proposed the equation for the in-plane free vibration of a cylindrical shell:

$$\frac{\partial^2 x}{\partial \theta^2} + \frac{\delta y}{\delta \theta} + k \left(\frac{\partial^2 x}{\partial \theta^2} - \frac{\partial^3 y}{\partial \theta^3} \right) = -\frac{R^2}{K} \rho h \frac{\partial^2 x}{\partial t^2}, \quad (1)$$

$$\frac{\partial x}{\partial \theta} + y - k \left(\frac{\partial^3 x}{\partial \theta^3} - \frac{\partial^4 y}{\partial \theta^4} \right) = -\frac{R^2}{K} \rho h \frac{\partial^2 y}{\partial t^2}, \quad (2)$$

where $K = \frac{E}{(1-\mu^2)}$ is the short cylinder stiffness, $k = \frac{h^2}{12R^2}$ is the short cylinder correlation constant, h is the radial thickness, R is the neutral plane radius, ρ is the material density, μ is the material Poisson's ratio, E is the material Young's modulus.

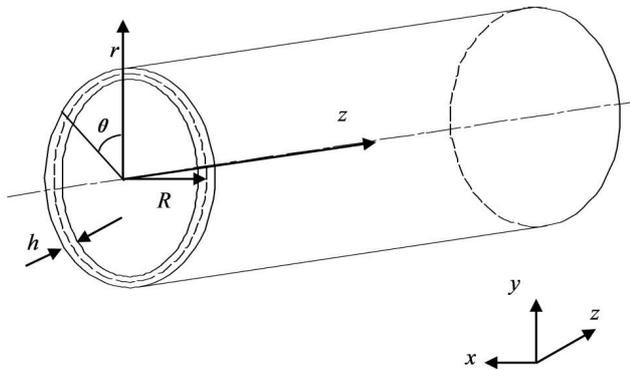


Fig. 2. Coordinates of cylindrical shells.

According to the periodicity of the ring structure, there are solutions of the following form:

$$\begin{pmatrix} x_{n1} \\ y_{n1} \end{pmatrix} = A_{n1} \begin{pmatrix} \sin n\theta \\ B_n \cos n\theta \end{pmatrix} \cos \omega t, \quad (3)$$

$$\begin{pmatrix} x_{n2} \\ y_{n2} \end{pmatrix} = A_{n2} \begin{pmatrix} \cos n\theta \\ -B_n \sin n\theta \end{pmatrix} \cos \omega t, \quad (4)$$

where ω is the circular frequency of the short cylinder, A_{n1} , A_{n2} , B_n are the amplitude coefficients. The aforementioned formula is substituted into the vibra-

tion equation to obtain:

$$\begin{vmatrix} \frac{K}{\rho h R^2} (1+k)n^2 - \omega_n^2 & \frac{K}{\rho h R^2} n(1+kn^2) \\ \frac{K}{\rho h R^2} n(1+kn^2) & \frac{K}{\rho h R^2} (1+kn^4) - \omega_n^2 \end{vmatrix} = 0. \quad (5)$$

Solving Eq. (5) yields:

$$\omega_{n1}^2 = \frac{1}{2} \frac{K}{\rho h R^2} [(1+k)n^2 + kn^4 + \sqrt{a^*}], \quad (6)$$

$$\omega_{n2}^2 = \frac{1}{4} \left(\frac{K}{\rho h R^2} \right)^2 [(1+k)n^2 + kn^4 - \sqrt{a^*}], \quad (7)$$

where

$$a^* = [(1+k)n^2 + kn^4]^2 + 4n^2(1+kn^2)^2 - 4n^2(1+k)(1+kn^4).$$

Combining Eqs. (3) and (4) yields:

$$B_{n1} = \frac{\rho h \omega_{n1}^2 R^2 - n^2 K (1+k)}{K n (1+n^2 k)}, \quad (8)$$

$$B_{n2} = \frac{\rho h \omega_{n2}^2 R^2 - n^2 K (1+k)}{K n (1+n^2 k)}, \quad (9)$$

where ω_{n1} is the intrinsic frequency of the n -th-order in-plane expansion mode, ω_{n2} is the intrinsic frequency of the n -th-order in-plane bending mode. B_{n1} , B_{n2} are the in-plane bending modes of the short cylinder.

2.3. Principle of operation

Figure 3 shows the working principle of the inner cone ultrasonic motor proposed in this paper. The internal cone type vibrator structure has a certain symmetry. When the two-phase piezoelectric ceramic sheets arranged at 90° intervals are excited by the sine and cosine excitation voltages, respectively, the vibrator will generate a third-order bending resonance, and the vibrator will be excited:

$$w_A = W \cos n\theta \cos \omega t, \quad (10)$$

$$w_B = W \sin n\theta \sin \omega t. \quad (11)$$

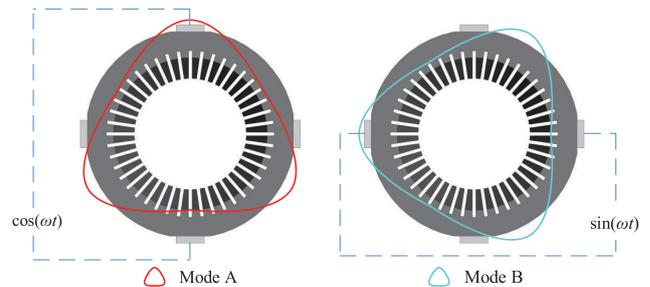


Fig. 3. Working principle of internal cone type ultrasonic motor.

A and B two-phase standing waves are superimposed on the vibrator to obtain bending traveling waves:

$$w = w_A + w_B = W \cos(n\theta - \omega t), \quad (12)$$

where W is the amplitude of the vibration of the A and B phases, n is the modal order of the bending vibration, θ is the angular coordinate along the circumferential direction, ω is the natural frequency of the third-order bending mode.

For the third-order bending mode, the two-phase ceramic sheets are separated by three-quarter wavelengths in this paper. When the two standing waves with equal amplitudes excited by the two modes A and B have a phase difference of $\pi/2$ in time, they will be superimposed on the internal cone type vibrator to form a traveling wave running in the circumferential direction. After the traveling wave is formed on the internal cone type vibrator, the two orthogonal in-plane third-order bending modes of the same frequency are superimposed on each other to generate an elliptical motion trajectory on the particle on the inner tooth surface. Finally, under the action of a certain pre-pressure, the rotary motion of the tapered rotor is realized through the friction coupling between the inner teeth and the tapered rotor.

3. Finite element simulation of piezoelectric vibrator

In this paper, modal and harmonic response analyses were performed with the help of the ANSYS finite element software to design and build an internal cone type vibrator model. Figure 4 shows the two third-order bending vibration patterns of the designed tapered vibrator under free boundary conditions.

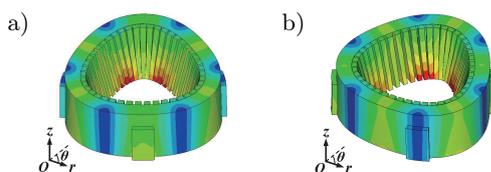


Fig. 4. Third-order bending mode of piezoelectric vibrator: a) B03 A-phase mode; b) B03 B-phase mode.

In selecting the vibrator vibration mode, the modal analysis results show that the vibrator is not only orthogonal but also similar in frequency in the third-order bending resonance mode. At the same time, the amplitude of the low-order mode is larger than that of

the high-order mode in terms of vibration strength, so the low-order mode is usually chosen. Finally, the two orthogonal third-order bending modes of the piezoelectric vibrator are selected to have intrinsic frequencies of $f_A = 22959.7$ Hz and $f_B = 22960.8$ Hz.

In order to ensure that the vibrator does not have interference modes in a certain wide working frequency band, the ANSYS finite element software is used to analyze the harmonic response of the ultrasonic motor vibrator. An excitation signal with a peak value of 40 V and the frequency range of 20000 Hz to 25000 Hz was applied to the two sets of ceramic sheets, respectively. The amplitude-frequency characteristics of the vibrator are obtained through the analysis and solution of the post-processing module of the ANSYS finite element software. The amplitude displacement peak appeared at the frequency of 22960 Hz, and no other amplitude displacement peaks appeared in the frequency range 20000~25000 Hz. The results show that the vibrator has no interference mode in the frequency range, which verifies that the motor has good stability in a wide frequency band. The analysis results of the A and B phases are shown in Fig. 5.

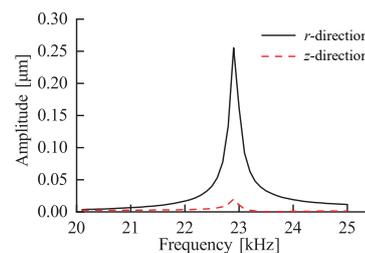


Fig. 5. Harmonic response analysis of vibrator.

4. Experimental study of ultrasonic motor

4.1. Prototype ultrasonic motor

The prototype of the cone type ultrasonic motor was made according to the structural dimensions given in Fig. 1. The vibrator material is 45# steel (high quality carbon structural steel with a carbon content of 0.45%), and the motor vibrator is boiled black in order to prevent the vibrator from being corroded by long working hours. Under certain pre-pressure, four rectangular PZT-81 piezoelectric ceramic sheets polarized along the thickness direction were attached to the four positioning slots on the outer cylindrical surface of the vibrator using epoxy resin. The length of PZT-81 piezoelectric ceramic sheet is 8 mm, the width is 4 mm, and the thickness is 1 mm. The detailed parameters are shown in Table 1. The bottom edge

Table 1. Properties of the piezoelectric ceramic sheet.

Ceramic	Piezoelectric strain factor d_{33} [$\text{pC} \cdot \text{N}^{-1}$]	Relative permittivity ϵ_r	Planar electromechanical coupling coefficient k_p [%]	Thickness electromechanical coupling coefficient k_t [%]	Mechanical quality factor Q_m
PZT-81	300	1350	0.58	0.50	1000

of the rectangular piezoelectric ceramic sheet is aligned with the end of the small aperture of the vibrator. In order to reduce the wear on the vibrator during the long working hours of the motor, the tapered rotor material is 2A12 (series 2 aluminum alloy with serial number 12) with a weight of 10 g. The prototype is shown in Fig. 6.

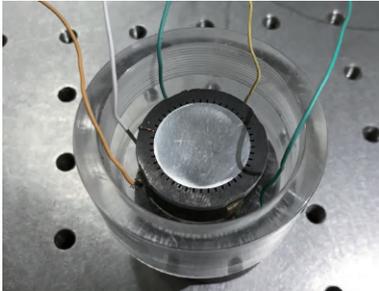


Fig. 6. Photograph of internal cone type ultrasonic motor principle prototype.

4.2. Ultrasonic motor vibrator test experiment

The vibration characteristics of the internal cone type vibrator were tested by the arbitrary waveform/function signal generator Tektronix AFG320, the 2713 Power Amplifier from B&K Denmark, the Germany Polytec OFV-505/5000 Laser Vibrometer, the multi-channel high frequency digital storage oscilloscope Agilent DS06014A and the precision vibration isolation platform as shown in Fig. 7. The frequency sweep test was carried out on the amplitude distribution of the midpoint P of the tooth structure end face of the inner tooth surface of the vibrator, as shown in Fig. 8.

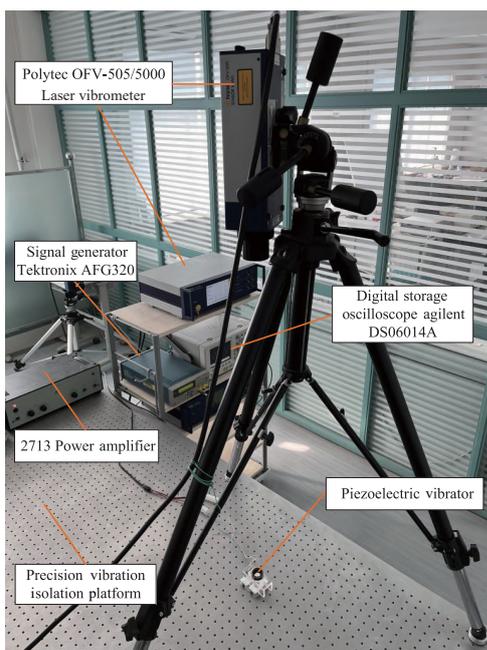


Fig. 7. Photograph of the experimental test instrument.

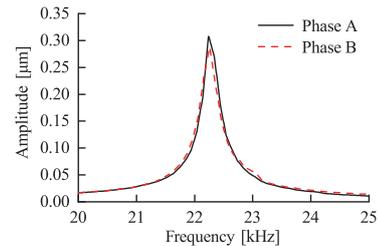


Fig. 8. Amplitude and frequency characteristics of the vibrator near the resonance point.

The experimental results show that the resonance frequencies of the two third-order bending modes of the vibrator are 22248.5 Hz and 22260 Hz, respectively, while the resonance frequencies obtained by modal analysis are 22959.7 Hz and 22960.8 Hz, respectively. The frequency difference between the two is 711.7 Hz and 700.8 Hz respectively, and the errors are 3.09% and 3.05%, respectively. The frequency of the third-order bending mode is basically consistent with the numerical simulation results of ANSYS software. The vibrator has no other interference modes in the frequency range of 20000~25000 Hz.

The amplitude distribution of the midpoint P of the tooth structure end face of the inner tooth surface of the vibrator was tested by a vibration testing instrument. The actual vibration measurement results are shown in Fig. 9. When the excitation frequency is 22260 Hz, the in-plane third-order bending mode of the vibrator can be well excited, and the vibrator can realize the expected traveling wave motion, which also proves the feasibility of the motor.

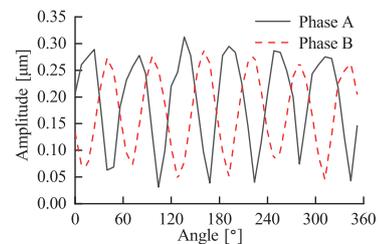


Fig. 9. 360° amplitude distribution of piezoelectric vibrator.

4.3. Ultrasonic motor vibrator test experiment

The output characteristics test rig was built (Fig. 10). The output characteristics of the motor are experimentally tested when the excitation voltage peak-to-peak value is 300 V and the excitation frequency is 22260 Hz using the multi-function driver. In the experimental test, a photoelectric tachometer was used to measure the rotational speed of the tapered rotor under different excitation voltages. When the excitation voltage peak-to-peak value is 300 V, the pre-pressure is 0.1 N, the excitation frequency is 22260 Hz, and the excitation voltage is increased to 300 V, the no-load speed of the ultrasonic motor can reach up to 441.2 rpm, as shown in Fig. 11.

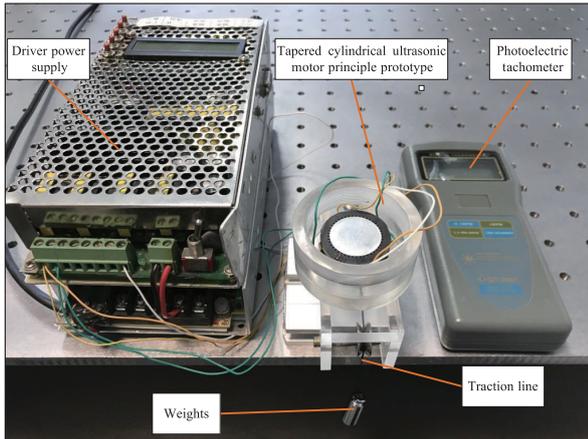


Fig. 10. Photograph of the output characteristics test rig.

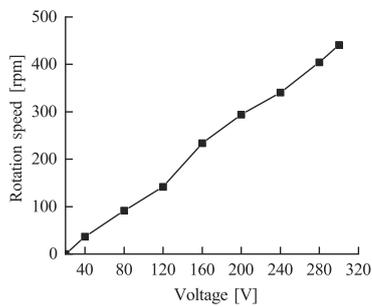


Fig. 11. No-load speed at different excitation voltages.

In the experimental test of torque and rotational speed, the magnitude of the torque is adjusted by lifting weights of different masses by the tapered rotor, while the rotational speed is still measured with a photoelectric tachometer. When the peak-to-peak value of the excitation voltage is 300 V, the pre-pressure is 0.1 N, and the excitation frequency is 22260 Hz, the motor speed decreases smoothly with the increase of torque, which is approximately linear. The maximum output torque of the motor is 1.06 N·mm, as shown in Fig. 12.

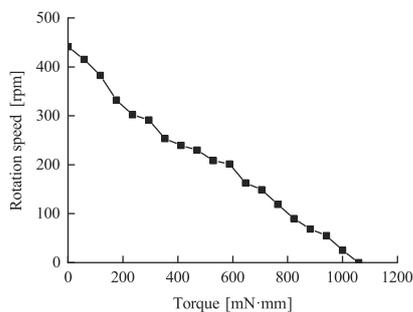


Fig. 12. Motor speed at different torques.

4.4. Ultrasonic motor pre-pressure analysis

The optimum ultrasonic motor pre-pressure depends on the design parameters and operating torque of the motor. When assembling a motor, choosing different pre-pressure for specific operating conditions and load

torques will positively affect its efficiency and performance. The test was performed with the output characteristics test rig above (Fig. 10). The platform is capable of applying loads and pre-pressure forces and testing the corresponding speeds. In the experimental tests of pre-pressure, load torque and rotational speed, the pre-pressure was adjusted by changing the weight and load of the tapered rotor. The torque is regulated by tapered rotors that lift different masses, while speed is still measured by a photoelectric tachometer.

The motor speed decreases as the load torque increases until the motor is locked. On the other hand, the motor speed increases as the pre-pressure increases and then decreases, as shown in Fig. 13. As can be seen from the figure, the pre-pressure and load torque do not affect the motor speed independently; the coupling between pre-pressure and load torque is as follows: as the load torque increases, the value of the optimal motor pre-pressure corresponds to the inflection point of the speed increase.

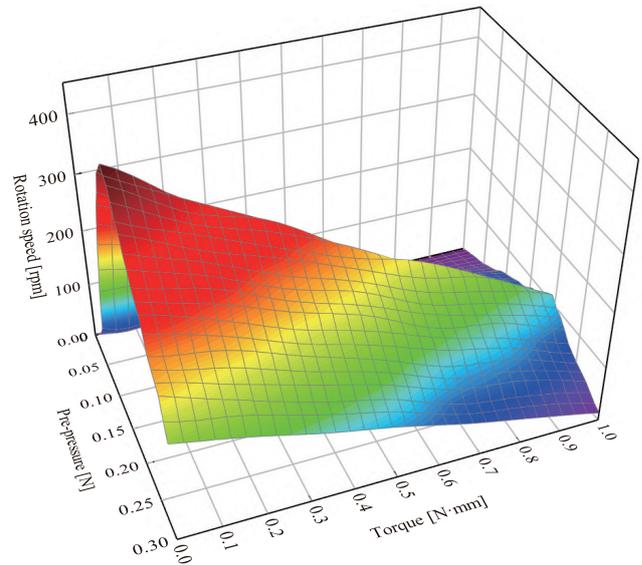


Fig. 13. Relationship between the experimentally found motor speed, pre-pressure, and load torque.

At present, the speed regulation methods of ultrasonic motors mainly include frequency regulation, voltage regulation, and phase regulation. The existing speed regulation methods often have the problem of coupling the speed and torque, as well as the narrow adjustment range. For such problems, we propose to change the pre-pressure speed regulation scheme and conduct a pre-pressure speed regulation experiment for this motor. According to the experimental results, the relationship between motor speed and pre-pressure under different loads can be obtained, as shown in Fig. 14.

As can be seen from Fig. 14, the motor speed increases and then decreases with increasing pre-pressure for all cases with different load torques. By using this monotonic relationship fragment before and after the

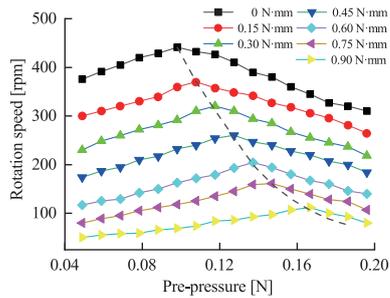


Fig. 14. Relationship between speed and pre-pressure under different loads.

pre-pressure reaches a specific value, the motor speed can be adjusted. In addition, for fast speed regulation and to avoid non-monotonic relationship in speed regulation, the pre-pressure should be gradually increased from the left side for small load torque and gradually decreased from the right side for large load torque until the motor reaches the desired speed. The dashed line marks the trend of the pre-pressure corresponding to the maximum speed of the motor at different load torques. It can also be seen that there is no sudden blocking of the motor when increasing the pre-pressure; theoretically a full range of speed regulation can be achieved.

5. Conclusion

With the help of the ANSYS finite element software, a parametric model of an internal cone type rotating traveling wave ultrasonic motor with trapezoidal teeth was established. The modal analysis and harmonic response analysis of the motor vibrator were carried out, and the structural parameters and working modes were determined. A prototype was fabricated, and the vibration characteristics of the motor vibrator were tested by the laser vibration measurement system, and the excitation frequency of the two orthogonal modes with the same frequency was 22260 Hz. An output performance test device was built, and the output characteristics of the prototype were tested experimentally. The prototype runs stably, has a high-speed output, and has good motion and power adjustment characteristics. When the excitation voltage peak-to-peak value is 300 V, the pre-pressure is 0.1 N, and the excitation frequency is 22260 Hz, the maximum output torque of the ultrasonic motor is 1.06 N·mm, and the maximum no-load speed is 441.2 rpm. The optimal pre-pressure of the motor under different loads is studied and analyzed. There is a coupling relationship between the influence of pre-pressure and load torque on the speed of the ultrasonic motor. Adjusting the pre-pressure according to the load and rotational speed can improve the output efficiency of the ultrasonic motor. This has important implications for the practical use of this ultrasonic motor.

Acknowledgments

This research was funded by the Natural Science Foundation of Liaoning Provincial Department of Science and Technology (grant no. 2022-BS-307) and the Basic Research Project of Education Department of Liaoning Province (grant no. JJJL202015401).

References

1. ČEPOŃIS A., MAŽEIKĀ D., VASILJEV P. (2020), Flat cross-shaped piezoelectric rotary motor, *Applied Sciences*, **10**(14): 5022, doi: [10.3390/app10145022](https://doi.org/10.3390/app10145022).
2. LI H., DENG J., ZHANG S., YU H., LIU Y. (2021), Design and experiment of a three-feet linear ultrasonic motor using third bending hybrid modes, *Sensors and Actuators A: Physical*, **331**: 112990, doi: [10.1016/j.sna.2021.112990](https://doi.org/10.1016/j.sna.2021.112990).
3. LI S. *et al.* (2019), Tailoring friction interface with surface texture for high-performance ultrasonic motor friction materials, *Tribology International*, **136**: 412–420, doi: [10.1016/j.triboint.2019.03.072](https://doi.org/10.1016/j.triboint.2019.03.072).
4. LIU R. *et al.* (2022), A precision positioning rotary stage driven by multilayer piezoelectric stacks, *Precision Engineering*, **76**: 226–236, doi: [10.1016/j.precisioneng.2022.03.013](https://doi.org/10.1016/j.precisioneng.2022.03.013).
5. LU D., LIN Q., CHEN B., JIANG C., HU X. (2020), A single-modal linear ultrasonic motor based on multi vibration modes of PZT ceramics, *Ultrasonics*, **107**: 106158, doi: [10.1016/j.ultras.2020.106158](https://doi.org/10.1016/j.ultras.2020.106158).
6. MAKAREM S., DELIBAS B., KOC B. (2021), Data-driven tuning of PID controlled piezoelectric ultrasonic motor, *Actuators*, **10**(7): 148, doi: [10.3390/act10070148](https://doi.org/10.3390/act10070148).
7. MASHIMO T., OBA Y. (2022), Performance improvement of micro-ultrasonic motors using the thickness shear mode piezoelectric elements, *Sensors and Actuators A: Physical*, **335**: 113347, doi: [10.1016/j.sna.2021.113347](https://doi.org/10.1016/j.sna.2021.113347).
8. MISHRA J.P., XU Q., YU X., JALILI M. (2018), Precision position tracking for piezoelectric-driven motion system using continuous third-order sliding mode control, [in:] *IEEE/ASME Transactions on Mechatronics*, **23**(4): 1521–1531, doi: [10.1109/TMECH.2018.2853737](https://doi.org/10.1109/TMECH.2018.2853737).
9. OH J.-H. *et al.* (2009), Design and performances of high torque ultrasonic motor for application of automobile, *Journal of Electroceramics*, **22**(1): 150–155, doi: [10.1007/s10832-008-9434-1](https://doi.org/10.1007/s10832-008-9434-1).
10. OLSSON P., NYSJÖ F., CARLBOM I.B., JOHANSSON S. (2016), Comparison of walking and traveling-wave piezoelectric motors as actuators in kinesthetic haptic devices, [in:] *IEEE Transactions on Haptics*, **9**(3): 427–431, doi: [10.1109/TOH.2016.2537803](https://doi.org/10.1109/TOH.2016.2537803).
11. PUOZA J.C., SAKTHIVELSAM Y. R. (2021), Ultrasonic motors structural design and tribological performance – A review, *Tribology Online*, **16**(4): 286–298, doi: [10.2474/trol.16.286](https://doi.org/10.2474/trol.16.286).

12. RYNDZIONEK R., SIENKIEWICZ Ł. (2021), A review of recent advances in the single- and multi-degree-of-freedom ultrasonic piezoelectric motors, *Ultrasonics*, **116**: 106471, doi: [10.1016/j.ultras.2021.106471](https://doi.org/10.1016/j.ultras.2021.106471).
13. SOEDEL W. (2004), *Vibrations of Shells and Plates*, CRC Press, USA.
14. TIAN X., LIU Y., DENG J., WANG L., CHEN W. (2020), A review on piezoelectric ultrasonic motors for the past decade: Classification, operating principle, performance, and future work perspectives, *Sensors and Actuators A: Physica*, **306**: 111971, doi: [10.1016/j.sna.2020.111971](https://doi.org/10.1016/j.sna.2020.111971).
15. WANG H., PAN Z., ZHU H., GUO Y. (2020), Pre-pressure influences on the traveling wave ultrasonic motor performance: A theoretical analysis with experimental verification, *AIP Advances*, **10**(11): 115211, doi: [10.1063/5.0028282](https://doi.org/10.1063/5.0028282).
16. WANG P., XU Q. (2017), Design and testing of a flexure-based constant-force stage for biological cell micro-manipulation, [in:] *IEEE Transactions on Automation Science and Engineering*, **15**(3): 1114–1126, doi: [10.1109/TASE.2017.2733553](https://doi.org/10.1109/TASE.2017.2733553).
17. XU D., YANG W., ZHANG X., YU S. (2021), Design and performance evaluation of a single-phase driven ultrasonic motor using bending-bending vibrations, *Micro-machines*, **12**(8): 853, doi: [10.3390/mi12080853](https://doi.org/10.3390/mi12080853).
18. ZHANG P., NIU J., ZHANG X., MAO S., LIU J., YANG B. (2022), Thick film ultrasonic micromotor based on chemical mechanical thinning and polishing process, [in:] *IEEE Electron Device Letters*, **43**(9): 1547–1550, doi: [10.1109/LED.2022.3195349](https://doi.org/10.1109/LED.2022.3195349).
19. ZHAO C. (2011), *Ultrasonic Motors: Technologies and Applications*, Springer, China.
20. ZHAO Y., YUAN S., CHU X., GAO S., ZHONG Z., ZHU C. (2016), Ultrasonic micro-motor with multilayer piezoelectric ceramic and chamfered driving tips, *Review of Scientific Instruments*, **87**(9): 095108, doi: [10.1063/1.4963662](https://doi.org/10.1063/1.4963662).

Research Paper

Relationship Between the Sound Transmission Through the Finite Double-Panel Structure with a Cylindrical Shell Array and the Vibro-Acoustic Characteristics of its Constituents

Song-Hun KIM, Myong-Jin KIM*

*Institute of Acoustics, Department of Physics, Kim Il Sung University
Pyongyang, Democratic People's Republic of Korea*

*Corresponding Author e-mail: mj.kim0903@ryongnamsan.edu.kp

(received February 23, 2022; accepted February 7, 2023)

Sound insulation of the finite double-panel structure (DPS) inserted with a cylindrical shell array is investigated by varying the sound incidence direction to improve its applicability. The effects of the vibro-acoustic characteristics of its constituents on the sound transmission loss (STL) are estimated in one-third octave bands from 20 Hz to 5 kHz for different incidence conditions. It shows that the first acoustic mode in the direction parallel to two panels (longitudinal modes) produces both the sudden variation of sound insulation with frequency and a large dependency on the incidence angle. Mineral wools are placed on two boundaries perpendicular to the panels, and the sound insulation is explored for different thicknesses of the porous materials. An absorbent layer with a certain thickness (more than 30 mm in our work) sufficiently eliminates the longitudinal mode, resulting in the improvement in the sound insulation by more than 15 dB and the decrease of its large variation with incidence direction. STLs with varying shell thicknesses are also assessed. It shows that the natural vibrations of the thin shells can give an enhancement in sound insulation by more than 10 dB in the frequency range of 1600–3700 Hz, corresponding to constructive interference.

Keywords: sound transmission loss; double-panel structure; eigenmode vibration; sonic crystal.



Copyright © 2023 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

It was first found experimentally in 1995 that the periodic arrangement of scatterers gives a high acoustic attenuation in the frequency range centered around Bragg's frequency (MARTÍNEZ-SALA *et al.*, 1995). Since then, several studies have been performed to control the noise using a cylindrical solid array.

The acoustical band structure of the regularly arranged solid cylinders in the free field was predicted by the plane wave expansion (PWE) method (CHEN, YE, 2001; ROMERO-GARCÍA *et al.*, 2011) and multiple scattering theory (MST) (KRYNKIN *et al.*, 2011) and validated experimentally. MARTÍNEZ-SALA *et al.* (2006) conducted a free field measurement on the tree arrangement similar to the periodic lattice and showed that it behaves like a sonic crystal (SC), its acoustic attenuation is higher than those for the ordinary forest and grass in the frequency range below 500 Hz, and it has the practical possibility to be used as a green noise

barrier. Sound attenuations through the solid cylinder arranged in periodic arrays were investigated using the coupling of Bragg-type reflection and the sound absorption caused by placing the absorbents, such as porous material, on the scatterers or filling the gaps between them (UMNOVA *et al.*, 2006; SÁNCHEZ-DEHESA *et al.*, 2011).

Inserting an SC between two panels can improve the sound insulation of the structure in the audible frequency range. TANG (2018) reported that installing cylinder columns into a plenum window gave rise to a broadband improvement in noise reduction. The author also investigated the effect of parameters such as the number and location of columns, lattice arrangement, and so on. GULIA and GUPTA (2018) estimated the individual and coupling effects of SC and porous material inserted into double-panel structure (DPS) and found that the coupling of two elements gave the highest sound insulation. KIM (2019a) extended the work of GULIA and GUPTA (2018) to the triple-panel

structure and showed that it exhibited higher sound insulation than DPS with the same total thickness and mass over the entire frequency range, even giving a potential to reduce the overall weight.

Using the local resonance effect is one way to reduce sound transmission in the low-frequency range or in the frequency range of interest without increasing the thickness and weight. CHALMERS *et al.* (2009) suggested two types of SCs comprised of C-shaped Helmholtz resonators: the first is simply mixing two columns with different resonance frequencies, and the second is the arrangement of concentric resonators. ELFORD *et al.* (2011) presented a six concentric Matryoshka configuration of slit cylinders. Simulation results demonstrated the overlapping of the individual resonance peaks in the low-frequency range and the broadening of the resonant band gap. Local resonant SCs comprised of thin elastic shells were also presented, in which the breathing mode resonance of the shell was used. KRYNKIN *et al.* (2010) showed that the periodic arrangement of the thin elastic shells made of soft rubber exhibits a wide band gap due to the axisymmetric resonance in the shell. FUSTER-GARCIA *et al.* (2007) demonstrated experimentally the resonant behavior of an array formed with pressurized gas-filled balloons and the possibility of achieving a full band gap in a desired frequency range. CHONG (2012) presented a study on the sound insulation of a noise barrier consisting of a concentric arrangement of an outer 4-slit rigid cylinder and an inner elastic cylindrical shell.

Local resonant SCs were applied to the multiple-panel structure to enhance the sound insulation. KIM (2019b) proposed an array of C-shaped scatterers with different Helmholtz resonance frequencies and showed that the DPS with the array gave a relatively wide band gap in the low-frequency range. GULIA and GUPTA (2019) showed that coupling the effects of the local resonant SC and porous material in the triple panel gave better sound transmission loss (STL) in the low-frequency range as well as in the high-frequency range. KIM *et al.* (2021) formed the resonant band gap in the low-frequency range by using the local resonant SC comprised of slot-type resonators, in which the slots were elongated into the cavity and had different depths.

Normal incidence condition is usually applied to the inlet of SC for numerical and analytical research, including the calculation of the acoustic band structure, the prediction of STL and the optimization. In addition, its dimension is considered infinite and the periodicity condition is used to reduce the computing cost. For the cylindrical scatterers, sound hard condition is applied to their surfaces, leading to no accounting for the effect of shell vibration on the sound propagation.

In this work, variation of the sound insulation through a DPS with a periodic arrangement of cylindrical shells is systematically investigated with vary-

ing the incidence angle of sound waves. Particularly, the dependency of sound insulation on the incidence direction is analyzed in relation to the acoustic mode of the finite air cavity. By weakening the longitudinal acoustic mode using the sound-absorbing material of a certain thickness, the large variation in STL with frequency is eliminated and the sound insulation is improved. Varying the shell thickness, the sound insulations are estimated in the frequency range where the scattered waves interfere constructively. Structural and acoustical characteristics of the entire structure and its constituents are analyzed using the VA One v.2015 software for analysis and design of vibro-acoustic systems.

2. Numerical modeling

A schematic diagram of DPSs with a periodic arrangement of the cylindrical shells is shown in Fig. 1. Figure 1a presents the DPS without any sound absorbent and Fig. 1b shows the one with the absorbent on two boundaries in the longitudinal direction (upper and lower in the figure).

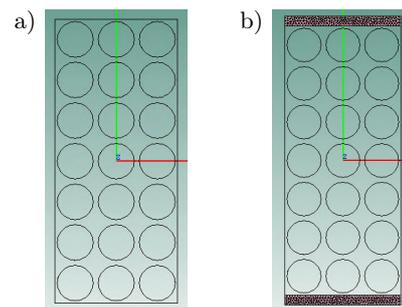


Fig. 1. DPSs with a cylindrical shell array: a) without absorbent; b) with absorbent.

Cylindrical shells are arranged on a 7×3 square lattice, and they are inserted between two parallel panels, as shown in Fig. 1. Distance between the two centers of the neighboring shells (lattice constant) is 40 mm and the outer diameter of the shell is 35 mm. The left side of the DPS is the aluminum panel with a thickness of 4 mm, and the right side is the plywood panel with a thickness of 5 mm. Cylindrical shells are considered to be made of aluminum. The sound speed in air is $c_0 = 343$ m/s, the density of air $\rho_0 = 1.25$ kg/m³, and the material properties of aluminum and plywood are presented in Table 1.

Table 1. Material properties of aluminum and plywood.

	Density [kg/m ³]	Young's modulus [GPa]	Poisson's ratio
Aluminum	2700	71	0.33
Plywood	700	6	0.25

Table 2. Non-acoustic properties of mineral wool.

Flow resistivity [N · s/m ⁴]	Porosity	Tortuosity	Viscous characteristic length [m]	Thermal characteristic length [m]
6×10^4	0.95	3.2	5×10^{-5}	1.06×10^{-4}

The DPS is mounted in an infinite baffle and sound waves are incident on the left side (i.e., aluminum panel with a thickness of 4 mm). The acoustic velocity is zero in the infinite baffle. The infinite air regions on both sides of the DPS are modeled using the boundary element method (BEM). Sound hard condition is applied to the upper and lower boundaries of the air region between the two panels. Sound waves are completely reflected by such boundaries, and the particle velocity perpendicular to the boundaries is zero as expressed by Eq. (1) (GULIA, GUPTA, 2018):

$$\left(-\frac{\nabla p}{\rho}\right) \cdot \mathbf{n} = 0, \quad (1)$$

where p and \mathbf{n} denote the sound pressure and normal vector, respectively.

For harmonic excitation, the vibration displacement of the panel can be expanded on the basis of eigenfunctions ψ_{mn} , which satisfy the following equation (BRUNEAU, 2006):

$$\left(\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2 \partial y^2} + \frac{\partial^4}{\partial y^4} - k_{mn}^2\right) \psi_{mn} = 0, \quad (2)$$

where $k_{mn}^2 = \omega_{mn}^2 M_s / B$, $M_s = \rho_p h$, $B = Eh^3 / 12(1 - \nu^2)$, ω_{mn} , M_s , and B represent the natural angular frequency, the mass per unit area and the bending stiffness, respectively, and ρ_p , ν , h , and E are the density, Poisson's ratio, the thickness, and Young's modulus of the panel, respectively.

For the simply supported rectangular panel, the displacement and bending moment are zero at the boundaries and the natural angular frequencies are as follows:

$$\omega_{mn} = \sqrt{\frac{B}{M_s} \left[\left(\frac{m\pi}{w}\right)^2 + \left(\frac{n\pi}{l}\right)^2 \right]}, \quad (3)$$

where w and l are the width and length of the panel, respectively.

Because the identical shells are arranged periodically, a high and wide band gap is formed around the Bragg frequency (Eq. (4)) due to the destructive interference of the scattered waves:

$$f_B = \frac{c}{2a}, \quad (4)$$

where a is the lattice constant and c is the sound speed in air. On the contrary, the sound insulation around the frequency corresponding to half the Bragg frequency tends to get worse compared to that without the scatterer array due to constructive interference (GULIA, GUPTA, 2018; KIM, 2019a; KIM *et al.*, 2021).

A porous material is one of the most widely used absorbents for noise control. In this work, porous materials with different thicknesses were placed on the upper and lower boundaries of the DPS. Mineral wool is taken as a porous material and an equivalent fluid model-limp porous model (ALLARD, ATALLA, 2009) is used to describe the sound propagation in the mineral wool. The limp model of the porous material is valid when the thin light foam is decoupled with an air gap from a vibrating system. This model requires the evaluation of the acoustic and non-acoustic (or microscopic) properties. The acoustic property is sound absorbing efficiency, while there are five non-acoustic properties, including static airflow resistivity, porosity, tortuosity, viscous characteristic length, and thermal characteristic length.

Similar to other porous materials, the absorption coefficient of mineral wool also increases with increasing thickness. The density of the mineral wool is 50 kg/m³, and its non-acoustic properties are presented in Table 2.

Harmonic analyses are performed to estimate STLs in the 1/3 octave bands covering the range from 20 Hz to 5 kHz. For the sake of reducing the computing cost, two-dimensional analyses are conducted, considering the length of shells to be large enough compared to the outer diameter and thickness. Acoustical and structural constituents are modeled using the finite element method (FEM), while the fluid medium (air) is modeled using the BEM. STL is taken as the difference in sound power levels at the 4 mm-thick aluminum panel and the 5 mm-thick plywood panel.

3. Results and discussion

Figure 2 shows the STLs in 1/3 octave bands for two cases of the incidence angles, 10 and 90°, between the panel surface and the sound incidence direction. For the incidence angles from 10 to 90°, STLs lie mainly between the two curves. In the figure, STL for diffused field incidence is also plotted.

The frequency range below 100 Hz can be considered as the stiffness-controlled portion in Fig. 2. The range from 100 to 200 Hz corresponds to the first resonances of the two panels. The first resonance frequencies of two panels are 84.6 and 118.7 Hz, respectively, leading to the high sound transmission in the frequency range.

STLs increase rapidly for the frequency range over 4 kHz and then become higher than 110 dB at 5 kHz.

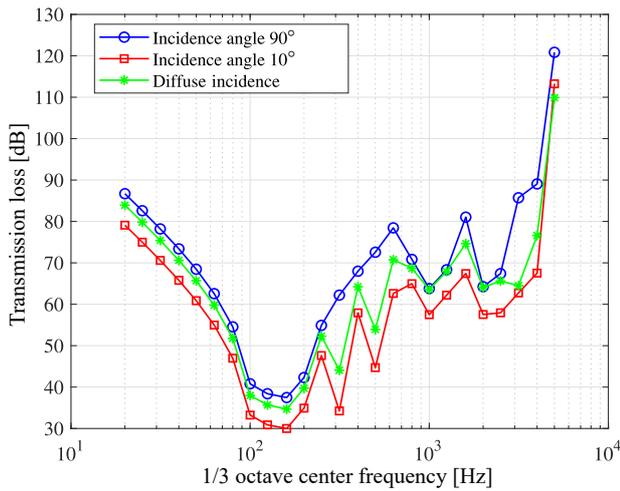


Fig. 2. STLs in 1/3 octave bands.

This corresponds to the Bragg band gap due to the periodic arrangement of cylindrical shells between two panels with a lattice constant of 40 mm. According to Eq. (4), the Bragg band gap is centered at 4287.5 Hz.

The sound insulation of the DPS gets higher when the incidence angle increases, and the STL for the diffused incidence falls approximately in between the values observed for two extreme cases (10 and 90°).

The variation width in STL with the incident angle is shown in Fig. 3. The variation in STL is uniform around 7.5 dB in the frequency range from 20 to 250 Hz, but large above that. Especially it is particularly large in the frequency ranges of 250 to 630 Hz and 2 to 4 kHz.

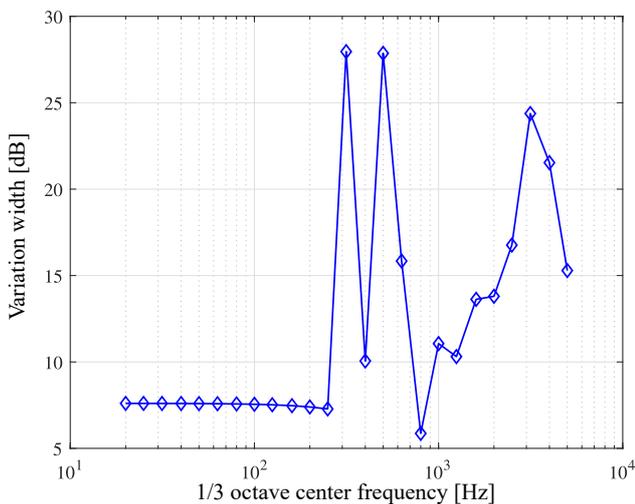


Fig. 3. Variation width in STL as a function of 1/3 octave band.

Since both the acoustic and structural elements of the DPS have finite dimensions, there are acoustic and structural eigenmodes in different directions.

Figure 4 shows the number of acoustic modes generated in the air region surrounded by cylindrical shells,

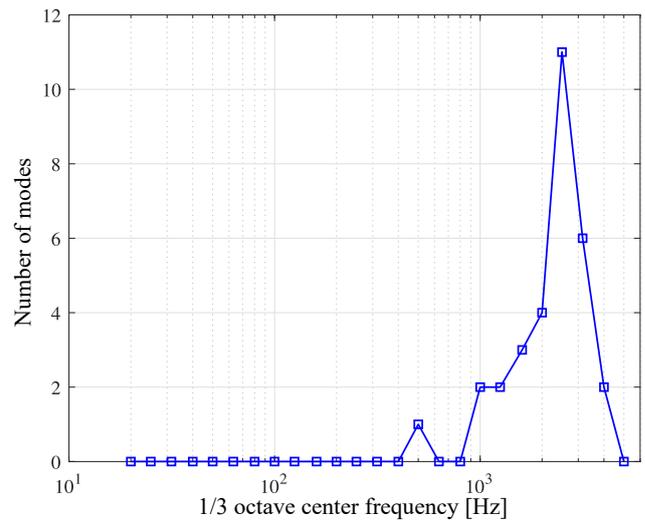


Fig. 4. Numbers of acoustic modes in 1/3 octave bands.

two parallel panels, and upper and lower boundaries. There is no acoustic mode in the frequency range below 400 Hz, and one mode exists in the 1/3 octave band with a center frequency of 500 Hz. For the frequencies above 1 kHz, the number of acoustic modes increases gradually, reaching a maximum of 11 in the band centered at 2500 Hz. Then, it decreases rapidly and becomes zero again at 5 kHz.

There are three types of acoustic eigenmodes: transverse, longitudinal, and coupled ones. Figure 5 shows four natural acoustic modes with frequencies of 469.9, 929.9, 1079.2, and 1182 Hz, respectively.

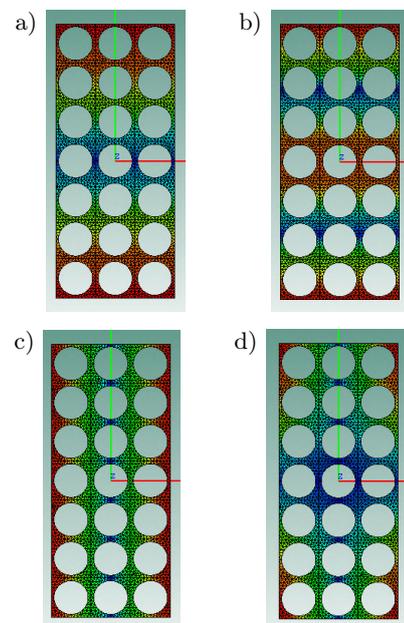


Fig. 5. Acoustic modes: a) 469.9 Hz; b) 929.9 Hz; c) 1079.2 Hz; d) 1182 Hz.

In the figure, 469.9 Hz corresponds to the first longitudinal mode, 929.9 Hz to the second longitudinal mode,

1079.2 Hz to the first transverse mode, and 1182 Hz to the coupled mode. Such acoustic modes provide different contributions to sound transmission through the DPS according to the incidence angle.

For a given incidence angle of 10° , four acoustic modes are suppressed one after another and the results are presented in Fig. 6. As presented in the figure, suppressing the first longitudinal mode eliminates the large fluctuations in sound insulation in the range from 250 to 630 Hz and STL increases monotonically. Considering the fact that the first longitudinal mode is in the range from 250 to 630 Hz, it can be seen that the large reduction and strong fluctuation in sound insulation are caused by the first eigenmode. The more the incident sound wave is inclined, the stronger the excitation of the first longitudinal mode becomes. Then, the scattering of sound waves by the cylindrical shells becomes stronger, resulting in an increase in sound energy radiated from the output surface of the DPS.

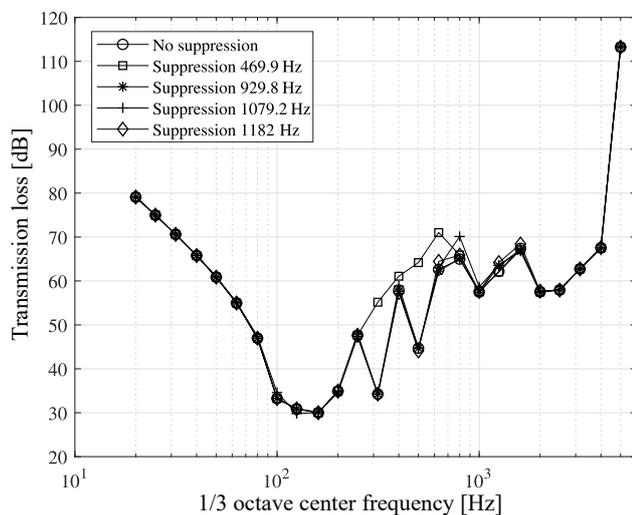


Fig. 6. STLs in 1/3 octave bands for an incidence angle of 10° .

For the other cases of 929.9, 1079.2, and 1182 Hz, the sound insulation is improved by about 1 dB in the frequency range from 630 to 1250 Hz and there is no change in STL in the remaining range. As shown in Fig. 4, there are a relatively large number of natural acoustic modes in the frequency range from 2 to 4 kHz, so it can be understood that the addition of effects of the individual modes results in a large variation in STL in the frequency range. From these behaviors, it can be concluded that, for a sound wave obliquely incident on the boundary of DPS, the first longitudinal acoustic mode produces the largest influence on its sound insulation. On the other hand, the influences of the remaining modes are very small compared to that of the first longitudinal mode.

In order to eliminate the remarkable variation and reduction in STL in the frequency range from 250 to 630 Hz, it is necessary to reduce the intensity of

the first longitudinal mode using the sound-absorbing materials. To confirm this, mineral wools with non-acoustic properties, as shown in Table 2, are placed on the upper and lower boundaries, as in Fig. 1b, and then STLs are evaluated for three different thicknesses of mineral wool. The absorption coefficients of the mineral wool are shown in Fig. 7.

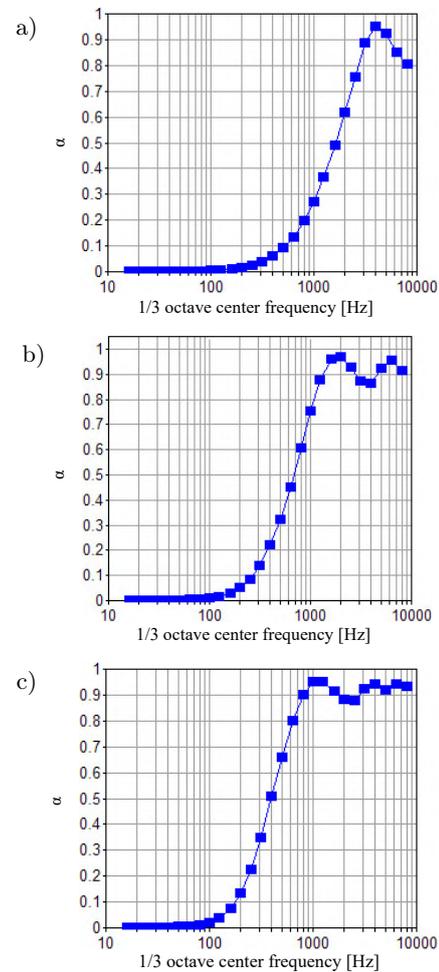


Fig. 7. Absorption coefficient of mineral wool thickness: a) 10 mm; b) 20 mm; c) 30 mm.

Similar to other porous materials, the sound absorption coefficient of mineral wool is very small at low frequencies and relatively great at high frequencies. As shown in Fig. 7, the absorption coefficient is smaller than 10% at frequencies below 500 Hz for a thickness of 10 mm. With increasing thickness, the absorption peak shifts towards the lower frequencies, and for a thickness of 30 mm, the absorption coefficient is greater than 10% at 200 Hz and increases rapidly beyond that frequency.

Figure 8 shows STLs with various thicknesses of mineral wool for an incidence angle of 10° .

