POLISH SOCIETY OF THEORETICAL AND APPLIED MECHANICS

JOURNAL OF THEORETICAL AND APPLIED MECHANICS

No. 1 • Vol. 63
Quarterly
eISSN: 2543-6308
ISSN: 1429-2955

WARSAW, February 2025

JOURNAL OF THEORETICAL AND APPLIED MECHANICS
(until 1997 Mechanika Teoretyczna i Stosowana, ISSN 0079-3701)

Beginning with Vol. 45, No. 1, 2007, *Journal of Theoretical and Applied Mechanics* (JTAM) has been selected for coverage in Thomson Reuters products and custom information services. Now it is indexed and abstracted in the following:

- **Science Citation Index Expanded** (also known as SciSearch®)
- **Journal Citation Reports/Science Edition**

Advisory Board

Michał Kleiber (Poland) – Chairman
Jorge A.C. Ambrosió (Portugal) ✶ Angel Baltov (Bulgaria)
✶ Romesh C. Batra (USA) ✶ Alain Combescure (France)
✶ Jüri Engelbrecht (Estonia) ✶ Józef Kubik (Poland)
✶ Włodzimierz Kurnik (Poland) ✶ Zenon Mróz (Poland)
✶ Wiesław Nagórko (Poland) ✶ Ryszard Parkitny (Poland)
✶ Ekkehard Ramm (Germany) ✶ Meir Shillor (USA)
✶ Andrzej Styczek (Poland) ✶ Eugeniusz Świtoński (Poland)
✶ Hisaaki Tobushi (Japan) ✶ Andrzej Tylikowski (Poland)
✶ Dieter Weichert (Germany) ✶ Jose E. Wesfreid (France)
✶ Joseph Zarka (France) ✶ Vladimir Zeman (Czech Republic)

Editorial Board

Editor-in-Chief – **Piotr Kowalczyk**
Section Editors: Iwona Adamiec-Wójcik, Piotr Cupiał, Krzysztof Dems,
Witold Elsner, Eric Florentin (France), Elżbieta Jarzębowska,
Oleksandr Jewtuszenko, Zbigniew Kowalewski, Tomasz Krzyżyński,
Anna Kucaba-Piętal, Stanisław Kukla, Tomasz Łodygowski,
Ewa Majchrzak, Janusz Narkiewicz, Michał Nowak, Piotr Przybyłowicz,
Błażej Skoczeń, Jacek Szumbarski, Krzysztof Tajduś,
Utz von Wagner (Germany), Jerzy Warmiński
Language Editor – Piotr Przybyłowicz
Technical Editor – Ewa Koisar
Secretary – Urszula Kowalczyk

Articles in JTAM are published under Creative Commons Attribution 4.0 International. Unported License https://creativecommons.org/licenses/by/4.0/deed.en. By submitting an article for publication, the authors consent to the grant of the said license.

The journal content is indexed in Similarity Check, the Crossref initiative to prevent plagiarism.

* * * * *

Editorial Office

Al. Armii Ludowej 16, room 650; 00-637 Warsaw, Poland
phone (+48) 664 099 345, e-mail: biuro@ptmts.org.pl

www.jtam.pl

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 3-12, Warsaw 2025
https://doi.org/10.15632/jtam-pl/194340

INVESTIGATION OF OUT-OF-PLANE COMPRESSION MECHANICAL PROPERTY FOR NOVEL HIERARCHICAL REENTRANT HONEYCOMB STRUCTURES

Dongquan Wu, Zhenyi Xu, Zhuo Yuan, Yupeng Li
Sino-European Institute of Aviation Engineering, Civil Aviation University of China, Tianjin, China
corresponding author Dongquan Wu, e-mail: dqwu@cauc.edu.cn
Zhiqiang Zhang
Institute of Aviation Engineering, Civil Aviation University of China, Tianjin, China

In this study, three novel types of hierarchical reentrant honeycomb structures were designed, which comprised subunits with a semi-reentrant, reentrant, and hexagonal honeycomb. These subunits exhibited zero/negative/positive Poisson's ratios, respectively. Geometric relationships of those structures were established. The out of plane compression behavior and deformation characteristics were investigated through finite element simulations and experiments. It was shown that the hierarchical structure composed of reentrant subunits had the best mechanical properties among three structures. The hierarchical structure composed of the classic hexagonal honeycomb subunit exhibited plastic deformation characteristic spreading from the middle and upper layers to the whole region during compression.

Keywords: hierarchical reentrant structure, honeycomb structures, out of plane, 3D printing, compression performance

1. Introduction

Honeycomb structures frequently appear in porous materials and are commonly regarded as effective energy-absorbing materials in many critical engineering applications. In the past research, various types of honeycombs such as hexagonal (Li *et al.*, 2019; Liu *et al.*, 2021; Ahmed and Xue, 2019), Kagome (Zhang and Zhang, 2013), circle (Ahmed and Xue, 2019) and Nomex honeycomb (Liu *et al.*, 2015) have been proposed and extensively studied according to specific application scenarios. Scholar Jamal *et al.* performed experimental and numerical modelling investigations of a typical four-point bending test of single and multi-layer honeycomb sandwich structures (Arbaoui *et al.*, 2014). Due to their high out-of-plane stiffness, low mass, and excellent mechanical properties, honeycomb structures have found wide application in traditional architectural design (Lei *et al.*, 2022; Davalos *et al.*, 2001), various airfoil aircraft(Gong *et al.*, 2022; Heo *et al.*, 2013; Solak *et al.*, 2023), and fields of nanotechnology and biomedicine Gao *et al.*, 2022; Gadkaree, 1998; Masuda and Fukuda, 1995).

To further optimize the out-of-plane mechanical performance of the traditional hexagonal honeycomb structure, researchers restructured the hexagonal honeycomb into a reentrant honeycomb with negative Poisson's ratio characteristics by concaving the cell walls on both sides. Different from hexagonal honeycomb structures, auxetic structures with negative Poisson's ratios, such as reentrant structures, exhibit synclastic curvature and many superior mechanical properties (Wu *et al.*, 2023). This structure is characterized by lateral contraction when compressed longitudinally. Due to its superior crashworthiness, it has garnered significant attention in recent years. Studies have shown that reentrant honeycombs outperform traditional honeycombs in terms of structural strength, stiffness, and energy absorption capacity (Garg *et al.*, 2023; Günaydin *et al.*, 2022; Ha *et al.*, 2022; Teng *et al.*, 2022; Hamzehei *et al.*, 2022).

However, all the aforementioned design concepts are limited to honeycomb structures with single-level cellular configurations, which have the disadvantage of low structural stiffness, restricting their potential applications. In this study, to further enhance their out-of-plane mechanical properties, hierarchical structures inspired by nature are introduced (Wang *et al.*, 2021; Yang *et al.*, 2021), such as hierarchical structures of osteons and lamellae with their length scale varying from macroscopic to microscopic. In a hierarchical structure, each level of the honeycomb structure can bear part of the load, thereby increasing the overall strength and stiffness of the structure. Different from traditional honeycomb structures, the novel self-similar hierarchical honeycomb (CVH) proposed by Liang *et al.* (2021) exhibits high energy absorption capacity and stable deformation characteristics under quasi-static compression. Unlike the traditional five-variable shear deformation theory, the four-variable hierarchical plate shear theory proposed by Bouazza *et al.* (2019) based on the principle of virtual work no longer requires shear correction factors and has been shown to have similarities with classical plate theories in many aspects. Ha *et al.* (2021) investigated energy absorption characteristics of bionic hierarchical multi-cell square tubes under axial compression, while Wang *et al.* (2021) examined compressive behavior of reinforced hierarchical lattice structures through both experiments and simulations. These studies demonstrate that hierarchical structures significantly improve mechanical performance and energy absorption capacity.

Previous studies have primarily focused on mechanical properties of honeycomb structures with negative Poisson's ratios and hierarchical structures, but those studies mostly concentrated on single-type honeycomb structures. This paper innovatively integrates nested honeycomb subunits with negative, zero, and positive Poisson's ratios, proposing a novel hierarchical reentrant honeycomb structure with superior mechanical performance, filling the gap in current research. Through a combination of finite element simulations and 3D printing experiments, the mechanical properties and deformation mechanisms of the proposed structure under out-of-plane compression were systematically studied. Through optimized design, the proposed hierarchical reentrant honeycomb structure not only enhances the mechanical properties of the material but also exhibits significant energy absorption and deformation control capabilities, making it highly relevant for practical applications. The results reveal the potential applications of this novel structure in the aerospace and civil engineering fields, providing new insights for the design of high-strength, lightweight structures.

Nomenclature of abbreviation

SEA – specific energy absorption,
RH – novel hierarchical reentrant honeycomb structures with reentrant honeycomb subunit,
SRH – novel hierarchical reentrant honeycomb structures with semi-reentrant honeycomb subunit
CH – novel hierarchical reentrant honeycomb structures with classic hexagonal honeycomb subunit.

2. Novel hierarchical reentrant honeycomb structure

2.1. Geometric description and material

In this study, novel hierarchical reentrant honeycomb structures are constructed by using nested substructures in place of the honeycomb walls of conventional recessed honeycombs. Here, three types of hierarchical inner-concave honeycombs are considered consisting of a reentrant honeycomb, a semi-concave honeycomb, and a classical hexagonal honeycomb substructure. The specific dimensions of the structure as well as the sub-cells were determined at the beginning of the study. The three novel hierarchical reentrant honeycomb structures are $c = 142\,\mathrm{mm}$ long and $b = 141\,\mathrm{mm}$ wide, and the cell wall thickness of all three cellular metamaterials is 0.6 mm, as shown in Fig. 1.

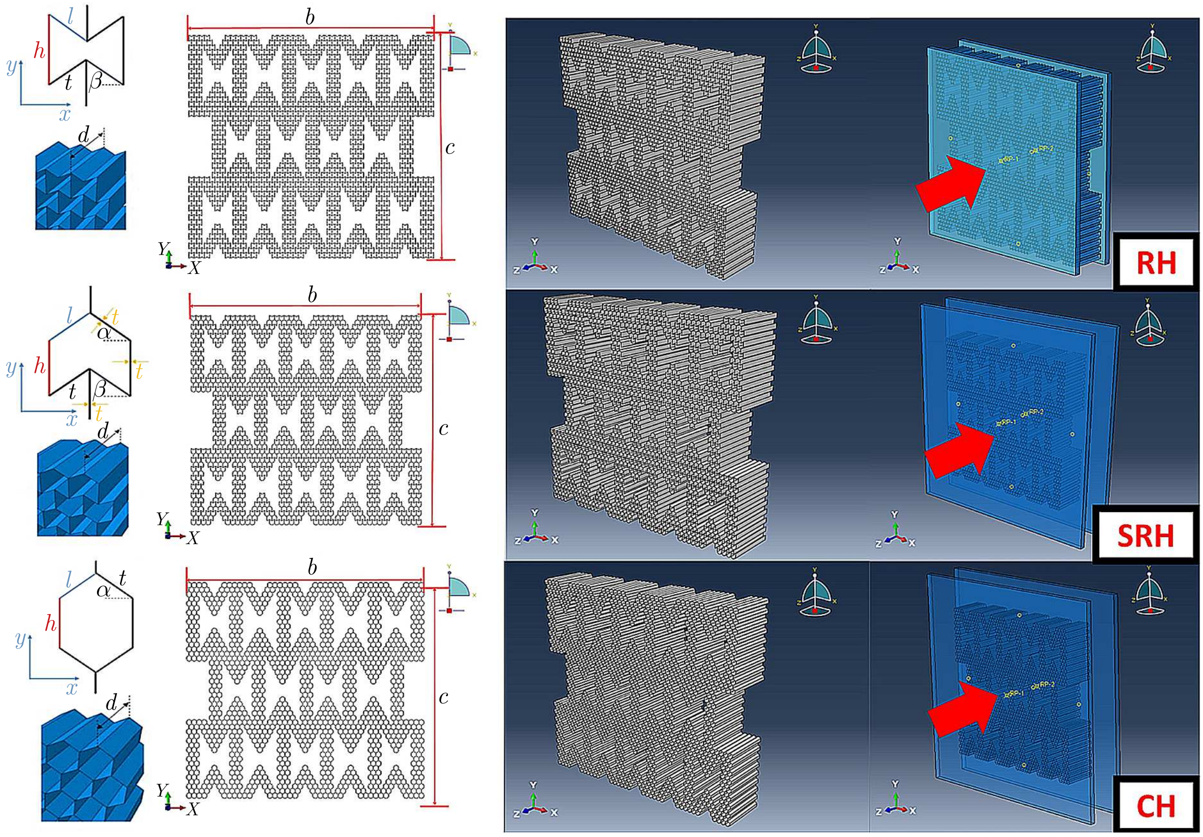


Fig. 1. Novel hierarchical reentrant honeycomb structure

A conventional hexagonal positive Poisson's ratio honeycomb is usually positive in the x-direction when pressurized in the y-direction. A positive Poisson's ratio indicates that the material tends to expand transversely when stretched in the longitudinal direction. In this study, the classical hexagonal honeycomb is utilized as a substructure, and the wall length $h = 2\,\mathrm{mm}$, $l = 1.8\,\mathrm{mm}$, turning angle $\alpha = 33.69^\circ$, and thickness $h = 2\,\mathrm{mm}$ of this substructure are determined.

The reentrant honeycomb structure exhibits negative Poisson's ratio (NPR) properties as a tensile extension material. In contrast to conventional materials, NPR materials contract in the transverse direction when they are stretched in the longitudinal direction. The reentrant honeycomb substructure designed in this study has wall length $h = 4\,\mathrm{mm}$, $l = 1.8\,\mathrm{mm}$, turning angle $\beta = 33.69^\circ$, and thickness $h = 2\,\mathrm{mm}$. The semi-reentrant honeycomb structure is a new type of hexagonal honeycomb structure that combines the characteristics of the traditional hexagonal honeycomb structure and the reentrant honeycomb structure, which exhibits a semi-reentrant appearance. Unlike the previous two structures, this structure exhibits zero Poisson's ratio (ZPR) properties (Ingrole etal, 2017; Farrokhabadi *et al.*, 1022) in tensile experiments, and this unique behavior is attributed to the fact that the structure combines the specific deformation mechanisms of the conventional hexagonal honeycomb with a positive Poisson's ratio (PPR) and the concave honeycomb negative Poisson's ratio (NPR) (Shukla and Behera, 2022). The semi-reentrant honeycomb substructure designed in this study has a wall length $h = 3\,\mathrm{mm}$, $l = 1.8\,\mathrm{mm}$, and thickness $d = 30\,\mathrm{mm}$. In this study, PLA (polylactic acid) was used as the main material to provide a feasible basis for the study with its good plasticity and biodegradation properties, and the specific material property parameters are shown in Table 1.

Table 1. Mechanical properties of PLA

Materials	Elastic modulus [MPa]	Yield strength [MPa]	Fracture strain	Poison's ratio	Mass density ρ_A [kg/m^3]
PLA	1451.14	39.14	0.18	0.3	1378

2.2. Specimen manufacturing

In this study, a polylactic acid (PLA) filament was chosen as the printing material based on the fused deposition principle. It has a low melting point, which allows it to melt and flow rapidly at a certain printing temperature, contributing to efficient printing. Three novel hierarchical reentrant honeycomb structures were used to print. Meanwhile, in order to ensure the quality of printed specimens, the printing speed of the printer was set to 50 mm/s, the nozzle diameter was set to 0.3 mm and the printing temperature was set to 210°C, while the temperature of the printing platform was set to 40°C in this study. The solid experimental samples of the three novel hierarchical reentrant honeycomb substructures are shown in Fig. 2.

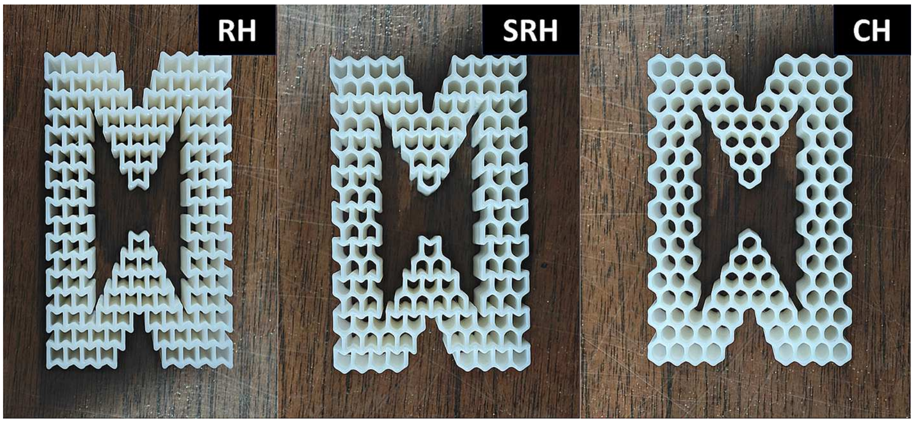


Fig. 2. 3D printed specimen of a novel hierarchical reentrant honeycomb substructure

2.3. Experiment and simulation setting

In the experimental phase, quasi-static compression tests were conducted using an Instron-5892 testing machine. A VIC-3D strain measurement system (model ProsilicaGained900) recorded deformation characteristics with a loading rate of 2 mm/min and compression images captured every 4 seconds until densification.

In the numerical simulation phase, detailed analysis was performed using Abaqus 2023 based on finite element theory. The compression and deformation mechanisms of three novel hierarchical reentrant honeycomb structures were studied using Abaqus 2023. Shell models were constructed based on previously designed geometric structures and meshed using four-node surface shell elements (S4R). To replicate the experimental compression state, rigid plates were placed on the front and rear surfaces of the structure, as shown in Fig. 1. To ensure simulation accuracy, a displacement boundary condition in the out-of-plane compression direction (Z direction) was applied to the front rigid plate, while the rear rigid plate was completely fixed. General contact was used as the interaction condition, with "hard contact" set in the normal direction and a friction coefficient of 0.3 set in the tangential direction. The mesh element size was set to 2 mm, with dynamic linear element types and element deletion enabled. Ultimately, each hierarchical structure was divided into approximately 90 000 elements.

3. Out of plane compression of novel hierarchical reentrant honeycomb structures

3.1. Simulation analysis

The stresses are calculated by the reaction forces F of the rigid grip section S is the area of grip section, so the stress is σ_1 or $\sigma_2 = F/S$. Strains $\varepsilon = U/d$ are computed from the displacements U of top grip divided by thickness d of the structure.

In this study, four compressive strain nodes of 0, 0.2, 0.4 and 0.6 are intercepted in order to investigate the deformation characteristics of the structure. The compressive deformation images of the overall structure are shown in Fig. 3. At a compressive strain of 0.2, the stress concentration of the RH structure is obviously in the inner recess of the novel hierarchical reentrant honeycomb structure, while the stress concentration range of the SRH structure and the CH structure is mainly characterized in the two sides of the wall of the hierarchical reentrant honeycomb structure, and it is worth noting that at this time, the three structures do not have a significant wave-shaped region as shown in Fig. 3 variable rows. At a compressive strain of 0.4, the reentrant of the novel hierarchical reentrant honeycomb structure that forms the whole has a tendency to shrink towards the center, whereas the shrinkage tendency of the RH structure is the smallest, which represents that the structure has a much higher stiffness and stability during the compression process. When the compressive strain reaches 0.6, all three structures show the phenomenon of collapse and fracture, and the compression of the overall structure tends to densify stability.

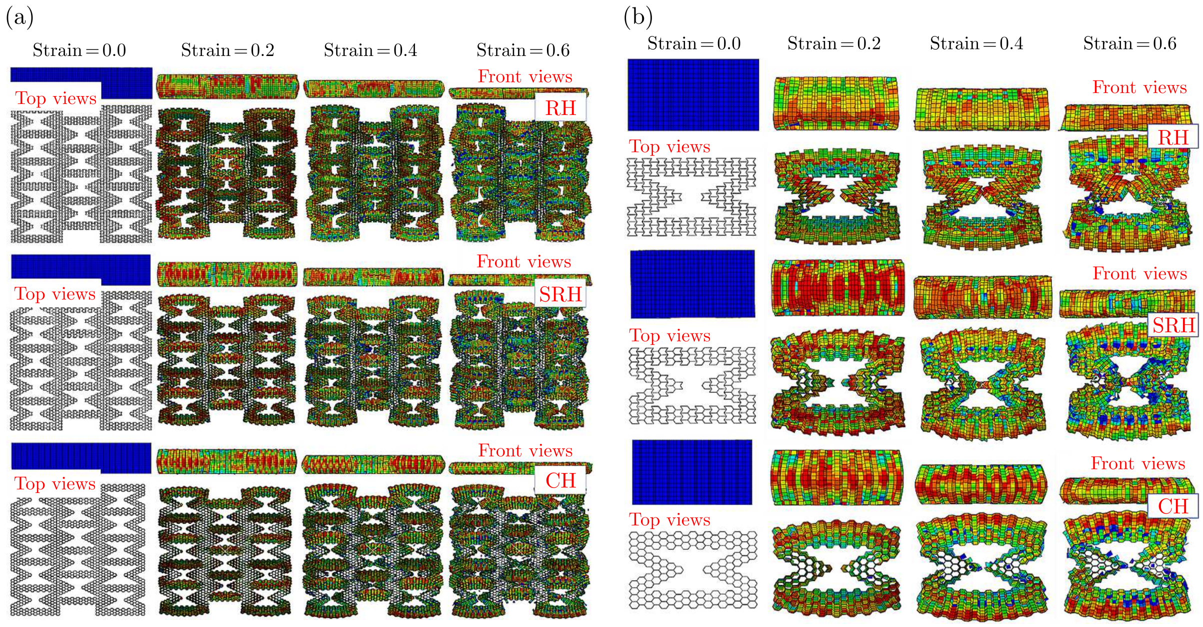


Fig. 3. Deformation contours of (a) entire structure, (b) substructure

To deeply investigate the deformation characteristics, this study was carried out to analyze characteristics of the novel hierarchical reentrant honeycomb structure. As shown in Fig. 3, the deformation characteristics of the substructure at four compressive strain nodes of 0, 0.2, 0.4 and 0.6 are plotted. The results show that the substructure has the same strain concentration region as the overall structure at a compressive strain of 0.2, while there is a phenomenon that the honeycomb subunits in the reentrant shrink towards the center region. In particular, in observing the front views process, both the RH and SRH structures have a stress concentration from the center-lower region and extend to the whole wall during compression. In contrast, the range of stress concentration for the CH structure is in the upper-middle region.

As shown in Fig. 4, the out of plane compressive stress-strain curves are shown for the overall structure and the substructure, respectively. The expression for the deformation energy per unit volume E corresponding to any point $(\varepsilon_a, \sigma_a)$ in this curve is given by

$$E = \int_0^{\varepsilon_a} \sigma_a \, d\varepsilon \tag{3.1}$$

The expression for calculating the energy absorption efficiency $\eta(\varepsilon_a)$ at any point is

$$\eta(\varepsilon_a) = \int_0^{\varepsilon_a} \frac{\sigma(\varepsilon)}{\sigma_a} \, d\varepsilon \qquad 0 \leqslant \varepsilon_a \leqslant 1 \tag{3.2}$$

The densification strain of this structure in out-of-plane compression was obtained by intercepting the highest point of the energy absorption efficiency-strain curve as shown in Fig. 4. Based on this value and the stress-strain curve, the mechanical property parameters associated with this structure were obtained as shown in Table 2.

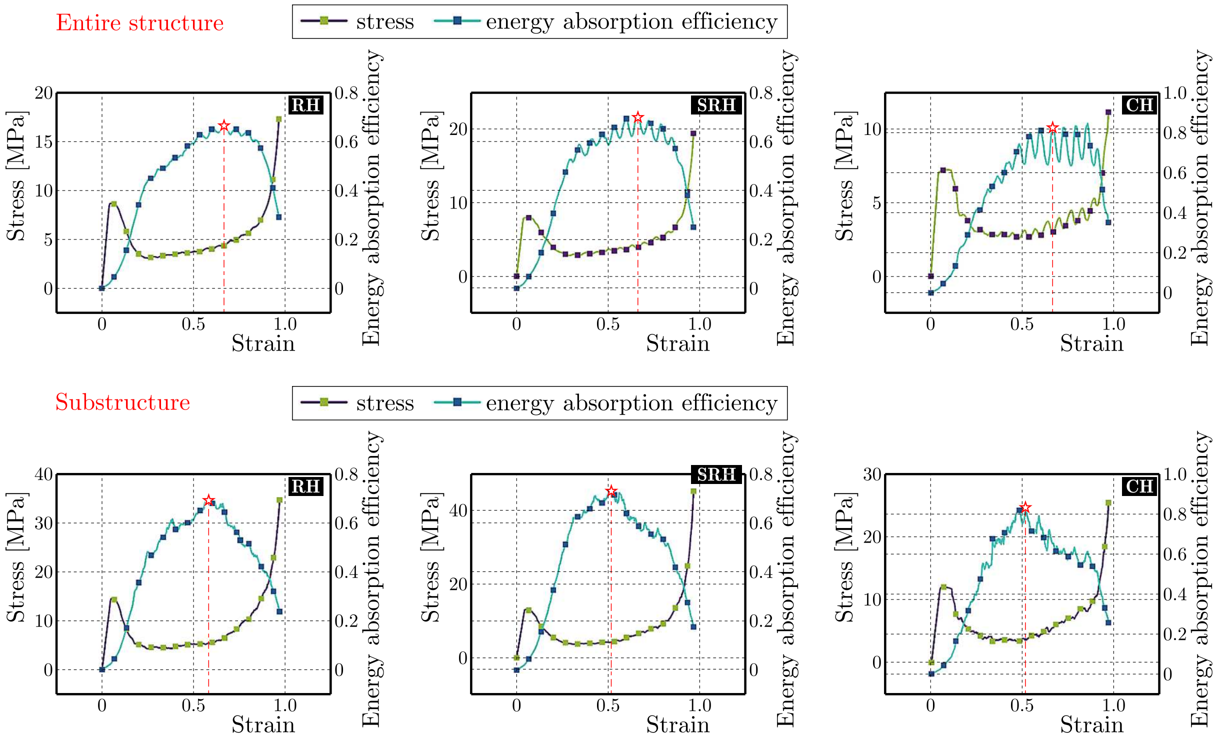


Fig. 4. Simulated out-of-plane compressive stress-energy absorption efficiency-strain curves

The results of the study show that the overall structure and the substructures are highly correlated and synchronized in the conclusive characterization of the mechanical property parameters and that the densification strains of the three structures are approximate and have characteristics of RH>SRH>CH in representation of the plateau stresses, the specific absorbed energies, and the modulus of elasticity, which means that the RH structure has an even better load-bearing capacity as well as stability. Meanwhile, the results of the study show that there is a positive relationship between the platform stress and the specific absorption energy.

Table 2. Mechanical properties of the novel hierarchical reentrant honeycomb structure in out-of-plane compression (simulation analysis)

Type	Plateau stress [MPa]		Energy absorption [J/g]		Densification strain		Elastic modulus [MPa]	
	Entire	Sub	Entire	Sub	Entire	Sub	Entire	Sub
RH	4.24	6.13	9.01	9.17	0.66	0.58	218.03	312.14
SRH	4.19	5.88	8.19	8.51	0.67	0.51	196.16	305.38
CH	4.15	5.35	7.24	8.02	0.67	0.51	174.21	255.59

3.2. Mechanical experiment

As shown in Fig. 5, the deformation characteristics at compressive strains of 0, 0.2, 0.4 and 0.6 reveal distinct stages of deformation, including initial plastic buckling, collapse fracture, and densification. At a compressive strain of 0.2, all three structures exhibit milky-white, worm-like plastic regions. The plastic regions of the RH and SRH structures are localized in the lower layers of the structure, while the plastic regions of the CH structure are more uniformly distributed. Upon reaching a compressive strain of 0.4, all three structures display varying degrees of collapse and fracture phenomena, with the plastic and fracture regions of the CH structure situated in the middle and upper layers of the structure. When the compressive strain reaches 0.6, the overall structure exhibits a tendency towards densification. This observation aligns entirely with the deformation characteristics identified in the simulation analysis above, providing a conclusive evidence of the synchronicity between finite element simulation and physical experiments in terms of deformation behavior.

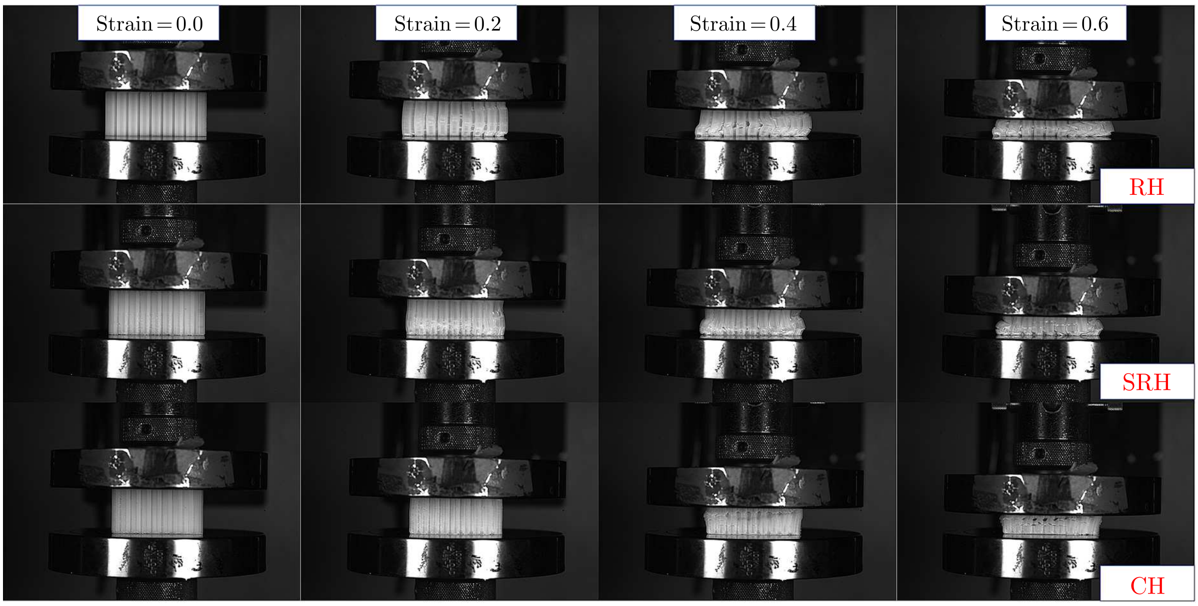


Fig. 5. Deformation contours of the novel hierarchical reentrant honeycomb structure

Plots of the stress-energy absorption efficiency-strain curves for this solid experiment are shown in Fig. 6. Based on the densification strain corresponding to the highest point of energy absorption efficiency, a detailed analysis of the mechanical property parameters was carried out in this study, and the results are summarized in Table 3. The results show that the three novel hierarchical reentrant honeycomb structures perform close to each other in terms of the values of densification strain. In the characterization of platform stress, modulus of elasticity, and specific absorption energy, the order is RH>SRH>CH, which indicates that the RH structure possesses superior mechanical properties in the out-of-plane compression solid experiment. In addition, the platform stress and specific absorption energy showed a positive proportional relationship in this out-of-plane compression solid experiment. Finally, this study verifies the synchronicity and consistency between the solid experiment and the finite element simulation analysis in the characterization of mechanical properties.

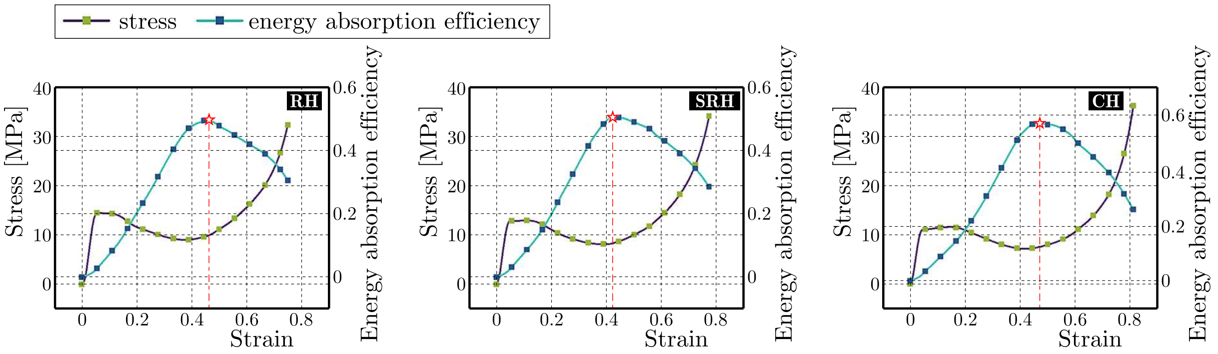


Fig. 6. Physical experiment for the out-of-plane compressive stress-energy absorption efficiency-strain curves

Table 3. Mechanical properties of the novel hierarchical reentrant honeycomb structure in out-of-plane compression (physical experiment)

Type	Plateau stress [MPa]	Energy absorption [J/g]	Densification strain	Elastic modulus [MPa]
RH	11.07	9.16	0.47	427.89
SRH	10.44	8.71	0.42	401.39
CH	9.45	8.5	0.47	366.53

4. Conclusion

In this study, we designed novel hierarchical reentrant honeycomb structures with different Poisson's ratio characteristics, including positive, zero, and negative Poisson's ratios. Using finite element simulations and 3D printing experiments based on fused deposition technology, we investigated the mechanical properties and deformation characteristics of those structures under out-of-plane compression. The results showed that the stress concentration in RH structure was located in the inner recess, while in SRH and CH structures, the stress concentration was mainly on the sides of the walls. The RH structure exhibited the smallest contraction tendency, indicating higher stiffness and stability.

Both experiments and simulations confirmed that the substructures had the same strain concentration regions as the overall structure. From the front view, the reentrant honeycomb subunits also showed a tendency to shrink towards the center. During compression, the RH and SRH structures generated worm-like plastic regions from the lower center, extending to the entire wall, while the plastic regions in the CH structure were located in the upper center.

Mechanical performance analysis indicated that the RH structure had the highest stiffness, stability, load-bearing capacity, and energy absorption capability, followed by the SRH and CH structures. This study verifies the synchronicity between simulation and experimental results in characterizing the mechanical properties under out-of-plane compression.

The enhanced mechanical properties and energy absorption capabilities of these novel hierarchical reentrant honeycomb structures suggest significant practical applications in aerospace, automotive, and civil engineering. They are suitable for lightweight, high-strength components such as fuselage panels in aerospace, impact-resistant structures in automotive applications, and improved building materials in civil engineering to withstand dynamic loads like earthquakes. This study highlights the potential of these innovative honeycomb structures in advancing engineering design and materials science.

Acknowledgement

The authors would like to acknowledge the support from the National Natural Science Foundation of China (Grant No. 52405441), the Fundamental Research Funds for the Central Universities (Grant No. 3122024041) and the Graduate Research Innovation Program of Civil Aviation University of China (Grant No. 2023YJSKC12002).

References

1. Ahmed N., Xue P., 2019, Governing the in-plane axial crushing of honeycomb with regular hexagonal symmetric division cells using cross-hinge inserts, *International Journal of Mechanical Sciences*, **161-162**, 105062
2. Arbaoui J., Schmitt Y., Pierrot J.-L., Royer F.-X., 2014, Numerical simulation and experimental bending behaviour of multi-layer sandwich structures, *Journal of Theoretical and Applied Mechanics*, **52**, 2, 431-442
3. Bouazza M., Becheri T., Boucheta A., Benseddiq N., 2019, Bending behavior of laminated composite plates using the refined four-variable theory and the finite element method, *Earthquakes and Structures*, **17**, 3, 257-270
4. Davalos J.F., Qiao P., Xu X.F., Robinson J., Barth K.E., 2001, Modeling and characterization of fiber-reinforced plastic honeycomb sandwich panels for highway bridge applications, *Composite Structures*, **52**, 3-4, 441-452
5. Farrokhabadi A., Ashrafian M.M., Behravesh A.H., Hedayati S.K. 2022, Assessment of fiber-reinforcement and foam-filling in the directional energy absorption performance of a 3D printed accordion cellular structure, *Composite Structures*, **297**, 115945
6. Gadkaree K.P., 1998, Carbon honeycomb structures for adsorption applications, *Carbon*, **36**, 7-8, 981-989
7. Gao W., Zhang W., Yu H., Xing W., Yang X., *et al.*, 2022, 3D CNT/MXene microspheres for combined photothermal/photodynamic/chemo for cancer treatment, *Frontiers in Bioengineering and Biotechnology*, **10**, 996177
8. Garg A., Belarbi M.O., Chalak H.D., Li L., Sharma A., *et al.*, 2023, Buckling and free vibration analysis of bio-inspired laminated sandwich plates with helicoidal/Bouligand face sheets containing softcore, *Ocean Engineering*, **270**, 113684
9. Gong X., Ren C., Sun J., *et al.*, 2022, 3D zero Poisson's ratio honeycomb structure for morphing wing applications, *Biomimetics*, **7**, 4, 198
10. Günaydin K., Rea C., Kazanci Z., 2022, Energy absorption enhancement of additively manufactured hexagonal and re-entrant (auxetic) lattice structures by using multi-material reinforcements, *Additive Manufacturing*, **59**, 103076
11. Ha N.S., Pham T.M., Hao H., Lu G., 2021, Energy absorption characteristics of bio-inspired hierarchical multi-cell square tubes under axial crushing, *International Journal of Mechanical Sciences*, **201**, 106464
12. Ha N.S., Pham T.M., Tran T.T., Hao H., Lu G., 2022, Mechanical properties and energy absorption of bio-inspired hierarchical circular honeycomb, *Composites Part B: Engineering*, **236**, 109818

13. Hamzehei R., Zolfagharian A., Dariushi S., Bodaghi M., 2022, 3D-printed bio-inspired zero Poisson's ratio graded metamaterials with high energy absorption performance, *Smart Materials and Structures*, **31**, 3, 035001
14. Heo H., Ju J., Kim D.M., 2013, Compliant cellular structures: Application to a passive morphing airfoil, *Composite Structures*, **106**, 560-569
15. Ingrole A., Hao A., Liang R., 2017, Design and modeling of auxetic and hybrid honeycomb structures for in-plane property enhancement, *Materials and Design*, **117**, 72-83
16. Lei Z., Sun X., Zhu S., Dong K., Liu X., *et al.*, 2022, Nature inspired MXene-decorated 3D honeycomb-fabric architectures toward efficient water desalination and salt harvesting, *Nano-Micro Letters*, **14**, 1, 10
17. Li Z., Liu D., Qian Y., Wang Y., Wang T., Wang L., 2019, Enhanced strength and weakened dynamic sensitivity of honeycombs by parallel design, *International Journal of Mechanical Sciences*, **151**, 672-683
18. Liang H., Wang Q., Pu Y., Zhao Y., Ma F., 2021, In-plane compressive behavior of a novel self-similar hierarchical honeycomb with design-oriented crashworthiness, *International Journal of Mechanical Sciences*, **209**, 106723
19. Liu J., Chen W., Hao H., Wang Z., 2021, In-plane crushing behaviors of hexagonal honeycombs with different Poisson's ratio induced by topological diversity, *Thin-Walled Structures*, **159**, 107223
20. Liu L., Wang H., Guan Z., 2015, Experimental and numerical study on the mechanical response of Nomex honeycomb core under transverse loading, *Composite Structures*, **121**, 304-314
21. Masuda H., Fukuda K., 1995, Ordered metal nanohole arrays made by a two-step replication of honeycomb structures of anodic alumina, *Science*, **268**, 5216, 1466-1468
22. Shukla S., Behera B.K., 2022, Auxetic fibrous structures and their composites: A review, *Composite Structures*, **290**, 115530
23. Solak A., Aşçioğlu Temiztaş B., Bolat B., 2023, Numerical investigation of the mechanical behavior of the vertical stabilizer leading edge with wavy honeycomb sandwich structure under bird strike, *Journal of Sandwich Structures and Materials*, **25**, 3, 387-400
24. Teng X.C., Ren X., Zhang Y., Jiang W., Pan Y., 2022, A simple 3D re-entrant auxetic metamaterial with enhanced energy absorption, *International Journal of Mechanical Sciences*, **229**, 107524
25. Wang Z., Zhou Y., Wang X., Wei K., 2021, Compression behavior of strut-reinforced hierarchical lattice – Experiment and simulation, *International Journal of Mechanical Sciences*, **210**, 106749
26. Wu D., Li D., Zhang Z., Chen J., 2023, Numerical study on compression properties of semi--reentrant filled tubular structures, *Journal of Theoretical and Applied Mechanics*, **61**, 2, 233-244
27. Yang M., Han B., Su P., Zhang Q., Zhang Q.C., *et al.*, 2021, Crashworthiness of hierarchical truncated conical shells with corrugated cores, *International Journal of Mechanical Sciences*, **193**, 106171
28. Zhang X., Zhang H., 2013, Theoretical and numerical investigation on the crush resistance of rhombic and kagome honeycombs, *Composite Structures*, **96**, 143-152

Manuscript received January 21, 2024; accepted for publication October 8, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 13-26, Warsaw 2025
https://doi.org/10.15632/jtam-pl/194501

INVESTIGATION ON IMPACT PROTECTION EFFECTIVENESS OF CUSHIONING MATERIAL IN PACKAGING CONTAINER UNDER LATERAL CONSTRAINT CONDITION

Zexiong Zhang
University of Science and Technology of China, Department of Modern Mechanics, Hefei, Anhui, China, and
Shock and Vibration of Engineering Materials and Structures Key Laboratory of Sichuan Province, Mianyang, China
e-mail: jz1108@mail.ustc.edu.cn (Zexiong Zhang); lijiaxing21@gscaep.ac.cn (Jiaxing Li)

Weizhou Zhong, Jiaxing Li
Shock and Vibration of Engineering Materials and Structures Key Laboratory of Sichuan Province, Mianyang, China
corresponding author Weizhou Zhong, e-mail: zhongwzcaep@126.com

Jingrun Luo
University of Science and Technology of China, Department of Modern Mechanics, Hefei, Anhui, China, and
Shock and Vibration of Engineering Materials and Structures Key Laboratory of Sichuan Province, Mianyang, China
e-mail: luojr@caep.cn

A theoretical model is proposed to evaluate the impact protection effectiveness of a porous cushioning material in a packaging container under the lateral constraint condition. An acceleration-displacement equation of the protected product in the packaging container is derived. The reliability of the equation is validated by numerical simulation. Subsequently, the equation is applied to analyse the effect of strain rate on impact protection effectiveness of three polymer foams under the lateral constraint condition, and to design the thicknesses of cushioning materials in the packaging container.

Keywords: impact protection, packaging container, cushioning material, lateral constraint, numerical simulation

1. Introduction

The impact-resistant packaging container generally consists of the outer shell, cushioning material and protected product. The cushioning material is usually filled between the protected product and the outer shell. The mechanical behaviour of the cushioning material determines the impact protection effectiveness of the packaging container. Therefore, evaluating the energy absorption of the cushioning material and optimizing cushioning material configuration are feasible approaches to improve the protective performance of the packaging container.

The mechanical behaviour and energy dissipation of the cushioning material greatly affect the protection effectiveness of the packaging container. Many scholars focus on promoting the energy absorption of cushion materials in recent years. For pre-designing configuration of cushioning materials, Meng *et al.* (2020) examined the mechanical response of composite materials made of honeycomb aluminium under explosive impact loads, and proposed an empirical formula to characterize the compression depth of double-layer honeycomb aluminium composite materials. Kader *et al.* (2016) assessed the shock propagation and elastic-plastic deformation behaviour of a closed-cell aluminium foam by experiments and numerical simulations. The results showed the significant influence of the foam material topology on the shortest path for stress wave propagation in a porous material. Estrada *et al.* (2017) analysed the sizing effect of discontinuities on energy absorption characteristics of a steel square profile. It indicated that pre-designed discontinuities effectively enhanced energy absorption in square-section steel tubes, whereas reduced the peak load. Baroutaji *et al.* (2016) analysed the energy absorption characteristics of nested circular tube structures under lateral loads, and explored the effects of geometrical and loading parameters on the responses of the best nested tube system. Sek and Rouillard (2006) investigated the response characteristics of multilayer corrugated cardboard cushioning systems under impact loads, and demonstrated that adding anti-wrinkle liners to soft cardboard cushion pads significantly expanded their protective coverage and capability to withstand extreme conditions. Guo *et al.* (2021) investigated the mechanical behaviour and cushioning energy absorption of paper composite sandwich structures with sinusoidal corrugation and hexagonal honeycomb cores. They showed that the influence of static compression rate on the yield strength and cushioning energy absorption of the paper composite sandwich structure was not obvious, and the composite sandwich structures with large inertia moment and large thickness had more excellent cushioning energy absorption.

Moreover, cushion materials with gradient density can be designed to improve energy absorption. Baertsch *et al.* (2021) optimized an auxetic re-entrant structure with a stiffness gradient for enhanced energy absorption with a low acceleration peak, and found that concave honeycomb structures with gradient stiffness variations exhibited superior energy absorption capability. Zeng *et al.* (2010) investigated the influence of the density gradient profile on the mechanical response of graded polymeric hollow sphere agglomerates under an impact loading. They showed that foam materials with gradient density or stiffness had a stronger cushioning effectiveness along the direction of gradient variation. Koohbor and Kidane (2016) proposed a semi-analytical approach to study the effect of density gradation in graded polymeric foams. A strong concave gradient was shown to promote substantially lighter structural weights with superior energy absorption. Gputa (2007) found that the compressive modulus, strength, and total energy absorption of syntactic foams could be controlled by an appropriate type and volume fraction of microballoons. The compressive strength and modulus of a functionally graded syntactic foam were dependent on the weakest layer in their structure. The literature indicates that optimization of structure topology and gradient configuration is a feasible approach to enhance the energy absorption of the cushioning material.

Apart from the studies on mechanical behaviour and energy dissipation of cushioning materials, many scholars (Lye *et al.*, 2004; Chang *et al.*, 2017; Zhang *et al.*, 2011; Pan and Yu, 2022) have performed impacting experiments to evaluate the impact protection effectiveness of packaging container systems. Luong *et al.* (2021) analysed the effects of repeated impacts on corrugated paper packaging by numerical simulations and experimental tests. An elastoplastic homogenization model was proposed to replace a corrugated-core sandwich panel by a homogeneous plate. Hussain *et al.* (2021) investigated the energy absorption of composite material automotive crash boxes using drop-weight impact tests. The strength-to-weight ratio was considered as the important factor for the packaging boxes made of different cross-sections cushion materials. An and Shi (2022) proposed an innovative reliability optimization method for cushioning design based on the dynamic stress-strain curve and an active set strategy. It provides a new design method for cushioning packaging based on the dynamic cushioning material stress-strain curve. In the aforementioned literatures, the cushioning materials are usually taken as fully compressible materials, ignoring the influence of lateral constraint conditions. A porous material volume exhibits incompressible characteristics in the compaction stage. The lateral constraint makes the cushion material stress increasing greatly with strain in the compaction process. It is necessary to investigate the compressive behaviour of cushioning materials under lateral constraint conditions, which contributes to improve the cushion evaluation and the optimization design of the packaging container.