Increasing the absorbent thickness not only reduces the unexpected fluctuation in STL with frequency, but also significantly enhances the sound insulation in the

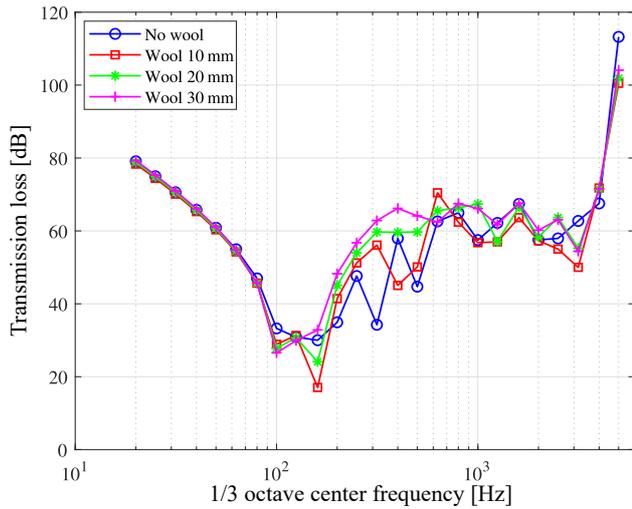


Fig. 8. STL with various thicknesses of absorbing material for an incidence angle of 10°.

frequency range of interest. Especially, the sound insulation for a thickness of 30 mm is higher than that without absorbing materials approximately at all the bands and the variation width in SPF with the incidence angle decreases to less than 10 dB.

Figure 9 shows STLs with varying thicknesses of the mineral wool for diffused field incidence. Similar to the above case of the incidence angle of 10°, placing the absorbent of thickness 30 mm improves the overall sound insulation and completely eliminates the abrupt change in STL in the frequency range from 250 to 630 Hz. From these results, it can be concluded that the variation in STL through DPS with the incidence direction of a sound wave is primarily caused by the longitudinal acoustic modes in the direction parallel to the panel surface of DPS. And a certain thickness of absorbent is required not only to decrease the strong

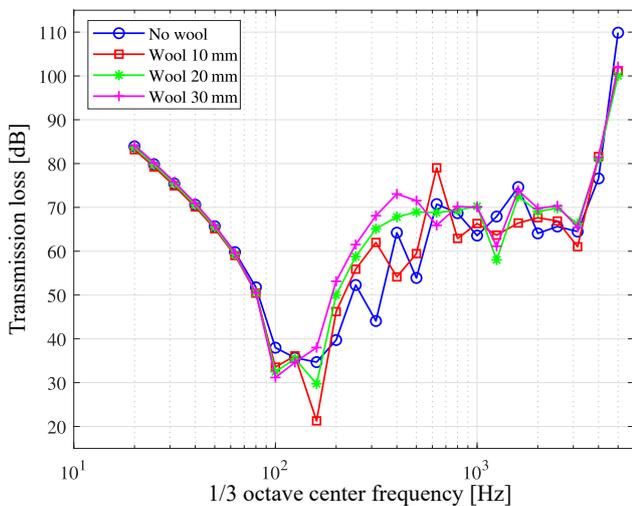


Fig. 9. STL with various thicknesses of absorbing material for diffused incidence.

dependency of STL on the incidence angle but also to improve the overall sound insulation.

The cylindrical shell can be made either of polyvinyl chloride (PVC) or sheet metal. In the case of cylindrical shells with finite thickness, they will oscillate by acoustic loading and consequently would have some influences on the sound propagation through DPS.

Figure 10 shows the overall sound power transmitted through the DPS in the frequency range from 20 Hz to 5 kHz when the thickness of the cylindrical shell, made of aluminum, varies from 0.2 to 2 mm. The sound wave is diffusely incident on the DPS and the amplitude of the sound pressure in all directions is 1 Pa.

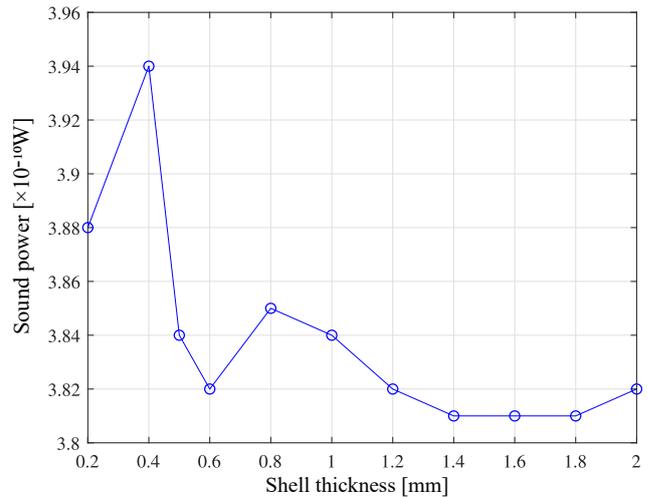


Fig. 10. Overall transmitted sound power as a function of shell thickness.

In Fig. 11, the overall transmitted sound power is a maximum for the shell thickness of 0.4 mm and approximately decreases with increasing the shell thickness. It has little variation above 1.4 mm, which means that vibration displacement can be regarded as almost zero for acoustic loading. Although the overall trans-

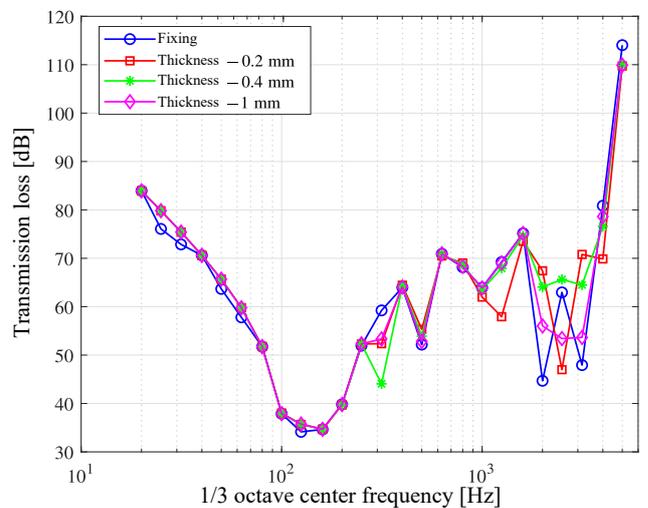


Fig. 11. STLs with various shell thicknesses.

mitted sound power varies with the shell thickness, its variation width is negligibly small on a logarithmic scale (about 0.15 dB). But, the individual ones in 1/3 bands might be significantly different from each other when the shell thickness varies.

Figure 11 shows STLs for three cases with shell thicknesses of 0.2, 0.4, and 1 mm. STL for the DPS with perfectly fixed shells is also shown in the figure. Differences in the STLs between the four cases are small except in the frequency ranges below 1600 and above 3700 Hz, where the sound insulation varies greatly with the shell thickness and is lower than those in the two adjacent frequency bands, even lower than that without the shell array. This result was already reported in previous research (GULIA, GUPTA, 2018; 2019; KIM, 2019a; KIM *et al.*, 2021).

The reason for this is that the range is around the frequency at which the lattice constant is equal to 1/4 of the wavelength and thus the acoustic waves scattered by periodic cylindrical shells constructively interfere with each other.

In the figure, the sound insulation in the frequency range from 1600 to 3700 Hz is the highest for the shell thickness of 0.4 mm, and it is up to 10 dB higher than the other cases in the frequency range.

The natural frequencies in the entire frequency range from 20 Hz to 5 kHz are shown in Table 3 for different shell thicknesses. In all the cases, the cylindrical shells exhibit one or two natural frequencies in the frequency range from 1600 to 3700 Hz.

Table 3. Natural frequencies of shells with various shell thicknesses.

Shell thickness [mm]	Natural frequency [Hz]
0.2	431, 1220, 2344, 3796
0.4	852, 2413, 4634
0.6	1269, 3593
0.8	1685, 4769
1	2100

As shown in Fig. 11, the even-order mode (2413 Hz for shell thickness of 0.4 mm) disturbs the constructive interferences of the acoustic waves scattered by the shells, resulting in an improvement of the sound insulation in the frequency range from 1600 to 3700 Hz. On the contrary, the odd-order modes (2344 Hz for shell thickness of 0.2 mm and 2100 Hz for shell thickness of 1 mm) exhibit no improvements in the sound insulation; instead, they negatively contribute to the sound insulation, as shown in Fig. 11.

4. Conclusions

The effects of the structural and acoustic modes on sound transmission through a finite-size DPS were

evaluated and analyzed. For the inclined incidence, STLs strongly vary around the frequency (489.9 Hz in this work) corresponding to the first longitudinal mode in the direction parallel to the panel surface of DPS, resulting in poor sound insulation. The variation in the sound insulation with incidence angle is also the largest around that frequency. With varying the incidence angle of sound waves from 10 to 90°, the sound insulation changes within 7 to 8 dB in most 1/3 octave bands, but it has a variation width of even more than 25 dB in some bands around that frequency. The first longitudinal acoustic mode gives the largest effect on the sound insulation and those of the other acoustic modes are very small. Placing the mineral wool of a certain thickness (30 mm in this work) on two longitudinal boundaries can significantly suppress the intensity of the first longitudinal acoustic mode. This exhibits a reduction of the variation in STL with an incident angle to less than 10 dB and an increase of the sound insulation by more than 15 dB in the frequency range around the mode (250 to 630 Hz in this work). Cylindrical shells can have their natural vibration modes in the frequency range where the sound waves scattered by the shells interfere constructively. Even-order modes of them behave to disturb the constructive interferences of the scattered waves, which can greatly enhance the sound insulation through the DPS in the frequency range (1600 to 3700 Hz in this work). On the other hand, the odd-order modes have negative effects on sound insulation.

Acknowledgments

The authors would like to express their gratitude to professors and researchers in the Department of Physics at Kim Il Sung University for their great opinions, discussions, and constructive assistance in improving the scientific value and quality of this work.

References

1. ALLARD J.F., ATALLA N. (2009), *Propagation of Sound in Porous Media: Modeling Sound Absorbing Materials*, 2nd ed., John Wiley & Sons, doi: [10.1002/9780470747339](https://doi.org/10.1002/9780470747339).
2. BRUNEAU M. (2006), *Fundamentals of Acoustics*, ISTE Ltd Press, London.
3. CHALMERS L., ELFORD D.P., KUSMARTSEV F.V., SWALLOWE G.M. (2009), Acoustic band gap formation in two-dimensional locally resonant sonic crystals comprised of Helmholtz resonators, *International Journal of Modern Physics B*, **24**(20n21): 4231–4243, doi: [10.1142/S0217979209063390](https://doi.org/10.1142/S0217979209063390).
4. CHEN Y.-Y., YE Z. (2001), Acoustic attenuation by two-dimensional arrays of rigid cylinders, *Physical Re-*

- view *Letters*, **87**(18): 184301, doi: [10.1103/PhysRevLett.87.184301](https://doi.org/10.1103/PhysRevLett.87.184301).
5. CHONG Y.B. (2012), *Sonic crystal noise barriers*, Ph.D. Thesis, The Open University, Milton Keynes, UK, doi: [10.21954/ou.ro.0000add6](https://doi.org/10.21954/ou.ro.0000add6).
 6. ELFORD D.P., CHALMERS L., KUSMARTSEV F.V., SWALLOWE G.M. (2011), Matryoshka locally resonant sonic crystal, *The Journal of the Acoustical Society of America*, **130**(5): 2746–2755, doi: [10.1121/1.3643818](https://doi.org/10.1121/1.3643818).
 7. FUSTER-GARCIA E., ROMERO-GARCÍA V., SÁNCHEZ-PÉREZ J.V., GARCÍA-RAFFI L.M. (2007), Targeted band gap creation using mixed sonic crystal arrays including resonators and rigid scatterers, *Applied Physics Letter*, **90**(24): 244104, doi: [10.1063/1.2748853](https://doi.org/10.1063/1.2748853).
 8. GULIA P., GUPTA A. (2018), Enhancing the sound transmission loss through acoustic double panel using sonic crystal and porous material, *The Journal of the Acoustical Society of America*, **144**(3): 1435–1442, doi: [10.1121/1.5054296](https://doi.org/10.1121/1.5054296).
 9. GULIA P., GUPTA A. (2019), Sound attenuation in triple panel using locally resonant sonic crystal and porous material, *Applied Acoustics*, **156**: 113–119, doi: [10.1016/j.apacoust.2019.07.012](https://doi.org/10.1016/j.apacoust.2019.07.012).
 10. KIM M.-J. (2019a), Improving sound transmission through triple-panel structure using porous material and sonic crystal, *Archives of Acoustics*, **44**(3): 533–541, doi: [10.24425/aoa.2019.129268](https://doi.org/10.24425/aoa.2019.129268).
 11. KIM M.-J. (2019b), Numerical study for increase-ment of low frequency sound insulation of double-panel structure using sonic crystals with distributed Helmholtz resonators, *International Journal of Modern Physics B*, **33**(14): 1950138, doi: [10.1142/S0217979219501388](https://doi.org/10.1142/S0217979219501388).
 12. KIM M.-J., RIM C.-G., WON K.-S. (2021), Broadening low-frequency band gap of double-panel structure using locally resonant sonic crystal comprised of slot-type Helmholtz resonators, *Archives of Acoustics*, **46**(2): 335–340, doi: [10.24425/aoa.2021.136587](https://doi.org/10.24425/aoa.2021.136587).
 13. KRYNKIN A., UMNova O., YUNG BOON CHONG A., TAHERZADEH S., ATTENBOROUGH K. (2010), Predictions and measurements of sound transmission through a periodic array of elastic shells in air, *The Journal of the Acoustical Society of America*, **128**(6): 3496–3506, doi: [10.1121/1.3506342](https://doi.org/10.1121/1.3506342).
 14. KRYNKIN A., UMNova O., SÁNCHEZ-PÉREZ J.V., YUNG BOON CHONG A., TAHERZADEH S., ATTENBOROUGH K. (2011), Acoustic insertion loss due to two dimensional periodic arrays of circular cylinders parallel to a nearby surface, *The Journal of the Acoustical Society of America*, **130**(6): 3736–3745, doi: [10.1121/1.3655880](https://doi.org/10.1121/1.3655880).
 15. MARTÍNEZ-SALA R., RUBIO C., GARCÍA-RAFFI L.M., SÁNCHEZ-PÉREZ J.V., SÁNCHEZ-PÉREZ E.A., LLINARES J. (2006), Control of noise by trees arranged like sonic crystals, *Journal of Sound and Vibration*, **291**(1–2): 100–106, doi: [10.1016/j.jsv.2005.05.030](https://doi.org/10.1016/j.jsv.2005.05.030).
 16. MARTÍNEZ-SALA R., SANCHO J., SÁNCHEZ J.V., GÓMEZ V., LLINARES J., MESEGUER F. (1995), Sound attenuation by sculpture, *Nature*, **378**: 241, doi: [10.1038/378241a0](https://doi.org/10.1038/378241a0).
 17. ROMERO-GARCÍA V., GARCIA-RAFFI L.M., SÁNCHEZ-PÉREZ J.V. (2011), Evanescent waves and deaf bands in sonic crystals, *AIP Advances*, **1**(4): 041601, doi: [10.1063/1.3675801](https://doi.org/10.1063/1.3675801).
 18. SÁNCHEZ-DEHESA J., GARCIA-CHOCANO V.M., TORRENT D., CERVERA F., CABRERA S., SIMON F. (2011), Noise control by sonic crystal barriers made of recycled material, *The Journal of the Acoustical Society of America*, **129**(3): 1173–1183, doi: [10.1121/1.3531815](https://doi.org/10.1121/1.3531815).
 19. TANG S.K. (2018), Reduction of sound transmission across plenum windows by incorporating an array of rigid cylinders, *Journal of Sound and Vibration*, **415**: 25–40, doi: [10.1016/j.jsv.2017.11.027](https://doi.org/10.1016/j.jsv.2017.11.027).
 20. UMNova O., ATTENBOROUGH K., LINTON C.M. (2006), Effects of porous covering on sound attenuation by periodic arrays of cylinders, *The Journal of the Acoustical Society of America*, **119**(1): 278–284, doi: [10.1121/1.2133715](https://doi.org/10.1121/1.2133715).

Research Paper

An Improved EMD Method Based on Utilizing Certain Inflection Points in the Construction of Envelope Curves

Mohsen KAFIL^{(1),(2)*}, Kaveh DARABI⁽²⁾, Saeed ZIAEI-RAD⁽³⁾

⁽¹⁾ *Mechanical Engineering Group, Pardis College, Isfahan University of Technology*
Isfahan, Iran

⁽²⁾ *Mobarakeh Steel Company*
Isfahan, Iran

⁽³⁾ *Department of Mechanical Engineering, Isfahan University of Technology*
Isfahan, Iran

*Corresponding Author e-mail: mohsen.kafil@gmail.com

(received July 17, 2022; accepted February 11, 2023)

The empirical mode decomposition (EMD) algorithm is widely used as an adaptive time-frequency analysis method to decompose nonlinear and non-stationary signals into sets of intrinsic mode functions (IMFs). In the traditional EMD, the lower and upper envelopes should interpolate the minimum and maximum points of the signal, respectively. In this paper, an improved EMD method is proposed based on the new interpolation points, which are special inflection points (SIP_{*n*}) of the signal. These points are identified in the signal and its first ($n - 1$) derivatives and are considered as auxiliary interpolation points in addition to the extrema. Therefore, the upper and lower envelopes should not only pass through the extrema but also these SIP_{*n*} sets of points. By adding each set of SIP_{*i*} ($i = 1, 2, \dots, n$) to the interpolation points, the frequency resolution of EMD is improved to a certain extent. The effectiveness of the proposed SIP_{*n*}-EMD is validated by the decomposition of synthetic and experimental bearing vibration signals.

Keywords: empirical mode decomposition (EMD); interpolation points; envelope curve; inflection points; rolling element bearing fault diagnosis.



Copyright © 2023 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

One of the most powerful methods in dealing with nonlinear and non-stationary time series is the Hilbert–Huang transform (HHT), introduced by HUANG *et al.* (1998). The HHT procedure consists of two parts. At the first step, the empirical mode decomposition (EMD) process adaptively decomposes a complicated multi-component signal into sets of mono-component intrinsic mode functions (IMFs), which can admit well-behaved Hilbert transform. In the second step, the Hilbert spectral analysis (HSA) is performed on each IMF and the instantaneous amplitude and frequency of all IMFs are computed.

The major deficiency of the traditional EMD is the mode mixing phenomenon, which is defined as the ap-

pearance of signals at very different scales in one IMF or the distribution of a specific scale signal among different IMFs. Some researchers have proposed creative approaches to solve this problem, including ensemble empirical mode decomposition (EEMD) (WU, HUANG, 2009), complete ensemble empirical mode decomposition (CEEMD) (YEH *et al.*, 2010), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) (TORRES *et al.*, 2011), ensemble noise-reconstructed EMD method (ENEMD) (YUAN *et al.*, 2013), high-fidelity noise-reconstructed empirical mode decomposition (HNEMD) (YUAN *et al.*, 2022), EMD manifold (WANG *et al.*, 2020), and other.

EMD is an iterative process with interpolation as its central part. At each step of the iteration, all local maxima and minima of the signal are obtained and

then interpolated to construct the envelope curves. Although the above-mentioned methods have shown great improvement in the mode mixing problem, utilizing maxima and minima as the interpolation points remains an unchanged rule among them. Little research has been so far carried out on utilizing other interpolation points to construct envelope curves in the EMD process, so further studies are still needed in this area. Since there is no difference between the traditional EMD, EEMD, and CEEMD methods in terms of the selection of extrema as the interpolation points, this paper only focuses on the traditional EMD method. Furthermore, the traditional EMD has another advantage, i.e., the main signal is not repeatedly affected by different random noises.

By using the genetic algorithm, KOPSINIS and McLAUGHLIN (2007; 2008) showed that the performance of EMD will be improved if envelope curves pass through the extrema of the highest instantaneous frequency component of the signal instead of the extrema of the signal itself. XU *et al.* (2010) presented a piecewise-defined quadratic interpolation in which the position information of minimum points (maximum points) is used in the construction of the upper envelope (lower envelope). Hence the upper envelope (lower envelope) not only passes through the maxima (minima) but also through the intersection points of the lines connecting the maxima (minima) with the vertical lines going through the minima (maxima). The extrema detection method is an important matter, especially in the presence of noise, so BOUCHIKHI and BOUDRAA (2012) introduced a new version of EMD, which is more robust to noise since the envelopes of the signal are constructed using smoothing B-splines instead of exact interpolation methods. CHU *et al.* (2012) proposed an envelope curve that interpolates the main and pseudo extrema by the quartic Hermitian polynomial interpolation method. SINGH *et al.* (2014) introduced an EMD algorithm based on pseudo extrema and nonpolynomial spline interpolation. Instead of calculating the mean of the upper and lower envelopes, WANG and LI (2013) proposed to calculate the mean values directly by interpolating the midpoints of line segments, connecting each maximum (minimum) to its two neighboring minima (maxima). LI *et al.* (2015a) used a similar method but replaced the cubic spline interpolation with the cubic B-spline. ZHAO *et al.* (2017) put forward adjustable cubic trigonometric cardinal spline interpolation for the direct construction of the mean curve. Efforts have also been made to define a suitable envelope for the EMD method. For instance, YANG *et al.* (2014) presented a new model of the envelope by using convex constraint optimization.

In this paper, we introduce a special set of inflection points as the auxiliary interpolation points for the construction of the envelope curves. These special inflection points (SIPs), as well as how to distinguish

them from other inflection points, will be discussed later in Sec. 3. The popular cubic spline is applied as the interpolation function in this paper. The cubic spline is a continuous function and has continuous first and second derivatives. Two boundary conditions should also be met at the left and right sides of the signal. We utilize here the not-a-knot boundary condition, which requires the continuity of the third derivative at the second and second to last points.

The rest of this paper is organized as follows: the main procedure and adjustable parameters of the traditional EMD algorithm are explained in Sec. 2. Sections 3 and 4 describe the theoretical aspects of our new method. In Sec. 3, the SIP_n points of a signal are defined and the differences between the traditional and new envelopes are shown. The additional required steps for the proposed EMD method are described in Sec. 4. In Sec. 5, the performance of the proposed EMD method is verified for numerous two-tone signals and a real vibration signal of a rolling element bearing and the decomposition results are compared with the traditional EMD. Finally, a brief conclusion is provided in Sec. 6.

2. Empirical mode decomposition (EMD) algorithm

The EMD algorithm decomposes a signal into several IMFs and a final residual, which is the trend of the signal. The extraction process for each IMF is called the sifting process. The IMFs are expected to simultaneously meet the following two conditions (HUANG *et al.* 1998): 1) in the whole data set, the number of extrema and the number of zero crossings should be equal or differ at most by one; 2) the mean value of the envelope defined by the local maxima (upper envelope) and the envelope defined by the local minima (lower envelope) should be approximately zero, at any point. In practice, the definition of the term “zero” in the second condition should be defined more clearly and it is better to evaluate the amplitude of the mean compared to the amplitude of the extracted mode. The traditional EMD steps are as follows:

- 1) Initialize $r_0 = y$, where y is the main signal.
- 2) Initialize $i = 1$ and $k = 1$, where i counts the number of IMFs, and k counts the number of sifting iterations for each IMF.
- 3) Set $h_{i(k-1)} = r_{(i-1)}$.
- 4) Identify all local maxima of $h_{i(k-1)}$ and interpolate them by cubic spline to construct the upper envelope. Similarly, construct the lower envelope by interpolating minima.
- 5) Calculate the mean $m_{i(k-1)}$ of the upper and lower envelopes and subtract it from $h_{i(k-1)}$ to obtain $h_{ik} = h_{i(k-1)} - m_{i(k-1)}$.

- 6) If h_{ik} satisfies the stop-sifting criteria, set $\text{IMF}_i = h_{ik}$; else set $k = k + 1$ and go to step (4).
- 7) Set $r_i = r_{(i-1)} - \text{IMF}_i$.
- 8) If the number of extrema of r_i is less than three or the desired number of IMFs is extracted, the decomposition process is finished. Otherwise, first set $i = i + 1$, $k = 1$, and then go to step (3).

Therefore, the given signal $y(t)$ is decomposed into several IMFs and a final residual r_N as follows:

$$y(t) = \sum_{i=1}^N \text{IMF}_i(t) + r_N(t). \quad (1)$$

The EMD method is an adjustable algorithm and can be improved by several approaches, such as optimizing the method of interpolation, adjusting the stopping criteria for the sifting process, changing the end conditions of the signal, and selecting new sets of interpolation points instead of signal's extrema.

Various interpolation methods have been proposed so far. The well-known cubic spline is the first and most widely used one (HUANG *et al.*, 1998), but other methods such as B-splines (BOUCHIKHI, BOUDRAA, 2012; CHEN *et al.*, 2006; WANG, LI, 2013), cubic and quartic Hermite (CHU *et al.*, 2012; LI *et al.*, 2015b; GUO, DENG, 2017), Akima interpolation (EGAMBARAM *et al.*, 2016), power function (QIN, ZHONG, 2006), rational splines (PEGRAM *et al.*, 2008), cubic trigonometric cardinal spline (ZHAO *et al.*, 2017), cubic trigonometric B-spline (LI *et al.*, 2018) and nonpolynomial spline (SINGH *et al.*, 2014) have also been employed.

To generate physically meaningful IMFs, an accurate stopping criterion should be defined. At least four stopping criteria for the sifting process have been introduced so far (HUANG *et al.*, 1998; 2003; WU, HUANG, 2009; RILLING *et al.*, 2003). In this paper, the criterion introduced by RILLING *et al.* (2003) is used. They substituted the second condition of the IMF for a more practical one by defining the relative mean as follows:

$$\sigma(t) = \left| \frac{\text{env}_{\max}(t) + \text{env}_{\min}(t)}{\text{env}_{\max}(t) - \text{env}_{\min}(t)} \right|, \quad (2)$$

where env_{\max} and env_{\min} are the upper and lower envelopes, respectively and $\sigma(t)$ is the ratio of the mean of the upper and lower envelopes to the half distance of them. The sifting is iterated until (1) $\sigma(t) < \theta_1$ for a predetermined fraction α of the signal and (2) $\sigma(t) < \theta_2$ for the total length of the signal. Common values for the thresholds and fraction are $\theta_1 = 0.01$, $\theta_2 = 0.1$, and $\alpha = 0.99$. The first condition guarantees a globally small fluctuation of the mean because the relative mean becomes smaller than 0.01 for at least 99% of the signal. The second condition prevents locally large deviation of the mean because there is no part of the signal in which the relative mean exceeds the value 0.1.

The endpoints of the signal are not necessarily its extrema. Since the envelope curves should pass through the extrema, some boundary extrema should be extended beyond the existing data range. In this paper, we use the reflection method proposed by RILLING *et al.* (2003). Without mirroring, considerable distortions at two ends of IMFs may emerge. This phenomenon is called the end effect of the EMD method (LEI *et al.*, 2013).

3. Special inflection points (SIP)

The EMD algorithm considers a signal as the superposition of some fast and some slow oscillating parts. After extracting the fastest oscillating part, the algorithm considers the remainder as a new signal. This process continues until just the slowest oscillating part remains. Oscillation, in signal processing terminology, refers to a repeated variation of a measured quantity. Repetition immediately brings into mind the concept of extrema. On this basis, the EMD algorithm first detects the extrema and then passes the upper and lower envelopes just through the detected extrema. By subtracting the mean of the lower and upper envelopes, the EMD indeed separates the slower detectable oscillating parts from the signal, step by step. The appearance of extrema plays a very important role in the EMD method. Some researchers believe that EMD can only extract a component if it can detect extrema that are related to it (RILLING, FLANDRIN, 2007). However, the origins of some oscillations are difficult to trace with the extrema alone. In these cases, the inflection points can be used as auxiliary points along with extrema to detect the oscillations. To investigate this claim, consider a simple composite signal, which consists of linear and harmonic components as follows:

$$y(t) = \alpha t + \cos(2\pi t), \quad (3)$$

where α is an arbitrary constant. The time of occurrence of the extrema can be calculated by:

$$t_E = \frac{1}{2\pi} \sin^{-1} \left(\frac{\alpha}{2\pi} \right) + n, \quad n = 0, 1, 2, \dots \quad (4)$$

If $\alpha \leq 2\pi$, the signal can be easily decomposed into its two components, but if $\alpha > 2\pi$, EMD cannot decompose the signal because y is a monotonically increasing function and it does not have any extrema. Although there is no fundamental difference between the two cases with $\alpha = (2\pi)^+$ and $\alpha = (2\pi)^-$, the traditional EMD algorithm treats them entirely differently. This is because of the insistence on using the first derivative of the signal (extremum points) in the traditional EMD and neglecting higher-order derivatives (for example, inflection points). Figure 1 shows this signal for eight different values of α . The minimum, maximum, and inflection points in the time range from 1 to 1.5 s are

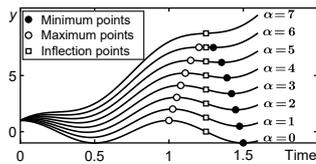


Fig. 1. Convergence of the maximum and minimum points in the signal $y(t) = \alpha t + \cos(2\pi t)$ with increasing α .

demonstrated by the black circles, white circles, and white rectangles, respectively. As α increases, the maximum and minimum points converge on each other and meet at the inflection point. In the uppermost curve, where α equals 7, there are no extrema, but from the oscillation point of view, the signal behaves similarly to the other cases. A closer look at Fig. 1 reveals that the inflection point of the uppermost curve differs from those of the other curves because it is not surrounded by two extrema. This leads us to the concept that in the local absence of extrema, the oscillatory motion can still be realized by the inflection points. However, only finite sets of inflection points should be determined as auxiliary points along with extrema to construct the envelope curves. By considering the arrangement pattern of extrema and inflection points, these special inflection points are identified. The first SIP_1 of a signal can be defined as follows.

An inflection point of a signal is its SIP_1 if, by moving away from that point toward either the left or right directions, we first reach another inflection point before reaching an extremum. We can generalize this definition to SIP_n in the same fashion. The SIP_n of a signal is the SIP_1 of the $(n - 1)$ -th derivative of the original signal.

For example, in a simple sine function, after each maximum, the signal starts to decrease and is forced to change its curvature sign before reaching the minimum point. Hence, in the sine function, every inflection point is always located between two extrema and thus there is not any SIP_n in the simple sine function. There should be more than two inflection points between two successive extrema to generate SIP_1 . To demonstrate SIP_n graphically, consider a two-component signal in the form of

$$y(t) = \cos(2\pi t) + a_r \cos(2\pi f_r t), \quad (5)$$

where a_r is the amplitude ratio and f_r is the frequency ratio of the two components. We have plotted these signals and their first and second derivatives for $a_r = 2$ and $a_r = 4$ in the left and right columns of Fig. 2, respectively. The frequency ratio is considered as $f_r = 0.65$ for both of them. These signals are two simple numerical examples to clarify the graphical interpretation of SIP_1 and SIP_2 . By determining the location of extrema and inflection points of the first signal, i.e., $a_r = 2$, one SIP_1 is detected at $t \cong 1.357$ and

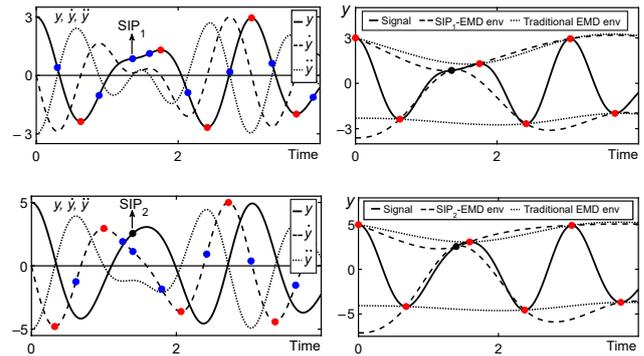


Fig. 2. Appearance of SIP_1 and SIP_2 points in two simple signals. In all curves, red, blue, and black circles represent extrema, inflection points, and SIP_n , respectively. Top row: the two-component signal with $f_r = 0.65$ and $a_r = 2$ in the left and with $f_r = 0.65$ and $a_r = 4$ in the right column with their first and second derivatives. Bottom row: the traditional EMD envelopes (dotted line), SIP_1 -EMD envelopes in the left (dashed line), and SIP_2 -EMD envelopes in the right column (dashed line).

shown in the left top diagram. It is a SIP_1 because it is surrounded by two other inflection points (blue circles) and not two extrema (red circles). For the second signal, i.e., $a_r = 4$, the extrema and inflection points of its first derivative are determined, which leads to detecting one SIP_2 at $t \cong 1.4$, shown in the right top diagram. It is a SIP_2 of the main signal because it is a SIP_1 of the first derivative of the signal. It should be noted that the frequency and amplitude ratios of these signals are selected in such a way that the first signal has just one SIP_1 and the second has just one SIP_2 , but in general, there may be any number of SIP_n points in an arbitrary signal.

The bottom row of Fig. 2 compares the envelopes of the traditional EMD with those of the SIP_n -EMD. In the SIP_n -EMD method, both the upper and lower envelopes should pass through all the SIP_n points. It may seem a little unreasonable that both envelopes pass through a common point. However, Fig. 1 shows how two successive minimum and maximum points converge to one SIP_1 , so we can suppose one SIP_1 as two infinitely close maximum and minimum points that coincide with each other.

The performance of the SIP_1 -EMD and the traditional EMD in separating the first component of the previous two-component signal with $a_r = 2$ and $f_r = 0.65$ has been shown in the left and right column diagrams of Fig. 3, respectively. To provide a better comparison, the main signal ($\cos(2\pi t) + 2 \cos(2\pi \times 0.65t)$) and the higher frequency component ($\cos(2\pi t)$) have been plotted in all diagrams with black continuous and red dotted lines. The output of the decomposition (h_{1k}) after the first, tenth, and hundredth sifting iterations ($k = 1, 10, 100$) has been drawn with the blue lines in the first, second, and third rows, respec-

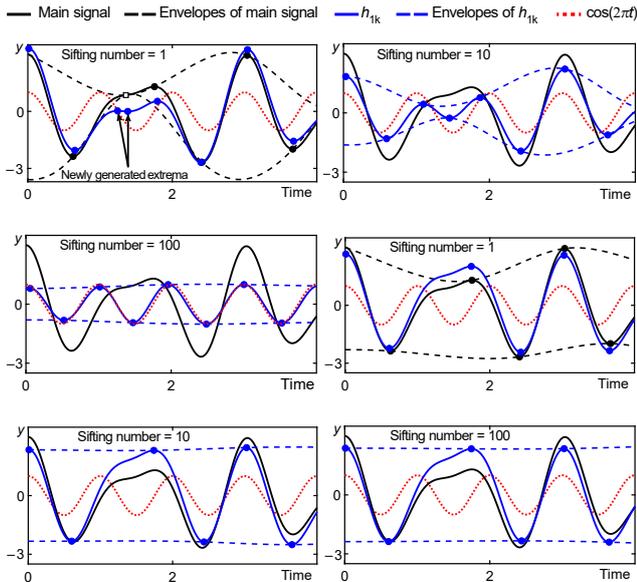


Fig. 3. Performance comparison of the SIP_1 -EMD (left column) and the traditional EMD (right column) methods for the case $y = \cos(2\pi t) + 2\cos(2\pi \times 0.65t)$. In all diagrams, the continuous and dashed black lines represent the main signal and its envelopes, the continuous and dashed blue lines represent h_{1k} and its envelopes and dotted red lines represent the first component ($\cos(2\pi t)$). Black and blue circles demonstrate extrema in all curves, and the SIP_1 point is shown by one white rectangle. The top, middle, and bottom rows are related to the first, tenth, and hundredth sifting iterations ($k = 1, 10$, and 100), respectively.

tively. The main signal has one SIP_1 at $t \cong 1.357$, which is demonstrated with a white rectangle in the left column, top row diagram. Both upper and lower envelopes pass through it and immediately after one sifting iteration, one maximum and one minimum point are generated in h_{11} before and after the SIP_1 at $t \cong 1.242$ and $t \cong 1.383$. These two newly generated extrema help the h_{1k} get closer and closer to the first component. In the third row of the left column, after 100 iterations, the h_{1-100} almost coincides with the first component, which is the desired result. As can be seen in the right column, the traditional EMD is completely unable to decompose this signal. As the number of sifting iterations increases, instead of approaching the first component, h_{1k} moves away from it. Since the upper and lower envelopes of h_{1-100} in the right column are two parallel horizontal lines, there would be no point in continuing the sifting process.

4. SIP_n -EMD method

The main difference between our proposed EMD method in this paper and the traditional EMD is in step (4) of the previously mentioned steps. We develop the SIP_n -EMD method by extending step (4) of the traditional EMD as:

- Identify all local maxima and local minima of $h_{i(k-1)}$, similar to the traditional EMD.
- Take the first derivative of $h_{i(k-1)}$ and find all local minima and local maxima of $\dot{h}_{i(k-1)}$. These points are the inflection points of $h_{i(k-1)}$ and the $SIPs$ of $h_{i(k-1)}$ should be selected from them. This set of points will be referred to as SIP_1 .
- Take the second derivative of $h_{i(k-1)}$ and find all local minima and local maxima of $\ddot{h}_{i(k-1)}$. These points are the inflection points of $\dot{h}_{i(k-1)}$ and the $SIPs$ of $\dot{h}_{i(k-1)}$ should be selected from them. This set of points will be referred to as SIP_2 .
- Repeat this process n times. In the n -th step, take the n th derivative of $h_{i(k-1)}$ and find all local minima and local maxima of $\frac{d^n h_{i(k-1)}}{dt^n}$. These points are the inflection points of the $(n-1)$ -th derivative of $h_{i(k-1)}$ and the $SIPs$ should be selected from them. This set of points will be referred to as SIP_n .
- Pass the upper envelope through the maxima and $SIP_1, SIP_2, \dots, SIP_n$. Similarly, the lower envelope should pass through the minima and $SIP_1, SIP_2, \dots, SIP_n$. Go to step (5) of the traditional EMD.

Theoretically, the SIP_n -EMD method can be extended to n -th derivatives and each new set of SIP_i ($i = 1, 2, \dots, n$) should improve the performance of this method. But since differentiation is a noise amplifying process, in practice, after some order of derivatives, the disadvantages of the additive noise outweigh the advantages of new SIP_n points. Hence, in this paper, we have restricted the procedure to extract SIP_1 and SIP_2 for the synthetic signals and SIP_1 for the real-world signals. When extracting the SIP_2 at step (c), we use a simple low pass filter in order to eliminate the additive noise induced by an extra differentiation made in this step.

5. Practical application

In this section, the aim is to compare the ability of the traditional and proposed EMD method to decompose some synthetic and real signals, including a general form of two-tone signals and a vibration signal of a rolling bearing. All calculations have been performed using MATLAB software.

5.1. Two-tone signals

We rewrite the general form of a two-tone signal here again as:

$$y(t) = C_1(t) + C_2(t) = \cos(2\pi t) + a_r \cos(2\pi f_r t), \quad (6)$$

where $t = [0 : 0.001 : 15]$ is the time vector, a_r is the amplitude ratio, f_r is the frequency ratio ($0 < f_r < 1$),

and C_1 and C_2 are two components of the signal. Because the frequency ratio is limited to change in the range from zero to one, C_1 and C_2 are the higher frequency (HF) and lower frequency (LF) components, respectively. The performance of the traditional EMD algorithm on the decomposition of this signal has been previously studied (RILLING *et al.*, 2003; RILLING, FLANDRIN, 2007).

At first sight, it may be expected that the first and second IMFs should be matched with the HF and LF components, respectively. However, this assumption will not be true for a wide range of frequency and amplitude ratios. In practice, depending on the adjustable parameters of the EMD, one, two, or more IMFs may be extracted. If only one IMF is obtained, it means that the EMD method cannot decompose the signal, and both components are considered as a whole. This is not always a weak point because, in the EMD algorithm, there is a greater emphasis on extracting zero-mean components versus extracting sinusoidal functions. For example, if $a_r = 1$ and $f_r \cong 1$, the superposition of two close tones may be interpreted as a zero-mean amplitude modulation signal, which is not a false interpretation, and in some cases (such as beat phenomenon), may even make more sense (RILLING, FLANDRIN, 2007; DEERING, KAISER, 2005). If two IMFs are extracted, the decomposition has probably been well conducted, and the first and second IMFs are related to the HF and LF components, respectively. But there is still a possibility of mode mixing. If more than two IMFs are obtained, mode mixing will certainly occur, and at least one, if not all, of the extracted IMFs, will deviate from the signal's components. So, even for this simple signal, several possibilities may arise after the EMD analysis. To eliminate the ambiguity caused by the number of extracted modes, only the first IMF will be extracted here for a two-tone signal. The tone separation error (TSE) will be assessed based only on the first IMF according to the following formula (RILLING, FLANDRIN, 2007):

$$\text{TSE} = \frac{\|I_1 - C_1\|_2}{\|C_2\|_2}, \quad (7)$$

where I_1 is the first IMF and $\|\cdot\|_2$ stands for the Euclidean norm. In the case of perfect separation of the two components, the first IMF matches the first component thoroughly, so TSE becomes zero. On the other hand, if the EMD method cannot detect the HF component at all and no decomposition is possible, the first IMF matches the main signal thoroughly, so TSE becomes one. It should be noted that, by the definition of TSE, it is not limited to changes between zero and one. However, it would not become much greater than one. This claim is based on experience and there is not any mathematical constraint on the upper value of TSE.

Figure 4 illustrates TSE values on the $(a_r - f_r)$ plane with more than 100 000 different signals. The white

and black areas present the minimum (zero) and maximum (a little more than one) errors, respectively. The decomposition process of the top, middle and bottom diagrams, has been simulated by the traditional EMD, SIP₁-EMD, and SIP₂-EMD methods, respectively. For all diagrams, 10 fixed sifting iterations are considered stop-sifting criteria. Moreover, all steps of the EMD method have been carried out in the time range from 0 to 15 s, but the values of TSE have been calculated from 0 to 10 s, so relatively large end effect errors at the right side of the signal vanished in the TSE calculations.

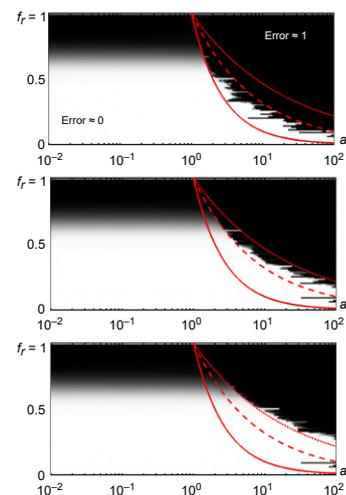


Fig. 4. TSE values on the $(a_r - f_r)$ plane with 10 fixed number of sifting iterations and three border curves, plotted by the continuous $(a_r f_r = 1)$, dashed $(a_r f_r^2 = 1)$, and dotted $(a_r f_r^3 = 1)$ red lines. Top – traditional EMD; middle – SIP₁-EMD; bottom – SIP₂-EMD.

The top diagram, which corresponds to the traditional EMD algorithm, has been previously presented by RILLING and FLANDRIN (2007). We show it here to compare it with diagrams of our proposed EMD method. As can be seen, the EMD always performs well at the beginning of the frequency ratio range ($f_r < 0.2$) and returns the main signal as the first IMF at the end of the frequency ratio range ($f_r > 0.8$). In fact, as the LF component gets closer to the HF component by increasing the frequency ratio, both IMFs start to deviate from their corresponding components, and simultaneously, the first IMF commences to approach the main signal. There is always a frequency ratio in which the EMD behavior is changed and tends to give back the main signal instead of the signal's components. This frequency ratio is referred to as the cut-off frequency ratio and can be seen as a color change from white (lower error region) to black (higher error region). Generally, the larger cut-off frequency ratio indicates a higher tone separation ability. For $a_r < 1$, the cut-off frequency ratio is almost independent of the amplitude ratio and is near 0.67, but for $a_r > 1$, it decreases with the amplitude ratio increasing. Although

in this region, the cut-off frequency ratio has severe fluctuations, RILLING and FLANDRIN (2007) proposed two border curves that bound the cut-off frequency ratio. These curves are $a_r f_r = 1$ and $a_r f_r^2 = 1$, which are plotted by the continuous and dashed red lines, respectively. We also add another curve $a_r f_r^3 = 1$, which is represented by the dotted red line.

The middle diagram illustrates the superiority of the SIP₁-EMD over the traditional EMD. While the cut-off frequency ratio (the border between black and white regions) of the traditional EMD lies between curves of $a_r f_r = 1$ and $a_r f_r^2 = 1$, it completely crosses the $a_r f_r^2 = 1$ curve, applying the SIP₁-EMD, so in the middle diagram, the cut-off frequency ratio is located between curves of $a_r f_r^2 = 1$ and $a_r f_r^3 = 1$. Similarly, the SIP₂-EMD leads the cut-off frequency ratio beyond the curve $a_r f_r^3 = 1$ in the bottom diagram. Considering that the horizontal axes of these diagrams are logarithmic, SIP₁-EMD and SIP₂-EMD methods significantly increase the percentage of the white-colored area. If we define the frequency resolution as the power of distinguishing between two adjacent spectral components, the growth of the white-colored area by the SIP₁-EMD and SIP₂-EMD methods can be interpreted as the improvement of the frequency resolution. It should be noted again that Fig. 4 consists of 100 000 different signals, so the improvements in the results are not random.

To conduct an in-depth analysis of the impact of SIP on EMD, we assign three different constant values to the frequency ratio ($f_r = 0.6, 0.65,$ and 0.7) and investigate the TSE as a function of the amplitude ratio. Abrupt change of the TSE value occurs in a specific amplitude ratio. We define the n -th critical amplitude ratio as:

$$a_{r,c_n} = \frac{1}{f_r^n}, \quad n = 1, 2, 3, \dots \quad (8)$$

Figure 5 shows that the SIP₁-EMD and SIP₂-EMD extend the acceptable error range at least to $[01/f_r^2]$ and $[01/f_r^3]$, respectively.

A question that should be answered here is whether increasing the number of sifting iterations in the traditional EMD can always reduce the TSE. In other words, could increasing the number of sifting iterations substitute for SIP₁-EMD and SIP₂-EMD methods? To answer this question, three two-component signals with the same frequency ratio ($f_r = 0.65$) and different amplitude ratios ($a_r = 2, 3,$ and 4) are selected and their TSE are computed for different sifting iterations.

Figure 6 represents the TSE in terms of the number of sifting iterations. In the first case ($a_r = 2$), the traditional EMD does not return an acceptable response even after 10000 iterations, while both SIP₁-EMD and SIP₂-EMD converge to low error just after 60 iterations. The coincidence of the SIP₁-EMD and SIP₂-EMD curves in the top diagram comes from

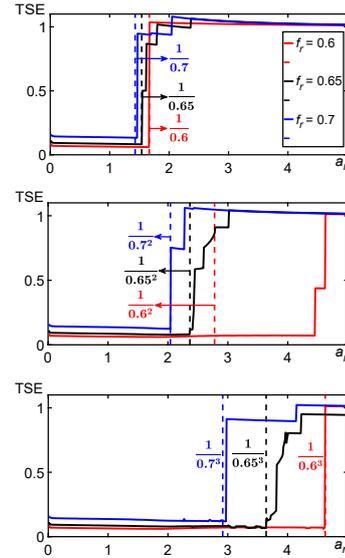


Fig. 5. TSE values as a function of a_r for $f_r = 0.6$ (red), 0.65 (black), and 0.7 (blue), with 100 fixed number of sifting iterations. The critical amplitude ratios are plotted by vertical dashed lines in each diagram. Top – TSE of the traditional EMD, the critical amplitude ratio is $a_{r,c_1} = 1/f_r$; middle – TSE of the SIP₁-EMD, the critical amplitude ratio is $a_{r,c_2} = 1/f_r^2$; bottom – TSE of the SIP₂-EMD, the critical amplitude ratio is $a_{r,c_3} = 1/f_r^3$.

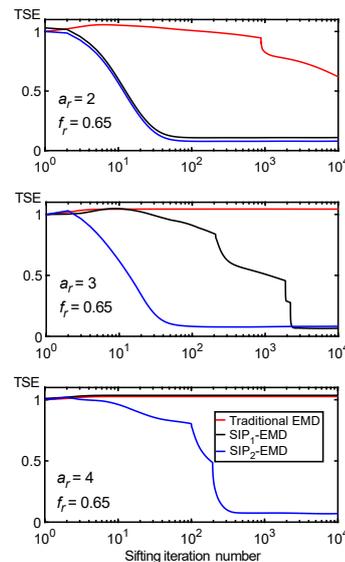


Fig. 6. TSE of three two-tone signals as a function of the number of sifting iterations. Signals have the same frequency ratio ($f_r = 0.65$) but different amplitude ratios (top diagram – $a_r = 2$; middle diagram – $a_r = 3$; bottom diagram $a_r = 4$). TSE values are calculated using the traditional EMD (red), SIP₁-EMD (black), and SIP₂-EMD (blue).

the fact that in the first case, the signal has some SIP₁, but it does not have any SIP₂, so the SIP₂-EMD gains no more benefits here. In the middle diagram ($a_r = 3$), the traditional EMD cannot decompose the signal at all. The SIP₁-EMD and SIP₂-EMD methods

converge to low error levels after 2500 and 70 sifting iterations, respectively. The benefit of the SIP₂-EMD, in this case, is its lower computational cost. In the third case ($a_r = 4$), which is plotted in the bottom diagram, neither the traditional EMD nor the SIP₁-EMD can decompose the signal, but SIP₂-EMD separates the signal's components after 400 iterations. The advantage of the SIP₂-EMD is more apparent here.

5.2. Vibration signal of a rolling bearing

Rolling element bearings are one of the most important components of rotating machines, and their failure may cause machinery breakdowns or safety hazards. Among the existing bearing fault diagnosis methods, vibration signal processing stands out for its accuracy, applicability, and simplicity.

Using EMD (traditional or modified variants) as the main decomposition method or along with other methods for bearing fault diagnosis has been the subject of many studies in recent years (LEI *et al.*, 2013; GUO, DENG, 2017). SUN *et al.* (2021) presented a novel bearing fault diagnosis on the basis of the EMD and improved Chebyshev distance. SHU *et al.* (2022) used modified CEEMDAN and the modified hierarchical amplitude-aware permutation entropy (MHAAPE) to decide whether the bearing was healthy and accurately identify different fault states in the bearings. WANG *et al.* (2022) studied a fault feature extraction method of variable speed rolling bearings based on statistical complexity measures (SCM). The SCM selects the optimal IMF component and evaluates an index for the optimal response of stochastic resonance. ZHENG *et al.* (2022) proposed a spectral envelope-based adaptive empirical Fourier decomposition (SEAEFD) method and compared its performance with some other methods, including EMD.

The occurrence of any fault in any part of the bearing, including the inner race, outer race, rolling elements, and the cage, can be diagnosed by its specific frequency.

In the top row of Fig. 7, a typical acceleration signal $y_{FB}(t)$ of a faulty bearing is shown as a function of time. This signal is taken from the bearing data center of Case Western Reserve University (Case Western Reserve University [CWRU], n.d.). The CWRU Bearing Data Center has provided access to valuable datasets of normal and faulty bearing acceleration signals. These datasets are frequently used as reference signals in the majority of bearing diagnostic studies. The ball pass frequency of the inner race (BPFI) of this bearing is approximately 162 Hz (SKF, n.d.). Further information about the categories of datasets, test stand, bearing details and experimental conditions are presented in (CWRU, n.d.; SMITH, RANDALL, 2015).

To assess the ability of the proposed EMD method to separate different modes, a simple sinusoidal func-

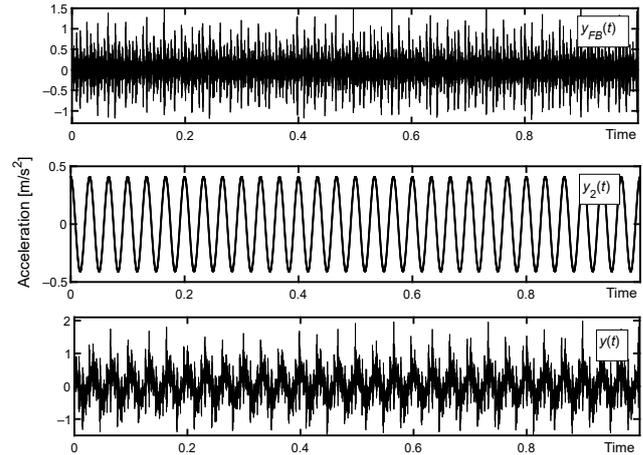


Fig. 7. Acceleration signal of the faulty bearing, $y_{FB}(t)$ (top), simple cosine function with $a_r = 1$ and $f_2 = 30$ Hz, $y_2(t)$ (middle), and a combination of them, $y(t)$ (bottom).

tion $y_2(t)$ with variable amplitude and frequency is added to the faulty bearing signal as follows:

$$y(t) = y_1(t) + y_2(t) = y_{FB}(t) + a_r \left(\sqrt{2} \text{RMS}(y_{FB}(t)) \right) \cos(2\pi f_2 t), \quad (9)$$

where f_2 is the frequency of $y_2(t)$, and a_r is the ratio of the root mean square (RMS) of $y_2(t)$ to the RMS of $y_1(t)$, e.g., $a_r = 1$ means that the energy of the two terms is the same. Thus, the vibration signal $y(t)$ may be considered the superposition of two distinct defects, $y_1(t)$, which is the signal of the faulty bearing with inner race fault, and $y_2(t)$, which can be, for example, a signal of an unbalanced rotor ($f_2 = 30$ Hz). The parameter a_r explains the severity of two defects relative to each other. These three time signals $y_1(t)$, $y_2(t)$, and $y(t)$ are plotted in the top, middle, and bottom rows of Fig. 7, respectively. The duration of the original $y_{FB}(t)$ is more than 10 s, but only the first second of it is used to reduce the computational time in all calculations of this paper.

Unlike the deterministic synthetic signals of the previous sections, the components of the bearing non-deterministic signal are unknown. Hence, we will define some criteria for evaluating the performance of the traditional and SIP-EMD.

Figure 8 shows the first nine IMFs of the vibration signal $y(t)$ (with $a_r = 1$ and $f_2 = 30$ Hz) and their residual after extraction of the ninth IMFs. The results of the traditional EMD and the SIP₁-EMD are displayed in the left and right columns, respectively. In this figure, only the first tenth of the second of IMFs are drawn for better clarity of the graphs. But all calculations are performed within one second, as mentioned before. Due to the noise and sampling frequency restrictions in real-world signals, the traditional EMD is only compared with the SIP₁-EMD in this section.

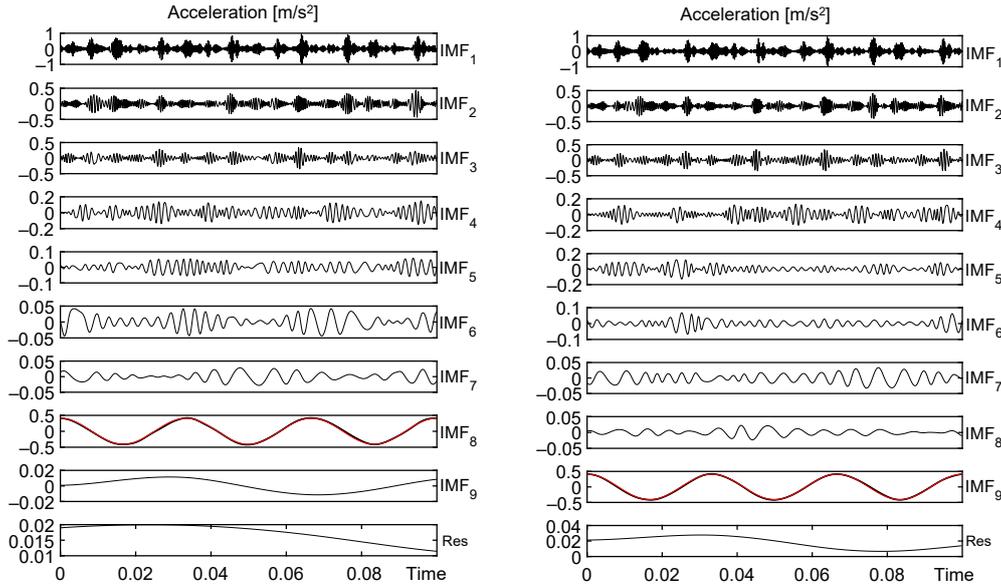


Fig. 8. First nine IMFs and residual of the vibration signal $y(t)$ (with $a_r = 1$ and $f_2 = 30$ Hz) for the traditional EMD (left) and the SIP₁-EMD (right). All IMFs and residuals are plotted with black lines, while $y_2(t)$ is plotted with the red line.

After the extraction of all IMFs, one of them shows the highest similarity with the cosine function $y_2(t)$. From now on, this IMF is called IMF*, which is plotted with black lines and fits well with the ideal cosine function $y_2(t)$ plotted with red lines in the eighth and ninth rows of the left and right columns of Fig. 8, respectively.

The SIP₁-EMD method extracts one more IMF than the traditional EMD. Although it should not be listed as an advantage by itself for either method, it must be taken into account in interpreting the results. The RMS of the difference between the IMF* with the ideal cosine function $y_2(t)$ will be applied as the first evaluation criterion (EC) as follows:

$$EC_1 = RMSE = RMS(IMF^* - y_2). \quad (10)$$

The errors of the traditional and SIP₁-EMD in this example are 0.013 and 0.009, respectively, which shows the better coincidence of the IMF* and y_2 in the SIP₁-EMD method.

To have an idea of energy conservation after decomposition, the index of energy conservation (IEC) is defined as the sum of the energies of all IMFs divided by the energy of the signal without considering the residual (CHEN *et al.*, 2006). In this paper, the parameter IEC is considered as the second EC:

$$EC_2 = IEC = \frac{\sum_{i=1}^N \sum_t |IMF_i(t)|^2}{\sum_t |y(t) - Res(t)|^2}, \quad (11)$$

where N is the number of extracted IMFs. Values greater or less than one indicate the generation or leakage of energy during the sifting process. In this example,

the IECs for the extracted IMFs of the traditional and SIP₁-EMD are 1.022 and 1.005, respectively, which shows better energy conservation in the SIP₁-EMD method.

The existence of spectral peaks at the bearing fault frequencies, their harmonics, and sidebands should be examined to detect the bearing faults. For example, a faulty bearing with a fault on its inner race should have a clear spectral peak at the BPFI, its harmonics, and sidebands. The spectrums of the most important IMFs of the traditional and SIP₁-EMD are displayed in the left and right columns of Fig. 9, respectively. The spectrums of other IMFs are omitted for the sake of brevity. In this special case with $f_2 = 30$ Hz, however, it can be shown empirically that the BPFI (162 Hz) content emerges in one or two IMFs before the IMF* (30 Hz). Here, the amplitudes of frequency spectrums are investigated at the BPFI (162 Hz), its harmonics ($162k, k = 1, 2, 3, \dots$), and sidebands of the BPFI and its harmonics ($162k \pm 30, k = 1, 2, 3, \dots$). The locations of the BPFI harmonics, their sidebands, and the shaft rotation frequency are indicated by the vertical continuous red lines, dashed red lines, and blue lines in Fig. 9, respectively. All amplitudes of this figure are normalized to the RMS of amplitudes for each frequency spectrum. This normalized amplitude is defined as the third EC as follows:

$$EC_3 = \text{Max} \left\{ \frac{IMF_i(\text{BPFI})}{RMS(IMF_i)} \right\}, \quad i = 1, 2, \dots, N, \quad (12)$$

where IMF_i is the frequency spectrum of the i -th IMF. The first three rows of Fig. 9 show that the SIP₁-EMD produces higher spectral peaks at the BPFI than the traditional EMD method.

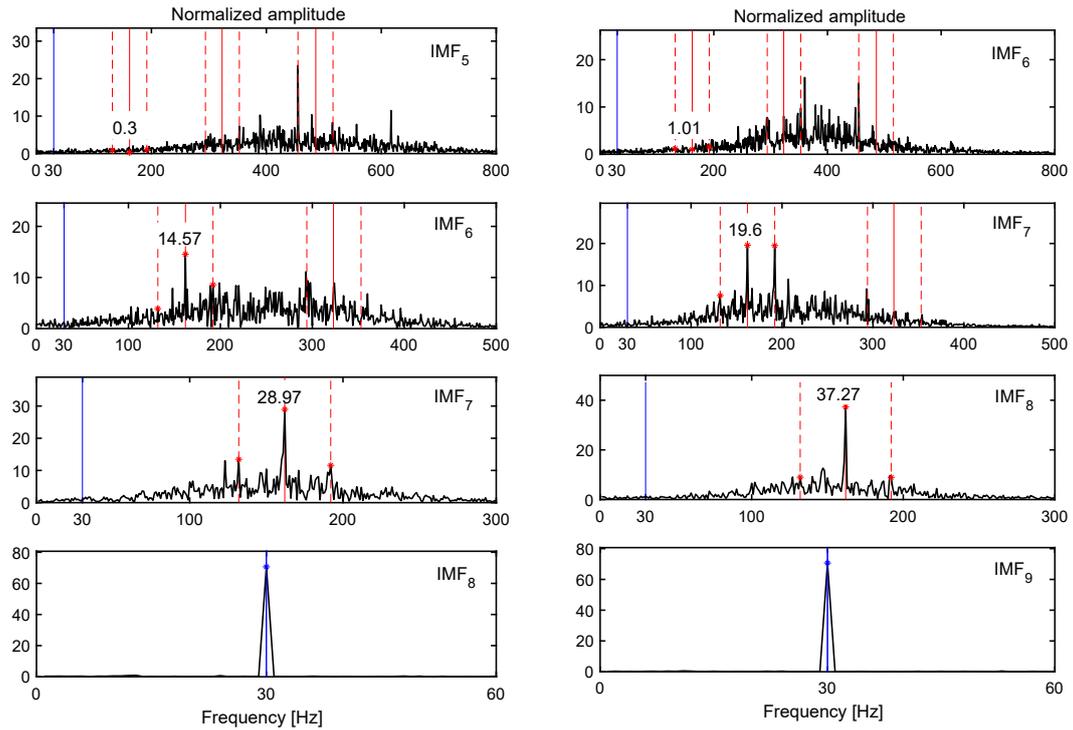


Fig. 9. Spectrums of the most important IMFs of the vibration signal $y(t)$ (with $a_r = 1$ and $f_2 = 30$ Hz) for the traditional EMD (left) and SIP₁-EMD (right). The shaft rotation frequency (30 Hz) is indicated with the vertical blue lines, the BPF (162 Hz) and its harmonics are indicated with the vertical continuous red lines, and the sidebands (± 30 Hz) of the BPF and its harmonics are plotted with the vertical dashed red lines.

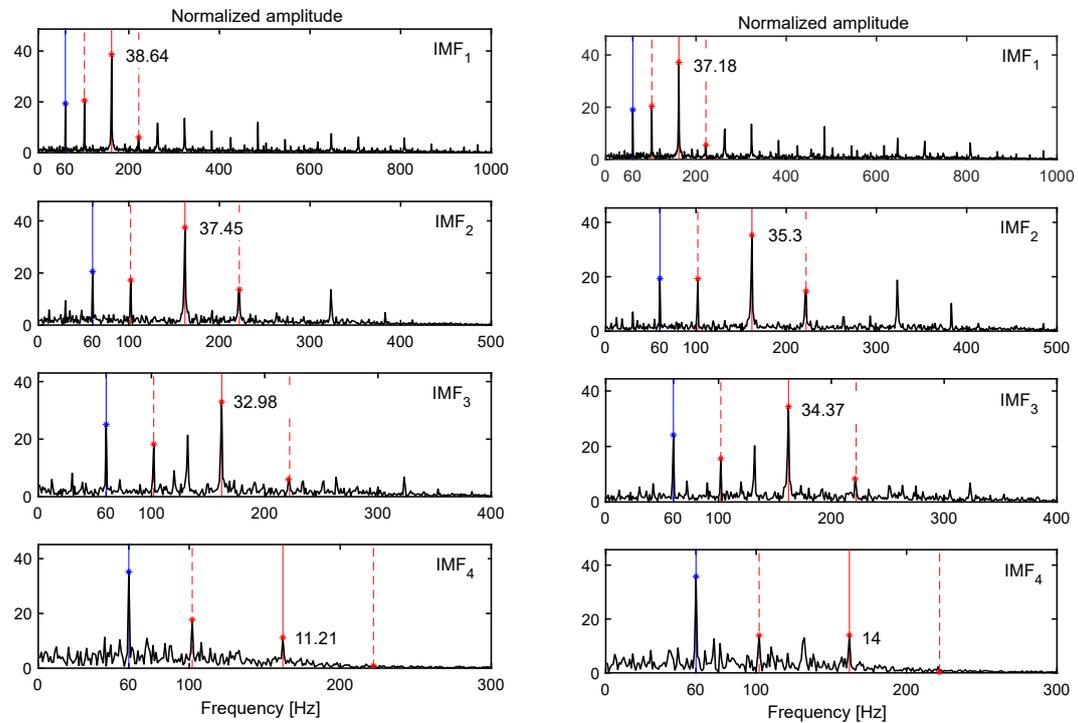


Fig. 10. Spectrums of envelopes for the first four IMFs of the vibration signal $y(t)$ (with $a_r = 1$ and $f_2 = 30$ Hz) extracted by the traditional EMD (left) and SIP₁-EMD (right). The shaft twice rotation frequency (60 Hz) is indicated with the vertical blue lines, the BPF (162 Hz) is indicated with the vertical continuous red lines, and the sidebands of the BPF (162 ± 60 Hz) are plotted with the vertical dashed red lines.

The envelope of the raw signal contains more diagnostic information than the raw signal itself. Hence, the spectrum of the envelope curves for the four IMFs of the traditional and the SIP₁-EMD are plotted in the left and right columns of Fig. 10, respectively. The spectral peaks are clearly evident at the BPF1 (162 Hz) and its sidebands (132 and 192 Hz), especially in the first four IMFs. Then they continue to become shorter until they vanish at the final envelope spectrums. The fourth EC is defined as follows:

$$EC_4 = \text{Max} \left\{ \frac{\text{env}_i(\text{BPF1})}{\text{RMS}(\text{env}_i)} \right\}, \quad i = 1, 2, \dots, N, \quad (13)$$

where env_i is the spectrum of the envelope of the IMF_{*i*}. As shown in Fig. 10, the traditional EMD method produces higher spectral peaks at the first and second IMF envelopes, while the SIP₁-EMD method produces higher spectral peaks at the third and fourth IMF envelopes. This is partly due to the abovementioned additional IMF of the SIP₁-EMD method; consequently, the vibratory energy of the faulty bearing is distributed into one more IMF. Hence, the final EC is defined according to the average of spectral peaks:

$$EC_5 = \text{Mean} \left\{ \frac{\text{env}_i(\text{BPF1})}{\text{RMS}(\text{env}_i)} \right\}, \quad i = 1, 2, \dots, 4. \quad (14)$$

The evaluation criteria of the traditional EMD and SIP₁-EMD methods for the decomposition of the faulty bearing vibration signal with and without the added sinusoidal function are presented in Table 1. In the absence of the added cosine function ($a_r = 0$), the first evaluation criterion does not make sense. The better performance in each case is emphasized with a bold font in the table. The results show the superiority of the SIP₁-EMD method to the traditional EMD method in all items except for EC₄, the reason for which was mentioned before.

To investigate the performance of the proposed SIP₁-EMD method in a wider frequency range, the frequency of the sinusoidal term (f_2) is assumed to be a variable parameter that varies from 10 Hz to 100 Hz. In this case, the amplitude ratio is considered a constant value of one ($a_r = 1$). The values of all evaluation criteria are then calculated and plotted with blue and red lines for the traditional EMD and SIP₁-EMD methods, respectively, in Fig. 11. The horizontal axis of all graphs in this figure is the frequency of the sinusoidal term in Hz. The results confirm the superiority of the SIP₁-EMD method to the traditional EMD through the specified frequency range for all evaluation criteria except for EC₄.

Table 1. Evaluation criteria of the traditional EMD and SIP₁-EMD for the decomposition of the vibration signal of the faulty bearing with (right column) and without (middle column) the added cosine term.

Evaluation criteria	$a_r = 0$		$a_r = 1, f_2 = 30 \text{ Hz}$	
	Traditional EMD	SIP-EMD	Traditional EMD	SIP-EMD
EC ₁ : RMS of error	–	–	0.013	0.009
EC ₂ : Index of energy conservation	1.052	1.011	1.022	1.005
EC ₃ : Maximum spectral peak of IMFs	27.82	28.27	28.97	37.27
EC ₄ : Maximum spectral peak of IMF envelopes	38.68	37.23	38.64	37.18
EC ₅ : Average of spectral peaks of IMF envelopes	29.07	30.26	30.07	30.21

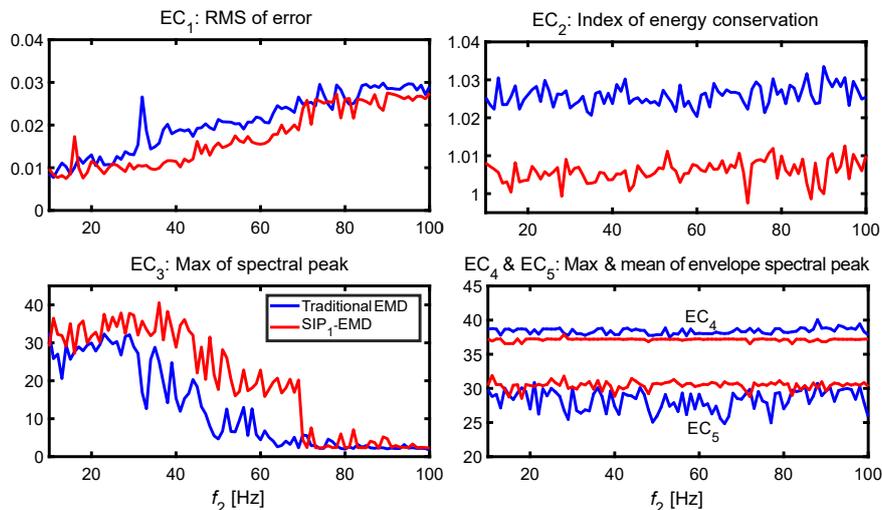


Fig. 11. Evaluation criteria for different signals with a variable frequency of $f_2 = [10 \text{ to } 100] \text{ Hz}$ and a constant amplitude ratio of $a_r = 1$. The curves of the traditional-EMD and SIP₁-EMD methods are plotted with blue and red lines, respectively.

6. Conclusions

This paper evaluated how the interpolation points of the envelope curve affect the EMD results. In the traditional EMD, the extremum points of the signal are used as the interpolation points for the construction of the envelope curves. We propose to add new sets of interpolation points to the existing one. These new points are special inflection points of the signal (SIP₁), special inflection points of the first derivative of the signal (SIP₂), and so on to the special inflection points of the $(n - 1)$ derivative of the signal (SIP _{n}). The results indicate that adding each of these sets to the previous interpolation points makes an obvious improvement in the tone separation ability of the EMD method. This claim was validated by the 2D diagram of the tone separation error on the $(a_r - f_r)$ plane, where the tone separation errors of more than 100 000 two-tone signals are plotted with different frequencies and amplitude ratios. If a signal itself does not have any SIP _{n} and no SIP _{n} appears in h_{ik} during the sifting process, the results of the traditional EMD will coincide with the SIP _{n} -EMD method. However, as soon as at least one SIP _{n} appears in the signal or during the sifting process, the SIP _{n} -EMD obtains a significant reduction in the error diagram, while the traditional EMD method either does not reduce the error at all or reduces the error after many times of sifting iterations. A vibration signal of a rolling element bearing with the inner race fault was analyzed to evaluate the effectiveness of both the traditional and SIP _{n} -EMD methods in the decomposition of a real-world signal. The results show a better performance of the proposed method in the fault feature extraction of the bearing.

Acknowledgments

This work was supported by the Mobarakeh Steel Company.

References

- BOUCHIKHI A., BOUDRAA A.-O. (2012), Multicomponent AM-FM signals analysis based on EMD-B-splines ESA, *Signal Processing*, **92**(9): 2214–2228, doi: [10.1016/j.sigpro.2012.02.014](https://doi.org/10.1016/j.sigpro.2012.02.014).
- Case Western Reserve University (n.d.), *Bearing Data Center*, <https://engineering.case.edu/bearingdatacenter/download-data-file> (access date: 21.02.2023).
- CHEN Q., HUANG N., RIEMENSCHNEIDER S., XU Y. (2006), A B-spline approach for empirical mode decompositions, *Advances in Computational Mathematics*, **24**(1): 171–195, doi: [10.1007/s10444-004-7614-3](https://doi.org/10.1007/s10444-004-7614-3).
- CHU P.C., FAN C., HUANG N. (2012), Compact empirical mode decomposition: An algorithm to reduce mode mixing, end effect, and detrend uncertainty, *Advances in Adaptive Data Analysis*, **4**(3): 1250017, doi: [10.1142/S1793536912500173](https://doi.org/10.1142/S1793536912500173).
- DEERING R., KAISER J.F. (2005), The use of a masking signal to improve empirical mode decomposition, [in:] *Proceedings (ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing*, **4**: 485–488, doi: [10.1109/ICASSP.2005.1416051](https://doi.org/10.1109/ICASSP.2005.1416051).
- EGAMBARAM A., BADRUDDIN N., ASIRVADAM V.S., BEGUM T. (2016), Comparison of envelope interpolation techniques in empirical mode decomposition (EMD) for eyeblink artifact removal from EEG, [in:] *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 590–595, doi: [10.1109/IECBES.2016.7843518](https://doi.org/10.1109/IECBES.2016.7843518).
- GUO T., DENG Z. (2017), An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing, *Applied Acoustics*, **127**: 46–62, doi: [10.1016/j.apacoust.2017.05.018](https://doi.org/10.1016/j.apacoust.2017.05.018).
- HUANG N.E. *et al.* (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**(1971): 903–995, doi: [10.1098/rspa.1998.0193](https://doi.org/10.1098/rspa.1998.0193).
- HUANG N.E. *et al.* (2003), A confidence limit for the empirical mode decomposition and Hilbert spectral analysis, *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **459**(2037): 2317–2345, doi: [10.1098/rspa.2003.1123](https://doi.org/10.1098/rspa.2003.1123).
- KOPSINIS Y., McLAUGHLIN S. (2007), Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach, [in:] *IEEE Transactions on Signal Processing*, **56**(1): 1–13, doi: [10.1109/TSP.2007.901155](https://doi.org/10.1109/TSP.2007.901155).
- KOPSINIS Y., McLAUGHLIN S. (2008), Improved EMD using doubly-iterative sifting and high order spline interpolation, *EURASIP Journal on Advances in Signal Processing*, **2008**(1): 128293, doi: [10.1155/2008/128293](https://doi.org/10.1155/2008/128293).
- LEI Y., LIN J., HE Z., ZUO M.J. (2013), A review on empirical mode decomposition in fault diagnosis of rotating machinery, *Mechanical Systems and Signal Processing*, **35**(1–2): 108–126, doi: [10.1016/j.ymssp.2012.09.015](https://doi.org/10.1016/j.ymssp.2012.09.015).
- LI H., QIN X., ZHAO D., CHEN J., WANG P. (2018), An improved empirical mode decomposition method based on the cubic trigonometric B-spline interpolation algorithm, *Applied Mathematics and Computation*, **332**: 406–419, doi: [10.1016/j.amc.2018.02.039](https://doi.org/10.1016/j.amc.2018.02.039).
- LI H., WANG C., ZHAO D. (2015a), An improved EMD and its applications to find the basis functions of EMI signals, *Mathematical Problems in Engineering*, **2015**: 150127, doi: [10.1155/2015/150127](https://doi.org/10.1155/2015/150127).
- LI Y., XU M., WEI Y., HUANG W. (2015b), An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis, *Measurement*, **63**: 330–345, doi: [10.1016/j.measurement.2014.12.021](https://doi.org/10.1016/j.measurement.2014.12.021).

16. PEGRAM G.G.S., PEEL M.C., MCMAHON T.A. (2008), Empirical mode decomposition using rational splines: An application to rainfall time series, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **464**(2094): 1483–1501, doi: [10.1098/rspa.2007.0311](https://doi.org/10.1098/rspa.2007.0311).
17. QIN S.R., ZHONG Y.M. (2006), A new envelope algorithm of Hilbert–Huang Transform, *Mechanical Systems and Signal Processing*, **20**(8): 1941–1952, doi: [10.1016/j.ymsp.2005.07.002](https://doi.org/10.1016/j.ymsp.2005.07.002).
18. RILLING G., FLANDRIN P. (2007), One or two frequencies? The empirical mode decomposition answers, [in:] *IEEE Transactions on Signal Processing*, **56**(1): 85–95, doi: [10.1109/TSP.2007.906771](https://doi.org/10.1109/TSP.2007.906771).
19. RILLING G., FLANDRIN P., GONÇALVÈS P. (2003), On empirical mode decomposition and its algorithms, [in:] *IEEE-EURASIP workshop on nonlinear signal and image processing*, **3**(3): 8–11, Grado.
20. SHU L., DENG H., LIU X., PAN Z. (2022), A Comprehensive working condition identification scheme for rolling bearings based on modified CEEMDAN as well as modified hierarchical amplitude-aware permutation entropy, *Measurement Science and Technology*, **33**(7): 075111, doi: [10.1088/1361-6501/ac5b2c](https://doi.org/10.1088/1361-6501/ac5b2c).
21. SINGH P., JOSHI S.D., PATNEY R.K., SAHA K. (2014), Some studies on nonpolynomial interpolation and error analysis, *Applied Mathematics and Computation*, **244**: 809–821, doi: [10.1016/j.amc.2014.07.049](https://doi.org/10.1016/j.amc.2014.07.049).
22. SKF (n.d.), *Bearing Select*, <https://www.skfbearingselect.com/#/bearing-selection-start> (access date: 21.02.2023).
23. SMITH W.A., RANDALL R.B. (2015), Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study, *Mechanical Systems and Signal Processing*, **64–65**: 100–131, doi: [10.1016/j.ymsp.2015.04.021](https://doi.org/10.1016/j.ymsp.2015.04.021).
24. SUN Y., LI S., WANG X. (2021), Bearing fault diagnosis based on EMD and improved Chebyshev distance in SDP image, *Measurement*, **176**: 109100, doi: [10.1016/j.measurement.2021.109100](https://doi.org/10.1016/j.measurement.2021.109100).
25. TORRES M.E., COLOMINAS M.A., SCHLOTTHAUER G., FLANDRIN P. (2011), A complete ensemble empirical mode decomposition with adaptive noise, [in:] *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4144–4147, doi: [10.1109/ICASSP.2011.5947265](https://doi.org/10.1109/ICASSP.2011.5947265).
26. WANG J., DU G., ZHU Z., SHEN C., HE Q. (2020), Fault diagnosis of rotating machines based on the EMD manifold, *Mechanical Systems and Signal Processing*, **135**: 106443, doi: [10.1016/j.ymsp.2019.106443](https://doi.org/10.1016/j.ymsp.2019.106443).
27. WANG J.-L., LI Z.-J. (2013), Extreme-point symmetric mode decomposition method for data analysis, *Advances in Adaptive Data Analysis*, **5**(3): 1350015, doi: [10.1142/S1793536913500155](https://doi.org/10.1142/S1793536913500155).
28. WANG Z., YANG J., GUO Y. (2022), Unknown fault feature extraction of rolling bearings under variable speed conditions based on statistical complexity measures, *Mechanical Systems and Signal Processing*, **172**: 108964, doi: [10.1016/j.ymsp.2022.108964](https://doi.org/10.1016/j.ymsp.2022.108964).
29. WU Z., HUANG N.E. (2009), Ensemble empirical mode decomposition: A noise-assisted data analysis method, *Advances in Adaptive Data Analysis*, **1**(1): 1–41, doi: [10.1142/S1793536909000047](https://doi.org/10.1142/S1793536909000047).
30. XU Z., HUANG B., LI K. (2010), An alternative envelope approach for empirical mode decomposition, *Digital Signal Processing*, **20**(1): 77–84, doi: [10.1016/j.dsp.2009.06.009](https://doi.org/10.1016/j.dsp.2009.06.009).
31. YANG L., YANG Z., ZHOU F., YANG L. (2014), A novel envelope model based on convex constrained optimization, *Digital Signal Processing*, **29**(1): 138–146, doi: [10.1016/j.dsp.2014.02.017](https://doi.org/10.1016/j.dsp.2014.02.017).
32. YEH J.-R., SHIEH J.-S., HUANG N.E. (2010), Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method, *Advances in Adaptive Data Analysis*, **2**(2): 135–156, doi: [10.1142/S1793536910000422](https://doi.org/10.1142/S1793536910000422).
33. YUAN J., HE Z., NI J., BRZEZINSKI A.J., ZI Y. (2013), Ensemble noise-reconstructed empirical mode decomposition for mechanical fault detection, *Journal of Vibration and Acoustics*, **135**(2): 021011, doi: [10.1115/1.4023138](https://doi.org/10.1115/1.4023138).
34. YUAN J., XU C., ZHAO Q., JIANG H., WENG Y. (2022), High-fidelity noise-reconstructed empirical mode decomposition for mechanical multiple and weak fault extractions, *ISA Transaction*, **129**(Part B): 380–397, doi: [10.1016/j.isatra.2022.02.017](https://doi.org/10.1016/j.isatra.2022.02.017).
35. ZHAO D., HUANG Z., LI H., CHEN J., WANG P. (2017), An improved EEMD method based on the adjustable cubic trigonometric cardinal spline interpolation, *Digital Signal Processing*, **64**: 41–48, doi: [10.1016/j.dsp.2016.12.007](https://doi.org/10.1016/j.dsp.2016.12.007).
36. ZHENG J., CAO S., PAN H., NI Q. (2022), Spectral envelope-based adaptive empirical Fourier decomposition method and its application to rolling bearing fault diagnosis, *ISA Transactions*, **129**(Part B): 476–492, doi: [10.1016/j.isatra.2022.02.049](https://doi.org/10.1016/j.isatra.2022.02.049).

Research Paper

VMD and CNN-Based Classification Model for Infrasound Signal

Quanbo LU, Mei LI*

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

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

(received November 4e, 2022; accepted April 1, 2023)

Infrasound signal classification is vital in geological hazard monitoring systems. The traditional classification approach extracts the features and classifies the infrasound events. However, due to the manual feature extraction, its classification performance is not satisfactory. To deal with this problem, this paper presents a classification model based on variational mode decomposition (VMD) and convolutional neural network (CNN). Firstly, the infrasound signal is processed by VMD to eliminate the noise. Then fast Fourier transform (FFT) is applied to convert the reconstructed signal into a frequency domain image. Finally, a CNN model is established to automatically extract the features and classify the infrasound signals. The experimental results show that the classification accuracy of the proposed classification model is higher than the other model by nearly 5%. Therefore, the proposed approach has excellent robustness under noisy environments and huge potential in geophysical monitoring.

Keywords: infrasound signal; variational mode decomposition; convolutional neural network; fast Fourier transform.



Copyright © 2023 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

Infrasound (≤ 20 Hz) is generated by natural disasters and human activities, including earthquakes, tsunamis, mudslides, tornados, volcano eruptions, nuclear explosions, missile launching, and ship navigation. Infrasound propagates through the atmosphere thousands of kilometers. The attenuation in the propagation process is small and its loss is less than a few thousandths (GI, BROWN, 2017; DE ANGELIS *et al.*, 2019). Consequently, infrasound can serve as monitoring properties for geological hazards.

Infrasound has been widely applied in recent years in geological hazard monitoring. Many scholars began to pay attention to this topic. LENG *et al.* (2017) presented a debris-flow monitoring approach, which relied on the characteristics of infrasound signal. ZHANG *et al.* (2020) designed an infrasound-based device for monitoring landslides. PERTTU *et al.* (2020) demonstrated the use of remote infrasound in estimating the height of volcanic plumes. WITSIL and JOHNSON (2020) extracted the infrasound features from the time and frequency domains. Then the infrasound features were clustered using the k-means algorithm. HAM *et al.*

(2008) used radial basis function (RDF) network as the subnetworks of a parallel neural network classifier bank to classify different infrasound events. With the wide application of data mining classification algorithm in the signal, feature extraction can be performed on the signal to achieve a better classification effect. ZHU *et al.* (2017) employed Hilbert-Huang transform (HHT) to extract the infrasound related to earthquakes. LIU *et al.* (2014) respectively used three kinds of feature extraction techniques (spectral entropy, discrete wavelet transformation (DWT) and HHT) to extract the feature vector of four kinds of infrasound signals. The signal feature extracted by back propagation neural network and support vector machines (SVM) for classification, all of which have higher classification accuracy, were studied (THÜRING *et al.*, 2015; IEZZI *et al.*, 2019).

LIU *et al.* (2021) constructed the feature vector set and applied the K-nearest neighbor (KNN) to identify the landslide infrasound signal. LI *et al.* (2016) performed spectral entropy on the infrasound signals to extract the features and classify the infrasound events. WITSIL *et al.* (2022) introduced a physics-based method that propagates infrasound sources through re-

alistic atmospheres. HUPE *et al.* (2022) provided infrasound data products that can serve as observational data for the atmospheric studies and scientific and civilian applications. WATSON *et al.* (2022) reviewed the progress and future direction of volcano infrasound research. GARCIA *et al.* (2022) presented the first detection of seismic infrasound from a large-magnitude earthquake on a balloon network of barometers. ECKEL *et al.* (2023) used pattern recognition techniques on infrasound signals to monitor volcanic activity. YANG *et al.* (2022) reviewed the research progress on infrasound generation mechanism, monitoring technology, and application. COFFERB and PARKER (2022) studied the infrasound signals in simulated nontornadic and pre-tornadic phase of the tornadic supercells. AS-MING *et al.* (2022) overviewed methods and algorithms for detecting, locating, and discriminating seismic and infrasound events. ZHANG *et al.* (2022) provided an infrasound source localization algorithm for improving the location accuracy of the gas pipeline leakage detection system. GARCIA *et al.* (2021) used coupled pressure/ground deformation methods to search for the infrasound signal. However, it can be difficult to automatically extract the features and classify the infrasound signals with all the above presented approaches.

To address it, a novel method based on variational mode decomposition (VMD) and convolutional neural network (CNN) is proposed for infrasound signal classification. Firstly, the infrasound signal is processed by VMD to eliminate the noise, and fast Fourier transform (FFT) is employed to convert the reconstructed signal into a frequency domain image. Then the obtained frequency domain image is used as the input of the CNN. Finally, a CNN model is established to automatically extract the features and classify the infrasound signals.

The rest of this paper is organized as follows. In Sec. 2, the basic theory of VMD and CNN used in this paper is shortly described. Section 3 compares the performance of the described methods in an experiment. Further, the experiment results are shown through the analysis of different methods in Sec. 4. Finally, conclusions are drawn in Sec. 5.

2. Methods

2.1. VMD

VMD decomposes an input signal f into a group of discrete modes u_k , and each mode is compressed to obtain a central frequency w_k (DRAGOMIRETSKIY, ZOZZO, 2014). The constrained variational model is shown in Eq. (1):

$$\min_{\{u_k\}, \{w_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \quad (1)$$

s.t. $\sum_k u_k = f,$

where $\{u_k\} = \{u_1, \dots, u_k\}$ are the k mode components obtained by decomposition and $\{w_k\} = \{w_1, \dots, w_k\}$ are the center frequencies of each mode; $\delta(t)$ is the Dirac delta function.

The augmented Lagrange function is introduced by Eq. (2), and the solution of Eq. (1) is obtained by the alternating direction method of multipliers:

$$L(\{u_k\}, \{w_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle, \quad (2)$$

where secondary penalty item α ensures the signal reconstruction accuracy under the Gaussian noise, the Lagrange multiplier λ is the constraint value, and $\langle \rangle$ is the inner product calculation.

The VMD algorithm process is as follows:

- Step 1. Set the number of decomposition modes. Initialize frequency domain $\{\widehat{u}_k^1\}$, center frequency $\{w_k^1\}$, and the Lagrange multipliers $\widehat{\lambda}^1$. Modal u_k and center frequency w_k are calculated by Eq. (3) and Eq. (4). Initialize, $\widehat{\lambda}^1$, $n \leftarrow 0$.
- Step 2. Set $n \leftarrow n + 1$, $k \leftarrow k + 1$ and execute the whole cycle. Update \widehat{u}_k and w_k for all $w \geq 0$ to reach the preset decomposition number. When $k = K$, the cycle ends. The updated formula of the narrow-band component and the corresponding center frequency is:

$$\widehat{u}_k^{n+1}(w) = \frac{\widehat{f}(w) - \sum_{i < k} \widehat{u}_i^{n+1}(w) - \sum_{i > k} \widehat{u}_i^n(w) + \frac{\widehat{\lambda}^n(w)}{2}}{1 + 2\alpha(w - w_k^n)^2}, \quad (3)$$

$$w_k^{n+1} = \frac{\int_0^\infty w |\widehat{u}_k^{n+1}(w)|^2 dw}{\int_0^\infty |\widehat{u}_k^{n+1}(w)|^2 dw}. \quad (4)$$

- Step 3. Update λ according to the formula:

$$\widehat{\lambda}^{n+1}(w) \leftarrow \widehat{\lambda}^n(w) + \tau \left(\widehat{f}(w) - \sum_k \widehat{u}_k^{n+1}(w) \right). \quad (5)$$

- Step 4. Return to step 2 and repeat the above process until the whole iterative process meets the constraints, and a series of narrow-band eigenmode component signals are obtained. Equation (6) is the constraint condition, where ε is set to 10⁻⁶:

$$\sum_k \left\| \widehat{u}_k^{n+1} - \widehat{u}_k^n \right\|_2^2 / \left\| \widehat{u}_k^n \right\|_2^2 < \varepsilon. \quad (6)$$

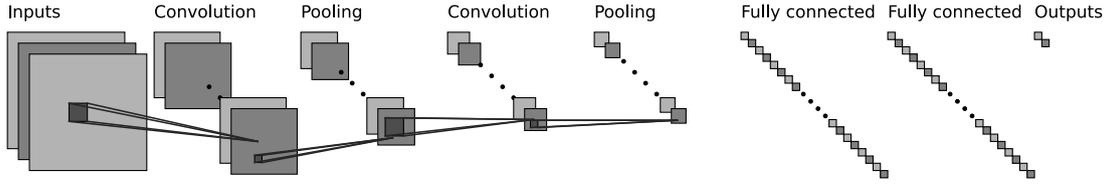


Fig. 1. Structure of CNN.

2.2. CNN

CNN is an important part of deep neural networks (LAWRENCE *et al.*, 1997). It is composed of a multi-level structure that can be trained. Due to its strong feature extraction ability, CNN has been widely used in the field of signal processing. Each level of CNN generally consists of a convolution layer, pooling layer, fully connected layer, and softmax layer. The feature extraction is obtained by multiple alternating operations. Finally, through the fully connected layer and classifier, the infrasound signal classification is realized. The structure of CNN is shown in Fig. 1.

2.2.1. Convolution layer

The convolution layer uses convolution kernels to perform convolution on input data and obtains the feature maps. Each convolution kernel outputs a feature map, which is conformed to a class of the extracted features. The mathematical expression of the convolution is as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right), \quad (7)$$

where l is the l -th convolution layer, x_i^l is the l -th output, x_i^{l-1} is the l -th input, k_{ij}^l is the weight matrix, b_j^l is the bias term, M_j is the j -th convolutional region of the $l-1$ -th feature map, and $f(\cdot)$ is the activation function. In the CNN model, the activation function usually uses the ReLU function. The activation function is represented as:

$$f(x) = \max(0, x). \quad (8)$$

2.2.2. Pooling layer

When the convolution layer finishes the convolution, the pooling layer performs downsampling on the input eigenvectors through the pooling kernels. It can reduce the dimension of the data and further highlight the extracted features. Generally, the pooling operations are divided into two types: max pooling and average pooling. The pooling is expressed as:

$$x_{i+1} = f(\beta^* \text{down}(x_i) + b), \quad (9)$$

where x_i is the input, $\text{down}(\cdot)$ is the pooling function, β is the multiplicative bias, b is the additive bias, and $f(\cdot)$ is the activation function.

2.2.3. Fully connected layer and softmax layer

The fully connected layer and softmax layer are applied in the classification stage of CNN. It can connect the feature maps obtained after a series of convolution and pooling operations into the one-dimensional feature vector. The classification results are gained by the softmax layer. The mathematic model of the fully connected layer and softmax layer can be described as:

$$y^k = \text{softmax}(w^k * x^{k-1} + b^k), \quad (10)$$

where x^{k-1} is the input of the fully connected layer, y^k is the output of the fully connected layer, w^k is the weight coefficient, b^k is the additive bias, and k is the k -th network layer.

2.3. The proposed approach

Infrasound signals collected from the International Monitoring System (IMS) usually exhibit non-linear and non-stationary characteristics. It makes feature extraction difficult and the classification performance unsatisfactory (MAYER *et al.*, 2020). This paper proposes an intelligent infrasound signal classification method based on VMD and CNN. The flowchart of the proposed method is shown in Fig. 2. The general procedures are summarized as follows:

- Step 1. Collect the infrasound signal with sensors.
- Step 2. The collected infrasound signal data is converted into U modes using VMD to eliminate the noise.
- Step 3. FFT is applied to convert the reconstructed signal into a frequency domain image.
- Step 4. The preprocessing data is separated into the training and testing samples. The proposed approach is used to extract deep features from the training samples based on CNN. The trained model is deployed for the infrasound signal classification.
- Step 5. The t -distributed stochastic neighbor embedding (t -SNE) is employed to visualize features in softmax layers (VAN DER MAATEN, HINTON, 2008).
- Step 6. The classification results are presented to compare the performance of different methods.

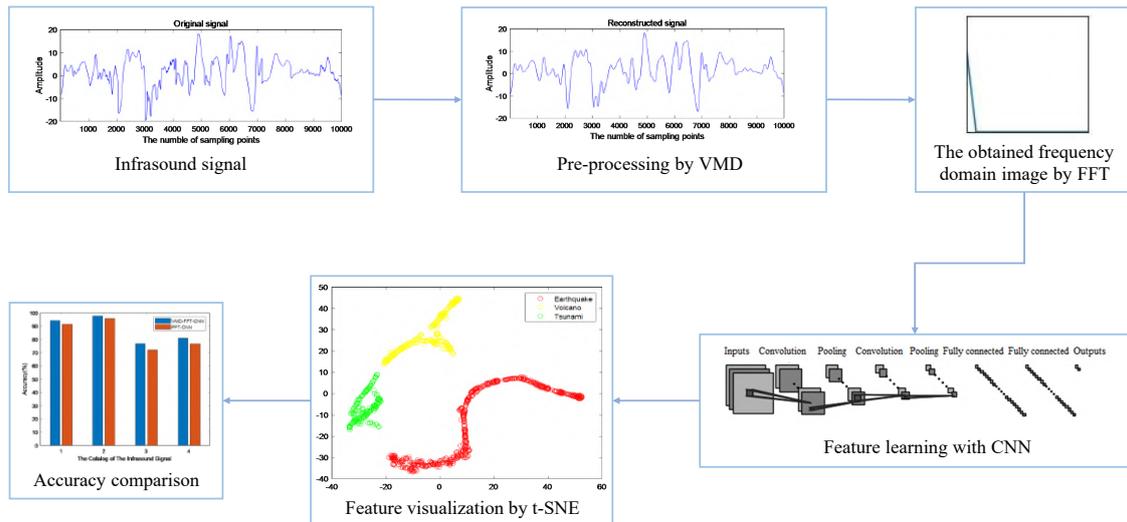


Fig. 2. Flowchart of the proposed approach.

2.4. Data set

The data used in this study comes from IMS with the help of the Comprehensive Nuclear-Test-Ban Treaty Beijing National Data Center. Three categories of infrasound events are classified in this study. The data are collected from six different infrasound sensor arrays with different locations around the world. This study uses 611 sets of data. Table 1 shows the details of infrasound data collected from different areas. The three categories of infrasound events are earthquake, tsunami, and volcano. The sampling frequency of all 611 infrasound signal recordings is 20 Hz. The map of the infrasound stations is showed in Fig. 3.

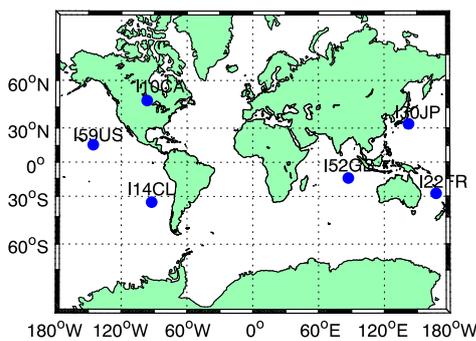


Fig. 3. Map of the infrasound stations.

3. Experiments

3.1. Experiments setup

According to the description of CNN in Subsec. 2.2, the main parameters of the CNN are summarized in Table 2. The simulation verification is devoted to applying the infrasound signal data mentioned above to evaluate the feature learning performance of the proposed CNN model. Every infrasound signal contains 10 400 data points. The data sets are divided into training samples and testing samples. The size of the input

Table 2. The parameters of CNN.

No. of layer	Layer type	Kernel size	Filters
1	Convolution 1	10 × 10	4
2	Maxpooling 1	3 × 3	–
3	Convolution 2	5 × 5	4
4	Maxpooling 2	3 × 3	–
5	Convolution 3	3 × 3	8
6	Convolution 4	3 × 3	8
7	Convolution 5	3 × 3	8
8	Maxpooling 3	3 × 3	–
9	Flattened	–	–
10	Fully-connected	–	–
11	Softmax	–	–

Table 1. Infrasound data summary.

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

map to the CNN model is $129 \times 129 \times 1$. The iteration number is 50.

3.2. Data preprocessing

The VMD is employed to decompose the infrasound signal. When the mode number U is different, the center frequency is different. The relationship between U and the center frequency is depicted in Fig. 4.

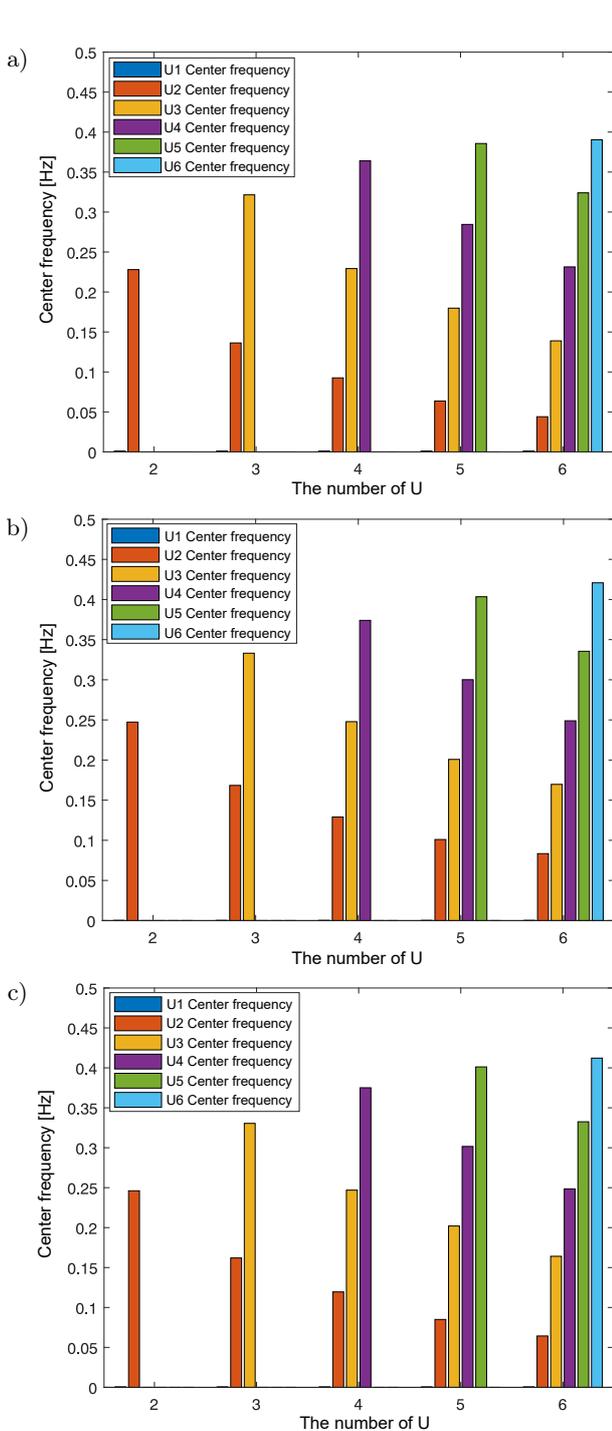


Fig. 4. Relationship between mode number and center frequency: a) earthquake after VMD; b) tsunami after VMD; c) volcano after VMD.