For a packaging container with a high strength outer shell, the cross-sectional area of the porous cushioning material can be taken as constant during the impact compression process. Based on the stress wave theory and propagation characteristics of stress waves in a porous material (Reid and Peng, 1997; Karagiozova, 2011), the acceleration of the protected product can be derived to evaluate the cushioning protection effectiveness of the packaging container. The theoretical model of the drop impact for cushioning material packaging containers is established, and the acceleration-displacement equation for the protected product is given in the paper. The equation indicates the peak acceleration of the protected product, which is influenced by density, platform stress and compaction strain of the cushioning material. Combining the mechanical properties of an expanded polystyrene foam (EPS), flexible polyurethane foam (FPUF) (Zhang *et al.*, 2023), and rigid polyurethane foam (RPUF) (Zhang *et al.*, 2022), the theoretical equation is verified by numerical simulations. The influence of the strain rate on the impact protection effectiveness of the three polymer foam materials under the lateral constraint condition is analysed, then the equation for computing the optimal thickness of the cushioning material in the packaging container is derived.

2. Theoretical model of the packaging container drop impact

2.1. Fundamental equation for the stress wave in a porous cushioning material

The stress-strain curve of porous materials exhibits a concave characteristic in the plastic stage. When a high amplitude stress wave propagates in a porous material, the material undergoes compaction and "shock waves" are generated. Ignoring the stress reflection or "shock waves" interaction, the stress in the porous material can be simplified in Eq. (2.1) (Karagiozova, 2011)

$$\sigma(\varepsilon)=\begin{cases}\sigma_p & 0\leqslant\varepsilon<\varepsilon_D\\ +\infty & \varepsilon_D\leqslant\varepsilon\end{cases}\tag{2.1}$$

where σ, ε is stress and strain, σ_p is plastic platform stress and ε_D is compaction strain.

Figure 1 shows a schematic of a strong discontinuity wave front. The physical variables S^+ and S^- show the states in ahead and behind wave front, respectively. The difference between the value is denoted as $[S]$. The C, A, v and x are wave propagation velocity, section area, velocity and displacement, respectively.

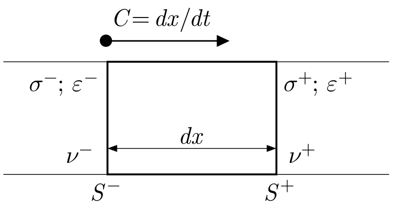


Fig. 1. Schematic of strong intermittent wave front propagation

The momentum conservation condition of the wave front is as follows

$$(\sigma^+-\sigma^-)Adt=\rho AdX(v^--v^+)\qquad\qquad [\sigma]=-\rho C[v]\tag{2.2}$$

The displacement continuity condition is expressed as

$$\frac{d}{dt}[u]=[X_{,t}]+C[X_{,x}]=[v]+C[\varepsilon]=0\qquad\qquad [v]=-C[\varepsilon]\tag{2.3}$$

The energy conservation equation is given below, where e is the internal energy per unit mass of the porous material

$$
\begin{aligned}
&(\sigma^+v^+ - \sigma^-v^-)Adt = (e^- - e^+)\rho AdX + \frac{1}{2}\rho AdX\{(v^-)^2 - (v^+)^2\} \\
&[\sigma v] = -\rho C[e] - \frac{1}{2}C\rho[v^2] \\
&-\rho C[e] = [\sigma v] + \frac{1}{2}C\rho[v^2] = [\sigma v] + \frac{1}{2}C\rho[v](v^+ + v^-) = [\sigma v] - \frac{1}{2}[\sigma](v^+ + v^-) \\
&-\rho C[e] = \sigma^+v^+ - \sigma^-v^- - \frac{1}{2}(\sigma^+v^+ + \sigma^+v^- - \sigma^-v^+ - \sigma^-v^-) \\
&-\rho C[e] = \frac{1}{2}(\sigma^+ + \sigma^-)(v^+ - v^-) \\
&\rho C[e] = -\frac{1}{2}(\sigma^+ + \sigma^-)[v] = \frac{1}{2}(\sigma^+ + \sigma^-)C[\varepsilon] \\
&\rho[e] = \frac{1}{2}(\sigma^+ + \sigma^-)[\varepsilon]
\end{aligned}
\tag{2.4}
$$

2.2. Acceleration-displacement equation for the protected product

During impacting the packaging container, the cushioning material is usually constrained by the outer shell. The cross-section area of the cushion layer can be taken as constant. In order to analyse theoretically analyse the acceleration response of the protected product during impact process, the packaging container is simplified as shown in Fig. 2. It only consists of the protected product, cushioning material and target. The left subfigure shows the undeformed state of the packaging container at the initial moment. The middle subfigure represents the deformation state of the packaging container at a certain moment t. The right subfigure represents the deformation state of the packaging structure at a certain moment $t + \Delta t$. M_a, ρ_0 and A are mass, density and section area of the cushioning material in the original stage, respectively. σ_d and ρ_D are stress and density of the cushioning material in the densification stage, respectively. M_b is mass of the protected product and C is the stress wave propagation speed in the cushion material. It is assumed that the cushion material is compressed into the densification layer in the impact process. The white and blue areas denote the uncompacted part and the densification part of the cushion material, respectively.

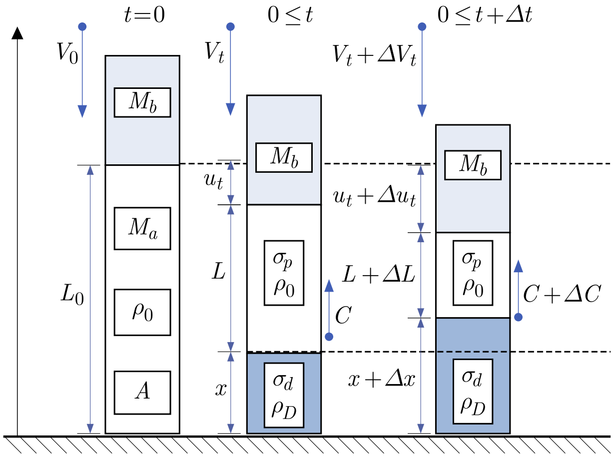


Fig. 2. Drop model of the packaging container

The variables ahead the wave front are as follows

$$
v^+ = V_t \qquad \varepsilon^+ = 0 \qquad \sigma^+ = \sigma_p \qquad \rho^+ = \rho_0 \tag{2.5}
$$

The variables behind the wave front are

$$v^- = 0 \qquad \varepsilon^- = \varepsilon_D \qquad \sigma^- = \sigma_d \qquad \rho^- = \rho_D = \frac{\rho_0}{1-\varepsilon_D} \tag{2.6}$$

where σ_d represents the dynamic stress. Based on Eqs. (2.2), (2.3), (2.5) and (2.6), the equations $[\sigma] = -\rho C[v]$ and $V_t = -C[\varepsilon]$ can be expressed as

$$\sigma_p - \sigma_d = -\rho_0 C(V_t - 0) = -\frac{\rho_0 V_t^2}{\varepsilon_D} \qquad \sigma_d = \sigma_p + \frac{\rho_0 V_t^2}{\varepsilon_D} \tag{2.7}$$

The length of the undeformed zone is as follows

$$L = L_0 - u_t - x \tag{2.8}$$

In Fig. 2, the porous cushioning material with initial length $u_t + x$ is compressed into the densification layer with length x. The strain ε_D can be defined as the ratio of the reduced length to the original length

$$\begin{aligned} &\varepsilon_D = \frac{u_t}{u_t + x} \qquad x = \frac{1-\varepsilon_D}{\varepsilon_D} u_t \\ &u_t + x = u_t + \frac{1-\varepsilon_D}{\varepsilon_D} u_t = \frac{u_t}{\varepsilon_D} \end{aligned} \tag{2.9}$$

According to Eq. (2.8) and Eq. $(2.9)_3$, the following equation can be given

$$\frac{dL}{dt} = -\frac{d(u_t + x)}{dt} = -\frac{1}{\varepsilon_D}\frac{du_t}{dt} = -\frac{V_t}{\varepsilon_D} \tag{2.10}$$

Based on displacement continuity and conservation of momentum and energy, the internal energy of the packaging container is constant. The compression of the porous cushioning material results in a change of internal energy ΔE_U. If the reduced length is ΔL, the internal energy variation is as follows

$$\begin{aligned} \Delta E_U &= E_{U2} - E_{U1} = \rho[e]A\Delta L = \frac{1}{2}(\sigma_d + \sigma_p)[\varepsilon]A\Delta L = \frac{1}{2}\Big(2\sigma_p + \rho_0 \frac{V_t^2}{\varepsilon_D}\Big)(0 - \varepsilon_D)A\Delta L \\ &= -\Big(\sigma_p + \frac{\rho_0 V_t^2}{2\varepsilon_D}\Big)\varepsilon_D A\Delta L \end{aligned} \tag{2.11}$$

The compaction part of the cushioning material has velocity 0, while the velocity of the protected product and the uncompressed part of the cushioning material is V_t. The change in kinetic energy ΔE_K of the packaging container can be obtained by following equations

$$\begin{aligned} \Delta E_K &= E_{K2} - E_{K1} = \frac{1}{2}\{M_b + \rho_0(L + \Delta L)A\}(V_t + \Delta V_t)^2 - \frac{1}{2}(M_b + \rho_0 LA)V_t^2 \\ &= \frac{1}{2}M_b(\Delta V_t^2 + 2V_t\Delta V_t) + \frac{1}{2}\rho_0 LA(\Delta V_t^2 + 2V_t\Delta V_t) + \frac{1}{2}\rho_0\Delta LA(V_t + \Delta V_t)^2 \\ &= M_b V_t \Delta V_t + \rho_0 LAV_t\Delta V_t + \frac{1}{2}\rho_0 \Delta LAV_t^2 + O\{(\Delta V_t^2 + 2\Delta L\Delta V_t + \Delta L\Delta V_t^2\} \end{aligned} \tag{2.12}$$

Ignoring higher order infinitesimals, the formula is simplified to

$$\Delta E_K = M_b V_t \Delta V_t + \rho_0 LAV_t \Delta V_t + \frac{1}{2}\rho_0 \Delta LAV_t^2 = \frac{1}{2}\{2V_t(M_b + \rho_0 LA)\Delta V_t + \rho_0 V_t^2 \Delta L\} \tag{2.13}$$

The total energy change of the packaging container is zero, as in the following

$$\Delta E_U + \Delta E_K = 0 \qquad \Delta E_K = -\Delta E_U \tag{2.14}$$

The packaging container energy conservation equation is expressed as

$$M_b V_t \varDelta V_t + \rho_0 L A V_t \varDelta V_t + \frac{1}{2}\rho_0 \varDelta L A V_t^2 = \Big(\sigma_p + \frac{\rho_0 V_t^2}{2\varepsilon_D}\Big)\varepsilon_D A \varDelta L \tag{2.15}$$

Then, Eq. (2.8) and Eq. $(2.9)_1$ can be written as

$$\begin{aligned} &M_b V_t \varDelta V_t + \rho_0 (L_0 - u_t - x) A V_t \varDelta V_t + \frac{1}{2}\rho_0 \varDelta L A V_t^2 = \sigma_p \varepsilon_D A \varDelta L + \frac{1}{2}\rho_0 V_t^2 A \varDelta L \\ &M_b V_t \varDelta V_t + \rho_0 \Big(L_0 - \frac{u_t}{\varepsilon_D}\Big) A V_t \varDelta V_t = \sigma_p \varepsilon_D A \varDelta L \end{aligned} \tag{2.16}$$

Equation $(2.16)_2$ is divided by dt and the limit of dt is close to 0. The acceleration-displacement equation of the protected product is obtained

$$\Big(M_b V_t + \rho_0 L_0 A V_t - \frac{\rho_0 A}{\varepsilon_D} u_t V_t\Big)\frac{dV_t}{dt} = \sigma_p \varepsilon_D A \frac{dL}{dt} \tag{2.17}$$

According Eq. (2.10) and Eq. (2.17), the following equation can be obtained

$$\begin{aligned} &\Big(M_b V_t + \rho_0 L_0 A V_t - \frac{\rho_0 A}{\varepsilon_D} u_t V_t\Big)\frac{dV_t}{dt} = -\sigma_p \varepsilon_D A \frac{V_t}{\varepsilon_D} \\ &\Big(M_b + \rho_0 L_0 A - \frac{\rho_0 A}{\varepsilon_D} u_t\Big) a_t = -\sigma_p A \end{aligned} \tag{2.18}$$

and

$$a_t = \frac{\sigma_p \varepsilon_D}{\rho_0 u_t - m\varepsilon_D - \rho_0 L_0 \varepsilon_D} \qquad m = \frac{M_b}{A} \tag{2.19}$$

Taking parameter symbols of I, J and K into account, Eq. $(2.19)_1$ is equivalently expressed as

$$a = \frac{I}{Ju - K} \tag{2.20}$$

and

$$I = \sigma_p \varepsilon_D \qquad J = \rho_0 \qquad K = m\varepsilon_D + \rho_0 L_0 \varepsilon_D \tag{2.21}$$

Based on acceleration-displacement Eqs. (2.20) and (2.21), one can conclude that acceleration of the protected product is related with density, compact strain, and platform stress of the cushioning material. The acceleration-displacement equation of the protected product can be calculated by the parameters of I, J and K. I is the combination of the platform stress and compact strain, which is the maximum energy absorption per unit volume of the cushioning material. J is density of the cushioning material. K depens on the compact strain, density, section area, original length of the cushioning material, and the protected product mass.

3. Numerical validation of theoretical model

To verify the theoretical model in Eq. (2.20), three polymer foam cushioning materials (EPS, FPUF, RPUF) are selected for the packaging container. The geometric dimensions of the packaging container are shown in Fig. 3. The container shell is made of aluminium. The inner and outer diameters of the shell are 72 mm and 78 mm, respectively. The protected product is a steel cylinder. The diameter and height are 72 mm and 78 mm. The two cushioning layers are with

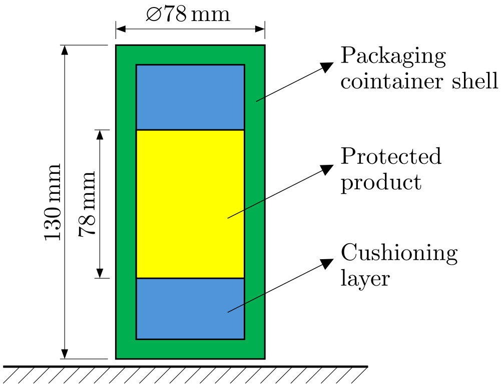


Fig. 3. Geometric diagram of packaging container

Table 1. Geometric paraneters of the packaging container

Parameter	Diameter [mm]	Height [mm]	Thickness [mm]
Outer shell	78	130	3
Protected product	72	78	–
Cushioning layer	72	20	–

the same size. The diameter and height of the cushioning layer are 72 mm and 20 mm. Detailed geometrical information is given in Table 1.

Taking density of the protected product as 7810 kg/m^3, then area density is 609.18 kg/m^2. The densities, equivalent platform stresses and compaction strains of EPS, FPUF and RPUF are cited from the references of Zhang *et al.* (2022, 2023). The initial impact velocity of the packaging container is 9.8 m/s. The parameters I, J, K for EPS, FPUF, and RPUF at the strain rate of 100/s are calculated by Eqs. (2.21), see Table 2.

Table 2. Cushioning material mechanical properties and design parameters

Material	Density [kg/m^3]	Platform stress [MPa]	Compaction strain	I	J	K
EPS-100/s	43.5	0.566	0.660	373390	43.5	402.633
FPUF-100/s	77.8	0.045	0.560	25431	77.8	342.012
RPUF-100/s	125.6	2.417	0.630	1522518	125.6	385.366

The finite element model of the packaging container is shown in Fig. 4. The material of the protected product and the target plate is steel. The container shell is made of aluminium. The mechanical properties of packaging container materials are shown in Table 3. An elastic-plastic constitutive model is adopted to describe mechanical behaviour of steel and aluminium. A crushable foam model is adopted to describe mechanical behaviour of EPS and RPUF. The mechanical behaviour of FPUF is described by the hyperfoam constitutive model. The compressive stress-strain curves of EPS, FPUF, RPUF at the strain rate of 100/s are shown in Fig. 5.

Figure 6 shows von Mises stress and equivalent plastic strain of the packaging container when the protected product velocity is 0. The packaging containers of FPUF, RPUF exhibit the highest and lowest equivalent plastic strain, respectively. It indicates that FPUF was a lower energy absorption, and energy absorption capacity of RPUF is stronger. The whole cushioning

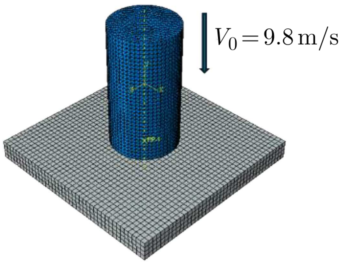


Fig. 4. Finite element model of the packaging container

Table 3. Material properties of the packaging container

Material	Density [kg/m^3]	Elastic modulus [MPa]	Poisson's ratio	Yield stress [MPa]
Aluminium	2700	69000	0.30	290
Steel	7810	212000	0.30	460
EPS-100/s	43.5	9.29	0.01	–
FPUF-100/s	77.8	0.91	0.01	–
RPUF-100/s	125.6	45.30	0.01	–

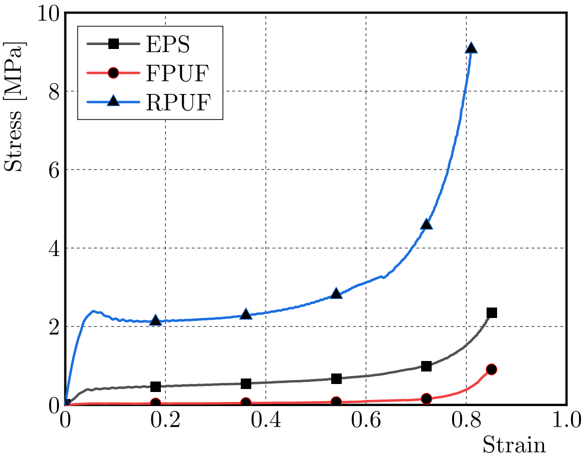


Fig. 5. Compressive stress-strain curves of cushioning materials

layer of FPUF is compacted, which results in that the kinetic energy of the protected product can not be absorbed fully by the cushioning material in the impact process. The cushioning material EPS has with a moderate energy absorption capacity. The numerical results show partial compaction of the EPS cushioning layer. In contrast, RPUF has a higher platform stress, which induces the highest energy absorption in the impact process.

The plastic platform stress is taken as constant in the deformation process in Eq. (2.20). The theoretical expression is only feasible in the platform stress stage, namely the protected product displacement is less than $l_0\varepsilon_D$. The expression is not suitable in the compaction stage. Figure 7 is the displacement-time curve comparison of the protected product by theoretical calculation and numerical simulation, respectively. It shows that the cushioning layers of EPS and FPUF are fully compacted in the drop impact process. The results of theoretical analysis and numerical simulation are consistent before the compaction stage, which verifies reliability of the theoretical analysis.

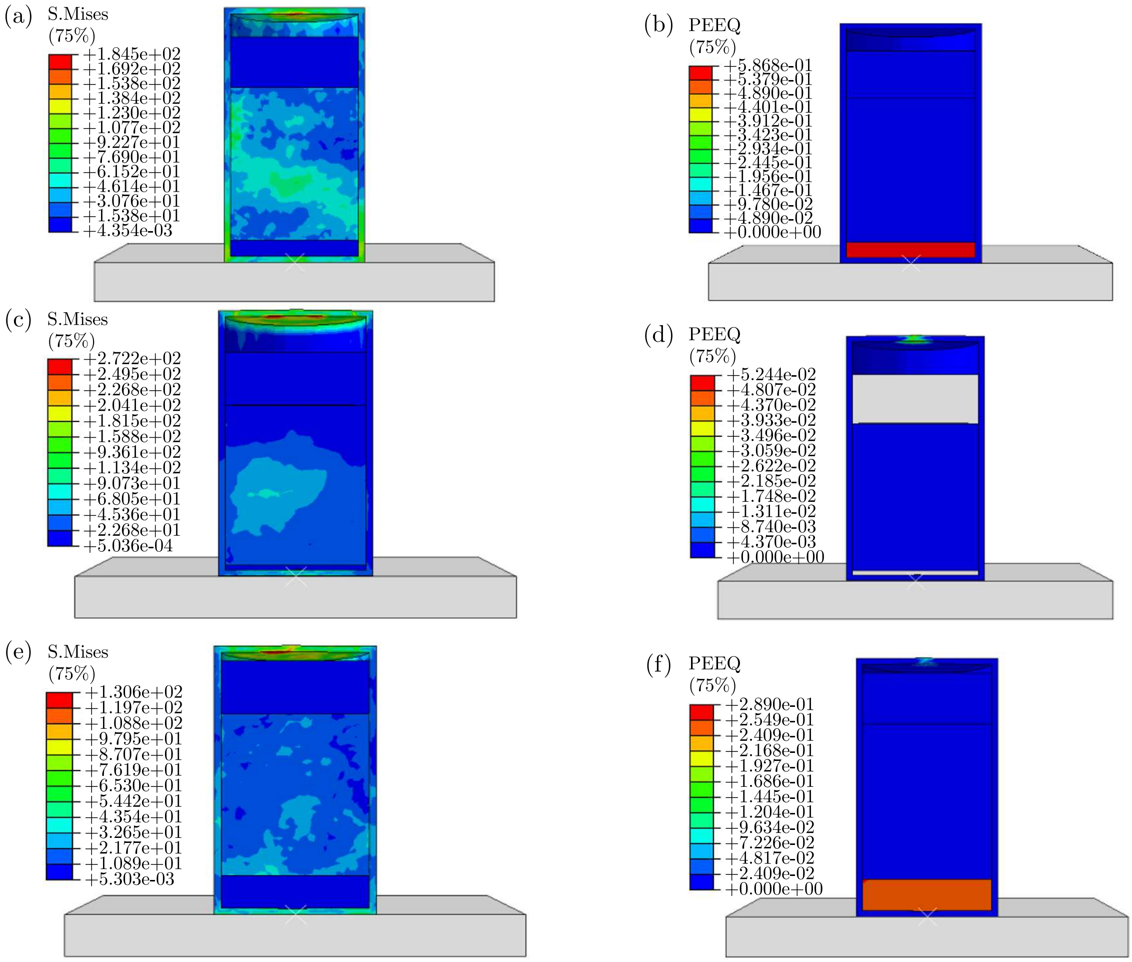


Fig. 6. Von Mises stress and equivalent strain of the packaging container: (a) EPS stress, (b) EPS plastic strain, (c) FPUF stress, (d) FPUF plastic strain, (e) RPUF stress, (f) RPUF plastic strain

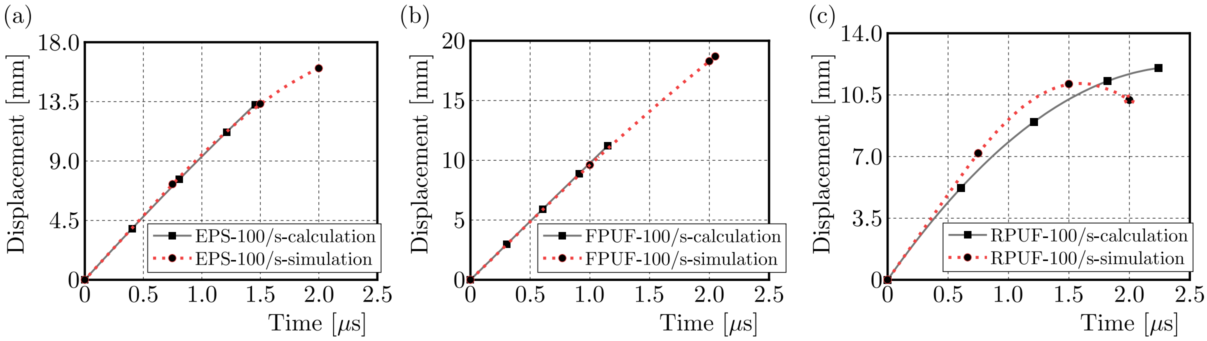


Fig. 7. Displacement-time curve comparison of the protected product by theoretical analysis and simulation: (a) EPS, (b) FPUF, (c) RPUF

For the RPUF cushioning material, the theoretical result is not consistent well with that of simulation in the large deformation stage. An assumption of the cushioning material with constant platform stress until reaching the densification stage is made. It means that elastic energy absorption is small and can be ignored comparing to the plastic deformation energy. Being different from EPS and FPUF, the elastic energy absorption of RPUF is relatively larger. It induces a discrepancy between the theoretical analysis and numerical simulation, which is increases with growth of the deformation.

4. Influence of the strain rate on the impact protection effectiveness under the lateral constraint condition

Polymer foam materials are usually sensitive to the strain rate (Hwang *et al.*, 2020; Tateyama *et al.*, 2016). In order to investigate the influence of the strain rate on the impact protection effectiveness of the cushioning material under the lateral constraint condition, packaging container (see Fig. 3) with initial impact velocity 9.8 m/s is analysed. During the packaging container impact process, the strain rate of the cushion layer is taken as constant. The mechanical properties of cushioning materials under different strain rate conditions are substituted into the theoretical model. The effect of strain rate on the impact protection effectiveness of cushioning materials under the lateral constraint condition is analysed. The stresses and compaction strains of EPS, FPUF and RPUF at strain rates of 1/s, 10/s, and 100/s are cited from the preliminary work (Zhang *et al.*, 2022, 2023). The parameters I, J, K of the three cushioning materials are calculated by Eqs. (2.21), see Table 4.

Table 4. Cushioning material design parameters at different strain rate conditions

Material	Density [kg/m^3]	Platform stress [MPa]	Compaction strain	I	J	K
EPS-1/s	43.5	0.415	0.660	273902	43.5	402.633
EPS-10/s	43.5	0.466	0.660	307291	43.5	402.633
EPS-100/s	43.5	0.566	0.660	373390	43.5	402.633
FPUF-1/s	77.8	0.025	0.560	13744	77.8	342.012
FPUF-10/s	77.8	0.034	0.560	19125	77.8	342.012
FPUF-100/s	77.8	0.045	0.560	25431	77.8	342.012
RPUF-1/s	125.6	2.029	0.630	1278195	125.6	385.366
RPUF-10/s	125.6	2.256	0.630	1421101	125.6	385.366
RPUF100/s	125.6	2.47	0.630	1522517	125.6	385.366

The energy absorbing capacity of the cushioning material and the maximum acceleration of the protected product are applied to evaluate the impact protection effectiveness of the packaging container. The acceleration of the protected product is equivalent to the stress propagating to the protected product. For the protected product safety, the stress is not allowed to exceed the bearing limit of the product material. A greater energy absorption of the cushioning material induces a lower acceleration of the protected product, which brings a better impact protection of the packaging container.

Figure 8 shows the displacement-time and velocity-time curves of the protected product found by theoretical calculations. The residual velocity of the protected product decreases with the increasing strain rate at the three strain rate conditions. It indicates that a high strain rate enhances the energy absorbtion of the three cushioning materials under lateral constraint conditions. Comparing with EPS and FPUF, the velocity of the protected products in RPUF containers is lower. It means that the EPS and FPUF cushioning layers are fully compacted, and the RPUF cushioning layer is partly compacted in the impact condition.

Figure 9 shows the maximum acceleration of the protected product during the packaging container impacting process. It is evident that a higher strain rate results in a higher acceleration of the protected product. The cushioning layer exhibits greater energy absorption at higher strain rate conditions. However, a higher strain rate loading brings leads to a higher stress of the protected product. The energy absorption and allowable stress should be considered designing in the packaging container.

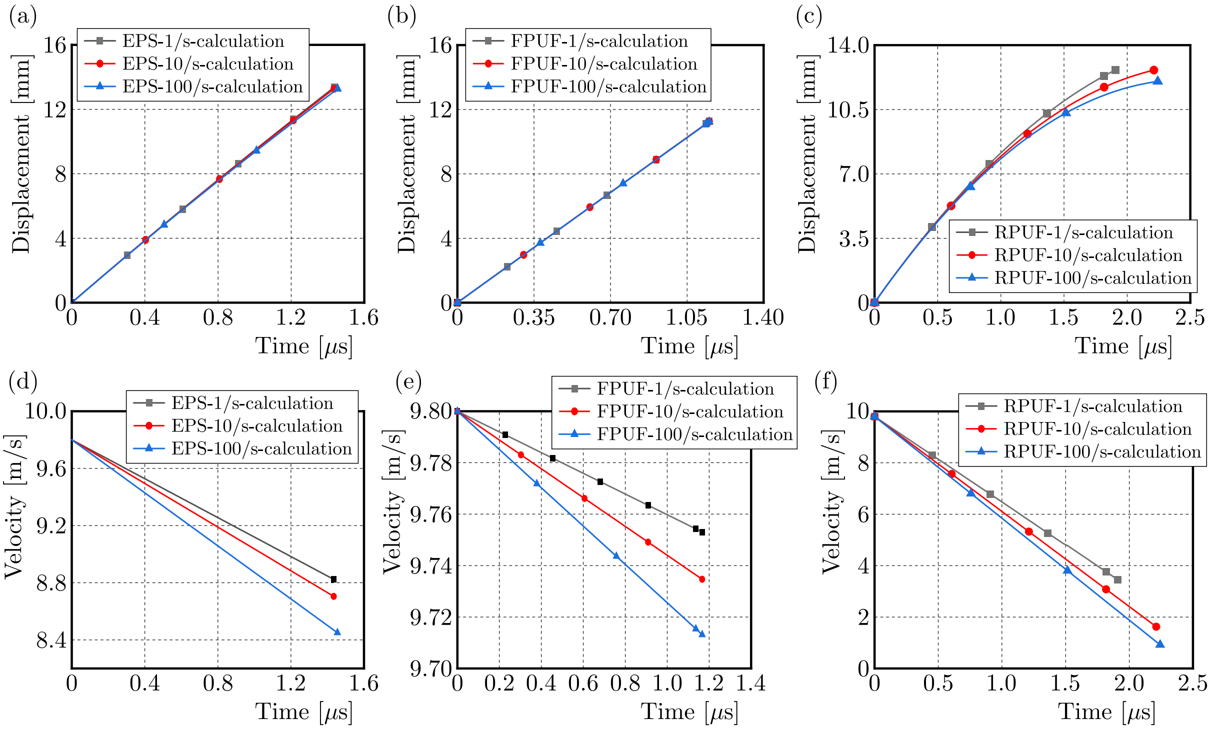


Fig. 8. Theoretical results for the packaging container: (a) EPS displacement, (b) FPUF displacement, (c) RPUF displacement, (d) EPS velocity, (e) FPUF velocity, (f) RPUF velocity

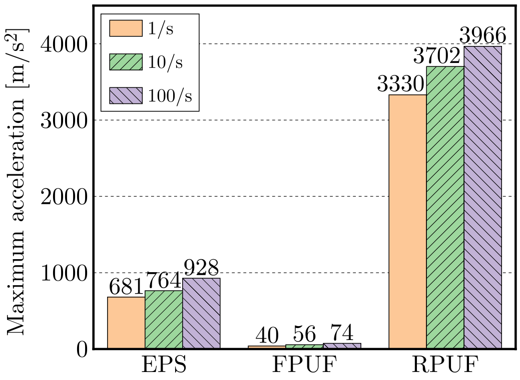


Fig. 9. Maximum acceleration of the protected product

5. Finding the cushioning material thickness in the packaging container

Once the cushioning material in the packaging container is compressed into the densification state, the internal space of the cushioning material is filled, and the stress increases sharply. For a high performance cushioning material, the platform stress should be stable under a large deformation condition. According to Eq. (2.20), the acceleration of the protected product in the packaging container increases with growing displacement. For ideal cushioning design, the cushioning material should be with little compaction rather than full compaction when the protected product reaches the maximum displacement. The maximum displacement is given as

$$u_{max} = \varepsilon_D L_0 \tag{5.1}$$

Taking Eq. (2.20) and Eq. (5.1) into account, the maximum acceleration a_{max} of the protected product is expressed as

$$a_{max} = -\frac{I}{K - Ju_{max}} = -\frac{\sigma_p \varepsilon_D}{m\varepsilon_D + \rho_0 L_0 \varepsilon_D - \rho_0 \varepsilon_D L_0} = -\frac{A\sigma_p}{m_b} \tag{5.2}$$

When the maximum acceleration is less than the allowance value of the protected product, the thickness of the cushioning material is given as follows

$$T = C_{min}\frac{Hg}{a_{max}} = C_{min}\frac{Hgm_b}{A\sigma_p} \tag{5.3}$$

C_{min} represents the minimum cushioning coefficient, H is the allowance drop height for the packaging container, and g is the gravity acceleration. Table 5 shows the minimum cushioning coefficients of EPS, FPUF and RPUF at strain rates of 1/s, 10/s and 100/s obtained from experimental tests (Zhang *et al.*, 2022, 2023). Based on the protected product (see Fig. 3, the mass is 2.45 kg and the section area is $4069\,\mathrm{mm}^2$) and the platform stresses in polymer foam materials (Table 3), the minimum cushioning coefficients of polymer foam materials are presented in Table 5. The theoretical minimum thickness of the cushioning materials under the impact condition of 9.8 m/s is listed in Table 6. It is contributed to optimize the cushioning layer in the packaging container and to enhance the impact protection performance.

Table 5. Minimum cushioning coefficient of the cushioning material

Strain rate	1/s	10/s	100/s
Minimum cushioning coefficient for EPS	2.323	2.429	2.369
Minimum cushioning coefficient for FPUF	3.324	3.239	3.056
Minimum cushioning coefficient for RPUF	2.380	2.369	2.215

Table 6. Theoretical design thickness of the cushioning material

Strain rate	1/s	10/s	100/s
Thickness for EPS [mm]	167	156	125
Thickness for FPUF [mm]	4042	2831	2009
Thickness for RPUF [mm]	35	31	27

6. Conclusion

- Based on the stress wave theory and energy conservation law, a theoretical model for packaging containers during the drop impact is proposed. The acceleration-displacement equation of the protected product is derived. It indicates that the product acceleration is affected by the cushioning material density, compaction strain, and platform stress.
- The displacement-time variation of the protected product in packaging containers filled with cushioning materials EPS, FPUF and RPUF, respectively, is obtained by theoretical analysis and numerical simulations. The two approaches achieve consistent results.
- Combining the acceleration-displacement equation of the protected product and the mechanical properties of EPS, FPUF, RPUF under different strain rates, the influence of the strain rate on the impact protection effectiveness of cushioning materials under the lateral constraint condition is analysed. It shows that a high strain rate brings container's improvement of the ultimate bearing capacity of the container, but induces higher stress in the protected product.

- The formula for thickness determination of the cushioning material is derived. The optimum thickness of the cushioning material is related with impact velocity, bearing area, and protected product mass.

Acknowledgement

The authors gratefully acknowledge the funding by National Natural Science Foundation of China under contract No.12172344.

References

1. An X., Shi D., 2022, An innovation active set strategy reliability optimization method for cushioning design based on dynamic stress-strain curve, *Packaging Technology and Science*, **35**, 2, 153-162
2. Baertsch F., Ameli A., Mayer T., 2021, Finite-element modeling and optimization of 3D-printed auxetic reentrant structures with stiffness gradient under low-velocity impact, *Journal of Engineering Mechanics*, **147**, 7, DOI: 10.1061/(ASCE)EM.1943-7889.0001923
3. Baroutaji A., Gilchrist M.D., Olabi A.G., 2016, Quasi-static, impact and energy absorption of internally nested tubes subjected to lateral loading, *Thin-Walled Structures*, **98**, Part B, 337-350
4. Chang J., Gong X., Sun Z., 2017, Analysis of mechanical properties of glass packaging, advanced graphic communications and media technologies, [In:] *Advanced Graphic Communications and Media Technologies, PPMT 2016, Lecture Notes in Electrical Engineering*, Zhao P., Ouyang Y., Xu M., Yang L., Ouyang Y. (Eds.), **417**, Springer, Singapore, 635-641
5. Estrada Q., Szwedowicz D., Majewski T., Oliver M., Cortes C,. Castro F., 2017, Effect of discontinuity size on the energy absorption of structural steel beam profiles, *Mechanics of Advanced Materials and Structures*, **24**, 1, 88-94
6. Guo Y., Han X., Wang X., Fu Y., Xia R., 2021, Static cushioning energy absorption of paper composite sandwich structures with corrugation and honeycomb cores, *Journal of Sandwich Structures and Materials*, **23**, 4, 1347-1365
7. Gupta N., 2007, A functionally graded syntactic foam material for high energy absorption under compression, *Materials Letters*, **61**, 4-5, 979-982
8. Hussain N.N., Regalla S.P., Rao Y.V. D., Dirgantara T., Gunawan L., Jusuf A., 2021, Drop-weight impact testing for the study of energy absorption in automobile crash boxes made of composite material, *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, **235**, 1, 114-130
9. Hwang B.K., Kim S.K., Kim J.H., Kim J.D., Lee J.M., 2020, Dynamic compressive behavior of rigid polyurethane foam with various densities under different temperatures, *International Journal of Mechanical Sciences*, **180**, 105657
10. Kader M.A., Islam M.A., Hazell P.J., Escobedo J.P., Saadatfar M., *et al.*, 2016, Modelling and characterization of cell collapse in aluminium foams during dynamic loading, *International Journal of Impact Engineering*, **96**, 78-88
11. Karagiozova D., 2011, Velocity attenuation and force transfer by a single- and double-layer claddings made of foam materials, *International Journal of Protective Structures*, **2**, 4, 417-437
12. Koohbor B., Kidane A., 2016, Design optimization of continuously and discretely graded foam materials for efficient energy absorption, *Materials and Design*, **102**, 151-161
13. Luong V.D., Bonnin A.S., Abbès F., Nolot J.B., Erre D., Abbès B., 2021, Finite element and experimental investigation on the effect of repetitive shock in corrugated cardboard packaging, *Journal of Applied and Computational Mechanics*, **7**, 2, 820-830
14. Lye S.W., Lee S.G., Chew B.H., 2004, Virtual design and testing of protective packaging buffers, *Computers in Industry*, **54**, 2, 209-221

15. Meng Y., Lin Y., Zhang Y., Li X., 2020, Study on the dynamic response of combined honeycomb structure under blast loading, *Thin-Walled Structures*, **157**, 107082
16. Pan Y., Yu K., 2022, Simulation analysis of packaging box for free fall airdrop, *2022 International Conference on Applied Physics and Computing (ICAPC)*, Ottawa, Canada, 374-378
17. Reid S.R., Peng C., 1997, Dynamic uniaxial crushing of wood, *International Journal of Impact Engineering*, **19**, 5-6, 531-570
18. Sek M.A., Rouillard V., 2006, Behaviour of multi-layered corrugated paperboard cushioning systems under impact loads, *Applied Mechanics and Materials*, **3-4**, 383-390
19. Tateyama K., Yamada H., Ogasawara N., 2016, Effect of strain rate on compressive properties of foamed polyethylene film, *Polymer Testing*, **52**, 54-62
20. Zeng H.B., Pattofatto S., Zhao H., Girard Y., Fascio V., 2010, Impact behaviour of hollow sphere agglomerates with density gradient, *International Journal of Mechanical Sciences*, **52**, 5, 680-688
21. Zhang G.Q., Huang J., Lin H.B., 2011, Dynamical simulation about buffer packaging of computer, *Applied Mechanics and Materials*, **55-57**, 1711-1724
22. Zhang Z., Zhong W., Chen J., Luo, J., 2022, Compressive properties and energy absorption of rigid polyurethane foam, *Journal of Physics: Conference Series*, **2368**, 012013
23. Zhang Z., Zhong W., Chen J., Luo J., Li J., Huang X., 2023, Influence of strain rate and temperature on compressive properties and energy absorption efficiency of expanded polystyrene and flexible polyurethane foam, *Mechanika*, **29**, 6, 445-453

Manuscript received July 7, 2024; accepted for publication October 11, 2024

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 27-36, Warsaw 2025
https://doi.org/10.15632/jtam-pl/194991

TACTICAL AND STRATEGIC AIR TRAFFIC SEQUENCING WITH MINIMUM-FUEL TRAJECTORIES

Adrian Pawełek, Piotr Lichota
Warsaw University of Technology, Institute of Aeronautics and Applied Mechanics, Poland
e-mail: adrian.pawelek.dokt@pw.edu.pl, piotr.lichota@pw.edu.pl

The constantly growing air traffic has a negative impact on the environment, flight safety, and the staff workload. The solution to these problems might be new techniques of air traffic management, especially automatic sequencing of the arriving aircraft combined with optimal flight trajectories. This work aims to analyze the feasibility and optimality of automatic air traffic sequencing by using MILP and large datasets of real air traffic data. The results show that the appropriate formulation of the problem may lead to both suboptimal solutions for tactical planning purposes and optimal solutions for strategic and pre-tactical planning purposes.

Keywords: air traffic, separations, flight dynamics, optimization, sequencing

1. Introduction

The problem of determining the required time of arrival (RTA) value for flights using continuous descent operations (CDO) has been the subject of research, among others in (Park and Clark 2015), and in (Takeichi, 2017) where the nominal flight time was optimized in terms of arrival time planning. Both studies assessed feasible arrival windows for predetermined waypoints, assuming that the arrival time was assigned before the aircraft reached the approach position. Furthermore, the feasible controlled time of arrival (CTA) windows was assessed in (Dalmau and Prats, 2016), assuming the use of energy-neutral CDOs. This study also assumed, unlike previous studies, that the arrival time is assigned after the descent begins. In recent years, researchers' efforts have also focused on the feasibility of operations using CTA and RTA. In (Houston and Barmore, 2009) an assessment of the feasibility of arrival procedures and an assessment of the feasibility of maintaining spacing between aircraft when the RTAs were used was performed. In (de Jong and Bussink, 2017) the use of energy-neutral trajectories was investigated, taking into account the human factor. Prats *et al.* (2016) considers CDOs supported by the use of an electronic flight bag (EFB), so the human-machine interface in the context of CDO is considered. The research results showed that EFB may be a potential solution for communication regarding CDO between the aircraft crew and the air traffic control tower. Mutiple methods for optimizing a set of flights were presented in (Durand *et al.*, 2016), where metaheuristics for optimal air traffic management were described. Methods for optimizing the air route network, managing airspace and airport traffic, allocating departure times, sequencing arrivals, and detecting and resolving conflicts in the airspace were presented. In the case of the arrivals sequencing problem, the need for separations management was highlighted and the problem was formulated in a general form. Also, an overview of possible ways to solve the problem was also presented.

By analyzing the literature, it can be noted that the previous research was aimed at developing theoretical foundations and confirming the correctness and feasibility of the concept and algorithms, but did not focus on using them on a full scale, i.e. using energy-efficient continuous descent trajectories along with assigning the required arrival time to each flight. Therefore, the full-scale feasibility and optimality of these concepts are unknown.

In this work, energy-efficient trajectories using only the idle engine thrust and elevator control, without the use of air brakes, were considered. These trajectories assigned to each flight allow for the determination of time windows using the RTA concept. The resulting time windows were combined with a large set of real arrival traffic data which were input to the aircraft sequencing problem formulated and solved with use of mixed integer linear programming (MILP) techniques.

2. Methodology

2.1. Aircraft sequencing problem

The main goal of the aircraft sequencing problem is to formulate a mathematical problem, the solution of which makes it possible to plan a set of flights in such a way that each aircraft is able to maintain safe separation from other aircraft at a selected navigation point. The problem must also take into account physical limitations, in the form of time windows in which the aircraft can be at a given point. Moreover, to ensure the optimality of the solution, a cost function must be defined that will allow to determine the optimality criterion.

Sequencing methods for aircraft landings had previously been used successfully (Beasley *et al.*, 2000; Briskorn and Stolletz, 2014; Kim *et al.*, 2014; Faye, 2015), but they did not address the concepts of future air traffic management and proposed assigning arrivals to different runways as a conflict resolution measure. In addition, the test scenarios were simplified, differed from reality, and considered a small number of days, usually with low traffic, where the goal was often to test the performance of the algorithm rather than the feasibility of the concept. It is worth noting that for the vast majority of airports, two-runway or three-runway, the issue of landing on multiple runways is not an optimal solution, especially in the contexts of maximizing throughput and minimizing operational complexity. The greatest operational efficiency can be achieved by continuous arrivals to one runway and continuous departures from the other runway. The exception are the largest airports, where dedicated operational methods must be developed (Kim *et al.*, 2014).

In this work, the feasibility of the aircraft sequencing task was examined taking into account the use of a combination of several futuristic concepts: CDO, RTA and time-based separations (TBS) based on the RECAT-EU concepts. Due to the objectives of the task, the following values were defined as the task parameters:

- ETA is the expected arrival time; this is the optimal arrival time obtained for the minimum-fuel trajectory using the energy-neutral CDO method;
- RTA^E is the earliest possible required time of arrival using the energy-neutral CDO method;
- RTA^L is the latest possible required time of arrival using the energy-neutral CDO method;
- $\left\langle \mathrm{RTA}^E, \mathrm{RTA}^L \right\rangle$ is the time window;
- t_{sep} is the minimum required time separation between a preceding and the following aircraft.

Additionally, a decision variable that defines the arrival time obtained after solving the aircraft sequencing problem, the required time of arrival, was defined as RTA. Sets are also defined:

- $\mathcal{A}$ is the set of all arrivals on a given day, where p is a single arrival, $p \in \mathcal{A}$;
- $\mathcal{P}$ is the set of all possible pairs of arrivals on a given day, where (p, q) is a single pair of arrivals, $(p, q) \in \mathcal{P}$, $p \neq q$.

As an optimization criterion, it was assumed that for each aircraft $p \in \mathcal{A}$ the required arrival time RTA should be as close as possible to the expected arrival time ETA. Moreover, it was assumed that the cost factor for the difference between the required and expected arrival time was linear and the same for all aircraft. Therefore, the optimization problem can be defined as

$$
\begin{aligned}
&\min \sum_{p\in\mathcal{A}} (t_p^+ + t_p^-)\\
&\mathrm{RTA}_p^E \leqslant \mathrm{RTA}_p \leqslant \mathrm{RTA}_p^L \qquad \forall p \in \mathcal{A}\\
&\mathrm{RTA}_p^E + t_{sep} \leqslant \mathrm{RTA}_q + L_{p,q}B_{p,q} \qquad \forall (p,q) \in \mathcal{P} \qquad \text{if } p > q\\
&\mathrm{RTA}_p^E + t_{sep} \leqslant \mathrm{RTA}_q + L_{p,q}(1 - B_{p,q}) \qquad \forall (p,q) \in \mathcal{P} \qquad \text{if } p < q\\
&B_{p,q} \in \{0,1\} \qquad \forall (p,q) \in \mathcal{P} \qquad \text{if } p < q\\
&\mathrm{RTA}_p - \mathrm{ETA}_p = t_p^+ + t_p^- \qquad \forall p \in \mathcal{A}\\
&t_p^+ \geqslant 0 \qquad \forall p \in \mathcal{A}\\
&t_p^- \geqslant 0 \qquad \forall p \in \mathcal{A}
\end{aligned}
\tag{2.1}
$$

Additional dependencies

$$
L_{p,q} = \max\{\mathrm{RTA}_p^L - \mathrm{RTA}_q^E, \mathrm{RTA}_q^L - \mathrm{RTA}_p^E\} \quad \forall (p,q) \in \mathcal{P} \quad \text{if } p > q \tag{2.2}
$$

2.2. Aircraft dynamics model in air traffic control

A set of ordinary differential equations can be used to describe the dynamics of aircraft flight (Etkin, 2005). These equations are non-linear and describe various physical phenomena that affect the dynamics of the aircraft, e.g. the influence of elasticity. However, deformations mainly concern the aircraft with significant aspect ratio, e.g. gliders. Therefore, to describe the flight dynamics of a passenger aircraft, a rigid body model with six degrees of freedom (DOF) is usually used (Fischenberg *et al.*, 2012; Seren *et al.*, 2006), making additional assumptions about geometric and mass symmetry in the vertical plane Oxz.

The plane of symmetry of geometry and mass and the assumption of a constant value of gravitational acceleration also enable the decomposition of the equations of motion into two sets: longitudinal and lateral-directional. The aircraft descends in the vertical direction, so only the longitudinal plane can be considered. To determine the aircraft condition for air traffic modeling and sequencing purposes, including calculation and analysis of the flight trajectory, it is necessary to determine such quantities as the route length, height, speed and angle of the flight path. The distance from the target point and the altitude along the flight path are used to determine the aircraft position on the descent path, and the speed and angle of the flight path are used by ATC to estimate how the position on the descent path will change before an update on the aircraft position is received. This means that the spatial orientation of the aircraft is not used, and therefore the equations of motion can be further simplified by adopting a two-dimensional point-mass model with two degrees of freedom, which are the speed and flight path angle (Stengel, 2004). This approach is often used in air traffic control problems, including aircraft sequencing (Vilardaga and Prats, 2015; Dalmau and Prats, 2015). The use of more complex flight dynamics models that would use additional physical quantities, the dynamics of system components (e.g. control system) or the influence of environmental factors would have a minor impact on accuracy. However, the calculation of optimal trajectories would be much longer, and finding a solution in a reasonable time scale would not be possible (Lichota and Ohme, 2014). With the adopted assumptions, the aircraft dynamics model can be expressed as follows

$$
m\dot{U} = F_T - F_D - mg\sin\gamma \tag{2.3}
$$

where F_T is the thrust force, m is mass of the aircraft, F_D is the aerodynamic drag force, g is the acceleration due to gravity, U is the longitudinal velocity, γ is the flight path angle.

2.3. Optimization problem parameters

In this work, a simplifying assumption was made, which consisted in considering one type of aircraft – the Airbus A320. The analyzes showed that approximately 80% of the flights in the dataset used were operated by Airbus A320 (ICAO: A320), Boeing 737-800 (ICAO: B738), Airbus A321 (ICAO: A321) and Airbus A319 (ICAO: A319). These aircraft have similar mass and geometric parameters, and therefore also dynamic properties. All of them are also in category D according to RECAT-EU. Therefore, a uniform time separation of $t_{sep} = 80\,\mathrm{s}$ was assumed.

The time parameters of energy-neutral trajectories were also determined for the Airbus A320 aircraft. For this purpose, Eq. (3) was used because the wing aspect ratio of the Airbus A320 is 10.47 (Airbus, 2016), and therefore the stiffness of the aircraft structure can be neglected. To obtain energy-neutral trajectories, it was assumed that the descent trajectory is controlled only by the flight path angle (in fact, this would be the result of using a control surface, e.g. elevator), and mass changes were neglected due to the fact that the mass change as a result of fuel consumption during descent at idle thrust is negligibly small compared to the weight of the aircraft.

Time windows and trajectories with the minimum fuel consumption as a function of the distance from the destination point for the Airbus A320 aircraft were obtained from the previous studies (Pawełek *et al.*, 2017, 2019). Values for the aircraft aerodynamic and propulsion variables were obtained using accurate performance data received from the manufacturer. Typically, these data are obtained as a result of experimental studies and specified in a tabular form. The data were approximated with continuous functions using B-splines to avoid numerical problems. In addition, the international standard atmosphere model is used, which relates density, pressure and temperature to altitude. Trajectories with minimum fuel consumption were obtained by minimizing the cost function

$$J = \int_{t^0}^{t^{fix}} FF(t)\, dt \tag{2.4}$$

where t^0 and t^{fix} are the relative start and end times of the trajectory, and FF is the fuel consumption function. The earliest possible arrival time was obtained by minimizing the arrival time at the destination, so the cost function was $J = t^{fix,E}$, and for the latest arrival time it was $J = -t^{fix,L}$.

In this work, the synchronization of air traffic at the IF point was considered, and the distance of 250 NM was assumed as the moment of calculating the time window (except for shorter trajectories, where correspondingly smaller distances were assumed), as in preliminary analyzes this value was considered the target value of future TMAs (Ky, 2021). This value is also consistent with the direction of development of the E-AMAN system, where the current horizon is 200 NM, taking into account the possibility of gradual expansion up to 500 NM. Therefore, the time parameters of the aircraft sequencing task were calculated in this work as follows for each aircraft p

- expected arrival time ETA_p was calculated as

$$\mathrm{ETA}_p = T_p^0 + (t_p^{fix} - t_p^0) \tag{2.5}$$

- earliest arrival time RTA_p^E

$$\mathrm{RTA}_p^E = T_p^0 + t^{fix,E} \tag{2.6}$$

- latest arrival time RTA_p^L

$$\mathrm{RTA}_p^L = T_p^0 + t^{fix,L} \tag{2.7}$$

where T_p^0 is the time of day when the plane was 250 NM from the airport threshold. In the case of shorter trajectories, a correspondingly shorter time is used and, therefore, smaller time windows. Additionally, the climb part of the trajectory is not included because CDOs cannot be considered at this stage.

3. Experimental scenario

3.1. Air traffic data

In this work, eleven European airports located in the countries associated with the European Civil Aviation Conference (ECAC) were taken into account. For the airports selected for research, the necessary data were obtained from Aeronautical Information Publications (AIP) of each of the countries in which the selected airports are located. The airports and their respective ICAO codes are presented in the Table 1.

Table 1. Airports selected for the study

Airport	ICAO Code	Airport	ICAO Code
Brussels Airport	EBBR	Barcelona El-Prat Airport	LEBL
Manchester Airport	EGCC	Palma de Mallorca Airport	LEPA
London-Stansted Airport	EGSS	Paris-Orly Airport	LFPO
Copenhagen-Kastrup Airport	EKCH	Roma-Fiumicino Airport	LIRF
Oslo-Gardermoen Airport	ENGM	Zurich-Kloten Airport	LSZH
Warsaw Chopin Airport	EPWA		

Air traffic data were obtained from the Data Demand Repository 2 (DDR2) (EUROCONTROL, 2015) database using the Network Strategic Tool (NEST) (EUROCONTROL, 2016) software. Both DDR2 and NEST are provided by EUROCONTROL.

Current trajectories, in NEST terminology designated CTFM (Computed Traffic Flight Model) or M3, were used for the experimental scenario. M3 is the initial trajectory (last filled flight plan) updated with available radar measurements whenever the flight deviates from the last filed flight plan by more than any of the predefined thresholds: 5 minutes, 7 FL or 20 NM; this trajectory is the best available estimate of the actual flight trajectory as managed by controllers on the day of operation.

In the sequencing problem, the arrival sequence for the entire day was used because outlier observations do not affect the solution, and at the same time, using the full arrival sequence increases the size of the problem, which results in more complex scenarios with a higher traffic intensity, the analysis of which was the purpose of this work.

Time separations at the intermediate fix (IF) point in which the aircraft should already be stabilized on the ILS glide path were analyzed, so IF was the fix variable present in Eq. (2.4). The time at a given distance to airport T_p^0, used in Eqs. (2.5)–(2.7), was calculated using the DDR2 data, by summing segments length from reverse for each trajectory, and where the sum reached the desired value, a linear approximation was used to find the time of day, as passenger aircraft velocity usually changes very slowly.

3.2. Solver configuration

The MILP problem defined by Eq. (2.1) was implemented in the *IBM ILOG CPLEX Optimization Studio 12.9* optimization environment using the Optimization Programming Language (OPL) and solved using the CPLEX solver, which is intended, among others, for solving large-scale mixed-integer linear optimization problems. To solve the problem, the *branch and cut* method was used, based on the *divide and cut* method. This method involves creating a tree of subproblems and searching it to find a solution. MILP problems can have multiple solutions located in different places in the solution tree, and finding all of them can be very time-consuming. The solution of the optimization problem is completed after searching the entire solution tree, and the *divide and cut* method implemented in the CPLEX solver uses several algorithms that cut off the areas of the tree that do not provide an integer solution in order to reduce the problem, heuristic algorithms supporting the solution search process, and other methods to improve the optimization process. A description of all methods can be found in the optimization environment documentation (IBM, 2017).

The calculations were performed on a computer with 16 GB RAM and a Intel® Core^{TM} i7-7500U processor with a base clock frequency of 2.7 GHz. All processor cores were used for calculations.

4. Results

The minimization problem defined by Eq. (2.1) was solved for all days and airports considered in this work in order to assess the feasibility of the concept. Moreover, it was solved also for the scenarios with highest air traffic intensity days in order to search for the optimal solution. The optimality of a solution for such scenarios implies the optimality for scenarios with lower air traffic volumes, because as air traffic volumes increase, the airspace becomes more restricted and therefore there is a greater risk of conflict.

4.1. Feasibility analysis results

To investigate the feasibility of nearly 4000 test scenarios, containing totally over million of flights, the solver was configured so that the calculations ended when the first solution was found. Moreover, in order to be able to analyze all scenarios, it was decided to introduce a time limit of 60 seconds to find a solution. This time was additionally justified by the fact that, as noted in (Durand *et al.*, 2016), for automatic real-time air traffic management, the air situation should be updated at least every 2–3 minutes.

Table 2 presents feasibility analysis results of the arrival traffic sequencing process including CDO, RTA, TBS and RECAT-EU concepts.

Table 2. Results of feasibility analysis

Airport [ICAO Code]	Scenarios with solution [%]	Scenarios without solution [%]
EBBR	92.08	7.92
EGCC	98.63	1.37
EGSS	98.36	1.64
EKCH	99.18	0.82
ENGM	99.18	0.82
EPWA	100.00	0.00
LEBL	99.18	0.82
LEPA	99.18	0.82
LFPO	100.00	0.00
LIRF	60.93	39.07
LSZH	99.73	0.27

It can be observed that the results vary depending on the airport, but for most airports the results can be considered as very good. For EPWA and LFPO all test scenarios had a solution, for

EGCC, EGSS, EKCH, ENGM, LEBL, LEPA and LSZH 98–99% of the scenarios were feasible, and for EBBR approximately 92% of the scenarios were feasible. For LIRF, the results differ from the rest by around 61% of days with the solution. However, it should be emphasized that the results refer to the solution obtained in one minute, so it can be assumed that the results would be better for a longer calculation time. Moreover, it should be noted that this work examined scenarios based on past air traffic in which the planning stages did not take into account new concepts. Quick calculations of feasible solutions can be beneficial at the tactical planning stage, as aforementioned, the air situation should be updated at least every 2–3 minutes, but ideally, to maintain highest levels of safety, it should be performed as frequently as possible. In the future, by solving the air traffic sequencing task in the earlier planning stages, i.e. strategic and pre-tactical, it is expected that tactical scheduling will be used for real-time air traffic management due to weather conditions, delays or other events which were not considered in the sequences planned in advance.

4.2. Optimality analysis

In order to find optimal or better suboptimal solutions, the aircraft sequencing problem was solved with a computation time limit of 21600 s (6 h) and without a limit of the solutions number. Also, the optimality criterion was defined as a relative gap value of less than 0.01%. The relative gap for MILP tasks is defined as the relative difference between the best integer solution found and the best solution to the problem if it were a linear optimization problem (IBM, 2017). For each airport, one day with a very high traffic scenario was selected for the analysis.

The results of the solution obtained after 21600 s of calculations and comparison with the first solution obtained in the feasibility analysis are presented in Table 3, which shows the values of the cost function obtained according to Eq. (2.1) for both cases. For the solution obtained after 21600 s, the value of the cost function which corresponds to the best solution found is presented. The gap value is also presented for this solution. Additionally to the data provided in the table: for EGSS, the optimization problem was solved in 8045 s and for EPWA, the optimization problem was solved in 2 s.

Table 3. Results of the optimality analysis

Scenario	First solution	Solution after 21600 s		
	Cost function value [s]	Number of solutions	Smallest cost function value [s]	Smallest gap [%]
EBBR	31341	38	15164	59.54
EGCC	107594	29	4946	3.17
EGSS	88201	28	3525	<0.01
EKCH	118218	50	8503	31.37
ENGM	108102	41	9057	33.07
EPWA	78924	18	2411	0.79
LEBL	132412	81	16138	51.44
LEPA	159813	68	12269	38.10
LFPO	124099	41	10357	45.26
LIRF	43606	48	14967	54.49
LSZH	119135	54	12673	53.59

By analyzing the results presented in Table 3, it can be seen that the first solution, although obtained very quickly, resulted in most cases in the value of the cost function several times or even several dozen times higher than in the case of a longer optimization time. The smallest relative difference in the cost function values occurred for EBBR and LIRF, where six-hour calculations allowed for obtaining results that were approximately 2 times and 3 times smaller, respectively. The largest relative difference in the value of the cost function occurred in the scenarios for EPWA, EGSS and EGCC, where the results were approximately 33 times, 25 times and 22 times better, respectively, than for the first solution.

It is worth noting that optimal or close to optimal solutions were obtained for EGCC and EPWA scenarios, hence the significant improvement in results. For the remaining scenarios, the solution obtained after six hours is suboptimal, as evidenced by the gap value, but despite this, the obtained solutions were about 10 times better. It should be emphasized here that the gap value of several dozen percent does not always mean that a significant change in the best result is still possible. In some cases, this may mean that there is still a significant part of the solution tree to search, but not necessarily containing feasible solutions to the MILP problem. Therefore, some of the presented solutions may be optimal solutions, but this is not known until a solution is obtained based on criteria other than time.

The significant amount of time needed to search for optimal solutions means that this activity can be carried out at the strategic and pre-tactical planning stages. It is worth noting that properly performed analyzes at the strategic and pre-tactical planning stages may result in tactical planning with solutions close to the solution with the lowest possible value of the cost function. In addition to the value of the cost function, attention should be focused on the number of solutions. For all scenarios, from several to several dozen of solutions were obtained. It was observed that a large part of the solutions were found in the first minutes of calculations, and over time the number of solutions found decreased, however, in some cases, after a long period without new solutions, there was a sudden increase in the number of solutions at various moments of calculations. This was due to the fact that the solver found a region of the tree that had combinations of values which allowed obtaining feasible arrival sequences. Multiple solutions with different characteristics located in different areas of the tree can be an operational advantage. Some solutions may be unsatisfactory, e.g. from the perspective of airline operators or airport operators. Obtaining several dozen solutions for scenarios with a high air traffic intensity makes it possible to consider extending the optimization task with additional constraints to take into account other air traffic management needs.

5. Conclusions

A method for automatic sequencing of arriving air traffic was presented, taking into account continuous descent trajectory techniques, required time of arrival and time-based separations. A feasibility study was carried out using an extensive set of real air traffic data. An optimality analysis was also performed for scenarios with the highest air traffic intensity.

Based on the analyzes performed and the results obtained, conclusions can be drawn. Firstly, in the vast majority of cases, it is possible to solve the problem of sequencing of the arrival traffic in the case of simultaneous use of CDOs, RTA, TBS and initiation of descent trajectory control at a distance of 250 NM from the destination point, through appropriate formulation and solution of a mixed integer linear optimization problem. Secondly, automatic sequencing of arrival traffic may have practical applications in air traffic control, where optimal sequences may be obtained at the strategic and pre-tactical planning stages, and suboptimal sequences may be obtained at the tactical planning stage. Finally, the use of time-based separations creates a room for the development of new robust methods for air traffic control.

Future research may aim to use dynamic models of larger numbers of aircraft to more accurately reflect the airspace situation. In order to assess the feasibility of methods in test scenarios, one target navigation point was considered, in which full air traffic synchronization should be achieved. The methods can be extended to many points on the descent trajectory, which will lead to more control of the situation in the airspace. The future research may also focus on extending the presented methods to include the influence of more factors, e.g. weather conditions.

References

1. Airbus, 2016, *A320 Aircraft Characteristics Airport and Maintenance Planning*, Airbus, Blagnac Cedex, France
2. Beasley J., Krishnamoorthy M., Sharaiha Y.M., Abramson D., 2000, Scheduling Aircraft Landings – The Static Case, *Transportation Science*, **34**, 2, 180-197
3. Briskorn D., Stolletz R., 2014, Aircraft landing problems with aircraft classes, *Journal of Scheduling*, **17**, 1, 31-45
4. Dalmau R., Prats X., 2015, Fuel and time savings by flying continuous cruise climbs: Estimating the benefit pools for maximum range operations, *Transportation Research Part D: Transport and Environment*, **35**, 62-71
5. Dalmau R., Prats X., 2016, Assessment of the feasible CTA windows for efficient spacing with energy-neutral CDO, *Proceedings of the ICRAT 7th International Conference on Research in Air Transportation*, Philadelphia, USA
6. de Jong P.M.A., Bussink F.J.L., Verhoeven R.P.M., de Gelder N., van Paassen M.M., Mulder M., 2017, Time and energy management during approach: A human-in-the-loop study, *Journal of Aircraft*, **54**, 1, 177-189
7. Durand N., Gianazza D., Gotteland J.B., Alliot J.M., 2016, *Metaheuristics for Air Traffic Management*, Wiley, Hoboken, NJ, USA
8. Etkin B., 2005, *Dynamics of Atmospheric Flight*, Dover Publications, Mineola, USA
9. EUROCONTROL, 2015, *DDR2 Reference Manual For Generic Users 2.1.2*, EUROCONTROL, Brussels, Belgium
10. EUROCONTROL, 2016, *NEST User Guide 1.5*, EUROCONTROL, Brussels, Belgium
11. Faye A., 2015, Solving the Aircraft Landing Problem with time discretization approach, *European Journal of Operational Research*, **242**, 3, 1028-1038
12. Fischenberg D., Mönnich W., Krag B., Jategaonkar R.V., 2012, Aspects of C-160 simulator model determination and validation on and close to the ground, *AIAA Flight Simulation Technologies Conference*, AIAA Paper 1994–3404, 22-31
13. Houston V.E., Barmore B., 2009, An exploratory study of runway arrival procedures: Time-based arrival and self-spacing, *9th AIAA Aviation Technology, Integration, and Operation Conference (ATIO)*, AIAA Paper 2009–7005
14. IBM, 2017, *IBM ILOG CPLEX Optimization Studio CPLEX User's Manual Version 12 Release 8*, IBM, Armonk, NY, USA
15. Kim B., Li L., Clarke J.P., 2014, Runway assignments that minimize terminal airspace and airport surface emissions, *Journal of Guidance Control and Dynamics*, **37**, 3, 789-798
16. Ky P., 2012, SESAR Release 1 Results, SESAR Joint Undertaking, Technical report, Brussels, Belgium
17. Lichota P., Ohme P., 2014, Design and analysis of new multi axis input manoeuvres for aircraft sys-ID, DLR, German Aerospace Center, Technical Report IB 111-2014/46, Braunschweig, Germany
18. Park S.G., Clarke J.P., 2015, Optimal control based vertical trajectory determination for Continuous Descent Arrival procedures, *Journal of Aircraft*, **52**, 5, 1469-1480

19. Pawełek A., Dalmau R., Lichota P., Prats X., 2017, Arrival traffic synchronisation with Required Time of Arrivals for fuel-efficient trajectories, [In:] A. Nijo, *et al.*, (Eds.), *17th AIAA Aviation Technology, Integration, and Operations Conference*, Denver, USA

20. Pawełek A., Lichota P., Dalmau R., Prats X., 2019, Fuel-efficient trajectories traffic synchronization, *Journal of Aircraft*, **56**, 2, 481-492

21. Prats X., Bendris B., Dalmau R., Montolio J., Day B., *et al.*, 2016, 4D continuous descent operations supported by an electronic flight bag: A human-in-the-loop study, *35th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, Sacramento, USA

22. Seren C., Bommier F., Bucharles A., Verdier L., Alazard D., 2006, Flight test protocol optimization using genetic algorithms, *14th IFAC Symposium on Identification and System Parameter Estimation, IFAC Proceedings Volumes*, **39**, 1, 642–647

23. Stengel F., 2004, *Flight Dynamics*, Princeton University Press, Princeton, NJ, USA

24. Takeichi N., 2017, Nominal flight time optimization for arrival time scheduling through estimation/resolution of delay accumulation, *Transportation Research Part C: Emerging Technologies*, **77**, 433-443

25. Vilardaga S., Prats X., 2015, Operating cost sensitivity to required time of arrival commands to ensure separation in optimal aircraft 4D trajectories, *Transportation Research Part C: Emerging Technologies*, **61**, 75-86

Manuscript received October 5, 2023; accepted for publication October 21, 2024

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 37-50, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195336

COMBINED ROOF INTERSTORY SHEAR ENERGY STORAGE MODEL AND ANALYSIS OF INFLUENCING FACTORS

Peng-Fei Zhang, Yan Tan, Lei Guo, Tong-Bin Zhao, Zhen-xin Wei

College of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao, China

e-mail: 15064221133@163.com (Peng-fei Zhang); tanyan.ty@qq.com (corresponding author Yan Tan);

1920365481@qq.com (Lei Guo); ztbwh2001@126.com (Tong-bin Zhao); 1484539823@qq.com (Zhen-xin Wei)

To reveal the shear deformation and energy storage mechanism of composite roof strata, and to quantify the interlayer shear energy storage characteristics of composite roof strata, this study establishes a shear energy mechanical model for layered composite roof structures under various conditions. The factors influencing the shear energy storage of rock strata are analyzed, and the energy release process as well as mechanism of layered composite roof strata are discussed, leading to the following conclusions: the shear strain energy linearly increases with an increase in bond layer thickness and quadratically increases with increasing external forces. During the bending deformation stage, when thick and hard layers exist in composite beams, shear failure dominated by interlayer shear slip occurs, and low-strength rock beams exhibit tensile crack initiation. In the overall instability stage, the composite beams mainly experience tensile fracture. When thick and hard layers exist in the composite beams, the fracture strength and released energy are higher. The presence of thick and hard layers in the roof is a key factor leading to severe structural damage and increased energy release.

Keywords: roof, interlayer shear, energy, composite beam model, influencing factors

1. Introduction

The hard roof is a type of roof above the coal seam that is characterized by strong integrity, high strength, and large thickness, accumulating a significant amount of elastic energy. It is a typical geological condition in coal mining engineering that can trigger dynamic ground pressure incidents (He *et al.*, 2012; Małkowski and Niedbalski, 2020; Tan *et al.*, 2018; Yang *et al.*, 2019). The movement and fracturing of the roof caused by coal mining lead to structural damage of the roof, releasing a large amount of stored energy, which is one of the root causes of induced dynamic disasters in mining areas. The complex and diverse characteristics of the roof above the coal seam lead to bending deformation and subsequent fracturing under the load of overlying strata. Throughout this process, the energy storage during bending deformation and the energy release during fracture of the roof structure are crucial for the prevention of dynamic ground pressure disasters (Tajdus *et al.*, 2018; Cui *et al.*, 2020; Xu *et al.*, 2021; Li *et al.*, 2022; Zhou *et al.*, 2023). Therefore, the mechanism of energy release in the destruction of hard roof structures has been an important research topic for mining scholars both domestically and internationally.

The bending and fracturing of the roof structure is a complex process influenced by various factors such as roof thickness (Lu *et al.*, 2019), physical and mechanical properties (Coggan *et al.*, 2012; Xu *et al.*, 2023), joint development characteristics (Bai and Tu, 2020), contact forms, and combined movement patterns (Shen *et al.*, 2019; Yu *et al.*, 2022). Extensive research has been conducted by scholars on the roof structure model in mining areas, proposing hypotheses such as the pressure arch hypothesis (Xia *et al.*, 2018), the "cantilever beam" hypothesis

(Han *et al.*, 2015), the hinged rock block hypothesis (Qin *et al.*, 2021), and the "natural balance arch" hypothesis (Han *et al.*, 2015), providing preliminary insights into the "zonation structure" of the roof. To further elucidate the mechanism of roof fracturing, scholars have developed mechanical structural models represented by beams (Ti *et al.*, 2021), plates (Wang *et al.*, 2022), and arch shells (Liu and Shi, 2021). For example, the "transmission rock beam", "masonry beam" structural models, and the "key layer theory" (Cao and Huang, 2021) have quantitatively solved the fracturing step distance of roof rock beams. In order to further study the coordinated movement laws of multi-layer roofs, scholars have introduced basic concepts such as the composite roof, composite key layer, and key layer group (Zhang *et al.*, 2021; Qi *et al.*, 2022; Wang *et al.*, 2023), investigating the fracturing laws of multi-layer composite roof structures under mining influence, including weakly bonded composite roofs, composite fragmentation, and multiple key layers of overlying rock (Liu *et al.*, 2021; Zhang *et al.*, 2023), revealing the composite effects and fracturing mechanisms of the roof.

In summary, valuable research findings have been obtained by domestic and foreign scholars regarding the bending and fracturing of composite roofs, including the roof structure models, fracturing mechanisms, and composite movement characteristics. However, the mechanical properties of interlayer contact surfaces in the roof significantly affect the integrity, flexural moment capacity, and fracturing distance of the composite roof. Most of the current research is limited to studying the fracture of the roof rock mass itself, while the shear instability mechanism of composite roofs has not been investigated, and a mechanical model describing the shear energy storage problem of composite roofs has not been established. This makes it difficult to quantitatively calculate the energy storage under such conditions. Therefore, this study establishes a shear energy storage mechanical model for layered composite roof structures under different conditions. From a theoretical perspective, it reveals the shear deformation and energy storage mechanisms of composite roofs and analyzes the factors influencing the shear energy storage of rock layers. This study provides a theoretical basis for the prevention and control of impact ground pressure caused by roof fracturing.

2. The stress analysis of the composite roof

In mining areas, certain relatively hard and thick rock layers are referred to as key strata (Lu *et al.*, 2020). These key strata are composed of single layers or adjacent multiple layers of hard rock, gradually forming a load-bearing structure in the mining area with the extraction of coal, playing a crucial control role in the overlying rock layers. When the key strata are formed by two layers of rock together, this load-bearing structure can be simplified as a composite beam that only bears the load of the overlying rock layers before the key strata fracture, as shown in Fig. 1.

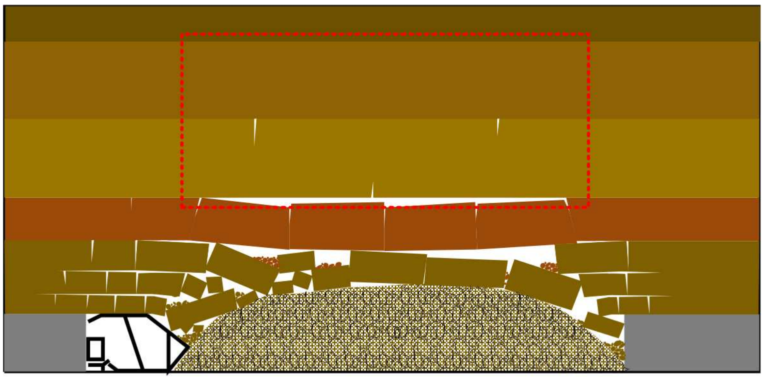


Fig. 1. Schematic diagram of the combined roof structure

Assuming the load provided by the overlying rock layers is a concentrated load, this segment of the composite beam becomes a lateral bending beam. According to the plane assumption, the normal stress on the cross-section of a rectangular section beam during bending deformation is

$$\sigma = \frac{My}{I_z} \tag{2.1}$$

where σ is the normal stress on the cross-section of a rectangular section beam, M is the equivalent bending couple exerted by the overlying rock, y is the distance from any point to the neutral axis, I_z is the moment of inertia of the rock beam.

The rock layers in the roof strata often exhibit variations in the lithology due to different geological formation periods. The interfaces between these rock layers are cemented together. Therefore, before relative sliding occurs between adjacent rock layers, the two adjacent roof layers can be treated as a double-layer composite beam structure. Assuming that the upper and lower layers of the beam consist of different lithologies with elastic moduli of E_1 and E_2, and thicknesses of h_1 and h_2, respectively, as shown in Fig. 2, the expressions for the shear force and bending moment in the beam are given by Eq. (2.2)

$$\begin{gathered}
F_S(x) = \begin{cases} \dfrac{P}{2} & 0 \leqslant x < \dfrac{S}{2} \\[2mm] -\dfrac{P}{2} & \dfrac{S}{2} \leqslant x < S \end{cases} \\[4mm]
M(x) = \begin{cases} \dfrac{Px}{2} & 0 \leqslant x < \dfrac{S}{2} \\[2mm] \dfrac{PS}{2} - \dfrac{Px}{2} & \dfrac{S}{2} \leqslant x \leqslant S \end{cases}
\end{gathered} \tag{2.2}$$

where $F_S(x)$ is the shear force distribution function in the beam, $M(x)$ is the bending moment distribution function in the beam, P is the concentrated load on the upper part of the rock beam, S is the span of the rock beam, and x is the horizontal position coordinate of any point in the beam.

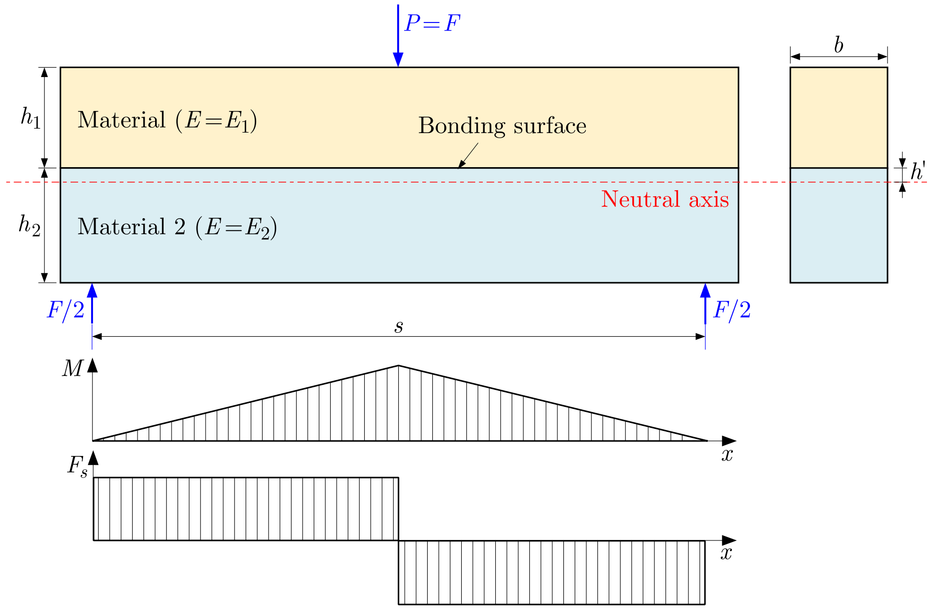


Fig. 2. Schematic diagram of force analysis of the composite beam

Based on the plane assumption, it can be inferred that the longitudinal strain at any point y along the height of the cross-section of the composite beam varies linearly. The longitudinal strain at point y is given by

$$\varepsilon = \frac{y}{\rho} \tag{2.3}$$

where ε is the longitudinal strain, ρ is the radius of curvature of the neutral axis.

3. Shear energy storage model for composite roof strata layers

3.1. Interlayer slip criterion

In order to analyze the mechanical properties at the interface of the composite beam, according to the method of equivalent sections, the cross-section of the composite beam is transformed equivalently into a beam with varying widths but made of the same material. This section is referred to as the equivalent section, as shown in Fig. 3.

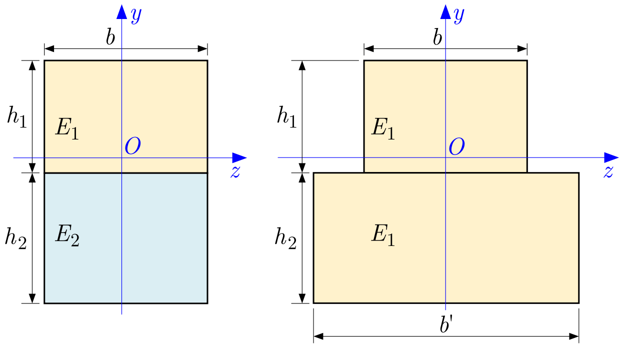


Fig. 3. Schematic diagram of the equivalent section of the composite beam

The moment of inertia of the equivalent section can be calculated as follows

$$I'_z = I_{z1} + \eta I_{z2} \tag{3.1}$$

where I'_z is the moment of inertia of the equivalent section; I_{z1} is the moment of inertia of the lower beam section, I_{z2} is the moment of inertia of the lower beam section, η is the ratio of elastic modulus between the lower and the upper beam, $\eta = E_2/E_1$.

Therefore, the static equilibrium equation for the equivalent section can be expressed as

$$\int\limits_{A_1} y_1\, dA + \int\limits_{A_2} \eta y_2\, dA = 0 \tag{3.2}$$

where y_1 is the distance from the upper beam point to the neutral axis, A_1 is the section area of the upper rock beam, y_2 is the distance from the upper beam point to the neutral axis, A_2 is the section area of the lower rock beam.

According to formula (3.2), the flexural shear stress on the upper and lower beam sections is

$$\tau_1 = \frac{F_S S_Z^*}{b I'_z} \qquad \tau_2 = \frac{\eta F_S S_Z^*}{b I'_z} \tag{3.3}$$

where b is the width of the composite beam, F_S is the shear force on the cross section, $F_S = F/2$, S_Z^* is the static moment of the area outside the cross section h' away from the neutral axis with respect to the neutral axis, $S_Z^* = b(h_1^2 - h'^2)/2$.

According to the theorem of mutual equivalence of shear stress, the shear stress at the interface of the composite beam is

$$\tau' = \tau_1 - \tau_2 = (1-\eta)\frac{F_S S_Z^*}{bI_z'} \tag{3.4}$$

where τ' is the shear stress at the interface of the composite beam, τ_1 is the flexural shear stress on the upper beam sections, τ_1 is the flexural shear stress on the lower beam sections.

In the test, the upper limit of τ' depends on the interfacial shear strength of the bond surface. In the actual stratum, τ' refers to the shear stress of the interlayer area of adjacent rock layers. If the shear strength of the cementing area is $[\tau]$, then according to the third strength theory, the criterion for shear slip of the bonding surface of the composite beam is as follows

$$\tau' \geqslant [\tau] \tag{3.5}$$

By bringing formula (3.4) into formula (3.5), the shear slip criterion of the bonded surface of the composite beam can be obtained as follows

$$P \geqslant \frac{4(I_{z1} + \eta I_{z1})[\tau]}{(1-\eta)(h_1^2 - h'^2)} \tag{3.6}$$

where h' is the distance from the boundary of the rock beam to the neutral axis, h_1 is the thickness of the upper beam.

When formula (3.6) is satisfied, the composite beam exhibits shear slip along the bonding surface between rock layers.

3.2. Interlayer shear energy storage

When composite beams are subjected to interlaminar shear stress, the stored shear strain energy mainly concentrates on the bonding surface. In the test, the energy storage area is the bonding layer between the composite beam rock layers, and in the real stratum, it is the interlayer area of the lithology change. According to the calculation formula of shear strain energy density, the shear strain energy density is

$$dU_{\varepsilon\tau} = \frac{\tau\, dx\, dz\, \gamma\, dy}{2} = \frac{\tau^2}{2G}\, dx\, dy\, dz \tag{3.7}$$

where $dU_{\varepsilon\tau}$ is the shear strain energy density stored in composite beams, τ is the shear stress of the bonded surface of the composite beam, γ is the shear strain of the bonded surface of the composite beam.

Therefore, the shear strain energy of the bonded surface of the composite beam is as follows

$$U_{\varepsilon\tau} = \iiint\limits_{v_\varepsilon} \frac{\tau^2}{2G}\, dv_\varepsilon = \int\limits_0^{h_b}\int\limits_0^{b}\int\limits_0^{S} \frac{\tau^2}{2G}\, dx\, dy\, dz \tag{3.8}$$

where $U_{\varepsilon\tau}$ is the shear strain energy stored in composite beams, v_ε is the volume of bonded surface of composite beam, h_b is the thickness of the bonding zone, G is the shear modulus of the bond.

Assuming that the shear stress on the bonding surface is uniformly distributed, the above formula can be simplified as follows

$$U_{\varepsilon\tau} = \frac{\tau^2}{2G} h_b b S = \frac{h_b b S}{2G}\Big[\frac{(1-\eta)(h_i^2 - h'^2)}{4(I_{z1} + \eta I_{z1})} P\Big]^2 \tag{3.9}$$

Formula (3.9) is the expression form of the interlayer shear energy storage model of the composite beam. When no shear slip occurs between the composite beams, the stored shear strain energy can be calculated according to Eq. (3.9). When interlayer shear slip occurs, the maximum shear strain energy stored by the composite beam is as follows

$$U_{\varepsilon\tau} = \frac{[\tau]^2}{2G} h_b b S \tag{3.10}$$

3.3. Analysis of influencing factors

In order to analyze the influence of various factors on the shear energy storage, a numerical model of composite beam FLAC3D was established. A total of 42 groups of numerical models were established. The schematic diagram of the model is shown in Fig. 4, and the simulation parameters are listed in Table 1.

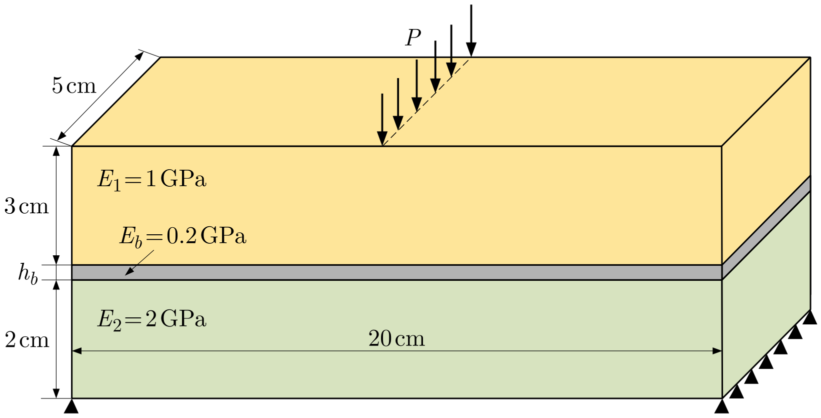


Fig. 4. Schematic diagram of the numerical model

Table 1. Parameter values from numerical simulation

Number	P [kN]	h_b [cm]	Number	P [kN]	h_b [cm]	Number	P [kN]	h_b [cm]
1	0.5	0.10	15	0.5	0.30	29	2.0	0.05
2	1.5	0.10	16	1.5	0.30	30	2.0	0.15
3	2.5	0.10	17	2.5	0.30	31	2.0	0.25
4	3.5	0.10	18	3.5	0.30	32	2.0	0.35
5	4.0	0.10	19	4.0	0.30	33	2.0	0.40
6	4.5	0.10	20	4.5	0.30	34	2.0	0.45
7	5.0	0.10	21	5.0	0.30	35	2.0	0.50
8	0.5	0.20	22	1.0	0.05	36	3.0	0.05
9	1.5	0.20	23	1.0	0.15	37	3.0	0.15
10	2.5	0.20	24	1.0	0.25	38	3.0	0.25
11	3.5	0.20	25	1.0	0.35	39	3.0	0.35
12	4.0	0.20	26	1.0	0.40	40	3.0	0.40
13	4.5	0.20	27	1.0	0.45	41	3.0	0.45
14	5.0	0.20	28	1.0	0.50	42	3.0	0.50

Figure 5 and Table 2 show the calculation results of shear strain energy in numerical simulation. When other conditions remain unchanged, the shear strain energy increases linearly with an increase of cementation layer thickness, indicating that the thicker the cementation

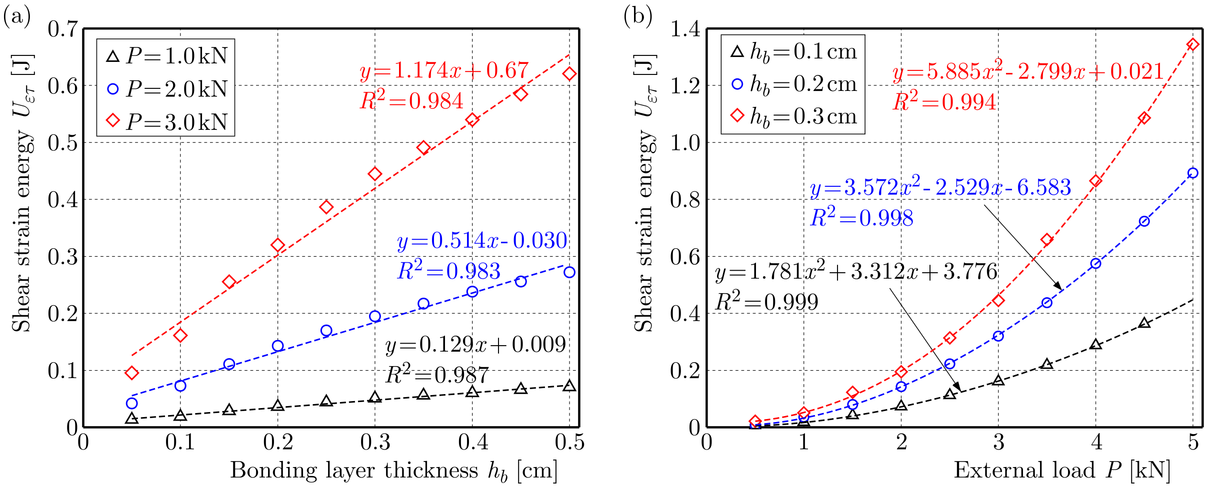


Fig. 5. Numerical simulation results: (a) influence of the bonding layer thickness, (b) influence of the external load

Table 2. Shear strain energy values

Number	Shear strain energy [J]	Number	Shear strain energy [J]	Number	Shear strain energy [J]
1	0.00464	15	0.02154	29	0.04183
2	0.04162	16	0.12246	30	0.11107
3	0.11230	17	0.31403	31	0.17074
4	0.21886	18	0.65948	32	0.21673
5	0.28728	19	0.86494	33	0.23807
6	0.36265	20	1.08598	34	0.25600
7	0.44718	21	1.34482	35	0.27199
8	0.00810	22	0.01333	36	0.09551
9	0.08012	23	0.02859	37	0.25554
10	0.22331	24	0.04398	38	0.38651
11	0.43689	25	0.05571	39	0.49136
12	0.57522	26	0.06003	40	0.54014
13	0.72299	27	0.06557	41	0.58462
14	0.89271	28	0.07057	42	0.62091

layer in the composite beam, the more shear strain energy can be stored. With an increase of the external force, the shear strain energy increases as a quadratic polynomial, indicating that the external force applied by the composite beam is greater, and the stored shear strain energy will increase accordingly.

4. Energy release process of the laminated composite roof

In order to analyze the energy storage characteristics of composite beam samples, a three-point bending test of a double-layer composite beam was carried out (Fig. 6). The samples are mainly composed of limestone and fine sandstone, and their basic mechanical properties are listed in Table 3. Before the test, the limestone and fine sandstone samples were cut into samples with a length of 200 mm and a width of 40 mm, and the sample heights were 20 mm/30 mm/40 mm, respectively. The cut samples were bonded by gypsum to make the samples of different combinations. The sample numbers and parameters are listed in Table 4.

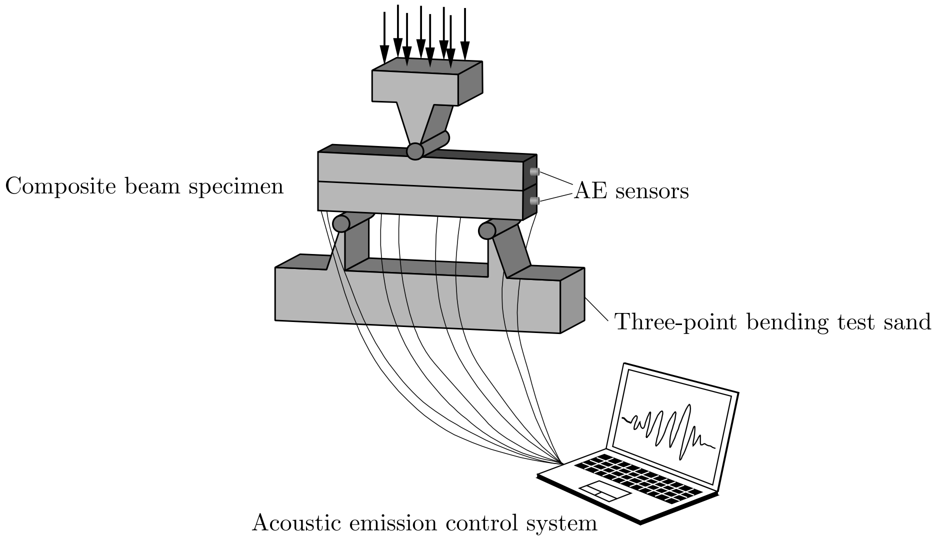


Fig. 6. Schematic diagram of three-point bending test of the composite beam

Table 3. Basic physical and mechanical parameters of composite beams

Lithology	Density ρ $[\mathrm{g\cdot cm^{-3}}]$	Compressive strength σ_c [MPa]	Tensile strength σ_t [MPa]	Modulus of elasticity E [GPa]	Shear wave velocity c_s $[\mathrm{m\cdot s^{-1}}]$	Longitudinal wave velocity c_p $[\mathrm{m\cdot s^{-1}}]$
Fine sandstone	2.65	78.23	7.64	7.51	2615.22	4102.69
Limestone	2.39	62.45	5.62	5.38	2203.17	3223.75

Table 4. Composite beam test plan

Sample number	Upper lithology	Upper layer thickness h_1 [mm]	Underlying lithology	Underlayer thickness h_2 [mm]
40S-20H	Fine sandstone	40	Limestone	20
40H-20S	Limestone	40	Fine sandstone	20
30S-30H	Fine sandstone	30	Limestone	30
30H-30S	Limestone	30	Fine sandstone	30
20S-40H	Fine sandstone	20	Limestone	40
20H-40S	Limestone	20	Fine sandstone	40

The loading system adopts the RLJW-2000 servo-controlled rock pressure testing machine of Shandong University of Science and Technology. The displacement loading method is adopted in the test. The loading rate is set at 0.05 mm/min, the sampling frequency is 10 Hz, and the span is set at 150 mm. In order to analyze the fracture characteristics of composite beams during loading, an acoustic emission test was carried out simultaneously. The acoustic emission testing equipment is AMSY-6 acoustic emission system produced by Vallen Company. Before the test, eight WS45-H acoustic emission probes were arranged on the surface of the sample to collect the acoustic emission signals during the failure process of the sample, and the acoustic emission positioning was carried out. The position of acoustic emission probe is shown in Fig. 7.