In the earthquake, tsunami, and volcano decomposition results, when the value of U starts from 6, the center frequency is close. This is an over-decomposition phenomenon. Hence, the U value taken in the test is 5. Based on VMD experience, the balance parameter constrained by data fidelity adopts the default value of 2000, and the time step of double rise is 0.1. Figure 5 demonstrates the results of the original signal

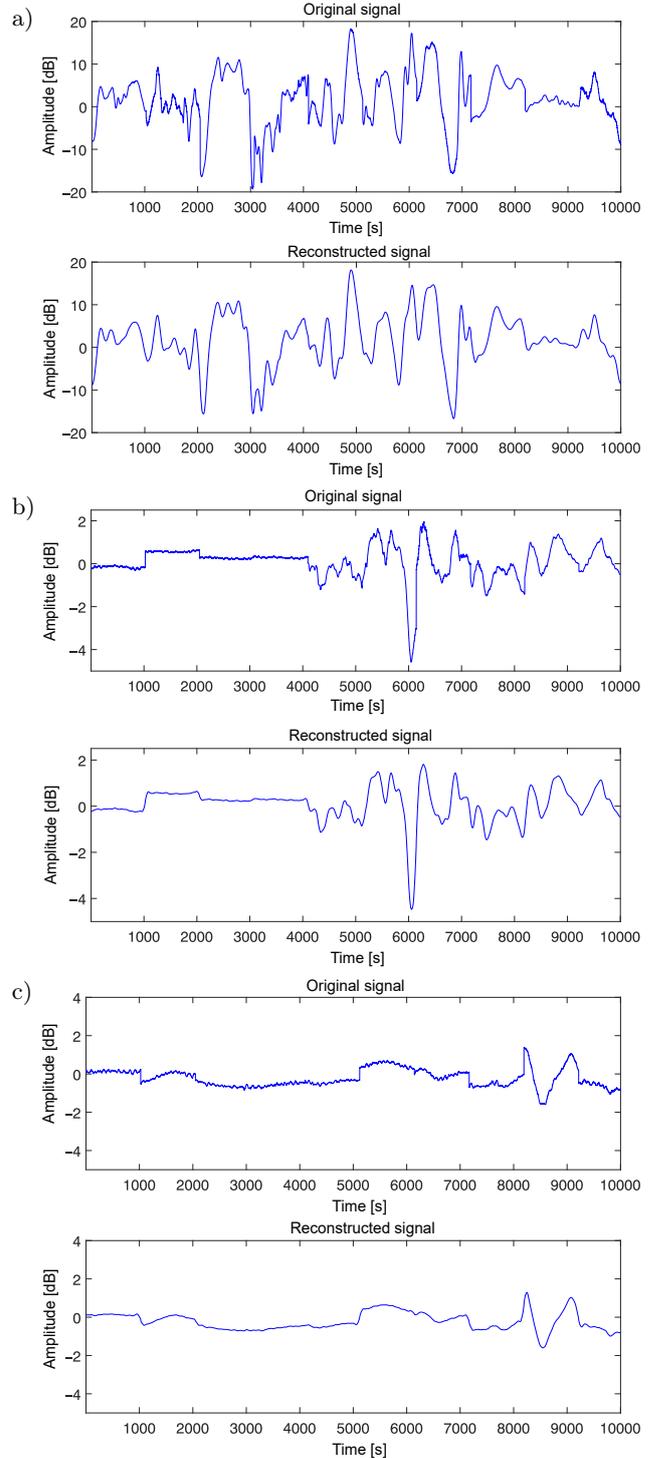


Fig. 5. Original signal and reconstructed signal: a) earthquake; b) tsunami; c) volcano.

and reconstructed signal. Compared with the original signal, the reconstructed signal eliminates the noise. Then, FFT is employed to convert the reconstructed signal into the frequency domain.

3.3. Model application results and analysis

In this study, the CNN structure contains five convolutional layers, three pooling layers, a flattened layer, a fully connected layer, and a softmax layer. The parameters on each layer are presented in Table 2, which are determined based on comparative trials and ex-

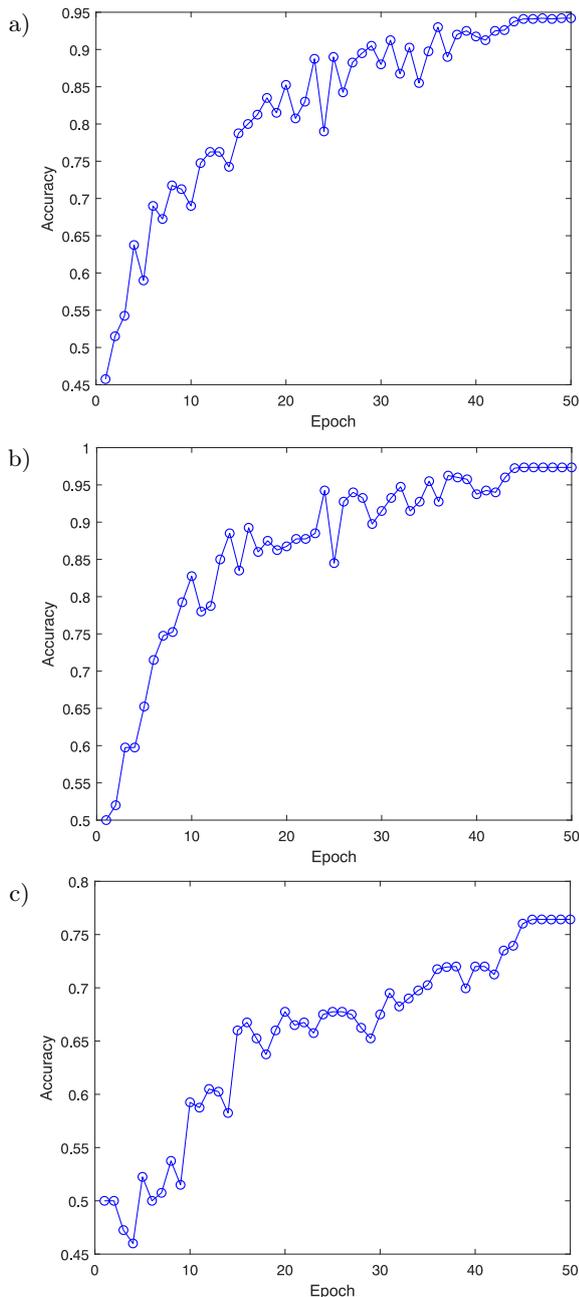


Fig. 6. Classification accuracy of the 2-class catalog after VMD-FFT-CNN: a) earthquake and tsunami; b) earthquake and volcano; c) tsunami and volcano.

perience. The CNN model is written in Python 3.5 and runs on Windows 64 with the Core (TM) i5-8250U CPU and 8G RAM.

The classification accuracy of the 2-class catalog after VMD-FFT-CNN is presented in Fig. 6. Figure 7 shows the classification accuracy of the 3-class catalog after VMD-FFT-CNN, respectively.

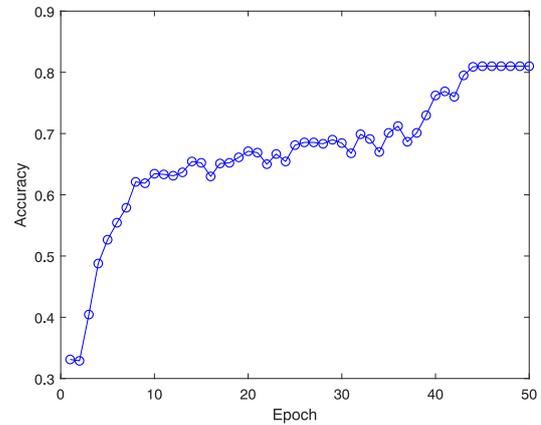


Fig. 7. Classification accuracy of the 3-class catalog after VMD-FFT-CNN.

In order to better illustrate the feature learning process of the proposed model, the t-SNE technique is applied to visualize the output of the softmax layer. It is a machine learning algorithm for high dimensional data visualization using a non-linear dimensionality reduction technique. The feature visualizations of the 2-class catalog and 3-class catalog after VMD-FFT-CNN are shown in Figs. 8 and 9, respectively. It can be seen that the distribution of the points with the same color is relatively closely grouped and easy to distinguish.

To better analyze the classification performance of the proposed method, the infrasound signals are processed by FFT-CNN. The classification accuracy of the 2-class catalog after FFT-CNN is presented in Fig. 10. Figure 11 shows the classification accuracy of the 3-class catalog after FFT-CNN. The feature visualizations of the 2-class catalog and 3-class catalog after FFT-CNN are shown in Figs. 12 and 13, respectively.

In order to verify the stability of the proposed method, the proposed model is tested ten times to derive the final classification result. The classification accuracies on the 2-class catalog consisting of earthquake and tsunami (1), earthquake and volcano (2), and tsunami and volcano (3) are shown in Fig. 14. The classification accuracies of the two architectures on the 3-class catalog consisting of signals from earthquake, tsunami, and volcano (4) are also presented in Fig. 14. As shown in Fig. 14, the classification accuracy of VMD-FFT-CNN is higher than the FFT-CNN model by nearly 5%, which shows that the VMD denoising process is effective. This implies that VMD-FFT-CNN has a good classification performance.

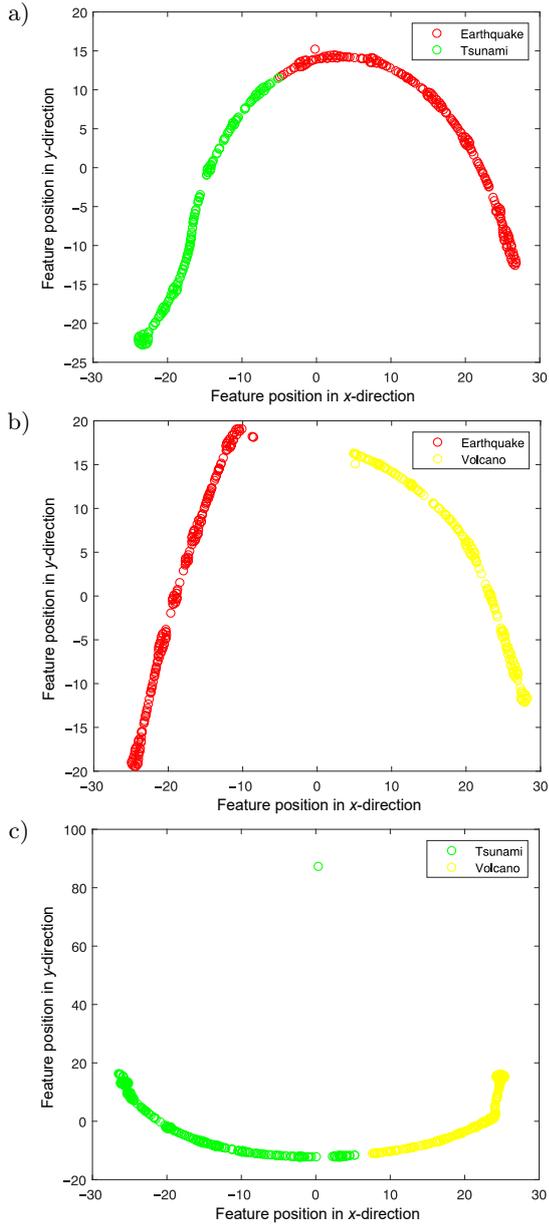


Fig. 8. Feature visualization of the 2-class catalog after VMD-FFT-CNN: a) earthquake and tsunami; b) earthquake and volcano; c) tsunami and volcano.

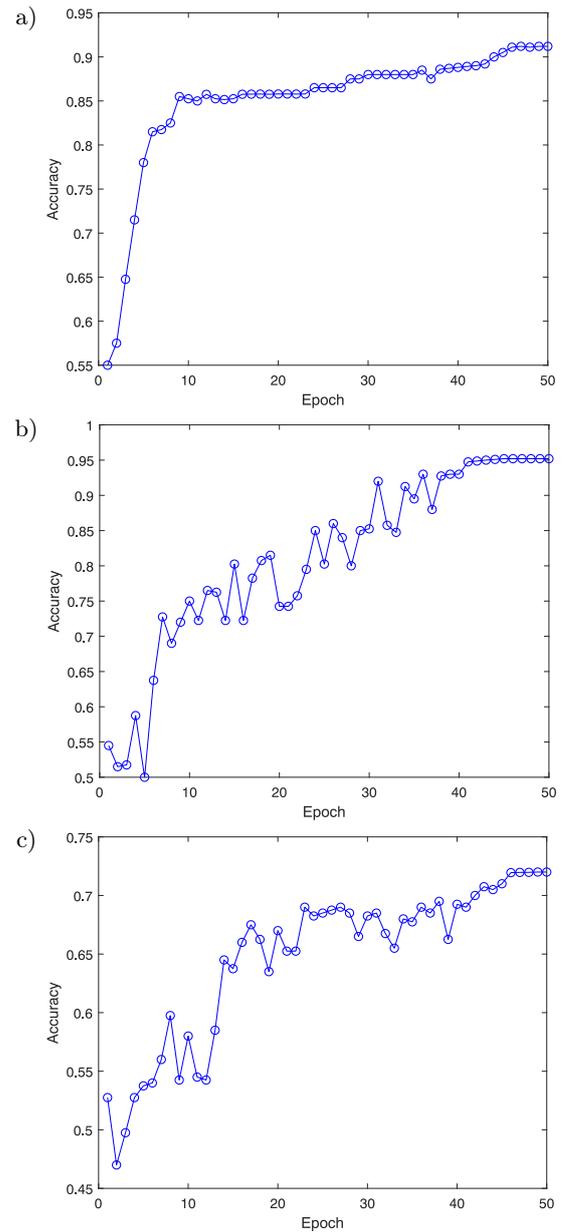


Fig. 10. Classification accuracy of the 2-class catalog after FFT-CNN: a) earthquake and tsunami; b) earthquake and volcano; c) tsunami and volcano.

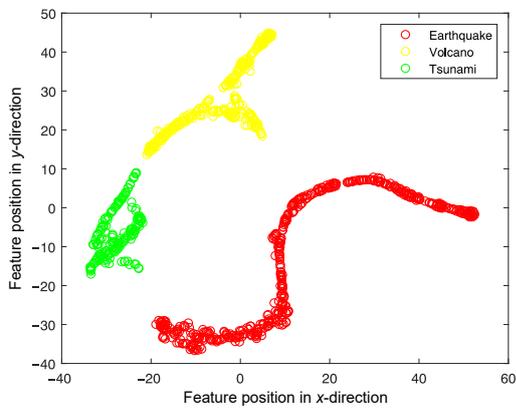


Fig. 9. Feature visualization of the 3-class catalog after VMD-FFT-CNN.

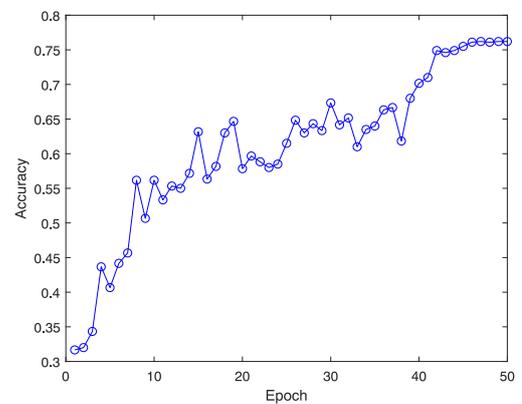


Fig. 11. Classification accuracy of the 3-class catalog after FFT-CNN.

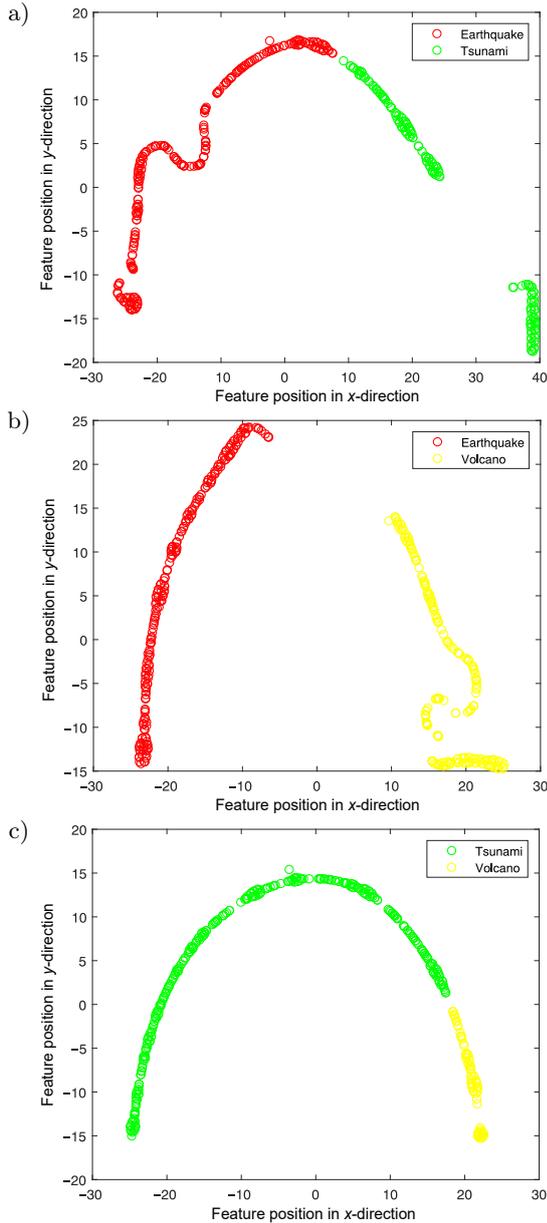


Fig. 12. Feature visualization of the 2-class catalog after FFT-CNN: a) earthquake and tsunami; b) earthquake and volcano; c) tsunami and volcano.

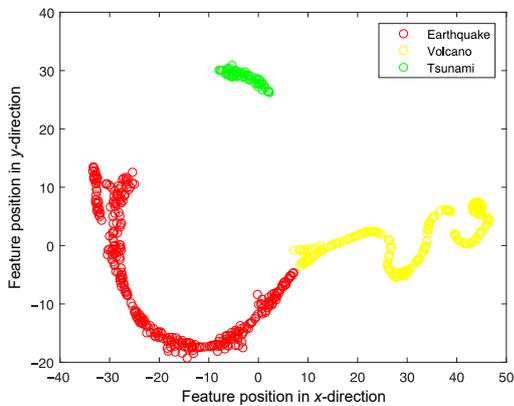


Fig. 13. Feature visualization of the 3-class catalog after FFT-CNN.

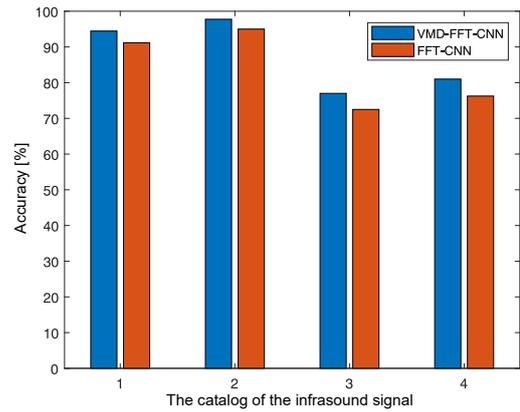


Fig. 14. Accuracy of two methods.

4. Discussion

VMD-FFT-CNN outperforms the FFT-CNN approach in the denoising process. VMD-FFT-CNN has high performance for infrasound signal identification and can achieve 97.75% and 81% of the 2-class catalog and 3-class catalog classification accuracies, respectively. The proposed approach shows excellent performance in classification accuracy compared with other methods and shows a good robustness under noisy environments. For example, the classification accuracy of VMD-FFT-CNN on the 2-class catalog consisting of earthquakes and volcanic increases by 23.25% compared with CNN (ALBERT, LINVILLE, 2020). This result demonstrates that the model presented in this paper has good accuracy for infrasound signal classification. As shown in Fig. 3, the source locations are widespread but their number is small. Due to the limitation of the data, the proposed approach may not be generalized for global hazard monitoring.

5. Conclusion and future work

This paper proposed a valid classification and identification method for the infrasound signal of disasters. The infrasound signal was processed by VMD to eliminate the noise. FFT was used to convert the reconstructed signal into a frequency domain image. A CNN model was constructed for automatically extracting the features and classifying the infrasound signals. The experiment results show that the proposed approach improves the accuracy of geophysical monitoring.

Due to the limitations of the existing conditions, tests can only use small samples and a few infrasound types, which will affect the reliability of the test results. In order to obtain more accurate results, more infrasound data and infrasonic event types should be analyzed. For future work, real-time infrasound signal classification will be carried out and further study on infrasound types will be performed. Deep learning

should be developed for global infrasound signal classification.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 41572347), Graduate Innovation Fund Project of China University of Geosciences, Beijing (YB2021YC020). Many thanks to the Comprehensive Nuclear-Test-Ban Treaty Beijing National Data Center for providing the data.

References

- ALBERT S., LINVILLE L. (2020), Benchmarking current and emerging approaches to infrasound signal classification, *Seismological Research Letters*, **91**(2A): 921–929, doi: [10.1785/0220190116](https://doi.org/10.1785/0220190116).
- ASMING V.E., FEDOROV A.V., FEDOROV I.S., ASMING S.V. (2022), Algorithms for the detection, location, and discrimination of seismic and infrasound events, *Izvestiya, Atmospheric and Oceanic Physics*, **58**(11): 1398–1417, doi: [10.1134/S0001433822110019](https://doi.org/10.1134/S0001433822110019).
- COFFERB B.E., PARKER M.D. (2022), Infrasound signals in simulated nontornadic and pre-tornadic supercells, *The Journal of the Acoustical Society of America*, **151**(2): 939–954, doi: [10.1121/10.0009400](https://doi.org/10.1121/10.0009400).
- DE ANGELIS S., DIAZ-MORENO A., ZUCCARELLO L. (2019), Recent developments and applications of acoustic infrasound to monitor volcanic emissions, *Remote Sensing*, **11**(11): 1302, doi: [10.3390/rs11111302](https://doi.org/10.3390/rs11111302).
- DRAGOMIRETSKIY K., ZOSSO D. (2014), Variational mode decomposition, *IEEE Transactions on Signal Processing*, **62**(3): 531–544, doi: [10.1109/TSP.2013.2288675](https://doi.org/10.1109/TSP.2013.2288675).
- ECKEL F., LANGER H., SCIOTTO M. (2023), Monitoring sources of volcanic activity at Mount Etna using pattern recognition techniques on infrasound signals, *Geophysical Journal International*, **232**(1): 1–16, doi: [10.1093/gji/ggac278](https://doi.org/10.1093/gji/ggac278).
- GARCIA R.F. *et al.* (2021), Search for infrasound signals in InSight data using coupled pressure/ground deformation methods, *Bulletin of the Seismological Society of America*, **111**(6): 3055–3064, doi: [10.1785/0120210079](https://doi.org/10.1785/0120210079).
- GARCIA R.F. *et al.* (2022), Infrasound from large earthquakes recorded on a network of balloons in the stratosphere, *Geophysical Research Letters*, **49**(15): e2022GL098844, doi: [10.1029/2022GL098844](https://doi.org/10.1029/2022GL098844).
- GI N., BROWN P. (2017), Refinement of bolide characteristics from infrasound measurements, *Planetary and Space Science*, **143**: 169–181, doi: [10.1016/j.pss.2017.04.021](https://doi.org/10.1016/j.pss.2017.04.021).
- HAM F.M., REKAB K., ACHARYYA R., LEE Y.-C. (2008), Infrasound signal classification using parallel RBF Neural Networks, *International Journal of Signal and Imaging Systems Engineering*, **1**(3–4): 155–167, doi: [10.1504/IJSISE.2008.026787](https://doi.org/10.1504/IJSISE.2008.026787).
- HUPE P., CERANNA L., LE PICHON A., MATOZA R.S., MIALLE P. (2022), International Monitoring System infrasound data products for atmospheric studies and civilian applications, *Earth System Science Data*, **14**(9): 4201–4230, doi: [10.5194/essd-14-4201-2022](https://doi.org/10.5194/essd-14-4201-2022).
- IEZZI A.M., SCHWAIGER H.F., FEE D., HANEY M.M. (2019), Application of an updated atmospheric model to explore volcano infrasound propagation and detection in Alaska, *Journal of Volcanology and Geothermal Research*, **371**: 192–205, doi: [10.1016/j.jvolgeores.2018.03.009](https://doi.org/10.1016/j.jvolgeores.2018.03.009).
- LAWRENCE S., GILES C.L., TSOI A.C., BACK A.D. (1997), Face recognition: A convolutional neural-network approach, *IEEE Transactions on Neural Networks*, **8**(1): 98–113, doi: [10.1109/72.554195](https://doi.org/10.1109/72.554195).
- LENG X.-P., LIU D.-L., WEI F.-Q., HONG Y., DAI D.-F. (2017), Debris flows monitoring and localization using infrasonic signals, *Journal of Mountain Science*, **14**(7): 1279–1291, doi: [10.1007/s11629-016-3836-3](https://doi.org/10.1007/s11629-016-3836-3).
- LI M., LIU X.Y., LIU X. (2016), Infrasound signal classification based on spectral entropy and support vector machine, *Applied Acoustics*, **113**: 116–120, doi: [10.1016/j.apacoust.2016.06.019](https://doi.org/10.1016/j.apacoust.2016.06.019).
- LIU D., TANG D., ZHANG S., LENG X., HU K., HE L. (2021), Method for feature analysis and intelligent recognition of infrasound signals of soil landslides, *Bulletin of Engineering Geology and the Environment*, **80**(2): 917–932, doi: [10.1007/s10064-020-01982-w](https://doi.org/10.1007/s10064-020-01982-w).
- LIU X.Y., LI M., TANG W., WANG S.C., WU X. (2014), A new classification method of infrasound events using Hilbert-Huang transform and support vector machine, *Mathematical Problems in Engineering*, **2014**(3): 1–6, doi: [10.1155/2014/456818](https://doi.org/10.1155/2014/456818).
- MAYER S., VAN HERWIJNEN A., ULIVIERI G., SCHWEIZER J. (2020), Evaluating the performance of an operational infrasound avalanche detection system at three locations in the Swiss Alps during two winter seasons, *Cold Regions Science and Technology*, **173**: 102962, doi: [10.1016/j.coldregions.2019.102962](https://doi.org/10.1016/j.coldregions.2019.102962).
- PERTTU A., TAISNE B., DE ANGELIS S., ASSINK J.D., TAILPIED D., WILLIAMS R.A. (2020), Estimates of plume height from infrasound for regional volcano monitoring, *Journal of Volcanology and Geothermal Research*, **402**: 106997, doi: [10.1016/j.jvolgeores.2020.106997](https://doi.org/10.1016/j.jvolgeores.2020.106997).
- THÜRING T., SCHOCH M., VAN HERWIJNENA A., SCHWEIZER J. (2015), Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays, *Cold Regions Science and Technology*, **111**: 60–66, doi: [10.1016/j.coldregions.2014.12.014](https://doi.org/10.1016/j.coldregions.2014.12.014).
- VAN DER MAATEN L., HINTON G. (2008), Visualizing data using t-SNE, *Journal of Machine Learning Research*, **9**(11): 2579–2605.
- WATSON L.M. *et al.* (2022), Volcano infrasound: progress and future directions, *Bulletin of Volcanology*, **84**(5): 44, doi: [10.1007/s00445-022-01544-w](https://doi.org/10.1007/s00445-022-01544-w).

23. WITSIL A., FEE D., DICKEY J., PEÑA R., WAXLER R., BLOM P. (2022), Detecting large explosions with machine learning models trained on synthetic infrasound data, *Geophysical Research Letters*, **49**(11): e2022GL097785, doi: [10.1029/2022GL097785](https://doi.org/10.1029/2022GL097785).
24. WITSIL A.J., JOHNSON J.B. (2020), Analyzing continuous infrasound from Stromboli volcano, Italy using unsupervised machine learning, *Computers & Geosciences*, **140**: 104494, doi: [10.1016/j.cageo.2020.104494](https://doi.org/10.1016/j.cageo.2020.104494).
25. YANG X., SU W., JIANG C., BIAN Y. (2022), Research progress on infrasound generation mechanism, monitoring technology and application, *Progress in Geophysics*, **37**(1): 78–93, doi: [10.6038/pg2022FF0081](https://doi.org/10.6038/pg2022FF0081).
26. ZHANG M., GAO L., ZHANG X., ZHANG S. (2022), An infrasound source localisation algorithm for improving location accuracy of gas pipeline leakage detection system, *International Journal of Embedded Systems*, **15**(1): 9–18, doi: [10.1504/IJES.2022.122042](https://doi.org/10.1504/IJES.2022.122042).
27. ZHANG S. *et al.* (2020), Model test: Infrasonic features of porous soil masses as applied to landslide monitoring. *Engineering Geology*, **265**: 105454, doi: [10.1016/j.enggeo.2019.105454](https://doi.org/10.1016/j.enggeo.2019.105454).
28. ZHU X., XU Q., LIU H.X. (2017), Using Hilbert-Huang transform (HHT) to extract infrasound generated by the 2013 Lushan earthquake in China, *Pure and Applied Geophysics*, **174**(3): 865–874, doi: [10.1007/s00024-016-1438-1](https://doi.org/10.1007/s00024-016-1438-1).

Research Paper

Numerical Investigation of the Propagation Characteristics of Surface Transverse Wave Considering Various Quartz Substrate and Electrode Configurations

Chao JIANG^{(1),(2),(3)*}, Xiaoli CAO^{(1),(2)}, Feng YANG^{(1),(2),(3)}, Zejun LIU⁽¹⁾

⁽¹⁾ *School of Computer Science and Information Engineering
Chongqing Technology and Business University
Chongqing, China*

⁽²⁾ *Chongqing Key Laboratory of Intelligent Perception and Blockchain Technology
Chongqing Technology and Business University
Chongqing, China*

⁽³⁾ *Chongqing Engineering Laboratory for Detection, Control and Integrated System
Chongqing Technology and Business University
Chongqing, China*

*Corresponding Author e-mail: chaojiang@ctbu.edu.cn

(received November 22, 2022; accepted March 24, 2023)

Featured with a higher velocity, increased power handling capability, and better aging behavior, surface transverse wave (STW) shows more promising prospects than Rayleigh wave nowadays in various sensing applications. The need to design, optimize, and fabricate the related devices motivates the development of modeling and simulation. For this reason, a three-dimensional (3D) finite element (FE) simulation of STW on quartz, considering the crystal cut angle and the electrode effects, is presented in this study. Firstly, we investigated the effects of quartz's cut angle on the generated waves. Here, the polarized displacements were analyzed to distinguish the wave modes. Secondly, the investigations of the electrode effects on the polarized displacement, phase velocity, and electromechanical coupling factor (K^2) were carried out, for which different material and thickness configurations for the electrodes were considered. Thirdly, to examine the excitation conditions of the generated waves, the admittance responses were inspected. The results showed that not only the crystal cut angle but also the density and the acoustic impedance of the interdigital transducer (IDT) material have a strong influence on the excited waves. This article is the first to analyze STWs considering quartz's cut angle and electrode effect through a 3D FE model. It could provide a helpful and easy way to design, optimize, and fabricate the related surface acoustic wave devices.

Keywords: surface transverse wave (STW); interdigital transducer (IDT); finite element analysis (FEA); quartz.



Copyright © 2023 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

Surface acoustic wave (SAW) devices have been widely investigated and used in signal processing, microfluidics, cell manipulation, and various sensors (FU *et al.*, 2017) due to their advantages in performance, miniaturization, durability, high sensitivity, reliability, and cost-effectiveness (FU *et al.*, 2016; YANTCHEV

et al., 2002). So far, several types of SAWs have been studied and employed, including Rayleigh wave, surface skimming bulk wave (SSBW), surface transverse wave (STW), and Love wave (FAN, JI, 2018; FU *et al.*, 2017; GASO ROCHA *et al.*, 2013).

Rayleigh wave was first theoretically predicted by Lord Rayleigh in 1885 and now it is commonly utilized. However, it suffers from a strong attenuation due to

the mechanical displacement components perpendicular to the substrate surface (JIANG *et al.*, 2019). SSBW is a horizontally polarized bulk shear wave that propagates just beneath the surface of the substrate. Because the wave radiates energy into the bulk of the crystal, it undergoes a considerable acoustic beam spreading loss as well. To minimize the effect, the method of depositing a waveguide on the surface to trap the energy is popularly applied (HASHIMOTO, 2000). As a result, the wave induced through a waveguiding medium using a thin continuous film with a low shear wave velocity is called a Love wave (GASO ROCHA *et al.*, 2013), while the one produced by periodic surface corrugations or gratings on the surface is named an STW (STRASHILOV, YANTCHEV, 2005). In practice applications, both two wave modes have attracted a great deal of interest in recent years. Thanks to the continuous waveguiding mechanism of the Love wave, the related sensors have been widely used for biochemical sensing (RANA *et al.*, 2018; ZHANG *et al.*, 2021). When it comes to STW, it has applications in a wide range of fields, including strain measurement (FU *et al.*, 2016), torque detection (JI *et al.*, 2016), etc. Unlike Rayleigh wave, STW's displacement components are parallel to the substrate surface and perpendicular to the propagation direction (AVRAMOV, 2000), so it is also very suitable for the applications in the field of fluid and liquid sensing. Besides, STWs have an approximately 1.6 times higher velocity than Rayleigh waves, making it possible to develop sensors with smaller dimensions while operating at higher frequencies. Furthermore, STW-based devices can handle a higher power with an improved noise suppression performance, so a strong signal response and a further enhanced working distance and detection precision can be achieved (STRASHILOV, YANTCHEV, 2005).

The development and study of STW have been ongoing for several decades. The earliest report can be traced back to the 1980s. During that period, investigations regarding the propagation of SSBW on shallow corrugation surfaces were widely conducted (AULD, GAGNEPAIN, 1976; AULD, YEH, 1979), establishing the fundamentals of STW. Later on, the investigations on the differences between STW and Rayleigh wave (AULD *et al.*, 1982), the propagation characteristics of STW under period and non-period structures (BAGHAI-WADJI *et al.*, 1988; DANICKI, 1983; RONNEKLEIV, 1986; THOMPSON, AULD, 1986), and the development and fabrication of STW-based resonators (AULD, THOMPSON, 1987; BAGWELL, BRAY, 1987; FLORY, BAER, 1987; STRASHILOV *et al.*, 1997) have been carried out as well. Recently, related micro-strain sensors (FU *et al.*, 2016), torque sensors (JI *et al.*, 2016), vapor sensors (STAHL *et al.*, 2018) and resonators (DOBERSTEIN, VEREMEEV, 2019) have been reported. It has been shown that the STW-based devices not only can achieve a higher Q-factor but also

have extraordinary sensitivity and nonlinearity than their counterparts.

To date, although the theory and application of STWs have been widely studied, to the authors' best knowledge, only a few reports on the modeling and simulation exist. And none of them has investigated the effects of the crystal cut of the piezoelectric substrate, as well as the interdigital transducer (IDT) material and thickness, on the propagation properties of the STW. Motivated by this, we conduct the study by means of three-dimensional (3D) finite element analysis (FEA) approach. Compared to other techniques, such as coupling-of-mode theory (STRASHILOV *et al.*, 1997) and Bloch-Floquet theorem (GAVIGNET *et al.*, 1995), this method not only provides a way for the models to quickly and easily check the results but also shows significant benefits for geometry, material, and physical field configurations.

In this paper, to fully analyze the results, the polarized displacement, wave mode, phase velocity, electromechanical coupling factor (K^2), and admittance of the excited waves are examined and discussed. The rest of the article is arranged as follows. Section 2 describes the methodology, including the considered materials and the simulation details for the FEA model. Section 3 provides comprehensive analyses and discussions of the obtained results. At last, we summarize our work in Conclusion.

2. Materials and methods

In this study, FEA through COMSOL Multiphysics was conducted to investigate the effects of crystal cut and IDTs. Instead of using a complete model, a 3D unit cell model was built to avoid large memory cost and time consumption. The piezoelectric constitutive equation to be solved is written as (JIANG *et al.* 2019):

$$\begin{cases} T_{ij} = c_{ijkl}^E S_{kl} - e_{kij} E_k, \\ D_j = e_{jkl} S_{kl} + \varepsilon_{kl}^S E_k, \end{cases} \quad (1)$$

where T_{ij} , D_j , S_{kl} , and E_k are the stress matrix, electric displacement vector, strain matrix, and applied electrical field vector, respectively. The constant c_{ijkl}^E [N/m²] is the elasticity matrix, ε_{kl}^S [F/m] is the permittivity matrix, and e_{kij} and e_{jkl} [C/m²] are the piezoelectric and inverse piezoelectric coupling matrix, respectively; the superscripts E and S indicate the components of the matrixes that are to be measured under constant electric field and strain, respectively, and the subscripts i , j , k , and l take values from 1, 2, and 3.

2.1. Considered materials

In STW-based devices, quartz is widely used as the piezoelectric substrate. As an anisotropic material, it presents different properties when its crystal cut

varies. On a rotated Y-cut quartz, a Rayleigh wave can be excited when the lengthwise direction of the electrodes is perpendicular to the x -axis of the wafer. On the other hand, the excitation of STWs is also achievable when the electrodes rotate 90° to follow the z -axis (YATSUDA, KOGAI, 2006). However, the rotated Y-cut quartz has many analogues, which can be expressed as $Y + \theta$ for short. Wherein the parameter θ denotes the cut angle. In order to excite STWs more effectively, the crystal cut of Y-cut quartz is examined in this study.

Apart from the piezoelectric substrate, the excited waves also strongly depend on the deposited electrodes, especially on the patterning, material, and thickness. Regarding the patterning, various IDT types have been reported (FU *et al.*, 2017; HASHIMOTO, 2000). Among them, single-electrode-type and split-electrode-type IDTs are frequently used. Thanks to structural simplicity and relatively wide strip width, the single-electrode-type IDTs are selected in this study. With respect to the material, the related electrode with a low mass to minimize the wave damping (FU *et al.*, 2017), a high acoustic impedance to confine the wave energy within the piezoelectric layer (JIANG *et al.*, 2019), and a high conductivity to reduce the series resistance in the transmission of the excitation signal, are preferred. Considering this, five types of frequently used materials were selected according to material properties and practical applications, that is, aluminum (Al), silver (Ag), gold (Au), platinum (Pt), and titanium (Ti). For convenience, Table 1 summarizes the materials with partial material constants (FU *et al.*, 2017).

2.2. Simulation details

In this paper, the velocity and the wavelength λ of STW were set to 5075 m/s and 11.6 μm , respectively,

resulting in a resonant frequency of 437.5 MHz. The metallization ratio of 0.5 was adopted, corresponding to the IDTs with a width of 2.9 μm . And the thickness h_{IDT} was initially set as 0.02λ . For the FEA model, a quartz substrate with a dimension of $\lambda \times 0.5\lambda \times 4\lambda$ was constructed, where the IDTs with a dimension of $0.25\lambda \times 0.5\lambda \times h_{\text{IDT}}$ were built on the top. To absorb the waves at the bottom, a perfect match layer (PML) (PARK, KAYNIA, 2017) with a dimension of $\lambda \times 0.5\lambda \times 0.5\lambda$ was employed as well. The constructed 3D geometry is shown in Fig. 1a, and the corresponding boundary conditions are listed in Table 2.

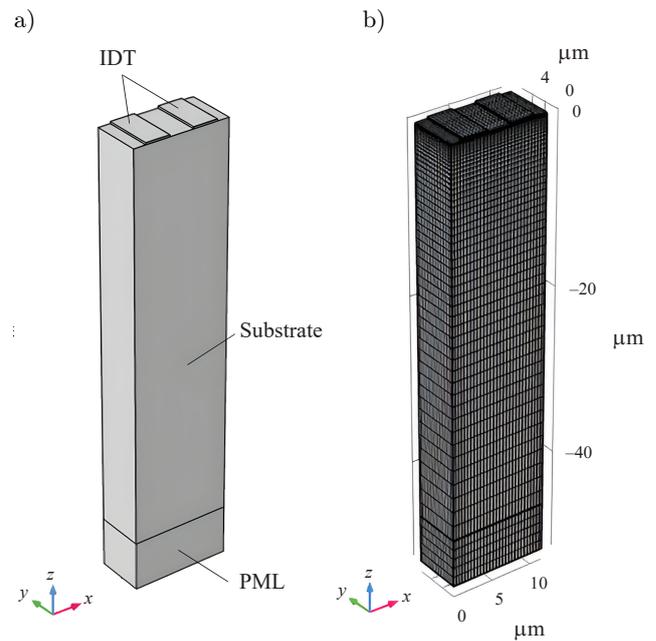


Fig. 1. The 3D unit cell model used for the simulation: a) without mesh; b) with mesh.

Table 1. Some material constants of Al, Ag, Au, Pt, and Ti.

Material constant	Al	Ag	Au	Pt	Ti
Young's modulus [GPa]	76	100	78	170	110
Poisson's ratio	0.33	0.37	0.42	0.38	0.32
Density [g/cm^3]	2.70	10.60	19.32	21.14	4.48
Shear velocity [m/s]	3120	1600	1200	1730	3100
Resistivity [$10^{-8} \Omega$]	2.82	1.59	2.44	10.5	42
Acoustic impedance [$\text{g}/\text{cm}^2 \cdot \text{s} \times 10^5$]	17.65	38.16	51.20	63.42	27.33

Table 2. Assigned mechanical and electrostatic boundary conditions for the simulation.

Geometry	Boundary	Solid mechanics	Electrostatics
Substrate	Left and right sides	$u_{\text{left}} = u_{\text{right}}$	$V_{\text{left}} = V_{\text{right}}$
	Front and back sides	$u_{\text{front}} = u_{\text{back}}$	$V_{\text{front}} = V_{\text{back}}$
	Top	Free	$n \cdot D = 0$
	Bottom	$u_{\text{bottom}} = 0$	$n \cdot D = 0$
IDTs	Left electrode	Free	+5 V
	Right electrode	Free	Ground

To inspect the effect of the crystal cut, the IDT material is firstly fixed and assigned as Al. In addition, a rotated coordinate system depicted by three Euler angles (α, β, γ) , according to the ZXZ convention (JIANG *et al.*, 2019), was also adopted to realize the desired cut for the quartz substrate in the software. The rotations of interest were set as $(90^\circ, 39^\circ, 0)$, $(90^\circ, 54^\circ, 0^\circ)$, $(90^\circ, 90^\circ, 0^\circ)$, $(90^\circ, 124^\circ, 0^\circ)$, $(90^\circ, 125^\circ, 0^\circ)$, $(90^\circ, 126^\circ, 0^\circ)$, $(90^\circ, 127^\circ, 0^\circ)$, $(90^\circ, 128^\circ, 0^\circ)$, $(90^\circ, 141^\circ, 0^\circ)$, and $(90^\circ, 180^\circ, 0^\circ)$. For simplicity, the corresponding cuts were expressed by $Y + \theta$, and hence, θ takes the values from -51° , -36° , 0° , $+34^\circ$, $+35^\circ$, $+36^\circ$, $+37^\circ$, $+38^\circ$, $+51^\circ$, and $+90^\circ$, respectively.

In FEA, the constructed 3D geometries were subdivided into small elements. The dimensions at least five times smaller than λ were initialized for acoustic wave problems (FU *et al.*, 2017). Because the energy of the generated waves mainly concentrates on the top surface, finer meshes were given to the boundaries rather than the bulk. The waves attenuate as the substrate depth increases, so the meshes for the lateral sides were created with gradually varied distribution nodes. Free quadrangular meshes were distributed to the surfaces while the remainder was swept mesh. By using these configurations, the meshes for the 3D unit cell model are created, see Fig. 1b, where 20 917 domain elements, 6013 boundary elements, and 612 edge elements are generated for further calculations.

3. Results and discussion

3.1. Effect of quartz's crystal cut

A notable feature of STW is that the main displacement components are perpendicular to the propagation direction and parallel to the surface. Therefore, close attention was paid to the y -direction component of the waves when searching for the dominant wave mode in a certain cut. Figure 2 presents the simulation results,

wherein the absolute amplitudes of the waves along the y -axis against the frequency with different cut angles are plotted.

Going through the curves, one can find that negligible amplitudes along the y -axis are present for the substrate with the crystal cuts of -51° , -36° , 0° , $+51^\circ$, and $+90^\circ$. While the counterparts on the substrate with a cut angle of $+34^\circ$, $+35^\circ$, $+36^\circ$, $+37^\circ$, and $+38^\circ$ were noticeable, which means that the related Y-cut quartz could excite STWs effectively. The highest peaks were observed at 435.2 MHz and 442.4 MHz, corresponding to the quartz substrate with a cut angle of $Y + 37^\circ$ and $Y + 34^\circ$, respectively. Inspecting the peaks, a movement, shifting from right to left, is also perceived as the cut angle increases. This results from the resonant frequency variation of the dominant wave modes in the substrate. In addition to these, crests with different amplitudes were shown in the spectrum, indicating the differences in wave mode.

To examine the details, sectional views with respect to displacement were reviewed for the $+34^\circ$, $+35^\circ$, $+36^\circ$, $+37^\circ$, and $+38^\circ$ cut substrates. However, due to the similarity of the excited waves on these substrates, only the results obtained using the quartz with a cut of $Y + 37^\circ$ were plotted. As shown in Fig. 3a, the generated waves at 435.2 MHz mainly concentrated on the top surface of the substrate. Wherein the polarized displacements were along the y -axis and in agreement with the feature of STW. Looking at the results posted in Fig. 3b, the maximum polarized displacements appear not only on the surface but also on the bulk of the substrate. This is an SSBW mode. To further examine the result shown in Fig. 3c, a wave analogous to a bulk acoustic wave is observed, where the energy mainly concentrates on the bulk of the substrate. Nevertheless, it is still an SSBW but just with a different mode. In addition to the wave mode, it is also very easy to find the wave energy variation on the surface as the wave mode changes. Compared to the excited STW,

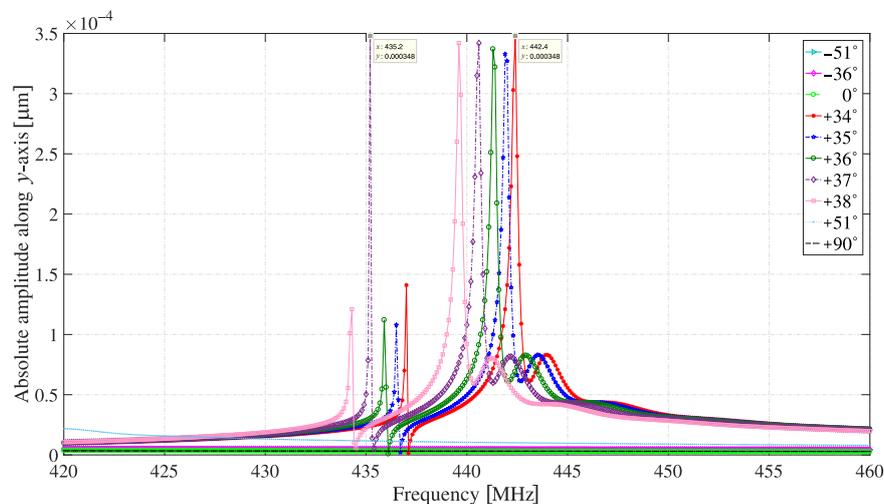


Fig. 2. Absolute amplitudes of the excited waves along the y -axis on different Y-cut quartz substrates.

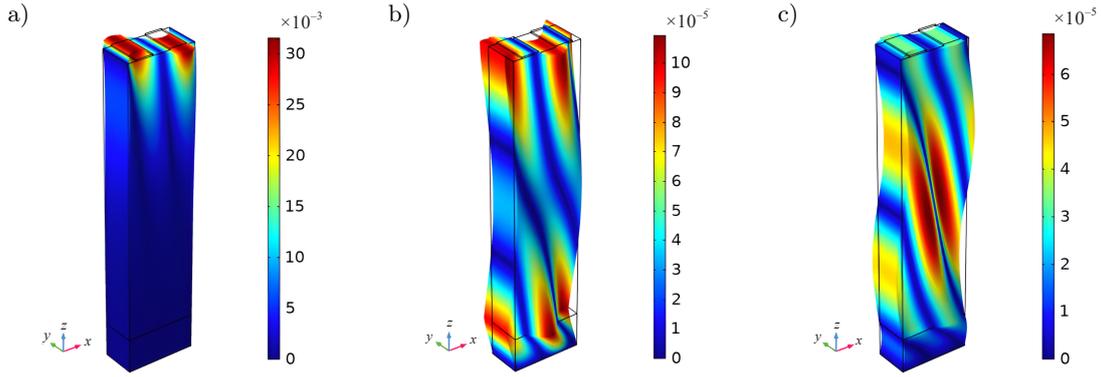


Fig. 3. Snapshots of sectional views from Y+37° cut quartz substrate: a) STW mode at 435.2 MHz; b) and c) SSBW mode at 440.6 MHz and 442.2 MHz, respectively. The unit of the color bars is micrometer $[\mu\text{m}]$.

the wave energy of the SSBW has permeated into the substrates and the generated transversal displacements have been reduced, which shows an attenuation when the wave mode changes from STW to SSBW.

Summarizing the results, it can be concluded that a Y-cut quartz substrate with a crystal cut varying from 34° to 38° could excite STWs effectively, while the counterpart with a cut of -51° , -36° , 0° , $+51^\circ$, and $+90^\circ$ showed less information in the spectrum, indicating futility for generating the waves.

3.2. Effect of electrode material and thickness

Based on the above conclusion, an AT-cut (Y + 35.25°) quartz with various IDT material and thickness con-

figurations was employed to investigate the effects of electrodes on the polarized displacement, velocity, K^2 value, and admittance. The thickness h_{IDT} was set to vary from 10 to 200 nm with a step of 10 nm.