The acoustic emission signals generated during the failure of composite beam specimens are caused by the fracture of specimens. The fracture in rock interior is mainly the tensile crack caused by the tensile stress, and the fracture at the interface is mainly the shear crack

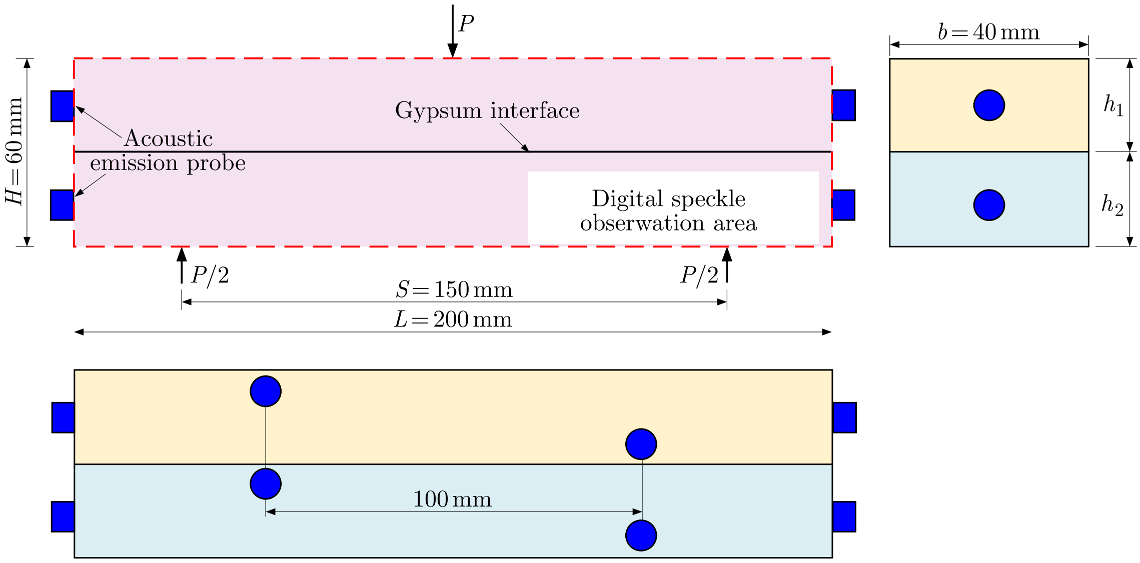


Fig. 7. Schematic diagram of sensor layout

caused by the bending shear stress. Therefore, it is of great significance to reveal the complete process of composite beam rupture by analyzing the distribution law of acoustic emission signals associated with the fracture of different mechanisms during specimen loading, and establishing the corresponding relationship between the fracture type, occurrence time and sound emission signal energy.

In the field of acoustic emission research, RA and AF distribution changes based on time-domain parameters can qualitatively describe the composition of shear cracks and tensile cracks in samples during the development stage of rock fracture, and then determine the evolution law of different types of cracks. RA and AF can be calculated according to formulas (4.1) and (4.2). In Japan's JC MS-III B5706 concrete building Code, the slope of AF/RA sector-line is defined as k, the signal of AF/RA $< k$ is defined as a shear fracture signal, and the signal of AF/RA $\geqslant k$ is defined as a tensile fracture signal. Although the partition parameter k of shear and tensile fracture is greatly different due to the influence of rock material and sensor type, the method of calculating the proportion of rock fracture type based on k value is valid according to the research results of many papers. In this Section, the research results from literature (Ohno and Ohtsu, 2010) are adopted, $k = 80$ is used as the dividing standard of shear and tensile fracture, that is, formula (4.3), and the test results of 6 groups are analyzed

$$\mathrm{AF} = \frac{C}{D} \tag{4.1}$$

where C is the acoustic emission count, D is the duration of acoustic emission

$$\mathrm{RA} = \frac{R}{A} \tag{4.2}$$

where R is the rise time of acoustic emission, A is the acoustic emission amplitude

$$\text{Crack type} = \begin{cases} \text{tensile crack} & \dfrac{\mathrm{AF}}{\mathrm{RA}} \geqslant 80 \\[2ex] \text{shear crack} & \dfrac{\mathrm{AF}}{\mathrm{RA}} < 80 \end{cases} \tag{4.3}$$

Figure 8 shows the number of two types of cracks and the corresponding energy distribution in the loading process of six groups of samples as can be seen from the figure.

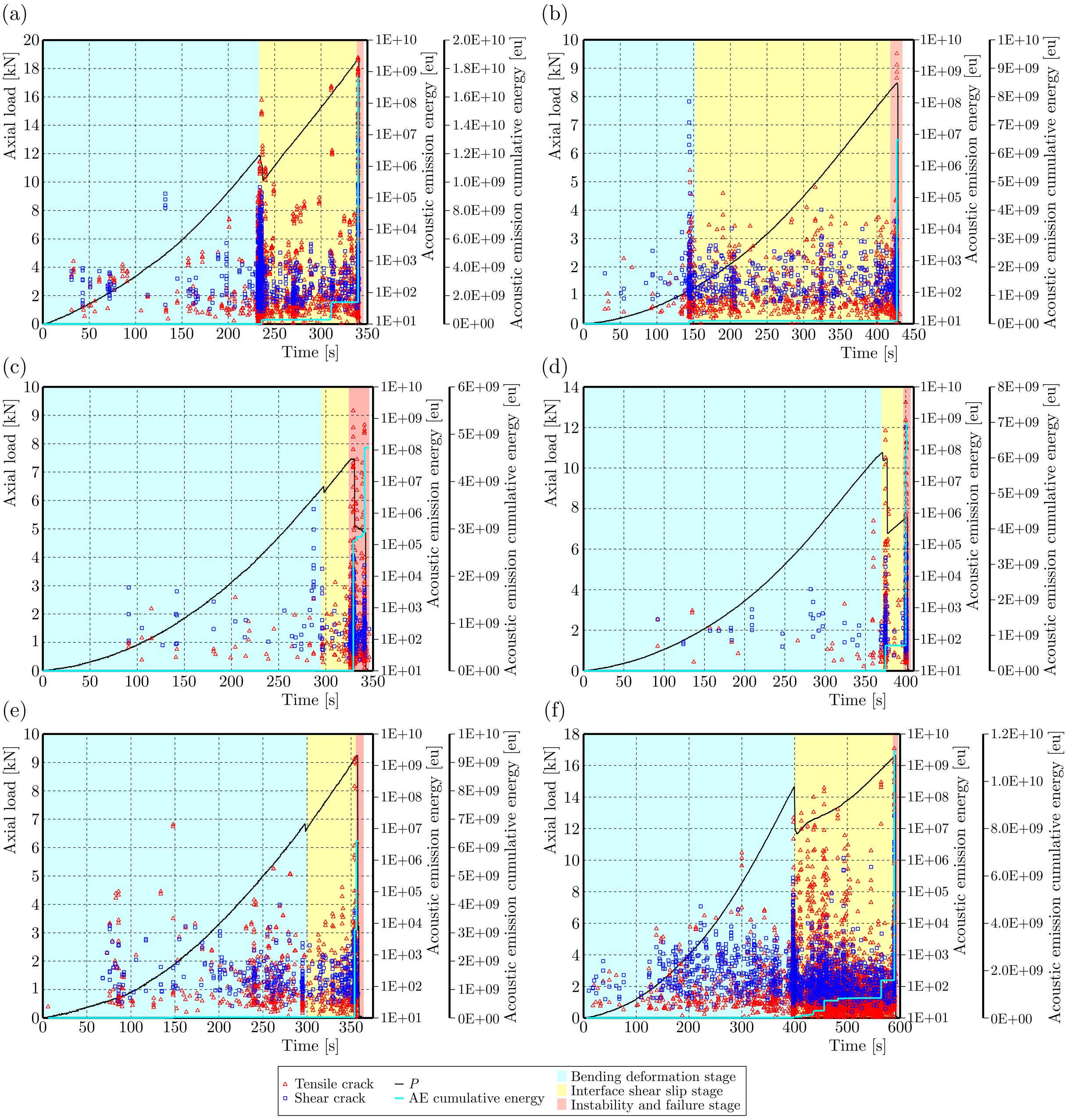


Fig. 8. Acoustic emission energy curve of composite beam: (a) 40S-20H, (b) 40H-20S, (c) 30S-30H, (d) 30H-30S, (e) 20S-40H, (f) 20H-40S

According to the loading time and acoustic emission curve in Fig. 8, it is divided into three stages: the overall bending deformation stage (blue background in the figure), the interface friction slip stage (yellow background in the figure), and the instability fracture stage (red background in the figure). In the six groups of tests, two concentrated areas of high energy acoustic emission events appeared almost simultaneously, which were located at the end of the overall bending deformation stage and the overall instability fracture stage, respectively, indicating that relatively severe macroscopic cracks appeared in the samples when the composite beams were loaded to two time nodes. In the early stage of loading, a small amount of two kinds of acoustic emission fracture signals appeared in the composite beam sample, indicating that the internal fracture of the sample in the early stage of loading is a mixture of internal tensile fracture and interface shear fracture.

The two types of composite beams have great differences in the acoustic emission energy. The total energy released by 40H-20S, 30S-30H and 20S-40H samples without thick and hard layers is 6.47×10^9eu, 4.71×10^9eu and 6.23×10^9eu, respectively. The total energy released by samples with thick hard layer in 40S-20H, 30H-30S and 20H-40S is 1.73×10^{10}eu, 7.00×10^{10}eu and 1.13×10^{10}eu, respectively. When there is a thick hard layer in the composite beam, the total energy of AE signal and the number of high energy AE signal increase significantly. The results show that the existence of the thick hard layer makes the internal fracture of the composite beam significantly increase the acoustic emission activity, rupture strength and energy release of the composite beam.

There are great differences in crack types between the two types of composite beams, which are mainly reflected in the following aspects: the cracks of composite beams without thick hard layers (samples 40H-20S, 30S-30H, 20S-40H) are mainly shear fracture in the overall bending deformation stage, indicating that shear failure occurs at the rock layer interface under the action of shear stress, and the interlayer shear slip occurs at this time. The composite beams containing thick hard layers (samples 40S-20H, 30H-30S, 20H-40S) show high energy tensile fracture at the same time in the concentration area of the first high energy acoustic emission event, indicating that when the composite beams contain thick hard layers, in addition to shear failure mainly caused by shear slip at the interface, tensile cracks in the rock also appear at this stage. This phenomenon occurs mostly in low-strength rock beams, and corresponds to the results of a digital speckle test. When the samples were loaded to the stage of overall instability, the acoustic emission signals were mainly of the tensile type. At this stage, the composite beam samples were overall unstable, and the two layers of rock beams both experienced the tensile fracture, indicating that the energy source in this stage was almost all from the tensile failure inside the rock mass.

To sum up, the thick hard layer in the roof is the key factor leading to a severe damage degree of the composite structure and an increase of energy release. In the prevention and control of rock burst, we should pay attention to this part of rock formation, and reduce the effective size and elastic modulus to diminish the purpose of weakening the fracture strength.

5. Discussion

According to the physical function relationship, the rock fracture process meets

$$U_s = W + V \tag{5.1}$$

where U_s is the total strain energy stored during rock loading, W is the fracture work done by rock crack formation, V is the energy released by rock failure.

Under the condition of composite beams, the total strain energy stored can be expressed as

$$U_s = U_{\varepsilon\tau} + U_{\varepsilon b} \tag{5.2}$$

where $U_{\varepsilon b}$ is bending strain energy stored in the composite beams.

Similarly, the energy V released by failure of the composite beam can be expressed as

$$V = V_b + V_\tau \tag{5.3}$$

where V_b is the energy release when bending and breaking, V_τ is the shear failure energy release.

Therefore, the functional relationship in the fracture process of composite beams can be expressed as

$$U_{\varepsilon\tau} + U_{\varepsilon b} = W + V_\tau + V_b \tag{5.4}$$

In fracture mechanics, the fracture work is a parameter to evaluate how much energy is consumed when a rock breaks, reflecting the energy required for fracture expansion, which can be obtained from the load-displacement curve (P-δ curve) measured by the three-point bending test and the area surrounded by the horizontal coordinate. Due to the limitation of test conditions, the posterior peak of the P-δ curve is generally difficult to measure directly, and it is approximately considered to conform to the law of power function, which can be expressed as

$$P = \beta\delta^{-\lambda} \tag{5.5}$$

where β and λ are curve coefficients, which can be determined by fitting the test data points after $P = P_{max}/3$, and the displacement at this time is set as δ_1.

Assuming that the end of the P-δ curve extends to an infinite displacement, the fracture work can be expressed as

$$\begin{aligned} W = W_0 + W_1 + W_G &= \int\limits_0^{\delta_1} P(\delta)\, d\delta + \int\limits_{\delta_1}^{+\infty} \beta\delta^{-\lambda}\, d\delta + mg\delta_0 \\ &= \int\limits_0^{\delta_1} P(\delta)\, d\delta + \frac{\beta}{1-\lambda}\delta_1 + mg\delta_0 \end{aligned} \tag{5.6}$$

where W_0 is the envelope area of the measured P-δ curve at $0 \leqslant \delta < \delta_1$ and the horizontal coordinate, W_1 is the envelope area of the measured P-δ curve at $\delta > \delta_1$ and the horizontal coordinate, W_G does work for the body weight of the specimen, δ_0 is the mid-span displacement of the composite beam.

6. Conclusion

To reveal the shear deformation and energy storage mechanism of composite roof strata, quantify the interlayer shear energy storage characteristics of composite roof strata, this study establishes a shear energy mechanical model of layered composite roof structures under different conditions. Factors influencing the rock strata shear energy storage are analyzed, the energy release process and mechanism of layered composite roof strata are discussed, leading to the following conclusions:

- The shear strain energy linearly increases with an increase in bond layer thickness and quadratically increases with increasing external forces.
- During the bending deformation stage, when thick and hard layers exist in the composite beams, shear failure dominated by interlayer shear slip occurs, and low-strength rock beams exhibit tensile crack initiation. In the overall instability stage, the composite beams mainly experience tensile fracture.
- When thick and hard layers exist in composite beams, the fracture strength and released energy are higher. The presence of thick and hard layers in the roof is a key factor leading to severe structural damage and increased energy release.
- In the prevention of dynamic rock burst, special attention should be paid to this part of thick and hard rock strata, aiming to weaken the fracture strength by reducing effective dimensions and decreasing elastic modulus.
- In the future research, a bending loading device suitable for triaxial conditions can be designed and developed according to the real ground stress environment, and corresponding tests can be carried out to reveal the influence of ground stress environment on the fracture characteristics of the roof.

Acknowledgments

We would like to thank Editage (www.editage.cn) for the English language editing. Furthermore, we are thankful for the support provided by the National Natural Science Foundation of China (Grant No. 52304095), the Major Program of Shandong Provincial Natural Science Foundation (No. ZR2019ZD13), the National Natural Science Foundation of China (52374097), and the Taishan Scholar Project in Shandong Province (Grant No. tstp20221126) for funding this research.

References

1. Bai Q., Tu S., 2020, Numerical observations of the failure of a laminated and jointed roof and the effective of different support schemes: A case study, *Environmental Earth Sciences*, **79**, 202
2. Cao J., Huang Q., 2021, Roof structure of shallow coal seam group mining in Western China, *Plos One*, **16**, 8
3. Coggan J., Gao F., Stead D., Elmo D., 2012, Numerical modelling of the effects of weak immediate roof lithology on coal mine roadway stability, *International Journal of Coal Geology*, **90-91**, 100-109
4. Cui F., Yang Y., Lai X., Jia C., Shan P., 2020, Experimental study on the effect of advancing speed and stoping time on the energy release of overburden in an upward mining coal working face with a hard roof, *Sustainability*, **12**, 1, 37
5. Han C.L., Zhang N., Li B.Y., Si G.Y., Zheng X.G., 2015, Pressure relief and structure stability mechanism of hard roof for gob-side entry retaining, *Journal of Central South University*, **22**, 4445-4455
6. He J., Dou L.M., Cao A.Y., Gong S.Y., Lü J.W., 2012, Rock burst induced by roof breakage and its prevention, *Journal of Central South University*, **19**, 1086-1091
7. Li W., Tu S., Tu H., Liu X., Miao K., *et al.*, 2022, A new method to assess thick, hard roof-induced rock burst risk based on mining speed effect on key energy strata, *Sustainability*, **14**, 22, 15054
8. Liu J., Li C., Shi Y., Zhang Y., 2021, Stability analysis and fracture patterns of hard main roof in longwall top coal caving with large mining height, *Shock and Vibration*, **2021**
9. Liu W., Shi Q., 2021, Study on evolution characteristics and instability model of overburden stress shell with the longwall mining along strike, *Shock and Vibration*, **2021**
10. Lu W., He C., Zhang X., 2020, Height of overburden fracture based on key strata theory in longwall face, *Plos One*, **15**, 1
11. Lu Y., Gong T., Xia B., Yu B., Huang F., 2019, Target stratum determination of surface hydraulic fracturing for far-field hard roof control in underground extra-thick coal extraction: A case study, *Rock Mechanics and Rock Engineering*, **52**, 8, 2725-2740
12. Małkowski P., Niedbalski Z., 2020, A comprehensive geomechanical method for the assessment of rockburst hazards in underground mining, *International Journal of Mining Science and Technology*, **30**, 3, 345-355
13. Ohno K., Ohtsu M., 2010, Crack classification in concrete based on acoustic emission, *Construction and Building Materials*, **24**, 12, 2339-2346
14. Qi F., Yang D., Li X., Li B., Zhai Y., Yu S., 2022, Study on roof presplitting mechanism and deformation control of reused roadway in compound soft rock by roof presplitting approach, *Geofluids*, **2022**, 1, 144937
15. Qin G., Cao J., Wang C., Wu S., Zhai M., 2021, Characteristics of stratum structure and fracture evolution in stratified mining of shallow buried high-gas-thick coal seam by similarity simulation, *Geofluids*, **2021**, 1, 555451

16. Shen W., Wang M., Cao Z., Su F., Nan H., Li X., 2019, Mining-induced failure criteria of interactional hard roof structures: A case study, *Energies*, **12**, 15, 3016

17. Tajdus A., Cala M., Tajdus K., 2018, Seismicity and rock burst hazard assessment in fault zones: A case study, *Archives of Mining Sciences*, **63**, 3, 747-765

18. Tan Y.L., Liu X.S., Shen B., Ning J.G., Gu Q.H., 2018, New approaches to testing and evaluating the impact capability of coal seam with hard roof and/or floor in coal mines, *Geomechanics and Engineering*, **14**, 4, 367-376

19. Ti Z., Li J., Wang M., Wang K., Jin Z., Tai C., 2021, Fracture mechanism in overlying strata during longwall mining, *Shock and Vibration*, **2021**

20. Wang H., Jiao J., Wu Y., Zheng K., Li Y., Wang T., Jiang B., 2023, Deformation characteristics and stress evolution law of composite hard roof under presplitting weakening, *Mining, Metallurgy & Exploration*, **40**, 839-850

21. Wang X., Lu M., Wei R., Wang Z., Li S., 2022, Mechanical mechanism analysis of roof fracture evolution in stope with variable length based on elastic-plastic structure theory, *Geofluids*, **2022**, 1, 475021

22. Xia B., Zhang X., Yu B., Jia J., 2018, Weakening effects of hydraulic fracture in hard roof under the influence of stress arch, *International Journal of Mining Science and Technology*, **28**, 6, 951-958

23. Xu C., Yang G., Sun H., Qin L., Wang K., *et al.*, 2021, Key strata inducing dynamic disasters based on energy condition: Criterion and application, *Geofluids*, **2021**, 1, 672020

24. Xu D.Q., Zhang P.S., Yan W., Zhang X.L., Dong Y.H., Niu H., 2023, Mechanical properties and energy of sandstone under cyclic loading in evolutionary pattern experimental studies, *Archives of Mining Sciences*, **68**, 2, 351-370

25. Yang Z., Liu C., Zhu H., Xie F., Dou L., Chen J., 2019, Mechanism of rock burst caused by fracture of key strata during irregular working face mining and its prevention methods, *International Journal of Mining Science and Technology*, **29**, 6, 889-897

26. Yu M., Zuo J., Sun Y., Mi C., Li Z., 2022, Investigation on fracture models and ground pressure distribution of thick hard rock strata including weak interlayer, *International Journal of Mining Science and Technology*, **32**, 1, 137-153

27. Zhang J., He Y., Yang T., Bai W., Wu J., *et al.*, 2023, Coevolution mechanism and branch of pillar-overburden fissures in shallow coal seam mining, *Energy Science & Engineering*, **11**, 5, 1630-1642

28. Zhang J., Wang B., Bai W., Yang S., 2021, A study on the mechanism of dynamic pressure during the combinatorial key strata rock column instability in shallow multi-coal seams, *Advances in Civil Engineering*, **2021**

29. Zhou K.Y., Małkowski P., Dou L.M., Yang K., Chai Y.J., 2023, Using elastic wave velocity anomaly to predict rockburst hazard in coal mines, *Archives of Mining Sciences*, **68**, 1, 141-164

Manuscript received April 3, 2024; accepted for publication October 28, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 51-58, Warsaw 2025
https://doi.org/10.15632/jtam-pl/194239

PIEZORESISTIVE SENSING MECHANISM OF GRAPHENE-BASED ELECTRONIC SKIN

Li Li
Department of Dermatology, Xi'an No.3 Hospital, Xi'an, PR China, and
The Affiliated Hospital of Northwest University, Xi'an, PR China

Jia Huang
Xi'an University of Technology, Xi'an, PR China

Wenbin Yang
Department of Dermatology, Xi'an No.3 Hospital, Xi'an, PR China, and
The Affiliated Hospital of Northwest University, Xi'an, PR China
e-mail: YangWB0910@126.com

The graphene-based sensing electronic skin has a broad application prospect, but its piezoresistive sensing mechanism still needs to be further studied. In this work, according to the microscopic characteristics of the graphene electronic skin, the Monte Carlo stochastic algorithm is used to generate randomly distributed graphene sheets, and then the piezoresistive sensor model of the graphene electronic skin is established. The relative resistance and gauge factor of the model are calculated by the finite element method. Meanwhile, the current density and potential contour under different graphene morphology are obtained. The results show that the graphene sensor with a high area fraction has a higher sensing range, and the graphene sensor with a low area fraction has a higher gauge factor. The piezoresistive effect of the model depends mainly on a change of graphene sheet density. With the change of strain, the variation of the overlap area and overlap number between graphene sheets will cause a change of graphene density and electron migration path. The separation between graphene sheets can lead to reduction in electron migration paths, resulting in nonlinear changes in the relative resistance of the model. The present work can provide technical support for design and preparation of the graphene-based electronic skin.

Keywords: graphene, electronic skin, finite element method, piezoresistive effect

1. Introduction

Skin is the largest organ of the human body, and it is a multi-functional sensor that can clearly perceive changes in the external environment, which is of great significance to the survival and development of human beings. Electronic skin, on the other hand, is a device that simulates the function of human skin (Ma and Khoo, 2024; Li *et al.*, 2023; Guo *et al.*, 2024; Fang *et al.*, 2021). It can mimic the tactile sensing function and flexible performance of human skin, and can attach to the surface of human skin or robots, sensing various stimuli such as pressure, temperature, etc. It has shown broad application prospects in fields such as intelligent medicine, human prosthetics, and health monitoring (Mudhulu *et al.*, 2023; Xiao *et al.*, 2023; Xu *et al.*, 2023; Hu *et al.*, 2024; Liao *et al.*, 2024).

The piezoresistive sensing electronic skin based on graphene has gained great attention due to its excellent sensitivity and flexibility (Chen *et al.*, 2023b; Fang *et al.*, 2023; Chen *et al.*, 2023c; Chen *et al.*, 2023a). Yun *et al.* (2023) prepared a waterproof and ultra sensitive wearable strain sensor using carbon black/graphene/carboxymethyl cellulose, which can easily capture signals of spatial strain during body movement. Feng *et al.* (2023) reported a new type of graphene-based resistance sensor with high sensitivity and response speed, and its blood pressure detection error value has passed the standards of the Association for the Advancement of Medical Instrumentation standard. Sharma *et al.* (2023) proposed a piezoresistive sensor based on graphene nanosheets and a PDMS substrate, which can detect both fine and large strains with good stability and repeatability. Wei *et al.* (2021) prepared a wearable piezoresistive sensor with high sensing performance through polypyrrole and reduced graphene aerogel. Hong *et al.* (2024) successfully prepared a dual biomimetic stretchable strain sensor with fingerprint patterns and a biomimetic lotus root fiber structure using silicone rubber, multi walled carbon nanotubes, and graphene, which could accurately achieve gesture recognition, human micro-expression monitoring. Combined with machine-learning algorithms, Ma *et al.* (2923) proposed a porous graphene-based flexible pulse sensor for cardiovascular disease diagnostics, which achieves high accuracy ($> 93\%$), showing a broad application prospect of the electronic skin.

In order to establish the connection between the macroscopic performance and microscopic structure of graphene-based composite sensors, and to study their sensing mechanism at a deeper level, numerical simulation is an economically effective method. Li and Yang (2020) established a two-dimensional random model through theoretical modeling and finite element simulation to study the piezoresistive behavior of stacked graphene composite sensors. Ren *et al.* (2022) used the finite element method to analyze the mechanical stress and deflection of the graphene piezoresistors, established a functional relationship between mechanical properties and size variables, and determined the optimal size of the structure. Lamba *et al.* (2022) used finite element analysis to design and simulate graphene piezoresistive sensors, which had the ability to perceive biaxial forces and were therefore suitable for microbiology, minimally invasive surgery, and healthcare applications.

In the present work, we establish a two-dimensional sensor model based on the microstructure characteristics of the graphene-based electronic skin, and explore the piezoresistive sensing mechanism of the graphene-based electronic skin and the influence of microstructure parameters on piezoresistive performance.

2. Graphene-based electronic skin sensing model

Based on the structural characteristics of randomly stacked graphene sheets in electronic skin (Li *et al.* (2016), the Monte Carlo random algorithm is used to generate randomly distributed graphene sheets, as shown in Fig. 1.

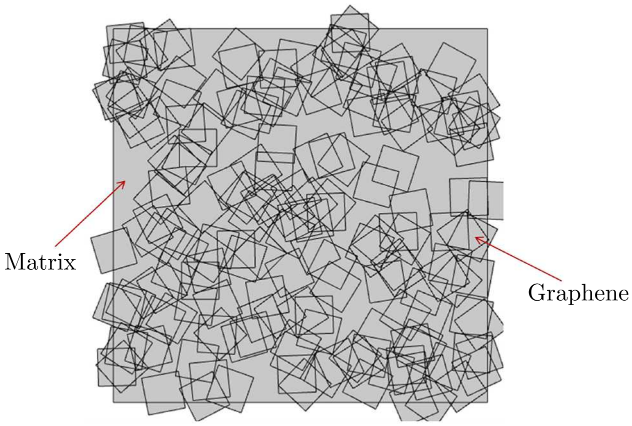


Fig. 1. Schematic diagram of the graphene sensor model

This work makes the following assumption: Firstly, since the modulus of graphene is much larger than that of the matrix, we can consider graphene sheets as rigid materials. Secondly, due to significant van der Waals forces between graphene sheets and polymer matrix materials, it is assumed that the matrix can fully transfer deformation to the graphene sheets. Finally, it is assumed that the sensor has uniform deformation, meaning that the center of the graphene sheet remains unchanged relative to the entire sensor position. If the model extends along the x-direction, the center coordinates of the i-th graphene at any time can be expressed as

$$x_i = x_i^0(1+\varepsilon) \qquad y_i = y_i^0(1-\nu\varepsilon) \tag{2.1}$$

where ε is the linear strain along the x-direction of the model, and ν is Poisson's ratio of the matrix material.

Commonly used electronic skin generally takes rubber as the matrix material. Under a large deformation, in order to determine Poisson's ratio of rubber, a small unit is taken out with an initial volume of $V_0 = dx\,dy\,dz$. The volume of the unit after deformation at unidirectional stress σ_x is $V_1 = (1+\varepsilon)(1-\nu\varepsilon)^2\,dx\,dy\,dz$. Due to the incompressibility of rubber, there is

$$1 = (1+\varepsilon)(1-\nu\varepsilon)^2 \tag{2.2}$$

Solving equation (2.2) yields

$$\nu = \frac{1}{\varepsilon} \pm \frac{1}{\varepsilon\sqrt{1+\varepsilon}} \tag{2.3}$$

Because $1-\varepsilon$ is greater than zero, the negative sign is taken from Eq. (2.3), which means that Poisson's ratio of rubber is

$$\nu = \frac{1}{\varepsilon} - \frac{1}{\varepsilon\sqrt{1+\varepsilon}} \tag{2.4}$$

Based on Eq. (2.1), we can rewrite the central coordinates of the i-th graphene sheet as

$$x_i = x_i^0(1+\varepsilon) \qquad y_i = y_i^0\frac{1}{\sqrt{1+\varepsilon}} \tag{2.5}$$

In order to analyze the resistance changes of the graphene sensor model, the left and right ends of the model are used as two electrodes, with constant current input and output respectively. Therefore, the analysis of resistance is transformed into the solution of the following control equation

$$\nabla\cdot J = Q_{j,V} \qquad J = \sigma E + J_e \qquad E = -\nabla V \tag{2.6}$$

where J represents current density, E represents electric field intensity, V represents electric potential, σ is electrical conductivity, $Q_{j,v}$ is an external current source, and J_e is the externally generated current density.

The boundary conditions of this model are insulation between the upper and lower boundaries, with terminals on the left and grounding on the right boundary. The basic solution approach is to use the finite element method to obtain the potential and current distribution of the model, and then use Ohm's law to determine the model resistance. The numerical calculations are performed with the COMSOL Multiphysics Software.

The above methods can be used to study the piezoresistive performance of graphene sensors. The gauge factor β is an important indicator for judging the piezoresistive performance of sensors, which represents the relative resistance per unit strain, i.e.

$$\beta = \frac{\Delta R}{R\varepsilon} \tag{2.7}$$

When the gauge factor β is a constant a indicates that the relative resistance varies linearly with strain, and the graphene sensor outputs a linear signal. Conversely, the sensor outputs a nonlinear signal.

3. Results and discussion

In this work, the electrical resistivity of single-layer graphene $\rho_0 = 4.2 \cdot 10^{-8}\,\Omega\text{m}$, and the in-plane resistivity of the overlapping area of multi-layer graphene sheets is equal to the parallel value of the resistances of each single-layer graphene sheet, i.e ρ_0/n, where n is the number of layers of graphene stacking. The resistivity of the rubber matrix is $1 \cdot 10^{13}\,\Omega\text{m}$, which is much higher than the resistivity of graphene. In the model, the content of graphene is represented by the area fraction $f = mA/A_0$, where m is the number of graphene sheets, A is the area of each graphene sheet, and A_0 is the area of the model.

Figure 2 shows the variation of the relative resistance and gauge factor of sensors with strain under different GNPs area fraction. The simulation results shown in Fig. 2a indicate that the initial stage of relative resistance changes approximately linearly, and as the strain increases, the resistance begins to increase rapidly. Because when the strain starts to increase, the GNPs mainly slips, causing only an increase in the length of the electron migration path. But when the deformation is large, the GNPs begin to separate, and the number of graphene sheets that do not participate in current transfer increases, reducing the electron migration path, resulting in a sudden change in relative resistance. Due to the larger slip range between graphene sheets with a high area fraction, the relative resistance of graphene sensors with the high area fraction will also exhibit better linear performance when the model is extended. At the same time, it can be seen that the sensing range of the high area fraction graphene sensor is higher, mainly due to better connectivity of the high aspect ratio graphene sensor when the model is extended. From Fig. 2b, it can be seen that sensors with a lower area fraction have a higher gauge factor, because for graphene sensors with lower area fraction, the increase in resistance caused by separation of graphene sheets can be more easily reflected. The above simulation results are consistent with the experimental conclusions of Son *et al.* (2019) and Sethy *et al.* (2019), which confirms that this model can simulate the piezoresistive effect of the graphene-based electronic skin.

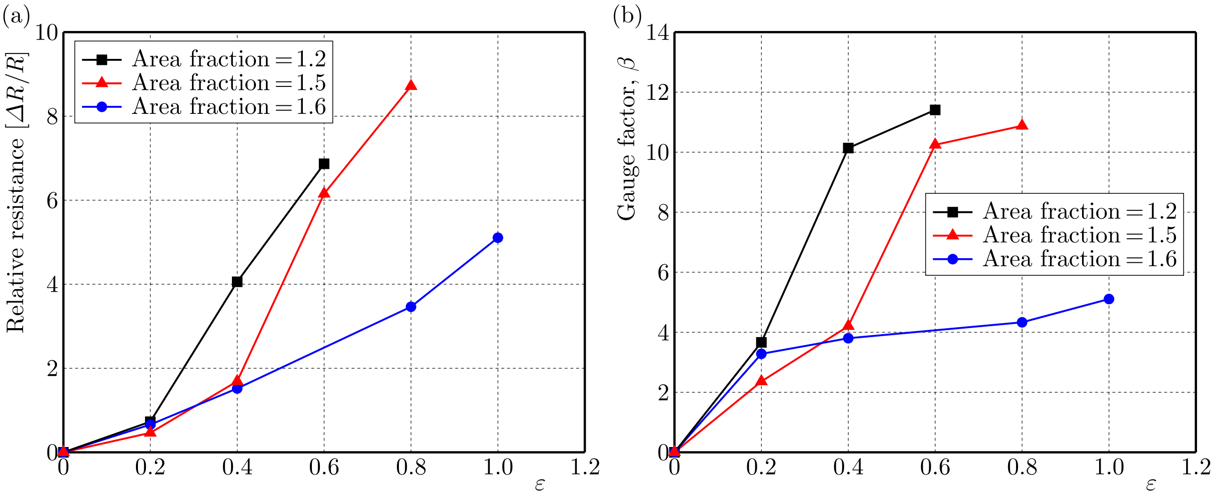


Fig. 2. Relative resistance $\Delta R/R$ and gauge factor β under different strains

Figure 3 shows the contour of current density under different strains. By observing Fig. 3a, three network branches with high current density are formed respectively at the upper, middle and lower positions of the sensor model. The low current density region in the model includes not only the rubber substrate (the rubber matrix has very low electrical conductivity, so the current density value is low), but also the graphene sheets separated from the network, which can be regarded as invalid sheets. Figure 3b shows the current density at a strain of 0.2, which forms two high current density network branches in the upper and middle positions of the model, reducing one branch compared to Fig. 3a. This indicates that the current density in the lower branch plummets, and the graphene sheet in the branch also becomes an ineffective sheet that does not participate in electron transfer, due to the separation of the graphene sheet in the lower branch. By comparing Fig. 3b and Fig. 3c, we find that the network branch in the middle region is also disconnected, which is manifested as a sharp drop in the current density in the middle region. The strain range from 0.2 to 0.4 not only involves the separation of longitudinal graphene sheets, but also the connection of transverse graphene sheets. Observing Figs. 3b and 3c, the current density originally located in the middle region of the left end has not been very high. However, due to the increase in strain, the graphene sheet appears to be getting closer and closer horizontally. Finally, when the strain is 0.4, the graphene sheet forms an electron migration channel in the horizontal direction. When the strain continued to increase to 0.6, as the graphene sheet separated, only a high current density electron migration path remained above the model.

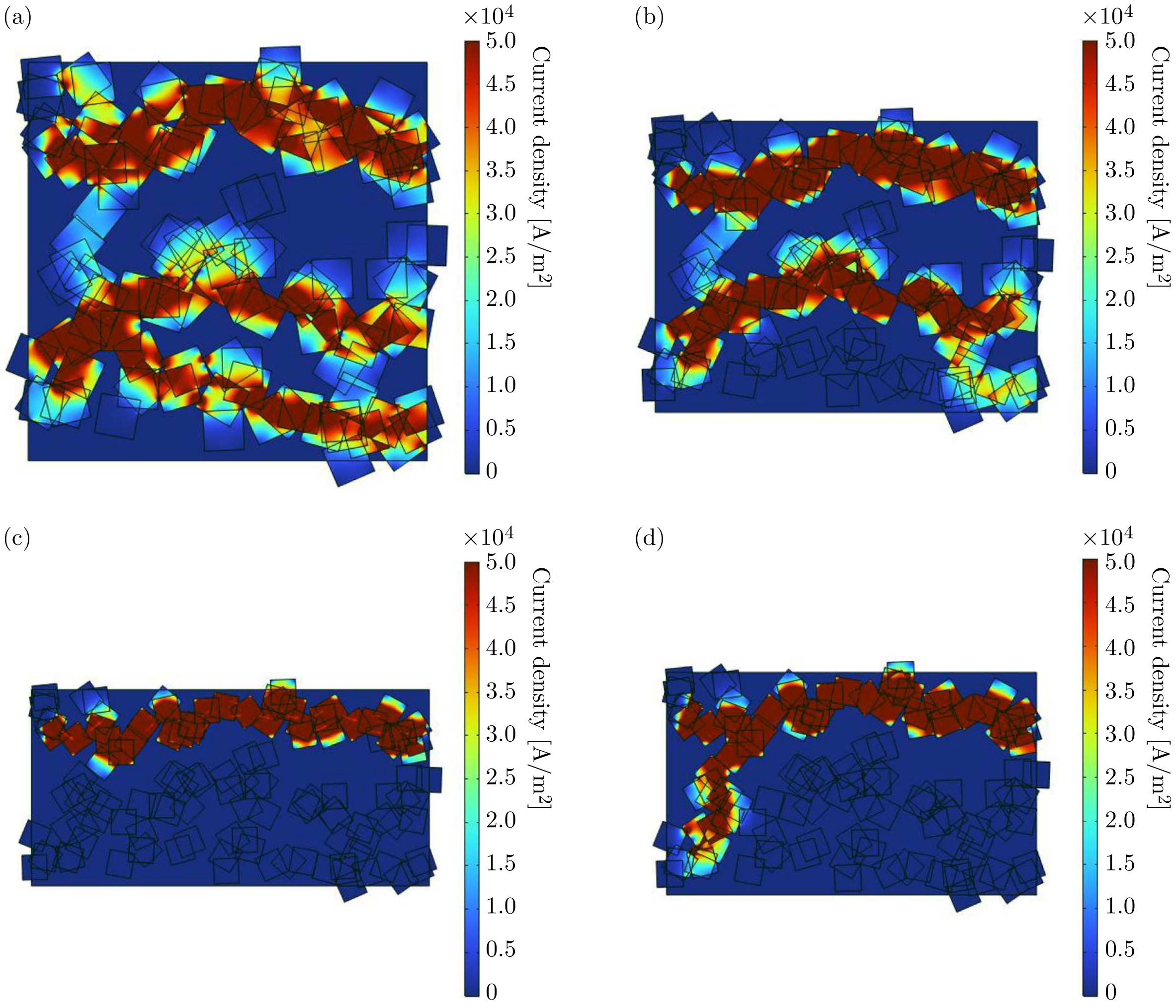


Fig. 3. Current density contour under different strains: (a) $\varepsilon = 0$, (b) $\varepsilon = 0.2$, (c) $\varepsilon = 0.4$, (d) $\varepsilon = 0.6$

Taking the model with an area fraction of 1.2 as an example, Fig. 4 shows the potential contour under different strains. When the strain is 0, it is basically in the low potential region. As shown in Fig. 3, at this time, there are more parallel branches and smaller resistance. But when the strain is increased to 0.2, the high potential region significantly increases, mainly because with the increase of strain, the graphene sheet begins to gradually slip, and the overlapping area of graphene gradually decreases, leading to a gradual increase in resistance. When the strain reaches 0.4, as the electron migration path decreases, the low potential region increases. However, the potential value in the high potential region increased, mainly due to the further increase in resistance with the slip of graphene. When the strain increcases to 0.6, this trend becomes more pronounced.

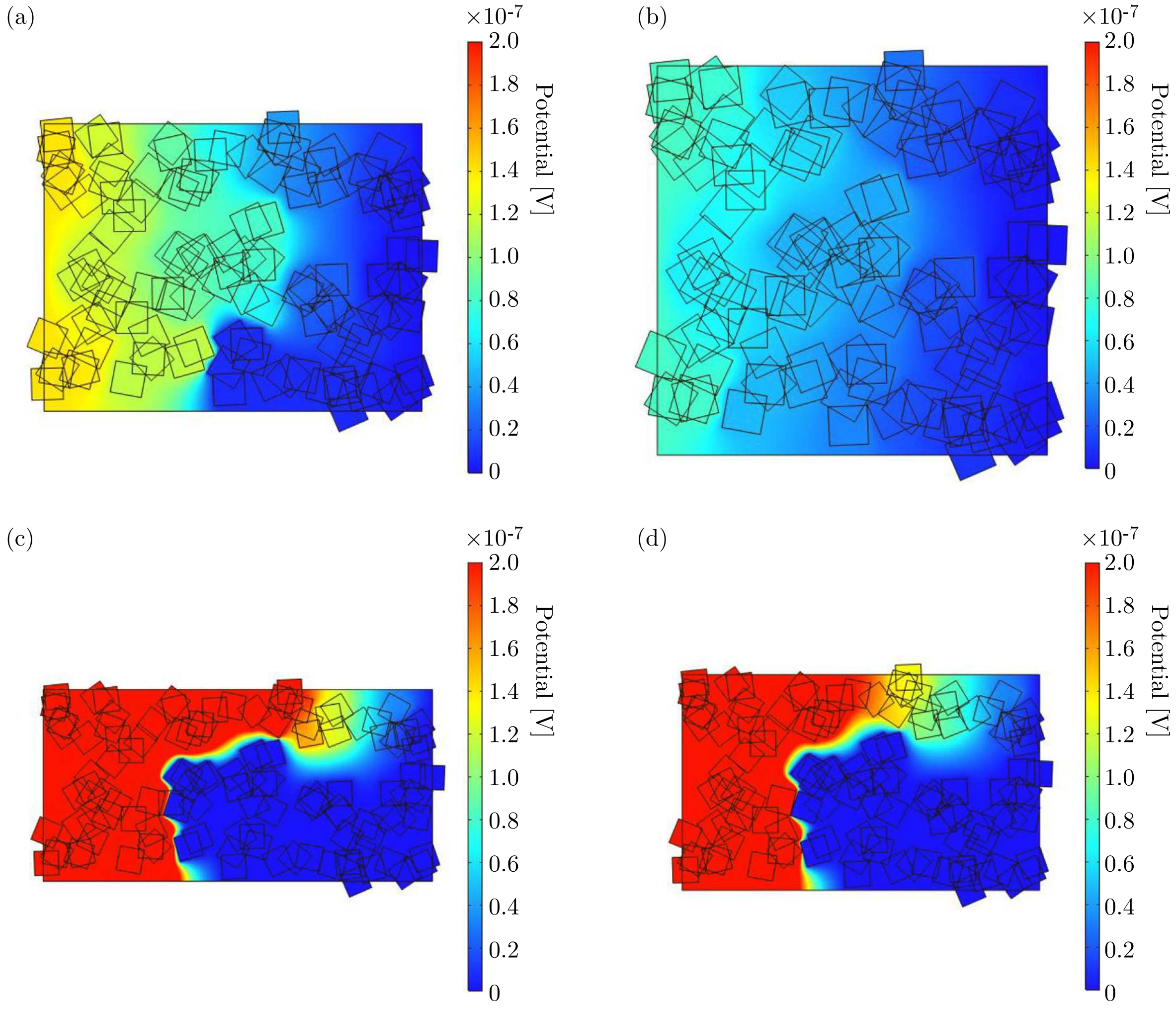


Fig. 4. Potential contour under different strains: a) $\varepsilon = 0$, (b) $\varepsilon = 0.2$, (c) $\varepsilon = 0.4$, (d) $\varepsilon = 0.6$

It should be noted that in addition to the morphological changes of graphene sheets (relative slip and separation of overlapping graphene sheets) affecting the piezoresistive properties of electronic skin, the tunneling effect between graphene sheets also has an impact on it (Wu *et al.*, 2019; Liu *et al.*, 2016). However, the sensing range where the tunneling effect plays a role is only a few nanometers. For graphene-based composites that can sense large deformations, the contribution of the tunneling effect to the piezoresistive performance of composites is very small. Therefore, to maintain simplicity of the model, the present work does not consider the influence of the tunneling effect on the piezoresistive performance of the graphene-based electronic skin.

4. Conclusions

The present work establishes a piezoresistive sensor model of the graphene-based electronic skin using a Monte Carlo random algorithm and finite element method. The results show that graphene-based sensors with a low area fraction have a higher gauge factor, while graphene sensors with a high area fraction have a higher sensing range. By comparing the current density and potential contour of different graphene morphology, it can be seen that the piezoresistive effect of the sensor mainly depends on the change in graphene sheet density. The initial deformation of the graphene sensing model is dominated by the slip of graphene, during which the resistance changes linearly. As the strain increases to a certain extent, the graphene sheets begin to separate, leading to a decrease in the network path and causing a sudden change in the resistance of the model (i.e. non-linear change).

References

1. Chen H., Zhuo F.L., Zhou J., Liu Y., Zhang J.B., *et al.*, 2023a, Advances in graphene-based flexible and wearable strain sensors, *Chemical Engineering Journal*, **464**, 142576
2. Chen K., Xu Y., Zhao Y., Li J., Wang X., Qu L.T., 2023b, Recent progress in graphene-based wearable piezoresistive sensors: From 1D to 3D device geometries, *Nano Materials Science*, **5**, 3, 247-264
3. Chen X.X., Yin Z.Z., Deng Y.T., Li Z.H., Xue M.S., *et al.*, 2023c, Harsh environment--tolerant and robust superhydrophobic graphene-based composite membrane for wearable strain sensor, *Sensors and Actuators A: Physical*, **362**, 114630
4. Fang H.Q., Yao D.H., Gao X.P., Sun Y.F., Shiwei A., *et al.*, 2023, Flexible sensors with tannin-modified vertical graphene arrays for the highly sensitive detection of humidity and strain, *Sensors and Actuators A: Physical*, **352**, 114213
5. Fang Y.S., Zou Y.J., Xu J., Chen G.R., Zhou Y.H., *et al.*, 2021, Ambulatory cardiovascular monitoring via a machine-learning-assisted textile triboelectric sensor, *Advanced Materials*, **33**, 41, 2104178
6. Feng H., Liu P., Guo X., Li J.L., Sun Y.F., *et al.*, 2023, PSS modified by 3-aminopropyltrimethoxysilane linking large-area GNPs/PSS to silicone rubber with stable interface combination for high sensitivity flexible resistive sensor, *Chemical Engineering Journal*, **465**, 143009
7. Guo W., Ma Z.Q., Chen Z., Hua H.J., Wang D., *et al.*, 2024, Thin and soft Ti_3C_2Tx MXene sponge structure for highly sensitive pressure sensor assisted by deep learning, *Chemical Engineering Journal*, **485**, 149659
8. Hong W.Q., Guo X.H., Zhang T.X., Zhu X.W., Su Z., *et al.*, 2024, Dual bionic-inspired stretchable strain sensor based on graphene/multi-walled carbon nanotubes/polymer composites for electronic skin, *Composites Part A: Applied Science and Manufacturing*, **179**, 108043
9. Hu X.H., Ma Z.Q., Zhao F.Q., Guo S., 2024, Recent advances in self-powered wearable flexible sensors for human gaits analysis, *Nanomaterials*, **14**, 1173
10. Lamba M., Chaudhary H., Singh K., Keshyep P., Kumar V., 2022, Graphene piezoresistive flexible MEMS force sensor for bi-axial micromanipulation applications, *Microsystem Technologies*, **28**, 7, 1687-1699
11. Li J., Jia H.L., Zhou J.K., Huang X.C., Xu L., *et al.*, 2023, Thin, soft, wearable system for continuous wireless monitoring of artery blood pressure, *Nature Communications*, **14**, 5009
12. Li X.M., Yang T.T., Yang Y., Zhu J., Li L., *et al.*, 2016, Large-area ultrathin graphene films by single-step Marangoni self-assembly for highly sensitive strain sensing application, *Advanced Functional Materials*, **26**, 9, 1322-1329
13. Li Z., Yang Q.S., 2020, Sensing mechanism of flexible and stretchable composites based on stacked graphene, *Materials and Design*, **187**, 108384
14. Liao J.C., Ma Z.Q., Liu S.Y., Li W., Yang X.D., *et al.*, 2024, Programmable microfluidic--assisted highly conductive hydrogel patches for customizable soft electronics, *Advanced Functional Materials*, **34**, 41, 2401930
15. Liu Y., Zhang D., Wang K., Liu Y.Y., Shang Y., 2016, A novel strain sensor based on graphene composite films with layered structure, *Composites Part A: Applied Science and Manufacturing*, **80**, 95-103

16. MA Z.Q., HUA H.J., YOU C.X., MA Z.H., GUO W., *et al.*, 2023, FlexiPulse: A machine-learning-enabled flexible pulse sensor for cardiovascular disease diagnostics, *Cell Reports Physical Science*, **4**, 12, 101690
17. MA Z.Q., KHOO B.L., 2024, Recent advances in laser-induced-graphene-based soft skin electronics for intelligent healthcare, *Soft Science*, **4**, 26
18. MUDHULU S., CHANNEGOWDA M., BALAJI S., KHOSLA A., SEKHAR P., 2023, Trends in graphene-based e-skin and artificial intelligence for biomedical applications – A review, *IEEE Sensors Journal*, **23**, 17, 18963-18976
19. REN X.C., LIU X.Y., SU X., JIANG X.F., 2022, Design and optimization of a pressure sensor based on serpentine-shaped graphene piezoresistors for measuring low pressure, *Sensors*, **22**, 13, 4937
20. SETHY D., MAKIREDDI S., VARGHESE F.V., BALASUBRAMANIAM K., 2019, Piezoresistive behaviour of graphene nanoplatelet (GNP)/PMMA spray coated sensors on a polymer matrix composite beam, *Express Polymer Letters*, **13**, 11, 1018-1025
21. SHARMA P., SHARMA R., JANYANI V., VERMA D., 2023, Development of a multi-modal graphene nanoparticles (GNP)-Polydimethylsiloxane (PDMS) flexible sensor for human activity monitoring and health assessment, *International Journal of Electrochemical Science*, **18**, 9, 100236
22. SON W., KIM K., LEE S., HYEON G., HWANG K.G., PARK W., 2019, Ecoflex-passivated graphene-yarn composite for a highly conductive and stretchable strain sensor, *Journal of Nanoscience and Nanotechnology*, **19**, 10, 6690-6695
23. WEI H.G., LI A., KONG D.S., LI Z.Z., CUI D., *et al.*, 2021, Polypyrrole/reduced graphene aerogel film for wearable piezoresisitic sensors with high sensing performances, *Advanced Composites and Hybrid Materials*, **4**, 1, 86-95
24. WU S., PENG S., YU Y., WANG C., 2019, Strategies for designing stretchable strain sensors and conductors, *Advanced Materials Technologies*, **5**, 2, 1900908
25. XIAO Y.C., CHEN Q.L.,YANG Z.M., XI M., ZHAO Y.L., *et al.*, 2023, Asymmetric and skin--mimicking hydrogels with wide temperature tolerance and superior elasticity for high-performance strain sensors, *ACS Omega*, **8**, 49, 46676-46684
26. XU Q.Y., LI M.M., ZHANG Y.X., GAO H.G., ZHANG L., *et al.*, 2023, Super anti-freezing, fast--responsive nanocomposite organohydrogels with excellent mechanical properties as multifunctional sensors for human motion monitoring, *Polymer*, **283**, 126248
27. YUN T.T., DU J., JI X.X., TAO Y.H., CHENG Y., *et al.*, 2023, Waterproof and ultrasensitive paper-based wearable strain/pressure sensor from carbon black/multilayer graphene/carboxymethyl cellulose composite, *Carbohydrate Polymers*, **313**, 120898

Manuscript received June 7, 2024; accepted for publication October 7, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 59-74, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195695

LAYERED MECHANICS MODEL FOR SURFACE MOVEMENT AND DEFORMATION DUE TO CONFINED AQUIFER PUMPING

Liangliang Zhang

School of Civil Engineering and Architecture, Anhui University of Science and Technology, Huainan, 232001, PR China
corresponding author e-mail: zllaust@163.com

Hua Cheng

School of Civil Engineering and Architecture, Anhui University of Science and Technology, Huainan, PR China and School of Resources and Environmental Engineering, Anhui University, Hefei, PR China

Surface movement and deformation caused by pumping from confined aquifers pose serious threats to safety and stability of surface buildings. A model for surface movement and deformation due to confined aquifer pumping is established based on the principle of coordinated strata deformation in which the effects of aquifer thickness, elastic modulus, pumping capacity, and overlying strata thickness are comprehensively considered. The model calculation results for surface subsidence caused by pumping from a confined aquifer in Dezhou City, Shandong Province, China, are compared with numerical simulation results to verify the accuracy and rationality of the model.

Keywords: confined aquifer; pumping well; layered mechanics; surface subsidence

1. Introduction

The long-term overexploitation of groundwater resources due to rapid urban development has resulted in decreasing groundwater levels and subsequent surface movement and deformation. This deformation seriously degrades the stability of the surface, which not only causes ground buildings to collapse and huge economic losses, but also serious damage to the natural environment and ecosystem (Loáiciga, 2013; Giang, 2022; Li *et al.*, 2021; Wang *et al.*, 2019). Surface subsidence was observed in Mexico City in 1891 (Ortega-Guerrero *et al.*, 1999). Overexploitation of groundwater in the Rafsanjan Plain in southeastern Iran has resulted in an annual subsidence rate of over 50 mm/year with local subsidence rates at some places exceeding 300 mm/year from 2006 to 2016. Since then, surface movement caused by groundwater pumping has been studied by researchers in the United States, Italy, Japan, and other countries (Abidin *et al.*, 2011; Holzer and Johnson, 1985). The identified land subsidence areas in China are mainly distributed in the Yangtze River Delta, North China Plain, and other regions, covering a total area exceeding 90000 km^2 (Cui and Tang, 2007).

Extensive in-depth research has been performed on ground subsidence caused by groundwater exploitation using two main types of methods (Xu *et al.*, 2015; Motagh *et al.*, 2017). In the first type, soil consolidation theory is used to calculate strata displacement and deformation caused by water loss in aquifers. Terzaghi proposed a well-known one-dimensional consolidation theory based on a series of assumptions and established the relationship between surface subsidence and time (Guido *et al.*, 2023). Giao (1997) proposed a subsidence calculation method based on one-dimensional consolidation theory combined with Green's theorem in which the dissipative pore water pressure is used as a variable. Peng *et al.* (2023) conducted high-pressure long-term consolidation joint tests on the bottom aquifer of thick alluvial layers based on the ETAS test system, and studied the evolution law of the permeability coefficient.

The second type of methods for calculating surface movement and deformation during groundwater extraction is based on probability integration. Drawing on the theory developed for coal seam mining subsidence, the vertical compression of an aquifer caused by groundwater extraction is treated as a coal seam mining unit of varying thickness. The probability integration method is then used to calculate the surface subsidence based on random media theory (Malinowska *et al.*, 2020; Szojda and Kapusta, 2023). Tang and Bai (2011) analyzed the soil mechanics mechanism of water loss consolidation in saturated loess layers caused by mining and established a model for the additional surface subsidence deformation caused by water loss consolidation in the mining loess layers using random medium theory. Cheng *et al.* (2022) established a three-dimensional model for calculating surface movement and deformation under the combined action of a thick loose layer, thin-bedrock coal seam mining, and aquifer drainage using soil consolidation theory and random medium theory.

Although the two types of methods for calculating surface subsidence mentioned above have been widely applied in practical engineering, they still suffer from some shortcomings that need to be addressed. Consolidation theory models do not accurately reflect the surface subsidence changes in the entire aquifer water loss area while the probability integration methods do not account for the effects of physical and mechanical parameters and layered distribution characteristics of strata above an aquifer on surface movement and deformation. In this study, the influence of physical and mechanical parameters of each stratum is considered using layered mechanics to establish a model for surface movement and deformation caused by pumping in a confined aquifer. The influence of pumping capacity, aquifer thickness, aquifer elastic modulus, and overlying strata thickness on surface subsidence and horizontal movement are analyzed. This study provides a theoretical basis for safety assessment in surface buildings for groundwater extraction and environmental protection.

2. Surface movement and deformation model

Heavy industrial and residential water use has led to frequent overexploitation of confined aquifers in strata. Before water pumping, the overlying strata on a confined aquifer are jointly supported by pore water and the soil skeleton. The stress and deformation are in equilibrium. Pumping results in a gradual decrease of the pore water pressure. A portion of the soil pressure originally borne by pore water is transferred to the soil skeleton, resulting in an increase in the effective stress and consolidation deformation of the aquifer (Galloway and Burbey, 2011; Zhou *et al.*, 2010). The continuous development of this deformation upward causes surface movement and deformation. Although the original equilibrium state is broken by pumping, the collaborative soil deformation and stress redistribution in the formation cause the entire formation to eventually reach a new equilibrium state. The differences in permeability characteristics, physical and mechanical parameters, and thicknesses of the confined aquifer and overlying strata result in varying deformation transmission and development across different strata. Therefore, the strata can be regarded as a layered stratum for calculating surface movement deformation caused by pumping from a confined aquifer. In our model, the self-consolidation deformation caused by pumping from a confined aquifer is first calculated using layered mechanics. The deformation at the top of the aquifer is then used as a disturbance source at the bottom of the overlying strata to obtain the surface movement deformation caused by the deformation at the bottom of the overlying strata based on deformation coordination conditions at the soil interface.

In layered mechanics, the displacement and stress at any stratum depth can be represented by the surface displacement and stress (Bobileva, 2015)

$$\left\{\begin{array}{c}\overline{u}(\zeta,H)\\ \overline{w}(\zeta,H)\\ \dfrac{\overline{\sigma}(\zeta,H)}{E}H\\ \dfrac{\overline{\tau}(\zeta,H)}{E}H\end{array}\right\}=\begin{bmatrix}A_{11} & A_{12} & A_{13} & A_{14}\\ A_{21} & A_{22} & A_{23} & A_{24}\\ A_{31} & A_{32} & A_{33} & A_{34}\\ A_{41} & A_{42} & A_{43} & A_{44}\end{bmatrix}\left\{\begin{array}{c}\overline{u}(\zeta,0)\\ \overline{w}(\zeta,0)\\ \dfrac{\overline{\sigma}_z(\zeta,0)}{E}H\\ \dfrac{\overline{\tau}_{zr}(\zeta,0)}{E}H\end{array}\right\} \tag{2.1}$$

where $\overline{u}(\zeta,H)$, $\overline{w}(\zeta,H)$, $\overline{\sigma}(\zeta,H)$, and $\overline{\tau}(\zeta,H)$ are the Hankel integral transformation equations for the horizontal displacement, subsidence, normal stress, and shear stress of the strata at depth H, respectively; $\overline{u}(\zeta,0)$, $\overline{w}(\zeta,0)$, $\overline{\sigma}(\zeta,0)$, and $\overline{\tau}(\zeta,0)$ are the Hankel integral transformation equations for the horizontal displacement, subsidence, normal stress, and shear stress at the surface, respectively; ζ is the integral parameter variable; and E is the elastic modulus.

The matrix coefficients are given by

$$\begin{aligned}
&A_{11}=A_{44}=\cosh\eta+\frac{\eta\sinh\eta}{2(1-\mu)}\\
&A_{12}=-A_{34}=\frac{1}{2(1-\mu)}\,[(1-2\mu)\sinh\eta+\eta\cosh\eta]\\
&A_{13}=-A_{24}=\frac{1+\mu}{2(1-\mu)}\sinh\eta \qquad A_{14}=\frac{1+\mu}{2\eta(1-\mu)}\,[(3-4\mu)\sinh\eta+\eta\cosh\eta]\\
&A_{21}=-A_{43}=\frac{1}{2(1-\mu)}\,[(1-2\mu)\sinh\eta-\eta\cosh\eta]\\
&A_{22}=A_{33}=\cosh\eta-\frac{\eta\sinh\eta}{2(1-\mu)}\\
&A_{23}=\frac{1+\mu}{2\eta(1-\mu)}\,[(3-4\mu)\sinh\eta-\eta\cosh\eta]\\
&A_{31}=-A_{42}=-\frac{\eta^2}{2(1-\mu^2)}\sinh\eta\\
&A_{32}=\frac{\eta}{2(1-\mu^2)}(\sinh\eta-\eta\cosh\eta) \qquad A_{41}=\frac{\eta}{2(1-\mu^2)}(\sinh\eta+\eta\cosh\eta)
\end{aligned} \tag{2.2}$$

where $\eta=\zeta H$, μ is Poisson's ratio, and sinh and cosh are hyperbolic sine and cosine functions, respectively

$$\begin{aligned}
u(r,H)&=\int\limits_0^\infty\overline{u}(\zeta,H)\zeta J_1(\zeta r)\,d\zeta \qquad & w(r,H)&=\int\limits_0^\infty\overline{w}(\zeta,H)\zeta J_0(\zeta r)\,d\zeta\\
\sigma(r,H)&=\int\limits_0^\infty\overline{\sigma}(\zeta,H)\zeta J_0(\zeta r)\,d\zeta \qquad & \tau(r,H)&=\int\limits_0^\infty\overline{\tau}(\zeta,H)\zeta J_1(\zeta r)\,d\zeta
\end{aligned} \tag{2.3}$$

where J_0 and J_1 are the zeroth- and first-order Bessel functions, respectively; and $u(r,H)$, $w(r,H)$, $\sigma(r,H)$, and $\tau(r,H)$ are the horizontal displacement, subsidence, normal stress, and shear stress of the strata at a horizontal distance r and depth H.

The pumping process of a confined aquifer is shown in Fig. 1. The part above the aquifer are the overlying strata and that below the stable aquifuge layer. Pumping changes only the water level of the aquifer but does not affect the deformation of its lower strata. The deformation of the upper strata is only dependent on the elastic modulus, Poisson's ratio, and thickness. Let the thickness of the confined aquifer be H_0, the thickness of the overlying loose layer H_w, the radius of the pumping well r_0, the pumping influence radius R, and the pumping capacity Q. After the

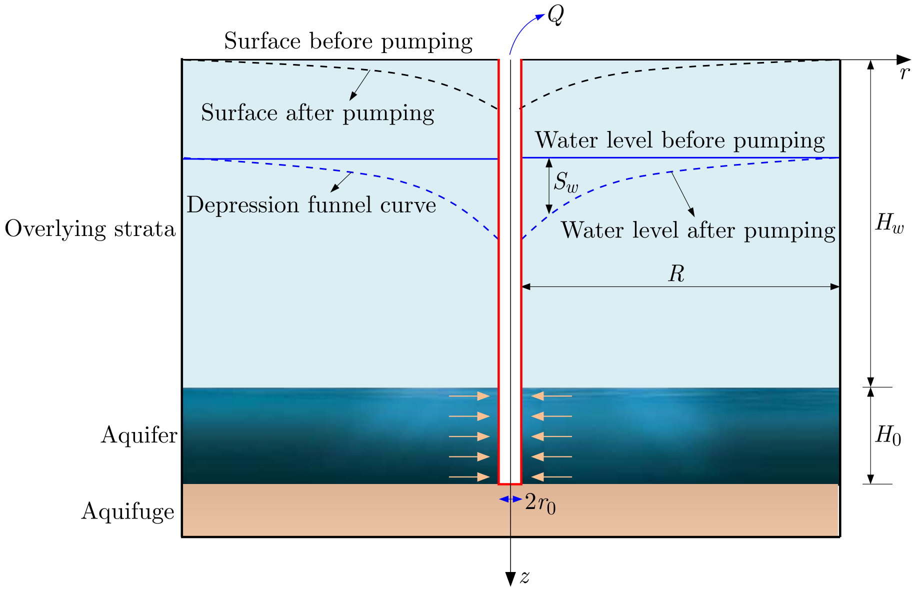


Fig. 1. Strata movement and deformation caused by pumping from a confined aquifer

pumping well has reached a stable state, the funnel curve of the confined water level drop can be expressed using the Dupuit formula (Bourel, 2024; Gamnitzer and Hofstetter, 2016)

$$s_w = \frac{Q}{2\pi k H_0} \ln \frac{R}{r} \tag{2.4}$$

where s_w is the depth of the water level drop and k is the permeability coefficient of the aquifer.

R is the pumping influence radius, which can be expressed using the Siechardt empirical formula (Jia *et al.*, 2016)

$$R = 10 s_w^{max} \sqrt{k} \tag{2.5}$$

where s_w^{max} is the maximum depth of the water level drop.

Focusing on the confined aquifer, the confined aquifer before pumping exerts a supporting force on the overlying strata with a magnitude directly proportional to the height of the confined water level. Pumping decreases the supporting force of the confined aquifer on the overlying strata and increases the effective stress of the soil (Liu *et al.*, 2023), which is directly proportional to the depth of the water level drop. This increase in the effective stress is equivalent to an additional load acting on the top of the aquifer with a funnel-shaped load distribution, which has the same funnel-shaped form as that of the aquifer level drop, as shown in Fig. 2.

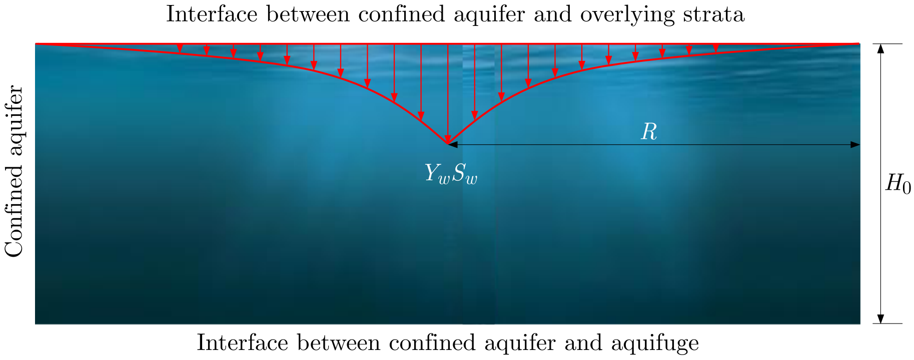


Fig. 2. Additional load acting on the top of the confined aquifer

The additional load at the top of the aquifer can be expressed as

$$\sigma_z(r,H_w)=\gamma_w s_w \qquad \tau_{zr}(r,H_w)=0 \tag{2.6}$$

where γ_w is the weight of water and $\sigma_z(r,H_w)$ and $\tau_{zr}(r,H_w)$ are the normal and shear stresses acting on the top of the aquifer, respectively.

Substituting Eq. (2.6) into Eq. (2.3) yields

$$\overline{\sigma}_z(\zeta,H_w)=\int\limits_0^R \gamma_w\Big[\frac{Q}{2\pi kH_0}\ln\frac{R}{r}\Big]rJ_0(\zeta r)\;dr \tag{2.7}$$

where $\overline{\sigma}_z(\zeta,H_w)$ is the Hankel integral transformation of $\sigma_z(r,H_w)$.

It is generally believed that hydrophobic consolidation in an aquifer does not cause bottom deformation. The displacement at the aquifer bottom can therefore be assumed to be 0 (Jia *et al.*, 2016)

$$\overline{u}(\zeta,H_w+H_0)=0 \qquad \overline{w}(\zeta,H_w+H_0)=0 \tag{2.8}$$

where $\overline{u}(\zeta,H_w+H_0)$ and $\overline{w}(\zeta,H_w+H_0)$ are the Hankel integral transformation equations for the horizontal displacement and subsidence at the bottom of the confined aquifer, respectively.

By substituting Eqs. (2.7) and (2.8) into Eq. (2.1), the displacement and deformation at the top of the aquifer are obtained as follows

$$\begin{aligned}\overline{u}(\zeta,H_w)&=-\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\times\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\\ \overline{w}(\zeta,H_w)&=\frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\times\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\end{aligned} \tag{2.9}$$

where E_2 is the elastic modulus of the confined aquifer and $\overline{u}(\zeta,H_w)$ and $\overline{w}(\zeta,H_w)$ are the Hankel integral transformation equations for the horizontal displacement and subsidence at the top of the confined aquifer, respectively.

Substituting Eq. (2.9) into Eq. (2.3) yields the horizontal displacement $u(r,H_w)$ and subsidence $w(r,H_w)$ at the top of the aquifer as follows

$$\begin{aligned}u(r,H_w)&=\int\limits_0^\infty -\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\times\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\zeta J_1(\zeta r)\;d\zeta\\ w(r,H_w)&=\int\limits_0^\infty \frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\times\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\zeta J_0(\zeta r)\;d\zeta\end{aligned} \tag{2.10}$$

Equation (2.10) expresses the deformation at the top of the confined aquifer caused by pumping. According to the principle of deformation coordination, the displacement distribution at the bottom of the overlying strata is the same as that at the top of the aquifer; the horizontal displacement and subsidence at the bottom of the overlying strata are therefore also given by $\overline{u}(\zeta,H_w)$ and $\overline{w}(\zeta,H_w)$, respectively. Therefore, the problem of determining the surface movement and deformation caused by pumping is transformed into the problem of determining the surface movement and deformation with a known displacement at the bottom of the overlying strata.

We focus on the overlying strata. If the surface does not bear loads

$$\overline{\sigma}(\zeta,0)=0 \qquad \overline{\tau}(\zeta,0)=0 \tag{2.11}$$

the corresponding surface movement deformation can then be obtained as follows by substituting Eq. (2.11) into Eq. (2.1)

$$\left\{\begin{array}{c}\overline{u}(\zeta,0)\\ \overline{w}(\zeta,0)\end{array}\right\}=\left[\begin{array}{cc}\dfrac{B_{22}}{B_{11}B_{22}-B_{21}B_{12}} & \dfrac{-B_{12}}{B_{11}B_{22}-B_{21}B_{12}}\\ \dfrac{B_{11}}{B_{11}B_{22}-B_{21}B_{12}} & \dfrac{-B_{21}}{B_{11}B_{22}-B_{21}B_{12}}\end{array}\right]\left\{\begin{array}{c}\overline{w}(\zeta,H_w)\\ \overline{u}(\zeta,H_w)\end{array}\right\} \tag{2.12}$$

where the matrix coefficients are given by

$$\begin{aligned}&B_{11}=\cosh m+\frac{m\sinh m}{2(1-\mu_1)} \qquad B_{12}=\frac{1}{2(1-\mu_1)}[(1-2\mu_1)\sinh m+m\cosh m]\\ &B_{21}=\frac{1}{2(1-\mu_1)}[(1-2\mu_1)\sinh m-m\cosh m] \qquad B_{22}=\cosh m-\frac{m\sinh m}{2(1-\mu_1)}\end{aligned} \tag{2.13}$$

where $m=\zeta H_w$ and μ_1 is Poisson's ratio of the overlying strata.

Equation (2.9) is substituted into Eq. (2.12) to obtain the surface movement deformation as follows

$$\begin{aligned}\overline{u}(\zeta,0)&=\Big(\frac{B_{22}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(-\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\\ &\quad+\Big(\frac{-B_{12}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(\frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\\ \overline{w}(\zeta,0)&=\Big(\frac{B_{11}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(\frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\\ &\quad+\Big(\frac{-B_{21}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(-\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\end{aligned} \tag{2.14}$$

By simplifying the coefficients in Eq. (2.14), we obtain

$$\overline{u}(\zeta,0)=(C_1+C_2)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0 \qquad \overline{w}(\zeta,0)=(C_3+C_4)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0 \tag{2.15}$$

where the coefficients are given by

$$\begin{aligned}C_1&=\Big(\frac{B_{22}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(-\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\\ C_2&=\Big(\frac{-B_{12}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(\frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\\ C_3&=\Big(\frac{B_{11}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(\frac{A_{21}A_{13}-A_{11}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\\ C_4&=\Big(\frac{-B_{21}}{B_{11}B_{22}-B_{21}B_{12}}\Big)\Big(-\frac{A_{22}A_{13}-A_{12}A_{23}}{A_{11}A_{22}-A_{12}A_{21}}\Big)\end{aligned} \tag{2.16}$$

The Hankel integral inverse transformation is performed on the expressions in Eq. (2.15) to obtain the surface movement deformation caused by pumping from the confined aquifer

$$\begin{aligned}u(r,0)&=\int\limits_0^{\infty}(C_1+C_2)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\zeta J_1(\zeta r)\,d\zeta\\ w(r,0)&=\int\limits_0^{\infty}(C_3+C_4)\frac{\overline{\sigma}_z(\zeta,H_w)}{E_2}H_0\zeta J_0(\zeta r)\,d\zeta\end{aligned} \tag{2.17}$$

Equation (2.17) implies that the surface movement and deformation caused by pumping from a confined aquifer are closely dependent on the thickness, elastic modulus, and Poisson's ratio of the strata and the stress distribution in the aquifer caused by pumping. Compared with the traditional probability integral models, more influencing factors are considered in our layer mechanics model. The total surface movement deformation caused by simultaneous pumping of multiple aquifers can be obtained by using our model to calculate the surface movement deformation caused by pumping from each confined aquifer and then summing the movement deformations caused by pumping over all the aquifers using the superposition principle.

3. Model validation

To verify the accuracy and rationality of the theoretical model, the confined aquifer pumping in Dezhou City, Shandong Province, China in the reference (Jia *et al.*, 2021) was taken as an example. The geological distribution of the study area is shown in Fig. 3, and the physical and mechanical parameters are listed in Table 1. The surface movement and deformation caused by extraction of the same volume of water from two confined aquifers at a pumping capacity of 2441.66 m^3/d, pumping well radius of 1 m, and water density of 10 $\mathrm{kN/m}^3$ were simulated in the reference (Jia *et al.*, 2021). Based on the data provided in the reference, our theoretical model was used to independently calculate the surface subsidence deformation caused by pumping from the two confined aquifers under the same pumping conditions and the results compared with those from the simulation. The results are shown in Fig. 4.

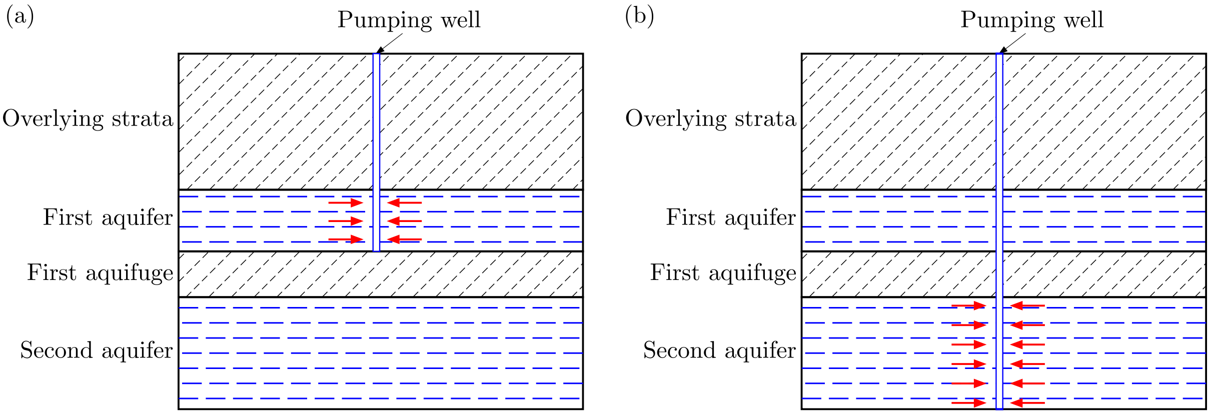


Fig. 3. Schematic of pumping from confined aquifers

Table 1. Calculation parameters

Stratum	Thickness [m]	E [MPa]	k [m/d]	μ	R [km]
Overlying strata	300	80	0.008	0.35	
First aquifer	150	120	1.8	0.25	2.0
First aquifuge	100	60	0.005	0.25	
Second aquifer	250	190	1.5	0.20	1.5

As shown in Fig. 4, the surface subsidence caused by pumping from a confined aquifer is symmetrically distributed with respect to the pumping well. The subsidence is larger closer to the pumping well. The theoretical curves for the surface subsidence caused by separate pumping of the same volume of water from the two confined aquifers are basically the same as the simulation curves. The high consistency of the theoretical calculation subsidence results with the simulation results verifies the correctness of the theoretical model.

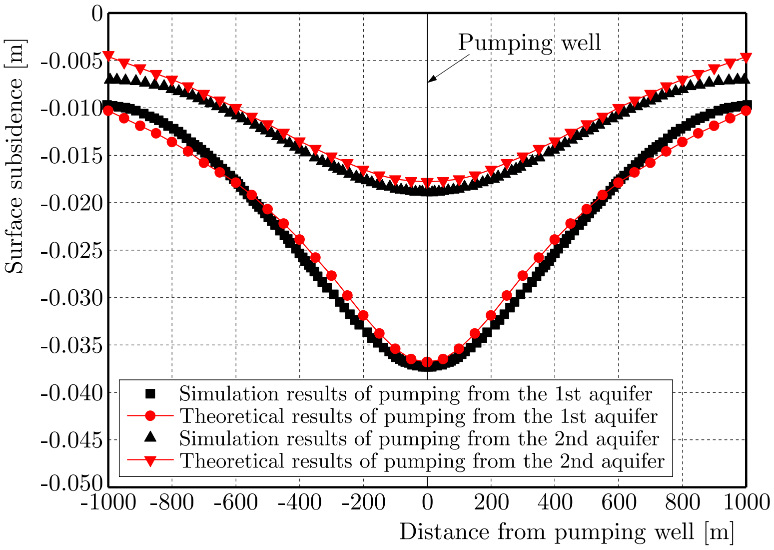


Fig. 4. Comparison of theoretical calculation and numerical simulation results

The maximum surface subsidence caused by individual pumping of the first confined aquifer in the simulation is 0.037 m while that in the theoretical calculation is 0.036 m. The error between them is only 0.01 m. The maximum surface subsidence caused by independent pumping of the second confined aquifer in the simulation is 0.019 m while that in the theoretical calculation is 0.018 m. The error of only 0.01 m between the two results indicates that the maximum surface subsidence in the theoretical calculation is essentially the same as that in the simulation. At a distance of 1000 m from the pumping well, the simulation values for the surface subsidence caused by the separate pumping of the first and second confined aquifers are 0.0097 and 0.0071 m, respectively, while the theoretically calculated values are 0.0103 and 0.0046 m, respectively. The respective errors of only 0.0006 and 0.0025 m are negligible. These error analysis results further confirm that the layered mechanics calculation for the surface subsidence caused by pumping of confined aquifers is reasonable and feasible.

4. Analysis of influencing factors

The first confined aquifer in Dezhou City, Shandong Province, China is taken as an example in this study for analyzing the effects of different pumping capacities and aquifer elastic moduli and thicknesses on the surface subsidence and horizontal displacement.

4.1. Pumping capacity

Equation (2.17) was used to calculate the surface subsidence and horizontal displacement at the pumping capacities of 1500, 2000, and 2500 m^3/d for a confined 150 m-thick aquifer with an elastic modulus of 120 MPa and overlying strata thickness of 300 m. The results are shown in Fig. 5.

The surface subsidence curve in Fig. 5a shows that the surface subsidence caused by pumping increases with the pumping capacity and is symmetrically distributed around the pumping well. The maximum surface subsidence occurs at the wellhead and gradually decreases along both its sides. At the pumping capacities of 1500, 2000, and 2500 m^3/d, the maximum surface subsidence values are 0.0226, 0.0302, and 0.0377 m, respectively, which correspond to increases of 0.0076 and 0.0075 m as the pumping capacity grows. At a distance of 1000 m from the pumping well, the surface subsidence for the three pumping capacities are 0.0063, 0.0085, and 0.0106 m, respectively, which correspond to increaments of 0.0022 and 0.0021 m as the pumping capacity grows.

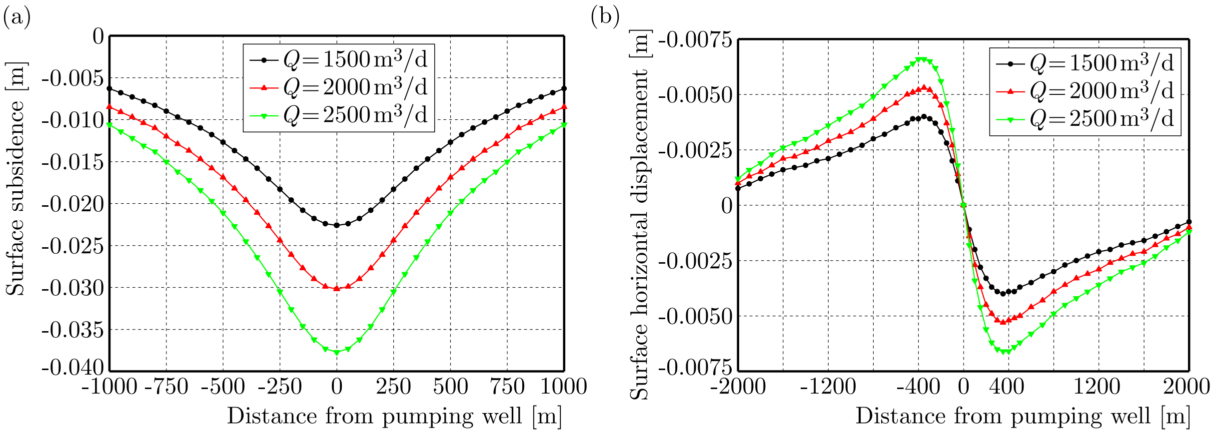


Fig. 5. Relationship between pumping capacity, surface movement, and deformation

These results indicate that increasing the pumping capacity causes a relatively uniform increase in the surface subsidence; however, the growth in the surface subsidence gradually decreases with increasing distance.

The surface horizontal displacement curve in Fig. 5b shows that the surface horizontal displacement caused by exploitation of the confined aquifer increases with the pumping capacity, and its shape is similar to that of a sine function. The surface horizontal displacement at the wellhead is approximately 0. The surface horizontal displacement gradually increases up to a range of approximately 310 m from the pumping well before decreasing with growing distance from the pumping well. Therefore, the maximum surface horizontal displacement occurs approximately 310 m away from the pumping well. For the pumping capacities of 1500, 2000, and 2500 m^3/d, the maximum surface horizontal displacements are 0.0040, 0.0053, and 0.0066 m, respectively, which correspond to increaments of 0.0013 and 0.0013 m. At a distance of 2000 m from the pumping well, the surface horizontal displacements are 0.0007, 0.0010, and 0.0012 m, respectively, which correspond to increaments of 0.0003 and 0.0002 m. These results indicate that increasing the pumping volume causes a relatively uniform growth in the horizontal displacement. The surface horizontal displacement is more sensitive to the pumping capacity nearer the pumping well. Figure 6 also shows that the surface subsidence is larger than the horizontal displacement, but the influence range of the latter is larger.

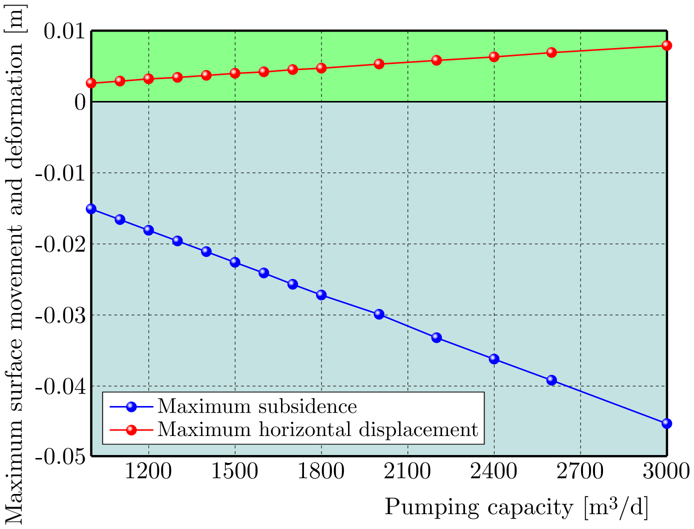


Fig. 6. Relationship between the maximum surface movement deformation and pumping capacity

Figure 6 shows the relationship between the maximum surface subsidence and horizontal displacement, and the pumping capacity. The maximum surface subsidence and horizontal displacement both increase linearly with the pumping capacity. As the pumping capacity increases from 1000 to 3000 m^3/d, the maximum surface subsidence increases from 0.0151 to 0.0453 m while the maximum surface horizontal displacement grows from 0.0026 to 0.0079 m. A higher pumping capacity results in a larger water level drop in the confined aquifer, which leads to a greater additional load acting on the interface between the top of the confined aquifer and the overlying strata. According to layered mechanics, the additional load results in a larger deformation of the confined aquifer, which is transmitted to the surface through the overlying strata and results in more significant surface movement and deformation.

4.2. Confined aquifer elastic modulus

Equation (2.17) was used to calculate the surface subsidence and horizontal displacement at the elastic modulus of 60, 90, and 120 MPa for a confined 150 m-thick aquifer with pumping capacities of 2441.66 m^3/d and overlying strata thickness of 300 m. The results are shown in Fig. 7.

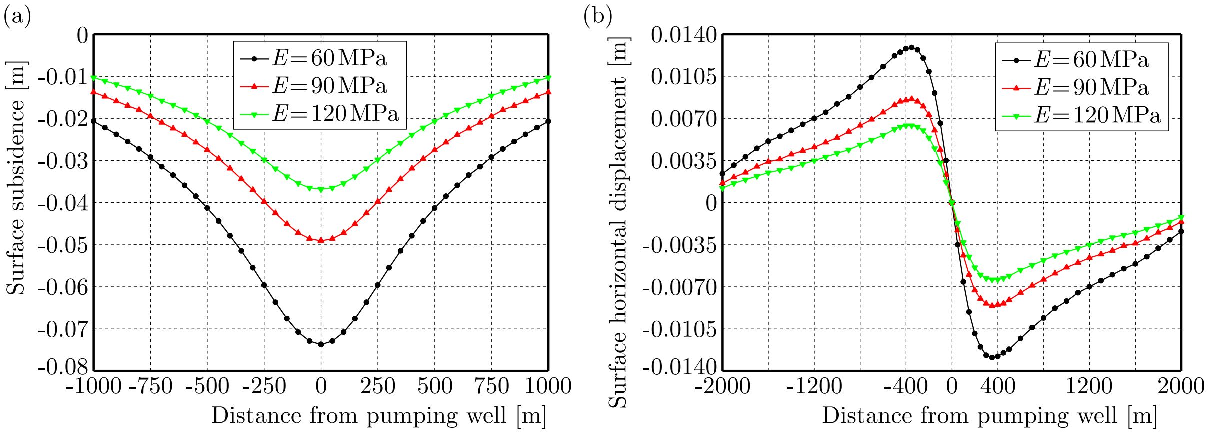


Fig. 7. Relationship between the elastic modulus, surface movement, and deformation

The surface subsidence curve in Fig. 7a shows that the surface subsidence caused by pumping decreases with the elastic modulus. At the elastic modulus of 60, 90, and 120 MPa, the maximum surface subsidence values are 0.0737, 0.0491, and 0.0368 m, respectively, which correspond to decreaments of 0.0246 and 0.0123 m as the elastic modulus increases. At a distance of 1000 m from the pumping well, the surface subsidence values for the three elastic modulus are 0.0207, 0.0138, and 0.0103 m, respectively, which correspond to decreaments of 0.0069 and 0.0035 m as the elastic modulus increases. These results indicate that increasing the elastic modulus causes a relatively uniform decrease in the surface subsidence.

The surface horizontal displacement curve in Fig. 7b shows that the surface horizontal displacement caused by the exploitation of the confined aquifer decreases with the elastic modulus, and its shape is similar to that of a sine function. The maximum surface horizontal displacement also occurs approximately 310 m away from the pumping well. For the elastic modulus of 60, 90, and 120 MPa, the maximum surface horizontal displacements are 0.0129, 0.0086, and 0.0064 m, respectively, which correspond to decreaments of 0.0043 and 0.0022 m. At a distance of 2000 m from the pumping well, the surface horizontal displacements are 0.0024, 0.0016, and 0.0012 m, respectively, which correspond to decreaments of 0.0008 and 0.0004 m. These results indicate that increasing the elastic modulus causes a smaller decrease in the surface horizontal displacement. The surface horizontal displacement is more sensitive to the elastic modulus nearer the pumping well.

Figure 8 shows the relationship between the maximum surface subsidence and horizontal displacement, and the elastic modulus. The maximum surface subsidence and horizontal displacement both decrease nonlinearly with the elastic modulus. As the elastic modulus increases from 10 to 200 MPa, the maximum surface subsidence decreases from 0.4423 to 0.0221 m while the maximum surface horizontal displacement decreases from 0.0772 to 0.0039 m. According to layered mechanics, a higher elastic modulus results in a smaller deformation of the confined aquifer, which is transmitted to the surface through the overlying strata and results in less significant surface movement and deformation.

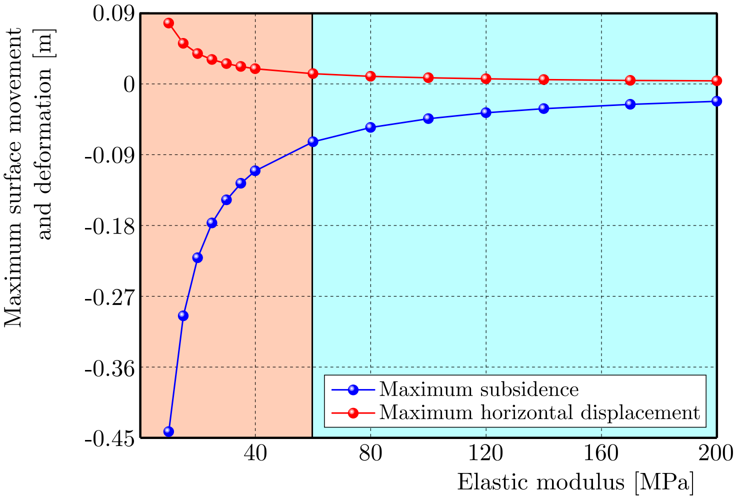


Fig. 8. Relationship between the maximum surface movement deformation and elastic modulus

4.3. Confined aquifer thickness

Equation (2.17) was used to calculate the surface subsidence and horizontal displacement at the thickness of 50, 100, and 150 m for a confined 2441.66 m^3/d – pumping capacities aquifer with an elastic modulus of 120 MPa and overlying strata thickness of 300 m. The results are shown in Fig. 9.

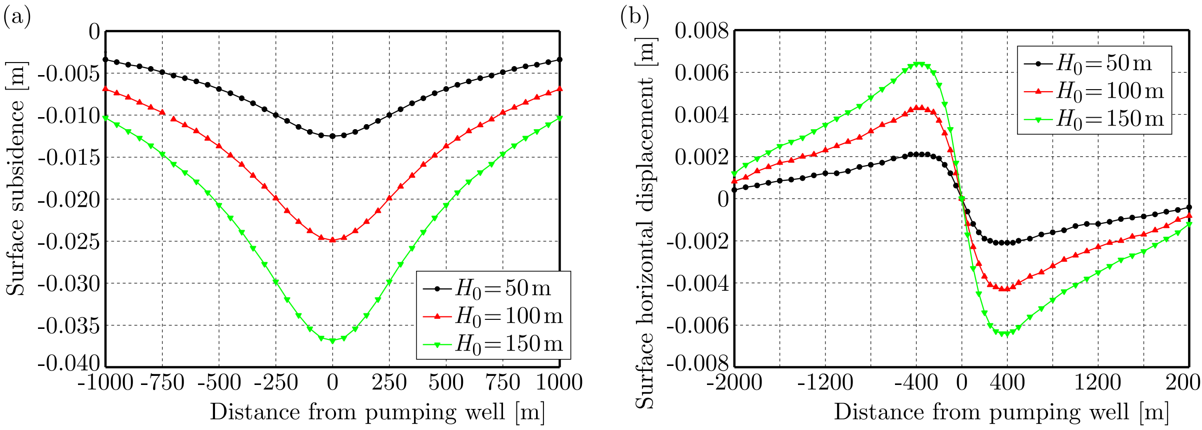


Fig. 9. Relationship between thicknesses, surface movement, and deformation

The surface subsidence curve in Fig. 9a shows that the surface subsidence caused by pumping increases with the aquifer thickness. At the aquifer thickness of 50, 100, and 150 m, the maximum surface subsidence values are 0.0125, 0.0249, and 0.0368 m, respectively, which correspond to increaments of 0.0124 and 0.0119 m as the aquifer thickness increases. At a distance of 1000 m from the pumping well, the surface subsidence values for the three aquifer thickness are 0.0034, 0.0069, and 0.0103 m, respectively, which correspond to increaments of 0.0035 and 0.0034 m as the aquifer thickness increases. These results indicate that increasing the aquifer thickness causes a relatively uniform increase in the surface subsidence.

According to the surface horizontal displacement curve in Fig. 9b, it can be seen that the surface horizontal displacement caused by exploitation of the confined aquifer increases with a growth of aquifer thickness, and the maximum surface horizontal displacement occurs about 310 m away from the pumping well. When the thickness of the aquifer is 50, 100, and 150 m, respectively, the maximum surface horizontal displacement is 0.0021, 0.0043, and 0.0064 m, with increaments of 0.0022 m and 0.0021 m, respectively. At a distance of 2000 m from the pumping well, the surface horizontal displacement is 0.0004, 0.0008, and 0.0012, respectively, with increaments of 0.0004 m and 0.0004 m, indicating that the increase in thickness of the aquifer causes a relatively uniform increase in surface horizontal displacement, and the elastic modulus of the aquifer is more sensitive to surface horizontal displacement near the pumping well than in the distance.

The surface horizontal displacement curve in Fig. 10b shows that the surface horizontal displacement caused by the exploitation of the confined aquifer increases with the aquifer thickness, and the maximum surface horizontal displacement occurs about 310 m away from the pumping well. For the aquifer thickness of 50, 1000, and 1500 m, the maximum surface horizontal displacements are 0.0021, 0.0043, and 0.0064 m, respectively, which correspond to increaments of 0.0022 and 0.0021 m. At a distance of 2000 m from the pumping well, the surface horizontal displacements are 0.0004, 0.0008, and 0.0012 m, respectively, which correspond to increaments of 0.0004 and 0.0004 m. These results indicate that increasing the aquifer thickness causes a relatively uniform increase in the surface horizontal displacement. The surface horizontal displacement is more sensitive to the pumping capacity nearer the pumping well.