3.2.1. Polarized displacement, velocity, and K^2 value

Figure 4 shows the absolute displacements of the excited waves plotted against varying IDT thicknesses with fixed materials. It can be observed directly that the components along the y -axis were predominant, which indicates that the excited waves are transverse. The maximum displacements of 4.70×10^{-4} , 1.96×10^{-3} , 2.96×10^{-3} , 2.60×10^{-3} , and 3.00×10^{-5} μm were obtained by using Al, Ag, Au, Pt, and Ti IDTs with a thickness of 0, 200, 200, 200, and 0 nm,

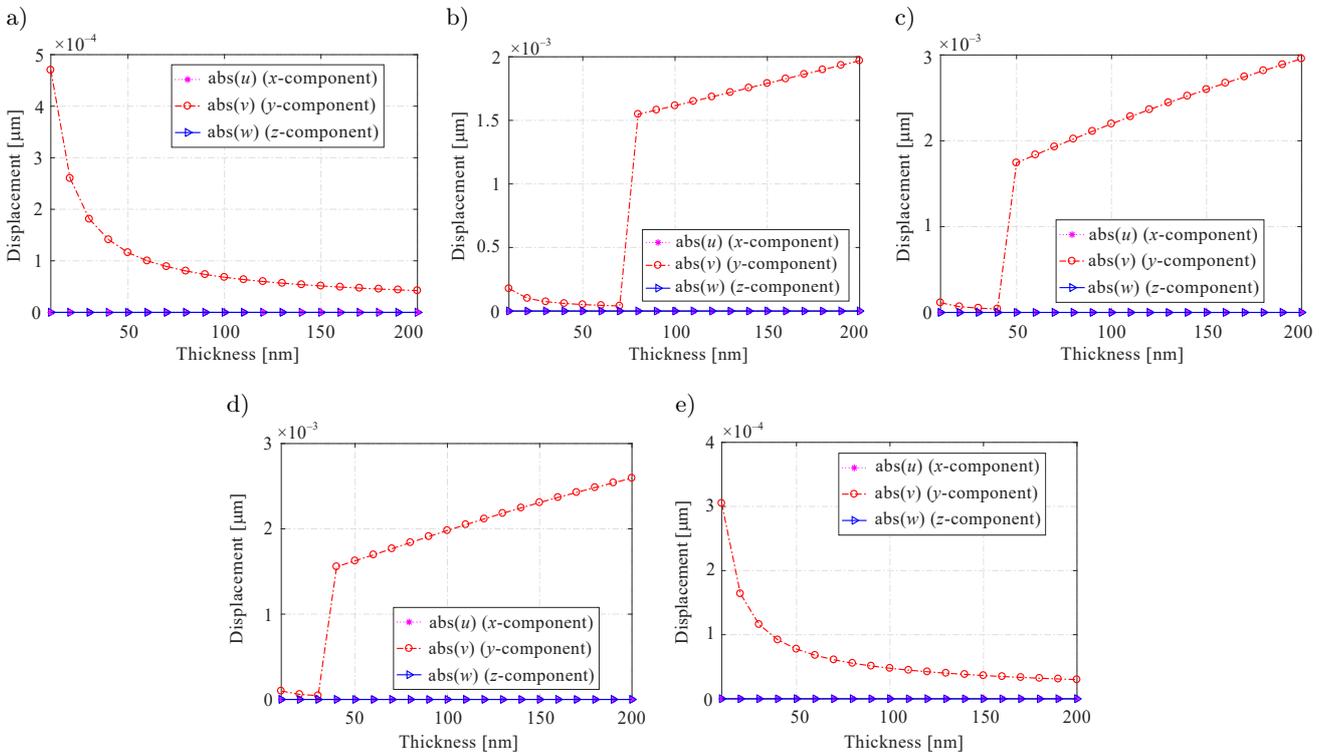


Fig. 4. The absolute displacements of the excited STWs versus varying IDT thickness using different IDT materials: a) Al; b) Ag; c) Au; d) Pt; e) Ti.

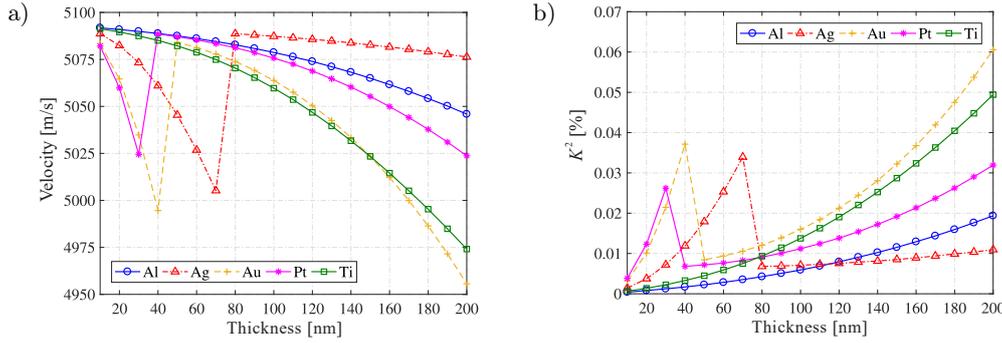


Fig. 5. The variations of velocity (a) and K^2 (b) value versus varying IDT materials and thicknesses.

respectively. Among these images, one can also see that the y -component curves presented in Figs. 4a and 4e showed a continuous descent as the thickness increased. While the ones plotted in the rest subfigures showed a discontinuous variation, that is, the displacement first reduced as the thickness increased, and then bumped as the thickness grew further. The corresponding discontinuous points were observed at 70, 40, and 30 nm, respectively.

The propagation velocity and K^2 with respect to different IDT material and thickness configurations are presented in Fig. 5. The phase velocity and K^2 are calculated from the following equations (FU *et al.*, 2017), respectively:

$$v = \lambda \frac{f_r + f_{ar}}{2}, \quad (2)$$

$$K^2 \approx \frac{2(f_{ar} - f_r)}{f_r}, \quad (3)$$

where f_r and f_{ar} denote the resonance frequency and anti-resonance frequency, respectively.

It is found that as the IDT thickness increased, the velocities showed a similar variation to that of the displacements when different materials were involved. That is to say, the velocities of the waves excited by Al and Ti IDTs gradually slowed down, while the ones of Ag, Au, and Pt IDTs got saltation at the same discontinuous points. Especially, when h_{IDT} approached zero, the maximum velocities were achieved, which were 5092, 5091, 5089, 5083, and 5082 m/s, respectively, corresponding to the virtual electrodes of Al, Ag, Au, Pt, and Ti. These agree with the result in (BIGLER *et al.*, 1991). Similar phenomena but with a reversed tendency are also observed in Fig. 5b. Wherein the maximum K^2 values of 0.011, 0.19, 0.32, 0.49, and 0.06% were obtained by using the Al, Ag, Au, Pt, and Ti electrodes with a thickness of 200 nm, respectively. It is noted that, however, the maximum K^2 values achieved by using Ag, Au, and Pt IDTs at this thickness relate to SSBW mode rather than STW mode, and the corresponding values for the latter mode are 0.034, 0.037, and 0.026%, respectively.

In Figs. 4 and 5, one can find that not only the polarized displacement but also the velocity and K^2 are strongly influenced by IDT material and thickness, especially for Ag, Au, and Pt electrodes. To inspect the details, the 3D mode shapes before and after the saltation are plotted, as shown in Fig. 6.

The images posted in the first row present the mode shapes before the saltation, while the analogues in the second row are the ones after. It can be clearly seen that the excited waves before the saltation are well confined to the substrate surface. They are STWs. While the ones shown in the second row are SSBWs (HASHIMOTO, 2000).

SSBWs can be converted into STWs when a proper IDT structure is used. However, the produced STWs also can be reversed back to SSBWs if the mass-loading effect of the IDTs increases satisfactorily (AVRAMOV, 2000). Note that the densities of Al and Ti are much smaller than those of Ag, Au, and Pt, see Table 1. So, with the same thickness, the mass-loading effects of Al and Ti IDTs are diminished, resulting in a lower coupling between the STWs and substrate. In other words, IDTs using Ag, Au, and Pt are more likely to convert STWs to SSBWs than Al and Ti IDTs when the same geometry is assigned. However, it is noted that, even though the thickness of the Ti IDTs is 200 nm, they did not realize the wave mode conversion yet. Whereas the IDTs using Ag, Au, and Pt have already converted the waves to SSBWs at a thickness of 70, 40, and 30 nm, respectively. To find the cause of this, we simply calculated the equivalent masses $m_e = \rho \cdot h_{\text{IDT}}$, of the Ti, Ag, Au, and Pt IDTs, as shown in Table 3.

Table 3. Equivalent masses of Ag, Au, Pt, and Ti IDTs with fixed thicknesses.

Material	Thickness [h_{IDT}]	Density [ρ]	Equivalent mass [m_e]
Ag	70	10.6	742
Au	40	19.32	772.8
Pt	30	21.14	634.2
Ti	200	4.48	896

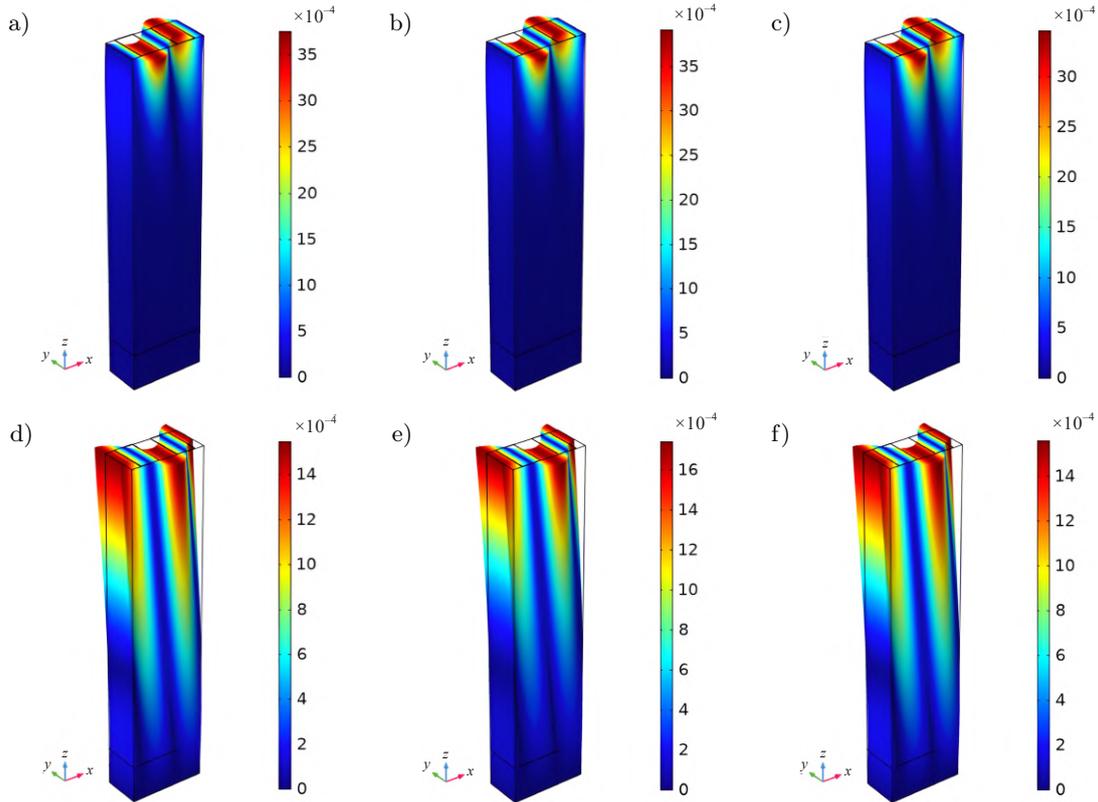


Fig. 6. The 3D mode shapes before (1st row) and after (2nd row) the saltation: a) and d) for Ag IDTs at 431.48 MHz (70 nm) and 438.68 MHz (80 nm); b) and e) for Au IDTs at 430.57 MHz (40 nm) and 438.29 MHz (50 nm); c) and f) for Pt IDTs at 433.14 MHz (30 nm) and 438.67 MHz (40 nm). The measurement unit of the color bars is micrometer [μm].

As one can see, the equivalent mass of the Ti IDTs $m_e = 896$ is larger than that of the rest electrodes. However, the IDTs did not cause any wave mode conversion, see Figs. 4e and 5b. On the contrary, the Pt IDTs with the lowest mass-loading effect $m_e = 634.2$ converted the STWs into SSBWs. This is caused by the difference in acoustic impedance. In Table 1, it can be found that the acoustic impedance of Pt is much higher than that of Ti. Hence, the reflection induced by Pt IDTs is much stronger, which is easier to create a stopband for the excited STWs (AVRAMOV, 2000). Consequently, wave mode conversion and energy attenuation occur. The same result also can be obtained with a comparison of the mass-loading effect, shear velocity, and acoustic impedance for Ag, Au, and Ti.

3.2.2. Admittance

In addition to the effects in mechanics and acoustics, the electrical property or, more clearly, the admittance response is inspected as well. For this, Al and Ti were only selected, due to their relatively lower electrode effect on the excited waves.

Figure 7 shows the admittance responses of the excited waves in $\text{IDT}_{\text{Al}}/\text{quartz}$ and $\text{IDT}_{\text{Ti}}/\text{quartz}$ structures as functions of frequency and IDT material and thickness. In Figs. 7a, 7b, and 7c, one can see that

two modes (M1 and M2) appear. M1 presents a larger amplitude than M2, indicating a stronger resonance in this mode. In addition, frequency shifts were also observed, when the IDTs had the same thickness but different materials. Looking at Figs. 7d and 7e, a similar phenomenon is found when the same material but different thicknesses are employed. This indicates that frequency adjustments through the thickness of the deposited IDTs could be realizable despite the width and interval of the electrodes being fixed. Besides, one can also see that the intervals of the resonant frequencies for these two modes become larger and larger as the IDT thickness increases, while the resonant frequency of M2 barely changes. This results from the widened stopband induced by the thickened IDTs. As the range of the stopband increases, it moves M1 further away from M2. So, for designing a resonator working in the manner of M1, the metallization parameter of the IDTs should be carefully controlled.

Based on the aforementioned admittance responses, the resonance frequencies (f_r) and anti-resonance frequencies (f_{ar}) for the modes are determined, and the corresponding mode shapes are presented, as shown in Figs. 8, 9, and 10. The images posted in the first and third columns are the mode shapes at the resonance frequencies (f_r) of these two waves, and the ones plotted in other columns relate to those at anti-resonance

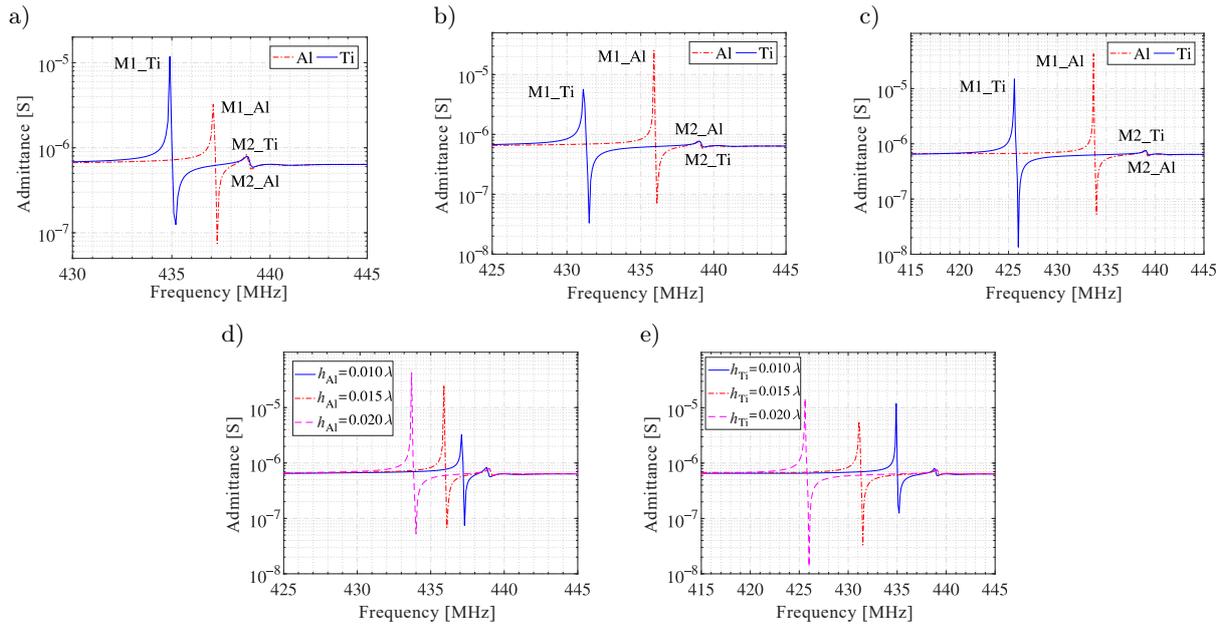


Fig. 7. Admittance responses of the unit cell using different IDT materials and thicknesses: a), b), and c) Al and Ti IDTs with a thickness of 0.010λ , 0.015λ , and 0.020λ , respectively; d) and e) Al and Ti IDTs with varying IDT thickness, respectively.

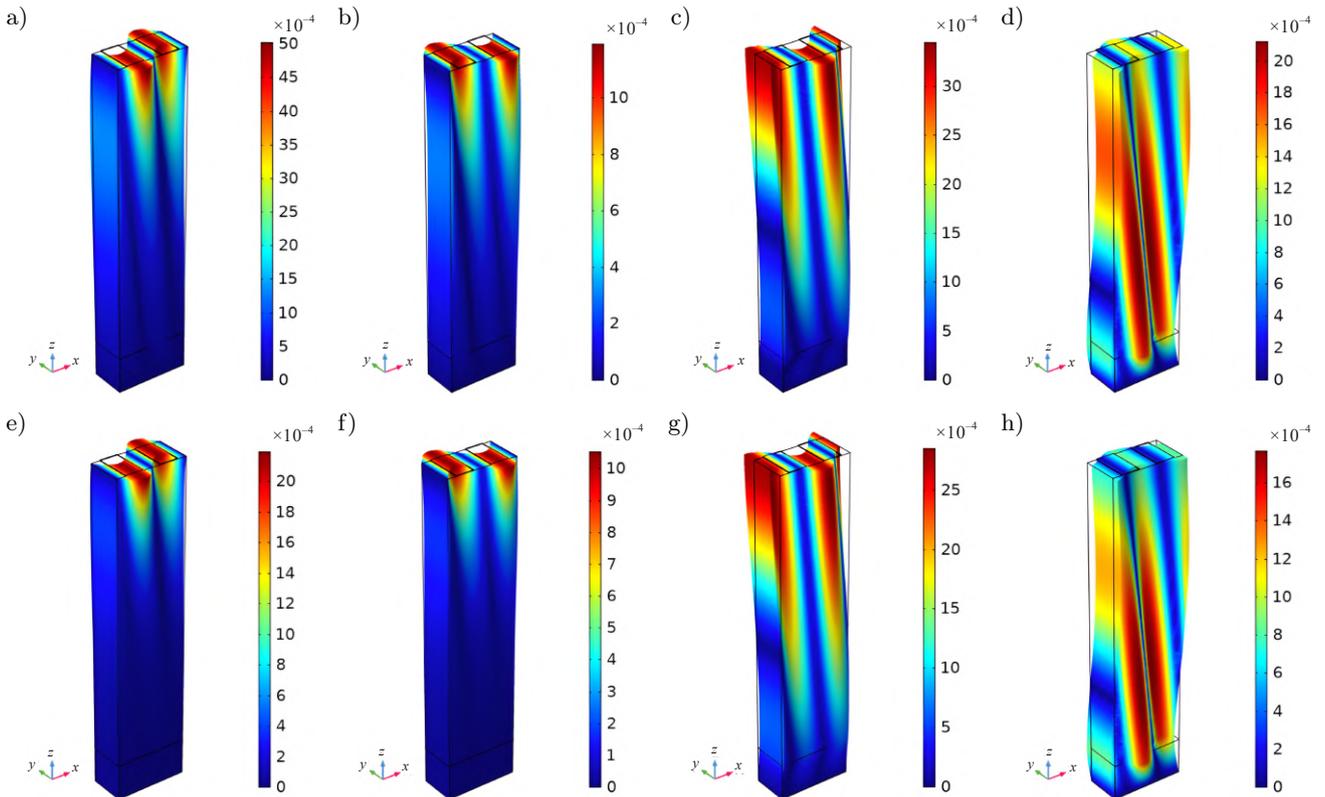


Fig. 8. The 3D mode shapes of the unit cells using 0.010λ -thickness Al IDTs and 0.010λ -thickness Ti IDTs: a)–d) for Al IDTs at 437.1, 437.3, 438.8, and 439.0 MHz, respectively; e)–h) for Ti IDTs at 434.9, 435.2, 438.8, and 439.1 MHz, respectively; the first two columns present the mode shapes for M1 and the remaining two columns for M2. The measurement unit of the color bars is micrometer $[\mu\text{m}]$.

frequencies (f_{ar}). From these images, one can see that the wave modes posted in the first two columns show strong energy concentrations and large transverse po-

larized displacements on the surface. They are STWs, corresponding to the M1 mode in Fig. 7. While the remaining ones shown in the other two columns are

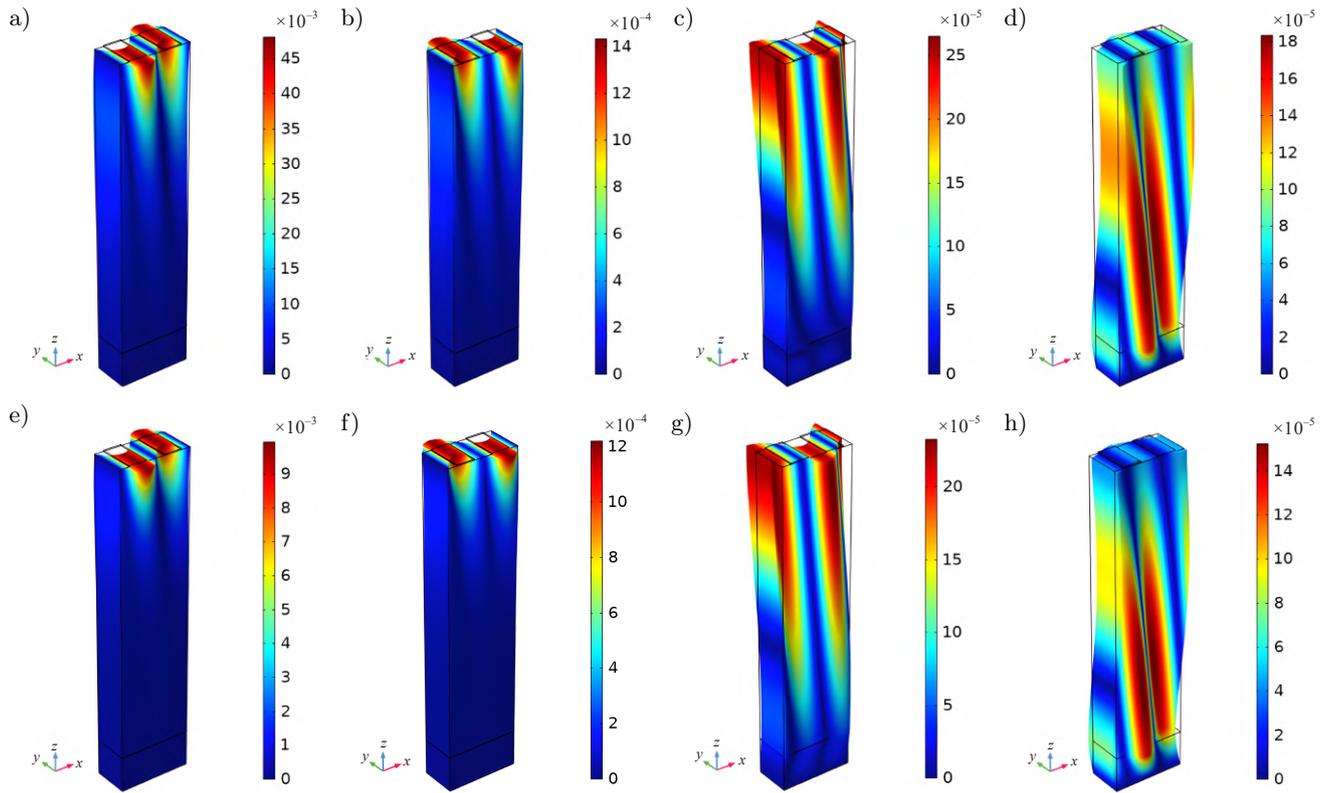


Fig. 9. The 3D mode shapes of the unit cells using 0.015λ -thickness Al IDTs and 0.015λ -thickness Ti IDTs; a)–d) for Al IDTs at 435.9, 436.1, 439, and 439.2 MHz, respectively; e)–h) for Ti IDTs at 431.1, 431.5, 439, and 439.3 MHz, respectively; the first two columns present the mode shapes for M1 and the remaining two columns for M2. The measurement unit of the color bars is micrometer [μm].

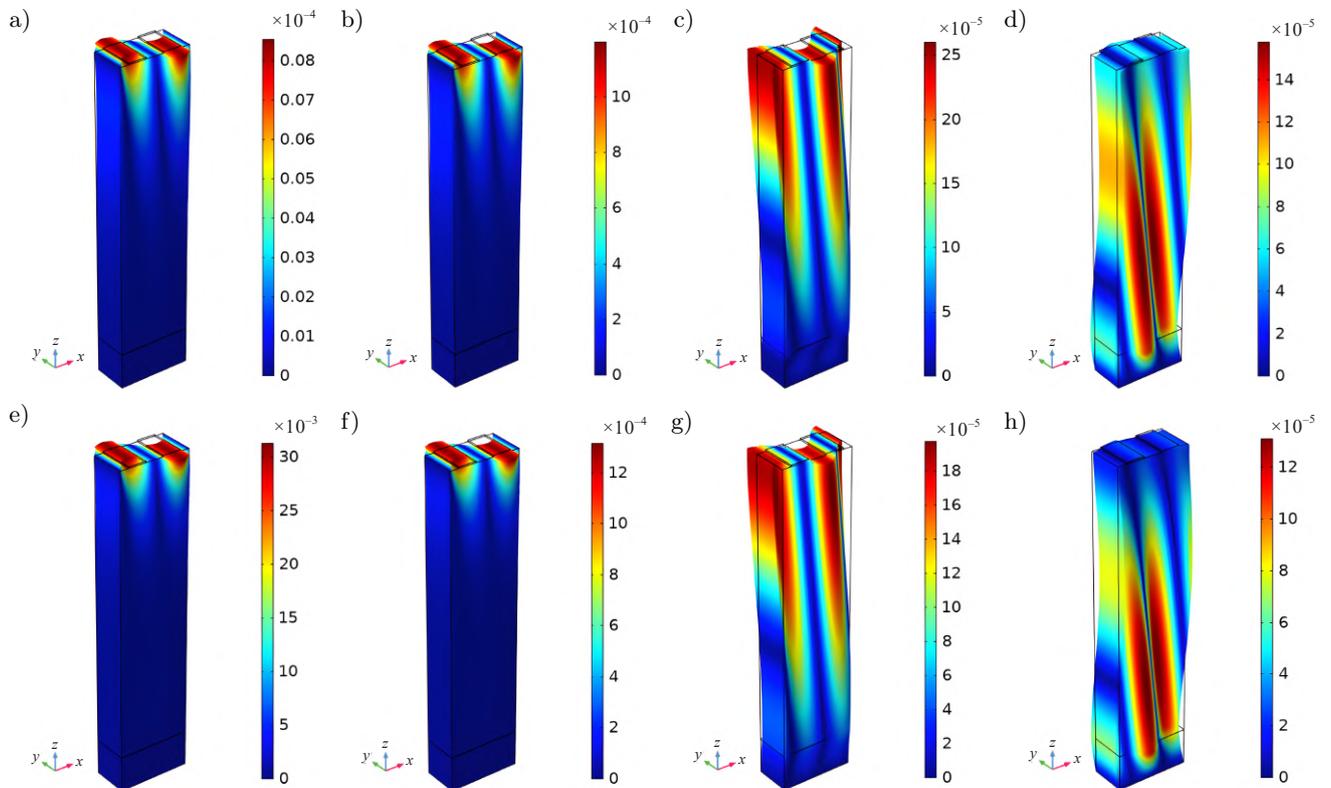


Fig. 10. The 3D mode shapes of the unit cells using 0.020λ -thickness Al IDTs and 0.020λ -thickness Ti IDTs; a)–d) for Al IDTs at 433.7, 434, 439, and 439.3 MHz, respectively; e)–h) for Ti IDTs at 425.6, 426, 439, and 439.4 MHz, respectively; the first two columns present the mode shapes for M1 and the remaining two columns for M2. The measurement unit of the color bars is micrometer [μm].

SSBWs, where smaller transverse displacements are posted. They are related to the M2 mode above. And as the frequency increases, it is found that the wave energy leakage and attenuation from the substrate surface to the bulk become increasingly severe.

4. Conclusions

In this paper, the propagation properties of STW on Y-cut quartz, considering electrode effects were simulated using 3D FEA. The polarized displacement, wave mode, phase velocity, K^2 value, and admittance were examined and analyzed. It is shown that the excitation of STW relates to not only the quartz's crystal cut but also the electrode material and thickness.

For an IDTs/Y-cut quartz structure, STWs can be effectively generated when the crystal cut varies from 34° to 38° . For an IDTs/AT-cut quartz structure, the wave mode, polarized displacement, phase velocity, resonant frequency, K^2 value, and admittance are all strongly influenced by the IDT material and thickness. However, no matter which factor is fixed, the induced effects increase as the mass of the IDTs grows. Hence, in order to alleviate IDTs' mass-loading effect, both the density and thickness should be taken into account. Especially when the IDTs of metal Al with a virtual zero thickness, the maximum phase velocity of 5092 m/s for the STWs is obtained. To increase the thickness of the Al IDTs, the polarized displacement, phase velocity and resonant frequency decrease while the K^2 value increases. Similar phenomena were also observed when Ag, Au, Pt, and Ti were involved. In addition, wave mode conversions (from STW to SSBW) also occur when the IDTs thickness reaches the critical values. When this occurred, the wave energy permeated and attenuated, resulting in a small polarized displacement on the substrate surface. By comparing the used materials, it was found that the mass-loading effect resulting from precious metals are more evident. So, in practical designs and fabrication processes, such behaviors induced by the IDT materials should be noticed. Besides, the acoustic impedance of the materials also have a strong modulation mechanism on the excited waves. The larger the value of acoustic impedance, the stronger the modulation effect on the excited waves. Hence, in order to design a proper STW device, both factors should be carefully considered.

This paper is the first to analyze STWs considering quartz's cut angle and electrode effect through a 3D FE model. It could provide a helpful and easy way to design, optimize, and fabricate the related surface acoustic wave devices. In our future work, experimental verification will be conducted, and an extensive investigation on the modeling of the related devices will be carried out as well.

Acknowledgments

This work was supported by the Science and Technology Research Program of Chongqing Municipal Education Commission under grant no. KJQN202100805, the Chongqing Elite, Innovation and Entrepreneurship Demonstration Team under grant no. CQYC201903246, and the Science and Technology Research Program of Chongqing Municipal Education Commission under grant no. KJZD-M202200801.

References

1. AULD B.A., GAGNEPAIN J.J. (1976), Horizontal shear surface waves on corrugated surfaces, *Electronics Letters*, **12**(24): 650–651, doi: [10.1049/el:19760499](https://doi.org/10.1049/el:19760499).
2. AULD B.A., RENARD A., HENAFF J. (1982), STW resonances on corrugated plates of finite thickness, *Electronics Letters*, **18**(4): 183–184, doi: [10.1049/el:19820126](https://doi.org/10.1049/el:19820126).
3. AULD B.A., THOMPSON D.F. (1987), Temperature compensation of surface transverse waves for stable oscillator applications, [in:] *IEEE 1987 Ultrasonics Symposium*, pp. 305–312, doi: [10.1109/ULTSYM.1987.198974](https://doi.org/10.1109/ULTSYM.1987.198974).
4. AULD B.A., YEH B.-H. (1979), Theory of surface skimming SH wave guidance by a corrugated surface, [in:] *1979 Ultrasonics Symposium*, pp. 786–790, doi: [10.1109/ULTSYM.1979.197312](https://doi.org/10.1109/ULTSYM.1979.197312).
5. AVRAMOV I.D. (2000), High-performance surface transverse wave resonators in the lower GHz frequency range, *International Journal of High Speed Electronics and Systems*, **10**(03): 735–792, doi: [10.1142/S0129156400000635](https://doi.org/10.1142/S0129156400000635).
6. BAGHAI-WADJI A.R., SEIFERT F., ANEMOGIANNIS K. (1988), Rigorous analysis of STWs in nonperiodic arrays including mechanical and electrical interactions, [in:] *IEEE 1988 Ultrasonics Symposium Proceedings*, pp. 303–306, doi: [10.1109/ULTSYM.1988.49388](https://doi.org/10.1109/ULTSYM.1988.49388).
7. BAGWELL T.L., BRAY R.C. (1987), Novel surface transverse wave resonators with low loss and high Q, [in:] *IEEE 1987 Ultrasonics Symposium*, pp. 319–324, doi: [10.1109/ULTSYM.1987.198976](https://doi.org/10.1109/ULTSYM.1987.198976).
8. BIGLER E., AULD B.A., RITZ E., SANG E. (1991), An analysis of the influence of design parameters on the resonant frequency and Q-factor of surface transverse wave (STW) resonators, [in:] *Proceedings of the 45th Annual Symposium on Frequency Control 1991*, pp. 222–229, doi: [10.1109/FREQ.1991.145906](https://doi.org/10.1109/FREQ.1991.145906).
9. DANICKI E.J. (1983), Propagation of transverse surface acoustic waves in rotated Y-cut quartz substrates under heavy periodic metal electrodes, *IEEE Transactions on Sonics and Ultrasonics*, **30**(5): 304–312, doi: [10.1109/T-SU.1983.31425](https://doi.org/10.1109/T-SU.1983.31425).
10. DOBERSTEIN S., VEREMEEV I. (2019), STW resonators with the high quality factor and reduced sizes,

- [in:] *2019 IEEE International Ultrasonics Symposium (IUS)*, pp. 691–694, doi: [10.1109/ULTSYM.2019.8926223](https://doi.org/10.1109/ULTSYM.2019.8926223).
11. FAN Y., JI X. (2018), A novel rotation speed measurement method based on surface acoustic wave, *Acoustical Physics*, **64**(1): 122–128, doi: [10.1134/S1063771018010074](https://doi.org/10.1134/S1063771018010074).
 12. FLORY C.A., BAER R.L. (1987), Surface transverse wave mode analysis and coupling to interdigital transducers, [in:] *IEEE 1987 Ultrasonics Symposium*, pp. 313–318, doi: [10.1109/ULTSYM.1987.198975](https://doi.org/10.1109/ULTSYM.1987.198975).
 13. FU C., LEE K.J., EUN K., CHOA S.-H., LEE K., YANG S.S. (2016), Performance comparison of Rayleigh and STW modes on quartz crystal for strain sensor application, *Journal of Applied Physics*, **120**(2): 024501, doi: [10.1063/1.4955419](https://doi.org/10.1063/1.4955419).
 14. FU Y. *et al.* (2017), Advances in piezoelectric thin films for acoustic biosensors, acoustofluidics and lab-on-chip applications, *Progress in Materials Science*, **89**: 31–91, doi: [10.1016/j.pmatsci.2017.04.006](https://doi.org/10.1016/j.pmatsci.2017.04.006).
 15. GASO ROCHA M.I., Y. JIMEÑEZ, LAURENT F.A., ARNAU A. (2013), Love wave biosensors: A review, [in:] *State of the Art in Biosensors – General Aspects*, Rincken T. [Ed.], IntechOpen, doi: [10.5772/53077](https://doi.org/10.5772/53077).
 16. GAVIGNET E., BALLANDRAS S., BIGLER E. (1995), Theoretical analysis of surface transverse waves propagating on a piezoelectric substrate under shallow groove or thin metal strip gratings, *Journal of Applied Physics*, **77**(12): 6228–6233, doi: [10.1063/1.359590](https://doi.org/10.1063/1.359590).
 17. HASHIMOTO K.-y. (2000), *Surface Acoustic Wave Devices in Telecommunications: Modelling and Simulation*, Springer, New York.
 18. JI X., FAN Y., CHEN J., HAN T., CAI P. (2016), Passive wireless torque sensor based on surface transverse wave, *IEEE Sensors Journal*, **16**(4): 888–894, doi: [10.1109/JSEN.2015.2499318](https://doi.org/10.1109/JSEN.2015.2499318).
 19. JIANG C., CHEN Y., CHO C. (2019), A three-dimensional finite element analysis model for SH-SAW torque sensors, *Sensors*, **19**(19): 4290, doi: [10.3390/s19194290](https://doi.org/10.3390/s19194290).
 20. PARK J., KAYNIA A.M. (2017), FE simulation of steady state wave motion in solids combined with a PML approach, *Procedia Engineering*, **199**: 1556–1561, doi: [10.1016/j.proeng.2017.09.054](https://doi.org/10.1016/j.proeng.2017.09.054).
 21. RANA L., GUPTA R., TOMAR M., GUPTA V. (2018), Highly sensitive Love wave acoustic biosensor for uric acid, *Sensors and Actuators B: Chemical*, **261**, 169–177, doi: [10.1016/j.snb.2018.01.122](https://doi.org/10.1016/j.snb.2018.01.122).
 22. RONNEKLEIV A. (1986), High Q resonators based on surface transverse waves, [in:] *IEEE 1986 Ultrasonics Symposium*, pp. 257–260, Williamsburg, doi: [10.1109/ULTSYM.1986.198748](https://doi.org/10.1109/ULTSYM.1986.198748).
 23. STAHL U. *et al.* (2018), Long-term capability of polymer-coated surface transverse wave sensors for distinguishing vapors of similar hydrocarbons, *Sensors and Actuators B: Chemical*, **274**: 560–564, doi: [10.1016/j.snb.2018.08.013](https://doi.org/10.1016/j.snb.2018.08.013).
 24. STRASHILOV V.L., DJORDJEV K.D., BOYANOV B.I., AVRAMOV I.D. (1997), A coupling-of-modes approach to the analysis of STW devices, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **44**(3): 652–657, doi: [10.1109/58.658322](https://doi.org/10.1109/58.658322).
 25. STRASHILOV V.L., YANTCHEV V.M. (2005), Surface transverse waves: properties, devices, and analysis, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **52**(5): 812–821, doi: [10.1109/TUFFC.2005.1503967](https://doi.org/10.1109/TUFFC.2005.1503967).
 26. THOMPSON D.F., AULD B.A. (1986), Surface transverse wave propagation under metal strip gratings, [in:] *IEEE 1986 Ultrasonics Symposium*, pp. 261–266, doi: [10.1109/ULTSYM.1986.198749](https://doi.org/10.1109/ULTSYM.1986.198749).
 27. YANTCHEV V.M., STRASHILOV V.L., RAPP M., STAHL U., AVRAMOV I.D. (2002), Theoretical and experimental mass-sensitivity analysis of polymer-coated SAW and STW resonators for gas sensing applications, *IEEE Sensors Journal*, **2**(4): 307–313, doi: [10.1109/JSEN.2002.804039](https://doi.org/10.1109/JSEN.2002.804039).
 28. YATSUDA H., KOGAI T. (2006), 3F-3 liquid sensor using SAW and SH-SAW on quartz, [in:] *2006 IEEE Ultrasonics Symposium*, pp. 552–555, doi: [10.1109/ULTSYM.2006.143](https://doi.org/10.1109/ULTSYM.2006.143).
 29. ZHANG J. *et al.* (2021), Real-time monitoring of HL-1 cell viscoelasticity for drug cardiotoxicity assessment using a Love wave biosensor, *Journal of the Electrochemical Society*, **168**(10): 107504, doi: [10.1149/1945-7111/ac29de](https://doi.org/10.1149/1945-7111/ac29de).

Research Paper

PZT Asymmetrical Shape Optimization in Active Vibration Reduction of Triangular Plates

Adam BRAŃSKI^{ORCID}, Romuald KURAS*^{ORCID}

Department of Electrical and Computer Engineering Fundamentals
Rzeszow University of Technology
Rzeszow, Poland

*Corresponding Author e-mail: r.kuras@prz.edu.pl

(received November 18, 2022; accepted March 11, 2023)

The article presents the new 2D asymmetrical PZT (a-PZT) and its effectiveness in the active reduction of triangular plate vibrations. The isosceles right triangular plate with simply supported edges was chosen as the research object. To determine the a-PZT asymmetry and its distribution on the plate, a maximum bending moment criterion for the beam was used. First of all, this criterion points out exact center location of the a-PZT. It was at the point, at which the plate bending moment has reached its maximum value. Next, at this point, it was assumed that the piezoelectric consists of active fibers located radially from the center. Each fiber acted on the plate as a separate actuator. Next, at each direction, the actuator asymmetry was found mathematically by minimizing the amplitude of the vibrations. By connecting the outer edges of individual fibers, the 2D a-PZT was obtained. It was quantitatively confirmed that the effectiveness of the new a-PZT was the best compared with the effectiveness of the standard square and the circular PZTs, adding the same exciting energy to the PZTs.

Keywords: triangular plate; actuator (PZT); active vibration reduction; vibration reduction coefficient; effectiveness coefficient.



Copyright © 2023 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

Active vibration reduction is applied to many common structural elements such as: beams, plates, shells, shafts, trusses. A significant problem is to increase the efficiency of active reduction. Recent publications on this subject have focused on the problem of optimizing the location and orientation of mainly square or circular (regular) PZTs on a structure. ARIDOGAN and BASDOGAN (2015) presented the review of active vibration and noise suppression of plate-like structures with PZT. This article shows numerical methods and experimental tools to study various aspects of controller architecture, i.e., the variety of control systems in active vibration control and their influence on the actuator configuration. ZHANG *et al.* (2018) investigated topology optimization of the electrode coverage of a laminated rectangular plate with piezoelectric patches attached. In this case, the optimization was

carried out in terms of minimizing energy consumption. ZORIĆ *et al.* (2019) presented research on piezo-fiber reinforced composite actuators (PFRC) and the optimization concerning the size, orientation and location of the actuator. GONÇALVES *et al.* (2017) considered active vibration reduction with embedded PZTs. Similarly to ZORIĆ *et al.* (2019), the topology optimization method was used to find the optimum design of actuators; the aim of the optimization was to determine the distribution of piezoelectric material that maximizes controllability for a given mode shape. The shape of the PZT layer was defined using the Sequential Linear Programming algorithm. GARDONIO and CASAGRANDE (2017) published the general guidelines for dimensioning the PZT patch and electrical shunt to maximize the electro-mechanical vibration absorption. Optimization of electrode distribution for two-dimensional structures was also presented by WANG (2003) and the results of the opti-

mization algorithms used confirm the correctness of the maximum bending moment criterion (BRAŃSKI, KURAS, 2022). DONOSO and SIGMUND (2009) considered optimization of the thickness of the PZT layers and its width profile on the beam structure. SUN *et al.* (2022) investigated analytically, numerically and experimentally the problem of active vibration control of a simply supported rectangular plate. The method was based on the plate deflection theory. In the case of homogeneous boundary conditions and regular structure (rectangular plate), this theory coincides with the theory based on the maximum bending moment criterion (BRAŃSKI, KURAS, 2022). TROJANOWSKI and WICIAK (2020) used ring-shaped sensor-actuator hybrid to investigate efficiency of such a system. The compared shapes of actuators are: square, disc, and ring with larger and smaller sensor part. KOZIEŃ and ŚCISŁO (2015) investigated bending vibrations of the beams. The PZT was located at the maximum bending moment of the structure and the parameter for control algorithm was the actual bending moment in the previous time step.

The methods used in the cited articles are most often based on regular PZTs, optimal distribution of piezoelectric material or optimal distribution of electrodes. But the parameter that would influence on the effectiveness of active vibration reduction could be an asymmetric shape of the PZT. In the 1D case (BRAŃSKI, KURAS, 2022), the geometric parameter is optimized which leads to the asymmetry of the PZT. Other parameters that can be also optimized, e.g., the PZT location, the voltage supplied to the transducer. However, for 2D structures, for example triangular, there is one more parameter that follows from the shape of the actuator and its orientation on the structure.

The aim of the paper is to find an asymmetrical shape of the PZT, i.e., a-PZT, and indicate its location on the triangular plate so that it ensures the maximum reduction of vibration. To achieve this aim, the maximum bending moment criterion is applied. Since the problem is more complex than in 1D problem, then the more complex shape and structure of the a-PZT is chosen into consideration, i.e., it consists of piezoceramic fibers. Hence, each fiber can be analyzed separately, leading to separate asymmetries of the individual fibers.

The common feature of all the fibers is that they have one common point (the basic a-PZT point), which is at the point where the plate bending moment reaches its absolute maximum. Each of the fibers can be considered separately as a one-dimensional PZT. The asymmetry of a separate fiber is determined separately and means that the forces on opposite edges have different values and the basic a-PZT point is not in the center of the fiber. Since this point is fixed, it leads to the determining different arms of forces on

opposite edges of separate fibers. It should be added that the values of these forces are determined assuming that the asymmetric moments of both pairs of forces are equal at all fibers. Connecting the outer edges of separate fibers, the 2D a-PZT is obtained. Minimizing amplitude of the plate vibration, forces values of the fibers are determined. The effect of the a-PZT acting, measured by the reduction of vibration amplitude, translates into reduction of the bending moment and shear force.

At the end, the effectiveness of a-PZT and the regular actuators, namely, square PZT (s-PZT) and circular PZT (c-PZT) are compared, assuming that the energy added to all PZTs is the same.

2. Forced vibration of the triangular plate with PZT actuator

The governing equation of transverse vibration of the triangular plate is based on Kirchhoff's classical small deflection theory. In the steady state the equation takes the form (FULLER *et al.*, 1997; HANSEN, SNYDER, 1997):

$$c^2 \Delta^2 w - \omega_f^2 w = \frac{f}{\rho h}, \quad (1)$$

where $f = f_E + f_P$ – the external forces, $f_E = f_E(x, y)$ – the exciting force, $f_P = f_P(x, y)$ – the forces interacting between PZT and the plate, w – the transverse displacement in a steady state, ω_f – the excited frequency, $c^2 = D/(\rho h)$, ρ – the mass density, h – the thickness, $D = \frac{Eh^3}{12(1-\nu^2)}$ – the flexural rigidity, E – Young's modulus, ν – Poisson's ratio, Δ – the Laplace operator, $\Delta w = D_x^2 w + D_y^2 w$.

The boundary conditions for the simply supported right triangular plate are defined as follows:

$$w = 0, \quad M_{nn} = 0, \quad (2)$$

where Eq. (2) relates to all of the edges, Fig. 1, M_{nn} – normal bending moment, n – normal to an edge.

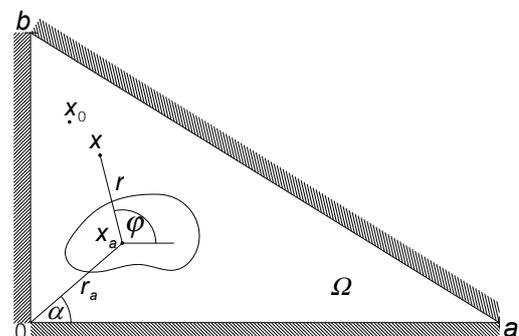


Fig. 1. Simply supported right triangular plate with global symbols: $x_a = (x_a, y_a)$; $x = (x, y) = (r_a + r) = (r_a \cos \alpha + r \cos \varphi, r_a \sin \alpha + r \sin \varphi)$.

An external excitation is in the following form:

$$f_E(x, y) = f_0 \delta(x - x_0, y - y_0), \quad (3)$$

where f_0 – amplitude of the exciting force, $x_0 = (x_0, y_0)$ – point of applying the exciting force.

The active forces of the PZT interaction depend on its shape, and for square and circular shapes they can be found, for example, in (BRAŃSKI, SZELA, 2008; 2010). However, these interaction forces for the a-PZT are derived below.

2.1. Free vibration problem by the superposition method

The superposition method was applied to solve the free vibration problem (GORMAN, 1983; 1999; RAO, 2007; SALIBA, 1990; 1996). It is the solution of the homogeneous Eq. (1), i.e., $\Delta^2 w - \lambda^4 w = 0$, λ – eigenvalue, $\lambda^4 = \omega_f^2 \rho h / D$, and simply supported boundary conditions along three edges (LEISSA, 1969), Eq. (2). First of all, as in standard, non-dimensional coordinates are introduced, i.e., $\xi = x/a$, $\eta = y/b$, and ψ – the plate aspect ratio. Then, the idea of the Lévy solution, based on two building blocks BB1 and BB2, is used, Fig. 2.

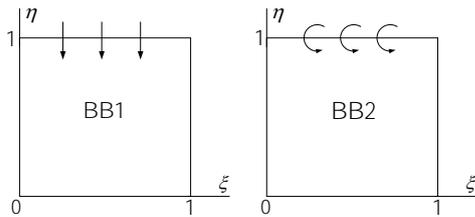


Fig. 2. Building blocks used in the solution.

The solution is assumed to be of the form satisfying SS-BC for each building block. This solution satisfies the SS-BC in advance along the edge $\xi = 0$, $\xi = 1$, and $\eta = 0$. According to the displacement and bending moments along the edge $\eta = 1$, the solution for the BB1 can be formulated as:

$$w_{BB1}(\xi, \eta) = \sum_m^{m^*} G_{1m} \theta_{11m} [a^* + \theta_{1m} \sin(\gamma_m \eta)] \sin(m\pi\xi) + \sum_{m^*+1}^{\infty} G_{1m} \theta_{22m} [a^* + \theta_{2m} \sinh(\gamma_m \eta)] \sin(m\pi\xi), \quad (4)$$

where

$$a^* = \sinh(\beta_m \eta),$$

and G_{1m} is the Fourier coefficient, $m = 1, 2, \dots, m^*$. The first sum is related to condition $\lambda^2 > (m\pi)^2$, the second sum ought to be used if $\lambda^2 < (m\pi)^2$, and

$$\theta_{1m} = \frac{[\beta_m^2 - \nu\psi^2(m\pi)^2] \sinh(\beta_m)}{[\gamma_m^2 + \nu\psi^2(m\pi)^2] \sin(\gamma_m)}, \quad (5)$$

$$\theta_{2m} = -\frac{[\beta_m^2 - \nu\psi^2(m\pi)^2] \sinh(\beta_m)}{[\gamma_m^2 - \nu\psi^2(m\pi)^2] \sinh(\gamma_m)},$$

$$\theta_{11m} = \frac{1}{\sinh(\beta_m) + \theta_{1m} \sin(\gamma_m)}, \quad (6)$$

$$\theta_{22m} = \frac{1}{\sinh(\beta_m) + \theta_{1m} \sin(\gamma_m)},$$

$$\beta_m = \psi \sqrt{\lambda^2 + (m\pi)^2}, \quad (7)$$

$$\gamma_m = \psi \sqrt{\lambda^2 - (m\pi)^2} \quad \text{or} \quad \gamma_m = \psi \sqrt{(m\pi)^2 - \lambda^2}.$$

The solution for the BB2 is obtained from Eq. (4) by replacing G_{1m} by G_{2m} and where

$$\theta_{1m} = -\frac{\sinh(\beta_m)}{\sin(\gamma_m)}, \quad (8)$$

$$\theta_{11m} = \frac{1}{\theta_{1m} \gamma_m^2 \sin(\gamma_m) - \beta_m^2 \sinh(\beta_m)},$$

$$\theta_{2m} = -\frac{\sinh(\beta_m)}{\sinh(\gamma_m)}, \quad (9)$$

$$\theta_{22m} = \frac{1}{\theta_{2m} \gamma_m^2 \sinh(\gamma_m) - \beta_m^2 \sinh(\beta_m)}.$$

The last step is to enforce the simply supported boundary condition along the hypotenuse. For this purpose, the BB1 and BB2 are overlapped. Hence, the SS-BC on two perpendicular sides of the triangle are satisfied. However, the boundary condition along the hypotenuse is enforced by adjusting the Fourier coefficients in $w_{BB1}(\xi, \eta)$ and $w_{BB2}(\xi, \eta)$, i.e., G_{1m} and G_{2m} . Thus along the hypotenuse, η may be written as $\eta = 1 - \xi$ and $w_{BB1}(\xi, 1 - \xi)$, $w_{BB2}(\xi, 1 - \xi)$ are obtained. Formulating the contributions of the BB1 and BB2 and adjusting the G_{1m} and G_{2m} to satisfy the bending moment along the hypotenuse, the formal solution of the problem is obtained. To solve the triangular plate free vibration problem, all boundary conditions should be simultaneously satisfied. The method leads to the system of homogeneous algebraic equations. The determinant of the coefficient matrix is equal to zero and eigenvalues λ^2 are calculated. It leads to the solution of the triangular plate free vibration problem SS-BC in the form:

$$w(\xi, \eta) = w_{BB1}(\xi, \eta) + w_{BB2}(\xi, \eta). \quad (10)$$

To verify the numerical code, the eigenvalues λ^2 for the first five mode shapes are calculated, and they are the same as in (SALIBA, 1990). To obtain the plate deflection in the (x, y) coordinates, in the above formulas the (ξ, η) dimensionless coordinates have to be transformed again.