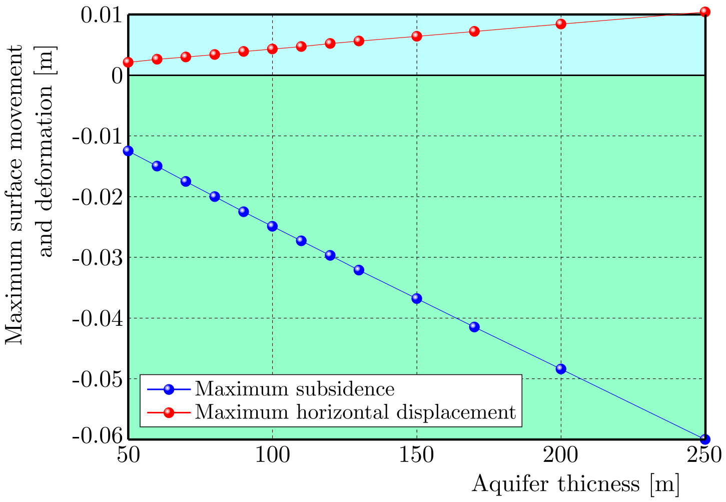


Fig. 10. Relationship between the maximum surface movement deformation and aquifer thickness

Figure 10 shows the relationship between the maximum surface subsidence and horizontal displacement, and the aquifer thickness. The maximum surface subsidence and horizontal displacement both increase linearly with the aquifer thickness. As the aquifer thickness increases from 50 to 250 m, the maximum surface subsidence increases from 0.0125 m to 0.0600 m while the maximum surface horizontal displacement increases from 0.0021 m to 0.0104 m. According to layered mechanics, a thicker aquifer results in a larger deformation of the confined aquifer, which is transmitted to the surface through the overlying strata and results in more significant surface movement and deformation.

5. Maximum displacement position

By analyzing the influencing factors, it was found that under different pumping volume and aquifer elastic modulus and thickness conditions, the maximum surface subsidence occurs at the wellhead, and the maximum horizontal displacement occurs at a distance of approximately 310 m from the pumping well. This indicates that the above influencing factors affect only the numerical value and range of the surface displacement deformation but not the position where the maximum surface displacement deformation occurs. Taking the first confined aquifer in Dezhou City as an example, the relationship between the position of the maximum surface horizontal displacement and the thickness of the overlying strata was analyzed. The results are shown in Fig. 11.

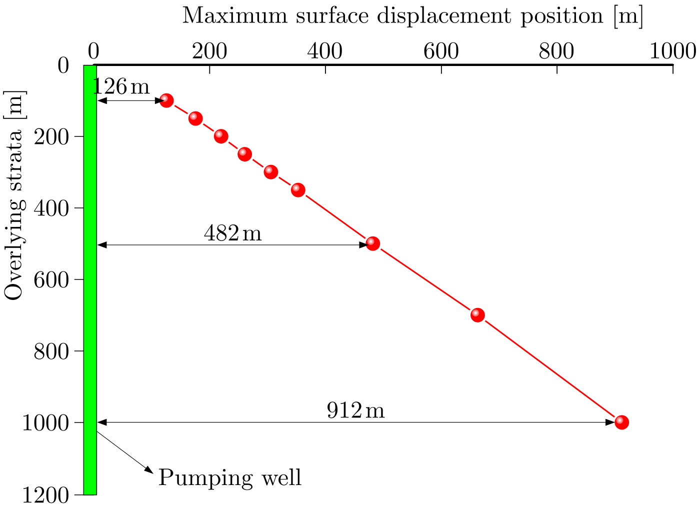


Fig. 11. Relationship between the maximum surface horizontal displacement position and overlying strata thickness

As shown in Fig. 11, the distance between the position of the maximum surface horizontal displacement and the pumping well increases linearly with the thickness of the overlying strata. At the overlying strata thicknesses of 100, 500, and 1000 m, the maximum surface horizontal displacements occur 126, 482, and 912 m away from the pumping well, respectively. The distance between the maximum surface horizontal displacement position and the pumping well increases with the thickness of the overlying strata at a rate of 0.87. In other words, for every 1 m increase in the thickness of the overlying strata, the maximum surface horizontal displacement occurs a further 0.87 m away from the pumping well. In contrast, the maximum surface subsidence occurs at the wellhead regardless of the aquifer or overlying strata parameters because the maximum water level drop caused by pumping in the confined aquifer occurs at the wellhead.

6. Discussion

Based on the principle of coordinated strata deformation, the paper uses the theory of layered mechanics to establish a calculation model for surface movement and deformation caused by pumping in confined aquifers. Compared with the consolidation theory model and the probability integration method in the calculation model, the model in this paper can consider not only the surface deformation within the range of pumping influence, but also reflect the influence of geological physical and mechanical parameters on surface movement and deformation. The theoretical model can comprehensively consider such factors as aquifer thickness, elastic modulus, pumping capacity, and overlying strata thickness, and is more reasonable compared to the existing computational models. By comparing with the monitoring data of surface deformation caused by pumping wells in Dezhou City, China, it is shown that the model is suitable for calculating surface movement and deformation caused by pumping in confined aquifers, and it also demonstrates that the model has high accuracy.

What is more, the theoretical model proposed in this paper is mainly used to calculate surface movement and deformation caused by pumping in confined aquifers, and has good applicability and high accuracy. However, in the field of coal resource extraction, surface movement and deformation often occur due to disturbances caused by coal mining, surface pumping, surface grouting, and other factors that affect the water level of the aquifer (Tajdus *et al.*, 2023; Dudek *et al.*, 2020), thereby affecting safety of surface buildings. The theoretical model in this paper is applicable for calculating surface movement and deformation caused by the above factors although still needs further modification and improvement.

7. Conclusion

(1) The effective stress increase caused by pumping from a confined aquifer is equivalent to an additional load acting on the top of the aquifer. A model for surface movement and deformation caused by pumping from a confined aquifer was established using layered mechanics in which the effects of the aquifer thickness, elastic modulus, pumping volume, and overlying strata thickness are considered. The consistency of the results for the surface subsidence caused by extraction of confined aquifers in Dezhou City, Shandong Province, China obtained using the model with numerical simulation results verifies the accuracy and rationality of the model.

(2) The surface subsidence and horizontal displacement caused by confined aquifer extraction are positively correlated with the pumping volume and aquifer thickness and negatively correlated with the aquifer elastic modulus. The maximum surface subsidence and horizontal displacement increase linearly with the pumping volume and aquifer thickness, and decrease nonlinearly with the increasing aquifer elastic modulus.

(3) The maximum surface subsidence caused by pumping from a confined aquifer occurs at the wellhead. The position of the maximum horizontal displacement is independent of the pumping volume, elastic modulus, and aquifer thickness. The distance between the position of the maximum horizontal displacement and the pumping well increases linearly with the thickness of the overlying strata.

Acknowledgments

This research was financially supported by the Natural Science Research Project of Anhui Educational Committee (Grant Numbers: 2023AH051203), the Scientific Research Start Funds of Anhui University of Science and Technology (Grant Numbers: 2022yjrc32), the National Natural Science Foundation of China (Grant Numbers: 52404069). The authors would like to thank the editor and reviewers for their contributions on the paper.

References

1. Abidin H.Z., Andreas H., Gumilar I., Fukuda Y., Pohan Y.E., Deguchi T., 2011, Land subsidence of Jakarta (Indonesia) and its relation with urban development, *Natural Hazards*, **59**, 3, 1753-1771

2. Bobileva T.N., 2017, Stress distribution in the stratified mass containing vertical alveole (in Russian), *Vestnik MGSU*, **12**, 8(107), 863-868

3. Bourel C., 2024, Water flow in shallow aquifers without the Dupuit hypothesis, *Computers and Mathematics with Applications*, **163**, 165-185

4. Cheng H., Zhang L.L., Yao Z.S., Peng S.L., Guo L.H., 2022, Mechanism of shaft deflection caused by asymmetric mining in thin bedrock and deep loose strata (in Chinese), *Journal of China Coal Society*, **47**, 1, 102-114

5. Cui Z.D., Tang Y.Q., 2007, Domestic and international recent situation and research of land subsidence disasters (in Chinese), *China Earthquake Engineering Journal*, **29**, 3, 275-278

6. Dudek M., Tajdus K., Misa R., Sroka A., 2020, Predicting of land surface uplift caused by the flooding of underground coal mines – A case study, *International Journal of Rock Mechanics and Mining Sciences*, **132**, 104377

7. Galloway D.L., Burbey T.J., 2011, Review: Regional land subsidence accompanying groundwater extraction, *Hydrogeology Journal*, **19**, 8, 1459-1486

8. Gamnitzer P., Hofstetter G., 2016, Fully coupled multi-phase modelling of pumping induced settlements, air- and water flow in multi-layered normally consolidated soils, *Computers and Geotechnics*, **79**, 10-21

9. Giang N.V., 2022, Application of artificial storage, recharge for semi-arid area in BacBinh Vietnam, *Archives of Mining Sciences*, **67**, 1, 37-53

10. Giao P.H., 1997, Artificial recharge of the Bangkok aquifer system for the mitigation of land subsidence, Dissertation No. GE-96-2, Asian Institute of Technology (AIT), Bangkok

11. Guido L., Libera E., Francesco F., 2023, Terzaghi's effective stress principle and hydrological deformation of Karst aquifers detected by GNSS measurements, *Rock Mechanics and Rock Engineering*, **57**, 4, 2365-2383

12. Holzer T.L., Johnson A.I., 1985, Land subsidence caused by ground water withdrawal in urban areas, *GeoJournal*, **11**, 3, 245-255

13. Jia C., Di S.T., Sun X.X., Zhang S.P., Ding P.P., Liu Z.T., 2021, Spatiotemporal evolution characteristics and transfer law of land subsidence in sand-clay interbed caused by exploiting the groundwater, *Arabian Journal for Science and Engineering*, **46**, 6, 5733-5753

14. Jia Y.J., Liang F.Y., Cui Z.D., Ye H., 2016, Analysis of soil deformation caused by decompression of confined water based on displacement coordination condition (in Chinese), *Rock and Soil Mechanics*, **37**, 42-48

15. Li M.G., Chen J.J., Xu Y.S., Tong D.G., Cao W.W., Shi Y.J., 2021, Effects of groundwater exploitation and recharge on land subsidence and infrastructure settlement patterns in Shanghai, *Engineering Geology*, **282**, 105995

16. Liu Z., Song Y., Zhou F., Wang L., 2023, Analytical solution for one-dimensional consolidation of unsaturated soils under dynamic load, *Journal of Theoretical and Applied Mechanics*, **61**, 1, 11-22

17. Loáiciga H.A., 2013, Consolidation settlement in aquifers caused by pumping, *Journal of Geotechnical and Geoenvironmental Engineering*, **139**, 7, 1191-1204

18. Malinowska A., Hejmanowski R., Dai H., 2020, Ground movements modeling applying adjusted influence function, *International Journal of Mining Science and Technology*, **30**, 2, 243-249

19. Motagh M., Shamshiri R., Haghighi M.H., Wetzel H.U., Akbari B., *et al.*, 2017, Quantifying groundwater exploitation induced subsidence in the Rafsanjan plain, southeastern Iran, using InSAR time-series and in situ measurements, *Engineering Geology*, **218**, 134-151

20. Ortega-Guerrero A., Rudolph D.L., Cherry J.A., 1999, Analysis of long-term land subsidence near Mexico City: field investigations and predictive modeling, *Water Resources Research*, **35**, 11, 3327-3341

21. Peng S., Li Z., Xu Y., Cao G., 2023, Study on the nonlinear permeability mechanism and pore structure characteristics of deep confined aquifers, *Applied Sciences*, **13**, 20, 11599

22. Szojda L., Kapusta L., 2023, Numerical analysis of buildings located on the edge of the post-mining basin, *Archives of Mining Sciences*, **68**, 1, 125–140

23. Tajdus K., Sroka A., Dudek M., Misa R., Hager S., Rusek J., 2023, Effect of the entire coal basin flooding on the land surface deformation, *Archives of Mining Sciences*, **68**, 3, 375-392

24. Tang F.Q., Bai F., 2011, Calculation method of surface subsidence caused by water loss in thick loess mining area (in Chinese), *Journal of Xian University of Science and Technology*, **31**, 4, 448-452

25. Wang X.W., Yang T.L., Xu Y.S., Shen S.L., 2019, Evaluation of optimized depth of waterproof curtain to mitigate negative impacts during dewatering, *Journal of Hydrology*, **577**, 123969

26. Xu Y.S., Yuan Y., Shen S.L., Yin Z.Y., Wu H.N., Ma L., 2015, Investigation into subsidence hazards due to groundwater pumping from Aquifer II in Changzhou, China, *Natural Hazards*, **78**, 1, 281-296

27. Zhou N.Q., Vermeer P.A., Lou R.X., Tang Y.Q., Jiang S., 2010, Numerical simulation of deep foundation pit dewatering and optimization of controlling land subsidence, *Engineering Geology*, **114**, 3-4, 251-260

Manuscript received July 10, 2024; accepted for publication November 7, 2024

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 75-89, Warsaw 2025
https://doi.org/10.15632/jtam-pl/196293

UTILIZING LINEAR QUADRATIC REGULATOR AND MODEL PREDICTIVE CONTROL FOR OPTIMIZING THE SUSPENSION OF A QUARTER CAR VEHICLE IN RESPONSE TO ROAD EXCITATION

Xhevahir Bajrami, Ahmet Shala, Ramë Likaj, Drin Krasniqi, Erjon Shala
University of Prishtina, Faculty of Mechanical Engineering, Departments of Mechatronics, Prishtina, Kosovo
corresponding author Drin Krasniqi, e-mail: drin.krasniqi@uni-pr.edu

Vehicle suspension systems are fundamental components designed to mitigate the adverse effects of road surface irregularities. These systems are typically categorized as passive, semi-active, or active suspensions. This study focuses on a quarter car suspension model to explore the application of two control methods, the Linear Quadratic Regulator (LQR) and the Model Predictive Control (MPC). Experimental data are collected using the Quanser active suspension experiment setup. Initially, the LQR controller is employed to optimize performance criteria related to the system state and input signals. Subsequently, the widely recognized MPC approach is used as an alternative control method. A comprehensive comparative analysis is conducted, taking into account various load conditions and parameter variations. Additionally, the study investigates system responses under varying road conditions, changes in plant characteristics, and the introduction of disturbances, to provide an exhaustive comparison of the two control methods. The results obtained with the MPC and the comparison with the findings of various authors to date allow us to emphasize that the presented results in this study significantly outperform the previous work. These outcomes have undergone rigorous validation on the physical model available in our mechatronics laboratory.

Keywords: vehicle suspension, quarter car model, linear quadratic regulator, model predictive control, experimental analysis

1. Introduction

In the realm of engineering and vehicle dynamics, the quest for optimal ride comfort and vehicle stability has driven the development of advanced suspension systems. Traditional passive suspensions have limitations in adapting to varying road conditions and vehicle dynamics. Active suspension systems, incorporating springs, dampers, and plate masses, have emerged as innovative solutions to overcome these limitations. These systems aim to dynamically adjust the suspension parameters in real time, offering the potential to enhance ride quality, handling, and overall vehicle performance.

This mathematical modelling study examines the intricacies of active suspension hardware, focusing on the integration of springs, dampers, and plate masses, to provide insights into the underlying principles and dynamics governing their behavior. In this paper, the focus is on advancing vehicle suspension systems to elevate ride quality, steering stability, passenger comfort, and mitigate concerns by Gandhi *et al.*, (2017). The study employs a Four Degrees of Freedom (4-DOF) half car active suspension model as the testing ground for a range of controllers (Likaj, 2005), including PID, LQR, Fuzzy Logic (Sahin, and Akalin, 2020). This paper explores the application of control strategies in the context of vehicle suspensions (Durmaz *et al.*, 2017). It examines three categories of suspensions (passive, semi-

-active, and active), with a specific focus on the quarter car suspension model. The study employs two prominent control methods, namely the Linear Quadratic Regulator (LQR) and Model Predictive Control (MPC), using the Quanser AS experiment set to gather experimental data.

Initially, the LQR is employed to optimize system performance based on predefined criteria related to the state and input signals. Subsequently, an MPC, a widely recognized industry-standard controller (Durmaz *et al.*, 2017), is utilized as a secondary control approach. Thus, we systematically investigate and compare the performance of LQR and MPC under various load conditions and parameter variations. Furthermore, this paper presents a straightforward yet effective control methodology for stabilizing the position of the sprung mass within the quarter car system by (Ovalle *et al.*, 2021).

These robust controllers guarantee exponential stability in the presence of non-vanishing disturbances, as substantiated by the Lyapunov function and stability analysis (Likaj *et al.*, 2016). Simulation and experimental Saha and Amrr (2020) results demonstrate the superiority of these control schemes over traditional linear methods, particularly within the Quanser ASS (Deshpande *et al.*, 2012; Fei *et al.*, 2022). Experimental results demonstrate that AIWPSO-tuned LQR significantly reduces vehicle body acceleration on uneven road surfaces, thereby ensuring passenger safety and enhanced ride comfort, as validated against ISO 2361-1 standards by (Reddipogu and Elumalai, 2020).

The system, requiring only the sprung mass position sensor, estimates states, uncertainties, and disturbances while ensuring overall stability (Wang and Zhou, 2019). A delay-dependent stability criterion is derived using the Lyapunov theory and the linear matrix inequality (LMI) method; this offers superior performance compared to traditional H-inf controllers by (Abdellahi *et al.*, 2000) as confirmed through simulation and experimental results. A novel Static Output Feedback (SOF) control approach is introduced for linear parameter-varying (LPV) systems, to ensure asymptotic stability and improved performance through gain-scheduled static output feedback (GS-SOF) controllers (Sereni *et al.*, 2020).

The design strategy incorporates a two-stage method, employing linear matrix inequalities (LMI) and Finsler's Lemma, and is demonstrated in practical applications for ASS control, showcasing its effectiveness (Pedro *et al.*, 2024). In this work, control objectives for nonlinear ASS with uncertain, time-varying constraints are addressed. The study demonstrates uniform boundedness and ultimate boundedness via Lyapunov analysis, supported by experimental and numerical simulations on a 2-DOF nonlinear ASS with uncertainties (Qin *et al.*, 2021; Zhang and Jing, 2021).

Experimental results demonstrate remarkable improvements in transient performance and energy efficiency, offering a fresh perspective on "robust and green" AS control for vehicles (Pusadkar *et al.*, 2019). In this paper, the focus is on addressing car body vibrations caused by track irregularities, to enhance ride quality (Eris *et al.*, 2015). In this work, a novel approach is introduced to enhance the performance of semi-active vehicle suspension systems (Ahmed and Svaricek, 2014; Deshpande *et al.*, 2017).

To minimize online computational requirements, control laws for various frequency ranges are derived from optimal controller data based on detected measured variable frequencies. Simulation and experimental results demonstrate improved ride comfort and road handling (Basturk, 2016). The study innovatively applies deep reinforcement learning to vehicle suspension control, emphasizing the adaptability and improved ride comfort. It introduces an enhanced DDPG algorithm, showcasing its efficiency with empirical samples, and establishes a physical model for comprehensive analysis (Liu *et al.*, 2020).

2. Mathematical modelling of active suspension hardware with springs, dampers, and plate masses

In this analysis, we will derive generalized dynamic equations of an active suspension system. The Free Body Diagram method is employed to capture the system dynamics as a dual-mass damper-spring model (Fig. 1). Within this approach, the two inputs to the system are considered to be commands of the AS control Fc, and the road surface position z_r (Apkarian and Abdossalami, 2013).

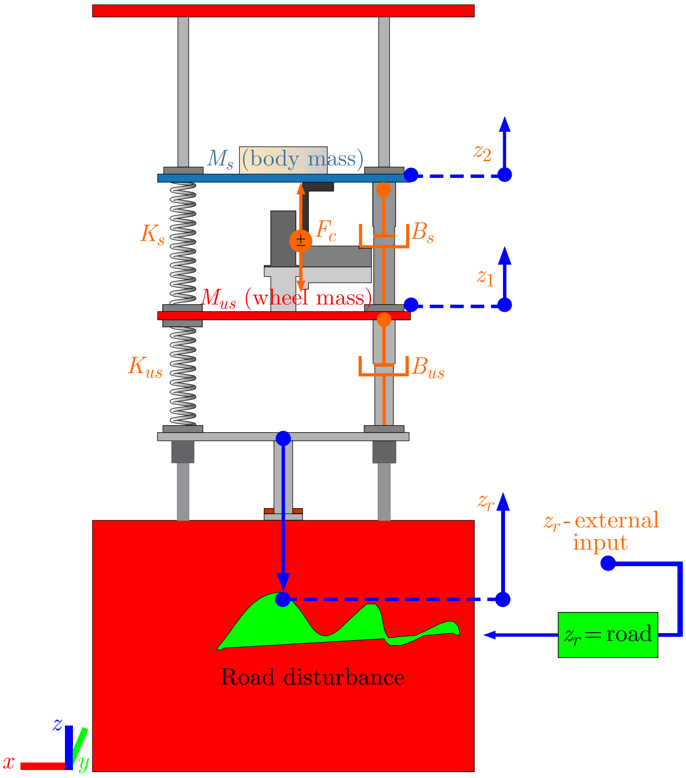


Fig. 1. System diagram of the active suspension system

Furthermore, the reference frames in Fig. 1 are utilized to define generalized coordinates, such as z_1 and z_2. The generalized coordinate z_1 represents the displacement of the suspension springs, while z_2 represents the displacement of the vehicle body.

Table 1. The numerical values linked with the system to be utilized for real and simulation conditions (Apkarian and Abdossalami, 2013)

Parameters	Name of parameters	Parameters values
M_s	Sprung mass	2.45 kg
M_{us}	Unsprung mass	1.0 kg
K_s	Suspension stiffness	900 N/m
K_{us}	Tire stiffness	1250 N/m
B_s	Suspension inherent damping coefficient	7.5 Ns/m
B_{us}	Tire inherent damping coefficient	5.0 Ns/m

2.1. Active suspension system EOMs

Within its operational range, we presume that the Active Suspension System can be accurately characterized as a linear system. The mathematical analysis is simplified by this linear approximation. We assume that the suspension displacements and vehicle body motions from their equilibrium positions are of small amplitude. Because of this assumption, we can linearize the system behavior and apply small-angle approximations. We consider a quasistatic analysis, which assumes that the dynamics of the Active Suspension System evolve at a slower timescale than the overall motion of the laboratory setup. Next, we individually analyze the masses of components to derive the equations of motion for accelerations, $\ddot{z}_1$ and $\ddot{z}_2$

$$\ddot{z}_2 = -g + \frac{F_c}{M_s} + \frac{B_s}{M_s}(\dot{z}_1 - \dot{z}_2) + \frac{K_s}{M_s}(z_1 - z_2) \tag{2.1}$$

The mass of the vehicle body M_s is presented below. The first system is described by its equations of motion (EOM). Likewise, the body diagram associated with the mass M_{us} is depicted in Fig. 2.

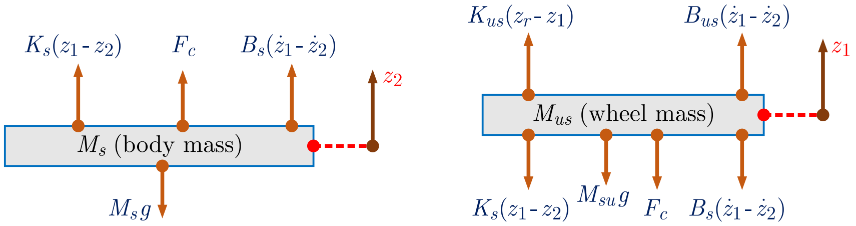


Fig. 2. System diagram of the upper plate and central plate, extracted from Fig. 1

The equations of motion (EOM) corresponding to this diagram are

$$\ddot{z}_1 = -g - \frac{F_c}{M_{us}} - \frac{B_s - B_{us}}{M_{us}}\dot{z}_1 - \frac{B_{us}}{M_{us}}\dot{z}_r + \frac{B_s}{M_{us}}\ddot{z}_2 - \frac{K_{us}}{M_{us}}z_r + \frac{K_s}{M_{us}}z_2 - \frac{K_s - K_{us}}{M_{us}}z_1 \tag{2.2}$$

We aim to demonstrate that the gravitational force solely alters the equilibrium points without influencing the system dynamics. When in the state of equilibrium, denoted by $z_1 = z_{q_1}$ and $z_2 = z_{q_2}$, the derivatives of z_1 and z_2 of any degree are all zero.

Additionally, the road surface z_r and all its derivatives, along with the control force F_c, are also zero in this equilibrium state. Substituting these conditions into Eqs. (2.1) and (2.2), we obtain

$$K_s z_{q_1} + K_{us} z_{q_1} - K_s z_{q_2} + M_{us} g = 0 \qquad K_s z_{q_2} - K_s z_{q_1} + M_s g = 0 \tag{2.3}$$

Subsequently, the equilibrium positions resulting from the influence of gravity are as follows

$$z_{q_1} = -\frac{M_{us} + M_s}{K_{us}} g \qquad z_{q_2} = \frac{K_s M_s + M_{us} K_s + M_{us} M_s}{K_s K_{us}} g \tag{2.4}$$

We implement the subsequent variable transformation to effectively eliminate the impact of gravitational forces from the equations

$$\begin{aligned} \dot{z}_1 &= \dot{z}_{us} \qquad & \dot{z}_2 &= \dot{z}_s \\ \ddot{z}_1 &= \ddot{z}_{us} \qquad & \ddot{z}_2 &= \ddot{z}_s \end{aligned} \tag{2.5}$$

By introducing the expression from Eq. (2.5) into Eqs. (2.1) and (2.2), we obtain

$$\begin{aligned} M_{us}\ddot{z}_{us} &= -F_c + B_{us}(\dot{z}_r - \dot{z}_{us}) + B_s(\dot{z}_s - \dot{z}_{us}) + K_s(z_s - z_{us}) + K_{us}(z_r - z_{us}) \\ M_s\ddot{z}_s &= F_c - B_s(\dot{z}_{us} - \ddot{z}_s) + K_s(z_{us} - z_s) \end{aligned} \tag{2.6}$$

Herein, the equations have been purged of gravitational influence, leaving only a shift in the system equilibrium point due to gravity. The state variables, coupled with the pair of inputs and the dual outputs, can be articulated as follows

$$z=\begin{bmatrix} z_s-z_{us}\\ \dot{z}_s\\ z_{us}-z_r\\ \dot{z}_{us}\end{bmatrix}\qquad u=\begin{bmatrix}\dot{z}_r\\ F_c\end{bmatrix}\qquad y=\begin{bmatrix} z_s-z_{us}\\ \ddot{z}_s\end{bmatrix} \tag{2.7}$$

By utilizing Eq. (2.6), it is possible to define A, B, C, and D in the subsequent manner

$$\begin{gathered}
A=\begin{bmatrix} 0 & 1 & 0 & -1\\ \frac{-K_s}{M_s} & -\frac{B_s}{M_s} & 0 & \frac{B_s}{M_s}\\ 0 & 0 & 0 & 1\\ \frac{K_s}{M_{us}} & \frac{B_s}{M_{us}} & \frac{-K_{us}}{M_{us}} & \frac{-B_s-B_{us}}{M_{us}}\end{bmatrix}\qquad D=\begin{bmatrix}0 & 0\\ 0 & \frac{1}{M_s}\end{bmatrix}\\
C=\begin{bmatrix}1 & 0 & 0 & 0\\ \frac{-K_s}{M_s} & \frac{-B_s}{M_s} & 0 & \frac{B_s}{M_s}\end{bmatrix}\qquad B=\begin{bmatrix}0 & 0\\ 0 & \frac{1}{M_s}\\ -1 & 0\\ \frac{B_{us}}{M_{us}} & -\frac{1}{M_{us}}\end{bmatrix}
\end{gathered} \tag{2.8}$$

3. Active suspension system using LQR

The Linear Quadratic Regulator (LQR) represents an unconstrained model-based control technique that discerns the optimal input $c_{(k_i)}=F_c$ through solution of an infinite horizon optimization problem

$$\begin{gathered}
\min J=\int\limits_0^\infty (z^{\mathrm{T}}_{(k_i)}Q_{z_{(k_i)}}+R^2_{c_{(k_i)}})\frac{dt}{dz}\qquad \text{subject to}\qquad z_{(k_i)}=Az_{(k_i)}+Bc_{(k_i)}\\
F_c=-K_{[k_i]}z_{(k_i)}
\end{gathered} \tag{3.1}$$

In this context, the gain $K_{[k_i]}$ is derived as follows

$$\begin{gathered}
K_{[k_i]}=[R+B^{\mathrm{T}}SB]^{-1}B^{\mathrm{T}}SA\\
A^{\mathrm{T}}SA-S(A^{\mathrm{T}}SB)(R+B^{\mathrm{T}}SB)^{-1}(B^{\mathrm{T}}SA)+Q_{z_{(k_i)}}=0
\end{gathered} \tag{3.2}$$

Given the assumption of linearity and time-invariance governing the two equations of motion (EOMs) associated with the ASS, their expression can be rendered in the state-space representation below

$$\begin{gathered}
Z_{(k_i+1)}=Az_{(k_i)}+Bc_{(k_i)}\\
c_{(k_i)}=-K_{[k_i]}z_{(k_i)}\\
Z_{(k_i+1)}=[A-BK_{[k_i]}]z_{(k_i)}=\varPhi z_{(k_i)})
\end{gathered} \tag{3.3}$$

The performance index J imposes penalties on the system state variables, namely the suspension displacement and tire deflection (both regarded as performance measures), as well as the velocities of the vehicle body and the tire. This is accomplished using the weighting matrix $Q_{z_{(k_i)}}$. The weighting matrix $Q_{z_{(k_i)}}$ is characterized by its symmetric, positive semidefinite nature, and it is essential for it to exhibit full rank. The performance index additionally encapsulates the constraints on control by subjecting the control input to penalties using the weighting coefficient R. The weighting matrices significantly influence the manner in which the LQR accomplishes its minimization objective, by essentially operating as tuning variables.

Algorithm 1 Parameters and description of calculation of $K_{[k]}$ based on LQR.

```
∀ parameters:
$A$ → System dynamics matrix, $B$ → Control input matrix, $C$ → Output matrix,
$Q$ → State cost matrix, $R$ → Control cost matrix, & $N$ → Time horizon
→ Initialization:
        $P_{[N-1]} = Q_f$ → Final state cost matrix.
        $K_{[N])}$→ Feedback control gain matrix.
Backward pass:
        for $k = N$ down to 1 do.
            $K_{[k]} = -\left\{(R + B^{\mathrm{T}} * P_{[k+1]} * B)^{-1} * B^{\mathrm{T}} P_{[k+1]} * A\right\}$
            $P_{[k]} = Q_z + A^{\mathrm{T}} * P_{[k+1]} * A - A^{\mathrm{T}} * P_{[k+1]} * B * K_{[k]}$
        end for
Forward pass:
        for $k = 1$ to $N$ do
            Calculate control input:
            $u_{[k]} = K_{[k]} * z_{[k]}$, where $z_{[k]}$ is the current state.
            → Apply the control input $u_{[k]}$ to the system.
            → Measure the new state $z_{[k+1]}$.
        end for
```

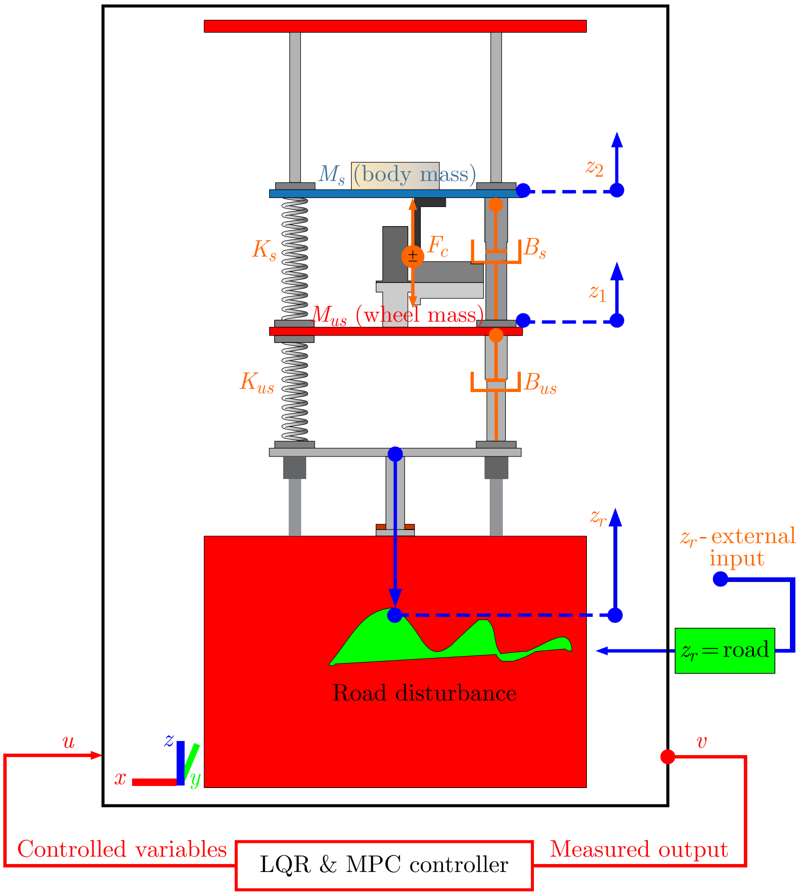


Fig. 3. Proposed system diagram of the LQR and MPC for the active suspension system

Devise this controller to effectively address a pulse road bump characterized by a 0.001 m amplitude and a frequency of 0.5 Hz. In Fig. 4, the transfer functions are extracted, representing all possible combinations of two inputs and two outputs from the ASS. This analysis provides a holistic perspective on the system behavior, which offers insights into how various inputs influence corresponding suspension outputs and assist in optimizing control strategies for enhanced performance.

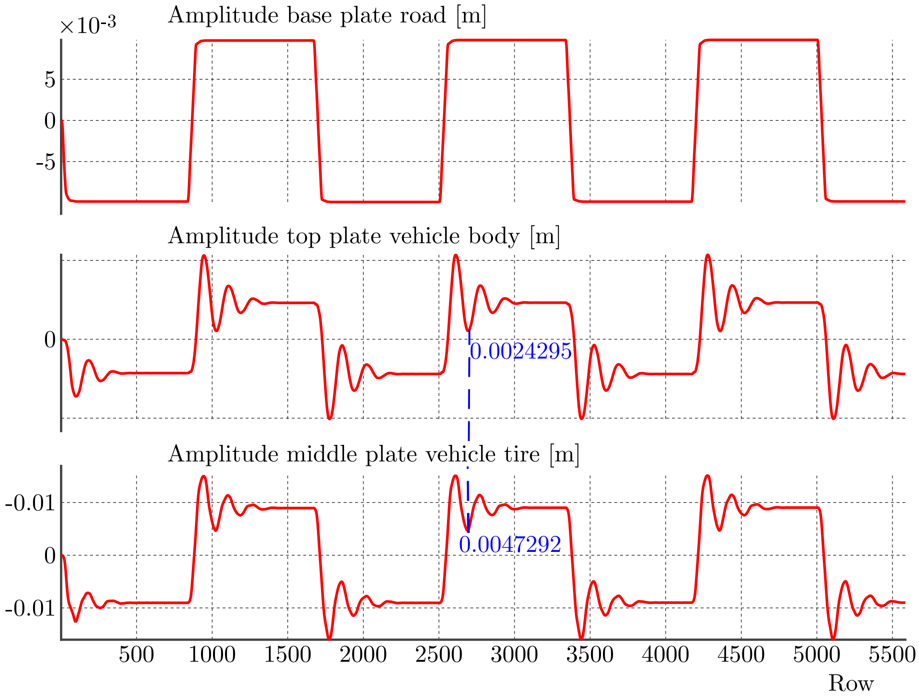


Fig. 4. The transfer functions for all combinations of two inputs, and two active hang outputs are extracted

Figure 5 shows the simulated active suspension using a road collision profile with a magnitude of 0.02 Hz and a phase of 5 Hz and demonstrates the system response to these parameters.

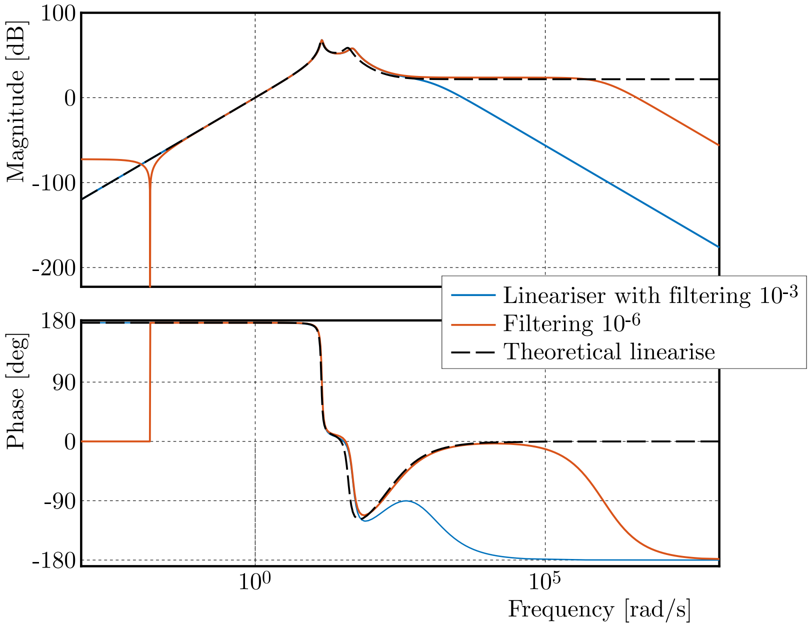


Fig. 5. Simulated active suspension with a road collision profile of magnitude 0.02 Hz and phase at 5 Hz

This graph offers insights into the dynamic behavior of the active suspension system, illustrating how the system control influences the simulated position and plate position. Such analyses are crucial for evaluating the efficacy of the active suspension system in maintaining the desired plate positions.

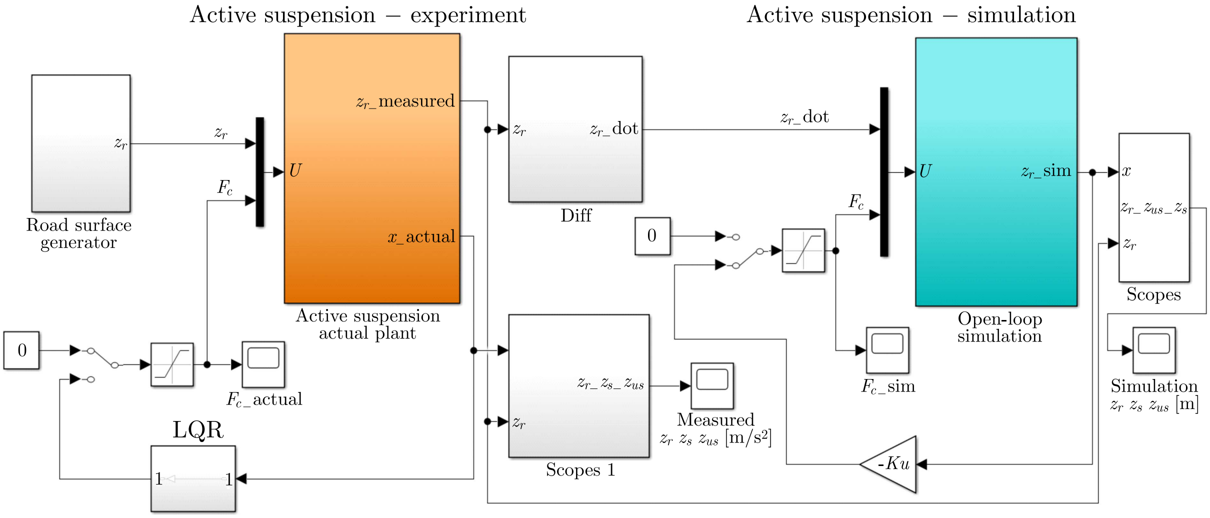


Fig. 6. Simulation scheme of the active suspension system: the real model is on the left side, and the simulation part on the right

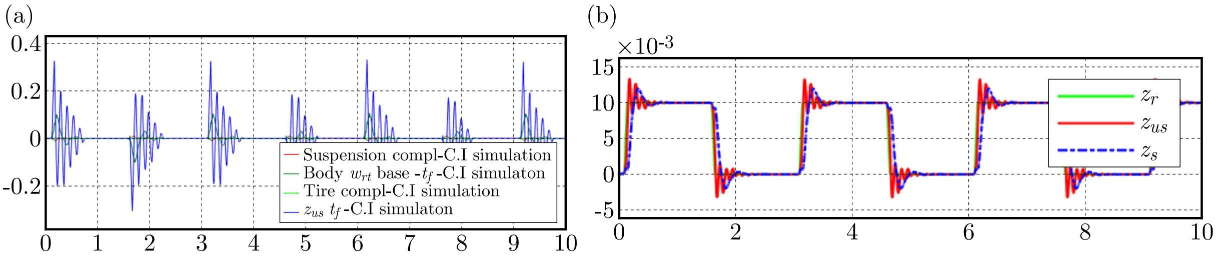


Fig. 7. Simulated closed-loop response for: (a) x_sim with the LQR, (b) plate position with the LQR

By observing the relationship between x_sim and $z_r\_z_{us}\_z_s$, engineers can optimize control strategies for improved ride quality and stability. This visual representation guides the refinement of active suspension algorithms and facilitates the creation of systems that effectively manage vehicle dynamics to ensure a smoother, more controlled driving experience. In Fig. 7a, a simulated closed-loop response is presented, depicting the correlation between x_sim and the plate position, represented as zr_zus_zs in Fig. 7b.

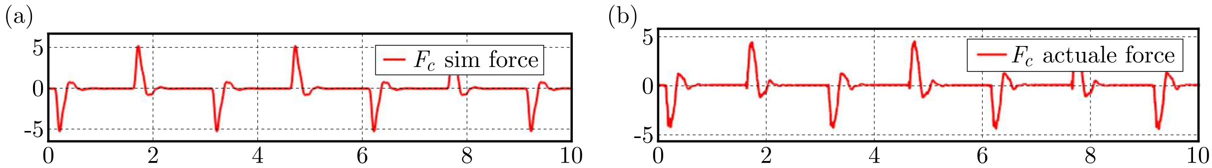


Fig. 8. (a) Simulated closed-loop response for control force with LQR. (b) Closed-loop response for control force measurement with LQR

Figure 8a demonstrates the simulated closed-loop response for the control force F_c [N], showcasing how the system control input correlates with the resulting force output. This offers insights into the control system efficacy. Figure 8b shows the control force measurement. This graph shows the correlation between the applied control force and the resulting measured control force, thus indicating the system responsiveness and accuracy in force regulation. The measured force closely matches the simulated force, validating the accuracy and responsiveness of the control system.

4. Active suspension system using MPC.

Model Predictive Control (MPC) is a control strategy that relies on the model of the system to make predictions and optimize control actions while considering constraints. A discrete-time representation of the plant is used to create an online optimization problem that incorporates constraints and is computationally manageable.

Figure 9 illustrates the receding horizon strategy; this involves computing an optimal input to ensure that the future evolution of outputs adheres to performance criteria and avoids constraint violations. While the optimal control calculates a sequence for the entire control horizon at each moment, only the initial component of this sequence is implemented in the plant, with the remaining components being disregarded. Based on the state-space model using matrices A, B, C and D in the future, the state variables are calculated sequentially using the set of future control parameters.

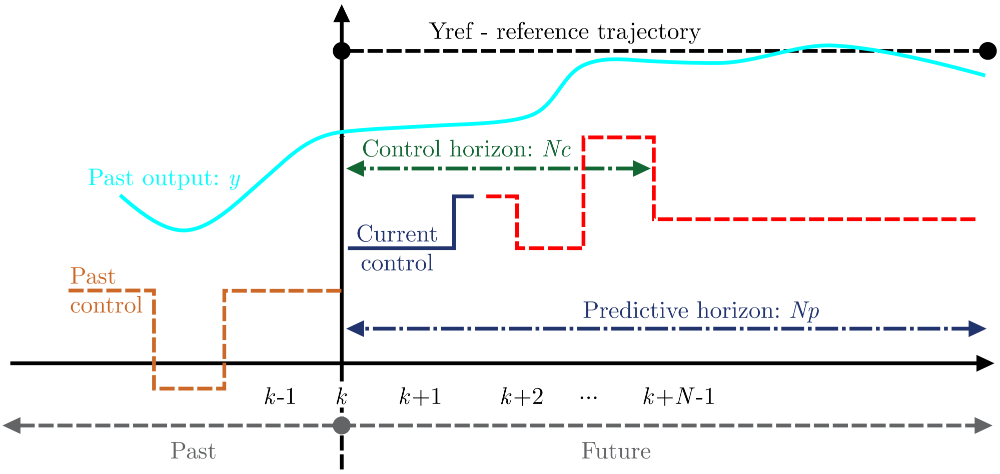


Fig. 9. Receding horizon strategy

The control horizon N_c and predictive horizon N_p for $z_{(k_i+N)}$ is the predictive state variable plan information $z_{(k_i)}$ for $i = 1, 2, \ldots$

$$\begin{aligned} z_{(k_i+1)} + z_{(k_i+2)} + \ldots + z_{(k_i+N_p)} &= Az_{(k_i)} + Bc_{(k_i)} + A^2 z_{(k_i+1)} + ABc_{(k_i+1)} + \ldots \\ &+ A^{N_p} z_{(k_i)} + A^{N_p-1} Bc_{(k_i)} + A^{N_p-2} Bc_{(k_i+1)} + \ldots + A^{N_p-N_c} Bc_{(k_i+N_c-1)} \end{aligned} \tag{4.1}$$

From the predicted state variables, the predicted output variable we define the vector $z_{(k+1)}$

$$\begin{aligned} z_{(k+1)} &= z_{(k_i+1)} + z_{(k_i+2)} + \ldots + z_{(k_i+N_p)} = ACz_{(k_i)} + BCc_{(k_i)} + A^2 Cz_{(k_i)} \\ &+ ABCc_{(k_i)} + ABCc_{(k_i)} + BCc_{(k_i+1)} + A^3 Cz_{(k_i)} + A^2 BCc_{(k_i)} + ABCc_{(k_i+1)} \\ &+ BCc_{(k_i+2)} + \ldots + A^{N_p} Cz_{(k_i)} + A^{N_p-1} CBc_{(k_i)} + A^{N_p-2} CBc_{(k_i+1)} + \ldots \\ &+ C^{N_p-N_c} CBc_{(k_i+N_c-1)} \qquad i = 1, 2, \ldots \end{aligned} \tag{4.2}$$

We collect equations x_1 and x_2 together in compact in the matrix form as

$$z_{(k_i+1)} = \begin{cases} \Phi z_{(k_i)} + \Gamma c_{(k_i)} & \text{for} \quad k_i \leqslant N_c \\ \Phi z_{(k_i)} & \text{for} \quad k_i > N_c \end{cases} \tag{4.3}$$

and

$$\vec{z}_{k_i+1} = \underbrace{\begin{bmatrix} AC \\ A^2C \\ A^3C \\ \vdots \\ CA^{N_p} \end{bmatrix}}_{\varPhi} + \underbrace{\begin{bmatrix} BC & 0 & 0 & \cdots & 0 \\ BAC & CB & 0 & \cdots & 0 \\ A^2BC & CAB & CB & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^{N_p-1}BC & A^{N_p-2}BC & A^{N_p-3}BC & \cdots & A^{N_p-N_c}BC \end{bmatrix}}_{\varGamma} \tag{4.4}$$

The incoming sequence has the formula

$$c_{(k)} = \begin{bmatrix} \varPhi(k_i|k_i) \\ \varPhi(k_i+1|k_i) \\ \vdots \\ \varPhi(k_i+N_p-1|k_i) \end{bmatrix} = \begin{bmatrix} c_{(k_i|k_i)}) & \cdots & c_{(k_i+N_p-1|k_i)} \end{bmatrix} \qquad \forall i = 1, 2, \ldots, N_c - 1 \tag{4.5}$$

In our case, we have taken the parameters: $N = 4$, and the matrices $\varPhi$ and $\varGamma$ are given below

$$\varGamma_{(N=4)} = \begin{bmatrix} A^{N-1}B & A^{N-2}B & A^{N-3}B & A^{N-4}B \end{bmatrix} = \begin{bmatrix} \varpi_1 & \varpi_2 & \varpi_3 & \varpi_4 \end{bmatrix} \tag{4.6}$$

Calculations of the parameters for ϖ_i, $i = 1, 2, 3, 4$ yield

$$\varpi_i = \begin{bmatrix} a_{11_{(ni)}} & a_{12_{(ni)}} \\ a_{21_{(ni)}} & a_{22_{(ni)}} \\ a_{31_{(ni)}} & a_{32_{(n1)}} \\ a_{41_{(ni)}} & a_{42_{(ni)}} \end{bmatrix} \qquad i = 1, 2, 3, 4 \tag{4.7}$$

Calculation of the performance index J, assuming that the forecasts are based on perturbations about the reaction $c_{k_i} = -K_{(k_i)} z_{k_i}$. The implemented MPC strategy involves using the initial element of the control trajectory. Observe the distinction between the Linear Quadratic Regulator (LQR) described in Eq. (3.1) and the MPC problem mentioned earlier. Initially, it is evident that the performance metric is now aggregated over a limited time horizon

$$\begin{gathered} [c_{k_i}, c_{k_i+1}, \ldots, c_{k_i+N}] = \arg(c_{k_i}, c_{k_i+1}, \ldots, c_{k_i+N}) \\ \min J = \sum_{k=0}^{N} c_{(k)}^{\mathrm{T}} H c_{(k)} + 2 z_{(k+1)}^{\mathrm{T}} F^{\mathrm{T}} \varPhi_{(k)} + z_{(k+1)}^{\mathrm{T}} G z_{(k)} z_{k_i+1} = A_s z_{k_i} + B_s c_k \end{gathered} \tag{4.8}$$

Taking into account

$$z_{min} \leqslant z_{k_i+1} \leqslant z_{max} \qquad c_{m}in \leqslant c_{k_i} \leqslant c_{max} \qquad k \in z_k^{\mathrm{T}} \qquad k = 1, 2, \ldots, N \tag{4.9}$$

Furthermore, the presence of supplementary constraints, related to the performance, provides bounds for both the system states and inputs. In summary, the performance index J becomes

$$J = c_{(k_i)}^{\mathrm{T}} H c_{(k_i)} + 2 z_{(k+1)}^{\mathrm{T}} F^{\mathrm{T}} \varPhi_{(k)} + z_{(k+1)}^{\mathrm{T}} G z_{(k)} \tag{4.10}$$

where the matrices are defined by

$$\begin{gathered} H = \varGamma^{\mathrm{T}} \widetilde{Q} \varGamma + \widetilde{R} \qquad F = \varGamma^{\mathrm{T}} \widetilde{Q} \varPhi \qquad G = \varPhi^{\mathrm{T}} \widehat{Q} \varPhi + Q \qquad Q = C^{\mathrm{T}} C \\ \nabla_{c_{(k_i)}} J = 2 H c_{(k_i)} + F z_{(k_i)} = \begin{bmatrix} \dfrac{\partial J}{\partial c_{(k_i|k_i)}} & \dfrac{\partial J}{\partial c_{(k_i+1|k_i)}} & \dfrac{\partial J}{\partial c_{(k_i+2|k_i)}} & \dfrac{\partial J}{\partial c_{(k_i+3|k_i)}} \end{bmatrix} \end{gathered} \tag{4.11}$$

```
Algorithm 2 Description of the calculation parameters for MPC.
∀ parameters:
$A$ → System dynamics matrix, $B$ → Control input matrix, $C$ → Output matrix,
$Q$ → State cost matrix,
$R$ → Control cost matrix, $N$ → Prediction horizon and $x_0$ → Initial state.
→Initialization:
        $z_{k+1} = z_0$ → Current state.
    for $t = 0$ to $T_{max}$ do
        → Solve the optimization problem to find the optimal control sequence $\{c_0, c_1, \ldots, c_{N-1}\}$.
        $\min J = \sum_{i=0\ to\ N-1} \{z_{[i]}^{\mathrm{T}} * Q * z_{[i]} + u_{[i]}^{\mathrm{T}} * R * c_{[i]}\}$
            subject to: $z_{[0]} = z_{k+1}$
            $z_{[i+1]} = A * z_{[i]} + B * c_{[i]}$, for $i = 0$ to $N - 1$
        → Constraints on states and inputs:
                $z_{min} \leqslant z_{k+1} \leqslant z_{max}$
                $c_{min} \leqslant c_k \leqslant c_{max}$
        ■ $z_{min}$ and $z_{max}a$ are vectors representing the lower and upper bounds
                for the state variables $(z_{k+1})$.
        ■ $c_{min}$ and $c_{max}a$ are vectors representing the lower and upper bounds
                for the input variables $(c_k)$.
        → Apply the first control input $u_0$ to the systems:
                $z_{k+1} = A * z_{k+1} + B * c_k$
        → Update the control input sequence for next iteration:
                $\{c_0, c_1, \ldots, c_{N-2} = c_1, c_2, \ldots, c_{N-1}, c_{N(x-1)}, u_{N-x}, \ldots\}$
    end for
```

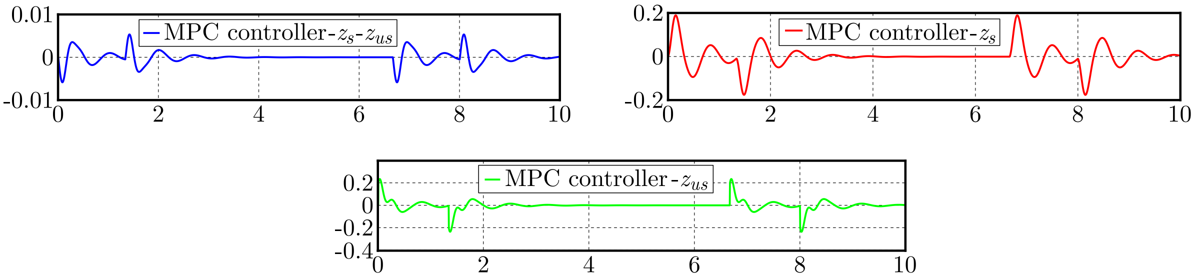


Fig. 10. Simulation model showing the dynamic behavior of the body-tire system, including the vertical displacement $(z_s - z_{us})$, body velocity $\dot{z}_s$, and tire velocity $\dot{z}_{us}$; all are effectively controlled and regulated using Model Predictive Control techniques

5. Experimental set-up of hardware-in-the-loop

Hardware-in-the-loop (HIL) simulation is a crucial method applied during the development and testing of intricate real-time embedded systems. This technique proves highly effective as it integrates the complexity of the process-actuator system, often referred to as the “plant”, into the testing environment. To accomplish this, a mathematical model representing all relevant dynamic systems, known as “plant simulation”, is incorporated. Consequently, the embedded system under evaluation interacts with this plant simulation, facilitating comprehensive testing and development (Hwang *et al.*, 2006). In the automotive sector, the active suspension technology plays a pivotal role in the ongoing management of the vertical wheel movement by employing actively controlled actuators positioned along the suspension axis. Furthermore, comparable technologies have found application in train bogies, where they serve to enhance the train handling during curves and reduce the perceived passenger discomfort resulting from abrupt accelerations (ASS, available online: accessed on 15 January 2024).

Table 2. Device specifications of Active Suspension System

$W \times L \times H$ [cm]	$30.5 \times 30.5 \times 61$
Mass [kg]	15
Range [mm]	[±22] (road), [±19] (tire), [±25.4] (car)
Position [mm/count]	0.002 (road), 0.005 (tire), 0.009 (body)
Stiffness [N/m]	0.4-2
Excitation frequency [Hz]	Up to 15
Resonant [Hz]	2 and 6

5.1. Numerical and experimental analyses of ASS hardware-in-the-loop

In the Mechatronics Laboratory at the Faculty of Mechanical Engineering (FME), University of Prishtina, we harnessed the versatile Quanser active suspension platform for seamless integration and exploration of both the LQR and MPC, Fig. 13. This encompassed a comprehensive process of system configuration, precise programming, rigorous testing, and meticulous fine-tuning. Our objective was to enable in-depth research and experimentation in advanced control strategies for suspension systems, to provide an enriched learning environment for our students.

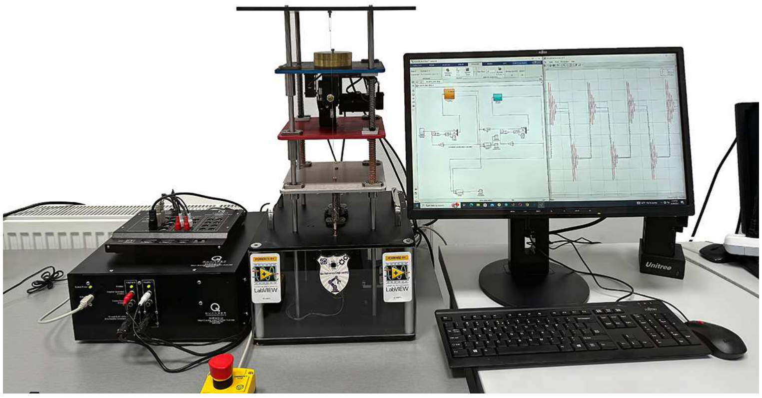


Fig. 11. Mechatronics laboratory: Physical representation of the ASS at the University of Prishtina

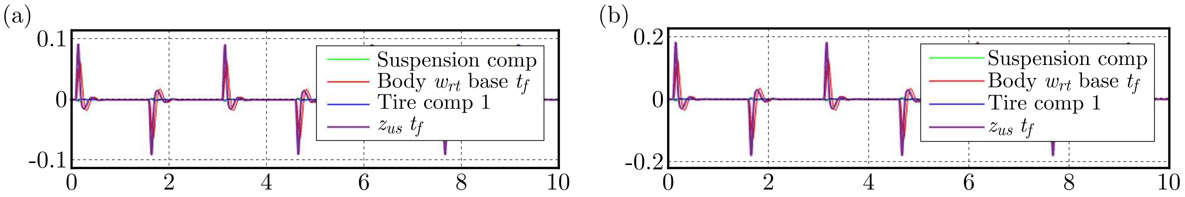


Fig. 12. (a) Closed-loop response for x_{sim} [mm/s]. (b) Measurements closed-loop response for x_{actual} [mm/s]

Figures 12a and 12b display the closed-loop response between the two variables: x_{sim} and x_{actual}, both measured in [mm/s]. These graphs visually portray the relationship between the simulated position x_{sim} and the actual position x_{actual} within a closed-loop control system. By observing the trends and correlations between these two positions, valuable insights can be derived regarding the accuracy and effectiveness of the control system. This graphical representation assists in the analysis and refinement of control strategies, offering a clear understanding of how closely the simulated position matches the actual position. This information is essential for evaluating the performance of control systems in scenarios where precise position tracking is critical and allows for adjustments that enhance system accuracy and reliability.

Figure 13a presents the closed-loop response for the actual position, referred to as x_{actual}, in the context of an active suspension system. This graph elucidates the behavior of the system position control under closed-loop conditions. The response illustrates how the ASS effectively manages and adjusts the actual position in response to various inputs or disturbances. This visual representation offers valuable insights into the system ability to maintain desired positions, enhance ride comfort, and mitigate the effects of external forces or road irregularities. Engineers and researchers can utilize these data to fine-tune the ASS, to ensure the optimal performance and responsiveness for a smoother and more controlled ride experience. Figure 13b depicts the closed-loop response of the plate position within an active suspension system. This graph provides a visual representation of how the ASS dynamically manages and adjusts the position of the plate in response to various inputs and external factors.

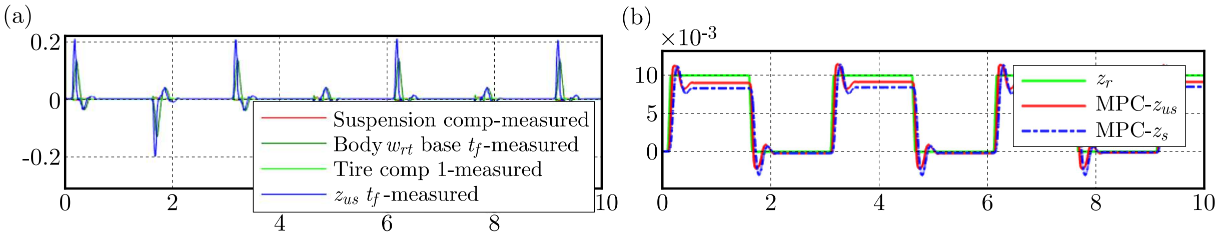


Fig. 13. (a) Closed-loop response for x_{actual} [m/s] with MPC. (b) Closed-loop response for plate position [m/s] using MPC

The response curve highlights the system ability to effectively control the plate position, thus contributing to improved vehicle stability and ride quality. This information is valuable for engineers and researchers working on active suspension systems, as it offers insights into the system performance characteristics and its capability to maintain desired plate positions even in the presence of disturbances or changing road conditions. By analyzing this response, developers can enhance the design and control strategies of the ASS, leading to better overall vehicle dynamics and passenger comfort.

Figure 14a illustrates the closed-loop response concerning the control force, denoted as F_c [N], and the resulting body acceleration in Fig. 14b, represented as z_r [m/s^2]. This response graphically portrays the relationship between the control force applied to guide the system in the rightward direction, and the consequent acceleration experienced by the body moving to the left. The graph provides insight into the dynamic behavior of the system under closed-loop control, highlighting how changes in the control force lead to corresponding variations in the body acceleration. This information is crucial for understanding and optimizing the performance of control systems in scenarios where precise manipulation of acceleration is paramount.

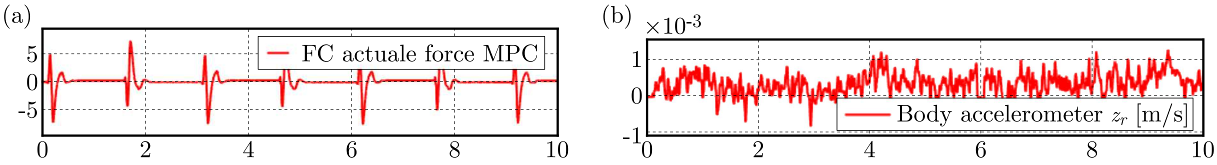


Fig. 14. Closed-loop response for: (a) control force F_c [N], (b) body acceleration z_r [m/s] using MPC

6. Conclusion

In this experimental study, we look at how the LQR and MPC control methods can be used in a quarter-car suspension model to optimize the system state and input signals performance. Using the Quanser active suspension setup, our experimental analysis provided crucial insights into these control strategies. The LQR controller demonstrated significant enhancements in the system performance under stable and well-understood dynamics. On the other hand, the MPC approach showed robustness in managing uncertain and time-varying parameters. The comparative study under varying load conditions, parameter variations, and disturbances highlighted individual strengths of the LQR and MPC. The LQR performs exceptionally well in certain scenarios with stable system dynamics, while the MPC effectively handles uncertainties and dynamic changes. The selection between the LQR and MPC for active suspension control should be based on specific system requirements and operational contexts. From the experimental results, we conclude that the MPC outperforms LQR in terms of time to achieve stability and performance. Ongoing research and development in these areas will surely contribute to the evolution of active suspension technology and shape the future where vehicles provide higher performance, comfort, and sustainability.

References

1. Abdellahi E., Mehdi D., M'saad M., 2000, Active suspension system by H_2 and H_∞: A weights choice procedure for vehicle performance, *IFAC Proceedings Volumes*, **33**, 14, 231-235

2. Active Suspension. Available online: https://www.quanser.com/products/active-suspension/ (accessed on 15 January 2024)

3. Ahmed M.M., Svaricek F., 2014, Preview optimal control of vehicle semi-active suspension based on partitioning of chassis acceleration and tire load spectra, *2014 European Control Conference (ECC)*, Strasbourg, France, 1669-1674

4. Apkarian J., Abdossalami A., 2013, *Laboratory Guide: Active Suspension Experiment for MATLAB/Simulink Users*, Quanser, Marhkam, Canada

5. Basturk H.I., 2016, A backstepping approach for an active suspension system, *2016 American Control Conference (ACC)*, Boston, USA, 7579-7584

6. Deshpande V.S., Bhaskara M., Phadke S.B., 2012, Sliding mode control of active suspension systems using a disturbance observer, *2012 12th International Workshop on Variable Structure Systems*, Mumbai, India, 70-75

7. Deshpande V.S., Shendge P.D., Phadke S.B., 2017, Nonlinear control for dual objective active suspension systems, *IEEE Transactions on Intelligent Transportation Systems*, **18**, 3, 656-665

8. Durmaz B.E., Kaçmaz B., Mutlu İ., Söylemez M.T., 2017, Implementation and comparison of LQR-MPC on active suspension system, *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, Bursa, Turkey, 828-835

9. Eris O., Ergenc A.F., Kurtulan S., 2015, A modified delayed resonator for active suspension systems of railway vehicles, *IFAC-PapersOnLine*, **48**, 12, 281-285

10. Fei J., Wang H., Fang Y., 2022, Novel neural network fractional-order sliding-mode control with application to active power filter, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **52**, 6, 3508-3518

11. Gandhi P., Adarsh S., Ramachandran K.I., 2017, Performance analysis of half car suspension model with 4 DOF using PID, LQR, FUZZY and ANFIS controllers, *Procedia Computer Science*, **115**, 2-13

12. Gupta S., Ginoya D., Shendge P.D., Phadke S.B., 2016, An inertial delay observer-based sliding mode control for active suspension systems, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **230**, 3, 352-370
13. Hwang T., Roh J., Park K., Hwang J., Lee K.H., *et al.*, 2006, Development of HILS systems for active brake control systems, *2006 SICE-ICASE International Joint Conference*, Busan, Korea (South), 4404-4408
14. Likaj R., 2005, Fuzzy Logic Control of Nonlinear Vehicle Suspension System, PhD Thesis, Prishtina, Kosovo
15. Likaj R., Bruqi M., Shala A., Bajrami X., 2016, Optimal design and analysis of quarter vehicle suspension system by using MATLAB, *Proceedings of the 27th DAAAM International Symposium*, B. Katalinic (Ed.), DAAAM International, Vienna, Austria, 0082-0090
16. Liu M., Li Y., Rong X., Zhang S., Yin Y., 2020, Semi-active suspension control based on deep reinforcement learning, *IEEE Access*, **8**, 9978-9986
17. Ovalle L., Ríos H., Ahmed H., 2022, Robust control for an active suspension system via continuous sliding-mode controllers, *Engineering Science and Technology, an International Journal*, **28**, 101026
18. Pedro J.O., Dangor M., Dahunsi O.A., Ali M.M., 2014, Intelligent feedback linearization control of nonlinear electrohydraulic suspension systems using particle swarm optimization, *Applied Soft Computing*, **24**, 50-62
19. Pusadkar U.S., Chaudhari S.D., Shendge P.D., Phadke S.B., 2019, Linear disturbance observer based sliding mode control for active suspension systems with non-ideal actuator, *Journal of Sound and Vibration*, **442**, 428-444
20. Qin W., Ge P., Liu F., Long S., 2021, Adaptive robust control for active suspension systems: targeting nonholonomic reference trajectory and large mismatched uncertainty, *Nonlinear Dynamics*, **104**, 4, 3861-3880
21. Reddipogu J.S.D., Elumalai V.K., 2020, Hardware in the loop testing of adaptive inertia weight PSO-tuned LQR applied to vehicle suspension control, *Journal of Control Science and Engineering*, **2020**, 1, 8873995
22. Saha S., Amrr S.M., 2020, Design of slip-based traction control system for EV and validation using co-simulation between Adams and Matlab/Simulink, *Simulation*, **96**, 6, 537-549
23. Sahin H., Akalin O., 2020, Articulated vehicle lateral stability management via active rear-wheel steering of tractor using fuzzy logic and model predictive control, *SAE International Journal of Commercial Vehicles*, **13**, 2, 115-128
24. Sereni B., Assunção E., Carvalho Minhoto Teixeira M., 2020, New gain-scheduled static output feedback controller design strategy for stability and transient performance of LPV systems, *IET Control Theory and Applications*, **14**, 5, 717-725
25. Wang G., Zhou Z., 2019, Design and implementation of H_∞ miscellaneous information feedback control for vehicle suspension system, *Shock and Vibration*, **2019**, 1, 3736402
26. Zhang M., Jing X., 2021, Switching logic-based saturated tracking control for active suspension systems based on disturbance observer and bioinspired X-dynamics, *Mechanical Systems and Signal Processing*, **155**, 107611

Manuscript received January 31, 2024; accepted for publication November 21, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 91-101, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195840

IMPROVED SHERWOOD-FROST PHENOMENOLOGICAL CONSTITUTIVE MODEL SUITABLE FOR POLYMETHACRYLIMIDE FOAM UNDER UNIAXIAL COMPRESSION

Qianying Cen, Ling Liu
School of Aerospace Engineering and Applied Mechanics, Tongji University, Shanghai, China
corresponding author Ling Liu, e-mail: lingliu@tongji.edu.cn

To investigate the stress-strain response of a polymethacrylimide (PMI) foam under uniaxial compression, an improved phenomenological constitutive model based on the Sherwood-Frost model is fitted from the compressive stress-strain curves of the PMI foam. Firstly, new function terms are proposed to describe the effects of temperature, density and strain-rate. Then, the model parameters are determined. Finally, compression experiments and numerical simulation are conducted on PMI foams at different temperatures, densities and strain-rates to verify the modified model. The results show that it can successfully predict the compressive mechanical response of the PMI foam if the effects of density, temperature and strain-rate are considered.

Keywords: polymethacrylimide (PMI) foam, constitutive model, modified Sherwood-Frost model

1. Introduction

Inspired by natural structures, such as hornbill beak, bird wing trabecular bone and balsa wood, lightweight sandwich structures with foams or honeycomb cores have emerged to meet the urgent demand for structural lightness and multifunctionality in engineering fields, which is of great significance for structural lightweight (Hedayati and Sadighi, 2018). Among many foams that can be selected as cores and simultaneously have the similar density, the polymethacrylimide (PMI) foam shows advantages in mechanical properties (such as high specific strength/modulus) and microstructure (rigid closed cell and isotropy) (Palamidi, 2010; Poxon, 2012). Meanwhile, compared with honeycomb, the PMI foam also exhibits excellent secondary processing performance, high interface bonding strength, and the ability to bear side pressure, which enables the PMI foam to support complex profiles and ensures the molding quality of sandwich structures (Maier *et al.*, 2006). In addition, the closed-cell structure of the PMI foam can avoid moisture absorption, which can reduce maintenance expenditure costs. Therefore, the PMI foam is widely used in sandwich structures of helicopter blades, aircraft pressure bulkheads, and rocket vehicles (Seibert, 2006).

In fact, the molding and service of PMI sandwich structures are usually affected by complex working conditions such as temperature and pressure. So, it is particularly important to understand mechanical properties and load-bearing behavior. Li *et al.* (2000) obtained the uniaxial tensile, compressive, pure shear, hydrostatic compressive, and shear-compression properties of PMI foams by experimental tests. Palamidi (2010) investigated the dynamic properties of PMI foams (Rohacell-51WF and 110WF) by using the split Hopkinson pressure bar (SHPB) and direct-impact (DI). Poxon (2012) conducted a comprehensive experimental study on PMI foams with different densities, temperatures and strain rates. Chai *et al.* (2020) performed quasi-static compression of 3D closed-cell PMI foams by in-situ X-ray micro computed tomography (μ-CT).

Huo *et al.* (2022) carried out systematic experimental characterization for elastic, plastic and fracture properties of PMI foams under quasi-biaxial stress loading conditions, and the yield surfaces of PMI foams were calibrated and analyzed. The afore-mentioned studies also show that the compression stress-strain curve of the PMI foam can be divided into three sections, linear elastic section, plateau section, and densification section, and the elastic modulus and plateau section stress increase with the rise of the strain rate and density, but the trend of temperature effect is opposite.

Despite the extensive work that has been done on PMI foams, there are relatively few studies on the constitutive model, because the complexity of the microstructure leads to numerous and complex control parameters of the theoretical model, which is difficult to implement in engineering practice (Li *et al.*, 2022). However, the PMI foam can be assumed as a macroscopic continuum when the size of foam is larger than the typical length of 6 cell bodies (Poxon, 2012), and its stress-strain relationship (including 3 sections) can be easily obtained by a phenomenological method, i.e., achieving a simple parameter identification by fitting experimental data (Sherwood and Frost, 1992). Rusch (1969) established the compression stress-strain function of the polyurethane (PU) foam, and proposed for the first time to use a dimensionless strain function to describe the shape of the compression curve, which was changed into a strain polynomial function by Meinecke and Schwaber (1970) and named the "shape function". Then, Sherwood and Frost (1992) developed a model to decouple the effects of density, temperature, and strain rate on stress, in order to predict the impact response of the PU foam under uniaxial compression. However, previous studies have mainly focused on other polymer foams, such as PU foams, while relatively little research has been done on PMI foams. Meanwhile, there are few studies on the constitutive model of the PMI foam, and often only a single variable was considered, such as density or temperature. Until now, a complete constitutive model that can describe the compressive mechanical behavior of the PMI foam simultaneously considering density, temperature and strain rate has not been publicly reported.

Therefore, the present work aims to investigate the compressive mechanical behavior of the PMI foam under different densities, temperatures and strain-rates. Based on the traditional Sherwood-Frost constitutive model, combined with the experimental data of the PMI foam, an improved model for the PMI foam is proposed, which contains new function terms of temperature, density, and strain rate, as well as the parameters of the new model are determined by fitting the compressive stress-stain curves of the PMI foam. Moreover, the improved model is compared with the model from the previous studies. Finally, compressive experiments and numerical simulation are conducted on PMI foams at different temperatures, densities and strain-rates to verify the improved model.

2. Sherwood-Frost model with modified function terms

In the Sherwood-Frost model, a decoupling function is used to describe the effects of temperature, density and strain rate on the uniaxial compressive stress-strain response of the PU foam (Sherwood and Frost, 1992). In this paper, the general form of this model is employed to predict the mechanical response of the PMI foam under uniaxial compression, as shown in

$$\sigma = H(T)G(\rho)M(\dot{\varepsilon})f(\varepsilon) \tag{2.1}$$

where σ [MPa], T [°C], ρ [kg/m^3], ε and $\dot{\varepsilon}$ [s^{-1}] represent stress, temperature, foam density, strain and strain rate, respectively. Meanwhile, in (Sherwood and Frost, 1992), $H(T)$, $G(\rho)$, $M(\dot{\varepsilon})$ and $f(\varepsilon)$, respectively, denote the bilinear temperature softening term, bilinear density term, exponential strain-rate hardening term, and shape function term which is expressed as a polynomial function

$$f(\varepsilon)=\sum_{i=0}^{n} A_i\varepsilon^i \tag{2.2}$$

where n and A_i represent the order of $f(\varepsilon)$ and the corresponding coefficients of each order, respectively.

On the basis of applying Eq. (2.2) to fit the compressive stress-strain curves of the PMI foam under the reference condition $(\rho_0, T_0, \dot{\varepsilon}_0)$ (Poxon, 2012), the bilinear temperature term $H(T)$ can be further improved into a new negative exponential function of strain, as seen in

$$H(T)=1+\frac{1-T}{T_0}\Big[c+b\exp\Big(-\frac{\varepsilon}{a}\Big)\Big] \tag{2.3}$$

where a, b and c are all material parameters, which can be found by approximating the experimental stress-strain data for several temperatures.

Similarly, the density function $G(\rho)$ is proposed to reflect the effect of foam density, which is initially a linear function of relative density $\rho^*=\rho/\rho_0$, as shown in Eq. (2.4) (Sherwood and Frost, 1992)

$$G(\rho)=B(\rho^*-1)+1 \tag{2.4}$$

This model only has one parameter B.

Subsequently, considering the influence of foam density on stress is also related to strain, $G(\rho)$ can also be expressed by a modified Freundlich model (Rahimidehgolan and Altenhof, 2023)

$$G(\rho,\varepsilon)=C(\rho)-D(\rho)\varepsilon^{F(\rho)} \tag{2.5}$$

This model has three parameters. The parameters C, D and F are all density dependent.

Because Eqs. (2.4) and (2.5) are suboptimal for the PMI foam, a new density function shown in Eq. (2.6) is constructed by fitting the compression stress-strain curves of the PMI foam in the reference (Poxon, 2012

$$G(\rho^*,\varepsilon)=1+(\rho^*-1)\Big[P(\rho^*)+Q(\rho^*)\exp\Big(\frac{\varepsilon}{R(\rho^*)}\Big)\Big] \tag{2.6}$$

which is an exponential function of relative density, where P, Q and R are all material parameters and related to relative density.

Besides, the strain-rate function $M(\dot{\varepsilon})$ mainly considers the influence of strain-rate, which can be calculated by using the Seeger model (Hu *et al.*, 1998), as shown as

$$M(\dot{\varepsilon})=1+K\lg\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} \tag{2.7}$$

The parameters K in this model can be fitted by experimental stress-strain data for several strain rates.

Because the influence of strain-rate on stress is also related to strain, thus the strain is coupled with the strain-rate, and so K in Eq. (2.7) is no longer a constant, which can be rewritten as (Jeong *et al.*, 2012)

$$M(\varepsilon,\dot{\varepsilon})=1+(p+q\varepsilon)\ln\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} \tag{2.8}$$

There are two parameters, p and q, in this model, which can be determined by data fitting.

In fact, Eq. (2.8) is also not well applicable to the PMI foam, so an improved strain-rate function is constructed by fitting the compression stress-strain curves in references (Poxon, 2012) and (Flores-Johnson *et al.*, 2008), as seen as

$$M(\varepsilon,\dot{\varepsilon})=1+(i+j\varepsilon^k)\ln\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} \tag{2.9}$$

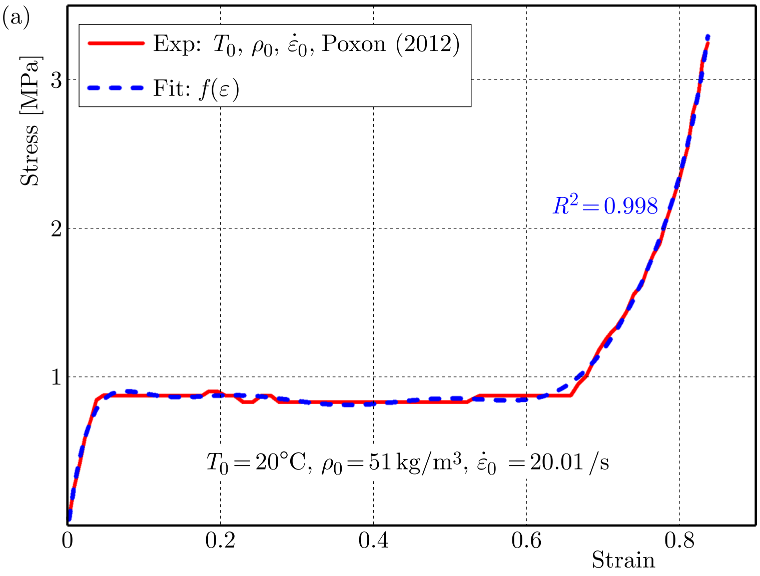
where i, j and k are all strain rate dependent parameters.

3. Determination of constitutive model parameters and verification

3.1. Constitutive model parameters

3.1.1. Determination of shape function parameters

Here, $\rho_0 = 51\,\mathrm{kg/m^3}$, $T_0 = 20^\circ\mathrm{C}$ and $\dot{\varepsilon}_0 = 0.01\,\mathrm{s}^{-1}$ are chosen as the reference condition. As shown in Fig. 1a, a polynomial fitting is used for the compressive stress-strain curve of the PMI foam (Rohacell 51WF) obtained from reference (Poxon, 2012). In order to fit more accurately, the least square method is used to fit a 9-order polynomial, i.e., the shape function $f(\varepsilon)$ of Eq. (2.2) ($n = 9$). The fitted curve and model parameters are respectively given in Figs. 1a and 1b, with a R-Square (R^2) of 0.998. Also, the fitting curve in Fig. 1a indicates that by selecting a 9th-order polynomial as the shape function $f(\varepsilon)$ a very good fitting result can be obtained.



(b)

A_0	A_1
-0.046	41.18
A_2	A_3
-689.91	5917.08
A_4	A_5
-29184.90	87122.53
A_6	A_7
-159554.92	174922.70
A_8	A_9
-105231.16	267707.07

Fig. 1. (a) Fitting curve for the shape function $f(\varepsilon)$ and (b) the obtained model parameters of $f(\varepsilon)$

3.1.2. Determination of temperature function parameters

The compressive stress-strain response of the PMI foam is sensitive to temperature. Keeping the density $\rho_0 = 51\,\mathrm{kg/m^3}$ and strain-rate $\dot{\varepsilon}_0 = 0.01\,\mathrm{s}^{-1}$ unchanged, the compressive stress-strain curves (Rohacell 51WF) at $-70^\circ\mathrm{C}$, $-40^\circ\mathrm{C}$ from (Poxon, 2012) and $40^\circ\mathrm{C}$ (testing in this study) were selected to fit the model parameters of the function $H(T)$. As seen in Fig. 2a, the PMI foam specimens ($100\,\mathrm{mm}\times100\,\mathrm{mm}\times50\,\mathrm{mm}$) are tested on a universal testing machine with an environmental chamber (MTS-5T, MTS, USA) according to ASTM D1621 standard with the same density $51\,\mathrm{kg/m^3}$ and loading rate $0.01\,\mathrm{s}^{-1}$ as these in (Poxon, 2012).

Figure 2b first presents the compressive stress-strain curves of the PMI foam at $-70^\circ\mathrm{C}$, $-40^\circ\mathrm{C}$ and $40^\circ\mathrm{C}$ obtained from the experiment (solid lines), respectively. It shows that the PMI foam will soften (decrease in stress) with an increase of temperature, and so the elastic modulus and plateau stress will decrease accordingly. By fitting the experimental curves using function terms of $f(\varepsilon)H(T)$, parameters of $H(T)$ are obtained, that is $a = 0.016$, $b = 0.26 + 0.0016T$, and $c = 0.042$. As shown in Fig. 2b, it indicates that the fitting curves at $-70^\circ\mathrm{C}$, $-40^\circ\mathrm{C}$ and $40^\circ\mathrm{C}$ are basically consistent with the experimental curves.

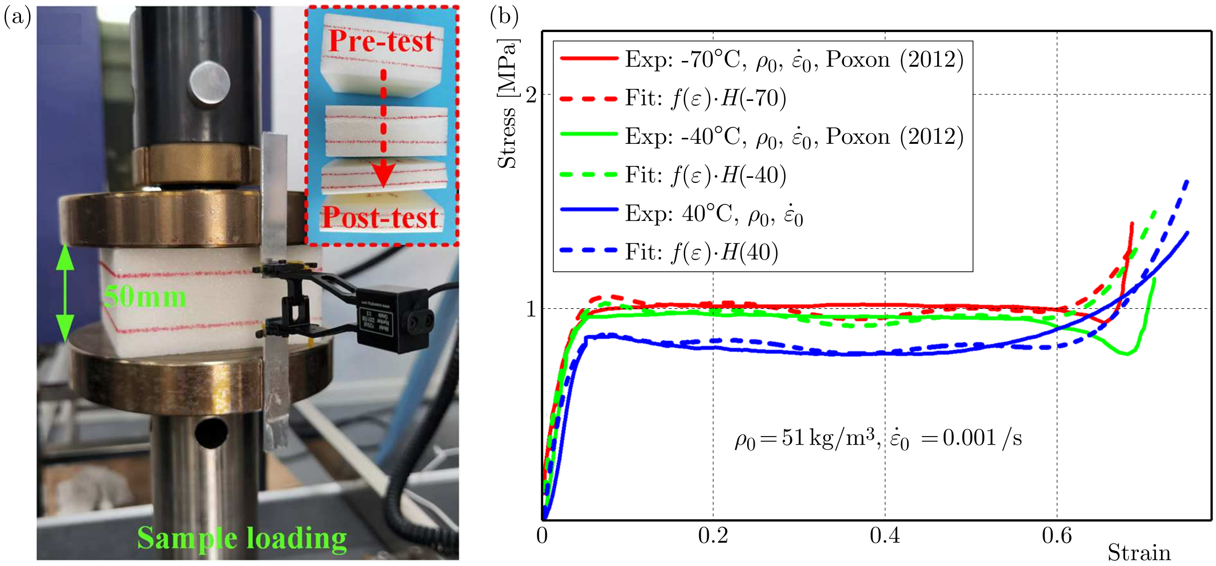


Fig. 2. (a) Compression test of the PMI foam, (b) experimental curves from testing and fitting curves with the temperature function $H(T)$

3.1.3. Determination of density function parameters

Similarly, the density of the PMI foam will also significantly affect its compressive mechanical behavior. Keeping the temperature $T_0 = 20°\mathrm{C}$ and strain rate $\dot{\varepsilon}_0 = 0.01\,\mathrm{s}^{-1}$ unchanged and only changing the density $\rho = 71$, 110 and $200\,\mathrm{kg/m^3}$, the experimental stress-strain curves (Rohacell 71WF, 110WF, 200WF) from (Poxon, 2012) are given in Figs. 3a-c, which (solid lines) exhibit that the higher the density of foam, the greater the yield stress, the shorter the yield plateau section, and the steeper the dense section. The linear function of relative density (G_1, Eq. (2.4)) and the modified Freundlich model (G_2, Eq. (2.5)), as well as the newly constructed density function (G_3, Eq. (2.6)) are used to fit the compressive stress-strain curves with different densities, and the fitted results are respectively given in Figs. 3a-c (dashed lines). Figure 3a indicates that the fitting results of the function G_1 to the stress-strain curve of a low-density PMI foam $\rho = 71$ and $110\,\mathrm{kg/m^3}$ is relatively good, but is seriously deviated for a high-density PMI foam (e.g., $200\,\mathrm{kg/m^3}$). That is, the linear function G_1 is only applicable to the low-density PMI foam. While, as shown in Figs. 3b and 3c, the G_2 and G_3 functions show highly consistent fitting results for these three densities (71, 110 and $200\,\mathrm{kg/m^3}$), respectively.

The fitting errors, $(\sigma_{fit} - \sigma_{exp})/\sigma_{exp}$, from these three density functions (G_1, G_2, G_3), are presented in Fig. 3d. It can be seen that the fitted results of the G_2 and G_3 functions show smaller errors than the G_1 function when ρ is 71 or $110\,\mathrm{kg/m^3}$ (the fitting error of the function G_1 at $\rho = 200\,\mathrm{kg/m^3}$ is too large, so the error curve is absent in Fig. 3d). In addition, Fig. 3d also indicates that the error curves obtained from the G_2 and G_3 functions basically coincide with each other when ρ is 71 or $110\,\mathrm{kg/m^3}$. Moreover, when ρ is $200\,\mathrm{kg/m^3}$, the error from the function G_3 is slightly smaller that of the function G_2. Therefore, for the function G_3 (Eq. (2.6)), the relationships between the model parameters and the relative density of the PMI foam are then fitted, as shown as

$$\begin{aligned}
&P(\rho^*) = 2.68 - 1.37\rho^* + 0.37\rho^{*2}\\
&Q(\rho^*) = 0.0094 - 0.011\rho^* + 0.0031\rho^{*2}\\
&R(\rho^*) = -0.0051 + 0.032\rho^*
\end{aligned} \tag{3.1}$$

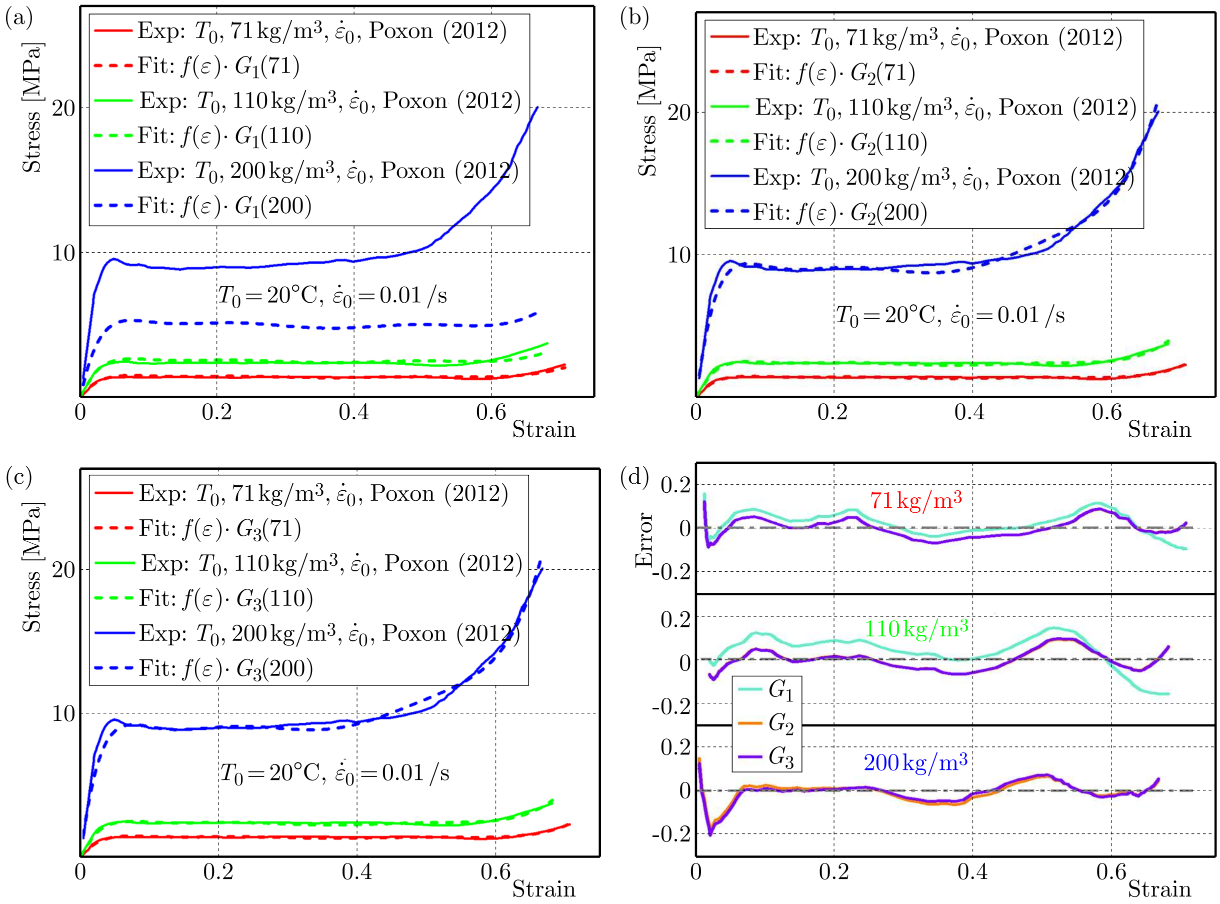


Fig. 3. Comparison between experimental curves and density function fitting curves of (a) G_1, (b) G_2 and (c) G_3, (d) comparison of fitting errors for different density functions

3.1.4. Determination of strain-rate function parameters

Subsequently, the influence of the strain rate is considered, since the compressive mechanical behavior of the PMI foam is also sensitive to it. Keeping $\rho_0 = 51\,\mathrm{kg/m^3}$ and $T_0 = 20°\mathrm{C}$ unchanged, and just changing the strain-rate ($\dot{\varepsilon} = 8.3\cdot 10^{-4}\,\mathrm{s^{-1}}$, 9.1 and $100\,\mathrm{s^{-1}}$), the strain-rate functions, M_1 Eq. (2.7), M_2 Eq. (2.8) and M_3 Eq. (2.9) are sequentially selected to fit the experimental curves (Rohacell 51WF) obtained from (Flores-Johnson *et al.*, 2008) and (Poxon, 2012). Figures 4a-c present a comparison between the fitted curves and the experimental curves, respectively, which show that the higher the strain-rate, the greater the yield stress and the shorter the yield plateau section. Among them, Fig. 4a indicates that the Seeger model M_1 fits well with the curve when $\dot{\varepsilon} = 8.3\cdot 10^{-4}\,\mathrm{s^{-1}}$ (quasi-static loading). However, when $\dot{\varepsilon} = 9.1$ and $100\,\mathrm{s^{-1}}$, the plateau section fitted by the function M_1 is higher, while the densification section is relatively lower. That is, the function M_1 is not suitable for high strain-rate situations. And as can be seen in Fig. 4b, M_2 also performs well under the quasi-static condition, and has improvements in the first half of the plateau section and the densification section when compared to M_1 function, but not sufficient. On the contrary, the function M_3 can fit the experimental stress-strain curves well for these three strain rates, as can be seen in Fig. 4c.