2.2. Determination of the a-PZT and its action on the plate

Standard (regular) piezoelectric actuators have square, rectangular and circular shapes. In a 2D analysis of the actuator acting on a plate, it is assumed that the actuator induces a bending moment along its

edge. If the piezoelectric material has the same properties in two directions, the induced bending moments are the same at each end of the cross-section of the actuator (FULLER *et al.*, 1997; HER, CHEN, 2020). More precisely, in the case of a square or rectangular actuator, the bending moments at opposite edges are the same. However, in the case of a circular actuator, the bending moment along the edge is the same. As shown in (BRAŃSKI, SZELA, 2008; 2010), bending moments in one cross-section can be replaced by two the same pairs of forces.

To form the a-PZT, the active fibers are considered, arranged radially at a certain point on the plate. Next, it is necessary to find this point that would connect all fibers, i.e., the point around which the actuator would be created. It turns out that at this point, marked by $x_a = (x_a, y_a)$, the bending moment of the plate reaches the absolute extreme. Then, the maximum bending moment criterion is used to find the x_a (it is a common point for all fibers, and the basic one of a-PZT).

To create the shape of the a-PZT, by analogy to one-dimensional structure, namely, to the beam (BRAŃSKI, KURAS, 2022), the plate bending moments in all directions with $\Delta\varphi_k = \varphi_k - \varphi_{k-1} = 1$ degree interval are calculated (Fig. 3). On separate fibers in both directions from point x_a , the points $x_k = (x_{1k}, y_{1k})$ with the same plate bending moments are found. Then the neighboring points at the ends of the fibers were connected with a line and this line created the edge of the a-PZT.

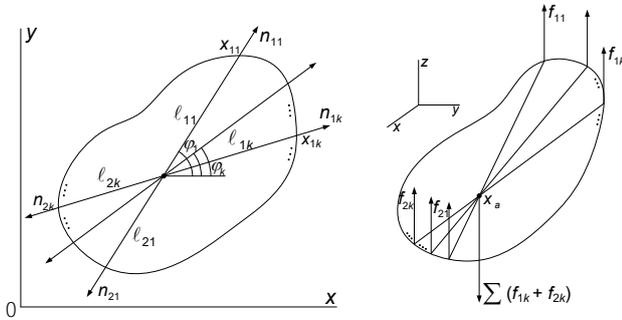


Fig. 3. The idea of constructing the a-PZT and global symbols.

An arbitrary fiber of the a-PZT acts on the plate with the bending moments:

$$2M_k = M_{1k} + M_{2k} = f_{1k}l_{1k} + f_{2k}l_{2k}, \quad (11)$$

where $l_k = \{l_{1k}, l_{2k}\}$ are the arm lengths of the asymmetric fiber.

Bending moments can be replaced by two pairs of forces:

$$\begin{aligned} f_{Pk} &= f_{Pk}(x, y) = f_{1k}\delta(x - x_{1k}, y - y_{1k}) \\ &\quad - (f_{1k} + f_{2k})\delta(x - x_a, y - y_a) \\ &\quad + f_{2k}\delta(x - x_{2k}, y - y_{2k}), \end{aligned} \quad (12)$$

where k – the number of PZT fibres in the actuator, $f_k = \{f_{1k}, f_{2k}\}$ – the forces due to a fibre, $\{x_{1k} = (x_{1k}, y_{1k}), x_a = (x_a, y_a), x_{2k} = (x_{2k}, y_{2k})\}$ – the points of applying forces of a fibre; for simplicity $x_k = (x_{1k}, x_{2k})$, $x = (x, y) = (r_a + r) = (r_a \cos \alpha + r \cos \varphi, r_a \sin \alpha + r \sin \varphi)$ – an arbitrary point of the plate.

Equation (12) defines the acting of one fiber of the a-PZT on a plate. To determine the acting of the a-PZT, it is necessary to sum the acting of all fibers:

$$f_P = f_P(x, y) = \sum_k f_{Pk}(x, y). \quad (13)$$

3. Forced vibration reduction by actuator

As we assumed the response of the plate and the excitation are harmonic, one can express these functions as

$$w_f(x, y) = \sum_n A_n w_n(x, y), \quad (14)$$

$$f(x, y) = \sum_n B_n w_n(x, y), \quad (15)$$

where A_n is the constant to be determined, and

$$B_n = \frac{1}{\zeta_n} \iint_{\Omega} f(x, y) w_n(x, y) dx dy, \quad (16)$$

$$\zeta_n = \iint_{\Omega} w_n^2(x, y) dx dy.$$

The force $f(x, y)$ represents both the exciting force and forces due to PZT, Eq. (1), so the integral in Eq. (16) can be written as $I_n = I_{n;E} + I_{n;P}$, where

$$\begin{aligned} I_{n;E} &= \iint_{\Omega} f_E w_n(x, y) dx dy \\ &= \iint_{\Omega} f_0 \delta(x - x_0, y - y_0) w_n(x, y) dx dy \\ &= f_0 w_n(x_0, y_0). \end{aligned} \quad (17)$$

According to Eq. (12)

$$\begin{aligned} I_{n;P} &= \iint_{\Omega} \sum_k f_{Pk}(x, y) w_n(x, y) dx dy \\ &= \sum_k [f_{1k} w_n(x_{1k}, y_{1k}) - (f_{1k} + f_{2k}) w_n(x_a, y_a) \\ &\quad + f_{2k} w_n(x_{2k}, y_{2k})]. \end{aligned} \quad (18)$$

Substituting Eq. (11) into square brackets and after some calculations, one can obtain

$$\begin{aligned} &\sum_k [f_{1k} w_n(x_{1k}, y_{1k}) - (f_{1k} + f_{2k}) w_n(x_a, y_a) \\ &\quad + f_{2k} w_n(x_{2k}, y_{2k})] \\ &= -\frac{1}{2D} \sum_k l_{1k} (l_{1k} + l_{2k}) M_k(x_a, y_a), \end{aligned} \quad (19)$$

where $M_k(x_a, y_a)$ is the plate bending moment at point (x_a, y_a) .

Equation (19) determines the direct relationship between transverse displacement of the plate caused by the actuator and the plate bending moment at the point (x_a, y_a) . Based on Eqs. (17)–(19), the B_n can be calculated. Substituting Eqs. (14)–(15) into Eq. (1) and taking into account the mode shape functions, Eq. (10), the A_n coefficients can be expressed as follows:

$$A_n = \frac{B_n}{(n^2\pi^2 - \beta_n^2)^2 - \omega_f^2},$$

$$A_n = \frac{B_n}{(n^2\pi^2 + \gamma_n^2)^2 - \omega_f^2} \quad \text{or} \quad A_n = \frac{B_n}{(n^2\pi^2 - \gamma_n^2)^2 - \omega_f^2} \quad (20)$$

whichever is real.

Substituting Eq. (20) into Eq. (14):

$$w_f(x, y) = \sum_n \frac{1/\zeta (I_{n;E} + I_{n;P})}{(n^2\pi^2 - \beta_n^2)^2 - \omega_f^2} w_n(x, y), \quad (21)$$

and considering only the part due to excitation, one can obtain a function that is an optimization objective function:

$$I_{n;f} = f_0 w_n(x_0, y_0) - \frac{1}{2D} \sum_k \ell_{1k} (\ell_{1k} + \ell_{2k}) M_k(x_a, y_a)$$

$$= I_{n;E} + I_{n;P}. \quad (22)$$

Equation (22) was used to model three types of PZT: square, circular, and asymmetrical. The $I_{n;f}$ reduction leads to a reduction of the plate vibrations. Equation (22) means that a total reduction of vibration is possible if the plate response to the exciting force f_E equals the plate response to the forces from all fibers composing the two-dimensional actuator, i.e., if $I_{n;f} = 0$. Thus, the objective function is trivial; the derived expression is minimized and can be written as follows:

$$J = \min (I_{n;f}). \quad (23)$$

4. Reduction effectiveness coefficient

The vibration reduction effectiveness is measured by analysing the vibration amplitude of the plate. The coefficient which is used to calculate the amount of vibration reduction is formulated as follows:

$$R_n = \frac{w_{n;E} - w_{n;f}}{w_{n;E}} \cdot 100\%, \quad (24)$$

where $w_{n;E}$ is the transverse displacement of the plate at the point of maximum amplitude (without PZT); vibrations are forced only by the force $f_E(x, y)$, $w_{n;f}$ – the transverse displacement of the plate at the point of maximum amplitude (with acting PZT); vibrations

are forced by $f_E(x, y)$ and reduced by forces from the PZT fibers.

Similarly to R_n , other coefficients can be formulated, based, for example, on the bending moment or the shear force. The paper presents the results of R_n , because the amplitude value is related to the values of the other two quantities.

5. Analytical calculations

The object of the calculations is an isosceles right triangular plate with simply supported edges. The following data were assumed in the calculations: $a = 1$ m, $b = 1$ m, $h = 1.59 \cdot 10^{-3}$ m, $E = 71.7 \cdot 10^9$ Pa, $\rho = 7169$ kg·m⁻³, $\nu = 0.33$, the lengths of each PZT fiber are the same and equal $0.2a$. The number of fibers was constant for all cases and equal 40. The amplitude of the exciting force was selected experimentally to obtain significant plate deflections for a given mode shape: $\{f_{0;n}\} = \{f_{0;1}, f_{0;2}, f_{0;3}, f_{0;4}, f_{0;5}\} = \{2, 20, 20, 40, 40\}$ N.

5.1. Vibration reduction of triangular plate via PZTs

Based on the assumptions made, the shapes of asymmetrical piezoelectric actuators are determined, and then the effect obtained after their application is compared with the effect of regular PZTs. Taking into account the asymmetric actuator determined by the exact method, its simplified version is presented, in which the asymmetry points of each fiber are arbitrarily set at the point of the maximum bending moment of the plate (Fig. 4). The quantitative results of the calculations are shown in Table 1 and Fig. 5.

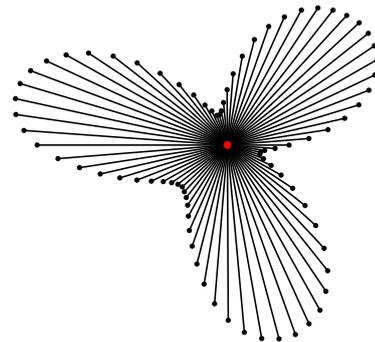


Fig. 4. a-PZT for first mode shape.

Table 1. Vibration reduction coefficient results.

Mode	R_n [%]		
	s-PZT	c-PZT	a-PZT
1	97.43	97.70	99.69
2	97.84	97.82	99.71
3	98.45	99.55	99.89
4	99.93	99.60	99.94
5	97.00	97.98	99.66

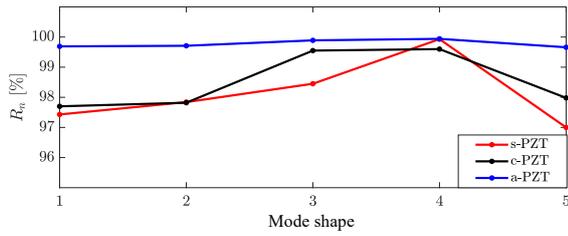


Fig. 5. R_n coefficient for first five mode shapes.

Figure 6 shows the effect of applying s-PZT, c-PZT, and a-PZT. Figures 7a–c corresponds to the first mode shape and they are compared with different shapes of actuators. Figures 7d–f shows the same for the second mode shape. In Fig. 8 there are a-PZTs for higher mode shapes.

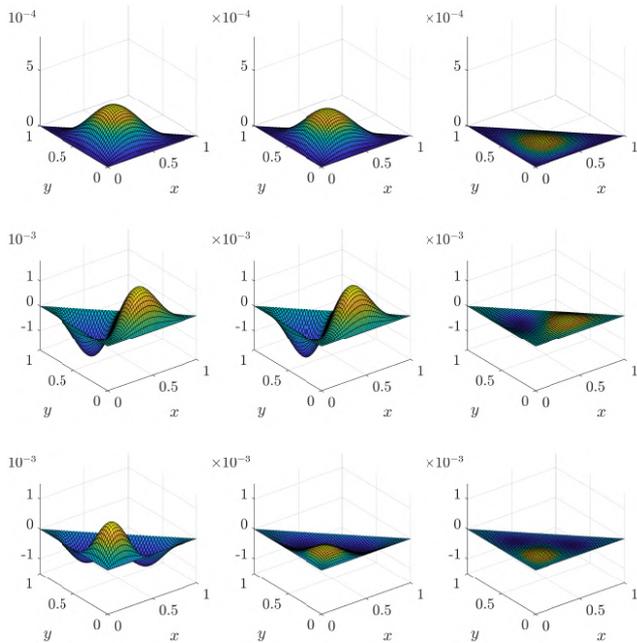


Fig. 6. Results of the PZTs acting for the first three mode shapes: s-PZT (left); c-PZT (center); a-PZT (right).

The results show a big advantage of the a-PZT over the s-PZT. The maximum difference in efficiency occurs for the 5th mode and it is almost 3%. The shape of the a-PZT closely corresponds to the shape of the nodal lines for a given mode. For the 1st, 2nd, and 5th mode shape, in which the nodal lines form triangles similar to the shape of the plate, one gets an irregular shape of the actuator with three vertexes. Moreover, for these mode shapes, a clear axial symmetry can be seen in the shape of the actuator. The axial symmetry can also be seen in the other two vibration modes, 3rd and 4th. In these cases the nodal lines in the area of the a-PZT location are close to a square with rounded vertices (one rounded vertex in the case of 3rd mode, Fig. 8a, two rounded vertices in the case

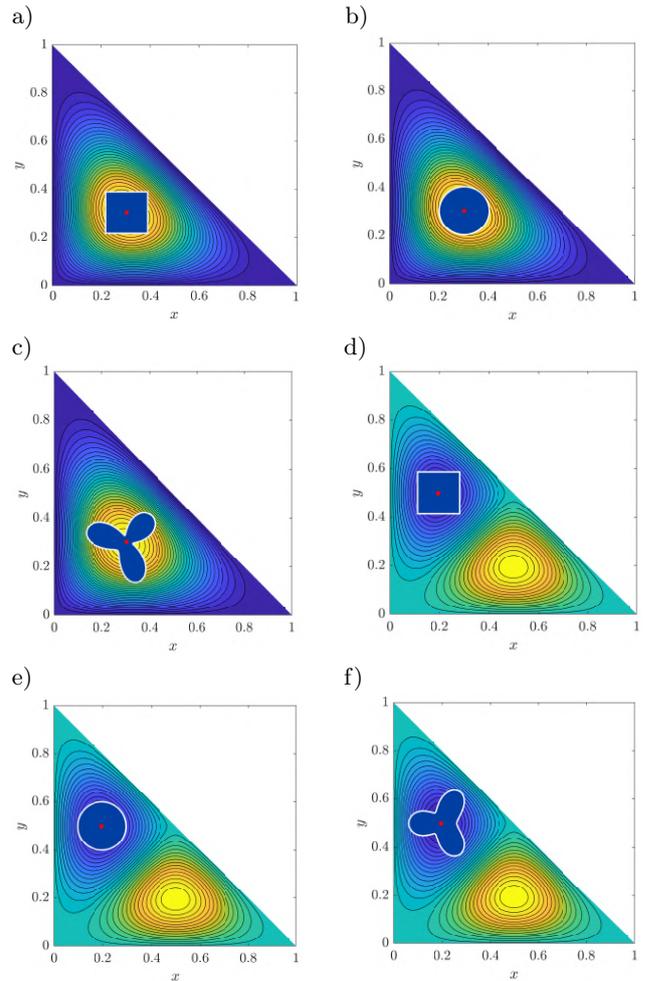


Fig. 7. PZTs shapes and locations for the first two mode shapes – 1st mode shape: a) s-PZT, b) c-PZT, c) a-PZT; 2nd mode shape: d) s-PZT, e) c-PZT, f) a-PZT.

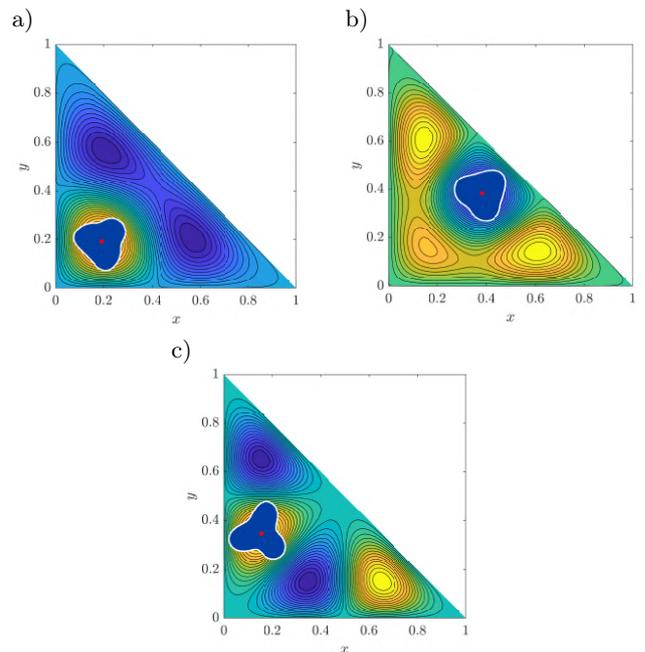


Fig. 8. a-PZTs shapes and locations for other mode shapes: a) 3rd mode shape; b) 4th mode shape; c) 5th mode shape.

of 4th mode, Fig. 8b). In these cases the axial symmetry is due to the plate geometry (isosceles triangle).

The shape of the nodal lines also affects the difference between the efficiency of the c-PZT and the a-PZT. It can be seen that for 1st, 2nd, and 5th mode shapes the difference is about 2% in favour of the a-PZT. However, for 3rd, 4th, and higher mode shapes, the reduction effectiveness coefficient for the c-PZT is very close to the a-PZT. The greater difference in efficiency for these mode shapes can be seen with the reduction of the PZT fiber length.

6. Conclusions

The article presents an active vibration reduction of triangular plate via asymmetrical PZT. Based on the principle of operation of the PZT, the problem of its optimal shape was solved. This led to the creation of the a-PZT. It consists of radial one-dimensional fiber actuators with a common point. This point is located at the maximum plate bending moment. Each actuator fiber was considered separately to optimize its arm lengths. By joining the adjacent ends of the separate fibers, an a-PZT shape was created. The a-PZT provides the most effective reduction of vibrations for a given structure. The results, presented in Table 1, show that the use of the a-PZT in the active reduction of vibrations of two-dimensional structures provides greater efficiency than regular actuators, both square and circular ones. Based on the calculations, the conclusions can be formulated:

- the a-PZT consists of the radial one-dimensional fiber actuators with a common point;
- this is the basic a-PZT point at which the plate bending moment reaches its maximum; it leads to the formulation of the maximum bending moment criterion, to obtain the a-PZT shape;
- the a-PZT reduces vibration more effectively than s-PZT and c-PZT assuming the same energy added to all systems;
- the a-PZT shape closely correlates with the nodal lines of a given mode, i.e., the type of asymmetry depends on the shape of the nodal lines around the maximum plate bending moment.

The idea of the a-PZT presented in the paper can be a starting point for considerations on the active reduction of vibrations of more complex two-dimensional structures, e.g., cylindrical or three-dimensional structures.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of conflicting interests

The Authors declare that there is no conflict of interest.

References

1. ARIDOGAN A., BASDOGAN I. (2015), A review of active vibration and noise suppression of plate-like structures with piezoelectric transducers, *Journal of Intelligent Material Systems and Structures*, **26**(12): 1455–1476, doi: [10.1177/1045389X15585896](https://doi.org/10.1177/1045389X15585896).
2. BRAŃSKI A., SZELA S. (2008), Improvement of effectiveness in active triangular plate vibration reduction, *Archives of Acoustics*, **33**(4): 521–530.
3. BRAŃSKI A., SZELA S. (2010), Quasi-optimal PZT distribution in active vibration reduction of the triangular plate with P-F-F boundary conditions, *Archives of Control Sciences*, **20**(2): 209–226, doi: [10.2478/v10170-010-0014-7](https://doi.org/10.2478/v10170-010-0014-7).
4. BRAŃSKI A., KURAS R. (2022), Asymmetrical PZT applied to active reduction of asymmetrically vibrating beam – semi-analytical solution, *Archives of Acoustics*, **47**(4): 555–564, doi: [10.24425/aoa.2022.142891](https://doi.org/10.24425/aoa.2022.142891).
5. DONOSO A., SIGMUND O. (2009), Optimization of piezoelectric bimorph actuators with active damping for static and dynamic loads, *Structural and Multidisciplinary Optimization*, **38**: 171–183, doi: [10.1007/s00158-008-0273-0](https://doi.org/10.1007/s00158-008-0273-0).
6. FULLER C.R., ELLIOT S.J., NIELSEN P.A. (1997), *Active Control of Vibration*, Academic Press, London.
7. GARDONIO P., CASAGRANDE D. (2017), Shunted piezoelectric patch vibration absorber on two-dimensional thin structures: Tuning considerations, *Journal of Sound and Vibration*, **395**: 26–47, doi: [10.1016/j.jsv.2017.02.019](https://doi.org/10.1016/j.jsv.2017.02.019).
8. GONÇALVES J.F., DE LEON D.M., PERONDI E.A. (2017), Topology optimization of embedded piezoelectric actuators considering control spillover effects, *Journal of Sound and Vibration*, **388**: 20–41, doi: [10.1016/j.jsv.2016.11.001](https://doi.org/10.1016/j.jsv.2016.11.001).
9. GORMAN D.J. (1983), A highly accurate analytical solution for free vibration analysis of simply supported right triangular plates, *Journal of Sound and Vibration*, **89**(1): 107–118, doi: [10.1016/0022-460X\(83\)90914-8](https://doi.org/10.1016/0022-460X(83)90914-8).
10. GORMAN D.J. (1999), *Vibration analysis of Plates by the Superposition Method*, World Scientific Publishing Co. Pte. Ltd., Singapore.
11. HANSEN C.H., SNYDER S.D. (1997), *Active Control of Noise and Vibration*, E & FN SPON, London.
12. HER S.-C., CHEN H.-Y. (2020), Deformation of composite laminates induced by surface bonded and embedded piezoelectric actuators, *Materials*, **13**(14): 3201, doi: [10.3390/ma13143201](https://doi.org/10.3390/ma13143201).

13. KOZIEŃ M.S., ŚCISŁO Ł. (2015), Simulation of control algorithm for active reduction of transversal vibrations of beams by piezoelectric elements based on identification of bending moment, *Acta Physica Polonica A*, **128**(1): A56–A61, doi: [10.12693/APhysPolA.128.A-56](https://doi.org/10.12693/APhysPolA.128.A-56).
14. LEISSA A.W. (1969), *Vibration of Plates*, Scientific and Technical Information Division NASA, Washington.
15. RAO S.S. (2007), *Vibration of Continuous Systems*, John Wiley & Sons Inc., Hoboken, New Jersey.
16. SALIBA H.T. (1990), Transverse free vibration of simply supported right triangular thin plates: A highly accurate simplified solution, *Journal of Sound and Vibration*, **139**(2): 289–297, doi: [10.1016/0022-460X\(90\)90889-8](https://doi.org/10.1016/0022-460X(90)90889-8).
17. SALIBA H.T. (1996), Free vibration of simply supported general triangular thin plates: An accurate simplified solution, *Journal of Sound and Vibration*, **196**(1): 45–57, doi: [10.1006/jsvi.1996.0466](https://doi.org/10.1006/jsvi.1996.0466).
18. SUN Y., SONG Z., LI F. (2022), Theoretical and experimental studies of an effective active vibration control method based on the deflection shape theory and optimal algorithm, *Mechanical Systems and Signal Processing*, **170**: 108650, doi: [10.1016/j.ymssp.2021.108650](https://doi.org/10.1016/j.ymssp.2021.108650).
19. TROJANOWSKI R., WICIAK J. (2020), Impact of the size of the sensor part on sensor-actuator efficiency, *Journal of Theoretical and Applied Mechanics*, **58**(2): 391–401, doi: [10.15632/jtam-pl/118948](https://doi.org/10.15632/jtam-pl/118948).
20. WANG W. (2003), *Electrode shape optimization of piezoelectric transducers*, Ph.D. Thesis, University of Florida.
21. ZHANG X., TAKEZAWA A., KANG Z. (2018), Topology optimization of piezoelectric smart structures for minimum energy consumption under active control, *Structural and Multidisciplinary Optimization*, **58**: 185–199, doi: [10.1007/s00158-017-1886-y](https://doi.org/10.1007/s00158-017-1886-y).
22. ZORIĆ N.D., TOMOVIĆ A.M., OBRADOVIĆ A.M., RADULOVIĆ R.D., PETROVIĆ G.R. (2019), Active vibration control of smart composite plates using optimized self-tuning fuzzy logic controller with optimization of placement, sizing and orientation of PFRC actuators, *Journal of Sound and Vibration*, **456**: 173–198, doi: [10.1016/j.jsv.2019.05.035](https://doi.org/10.1016/j.jsv.2019.05.035).

Research Paper

A Microscopic Prediction Model for Traffic Noise in Adjacent Regions to Arterial Roads

Ming LI⁽¹⁾, Jizhou LIU⁽²⁾*

⁽¹⁾ *School of Transportation Engineering, Shandong Jianzhu University
Jinan, China*

⁽²⁾ *School of Thermal Engineering, Shandong Jianzhu University
Jinan, China*

*Corresponding Author e-mail: liujizhou20@sdjzu.edu.cn

(received October 7, 2022; accepted December 30, 2022)

Traffic noise in big cities impacts the people who live and work in high-rise buildings alongside arterial roads. To determine this impact magnitude, this paper proposes and validates a microscopic level method that locally predicts the total noise level and the spectral characteristics of traffic flow in the near-road region. In the proposed method, the vehicles on the road are considered as multiple queues of moving point sound sources with ground reflection considered. To account for the flow of vehicles on the road, traffic field data, and individual vehicle noise source models are also employed. A field measurement is conducted to validate the proposed method. Results comparison shows that the predicted and the measured overall A-weighted sound pressure level and A-weighted noise spectra are within 3 dBA and 5 dBA, respectively. Based on the validated method, the spatial distribution of traffic noise near the arterial road is investigated for different traffic scenarios.

Keywords: traffic noise; arterial roads; noise prediction method; microscopic level.



Copyright © 2023 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

1.1. Background and motivations

In many big cities, arterial roads, and overhead freeways have been constructed in the urban areas or between the urban and suburban areas to meet the cities' transportation demands. Due to the high traffic speed and volume, the noise generated by traffic flows on such roads is much more important than that generated by traffic on secondary streets. In many densely populated cities, it is common for high-rise buildings, whether residential, educational, business or commercial, to be built just by the overhead freeways or no more than several hundred meters away from the arterial roads. People who live or work in those buildings experience daily traffic noise annoyance, which can be harmful to their health (LOKHANDE *et al.*, 2018; BARRIGÓN MORILLAS *et al.*, 2022; SANOK *et al.*, 2022).

To evaluate the traffic noise impact on existing or planned infrastructures, a series of traffic noise prediction methodologies have been established. Depending on the scope and degree of detail of the researched problem, these methods could be divided into three categories. The first methodology is the deterministic model. It relies on single vehicle noise emission model as well as traffic speed and volume data (NIELSEN *et al.*, 1996; BENDTSEN, 1999; KEPHALOPOULOS *et al.*, 2012; Directive EN, 2015; LIN *et al.*, 2012; PENG *et al.*, 2019). Although the prediction range is limited, detailed information, such as the third-octave spectra and higher accuracy, is ensured. The second modeling category, based on the noise map (HINTON *et al.*, 2005; POPP, 2003), extends the road traffic noise prediction to city regions (LEE *et al.*, 2008). By noise maps, the traffic noise exposure of the population could be estimated (KADDOURA *et al.*, 2017; LAN *et al.*, 2020; LOKHANDE *et al.*, 2017). However, for regional or even

national scale noise exposure estimation, incomplete data makes the noise map partially efficient. For such a scale, the statistical model (STAAB *et al.*, 2022) based on available data and linear land-use regression could be employed.

In view of the traffic noise issue alongside arterial roads in big cities, this paper targets to develop a traffic noise prediction method at a microscopic level, which is applicable to evaluating locally the impact of traffic noise from a segment of the arterial road on people living or working in nearby high rise buildings, and provides detailed spectral features and noise level spatial distributions. Under such circumstances, the width of the noise source could vary from 25 m to over 60 m depending on the road width, while the distance of the observer in nearby buildings from the road is mostly no more than 500 m, making the dimension of the noise source non-negligible as compared to the range of sound propagation.

The method proposed by this paper aims at predicting locally the traffic noise near a segment of big roads with the multi-lane configuration of the arterial road considered to avoid the linear source assumption (STEELE, 2001; QUARTIERI *et al.*, 2009). On each lane, the vehicles are modeled as a queue of moving point sound sources in half space. The time-dependent sound pressure field of each vehicle is derived by acoustic theory considering the translational movement of the vehicle and the reflection of the ground. At the receiver point, the total sound pressure at a certain moment is considered as the superposition of the sound pressure fields of all the vehicles on the road. The noise intensity of each vehicle as a point source follows the noise emission expression in (LIN *et al.*, 2012) and the noise spectral characteristics for different vehicle types follow the work of YANG *et al.* (2020). The sound attenuation effect of trees by the road is also considered. Therefore, the valid range of the method is in the vicinity of arterial roads where the road width makes simple line source assumption less suitable. The $1/3$ octave sound pressure level spectra could be predicted with different lane configurations, vehicle speeds, and vehicle types.

1.2. Related work

In predicting traffic noise alongside roads, three basic elements are needed: 1) the noise source characteristics: the noise level and spectral characteristics of a single vehicle depending on vehicle speed and type; 2) traffic flow data: including traffic speed data and traffic volume data; 3) the propagation module: the noise emission model that estimates the noise intensity at a certain distance away from the vehicle. Many researchers worked on these elements and have laid a well-established foundation for reference.

To obtain the noise source characteristics of a single vehicle, field tests have been conducted with different

vehicle types, speeds, accelerations, and road surface status. By collecting noise measurement results near the roads in the Nordic countries, the Nordic Prediction method (NIELSEN *et al.*, 1996; BENDTSEN, 1999) established the noise emission curves as a function of vehicle speed for light and heavy vehicles. LIN *et al.* (2012) measured the noise data of light, medium, and heavy vehicles at a traffic intersection and established the relationship of the sound pressure level of a single vehicle as a function of speed and acceleration at a reference location of 7.5 m from the first lane. In the report of Common Noise Assessment Methods in Europe (CNOSSOS-EU) (KEPHALOPOULOS *et al.*, 2012; Directive EN, 2015), the influence of frequency and noise source type was introduced to the source emission module. For instance, the vehicle noise is decomposed into rolling noise and propulsion noise. Their dependencies on vehicle speed and frequency were provided, respectively. To obtain detailed spectral noise characteristics, LUO *et al.* (2013) and YANG *et al.* (2020) collected and classified the $1/3$ octave noise spectral energy contribution of single vehicles based on vehicle type and speed, which allows the prediction of the traffic noise spectrum.

Another aspect of traffic noise prediction is to relate the noise emitted by the traffic flow with the receiver at a certain point. A variety of models have been established by researchers from different countries. For example, see the FHWA (BARRY, REAGAN, 1978), the CoRTN (CoRTN, 1975), the RLS 90 (RLS, 1990), the Nord2000 (NIELSEN *et al.*, 1996; BENDTSEN, 1999), the NMPB-Route-2008 (DUTILLEUX *et al.*, 2010), and the CNOSSOS-EU (KEPHALOPOULOS *et al.*, 2012; Directive EN, 2015; KHAN *et al.*, 2021). In these models, the propagation module adopts the energy type equation or the ray tracing theory to calculate the sound attenuation caused by geometrical divergence during the propagation. Influences of the absorption of air, the effect of obstacles, the reflection of the ground, etc., are also implemented in the models. By using these emission models, the equivalent sound pressure level of traffic flow could be estimated. For instance, STOILOVA and STOILOV (1998) applied this kind of model to study noise pollution control by traffic lights. In (YANG *et al.*, 2020) the noise spectrum of the traffic flow was calculated based on a localized noise emission model.

2. Methodology

This section describes the formulation of the proposed method from a single vehicle to the traffic flow on the road.

2.1. Acoustic field of a single vehicle on the road

The basic hypothesis in the current formulation is that the vehicle could be regarded as a point sound

source or acoustic monopole. This means that the directivity of the vehicle noise is dropped in the first place. However, unlike aircraft, the vehicle does not operate in free space but on the road and moves along a certain direction with variable speed. Therefore, the reflection effect of the ground and the moving effect of the vehicle on sound propagation should not be neglected.

With the above considerations, the vehicle on the road is modeled as a moving point sound source in half-space. Figure 1a sketches the geometric configuration of the observation point (X_o, Y_o, Z_o, t) and the i -th vehicle (in red) on the road (X_i, Y_i, Z_i, t) at time t . The formulation starts from the three-dimensional sound pressure field expression of a static point source at frequency f . At time t the acoustic pressure at the observation point is (SMITH III, 2010)

$$p_s(X_o, Y_o, Z_o, t) = \frac{p_a}{r_{io}} e^{i(2\pi ft - kr_{io})}, \quad (1)$$

where p_a represents the acoustic pressure amplitude of the point sound source, f and k are the frequency and the wavenumber of the sound wave, and r_{io} is the distance between the sound source and the observation point. In Eq. (1), the pressure variable whose amplitude is proportional to $1/r$ with propagation distance (SMITH III, 2010) is used instead of the acoustic intensity, which follows the inverse square law in order to consider the phase difference when superimposing the contribution of all the vehicles on the road in Eq. (3).

The moving effect means that the measurement of acoustic pressure at the observation point at the moment t does not come directly from the vehicle's current position at t (red vehicle in Fig. 1a) but from a certain moment prior to t at the position of $(X_i - V_i\tau, Y_i, Z_i, t - \tau)$ (upward pink vehicle in Fig. 1a). This difference in time is called the retarded time τ or delay. To estimate it, the simplified sketch in Fig. 1b should be referred to. The first step is to calculate the

orientation angle θ . According to the dot product of vector $\overrightarrow{A'A}$ and \overrightarrow{AO} , the orientation angle is

$$\theta = \arccos \left(\frac{\overrightarrow{A'A} \cdot \overrightarrow{AO}}{|\overrightarrow{A'A}| |\overrightarrow{AO}|} \right), \quad (2)$$

where \overrightarrow{AO} is the vector from the vehicle's current position to the observation point at time t . $\overrightarrow{A'A}$ is the distance traveled by the i -th vehicle during the retarded time. In Eq. (2), since the retarded time τ is unknown, the direction of the vector $A'A$ is sufficient for the calculation.

After obtaining the orientation angle θ , it is possible to derive the retarded time τ . During the retarded time, the vehicle has traveled a distance of $|\overrightarrow{A'A}| = V_i\tau$ if its speed was assumed constant during that period. Let us take the triangle $\Delta OAA'$, the cosine theorem says:

$$(c_0\tau)^2 = (V_i\tau)^2 + |\overrightarrow{OA}|^2 - 2(V_i\tau)|\overrightarrow{OA}|\cos(\theta), \quad (3)$$

where c_0 is the speed of sound and V_i is the speed of the i -th vehicle. By solving this second degree equation, the retarded time τ is obtained so that the exact position of the i -th vehicle at time $t - \tau$ is derived as $A' = (X_i - V_i\tau, Y_i, Z_i, t - \tau)$.

The ground effect means that the sound traveling downwards z -direction is reflected by the road surface. This effect is treated by adding an image point source beneath the road (HUDSON, 2008; MCLAUGHLIN *et al.*, 2008), as shown in Fig. 1b by an upside-down pink vehicle. Mathematically, the position of the image point source is $A'' = (X_i - V_i\tau, Y_i, -Z_i, t - \tau)$. The strength of the image source varies with different types of road surface due to different sound reflection and absorption coefficients.

With the moving point sound source and the ground effect, the acoustic pressure of frequency f at

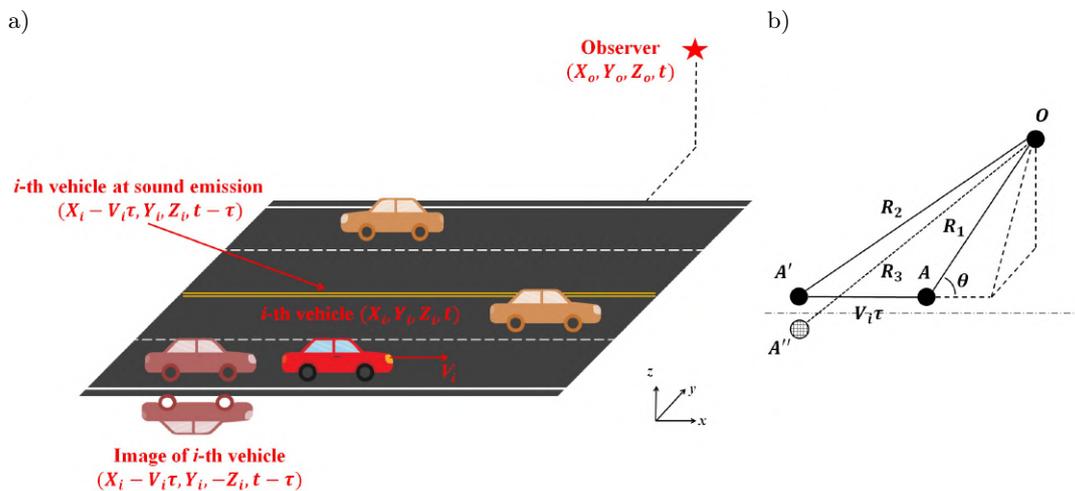


Fig. 1. Schematic representation of the formulation: a) geometry configuration; b) simplified sketch.

the observation point at time t caused by the i -th vehicle is the sum of the two point sources located at $A' = (X_i - V_i\tau, Y_i, Z_i, t - \tau)$ and $A'' = (X_i - V_i\tau, Y_i, -Z_i, t - \tau)$.

$$p_o^i(X_o Y_o Z_o t) = \frac{p_a^i}{R_2} e^{i(2\pi f(t-\tau) - kR_2)} + \frac{C_g(f)p_a^i}{R_3} e^{i(2\pi f(t-\tau) - kR_3)}, \quad (4)$$

where $R_2 = |OA'|$ is the distance between the i -th vehicle at the moment of sound emission $t - \tau$ and the observer point, and $R_3 = |OA''|$ is the distance between the image of the i -th vehicle and the observer point. The coefficient $C_g(f)$ stands for the frequency-dependent reflection coefficient of the ground surface. In the current paper, it is set according to the measured absorption coefficients of asphalt pavements in the work of KNABBEN *et al.* (2016) and LI *et al.* (2014). Detailed values of the reflection coefficient are listed in Appendix A.

2.2. Vehicle noise source characteristics

Subsection 2.1 derived a formulation that could calculate the acoustic pressure field of a single moving vehicle on the road. However, in Eq. (4), the acoustic pressure amplitude p_a is dependent on the strength of the sound source. The dependence of this variable on vehicle type, speed and frequency requires the support of field data. In this paper, we are not engaged in collecting vehicle noise source characteristics but in using the source model of LIN *et al.* (2012) and the single vehicle noise spectral characteristics in the work of YANG *et al.* (2020). A brief introduction and the use of the employed models are presented herein for completeness.

According to LIN *et al.* (2012), the sound pressure level of a vehicle at the reference location 7.5 m away from the first lane and 1.2 m in height is dependent on its speed and vehicle type by:

$$\begin{aligned} \text{light vehicle:} \quad & L = 27.96 + 24.91 \log 10(V), \\ \text{medium vehicle:} \quad & L = 28.36 + 29.73 \log 10(V), \quad (5) \\ \text{heavy vehicle:} \quad & L = 31.77 + 29.70 \log 10(V), \end{aligned}$$

where V is the speed of the vehicle in km/h. LIN *et al.* (2012) expression establishes the relationship between vehicle noise pressure level and vehicle speed based on vehicle type. It means that the acoustic pressure amplitude could be written as $p_a^i(\text{Type}_i, V_i)$. However, Lin's expression (LIN *et al.*, 2012) is only valid for the overall sound pressure level that does not contain spectral energy distribution.

In (YANG *et al.*, 2020) detailed noise spectral energy distribution at $1/3$ octave frequencies from 12.5 Hz to 20 kHz is provided for four types of vehicles: light vehicle, medium vehicle, heavy vehicle, and bus. Since

the vehicle speed also influences the spectral energy distribution, YANG *et al.* (2020) have also provided the noise spectral energy distribution at different velocity intervals. By collecting the spectral information the authors of this paper have established a database for four vehicle types and five vehicle speed intervals at 28 $1/3$ octave frequency bands. Detailed spectral distributions can be found in Appendix B.

To combine the model of LIN *et al.* (2012) and the data of YANG *et al.* (2020), the acoustic pressure amplitude p_a at frequency f_k for a certain type of vehicle at velocity V_i is the multiplication of the sound pressure (converted from sound pressure level) from Eq. (5) and the spectral energy distribution at frequency f_k . Finally, the acoustic pressure amplitude in Eq. (4) is dependent on vehicle type, speed and frequency $p_a^i(\text{Type}_i, V_i, f_k)$. It should be pointed out that although the field data in the literature was acquired by experimental measurements, the absolute amplitude of $p_a^i(\text{Type}_i, V_i, f_k)$ still needs to be calibrated for realistic application. This is because the exact values of the first terms in Eq. (5) (Lin's expression (LIN *et al.*, 2012)) depend on the reference location of the measurement.

2.3. Prediction of traffic flow noise on the road

With the formulation of the acoustic field of a single vehicle and the vehicle noise feature database, it is possible to evaluate the total noise spectrum of the whole traffic flow on the road. At a given observation point position and time (X_o, Y_o, Z_o, t) , the total acoustic pressure p_{total} at one of the $1/3$ octave frequency bands f_k is the sum of the contribution of each vehicle on the road:

$$p_{\text{total}}(X_o, Y_o, Z_o, t, f_k) = \sum_i \left(\frac{p_a^i(\text{Type}_i, V_i, f_k)}{R_2} e^{i(2\pi f_k(t-\tau) - kR_2)} + \frac{C_g p_a^i(\text{Type}_i, V_i, f_k)}{R_3} e^{i(2\pi f_k(t-\tau) - kR_3)} \right), \quad (6)$$

where i represents the i -th vehicle on the road, and k is the k -th $1/3$ octave frequency band. To obtain the noise spectrum at the observation point, the total acoustic pressure p_{total} at every $1/3$ octave frequency band should be calculated following Eq. (6).

To derive the sound pressure level spectra, the sound pressure level at each frequency band $\text{SPL}(f_k)$ is calculated first with the A-weighting factor. Then, considering that vegetation by the side of the arterial roads could attenuate sound that passes, sound attenuation through trees should be accounted for when calculating the sound pressure level spectra in the far field (VAN RENTERGHEM *et al.*, 2012). Figure 2 shows a schematic configuration when calculating the sound

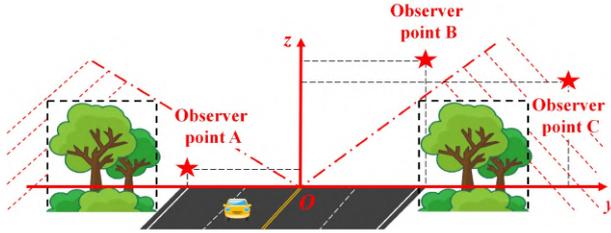


Fig. 2. Schematic configuration for sound attenuation calculation by vegetation.

attenuation by vegetation on a certain cross-section plane of the road. The origin of the cross-section plane is placed on the road’s centerline, where the noise source location is assumed. For the observer points in the dashed line region, the attenuation of the trees is accounted for because the trees would block the sound propagating directly from the source to the observers, such as the far-field observer point C in the figure. However, for far-field observer point B and near-field observer point A, the sound attenuation by the vegetation needs not to be counted. Detailed sound attenuation values employ those in (PRICE *et al.*, 1988), as listed in Appendix C.

2.4. Comments on the proposed traffic noise prediction method

The proposed traffic noise prediction method is based on the acoustic pressure field of moving point sources in half-space and the field data of vehicle noise from the literature, as explained in Subsecs. 2.1 and 2.2. These two elements are independent of the traffic flow status because they deal only with one vehicle. However, the traffic flow on arterial roads contains hundreds to thousands of vehicles. The summation in Eq. (6) is mathematically rigid but not easy to be applied in real circumstances since it demands the time-dependent position and velocity of every vehicle on the road.

In real circumstances, traffic noise from arterial roads is more important at rush hours during the day because of the very high traffic volume. During rush hours, the traffic volume is not only high, but also steady and uniform. This means that all the vehicles on the road are driven at a relatively high and steady speed, and the gaps between vehicles are nearly uniform. Under this condition, the lane-wise time-averaged vehicle speed and volume could be a good approximation of the realistic traffic scenarios. As it could represent the main feature of the traffic flow on the road and it does not demand every detail of the traffic flow. Therefore, to apply the proposed method to realistic traffic noise prediction problems, we could first obtain lane-wise traffic data by field measurements during rush hours. Then establish numerically an equivalent traffic flow calculated from the lane-wise measurement data. With the equivalent traffic flow es-

tablished, the position and the velocity of each vehicle in the traffic flow could be derived so that the formulation in Subsecs. 2.1, 2.2, and 2.3 could be applied to calculate the noise spectra from the whole segment of the road at a certain observation point.

3. Calibration and validation

In order to validate the proposed method, a field measurement is conducted, obtaining four groups of traffic flow and roadside noise data. The logic behind is to use one group of the data to calibrate the absolute value of the acoustic pressure amplitude $p_a^i(\text{Type}_i, V_i, f_k)$ in Eq. (6) and use the other three groups for validation.

3.1. Test site

The test site is chosen to be a segment of Jingshi Road, Jinan, China. It is a two-way 14-lane arterial road that is the most important road in Jinan city. Near Jingshi Road, three residential districts, business buildings, hotels, and hospitals are located. During rush hours, due to the very high traffic flow volume, the traffic noise is very explicit and impacts the people who live and work in the buildings by the road.

As shown in Fig. 3a, the length of the tested road segment is about 1.2 km. The observation points are

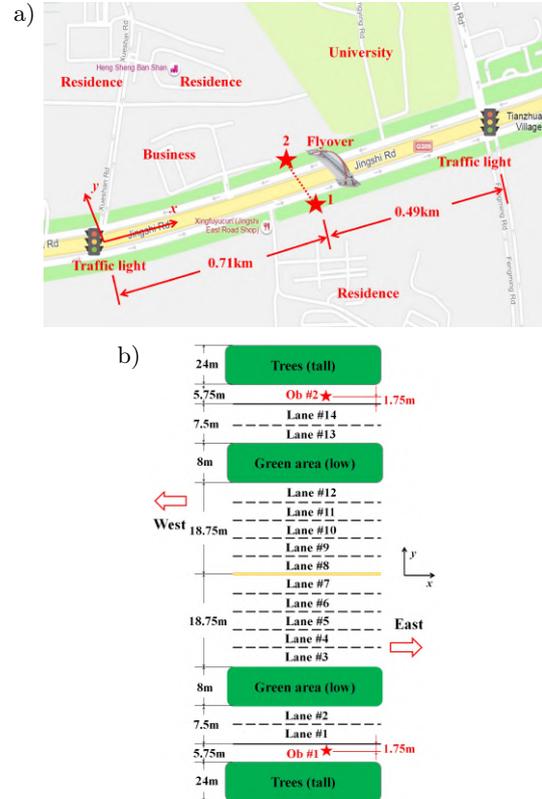


Fig. 3. Spatial configuration of the test site: a) overall geometry of the tested road segment; b) lane definition of the tested road segment.

chosen to be near the middle of the tested segment to make sure that the vehicles' speeds passing the observation points reached a quasi-steady state. A flyover bridge is located near the observation points where the video of the traffic flow is taken. The road has 14 lanes and 2 green areas. The width of each lane is 3.75 m and that of the green area is approximately 8 m. On both sides of the road, there are two belts of barrier trees. These barrier trees, consisting of a mixture of poplar (*Populus simonii*), cypress, and Chinese holly (*Ilex cornuta*), are 40 m away from the road's centerline and have a width of 24 m and a height of about 20 m. The total road width is 68.5 m including the automotive lanes only, and 128 m including the pedestrian lanes and barrier trees. The road direction is chosen to be along the x -direction in the formulation. The origin is at the west end of the road segment. The road width direction is the y -direction, originating at the road's centerline. The lane is numbered from south to north, as shown in Fig. 3b.

The test was conducted on May 30th, 2021, during rush hour from 5:30 pm to 6 pm (sunny, local temperature 25° (Hourly Historical Weather Data [OL])). Four groups of traffic flow and roadside noise data were collected. For group #1 and group #2, the noise measurement was conducted at point 1 (710 m, -36 m, 1.6 m), which is 1.75 m to the southern edge of lane #1 and at the height of 1.6 m, as shown in Fig. 3a. For group #3 and group #4, the observation location was on the other side of the road at point 2 (710 m, 36 m, 1.6 m), as shown in Fig. 3a. The sound pressure levels at $1/3$ octave frequencies were recorded at the observation points.

3.2. Lane-wise average traffic flow volume and speed

The traffic volume on the tested road segment is wavy in time due to the traffic lights being located at the two ends (Fig. 3a). In reality, the traffic noise is crucial only when the traffic volume is high. Therefore, we choose to record the noise and flow data only when the traffic flow volume is high and steady. This corresponds to a duration of about 40 s for each group of measurement where the green lights are on (the green cycle is 120 s) and the vehicles passing the observation points reach relatively high and steady speeds. By analyzing the videos of the traffic flow, the traffic volume and average speed on each of the 14 lanes could be derived. Detailed traffic flow data are listed in Table 1.

In the measured traffic flow data, group #1 and group #2, focusing on the traffic from west to east, have comparable traffic volumes of 12690 veh/h and 13050 veh/h. Group #3 and group #4, focusing on the other direction, have comparable traffic volumes of 14130 veh/h and 14490 veh/h. The aforementioned traffic volumes are the total volume of both directions, and the difference is due to the measurement time, where the first two groups were measured between 5:30 pm to 5:45 pm and the last two groups were measured between 5:45 pm to 6 pm. As for vehicle type, most vehicles moving on the road during this period of time are light vehicles and buses. No heavy trucks were found. In terms of the lane-wise data, traffic volume is high in middle lanes (#4, #5, #6 and #9, #10, #11) for both directions and the speed is high in the fast lanes (#7 and #8). For the lanes near the roadside (#1 and #14), the traffic volumes and speeds are

Table 1. Lane-wise traffic flow data of group #1, group #2, group #3, group #4 (separated by slash sign).

Lane	Group #1/Group#2/Group #3/Group #4			Lane speed [km/h]	Lane gap [m]
	Traffic volume [veh/h]				
	Light vehicle	Medium vehicle	Bus		
#1	90/90/90/90	0/0/0/0	0/0/0/0	27/20/20/20	300/222.2/222.2/222.2
#2	720/810/810/810	0/0/90/0	90/0/270/90	36/36/36/32	44.44/44.44/30.77/35.56
#3	900/810/540/1260	0/0/0/0	90/0/90/0	48/52/52/52	48.48/64.20/82.54/41.27
#4	1350/990/1350/1170	0/90/0/0	0/90/0/0	54/56/56/52	40.00/47.86/41.48/44.44
#5	1350/1530/1710/1440	0/0/0/0	0/0/0/0	54/56/56/56	40.00/36.60/32.75/38.89
#6	990/1170/1350/1530	90/0/0/0	0/0/0/0	54/56/56/56	50.00/47.86/41.48/36.60
#7	990/720/720/1080	0/0/0/0	0/0/0/0	60/54/54/60	60.61/75.00/75.00/55.56
#8	990/1260/900/1350	0/0/0/0	0/0/0/0	58/58/60/62	58.59/46.03/66.67/45.93
#9	1260/1170/1530/1080	0/0/0/0	0/0/0/0	52/54/56/54	41.27/46.15/36.60/50.00
#10	900/1080/1530/1350	90/0/0/0	0/90/0/0	52/58/54/54	52.53/49.57/35.29/40.00
#11	1530/1080/990/1350	0/90/0/0	0/0/90/180	58/54/54/51	37.91/46.15/50.00/33.33
#12	360/360/1080/720	0/0/0/0	0/450/90/90	50/48/52/50	138.9/59.26/44.44/61.73
#13	720/1080/720/810	0/0/0/0	180/90/90/90	40/42/48/48	44.44/35.90/59.26/53.33
#14	0/0/0/0	0/0/0/0	0/0/90/0	0/0/45/0	-/-/500/-
Total	12690/13050/14130/14490				

very low. This phenomenon has two reasons. The first reason is that during rush hours, these two lanes are reserved only for buses and other vehicles are allowed to use them only for taking turns. The other reason is that electric and traditional motorcycles on the road nearly block the side lanes at rush hour, so buses can only take the lanes next to the side lanes.

3.3. Equivalent traffic flow

Based on the time-averaged lane-wise traffic flow data in Table 1, the equivalent steady traffic flow could be established numerically. In this equivalent traffic flow, vehicles on the same lane have identical speed and vehicle gap. The total number of vehicles depends on the traffic volume values. The type of vehicle also depends on the lane-wise percentage.

As mentioned above, each vehicle is regarded as a moving point sound source in half space. The vehicles are set to be moving at constant speed along the centerline of each lane. The z -coordinates of the sound source of vehicles depend on its type. For light, medium, and heavy vehicles and buses, the heights of the point sound sources are set to $z = 0.5$ m, 0.7 m, 1.0 m, and 1.0 m, respectively. A simple geometric configuration of vehicles on the road and the observer in the far field is shown in Fig. 4. In accordance with the field measurements, the duration of the equivalent traffic flow in the predictions is also set to 40 s.

3.4. Roadside traffic noise

The roadside traffic noise was measured simultaneously with the traffic flow. By averaging the measured noise data, the time-averaged noise spectra and the overall sound pressure levels could be derived for each group of measurement. As previously pointed out, the data of group #1 is used to calibrate the method and the other three groups' data are used for validation.

3.4.1. Calibration of the formulation

Since the observer-to-road distances are different in Lin's equations (LIN *et al.*, 2012) and in the current measurement, the constant values need to be calibrated. By shifting the constant values in Eq. (5)

by -1.94 dB, the measured and predicted traffic noise spectra are shown in Fig. 5. It can be seen in the figure that the predicted spectrum approximates the measured one within an error band of 3 dB between 400–6000 Hz where the sound pressure level is high. The error band in the measurements is 3 dB, because during the field measurement, randomly-occurred noise sources such as motorcycles, walking-by pedestrians and wind would alter the sound levels at the observation points and thus influencing the consistency of the measurement results. The A-weighted overall sound pressure level calculated from the prediction is 67.3 dBA, while the measured value is 67.7 dBA. Therefore, it is believed that under this parameter setting, the prediction is close enough to the measurement where the shifting value in Eq. (5) is -1.94 dB. These parameters are fixed and are employed to check whether the predictions by the proposed method could well approximate the other field measurements.

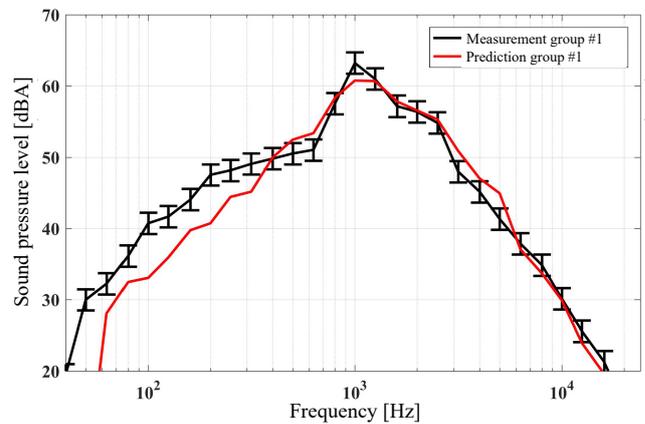


Fig. 5. Comparison of measured and predicted traffic noise spectrum for group #1 (error band for the measurement is 3 dB).

3.4.2. Validation of the method

With the calibrated formulation and the measured traffic flow data of group #2, group #3, and group #4, the noise generated by the traffic flow could be estimated and compared with the measured noise. Figure 6 compares the measured and predicted noise spectra.

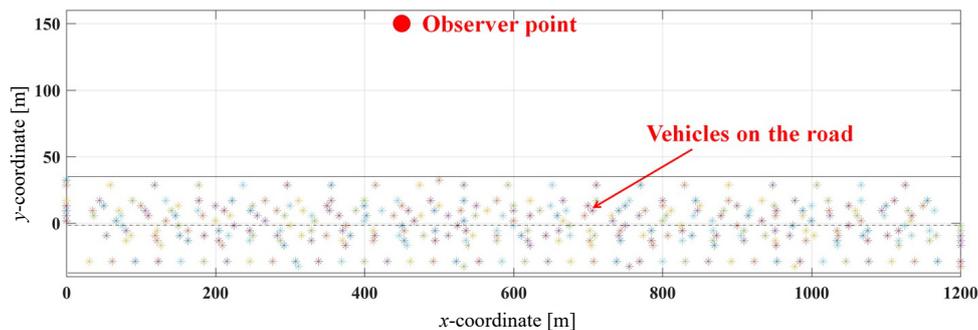


Fig. 4. Geometric configuration of the vehicles on the road and the observer point (each * represents one vehicle).

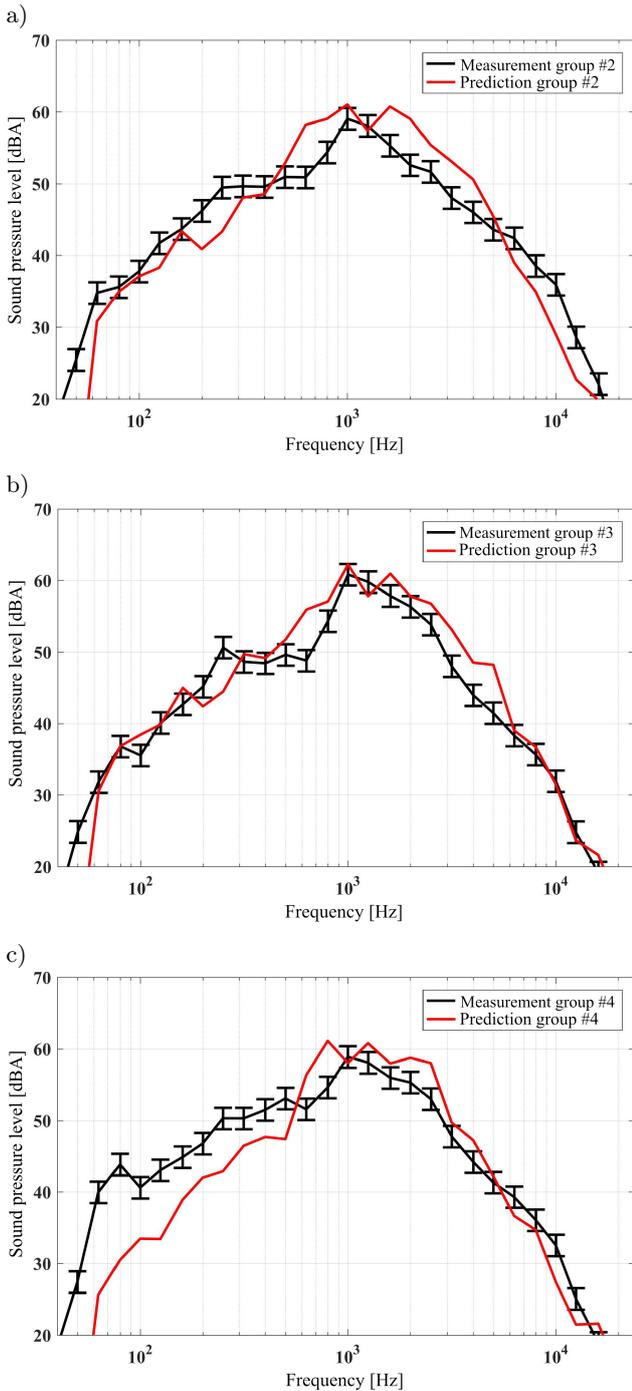


Fig. 6. Comparison of measured and predicted traffic noise spectrum for: a) group #2; b) group #3; c) group #4 (error band for the measurement is 3 dB).

For group #2 in Fig. 6a, the predicted spectrum approximates the measurement quite well. At some $1/3$ frequency bands from 100 Hz to 6000 Hz, the differences in sound pressure level exceed the error band of 3 dB but are mostly less than 5 dB. For group #3 and group #4, where the noise measurement point is on the other side of the road, the predicted spectra seem to be even closer to the measurements than the calibration group. Especially for group #3, the differ-

ences in sound pressure level from 50 Hz to 6000 Hz are mostly within the error band of 3 dB.

In terms of the overall sound pressure level, the predicted values are 67.3 dBA, 68.0 dBA, 68.0 dBA, and 67.7 dBA for group #1, group #2, group #3, and group #4, while the measured values are 67.7 dBA, 65.0 dBA, 66.5 dBA, and 65.3 dBA. The differences in the overall sound pressure level are -0.4 dB, 3.0 dB, 1.5 dB, and 2.4 dB, respectively. The four groups' overall sound pressure level differences are less than 3 dB.

Table 2. Comparison of the overall sound pressure level.

Purpose	Calibration		Validation	
Group number	#1	#2	#3	#4
Measurement	67.7 dBA	65.0 dBA	66.5 dBA	65.3 dBA
Prediction	67.3 dBA	68.0 dBA	68.0 dBA	67.7 dBA
SPL difference	-0.4 dB	3.0 dB	1.5 dB	2.4 dB

4. Prediction of roadside noise in different scenarios

In Sec. 3, the proposed traffic noise prediction method was calibrated and validated by field measurements. The predicted and the measured overall sound pressure levels and spectra show good coherence. On this basis, we intend to investigate the influence of traffic noise on people who live and work alongside arterial roads. This means more attention will be paid to the sound field distribution in the far field.

Of all the traffic scenarios, two are considered an obvious annoyance. The first one is during rush hours, when the impact of traffic noise is quite severe due to the high traffic volume on the road. Then it is interesting to know how the traffic noise evolves with traffic volume and vehicle average speed. The second scenario is near midnight when people are about to fall asleep. During this period of time, there are frequently fully-loaded heavy-duty trucks running on the road because trucks are not allowed to enter the urban area during daytime due to local traffic regulations. Since the traffic volume is low during that period, these heavy trucks tend to be driven at very elevated speeds generating high-level noise.

Therefore, this section intends to employ the previously established method to study the far-field traffic noise of the above two scenarios.

4.1. Roadside noise distribution during rush hours

In Subsec. 3.4.2, since the predicted and the measured spectra are the closest for group #3 (Fig. 6b) among group #2, group #3, and group #4, we employ the equivalent traffic flow established from the data of this group to represent the traffic flow during rush hours. Near the middle of the road segment ($x = 710$ m), a cross-section observation plane is set up,

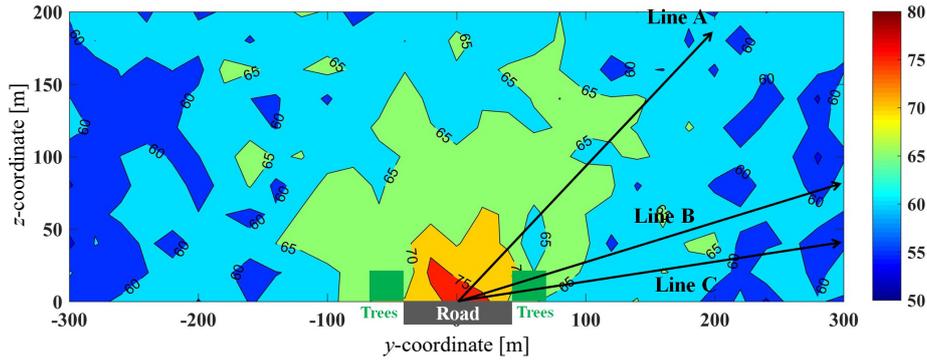


Fig. 7. Overall time-averaged A-weighted sound pressure level on cross-section observation plane (grid size 20×20 m).

which is 600 m in width (y -direction) and 200 m in height (z -direction).

With the equivalent traffic flow and the proposed traffic noise prediction method, the overall time-averaged A-weighted sound pressure level on the observation plane is calculated (Fig. 7). From the predicted results, it can be seen that traffic noise at rush hours on the tested road is quite strong. In the region between the two tree belts and within 100 m from the road centerline, the maximum overall sound pressure level could reach as high as 74.3 dBA, and the minimum value is no less than 65 dBA. The far-field noise level distribution has a sort of directivity due to the attenuation of the trees on both sides of the road. In the region “behind” the trees, the noise pressure level is reduced significantly.

To provide a detailed evolution of noise level with distance, the overall time-averaged A-weighted sound pressure level as a function of distance away from the road centerline is drawn in Fig. 8 along the directions of the arrows in Fig. 7. The angles between the arrow and positive y -axis for line A, line B, and line C are 45° , 18.43° , and 9.46° , respectively. It can be noted that the traffic noise level decreases with distance. Starting from 50 m, for the direction of line B and line C, due to the attenuation of the trees, the overall sound pressure levels are 2–3 dB lower than that along line A.

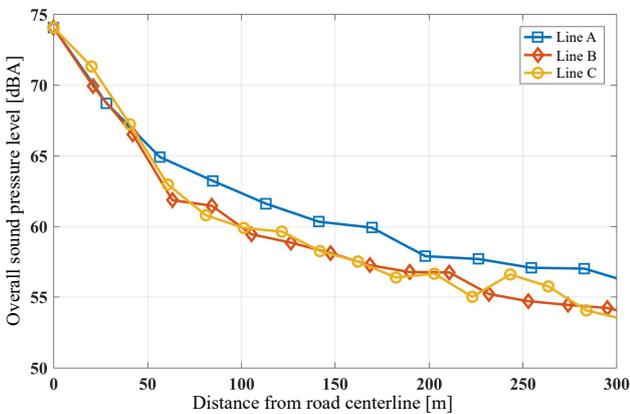


Fig. 8. Overall time-averaged A-weighted sound pressure level as a function of distance from the road centerline.

Figure 9 plots the predicted noise spectrum at a far-field observation point of (450, 350, and 100 m). This far-field point represent the location of the windows of the nearest high-rise residential buildings by the road. It can be seen that the sound pressure level exceeds 40.0 dBA during rush hours for a frequency range between 630 Hz to 2500 Hz. The overall predicted sound pressure level at this point is 53.2 dBA.

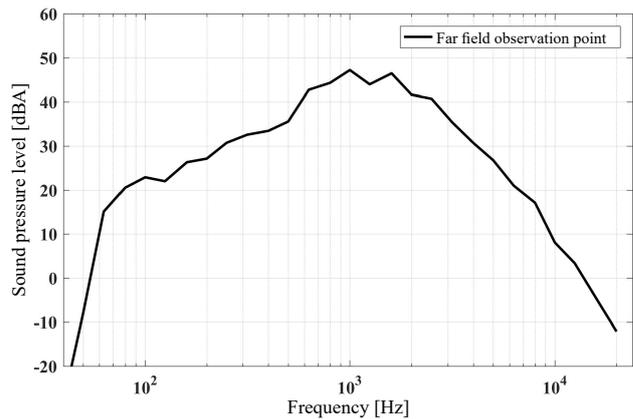


Fig. 9. Noise spectrum at far-field observation point (450, 350, and 100 m).

4.2. Traffic noise evolution with traffic volume

Besides rush hours, it is also interesting to know the traffic noise evolution when the traffic volume varies. In this section, we are going to test different traffic volumes. The equivalent traffic flow is also based on the data of group #3. The total traffic volume on the road is set to be 7065 veh/h, 10598 veh/h, 14130 veh/h, and 17663 veh/h, corresponding to 50%, 75%, 100%, 125% of that in group #3.

Figure 10 shows the predicted overall time-averaged A-weighted sound pressure level on the cross-section observation plane at $x = 710$ m for the traffic volume of 7065 veh/h, 10598 veh/h, and 17663 veh/h (the 14130 veh/h case is shown in Fig. 7). As the traffic volume increases, the maximum A-weighted overall sound pressure level near the road increases from 70.1 dBA to 73.5 dBA to 74.3 dBA and to 75.0 dBA.

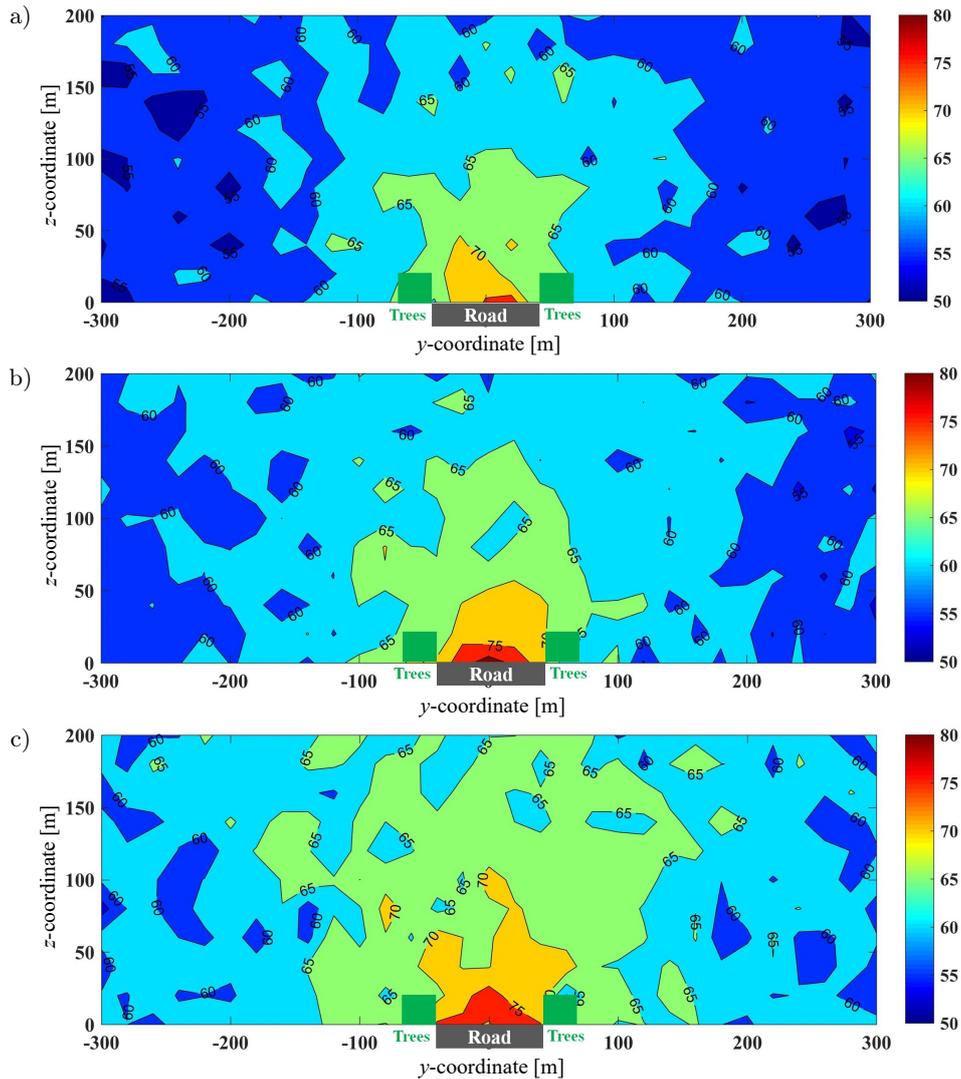


Fig. 10. Overall time-averaged A-weighted sound pressure level on cross-section observation plane at $x = 710$ m: a) 7065 veh/h; b) 10598 veh/h; c) 17663 veh/h (grid size 20×20 m).

Figure 11 compares the spatial evolution of the overall A-weighted sound pressure level at different

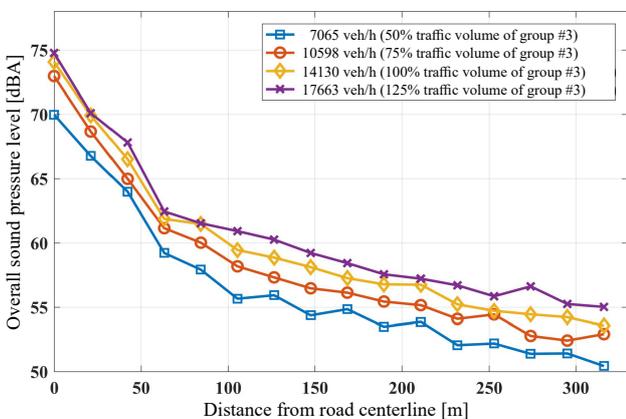


Fig. 11. Overall time-averaged A-weighted sound pressure level as a function of distance from the road centerline (along line B in Fig. 7) at different traffic volumes.

traffic volumes along line B in Fig. 7. From a traffic volume of 7065 veh/h, each increase of 3533 veh/h would bring an increase in the overall sound pressure level of about 1.8 dB, 1.3 dB, and 1.0 dB, respectively.

4.3. Traffic noise evolution with vehicle speed

In this section, the influence of the average vehicle speed on traffic noise is tested. To make a comparison, the lane-wise average vehicle speeds in group #3 are multiplied by a ratio of 0.75, 1, and 1.25, making an average speed of 37.45, 49.93, and 62.41 km/h.

Figure 12 shows the predicted overall time-averaged A-weighted sound pressure level on the cross-section observation plane at $x = 710$ m for the average speed of 37.45 km/h and 62.41 km/h (the 49.93 km/h case is shown in Fig. 7). It can be seen that with higher average vehicle speed, the traffic noise level becomes higher. The maximum A-weighted overall sound pres-

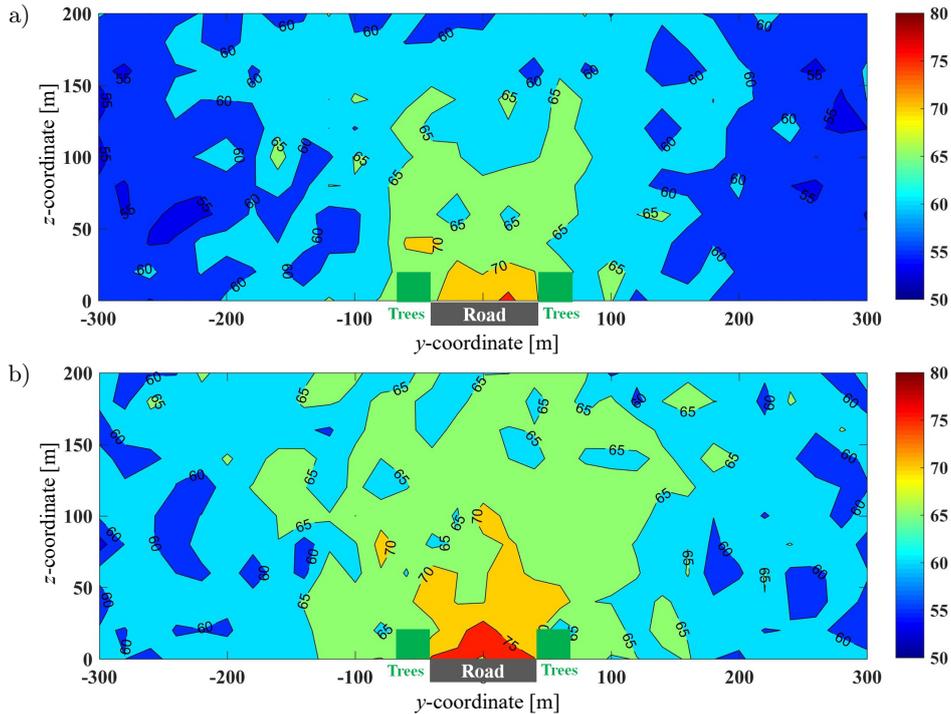


Fig. 12. Overall time-averaged A-weighted sound pressure level on cross-section observation plane at $x = 710$ m: a) 37.45 km/h; b) 62.41 km/h (grid size 20×20 m).

sure level near the road increases from 72.0 dBA to 74.3 dBA and to 77.5 dBA with increasing speed.

To quantify the influence of traffic flow speed, the spatial evolution of the overall A-weighted sound pressure level at different average vehicle speeds are compared in Fig. 13. The mean difference in decibel between the 37.45 km/h case and the 49.93 km/h case is 1.7 dB and that between the 50.31 km/h case and the 62.41 km/h case is 2.2 dB.

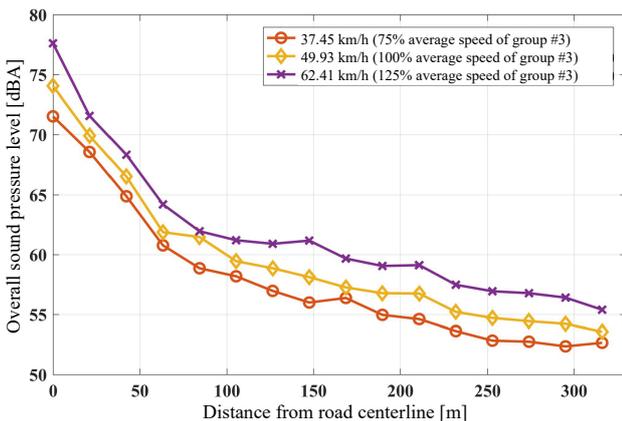


Fig. 13. Overall time-averaged A-weighted sound pressure level as a function of distance from the road centerline at different vehicle speeds.

4.4. Traffic noise from heavy trucks

Finally, the midnight scenario dominated by heavy-duty trucks' noise is studied. We name this case

group #5. During mid-night on Jingshi Road, except for heavy trucks, there are still occasionally light vehicles, but medium vehicles and buses are rare. Since the traffic volume is low, the vehicles on the road tend to be driven at full speeds. The speed limit on Jingshi Road is 80 km/h and the speed limit for trucks is 50 km/h within urban areas. Based on these considerations, the traffic flow for group #5 is presented in Table 3. In this case, heavy trucks having a lane-wise volume of 360 veh/h exist on lane #2, #3, #4, #11, #12, and #13 at the limit speed of 50 km/h. On the other lanes of the road exists a certain number of light vehicles at 80 km/h. The total traffic volume is 3420 veh/h for both directions.

Figure 14 plots the predicted overall time-averaged A-weighted sound pressure level distribution on the cross-section observation plane at $x = 710$ m. Compared with previously tested cases, the truck-dominated case is as 'noisy' as the group #3 case in Fig. 7. In the region near the road, the maximum A-weighted overall sound pressure level for group #5 is 73.5 dBA and that of group #3 is 74.3 dBA.

Since the noise characteristics depend on vehicle type, the noise spectral feature could be different even with the same value of sound pressure level. Figure 15 compares the spectra of the two cases. It is observed that for the truck-dominated case, the contribution at low frequencies is more important. Between 100 Hz and 200 Hz, the noise level emitted by trucks is 5 dB higher than for a mixed type of vehicles. The overall A-weighted sound pressure level for group #5 is

Table 3. Lane-wise traffic flow data of group #5.

Group #5						
Lane number	Traffic volume [veh/h]				Lane speed [km/h]	Lane gap [m]
	Light vehicle	Medium vehicle	Heavy vehicle	Bus		
#1	90	0	0	0	30	333.33
#2	0	0	360	0	50	138.89
#3	0	0	360	0	50	138.89
#4	0	0	360	0	50	138.89
#5	180	0	0	0	80	444.44
#6	180	0	0	0	80	444.44
#7	180	0	0	0	80	444.44
#8	180	0	0	0	80	444.44
#9	180	0	0	0	80	444.44
#10	180	0	0	0	80	444.44
#11	0	0	360	0	50	138.89
#12	0	0	360	0	50	138.89
#13	0	0	360	0	50	138.89
#14	90	0	0	0	30	333.33
3420 (total)						

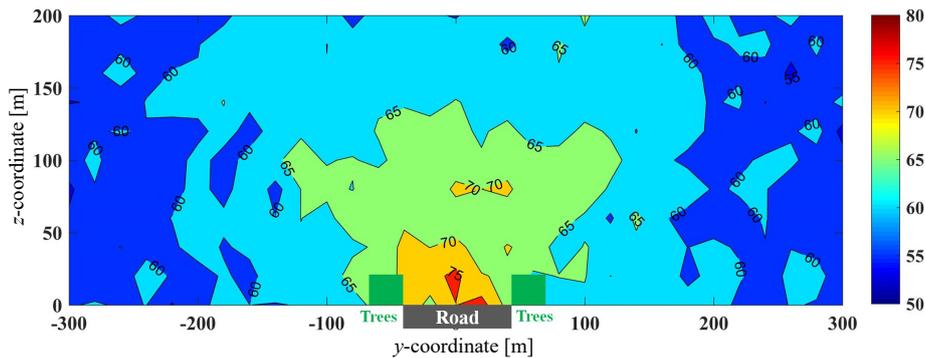


Fig. 14. Overall time-averaged A-weighted sound pressure level on cross-section observation plane at $x = 710$ m for the heavy truck-dominated case (grid size 20×20 m).

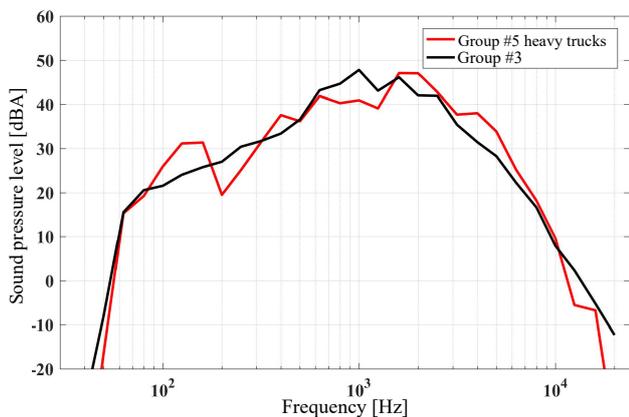


Fig. 15. Noise spectra at far-field observation point (450 m, 350 m, 100 m) for the truck-dominated case and that of group #3 case.

52.9 dBA, while that of group #3 at the same observation point is 53.2 dBA.

5. Conclusions

Focusing on the influence of traffic noise from arterial roads on people in nearby high-rise buildings, this paper proposes and validates a microscopic-level noise prediction method that could estimate the spatial distribution and spectral characteristics of traffic noise in the vicinity of multi-lane arterial roads.

In the prediction method, the sound pressure field for each vehicle could account for the motion of the vehicle and the reflection effect of the ground. With lane-wise field data as input, an equivalent traffic flow could be established to model the road as a multiple-lane noise source. Accompanied by the noise source model and spectral characteristics in the literature, the method is able to evaluate the spectral and total sound pressure level at a certain observer point by superimposing the individual contribution of all the vehicles on the road. The proposed method is only applicable

to noise issues near city arterial roads where the linear source assumption of the traffic flow does not hold.

To validate the proposed method, a field measurement on a segment of Jingshi Road, Jinan, China, was conducted, obtaining four groups of traffic flow and noise data. The data of group #1 was employed to calibrate the formulation, whereas the other three groups were used for validation. The validation shows that, after calibration, the predicted overall sound pressure level could approximate the measured values by no more than 3 dB and the predicted noise spectra are within an error band of less than 5 dB for the frequency range between 500 Hz to 6000 Hz.

Based on the validated method, the roadside noise of some scenarios was studied. First, the traffic noise during rush hours was investigated. Prediction results show that the overall sound pressure level of the traffic flow during rush hours is over 60.0 dBA within a distance of 100 m from the road centerline. At a distance of 300 m, the noise level could still be as high as over 55.0 dBA. Then the influence of traffic volume and speed is investigated. The investigation shows that from a total traffic volume of 7065 veh/h to 17663 veh/h, each increase of 3533 veh/h would cause the overall sound pressure level to increase by 1.8 dB, 1.3 dB, and 1.0 dB, respectively. From an average traffic speed of 37.45 km/h to 62.41 km/h, each increase of 12.48 km/h would cause the overall sound pressure level to increase by 1.7 dB and 2.2 dB.

At last, a heavy truck-dominated scenario was studied. Prediction results show that at a total traffic volume of 3420 veh/h, the heavy truck-dominated scenario during the night could be as noisy as a total traffic volume of 14130 veh/h for a mixture of vehicles in the daytime. It was also found that the spectral contribution in the low-frequency range of 100 Hz and 200 Hz is more important for the heavy truck-dominated case.

The current paper demonstrates that the proposed method predicts traffic noise distribution near a segment of arterial road with a short-term quasi-steady traffic flow with a reasonable accuracy. However, the applicability and feasibility of the proposed method for long-term dynamic scenarios with detailed traffic flow data have not been investigated. Besides, some of the details in the modeling could be further extended, including the effect of sound attenuation of the ground (compacted field, loose ground, soft forest floor, etc.) and the effect of noise shielding of the vegetation belts (VAN RENTERGHEM *et al.*, 2012).

Acknowledgments

This research was supported by the Shandong Provincial Natural Science Foundations, China (Grants no. ZR2021QG061 and ZR2021QG150), the Doctoral

Research Foundations of Shandong Jianzhu University, China (Grants no. X21108Z and X21107Z).

References

1. BARRIGÓN MORILLAS J.M., REY GOZALO G., MONTES GONZÁLEZ D., SÁNCHEZ-FERNÁNDEZ M., BACHILLER LEÓN A. (2022), A comprehensive experimental study of the influence of temperature on urban road traffic noise under real-world conditions, *Environmental Pollution*, **309**: 119761, doi: [10.1016/j.envpol.2022.119761](https://doi.org/10.1016/j.envpol.2022.119761).
2. BARRY T.M., REAGAN J.A. (1978), *FHWA highway traffic noise prediction model*, Washington: Department of Transportation, Federal Highway Administration National Technical Information Service.
3. BENDTSEN H. (1999), The Nordic prediction method for road traffic noise, *Science of the Total Environment*, **235**(1–3): 331–338, doi: [10.1016/S0048-9697\(99\)00216-8](https://doi.org/10.1016/S0048-9697(99)00216-8).
4. CoRTN (1975), *Calculation of road traffic noise*, United Kingdom Department of the Environment and Welsh Office Joint Publication.
5. Directive EN (2015), *Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council*, <http://data.europa.eu/eli/dir/2015/996/oj>.
6. DUTILLEUX G. *et al.* (2010), NMPB-Route-2008: The revision of the French method for road traffic noise prediction, *Acta Acustica United with Acustica*, **96**(3): 452–462, doi: [10.3813/AAA.918298](https://doi.org/10.3813/AAA.918298).
7. HINTON J., JELLYMAN A., HOWELL L.K. (2005), BUMP – The Birmingham updated noise mapping, *Forum Acusticum*, pp. 1003–1006.
8. Hourly Historical Weather Data [OL] (n.d.), <https://www.weatherbit.io/history/hourly> (access: 1.06.2021).
9. HUDSON S. (2008), Ground reflection, *Lecture reference of Engineering for Telecommunications of Washington State University*.
10. KADDOURA I., KRÖGER L., NAGEL K. (2017), An activity-based and dynamic approach to calculate road traffic noise damages, *Transportation Research Part D: Transport and Environment*, **54**: 335–347, doi: [10.1016/j.trd.2017.06.005](https://doi.org/10.1016/j.trd.2017.06.005).
11. KEPHALOPOULOS S., PAVIOTTI M., ANFOSSO-LÉDÉE F. (2012), *Common Noise Assessment Methods in Europe (CNOSSOS-EU)*, EUR 25379 EN, Publications Office of the European Union, Luxembourg.
12. KHAN J., KETZEL M., JENSEN S.S., GULLIVER J., THYSELL E., HERTEL O. (2021), Comparison of road traffic noise prediction models: CNOSSOS-EU, Nord 2000 and TRANEX, *Environmental Pollution*, **270**: 116240, doi: [10.1016/j.envpol.2020.116240](https://doi.org/10.1016/j.envpol.2020.116240).

13. KNABBen K.M., TRICHÈS G., GERGES S.N.Y., VERGARA E.F. (2016), Evaluation of sound absorption capacity of asphalt mixtures, *Applied Acoustics*, **114**: 266–274, doi: [10.1016/j.apacoust.2016.08.008](https://doi.org/10.1016/j.apacoust.2016.08.008).
14. LAN Z., HE C., CAI M. (2020), Urban road traffic noise spatiotemporal distribution mapping using multisource data, *Transportation Research Part D: Transport and Environment*, **82**: 102323, doi: [10.1016/j.trd.2020.102323](https://doi.org/10.1016/j.trd.2020.102323).
15. LEE S.-W., CHANG S.I., PARK Y.-M. (2008), Utilizing noise mapping for environmental impact assessment in a downtown redevelopment area of Seoul, Korea, *Applied Acoustics*, **69**(8): 704–714, doi: [10.1016/j.apacoust.2007.02.009](https://doi.org/10.1016/j.apacoust.2007.02.009).
16. LI M., VAN KEULEN W., VAN DE VEN M., MOLENAAR A., TANG G. (2014), Investigation on material properties and surface characteristics related to tyre-road noise for thin layer surfacings, *Construction and Building Materials*, **59**: 62–71, doi: [10.1016/j.conbuildmat.2014.02.050](https://doi.org/10.1016/j.conbuildmat.2014.02.050).
17. LIN Y., CAI M., LI F. (2012), Studies on traffic noise source intensity regard for the acceleration, *Applied Acoustics (Chinese version)*, **31**: 282–286.
18. LOKHANDE S.K., DHAWALE S.A., PATHAK S.S., GAUTAM R., JAIN M.C., BODHE G.L. (2017), Appraisal of noise level dissemination surrounding mining and industrial areas of Keonjhar, Odisha: A comprehensive approach using noise mapping, *Archives of Acoustics*, **42**(3): 423–432, doi: [10.1515/aoa-2017-0044](https://doi.org/10.1515/aoa-2017-0044).
19. LOKHANDE S.K., PATHAK S.S., KOKATE P.A., DHAWALE S.A. BODHE G.L. (2018), Assessment of heterogeneous road traffic noise in Nagpur, *Archives of Acoustics*, **43**(1): 113–121, doi: [10.24425/118086](https://doi.org/10.24425/118086).
20. LUO P., CAI M., WANG H. (2013), The noise spectral characteristics of various types of vehicles, *Noise and Vibration Control (Chinese version)*, **33**(5): 86–89.
21. McLAUGHLIN D.K., KUO C.-W., PAPAMOSCHOU D. (2008), Experiments on the effect of ground reflections on supersonic jet noise, [in:] *46th AIAA Aerospace Sciences Meeting and Exhibit*, doi: [10.2514/6.2008-22](https://doi.org/10.2514/6.2008-22).
22. NIELSEN H.L. *et al.* (1996), *Road Traffic Noise – Nordic Prediction Method*, TemaNord 1996:525, Nordic Council of Ministers, Copenhagen, Denmark.
23. PENG J., LIU D., PARNELL J., KESSISSOGLU N. (2019), Influence of translational vehicle dynamics on heavy vehicle noise emission, *Science of the Total Environment*, **689**: 1358–1369, doi: [10.1016/j.scitotenv.2019.06.426](https://doi.org/10.1016/j.scitotenv.2019.06.426).
24. POPP C. (2003), Noise abatement planning in German-experiences and consequence of the EU directive on the assessment of environmental noise, *Acta Acustica United with Acustica*, **Suppl. 1**: 65–67.
25. PRICE M.A., ATTENBOROUGH K., HEAP N.W. (1988), Sound attenuation through trees: Measurements and models, *The Journal of the Acoustical Society of America*, **84**(5): 1836–1844, doi: [10.1121/1.397150](https://doi.org/10.1121/1.397150).
26. QUARTIERI J. *et al.* (2009), A review of traffic noise predictive models, [in:] *Recent Advances in Applied and Theoretical Mechani Conference: Recent Advances in Applied and Theoretical Mechanics*, pp. 72–80.
27. RLS-90 (1990), *Guidelines for noise protection on roads* [in German: *Richtlinien für den Lärmschutz an Strassen*], Der Bundesminister Für Verkehr.
28. SANOK S. *et al.* (2022), Road traffic noise impacts sleep continuity in suburban residents: Exposure-response quantification of noise-induced awakenings from vehicle pass-bys at night, *Science of The Total Environment*, **817**: 152594, doi: [10.1016/j.scitotenv.2021.152594](https://doi.org/10.1016/j.scitotenv.2021.152594).
29. SMITH III J. (2010), Spherical waves from a point source, [in:] *Physical Audio Signal Processing for Virtual Musical Instruments and Audio Effects*, online version, https://ccrma.stanford.edu/~jos/pasp/Spherical_Waves_Point_Source.html.
30. STAAB J., SCHADY A., WEIGAND M., LAKES T., TAUBENBÖCK H. (2022), Predicting traffic noise using land-use regression – A scalable approach, *Journal of Exposure Science & Environmental Epidemiology*, **32**: 232–243, doi: [10.1038/s41370-021-00355-z](https://doi.org/10.1038/s41370-021-00355-z).
31. STEELE C. (2001), A critical review of some traffic noise prediction models, *Applied Acoustics*, **62**(3): 271–287, doi: [10.1016/S0003-682X\(00\)00030-X](https://doi.org/10.1016/S0003-682X(00)00030-X).
32. STOILOVA K., STOILOV T. (1998), Traffic noise and traffic light control, *Transportation Research Part D: Transport and Environment*, **3**(6): 399–417, doi: [10.1016/S1361-9209\(98\)00017-0](https://doi.org/10.1016/S1361-9209(98)00017-0).
33. VAN RENTERGHEM T., BOTTELDOOREN D., VERHEYEN K. (2012), Road traffic noise shielding by vegetation belts of limited depth, *Journal of Sound and Vibration*, **331**(10): 2404–2425, doi: [10.1016/j.jsv.2012.01.006](https://doi.org/10.1016/j.jsv.2012.01.006).
34. 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).

Appendix A. Frequency-dependent road reflection coefficient

The reflection coefficient $C_g(f)$ of the road employed in the model of the current paper takes the averaged values of the measurement data by LI *et al.* (2014) and KNABBEN *et al.* (2016) for asphalt pavements.

Table 4. Reflection coefficient $C_g(f)$ at $1/3$ octave frequencies.

Frequency [Hz]	Reflection coefficient	Frequency [Hz]	Reflection coefficient	Frequency [Hz]	Reflection coefficient
40	0.909	400	0.889	4000	0.689
50	0.908	500	0.883	5000	0.689
63	0.908	630	0.876	6300	0.689
80	0.907	800	0.867	8000	0.689
100	0.906	1000	0.856	10000	0.689
125	0.904	1250	0.842	12500	0.689
160	0.902	1600	0.822	16000	0.689
200	0.900	2000	0.800	20000	0.689
250	0.897	2500	0.772		
315	0.894	3150	0.736		

Appendix B. Vehicle spectral noise database

Based on the noise spectra provided by YANG *et al.* (2020) the authors of this paper have established a vehicle noise database for four vehicle types and five vehicle speed intervals at 28 $1/3$ octave frequencies.

Table 5. Spectral energy contribution at $1/3$ octave frequencies for light vehicle.

Velocity [km/h]	40 Hz	50 Hz	63 Hz	80 Hz	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz
0 ~ 18	0.001	0.005	0.007	0.014	0.019	0.022	0.025	0.045	0.048	0.055
18 ~ 36	0.001	0.005	0.007	0.013	0.019	0.022	0.024	0.044	0.048	0.054
36 ~ 54	0.001	0.005	0.007	0.012	0.018	0.021	0.023	0.043	0.047	0.053
54 ~ 72	0.001	0.004	0.007	0.010	0.017	0.020	0.022	0.043	0.046	0.052
72 ~ 90	0	0.001	0.001	0.001	0.002	0.002	0.005	0.011	0.02	0.013
Velocity [km/h]	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz
0 ~ 18	0.043	0.059	0.084	0.100	0.132	0.089	0.084	0.070	0.046	0.020
18 ~ 36	0.043	0.058	0.083	0.098	0.131	0.090	0.085	0.071	0.047	0.021
36 ~ 54	0.042	0.057	0.082	0.098	0.130	0.091	0.086	0.072	0.048	0.022
54 ~ 72	0.041	0.056	0.082	0.096	0.128	0.093	0.87	0.073	0.049	0.023
72 ~ 90	0.019	0.064	0.06	0.065	0.137	0.152	0.166	0.142	0.078	0.033
Velocity [km/h]	4000 Hz	5000 Hz	6300 Hz	8000 Hz	10000 Hz	12500 Hz	16000 Hz	20000 Hz		
0 ~ 18	0.008	0.004	0.004	0.005	0.003	0.003	0.003	0.002		
18 ~ 36	0.019	0.006	0.005	0.005	0.003	0.003	0.003	0.002		
36 ~ 54	0.011	0.007	0.006	0.006	0.004	0.003	0.003	0.002		
54 ~ 72	0.012	0.009	0.007	0.008	0.004	0.003	0.003	0.002		
72 ~ 90	0.013	0.008	0.004	0.002	0.001	0	0	0		

Table 6. Spectral energy contribution at $1/3$ octave frequencies for a medium vehicle.

Velocity [km/h]	40 Hz	50 Hz	63 Hz	80 Hz	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz
0 ~ 18	0.001	0.006	0.015	0.003	0.006	0.011	0.014	0.032	0.046	0.05
18 ~ 36	0.001	0.006	0.014	0.002	0.005	0.009	0.009	0.027	0.042	0.043
36 ~ 54	0.001	0.006	0.012	0.002	0.005	0.008	0.008	0.026	0.037	0.041
54 ~ 72	0.001	0.004	0.01	0.002	0.005	0.008	0.008	0.023	0.034	0.037
72 ~ 90	0.001	0.003	0.005	0.002	0.005	0.007	0.008	0.02	0.034	0.04
Velocity [km/h]	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz
0 ~ 18	0.072	0.091	0.08	0.078	0.067	0.071	0.088	0.077	0.076	0.052
18 ~ 36	0.073	0.096	0.086	0.081	0.067	0.072	0.092	0.077	0.077	0.054
36 ~ 54	0.07	0.09	0.086	0.81	0.07	0.079	0.096	0.081	0.078	0.054
54 ~ 72	0.066	0.086	0.084	0.083	0.074	0.081	0.1	0.084	0.08	0.056
72 ~ 90	0.064	0.08	0.081	0.084	0.079	0.087	0.103	0.086	0.082	0.059
Velocity [km/h]	4000 Hz	5000 Hz	6300 Hz	8000 Hz	10000 Hz	12500 Hz	16000 Hz	20000 Hz		
0 ~ 18	0.029	0.021	0.008	0.004	0.001	0.001	0	0		
18 ~ 36	0.03	0.022	0.009	0.004	0.001	0.001	0	0		
36 ~ 54	0.032	0.022	0.009	0.004	0.001	0.001	0	0		
54 ~ 72	0.034	0.024	0.01	0.004	0.001	0.001	0	0		
72 ~ 90	0.038	0.019	0.008	0.003	0.001	0.001	0	0		

Table 7. Spectral energy contribution at $1/3$ octave frequencies for a heavy vehicle.

Velocity [km/h]	40 Hz	50 Hz	63 Hz	80 Hz	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz
0 ~ 18	0	0.001	0.01	0.028	0.011	0.011	0.02	0.04	0.054	0.071
18 ~ 36	0	0.001	0.01	0.026	0.008	0.009	0.017	0.036	0.05	0.068
36 ~ 54	0	0.001	0.01	0.026	0.007	0.008	0.015	0.033	0.047	0.066
54 ~ 72	0	0.001	0.008	0.022	0.006	0.007	0.014	0.031	0.045	0.064
72 ~ 90	0	0.001	0.006	0.02	0.005	0.006	0.013	0.03	0.043	0.06
Velocity [km/h]	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz
0 ~ 18	0.076	0.079	0.073	0.067	0.072	0.071	0.075	0.071	0.064	0.04
18 ~ 36	0.073	0.078	0.073	0.069	0.074	0.074	0.078	0.073	0.067	0.041
36 ~ 54	0.07	0.075	0.074	0.074	0.079	0.077	0.079	0.074	0.068	0.045
54 ~ 72	0.068	0.072	0.071	0.075	0.085	0.081	0.084	0.077	0.071	0.048
72 ~ 90	0.065	0.07	0.071	0.079	0.089	0.086	0.09	0.08	0.073	0.05
Velocity [km/h]	4000 Hz	5000 Hz	6300 Hz	8000 Hz	10000 Hz	12500 Hz	16000 Hz	20000 Hz		
0 ~ 18	0.028	0.017	0.01	0.006	0.003	0.001	0.001	0		
18 ~ 36	0.03	0.019	0.012	0.007	0.004	0.002	0.001	0		
36 ~ 54	0.031	0.018	0.011	0.006	0.003	0.002	0.001	0		
54 ~ 72	0.033	0.018	0.01	0.005	0.002	0.001	0.001	0		
72 ~ 90	0.031	0.016	0.008	0.004	0.002	0.001	0.001	0		

Table 8. Spectral energy contribution at $1/3$ octave frequencies for a bus.

Velocity [km/h]	40 Hz	50 Hz	63 Hz	80 Hz	100 Hz	125 Hz	160 Hz	200 Hz	250 Hz	315 Hz
0 ~ 18	0	0.006	0.091	0.101	0.082	0.069	0.075	0.026	0.027	0.031
18 ~ 36	0	0.004	0.088	0.099	0.08	0.067	0.075	0.025	0.025	0.03
36 ~ 54	0	0.004	0.086	0.095	0.077	0.065	0.072	0.025	0.026	0.031
54 ~ 72	0	0.004	0.083	0.092	0.075	0.063	0.069	0.024	0.027	0.032
72 ~ 90	0	0.004	0.08	0.088	0.072	0.06	0.065	0.024	0.028	0.033
Velocity [km/h]	400 Hz	500 Hz	630 Hz	800 Hz	1000 Hz	1250 Hz	1600 Hz	2000 Hz	2500 Hz	3150 Hz
0 ~ 18	0.038	0.037	0.045	0.043	0.041	0.042	0.042	0.044	0.049	0.036
18 ~ 36	0.039	0.038	0.046	0.044	0.042	0.043	0.043	0.045	0.05	0.037
36 ~ 54	0.041	0.04	0.047	0.045	0.043	0.044	0.044	0.046	0.051	0.038
54 ~ 72	0.043	0.042	0.049	0.047	0.044	0.045	0.046	0.048	0.052	0.04
72 ~ 90	0.045	0.044	0.051	0.048	0.046	0.047	0.047	0.049	0.054	0.042
Velocity [km/h]	4000 Hz	5000 Hz	6300 Hz	8000 Hz	10000 Hz	12500 Hz	16000 Hz	20000 Hz		
0 ~ 18	0.029	0.022	0.011	0.008	0.003	0.001	0.001	0		
18 ~ 36	0.031	0.024	0.013	0.007	0.003	0.001	0.001	0		
36 ~ 54	0.032	0.025	0.012	0.006	0.003	0.001	0.001	0		
54 ~ 72	0.033	0.023	0.011	0.005	0.002	0.001	0	0		
72 ~ 90	0.035	0.022	0.01	0.004	0.001	0.001	0	0		

Appendix C. Frequency-dependent sound attenuation by vegetation

The amount of sound attenuation (in dB) by vegetation in the current paper takes the averaged values of the measurement data by PRICE *et al.* (1988).

Table 9. Sound attenuation by vegetation.

Frequency [Hz]	Amount of sound attenuation [dB]	Frequency [Hz]	Amount of sound attenuation [dB]	Frequency [Hz]	Amount of sound attenuation [dB]
40	2.000	400	2.000	4000	4.000
50	2.000	500	2.000	5000	4.667
63	2.000	630	2.000	6300	5.533
80	2.000	800	2.000	8000	6.667
100	2.000	1000	2.000	10000	8.000
125	2.000	1250	2.167	12500	9.667
160	2.000	1600	2.400	16000	12.000
200	2.000	2000	2.667	20000	14.667
250	2.000	2500	3.000		
315	2.000	3150	3.433		

Chronicle

51st Winter School on Wave and Quantum Acoustics Lth Winter School on Environmental Acoustics and Vibroacoustics Szczyrk, Poland, February 27 – March 3, 2023

On behalf of the Upper Silesian Division of the Polish Acoustical Society (the main organizer) and the Committee of Acoustics of the Polish Academy of Sciences (the co-organizer) we are pleased to announce that after many years of tradition, the above-mentioned Conferences known as the “Winter Schools” have been organized and completed with success.

This year we were celebrating the Lth Winter School on Environmental Acoustics and Vibroacoustics, a conference well-known among Polish acousticians. For this reason, the “Winter Schools” began with a special, joined session dedicated to the celebration of this jubilee. As it has always been over the past 50 years of the conference history, Lth Winter School on Environmental Acoustics and Vibroacoustics was concerned with all environmental and vibroacoustics fields, but particularly with the traffic noise, vibroacoustics of machines, room acoustics, building acoustics, noise protection and similar problems. Many persons well-known from previous editions of the Conference joined us, either on-site or on-line. During the Conference, a seminar on “Industrial and Impulsive Noise” was organized in association with SVANTEK, which attracted 30 participants.

51st Winter School on Wave and Quantum Acoustics constituted platforms for sharing the results and achievements obtained in different branches of physical acoustics, as molecular acoustics, quantum acoustics, acousto-optics, magnetoacoustics, photoacoustics, acoustics of solid state, acoustic emission, and others. Moreover, researches in some selected topics related to those mentioned above (e.g., optoelectronics, relaxation processes) were presented during the school. The Conference consisted of the 18th Workshop on Acoustoelectronics and the 18th Workshop on Molecular Acoustics, Relaxation and Calorimetric Methods. However, the organizers are opened to organizing workshops on other subjects in the future. We would like to invite scientific centers and other professional groups to cooperate in organizing workshops on the subjects of their interests.

In summary, 51 persons participated in the Conferences and seminars, presenting 46 lectures, reports, and posters. In this issue, one can find abstracts of some lectures and posters, which were presented during the Conferences.

Further information about Conferences is available on our website:

<https://ogpta.pl/index.php/en/>

Dariusz Bismor

Secretary of the Organizing Committee

Abstracts

Pickering droplets and liquid marbles as templates for the formation of capsules

BIELAS Rafał (Rafal.Bielas@amu.edu.pl),
KUBIAK Tomasz, JÓZEF CZAK Arkadiusz

Faculty of Physics, Adam Mickiewicz University
Poznań, Poland

Various types of pharmaceuticals can be delivered to the site of interest via encapsulation. The essential requirements for efficient capsule carriers include their durability, resistance against destabilization, and the robustness of the preparation technique. One of the possible approaches to fabricating such capsules, which can be used in biomedical applications, is using Pickering droplets, i.e., the droplets covered by solid particles, as templates and making their particle shell more rigid by sintering under high-amplitude alternating magnetic fields. In our work, we compared the capsules formed from Pickering droplets with those prepared using liquid marbles, i.e., droplets covered by particles residing at the liquid-air interface. The results suggest that both methods could be used for different purposes, depending on the possibility of locating magnetic nanoparticles as nano-heaters inside or outside the droplet. In the future, such fabricated capsules could be used in magnetically-responsive targeting of active substances, e.g., antibiotics.

The work was supported by the Polish National Science Center by the grant 2019/35/N/ST5/00402.

* * *

Structural active noise control using the leaky partial update LMS algorithms

BISMOR Dariusz (Dariusz.Bismor@polsl.pl)

Department of Measurements and Control Sciences
Silesian University of Technology
Gliwice, Poland

Structural Active Noise Control (ANC) systems are one of few solutions of the ANC problem which allow to obtain a global noise reduction effect. Unfortunately, due to high dimensionality of these multichannel adaptation systems, which use many sensors and actuators, structural ANC systems are systems with high computational power requirements. A promising group of methods allowing to reduce these requirements are partial update LMS algorithms. In this communication, a modification of partial update LMS algorithms with leakage is presented. The computational power savings are discussed, and two simulation setups are presented to test the leaky partial updates. The results of the simulations confirm the algorithms are stable and provide good attenuation results.

* * *

Ultra-long carbon nanotube-paraffin composites of record thermal conductivity

BONCEL Sławomir¹ (Sławomir.Boncel@polsl.pl),
KUZIEL Anna W.¹, DZIDO Grzegorz²,
TURCZYN Roman³, JEŃDRYSIAK Rafał G.¹,
KOLANOWSKA Anna¹, TRACZ Anna⁴,
ZIEBA Wojciech⁵, CYGANIUK Aleksandra⁵,
TERZYK Artur P.⁵

¹ Department of Organic Chemistry, Bioorganic Chemistry and Biotechnology, Silesian University of Technology
Gliwice, Poland

² Department of Chemical and Process Design
Silesian University of Technology
Gliwice, Poland

³ Department of Physical Chemistry and Technology of Polymers, Silesian University of Technology
Gliwice, Poland

⁴ Grupa Azoty, Zakłady Azotowe Kędzierzyn S.A.
Kędzierzyn-Koźle, Poland

⁵ Faculty of Chemistry, Physicochemistry of Carbon Materials Research Group
Nicolaus Copernicus University in Toruń
Toruń, Poland

Phase change materials (PCMs) are capable of storage considerably more energy than conventional systems based on sensible heat. Despite immense and global research, there is a continuous pursuit for high-performance PCMs among which carbon nanocomposites emerge as the most prospective ones. Here, by comprehensive analysis of carbon nanotubes (CNTs) of three various morphologies (crystallinity, number of walls, and aspect ratio), we report record-breaking characteristics of CNT-paraffin nanocomposites based on ultra-long (770 μm) in-house multi-wall CNTs (MWCNTs) as fully functional PCMs prepared by a melting technique. By systematic investigations covering scanning electron microscopy (SEM), transmission electron microscopy (TEM), X-ray diffraction (XRD), and analysis of thermophysical properties, we have constructed the most

promising MWCNT(0.5% wt.)-paraffin nanocomposite of 37%-enhanced thermal conductivity and 6.3%-higher enthalpy of the phase change (ΔH_m), in reference to the base paraffin, as well as excellent cycling stability (>50 heating/cooling cycles), and as low supercooling temperature ($\Delta T = T_m - T_c$) as 2.4°C. The superior characteristics derive from more rapid nucleation of larger crystallites by MWCNTs proceeding via short- and long-range templating, as well as intrinsic characteristics of individual and fibrous ultra-long MWCNTs. Additionally, even a 161%-enhancement in thermal conductivity is available for the long MWCNT-paraffin composite, but at the cost of preserving the remaining thermophysical characteristics of neat paraffin. The results clearly point out the potential of the elaborated PCMs for thermal energy storage.

* * *

System for monitoring and identification of vibroacoustic threats – SMIZW

CHMIELEWSKI Bartosz, NIERADKA Paweł

KFB Acoustics
Domasław, Poland

The presentation will discuss the System for Monitoring and Identification of Vibroacoustic Threats (SMIZW). The SMIZW system combines artificial intelligence and deterministic algorithms to automatically analyze measured noise for hazards that cause noise exceedances. The system is based on a system of independent sensors and performs automatic classification of acoustic events in real time and continuously creates a ranking of threats (noise sources). An integral part of SMIZW is a graphical user interface that allows simple operation of the system and generation of analysis reports.

* * *

Comparative analysis of acoustic emission signals generated by the on-load tap-changer with regard to the possibility of detecting non-simultaneous operation

CICHOŃ Andrzej (a.cichon@po.edu.pl),
WŁODARZ Michał

Faculty of Electrical Engineering, Automatic Control and Informatics
Opole University of Technology
Opole, Poland

The article presents the results of using acoustic method for diagnostics of the on-load tap changer. Measurements were carried out in laboratory conditions by simulating non-simultaneous operation on physical on-load tap-changer model. An oil-insulated tap-changer of the VEL-110 type, often used in polish power system, was used for the study. Measurements were made simultaneously using four piezoelectric transducers mounted on the outer wall of the tap-changer model. In addition, electrical measurements were carried out – oscylographic method was used as well as the motor current in the device's drive system was recorded. The main goal of the research was to determine the possibility of detecting the nonsimultaneous operation of the tap-changer using piezoelectric

transducers with different transmission characteristics. Obtained results were correlated with oscillographic measurements. Sensitivity of piezoelectric transducers was specified to changes in the system. The results were analyzed in the time and time-frequency domain. The article also attempts to verify the feasibility of using selected acoustic signal descriptors to conduct diagnostics of on-load tap-changers for detection of non-simultaneous operation.

* * *

Thermal properties of ionic liquids with functionalized multi-walled carbon nanotubes

CWYNAR Krzysztof¹ (krzysztof.cwynar@us.edu.pl),
JĘDRYSIAK Rafał², BONCEL Sławomir²,
DZIDA Marzena¹

¹ Institute of Chemistry, University of Silesia in Katowice
Katowice, Poland

² Department of Organic Chemistry, Bioorganic Chemistry
and Biotechnology, Silesian University of Technology
Gliwice, Poland

Ionic liquids (ILs) are systems containing ionic liquids (ILs) and dispersed nanoparticles therein. Our studies proved that the addition of non-functionalized multiwalled carbon nanotubes (MWCNTs) to ILs leads to an increase in thermal conductivity and has no effect on isobaric heat capacity (JÓŹWIAK *et al.*, 2020). On the other hand, functionalized MWCNTs may lead to an increase in isobaric heat capacity.

Thus, the main object of this study is to explore the thermal properties of ILs composed of different ILs and different functionalized MWCNTs. Additionally, the isobaric heat capacity of multi-ionic systems, a new class of ILs, is investigated.

The study shows that the addition of functionalized MWCNTs may lead to a slight increase in isobaric heat capacity, but at the same time it has a slight increase or no effect on thermal conductivity and viscosity. Proposed multi-ionic systems are promising but need to be further investigated.

This work was financially supported by the National Science Centre (Poland) Grant No.2021/41/B/ST5/00892.

JÓŹWIAK B. *et al.* (2020), Remarkable thermal conductivity enhancement in carbon-based ionic liquids: Effect of nanoparticle morphology, *ACS Applied Materials & Interfaces*, **12**(34): 38113–38123, doi: 10.1021/acsami.0c09752.

* * *

Impact of wind turbines on people-field surveys and a series of laboratory experiments

FELCYN Jan¹, EMCHE Martyna¹,
BUSZKIEWICZ Maciej¹, PREIS Anna¹,
CHACIŃSKA Patrycja² (patrycja@ios.edu.pl)

¹ Acoustics Department

Adam Mickiewicz University
Poznań, Poland

² Institute of Environmental Protection

– National Research Institute
Warsaw, Poland

As part of the project “Healthy society-towards optimal management of wind turbines’ noise (HETMAN)”, a num-

ber of experiments are being carried out with the aim of comprehensively determining key parameters for wind turbine noise perception and investigating people’s reactions to this type of noise. Several studies are currently underway, the preliminary results of which will be briefly discussed. We will present how the annoyance rating of this type of noise compares to road noise. We will describe the similarities and differences in conducting studies in situ and in the laboratory. We will show whether it is possible to mask turbine noise and discuss whether infrasound noise from turbines is indeed a significant perceptual factor.

* * *

Ultrasound absorption and some rheological properties of 1-ethyl-3-methylimidazolium hydrogen sulfate ionic liquid

GANCARZ Paweł¹ (gancarz.pw@gmail.com),
ZOREBSKI Edward¹, SZCZĘCH Marcin²

¹ Institute of Chemistry

University of Silesia in Katowice
Katowice, Poland

² Faculty of Mechanical Engineering and Robotics
AGH University of Sciences and Technology
Kraków, Poland

Ionic liquids have gained increasing interest over a dozen last years, but only very little effort was spent in the study of ultrasound absorption in this still promising class of compounds. In this study, we present the results of ultrasound absorption measurements for 1-ethyl-3-methylimidazolium hydrogen sulfate. The ultrasound absorption in the frequency range (10 to 35) MHz and at temperatures from (293.15 to 323.15) K was measured by means of a standard pulse technique (first traveling pulse in the variable path length). The results reveal that the sample shows very high absorption, e.g., the frequency normalized absorption at 10 MHz and 293.15 K reaches $12500 \cdot 10^{-15} \text{ s}^2 \cdot \text{m}^{-1}$. Over the whole temperature and frequency measurement range, the relaxation of ultrasound absorption is clearly visible. Moreover, the negative temperature coefficients of normalized absorption are observed. The non-Newtonian properties (viscoelasticity) are detected as well.

* * *

Numerical analysis and new formulas of the response stage of a new SAW structure with disappeared layer RR-P3HT in detection DMMP

HEJCZYK Tomasz¹ (thejczyk@ente.com.pl),
WROTNIK Jarosław², POWROŹNIK Paulina²,
JAKUBIK Wiesław²

¹ Akademia Rozwoju Kreatywnego
Marklowice, Poland

² Institute of Physics
Silesian University of Technology
Gliwice, Poland

The paper presents the results of numerical analyses of the SAW gas sensor in the response state. The effect of

SAW velocity changes vs. the surface electrical conductivity of the sensing layer is predicted. The conductivity of the roughness sensing layer above the piezoelectric waveguide depends on the profile of the diffused gas molecule concentration inside the layer. Numerical results for the gas DMMP (CAS Number 756-79-6) for disappeared layer (RR)-P3HT in the response steady state have been shown. The main aim of the investigations was to study thin film interaction with target gases in the SAW sensor configuration based on diffusion equation for polymers and formulas based on signal theory – splice of the signal. Numerical results for profile concentration in response state have been shown and mathematical formulas. The results of numerical analyzes allow to select the sensor design conditions, including the morphology of the sensor layer, its thickness, operating temperature and layer type. The numerical results basing on the code written in Python, are described and analyzed. The theoretical results were verified and confirmed experimentally.

* * *

Development of an autonomous, modular transport vehicle for performing logistics operations – vibroacoustic and dynamic issues at the design stage

JAKUBOWSKI Piotr¹ (piotr.jakubowski@cto.gda.pl),
CISAK Konrad² (k.cisak@ekolaser.pl)

¹ Maritime Advanced Research Centre
Gdańsk, Poland

² Eko-Laser Sp. z o.o. Sp. k.
Bożepole Wielkie, Poland

The aim of the NCBiR-subsidized project is to develop an autonomous, modular transport vehicle for performing logistics operations in indoor and outdoor environments with a vision, laser and radio-based navigation system. The AGV (Automated Guided Vehicle), with the working name T1500, will be characterized by its ability to perform transport operations efficiently in varied environmental conditions. Unlike most standard AGVs, which are only prepared to operate indoors, the T1500 AGV will also be able to perform transport tasks outdoor in adverse weather conditions such as frost, snow, fog or light rain.

Advanced vehicle control algorithms include not only terrain-map-based route optimization and effective obstacle avoidance strategies, but also methods based on learning and selecting the most appropriate route on the basis of statistics on previously made journeys. Particular attention was paid to safety of the vehicle, its working environment and the transported goods. For this purpose, the type of goods to be transported is identified acoustically and dynamically and the control algorithm of the vehicle is adjusted correspondingly with the goods particularly difficult to be transported.

The article describes the results from load dynamics tests and the concept of implementing the recognition of dangerous situations during passage on the basis of an acoustic signal developed with the MATLAB program, with the use of neural networks.

* * *

Ultrasound study of magnetic pickering emulsions

JAMEEL Bassam (basjam@amu.edu.pl), BIELAS Rafał,
HORNOWSKI Tomasz, JÓZEFCAK Arkadiusz

Faculty of Physics, Adam Mickiewicz University
Poznań, Poland

Ultrasound spectroscopy provides interesting results useful in material science, especially when studying suspensions and emulsions. In this research, we used this technique to investigate the stability of magnetic Pickering emulsion, an emulsion stabilized by nanoparticles accumulated at the droplet surface as a shell form. The ultrasound scattering theory based on the core-shell model was applied to determine the ultrasound attenuation and interpret the data from experiments. The implemented model considers the thermal, density, and compressibility contrast between emulsion phases during ultrasound wave propagation. Ultrasound attenuation in the function of frequency was calculated for different core radii and shell thicknesses and fitted to the experimental attenuation spectra. The results show that ultrasound attenuation spectroscopy can be used to study the size of magnetic Pickering droplets and the thickness of the particle shell when different volume fractions of magnetic nanoparticles were used to stabilize the emulsion. The knowledge of shell stability is crucial from the perspective of the application of Pickering emulsions in industry and biomedicine; therefore, the non-destructive technique that can efficiently characterize such a complex, three-phase system is desirable.

This work was supported by project no. 2019/35/O/ST3/00503 (PRELUDIUM BIS) of the Polish National Science Centre.

* * *

Salty carbon nanotubes

JĘDRYSIAK Rafał G. (Rafal.Jedrysiak@polsl.pl),
BONCEL Sławomir, RUCZKA Szymon, STAŃCO Katarzyna

Department of Organic Chemistry, Bioorganic
Chemistry and Biotechnology
Silesian University of Technology
Gliwice, Poland

Synthesized by catalytic chemical vapor deposition (c-CVD), ultra-long, multi-walled carbon nanotubes can be chemically modified to enhance and target chemical and physical properties. It seems that systems consisting of cationically/anionically modified carbon nanotubes will significantly modify the conductivity of hybrid and composite systems. In this work, examples of the creation of such modified systems with the participation of carboxylated and aminated carbon nanotubes are presented.

* * *

Fabrication, properties and applications of 1D microparticle structures

KACZMAREK Katarzyna (kk80742@amu.edu.pl)

Chair of Acoustics, Faculty of Physics
Adam Mickiewicz University
Poznań, Poland

1D micro-particle structures have been extensively studied over the past few decades. The selective distribu-

tion of microparticles in a form of predefined patterns found interest among others in the fields of electronics, micro-robotics, and biosensors.

Depending on the application purpose, complex morphologies can be created by the assembly of microparticles with different shapes i.e., spheres, Janus dimers, polyhedrons, rods, and ellipsoids. To assemble, such various building blocks into 1D structures, electric, magnetic, acoustic, and optic field-assisted assembly methods are commonly used. Our group led by prof. Rozynek recently developed a simple and efficient approach that combines electric and capillary interactions. The novelty of our method is the ability to efficiently assemble long 1D structures outside a liquid environment.

The most interesting assembly methods, properties, and applications of 1D microstructures presented over the years including our approach, colloidal caterpillars for cargo transportation, bending responsive hydrogels, colloidal origami, and microrobot assemblies will be discussed.

This work was supported by the Polish National Science Centre through the OPUS17 programme (2019/33/B/ST5/00935).

* * *

Contrast-enhanced magneto-motive ultrasound for sentinel lymph node identification

KACZMAREK Katarzyna^{1,2} (kk80742@amu.edu.pl), SJÖSTRAND S.², BACOU M.³, THOMSON A.⁴, JANSEN T.⁵, MOUG S.⁶, FARRINGTON S.³, MORAN C.M.⁴, MULVANA H.²

¹ Chair of Acoustics, Faculty of Physics
Adam Mickiewicz University

Poznań, Poland

² Department of Biomedical Engineering
University of Strathclyde
Glasgow, UK

³ Institute of Genetics and Cancer, University of Edinburgh
Edinburgh, UK

⁴ Centre for Cardiovascular Science
University of Edinburgh
Edinburgh, UK

⁵ Department of Clinical Sciences Biomedical Engineering
Lund University
Lund, Sweden

⁶ Department of Surgery, Royal Alexandra Hospital
Scotland, UK

Identification of cancerous lymph nodes is crucial for cancer staging. The absence of metastases in the sentinel node is a good prognostic factor, as it determines the lower probability of metastasis in other draining nodes. Localization of sentinel lymph nodes is not an easy task. Commonly used in clinical practice methods, such as blue dye staining or isotope staining, are lacking accuracy.

As an alternative tool for lymph node identification Contrast-Enhanced Magneto-Motive Ultrasound (CE-MMUS) has been proposed. CE-MMUS uses magnetic nanoparticles and microbubbles as contrast agents, and a low-frequency alternating magnetic field to induce oscillations of magnetic microbubbles to generate tissue-laden movement.

Results showed that magnetic microbubbles can be successfully used as bimodal contrast agents. Microbubble accumulation was verified using contrast-enhanced ultra-

sound, and their presence visibly increased the echogenicity of the sentinel lymph node. Tissue-laden movement was tracked and filtered out with a phase and frequency tracking algorithm. Mean tissue displacement caused by microbubbles increased compared to displacement caused by magnetic nanoparticles only. The studies were conducted on mice models in vivo.

This work was supported by CRUK-grant numbers A23333/24730.

* * *

Urban green infrastructure as a moderator of physical environmental factors – an acoustic-electromagnetic example of criteria for multi-factor assessment

KARPOWICZ Jolanta (jokar@ciop.pl),
MORZYŃSKI Leszek, PODLEŚNA Marlena,
PLEBAN Dariusz

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

Urban green infrastructure can, among other things, act as a moderator of various parameters of physical environmental factors (as their source or as a barrier for propagation). For this reason, a properly selected structure of green infrastructure can be used to create micro-environments in the work or non-professional life environments that are conducive to recreation during the working day or in leisure time. Evaluation of the quality of such micro-environments requires appropriate research methods and criteria for multi-factor assessment, regarding physical environmental factors which are relevant for health, well-being and efficiency at work.

The core elements of multi-factor assessing the barrier capacity of green urban infrastructure will be presented, based on the preliminary results of pilot study, for acoustic and electromagnetic factors (taking into account their most common urban sources and their relevant parameters recognized in Warsaw – traffic noise and radiocommunication networks).

National Program “Governmental Program for Improvement of Safety and Working Conditions”; supported from the resources of the National Centre for Research and Development; research task IV.PN.05.

* * *

Magnetic nanoparticles from nature – physical properties and possible applications in biomedicine and biotechnology

KOPČANSKÝ Peter (kopcan@saske.sk)

Institute of Experimental Physics
Slovak Academy of Sciences
Košice, Slovakia

Magnetic fluids as one of the pioneers of modern nanotechnologies are still attractive for basic as well as applied research. Magnetic fluids have prompted further research, especially in the field where nanoparticles, especially magnetic ones, give them the ability to manipulate them in an external magnetic field thus change the properties required in various applications. In recent years, development

of magnetic nanoparticles has been in the forefront of material research and development. Various methods have been developed to reach a better quality of samples for further application in technology. Almost all living systems, including bacteria, plants, animals, and humans, can create iron-based nanoparticles coated with apoferritin. Such stored iron forms ferritin, 10–12 nm metalloprotein magnetic particles. Magnetotactic bacteria create magnetosomes for example. Such bioinspired magnetic particles can be used in all applications that provide magnetic fluids, but they are unique in that they were created by nature or were inspired by nature. Due to their biocompatibility, they are much better usable in biomedical and biotechnological applications than artificially prepared magnetic particles. In lecture will be presented new results associated with these nanoparticles such as their structural and magnetic characterization as well as examples for biomedical, diagnostic or environmental engineering.

* * *

Measurement and analysis of wind turbine noise

KOZIOL Michał (m.koziol@po.edu.pl),
WIECZOREK Roman, BOCZAR Tomasz, CICHON Andrzej,
ZMARZLY Dariusz, NAGI Łukasz

Faculty of Electrical Engineering, Automatic Control
and Informatics
Opole University of Technology
Opole, Poland

The subject of the research concerns the analysis of noise emitted by a working wind turbine in field conditions. The studies carried out were intended to indicate the potential frequency range of emitted noise in the low frequency range and in the range heard by humans. Acoustic signals generated by the wind turbine were recorded using three independent measuring tracks, which were synchronized with each other. The obtained results will be used to determine the indicators for the studied frequency ranges, which will be the subject of further research work.

* * *

Innovative solutions of people and the environment protection against rail traffic noise

KSIĄŻKA Piotr (piotr.ksiazka@ios.edu.pl),
CHACIŃSKA Patrycja

Institute of Environmental Protection
– National Research Institute
Warsaw, Poland

The increase in the speed of rail vehicles and the increased volume of train traffic causes increased noise emissions, which can have a negative impact on people and the environment. To minimize this negative impact, various methods of reducing noise emissions into the environment are routinely used. However, traditional methods are not always applicable or bring satisfactory results, e.g., noise barrier. An alternative to them may be innovative solutions, such as dampers of the rails and track suppressors, which was the subject of the research grant “Innovative solutions to protect people and the environment against rail traffic noise – InRaNoS” The main goal of this project was

to develop guidelines for the use of noise mitigating devices on railway lines. The guidelines was based on tests results for two types of devices (dampers and track suppressors).

* * *

Detection of drone eviction from the hive based on audio signal

LIBAL Urszula,
BIERNACKI Paweł (pawel.biernacki@pwr.edu.pl)

Department of Acoustics, Multimedia
and Signal Processing
Wrocław University of Science and Technology
Wrocław, Poland

Eviction of drones from the hive in summer, may be the first sign of swarming. This is a potentially dangerous situation, as the queen and the entire hive may escape. To prevent this, we propose an early swarming detection system based on the behaviour of bees near the hive entrance. The system analyzes sound signals recorded at the entrance of the hive by an autoencoder neural network to detect ejected drones. Simulations using real signals have shown that it is possible to effectively detect drones ejected from the hive. The achieved detection probability of 86% makes it possible to create an effective alarm system for beekeepers.

* * *

Acoustic noise in hatching plants as element of environment and its influence on the animal welfare

LISOWSKA-LIS Agnieszka¹ (lisowskalis@anstar.edu.pl),
WIELGAT Robert¹, LIS Marcin W.²

¹ University of Applied Sciences in Tarnow
Tarnów, Poland

² University of Agriculture in Krakow
Kraków, Poland

Acoustic communication of chicks and ducks inside the egg before hatching is an important element of hatch synchronization in precocial species. Noise of various origins is characteristic of the artificial environment. This can negatively affect hatching. Measurements and registrations were carried out in the large poultry hatchery for chicks and ducks: setters, hatchers. Some noises of working machinery and animals were registered in: technological halls, corridors, storerooms, technical rooms. The registrations were analysed in Cool Edit Pro 2.1 tool in the wave form for the sound and noise presented in the signal for every measured points (FFT analysis were done).

* * *

Design of a matrix sound source with controlled directivity characteristics

MAKAREWICZ Grzegorz,
MORZYŃSKI Leszek (lmorzyns@ciop.pl)

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

Sound messages, including voice messages, play an important role in the work environment and in the living en-

vironment. With their help, information that is important at a given moment, in particular warnings about dangers, is provided to interested persons. For bystanders, however, a sound message can be an annoying noise, therefore it is recommended to construct electroacoustic systems for transmitting sound messages in such a way that the signal emitted by them reaches mainly interested people and minimizes noise pollution of the environment. One of the possible solutions is the use of sound sources with controlled directional characteristics. The paper presents the design of a sound source in the form of a matrix of dynamic loudspeakers with controlled directivity characteristics. The results of numerical simulations related to controlling the directionality characteristics of sound sources and related acoustic field distributions are presented. Simulation calculations were carried out using dedicated software developed for research purposes. Possibilities of controlling the directivity characteristic of the matrix source were discussed, and then the design of the electronic control system was proposed.

* * *

Monitoring and control of vibroacoustic hazards in the work environment using the Internet of Things – review of issues and system proposal

MORZYŃSKI Leszek, PODLEŚNA Marlena (mapod@ciop.pl), SZCZEPAŃSKI Grzegorz

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

The work environment is an environment in which factors harmful to the health of employees may occur. A physical factor's potential harmfulness to health on the intensity of a given factor in the work environment and on the exposure time of the employee. The intensity of the harmful factor can change over time, sometimes in a way that is difficult to predict, causing increased risks to the employees. Detection of the threat and quick and effective response aimed at limiting its impact on the employee's health is possible in such cases, assuming continuous monitoring of the parameters of the working environment. The paper discusses issues concerning wireless sensor networks and the Internet of Things used for monitoring and controlling environmental parameters. A proposal for the structure of the system for monitoring and remote control of vibroacoustic hazards in the work environment has been presented. The technical possibilities of implementing the components of the system were discussed.

* * *

Development of the SAFE website in the context of changes in the Polish labour market on the example of noise hazards

MORZYŃSKI Leszek (lmorzyns@ciop.pl), WŁUDARCZYK Anna, ŁADA Krzysztof

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

The SAFER website developed at CIOP-PIB is a rich source of knowledge about harmful physical factors of the

work environment. This website consists of content related materials, educational materials and small internet tools aimed at supporting the assessment of exposure and occupational risk, for each of the factors described on the website. To maintain high usability, the website must be constantly updated and expanded with new knowledge. The article discusses the content of the website and presents its current development on the example of the section devoted to noise in the work environment. The discussed development of the website takes into account changes on the Polish labor market related in particular to economic immigration.

* * *

Differences in acoustic features between the standard and the new variant of Polish /ɛ/

NAWROCKI Grzegorz,
WIELGAT Robert (rwielgat@poczta.onet.pl)

Polytechnic Department
University of Applied Sciences in Tarnow
Tarnów, Poland

The paper presents an examination of differences in selected acoustic features between two variants of the Polish alveolo-palatal voiceless sibilant, the standard realization of /ɛ/, and its fronted palatalized version [sj]. Four spectral parameters, namely standard deviation, skewness, kurtosis, and center of gravity (COG) are analyzed in search of significant and distinctive spectral features of the tested fricative segments. Differentiation in the use of the two variants among the speakers and their occurrence depending on the preceding vocalic contexts are focused on, as well. Statistical tests reveal significant differences between the two variants of the Polish phoneme in COG for the middle portion of the segment located after three non-high vowels: /ɛ, a, ɔ/.

* * *

Prediction of vibration transmission through technical connectors of various complexity

NIERADKA Paweł¹ (p.nieradka@kfb-acoustics.com), DOBRUCKI Andrzej²

¹ KFB Acoustics
Wrocław, Poland

² Department of Acoustics
Wrocław University of Science and Technology
Wrocław, Poland

The paper compares the results of simulations and measurements of vibration transmission through selected technical connectors. The tested joints were characterized by various degrees of complexity. The increasing advantage of experimental methods over simulation methods, when the degree of complexity of connections increases, has been indicated. Continuous joints (welded) and point joints (welded, riveted, bolted) were considered. During the experiments, the PIM (Power Injection Method) method was used to determine the CLF (Coupling Loss Factor) coefficients describing the vibration transmission. The results of the CLF simulations came from the developed FEM (Finite Element Method) models and from the theoretical dependencies derived from the wave theory. A good agreement between the simulations and measurements was ob-

tained for the welded joint, while a significant discrepancy occurred for point joints. Problematic features of point joints causing difficulties in developing their exact numerical model were indicated.

* * *

New trends in housing construction – a difficult challenge for building acoustics

NURZYŃSKI Jacek (j.nurzynski@itb.pl)

Building Research Institute – ITB
Warsaw, Poland

New trends in residential construction direct its further development towards a lightweight prefabricated structure that is thermally effective and meets the criteria of sustainable development. Acoustic properties of this type of objects are a big challenge. Due to the low mass of building partitions, it is necessary to use specific structural details and additional insulating layers limiting the transmission of airborne and impact sounds. The paper presents an analysis of the impact of these solutions on the acoustic parameters of walls and floors, carried out on the basis of laboratory test results. The main objective of the work was to assess the effectiveness of various types of cladding and structural dividers in relation to lightweight partitions with a frame structure and floors consisting of composite panels. Additional suspended ceilings and floor coverings as well as thermal insulation of external walls were considered. Typical floating floor systems gave a weaker acoustic effect than when they are used on traditional massive floors. In the case of wall claddings and suspended ceilings, the use of elastic connectors and flexible suspenders limiting the transmission of structural sound was of great importance. The individual layers of the partition mutually interacted, so their effectiveness was not subject to simple summation. The presented analysis, results and conclusions may be useful in the design of lightweight prefabricated buildings and in the development of construction details that increase the acoustic efficiency of new solutions.

* * *

Refraction analysis of divergent ultrasound beam transmitted through a coronal section of a female breast immersed in water

OPIEŁIŃSKI Krzysztof J. (krzysztof.opielinski@pwr.edu.pl),
BUŁKOWSKI Mariusz, GABRYEL Andrzej,
WIKTOROWICZ Andrzej

Department of Acoustics, Multimedia
and Signal Processing
Wrocław University of Science and Technology
Wrocław, Poland

The analysis of the phenomenon of ultrasound beam refraction is important in ultrasound imaging in developing methods to reduce the effects of this phenomenon. The propagation paths of ultrasound beam rays in soft tissue are not rectilinear, leading to distortion of the imaged structures. This is particularly important for imaging the breast structure by quantitative ultrasound transmission tomography, as the pixels in the image present absolute local values of propagation speed and attenuation, which are the basis

for diagnosing pathological changes. The phenomenon of refraction affects the deformation of visualized lesions and falsifies the distribution of the values of the acoustic parameters. The study modeled the refraction phenomenon of divergent ultrasound beam transmitted through a circular section of breast glandular tissue surrounded by subcutaneous fat, skin, and water. The developed model makes it possible to change the size of the section and layers, the ultrasound velocity in each medium, and the temperature of the water. With the help of the developed model, different scenarios of the refraction phenomenon were analyzed for the possibility of its reduction. Control tests of the same female breast in vivo at different water temperatures were also performed using the ultrasound tomography breast scanner prototype.

* * *

Wind turbine noise annoyance due to the possibility of the employee carrying out works requiring attention

PLEBAN Dariusz (daple@ciop.pl), SZCZEPAŃSKI Grzegorz,
RADOSZ Jan, KAPICA Łukasz, ALIKOWSKI Adrian,
ŁADA Krzysztof, WŁUDARCZYK Anna

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

The constant growth of energy demand, as well as the accompanying increase in environmental pollution resulting from the prevailing use of fossil fuels, has led to a rising use of energy from renewable sources. According to the data of the Statistics Poland of 2021, energy from renewable sources in Poland in 2020 mainly came from solid biofuels (71.61%), wind energy (10.85%) and liquid biofuels (7.79%). The use of wind turbines to generate electricity has many obvious advantages, such as lack of fuel costs during operation and lack of harmful pollutants, including CO₂. Despite advantages, the use of wind energy (wind turbines) constantly raises questions and concerns. The questions concerning the impact of wind farms on humans still remain valid. This impact includes many factors related to the operation of wind farms, and in particular noise emitted by these farms. The wind turbine noise impact on humans has been studied by the Central Institute for Labour Protection – National Research Institute, but unlike other studies it focused on the impact of this noise as an annoyance factor affecting the employees' capacity to perform their basic tasks. Taking into account the obtained results, it was concluded that wind turbine noise with an equivalent A-weighted sound pressure level of at least 50 dBA should be considered as a noise nuisance due to the possibility of the employee carrying out works requiring attention.

* * *

Effect of ultrasonic noise on humans

RADOSZ Jan (jarad@ciop.pl)

Central Institute for Labour Protection
– National Research Institute
Warsaw, Poland

Ultrasonic noise can be harmful to the hearing organ, as well as negatively affect the vestibular system, which may

be manifested by headaches, dizziness and balance disorders. The last revision of the limit values for ultrasonic noise took place in 2001 and was based, among others, on based on the WHO guidelines from 1982 and the IRPA from 1984. The limit values adopted in Poland and in other countries are largely based on the results of research from the 70s and 80s. It should be emphasized that many of these scientific reports have not been confirmed by other researchers. The paper presents current literature reports on the impact of ultrasonic noise on humans in the context of the applicable limit values.

* * *

From magnetic adsorbents to magnetically responsive nanozymes

SAFARIK Ivo^{1,2} (ivosaf@yahoo.com), PROCHAZKOVA Jitka¹, BALDIKOVA Eva¹, POSPISKOVA Kristyna Zelena², TIMKO Milan³, KOPČANSKÝ Peter³

¹ Department of Nanobiotechnology, Biology Centre, ISB Czech Academy of Sciences

Ceske Budejovice, Czech Republic

² RCPTM-CATRIN, Palacky University

Olomouc, Czech Republic

³ Institute of Experimental Physics

Slovak Academy of Sciences

Košice, Slovakia

Currently there is a high interest in the study and potential applications of magnetically responsive adsorbents for the removal of wide variety of inorganic and organic pollutants. Special attention is paid to low-cost (bio)sorbents including food and agricultural wastes, microbial cells, macroalgae, biochar or clays. Simple and inexpensive magnetic modification procedures, employing, e.g., magnetic fluid treatment, mechanochemical procedure or microwave assisted synthesis have been developed. In most cases nano- and microparticles of ferrimagnetic iron oxides (magnetite, maghemite or different types of ferrites) are used for (bio)sorbent modification. Both magnetite and maghemite exhibit peroxidase-like activity and belong to a wide group of nanozymes. Magnetically modified materials (adsorbents) can thus exhibit dual activity, where both adsorption and enzyme-like degradation can be applied for specific pollutants removal. Various types of magnetically modified materials developed by the authors will be presented in the lecture, together with their characterization and potential applications.

* * *

New two-step non-ultrasonication method of preparation of ionanofluids

SHELLER Łukasz¹, GREER Heather F.², JÓŹWIAK Bertrand³, DZIDO Grzegorz³, DZIADOSZ Justyna¹, KOLANOWSKA Anna³, JĘDRYSIAK Rafał³, BONCEL Sławomir³, DZIDA Marzena¹

¹ Institute of Chemistry, University of Silesia in Katowice Katowice, Poland

² Faculty of Chemistry, Silesian University of Technology Gliwice, Poland

³ Department of Chemistry, University of Cambridge Cambridge, UK

Multi-walled carbon nanotubes (MWCNTs) dispersed in ionic liquids (ILs), also known as ionanofluids (INFs), have greatly thermal conductivity (TC) increased compared to the base IL. Thus, INFs can be considered as new heat-transfer fluids (HTFs). INFs containing long MWCNTs have higher thermal conductivity than INFs with short MWCNTs (DZIDA *et al.*, 2020; JÓŹWIAK *et al.*, 2020) Only the long MWCNTs create subzipped networks and connect in multiple bridges (DZIDA *et al.*, 2020) which suggests that the length of MWCNTs plays a key role in the TC enhancement. However, the ultrasound sonication applied during INFs' preparation shortens the MWCNTs (JÓŹWIAK *et al.*, 2021). In order to preserve the MWCNTs' length we developed a new, repeatable method of preparation of INFs without applying the ultrasonication. This method allowed to obtain long-term stable INFs with the TC increase of 60% for 1.0% by weight of long in-house MWCNTs dispersed in 1-ethyl-3-methylimidazolium thiocyanate at 298.15 K which is larger by 20% than for the same INF prepared using ultrasonication method (JÓŹWIAK *et al.*, 2020). On the other hand, the viscosity is also higher which can be a disadvantage when designing an INF with optimal both thermal and rheological properties.

This work was financially supported by the National Science Center (Poland), grant no. 2021/41/B/ST5/00892.

DZIDA M. *et al.* (2022), High-performance ionanofluids from subzipped carbon nanotube Networks, *ACS Applied Materials & Interfaces*, **14**(45): 50836–50848, doi: [10.1021/acsami.2c14057](https://doi.org/10.1021/acsami.2c14057).

JÓŹWIAK B. *et al.* (2020), Remarkable thermal conductivity enhancement in carbon-based ionanofluids: effect of nanoparticle morphology, *ACS Applied Materials & Interfaces*, **12**(34): 38113–38123, doi: [10.1021/acsami.0c09752](https://doi.org/10.1021/acsami.0c09752).

JÓŹWIAK B. *et al.* (2021), Effect of ultrasonication time on microstructure, thermal conductivity, and viscosity of ionanofluids with originally ultra-long multi-walled carbon nanotubes, *Ultrasonics Sonochemistry*, **77**: 105681, doi: [10.1016/j.ultsonch.2021.105681](https://doi.org/10.1016/j.ultsonch.2021.105681).

* * *

The impact of alternating and rotating regimes on the heating characteristics of magnetic fluid

TIMKO Milan¹ (Timko@saske.sk), MOLCAN Matus¹, SKUMIEL Andrzej², KOPČANSKÝ Peter¹, SAFARIK Ivo³

¹ Institute of Experimental Physics

Slovak Academy of Sciences

Košice, Slovakia

² Faculty of Physics, Adam Mickiewicz University

Poznań, Poland

³ Department of Nanobiotechnology, Biology Centre

Ceske Budejovice, Czech Republic

Magnetic nanoparticles can be used in numerous therapeutic procedures, including controlled drug delivery and release, thermal therapy, and photo- and sonodynamic therapy. However, such approaches remain limited due to difficulties associated with accurately controlling their therapeutic efficiency and localized consequences. In this contribution, magnetite nanoparticles of various configurations (single, chains from bacteria, halloysite chains, aggregated in bacterial cellulose) are measured in hyperthermia experiments. The choice of a suitable set up for generating of

magnetic field can significantly affect the resulting thermal effect and thus the efficiency itself. The technical details of the generating RMF as a new one compared to traditional AC magnetic field will be presented. Comparison of specific absorption rate (SAR) values for various magnetically modified biocarriers can be one of the parameters for the selection of optimal materials for potential hyperthermia applications.

* * *

Calibration of vibration measuring instruments using comparative methods in Central Office of Measures

TWOROGOWSKI Piotr (piotr.tworogowski@gum.gov.pl)

Laboratory of Mechanical Vibrations
Central Office of Measures
Warsaw, Poland

Laboratory of Mechanical Vibrations in Central Office of Measures has an ability to calibrate vibration measuring instruments using laser interferometry method (on national measuring stand) and also using secondary calibration method according to *ISO 16063-21: 2003 Methods for the calibration of vibration and shock transducers – Part 21: Vibration calibration by comparison to a reference transducer*, which is carried out on independent measuring stand.

Purchasing in 2003 another measuring system for calibration using laser interferometry method (mentioned above) made it possible to expand calibrating possibilities of the Laboratory especially in range of obtained measurement uncertainties. Personnel changes in recent years in Laboratory caused interest in other possibilities of this measuring system. As a result, there is a possibility to calibrate vibration instruments using secondary method also on this measuring stand and there is also possibility to carry out calibration using SWEEP signal. This paper shows actual measuring systems of secondary method working in Laboratory of Mechanical Vibrations, presents the results of conducted measurements of vibration transducer with an interpretation using two different types of exciters

(two measuring stands) and at the end, presents calibration results using SWEEP signal.

* * *

Partial discharges research by acoustic emission method and by electric and acoustic emission methods used in parallel

WITOS Franciszek (franciszek.witos@polsl.pl),
OLSZEWSKA Aneta

Department of Optoelectronics
Silesian University of Technology
Gliwice, Poland

Original multi-channel author's measurement systems designed and built to conduct research on PDs phenomena using the AE method are presented. The systems provide real-time monitoring and recording of signals as well as analysis of recorded signals. The analysis of the recorded AE signals is carried out in the following domains: time, frequency, time-frequency and discrimination threshold. In particular, in the domain of the discrimination threshold, the AE descriptor has been defined with the ADC acronym, which ranks the signals according to the so-called degree of advancement. The conducted parallel tests carried out by the electrical method and the AE method showed that, for a single PD source, the ordering of signals using the ADC descriptor is identical to the ordering of sources according to the value of the introduced apparent charge.

The paper presents the author's method of locating and identifying PDs in selected real objects: coil rods of hydro-generators, power oil transformers. The results of PDs tests of power oil transformer at the test station, conducted in parallel with the electric method and author's AE method, and the results of PDs tests in 3 selected transformers carried out, during ongoing in-situ operation, using the AE method are presented. Based on the results of these tests, the authors made diagnoses of the condition of the insulation systems in the tested transformers. The inspections of these transformers confirmed the diagnoses.

* * *

AIMS AND SCOPE

Archives of Acoustics was founded in 1976 and since then publishes original research papers from all areas of acoustics including:

- acoustical measurements and instrumentation,
- musical acoustics,
- acousto-optics,
- architectural, building, and environmental acoustics,
- bioacoustics,
- electroacoustics,
- linear and nonlinear acoustics,
- noise and vibration,
- psychoacoustics,
- physical and chemical effects of sound,
- physiological acoustics,
- speech processing and communication systems,
- speech production and perception,
- sonochemistry,
- transducers,
- ultrasonics,
- underwater acoustics,
- quantum acoustics.

INFORMATION FOR AUTHORS

- Manuscripts intended for publication in *Archives of Acoustics* should be submitted in pdf format by an on-line procedure using **New Submission** link in **User Home** menu.
- Manuscript should be original and should not be submitted either previously or simultaneously elsewhere, neither in whole, nor in part.
- Submitted papers must be written in good English and proofread by a native speaker.
- In principle the papers should not exceed 40 000 typographic signs (about 20 typewritten pages).
- A brief abstract must be included. The keywords and subject classification for the article are required.
- Postal addresses, affiliations and email addresses for each author are required.
- The submitted manuscript should be accompanied by a cover letter containing the following information:
 - why the paper is submitted to *Archives of Acoustics*,
 - suggestion on the field of acoustics related to the topic of the submitted paper,
 - the statement that the manuscript is original, the submission has not been previously published, nor was it sent to another journal for consideration,
 - 3–5 names of suggested reviewers together with their affiliations, full postal and e-mail addresses; at least 3 suggested reviewers should be affiliated with other scientific institutions than the affiliations of the authors,
 - author's suggestion to classification of the paper as the research paper, review paper or technical note.
- Information regarding the preparation of manuscripts is provided on the journal's website <https://acoustics.ippt.pan.pl>.
- Reviews of manuscript is available online.
- Full information about the journal are available at <https://acoustics.ippt.pan.pl>.

Each manuscript is initially evaluated by the Editor and then allocated to an Section Editor according to specific subject area. Upon receiving the Section Editor sends it to reviewers. The Editor and the Section Editors make the final recommendation. Their opinions are crucial for the Editorial Committee to decide whether the paper can be accepted for publication or not.

The Word or L^AT_EX source and pdf files of the final version of articles will be requested after acceptance for publication. High-quality illustrations (format bitmap (gif, jpg) or PostScript, with resolution no lower than 300 dpi) should be sent along with the manuscript.

Editorial Board
Archives of Acoustics