Especially, compared to M_1 or M_2, M_3 shows superior fitting accuracy in the densification section, as shown in Fig. 4d. Therefore, the relationships between the model parameters of the function M_3, Eq. (2.9), and the strain-rate are then obtained, as seen in Eqs. (3.2)

$$
\begin{aligned}
i &= 0.037 - 0.0024 \ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} + 6.11 \cdot 10^{-5} \Big(\ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\Big)^2 \\
j &= 217.89 - 40.34 \ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} + 1.82 \Big(\ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\Big)^2 \\
k &= 55.61 - 8.22 \ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} + 0.32 \Big(\ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\Big)^2
\end{aligned}
\tag{3.2}
$$

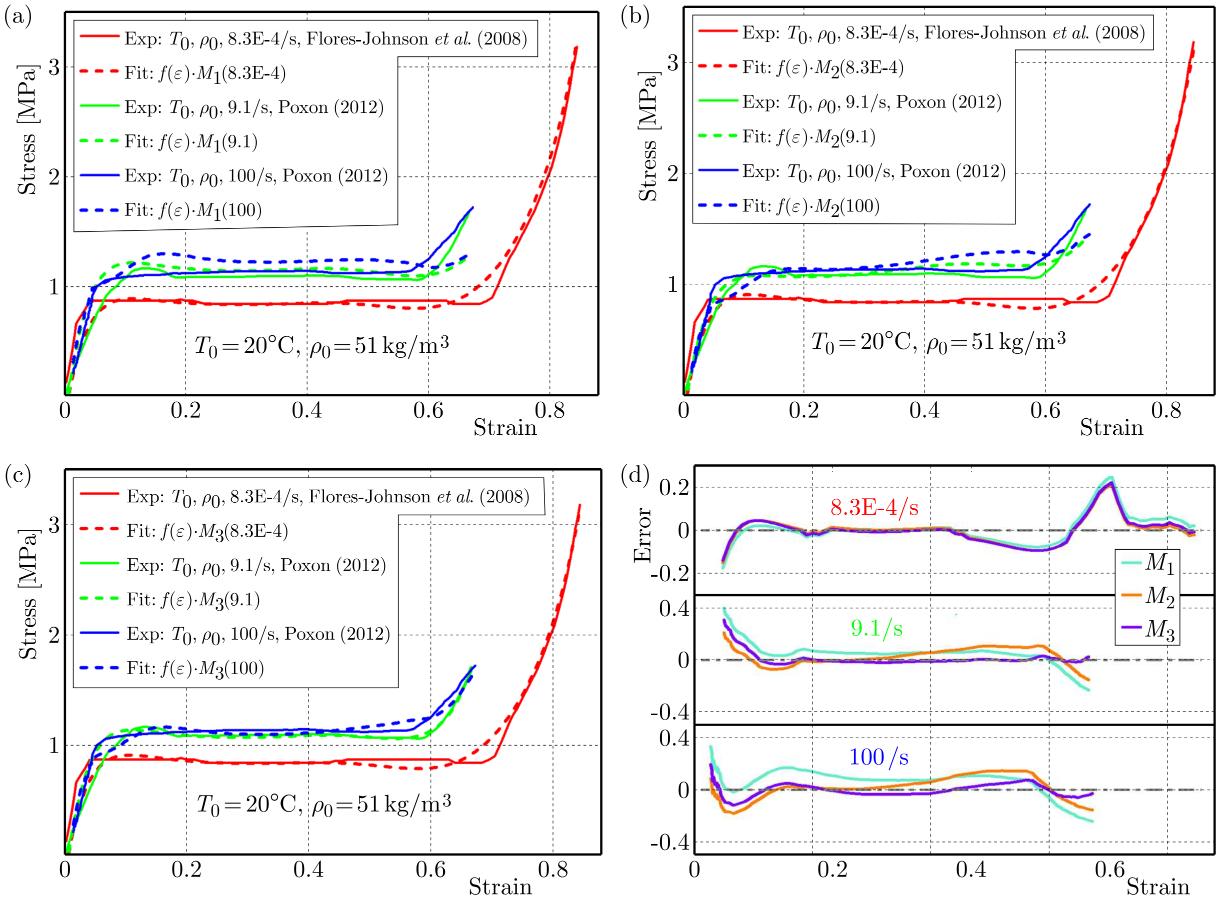


Fig. 4. Comparison between experimental curves and strain-rate function fitting curves of (a) M_1, (b) M_2 and (c) M_3, (d) comparison of fitting errors for different strain-rate functions

It can be seen from Figs. 2 and 4 that the effects of strain-rate and temperature on the compression stress-strain curve of the PMI foam are opposite, which can be simply linked through the time-temperature equivalence principle (Poxon, 2012). That is, the mechanical behavior of the PMI foam at a certain strain-rate can be mapped to the response at a certain temperature. For example, for the PMI foam, when $\rho = 51\,\text{kg/m}^3$, the mechanical behavior of $T = 20°\text{C}$ and $\dot{\varepsilon} = 100\,\text{s}^{-1}$ can be mapped to each other, and the same is true for the mechanical behavior of $T = -70°\text{C}$ and $\dot{\varepsilon} = 0.001\,\text{s}^{-1}$. Also, when $\rho = 200\,\text{kg/m}^3$, the state of $T = 20°\text{C}$ and $\dot{\varepsilon} = 1000\,\text{s}^{-1}$, and the state of $T = -70°\text{C}$ and $\dot{\varepsilon} = 0.001\,\text{s}^{-1}$ can be mapped to each other too, as can be referenced in the study carried by Poxon (2012).

3.2. Verification by experimental curves

The improved phenomenological constitutive model of the PMI foam based on the Sherwood-Frost model is obtained in Section 3.1, in which the effects of temperature, density and strain-rate on the compression stress-strain curves can be considered simultaneously. The stress-strain curves from case 1 ($\rho = 51\,\text{kg/m}^3$, $T = 80°\text{C}$, $\dot{\varepsilon} = 0.01\,\text{s}^{-1}$), case 2 ($\rho = 110\,\text{kg/m}^3$, $T = 20°\text{C}$,

$\dot{\varepsilon} = 0.01\,\mathrm{s}^{-1}$) and case 3 ($\rho = 51\,\mathrm{kg/m}^3$, $T = 20°\mathrm{C}$, $\dot{\varepsilon} = 0.1\,\mathrm{s}^{-1}$) are obtained by compressive testing (as seen in Fig. 2a), which are used to verify the $H(T)$, G_3 and M_3 functions in the improved model. As shown in Fig. 5a-c, the experimental curves (solid lines) and the verified curves (dashed lines) are quite consistent, and the error margins are -13.39%-$+4.76\%$ (case 1), -9.94%-$+9.52\%$ (case 2), and -3.69%-$+5.41\%$ (case 3), respectively, which indicates that the modified model can effectively predict the compressive stress-strain curves of the PMI foam. Among them, in case 1, due to the temperature effect ($T = 80°\mathrm{C}$), the transition zone between the plateau and densification section is becoming less apparent, resulting in slightly higher errors compared to the other two. Moreover, compressive testing under $\rho = 51\,\mathrm{kg/m}^3$, $T = 120°\mathrm{C}$ and $\dot{\varepsilon} = 0.00167\,\mathrm{s}^{-1}$ (case 4) is also conducted to characterize the fitting effect of the model on the coupling effect of strain-rate and temperature, as shown in Fig. 5d, indicating a relatively small fitting error. However, under low strain-rates, the fitting results of the transition section from the elastic stage to the plateau section (the curve knee) are poor and further improvement is needed. At the same time, when the temperature reaches 120°C, the transition zone between the plateau and densification section is no longer obvious (Xing *et al.*, 2024), so the fitting result of this zone is also poor.

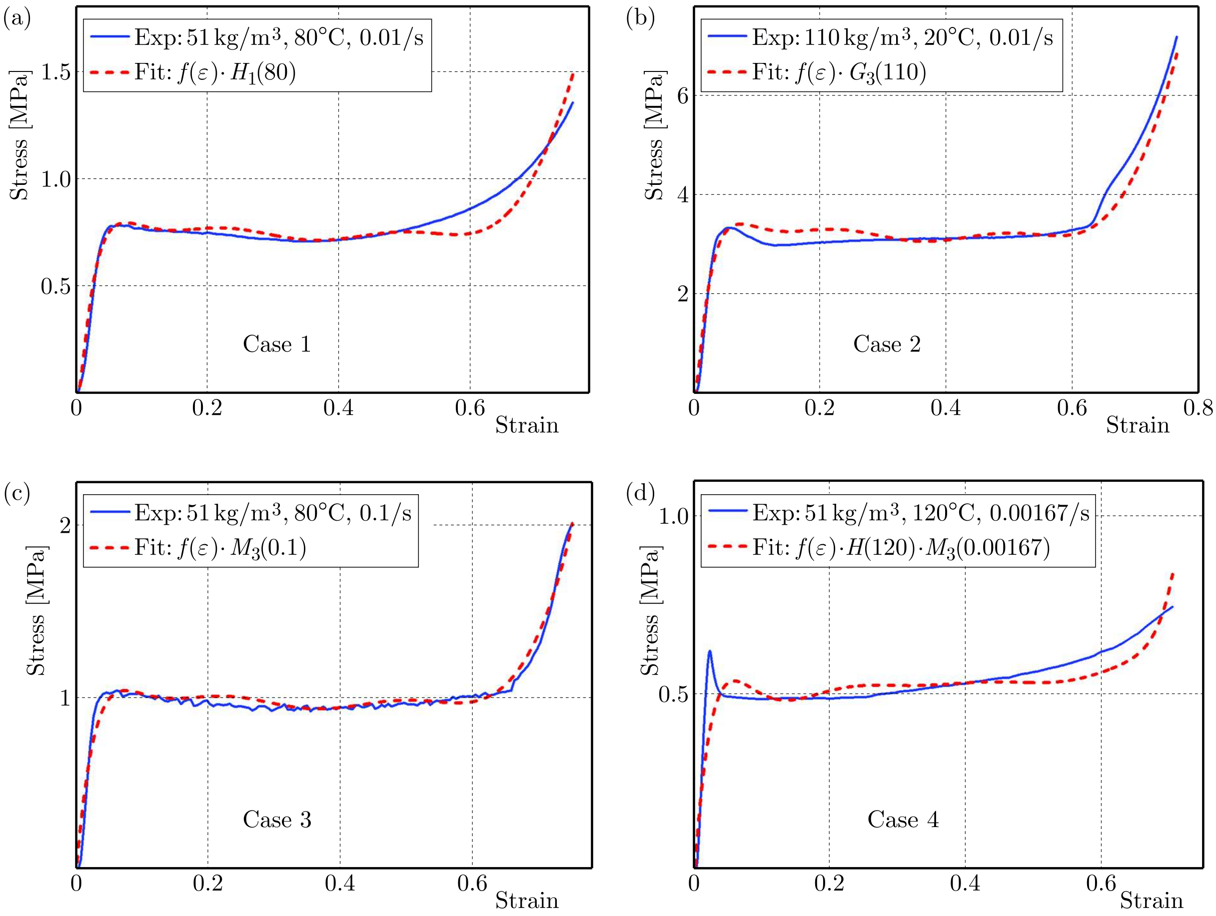


Fig. 5. Experiment and verification curves of (a) $H(T)$, (b) G_3, (c) M_3 and (d) $H(T)$ coupled with M_3

3.3. Finite element (FE) calculation verification

Abaqus software is used to model the uniaxial compressive mechanical behavior of the PMI foam (Rohacell 51WF). Figure 6a displays the FE model for compression specimens using element C3D8R, where, a simple model consists of a cuboid foam ($100\,\mathrm{mm}\times 100\,\mathrm{mm}\times 50\,\mathrm{mm}$, same size as specimens in Fig. 2a) and two rigid compression surfaces with boundary and loading conditions. The PMI foam properties are input through an elastic-plastic model in the property module. In the elastic term, the foam modulus of temperature effect $T = 80^\circ\mathrm{C}$ and strain rate effect $\dot{\varepsilon} = 0.1\,\mathrm{s}^{-1}$ are 44.68 MPa and 69.09 MPa, respectively, and the Possion ratio is 0.33. In the plastic term, 0.013 and 0.011 are respectively selected as yield strains for the temperature effect and strain rate effect, and the plastic strain and stress data generated by the improved model (Eqs. (2.2) and (2.3) with parameters are obtained in Section 3.1 for the temperature effect, as well as Eqs. (2.2), (2.9) and (3.2) for the strain rate effect) are used as the foam plastic property. The reference points coupled with the rigid surfaces are established for the purpose of applying loads and extracting the constraint reaction force, where the bottom one is constrained in all directions, while the upper one only unconstrained in the z-direction, and the foam compression process is simulated by moving it down. Moreover, universal contact (hard contact, penalty algorithm) is used between the surfaces and the foam. By conducting mesh sensitivity analysis on the model, an ideal mesh size of 2 mm is obtained.

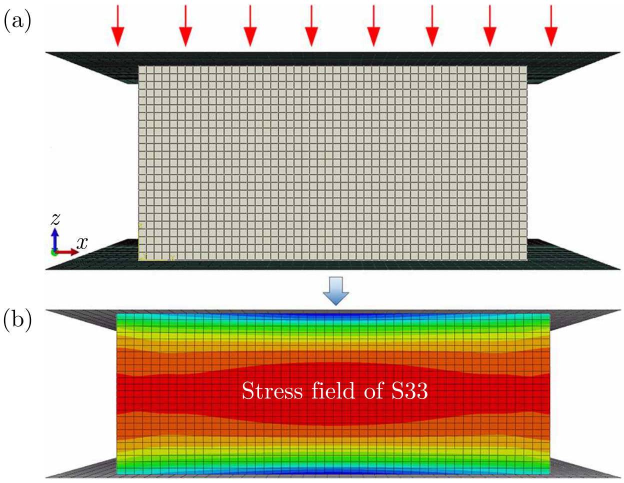


Fig. 6. Compressive simulation of the PMI foam (a) model with boundary and loading conditions, and (b) stress field of z-direction during the compression

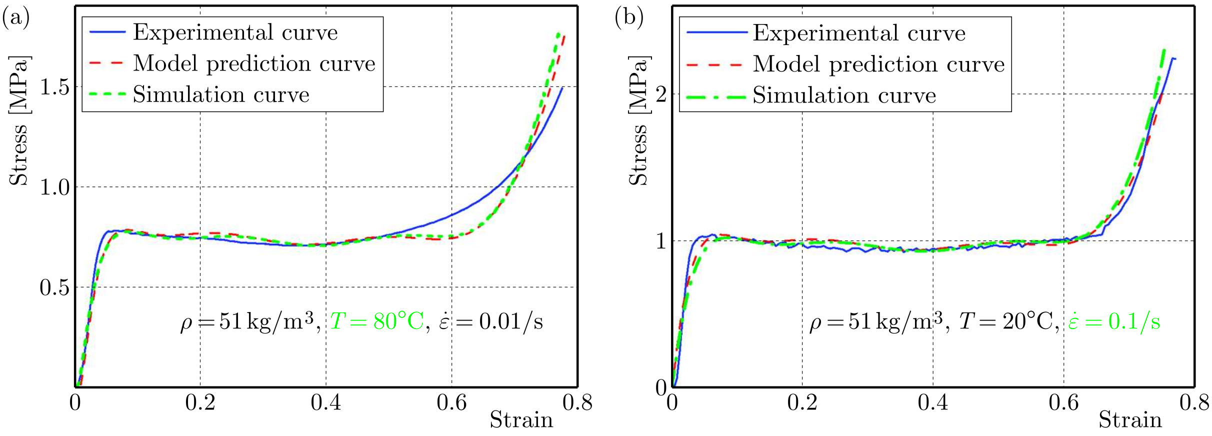


Fig. 7. Comparison curves of experimental, model prediction and simulation (a) temperature effect, (b) stress-rate effect

The simulation results show that the stress concentration first occurs in the center of the foam, as shown in Fig. 6b, which is consistent with the experimental results as given in Fig. 2a. The calculation results of temperature and strain-rate effects obtained by converting the force and displacement of the reference point are shown in Figs. 7a and 7b, respectively, which indicate a good agreement with the experimental results. Therefore, experimental and simulation verifications prove that the improved model proposed in the present work can well predict the mechanical response of a PMI foam during uniaxial compression when the effects of density, temperature and strain rate are considered separately or simultaneously.

4. Conclusions

The constitutive model of the foam plays a decisive role in the design and failure analysis of foam sandwich structures. There are few studies on the constitutive model of the PMI foam, and often only a single variable is considered, such as density or temperature. In this work, a constitutive model for the PMI foam has been proposed to describe the compressive mechanical behavior simultaneously considering density, temperature and strain rate. Key conclusions drawn from this study are summarized here:

- Based on the traditional Sherwood-Frost constitutive model and combined with the experimental data of the PMI foam, a new improved constitutive model containing new function terms of temperature, density, and strain rate for the PMI foam is proposed.
- The parameters of the temperature term, density term and strain-rate term in the new model are successively obtained by fitting the compressive stress-strain curves under different temperatures, densities and strain-rates. The new model shows a higher accuracy when comparing with the terms proposed in the previous studies.
- The new model is well verified by the compressive experimental data (error ranges, -13.39%-$+4.76\%$ for case 1, -9.94%-$+9.52\%$ for case 2, and -3.69%-$+5.41\%$ for case 3) and numerical simulation curves at different temperatures, densities and strain-rates, which can successfully predict the compressive mechanical response of the PMI foam if the effects of temperature, density and strain-rate are concurrently considered, and the error is within an acceptable range.

Acknowledgements

The authors would like to thank for the financial support from the National Natural Science Foundation of China (No. 12272268).

References

1. Chai H.W., Xie Z.L., Xiao X.H., Xie H.L., Huang J.Y., Luo S.N., 2020, Microstructural characterization and constitutive modeling of deformation of closed-cell foams based on in situ x-ray tomography, *International Journal of Plasticity*, **131**, 102730
2. Flores-Johnson E.A., Li Q.M., Mines R.A.W., 2008, Degradation of elastic modulus of progressively crushable foams in uniaxial compression, *Journal of Cellular Plastics*, **44**, 5, 415-434
3. Hedayati R., Sadighi M., 2018, Low-velocity impact behaviour of open-cell foams, *Journal of Theoretical and Applied Mechanics*, **56**, 4, 939-949
4. Hu S., Liu J., Wang W., 1998, Study of the constitutive relationship of rigid polyurethane foam (in Chinese), *Chinese Journal of Theoretical and Applied Mechanics*, **30**, 2, 151-156
5. Huo X., Jiang Z., Luo Q., Li Q., Sun G., 2022, Mechanical characterization and numerical modeling on the yield and fracture behaviors of polymethacrylimide (PMI) foam materials, *International Journal of Mechanical Sciences*, **218**, 107033
6. Jeong K.Y., Cheon S.S., Munshi M.B., 2012, A constitutive model for polyurethane foam with strain rate sensitivity, *Journal of Mechanical Science and Technology*, **26**, 7, 2033-2038
7. Li Q.M., Mines R.A.W., Birch R.S., 2000, The crush behaviour of Rohacell-51WF structural foam, *International Journal of Solids and Structures*, **37**, 43, 6321-6341

8. Li X.Q., Tao J.L., Landauer A.K., Franck C., Henann D.L., 2022, Large-deformation constitutive modeling of viscoelastic foams: Application to a closed-cell foam material, *Journal of the Mechanics and Physics of Solids*, **161**, 104807

9. Maier L., Hu P., Seibert H.F., 2006, PMI foam cored sandwich components produced by means of different manufacturing methods (in Chinese), *Journal of Materials Engineering*, 5, 37-40, 45

10. Meinecke E.A., Schwaber D.M., 1970, Energy absorption in polymeric foams. I. Prediction of impact behavior from Instron data for foams with rate-independent modulus, *Journal of Applied Polymer Science*, **14**, 9, 2239-2248

11. Palamidi E., 2010, Hopkinson bar testing of cellular materials, Ph.D. Thesis, University of Manchester, Manchester

12. Poxon S., 2012, The mechanical response of low to high density Rohacell foams, Ph.D. Thesis, University of Oxford, Oxford

13. Rahimidehgolan F., Altenhof W., 2023, Compressive behavior and deformation mechanisms of rigid polymeric foams: A review, *Composites Part B: Engineering*, **253**, 110513

14. Rusch K.C., 1969, Load-compression behavior of flexible foams, *Journal of Applied Polymer Science*, **13**, 11, 2297-2311

15. Seibert H.F., 2006, Applications for PMI foams in aerospace sandwich structures, *Reinforced Plastics*, **50**, 1, 44-48

16. Sherwood J.A., Frost C.C., 1992, Constitutive modeling and simulation of energy absorbing polyurethane foam under impact loading, *Polymer Engineering and Science*, **32**, 16, 1138-1146

17. Xing Z., Cen Q., Wang Q., Li L., Wang Z., Liu L., 2024, Compressive mechanical behavior and corresponding failure mechanism of polymethacrylimide foam induced by thermo-mechanical coupling, *Polymers*, **16**, 9, 1199

Manuscript received July 21, 2024; accepted for publication November 11, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 103-113, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195717

TOOL FOR ANALYSIS OF TRAFFIC OF VERTICAL TAKE-OFF AND LANDING AIRCRAFT IN URBAN AGGLOMERATIONS

Antoni Kopyt, Mateusz Sochacki, Kacper Kaczmarek
Warsaw University of Technology, Institute of Aeronautics and Applied Mechanics, Poland
e-mail: antoni.kopyt@pw.edu.pl; mateusz.sochacki@pw.edu.pl; kacper.kaczmarek3.stud@pw.edu.pl

The concept of Urban Air Mobility (UAM) assumes that a significant part of road traffic and services will be shifted to the airspace. The appearance of a large number of VTOLs over urban agglomerations must be preceded by a set of ATM and risks analysis. In order to support this process, a tool for aircraft traffic analysis in urban agglomerations has been developed. In the article, the software is presented, where all crucial features and configuration are shown. To prove its usability, a couple of simulations have been conducted. Finally, the results of test simulations and their analysis are described.

Keywords: Urban Air Mobility, air traffic analysis, flight simulation, vertical take-off and landing aircraft, unmanned aerial vehicles

1. Introduction

The growing complexity of advanced aviation systems brings new threats to safety and security and poses several regulatory and technical challenges. As UAV traffic increases, it is important to ensure the safe integration of UAV into an air traffic management system initially designed to support manned aircraft (Thipphavong *et al.*, 2018). The growing use of UAVs in urban environment will increase the risks of potential collisions, various incidents and accidents that can result in serious losses or injury. Since the use and applications of UAV are increasing and diversifying, it is necessary to develop effective risk management practices, methodologies, and processes in order to ensure the safety of operations. Identifying and understanding the risk factors (Castro and Garcia, 2021; Ellis *et al.*, 2020) is a key step in developing risk mitigation and response measures for UAV operating in urban environments.

To verify various approaches, methods and air traffic rules to the UAM, a dedicated simulation tool was developed. A virtual environment in which different scenarios could be tested may be used to identify and mitigate some risks in the UAM. The goal of the UAM TAT is to identify potential hazards and then determine which safety regulations and procedures applied are most reliable and safe.

From among the research (European Aviation Safety Agency, 2015; Meincke *et al.*, 2022) it seems that a centralised method for controlling UAV may not be efficient and it is needed to apply a decentralised approach. Conducting various case studies, it is possible that not all methods and regulations are scalable from general aviation. The UAM TAT allows one to simulate various platforms with different range of autonomy, dynamic properties (e.g. velocity, turn radius, weight, onboard sensors, ability to S&A). Consequently, the software and hardware onboard the UAVs needs to be unified and integrated, which is a significant challenge to overcome in the future. Conducting simulations by varying those parameters could result in defining the regulations for national aviation authorities regarding the minimum requirements for UAVs to be approved in the current U-Space. The developed tool is also capable of defining various methods of air traffic rules such as queuing, prioritization and emergency situations. This feature could result in determining effective and safe methods of controlling the U-space. In the following paper, the tool for traffic analysis and its features is presented, followed by some case studies.

2. Input data, modelling and simulation

Urban Air Mobility Traffic Analysis Tool is comprised of several modules which allow one to run the full simulation and obtain the desired output:

- input data generation script,
- input data preprocessing module,
- simulation module,
- plot and output files generation module.

2.1. Input data generation

The goal of the input data generation script is to create an input file containing data about each simulated VTOL based on parameters provided by the user which are presented in Tables 1 and 2. These parameters are stored as JSON-formatted text in population declaration files.

Table 1. Description of parameters used for invariant input data generation

Parameter name	Description
familyName	VTOL family name, it has no influence on their behavior during the simulation
numberOfDrones	Number of VTOLs of this family present in the simulation
failureRate	Failure rate parameter, it influences the frequency of malfunctions occurring in the simulation

VTOLs waypoints are generated by using population density data in the simulated region. They are published by Statistics Poland as one-kilometer grids (GIS Support, 2022). The no-fly zones for presented airspace are published by Polish Air Navigation Services Agency (2019, 2023). The user can introduce any To allow for testing new Urban Air Mobility traffic solutions, it is possible to add user-defined zones to the simulation.

The first step of preparing simulation is generation of parameters for particular families of VTOLs. The parameters shown in Table 1 such as `familyName` and `failureRate` are directly assigned to each VTOL from the family, whereas `numberOfDrones` is the number of VTOLs that should have their data generated. Parameters shown in Table 2 are randomly generated for each VTOL. For parameters containing a probability density function described by user-provided weights the program uses this data to randomly draw their values. Four types of routes were implemented in the program. The generated waypoints are ensured not to lie inside a no-fly zone.

2.1.1. Point-to-point

Based on the population density, the program generates take off and landing waypoints. It is the simplest type of route, comprising of two waypoints only. This type represents e.g. taxi routes.

2.1.2. Hub-to-point

VTOL takes off and lands in one of the hubs and flies via waypoints. This implementation simulates courier deliveries – parcels are picked up from large hubs and then delivered to cus-

Table 2. Description of parameters used for randomized input data generation

Parameter name	Description
autonomyLevel	Two-row matrix containing VTOL autonomy levels and corresponding weights. The autonomy level of a particular VTOL is drawn according to this probability density function
size	Vector containing lower and upper bounds of VTOL size. The size of a particular VTOL is drawn using a uniform distribution with these parameters applied
safetyZone	Vector containing lower and upper bounds of VTOL safety zone size. The size of a particular safety zone is drawn using a uniform distribution with these parameters applied
velocity	Vector containing lower and upper bounds of VTOL cruise speed. The speed of a particular VTOL is drawn using a uniform distribution with these parameters applied
range	Vector containing lower and upper bounds of VTOL range. The range of a particular VTOL is drawn using a uniform distribution with these parameters applied
altitude	Vector containing lower and upper bounds of VTOL cruise altitude. The cruise altitude of a particular VTOL is drawn using a uniform distribution with these parameters applied
departureTimes Distribution	Two-row matrix containing departure hours and corresponding weights. The departure time of a particular VTOL is drawn according to this probability density function
routeType	VTOL route type: "point-to-point", "hub-to-point", "hub-to-point-to-point-to-hub" or "loitering"
hubsList	Matrix containing geographic coordinates of VTOL hubs and weights corresponding to them

tomers. Deliveries statistics were estimated, based on a parcel locker used by a leading Polish parcel delivery company – InPost. It was estimated that:

- 32% of VTOLs would fly to between 1 and 2 waypoints,
- 44% of VTOLs would fly to between 3 and 4 waypoints,
- 23% of VTOLs would fly to between 5 and 8 waypoints.

The number of waypoints is drawn from a Weibull distribution whose probability density function is described as follows:

$$f(x;\lambda,k)=\frac{k}{\lambda}\Big(\frac{x}{\lambda}\Big)^{k-1}\mathrm{e}^{-(x/\lambda)^k} \tag{2.1}$$

The Weibull distribution parameters allowing for satisfying the estimated number of deliveries were calculated to be: $\lambda = 3.31506$ and $k = 1.88177$. The mean number of delivered parcels per VTOL for these parameters is 3.41.

After summing up the total number of waypoints for a given VTOL family, the program draws the waypoints locations based on the population density distribution. Then, one of the waypoints is chosen as the first point of the route. The next step is to pick the following waypoints as the ones closest to the first point. Moreover, the closest hub is selected as the takeoff location. It ensures optimal distribution of the "parcels" in the hubs. The last step is to apply a travelling salesman algorithm to the waypoints in order to determine their optimal order.

2.1.3. Hub-to-point-to-point-to-hub

In this type of route the VTOL takes off in one of the predetermined hubs, flies to two points and ends its flight in one of the hubs (possibly the first one). This type of route simulates food deliveries from restaurants to clients.

2.1.4. Loitering

This type of route simulates the flight of VTOLs which are used to fly in specific areas (e.g. for recreational or inspection purposes). They move chaotically over a limited area, take off and ending their flight in the same spot. The waypoints are drawn taking into account the VTOL range.

2.2. Input data preprocessing

Since the input data contains only the main waypoints, it is necessary to determine a path which the VTOL should follow in order to avoid entering the no-fly zones and take into account the limited-fly zones.

2.2.1. No-fly zones

If a route leg crosses a no-fly zone, the program determines intermediate waypoints in such a way that flying through this zone is avoided. For this purpose, the A* algorithm (Hart *et al.*, 1968) is used.

2.2.2. Limited-fly zones

The limited-fly zone, as opposed to the no-fly zone, is not entirely inaccessible for VTOLs. The VTOL can fly in the zone if its departure or destination point is located there. The intermediate waypoints are determined in such a way that the time of the VTOL presence in the zone is minimized.

2.3. Simulation module

The VTOL is modelled as a point particle with 3 degrees of freedom. In order to determine its position and velocity in consecutive steps of the simulation, the following equation of motion is used

$$\mathbf{r}_{i+1}^{k} = \mathbf{r}_{i}^{k} + \mathbf{v}^{k}\varDelta t \tag{2.2}$$

where: $\mathbf{r}$ – VTOL position vector, $\mathbf{v}$ – VTOL velocity vector, k – VTOL number, i – time step number, $\varDelta t$ – time step.

A constant time step is used in the simulation (usually set as $\varDelta t = 1\,\mathrm{s}$). Components of the velocity vector are determined based on the direction from the VTOL position in the current time step to the next waypoint position and the cruise speed of the VTOL

$$\mathbf{v}^{k} = \frac{\mathbf{w}^{k} - \mathbf{r}_{i}^{k}}{\|\mathbf{w}^{k} - \mathbf{r}_{i}^{k}\|}V^{k} \tag{2.3}$$

where: $\mathbf{w}^{k}$ – next waypoint position, V^{k} – cruise speed.

One of the most crucial functionality is the detection of collisions between the VTOLs and violations of safety zones (Żugaj *et al.*, 2016). The safety zones were implemented to help analyze cases in which it might be desirable to ensure a specific separation distance between the VTOLs. In the case of VTOLs collision, which are modelled as spheres in this part of the program, it is checked whether the distance between the centers of the spheres is smaller than the sum of their radii (which is a half of their size). Safety zone violation is checked in the same way, however the safety zones are generally larger. The program simulates drone malfunctions (GPS loss, bad data transmission, subsystems failure) based on a preassigned failure rate (Jankowski *et al.*, 2024). Three types of malfunctions were implemented:

- free fall,
- autonomy level drop,
- loitering.

As part of the program settings, the user has the ability to input probability levels for the occurrence of each type of malfunction. It is important to note that a VTOL with the "0" level of autonomy cannot experience a malfunction causing it to drop its autonomy level or start loitering.

2.3.1. Free fall

If the drone experiences the free fall, its horizontal velocity is zeroed. Its vertical velocity is calculated based on a predefined ratio of the free-fall to cruising velocity.

2.3.2. Autonomy level drop

After detecting an autonomy level drop, it is reduced by 1, and if it falls to 0, the VTOL takes off loitering.

2.3.3. Loitering

In the case of the loitering malfunction, the program determines the loitering route waypoints based on the current VTOL position, its range and the distance that it flew up to this moment.

2.4. Output data generation

The program is able to generate several types of output files which help analyzing the simulation results.

2.4.1. Animation

The first type of output file is an animation which visually presents the air traffic over the simulated region (see Fig. 1). The generated map presents the VTOLs that were airborne at the particular point in time. Different colors allow one to distinguish VTOLs with different autonomy levels. The no-fly and limited-fly zones are marked with red and yellow polygons, respectively. Moreover, the animation shows the current local simulation time and the number of airborne VTOLs.

2.4.2. Heat maps

Another type of output files are heat maps. They present the traffic density over a particular region. Before starting the simulation, the whole region is divided into bins. During the simulation, the number of VTOLs present in the bins at the particular time steps is summed up. It allows one to quickly evaluate where the most dense traffic occurs. Moreover, heat maps presenting the number of collisions generated, as well as the number of collision avoidance manoeuvres.

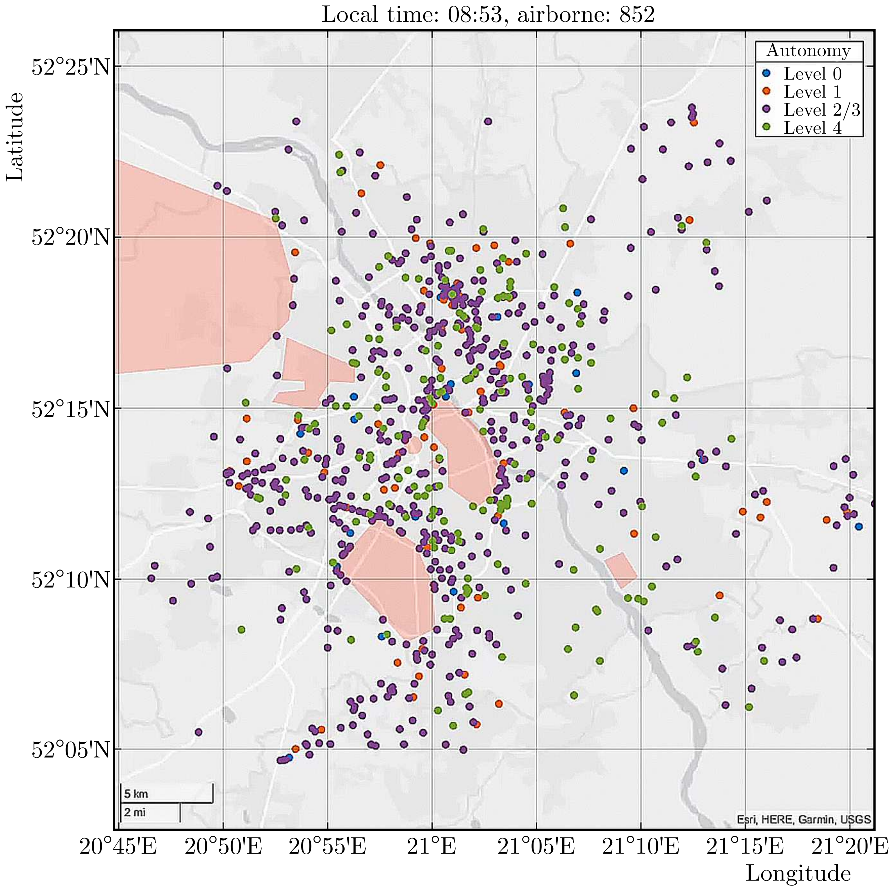


Fig. 1. Animation screenshot

2.4.3. Number of airborne VTOLs

In order to show the daily distribution of traffic, the program generates a plot of the number of airborne VTOLs as a function of time. The plot shows how the number of airborne VTOLs changes during a day, taking into account the division between the autonomy levels.

2.4.4. Collisions histogram

The next type of plot is the collisions histogram. The plot presents the number of collisions as a function of time.

2.4.5. Simulation report

The last type of output file is the simulation report which is the only output stored as a text file. It stores the most important metrics about the performed simulation. The file contains information such as: simulation start time, local simulation time at its beginning, simulation time duration, total number of VTOLs in the simulation, table containing the number of VTOLs present in the simulation divided based on the autonomy level, total number of collided VTOLs, total number of collision avoidance manoeuvres.

3. Case study

This Section covers several examples examined using UAM TAT. This case study investigates the influence of presence of no-fly zones and limited-fly zones over the area of Warsaw (Poland) on the traffic density and collision probability. The three presented cases use the same input data in terms of VTOL flight distribution and only differ in airspace configuration.

3.1. Input data

Four VTOL families were taken into account in the case study, i.e.: recreational drones, air taxis, parcel delivery and food delivery.

The recreational drone family consists of 100 VTOLs, all of which have the lowest autonomy level (0). Their route type is “loitering” with origins distributed according to the population density in Warsaw. The assumed departure times distribution is presented in Fig. 2a.

The taxi family consists of 20000 VTOLs, all with the highest autonomy level (4). They fly “point-to-point” route type with departure and destination points distributed according to population density. Departure times are distributed according to daily road traffic distribution (see Fig. 2b, source Warsaw Metropolitan Roads Administration, 2923).

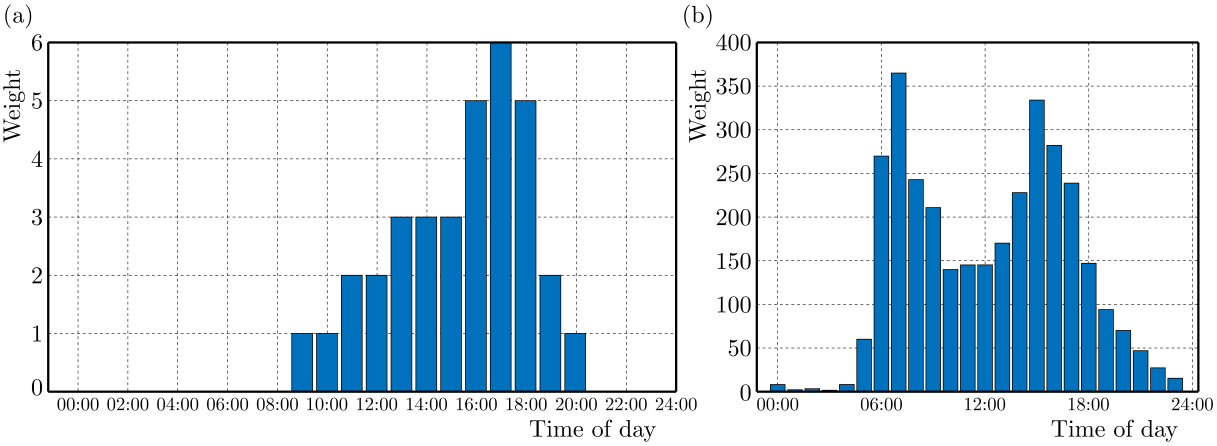


Fig. 2. (a) Recreational VTOLs family departure times distribution. (b) Air taxi family departure times distribution

The parcel delivery family consists of 45365 VTOLs with various autonomy levels (between 1 and 3). They fly “hub-to-point” route type with a 3.41 average number of stops distributed according to population density. Departure hubs were associated with the major parcel delivery hubs in the Warsaw area. The distribution of departures per delivery hub was determined based on the market share of the particular delivery company (InPost S.A., 2023; Laast Mile Experts, 2022). Parcel delivery departure times were assumed to be uniformly distributed between 07:00 and 22:00.

The food delivery family consists of 7706 VTOLs with various autonomy levels (between 1 and 3). They fly “hub-to-point-to-point-to-hub” route type with two stops. The second stops (associated with the delivery point) were distributed according to population density, while the first ones (associated with pick-up points) were randomly selected in the vicinity of the delivery point. Departure hubs were associated with the major shopping centres in Warsaw (as a suitable area for housing large VTOL maintenance facilities). The nearest hub is selected for departure. The daily distribution of food deliveries was taken from (Stava, 2023). The cruise altitude is randomly selected for each VTOL with the exception of recreational drones for which the altitude varies during the flight. All three simulations cover a period of 24 hours and the simulation time step was 1 second.

3.2. Results

Figures 3 and 4a present the heat maps of airborne VTOLs for three simulations (i.e. no airspace restrictions, no-fly zones, no-fly zones and limited-fly zones cases, respectively). The heat maps are expressed in drones per hectare which must be understood as the number of VTOLs visible over an area of one hectare for one day over a period of one day (i.e. 1 drone/ha means that over an area of one hectare there was one drone visible continuously for the duration of the entire day). VTOL hubs are clearly visible, as well as natural corridors of increased air traffic appearing after the introduction of airspace limitations.

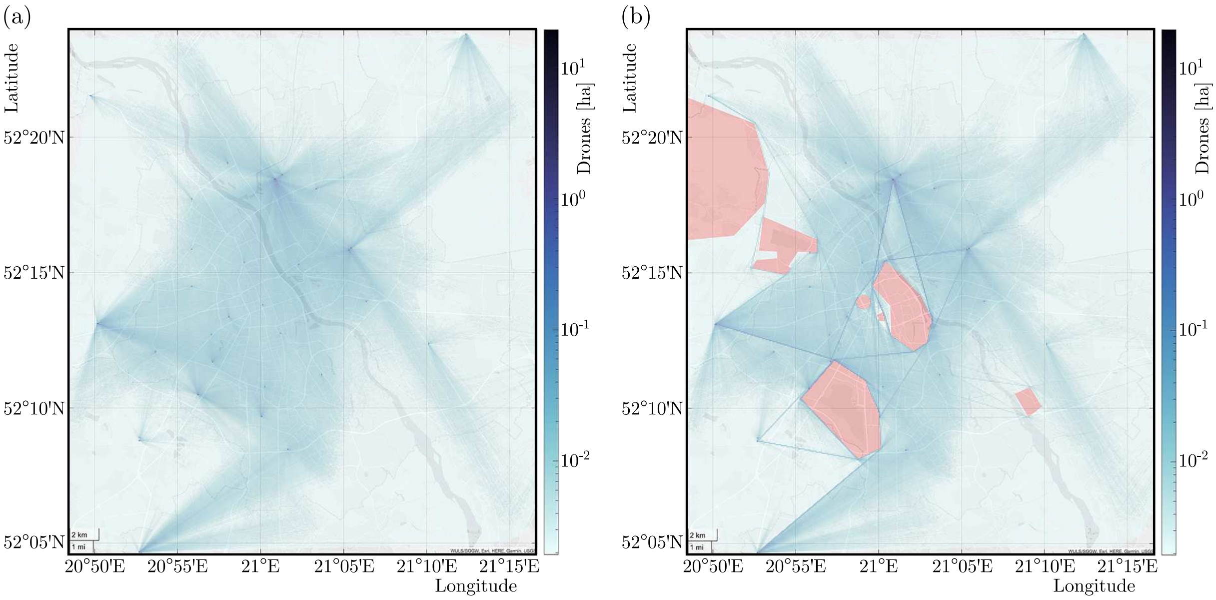


Fig. 3. Heat map of airborne VTOLs: (a) case 1, no airspace restrictions, (b) case 2, no-fly zones (red)

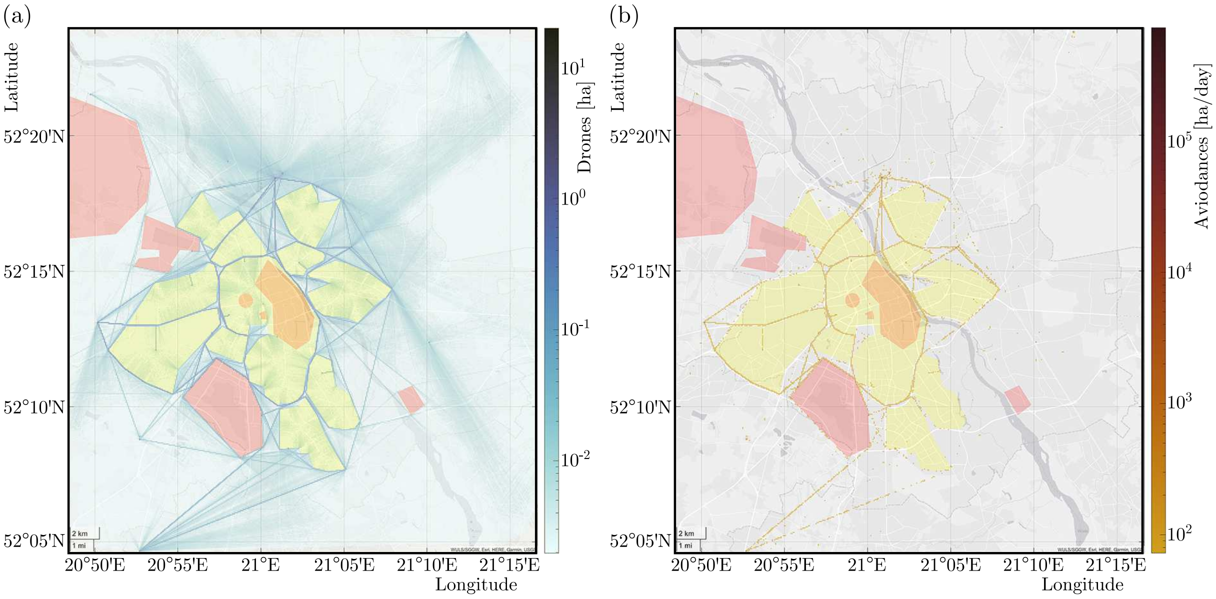


Fig. 4. (a) Heat map of airborne VTOL – case 3, no-fly zones (red) and limited-fly zones (yellow). (b) Heat map of VTOL collision avoidance manoeuvres – case 3, no-fly zones (red) and limited-fly zones (yellow)

Figure 4b presents the heat map of VTOLs collision avoidance manoeuvres for the third case (with most airspace restrictions). It should be noted that although there is a visible increase in the number of events in the flight corridors, almost all events occur at the hubs where VTOLs take off and land.

Figure 5 presents daily airborne VTOLs distribution for each autonomy level (only one graph, corresponding to case 3, is presented as the differences between cases were insignificant).

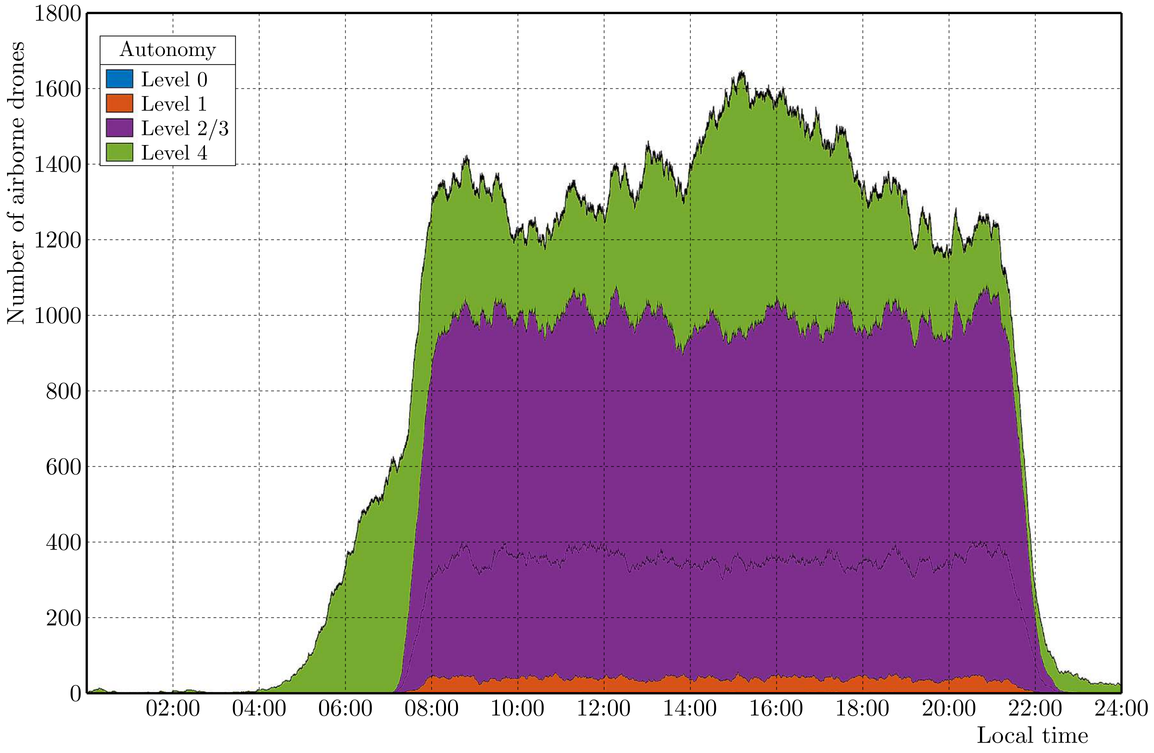


Fig. 5. Number of airborne VTOLs as a function of time – case 3, no-fly zones and limited-fly zones

Figure 6 presents the daily distribution of collisions and avoidance manoeuvres (only one graph, corresponding to case 3, is presented as the differences between cases were insignificant).

Table 3 summarises the results in terms of the number of collisions and avoidance manoeuvres. Counter-intuitively, it seems that the number of events does not increase with airspace restrictions but this effect is just dwarfed by the number of events in the vicinity of hubs.

Table 3. Number of collisions and collision avoidance manoeuvres for each simulation case

Case	Number of collisions	Number of collision avoidance manoeuvres
No airspace restrictions	897	738825
No-fly zones	865	716675
No-fly zones and limited-fly zones	832	748061

4. Conclusion and further research

The following paper demonstrates a tool dedicated for Urban Air Mobility simulation. It allows one to implement various families of VTOLs and different air traffic rules. Its modularity makes it easy to add new parameters or constraints. It is possible to extend and develop deeper the dynamics of the drones, their inertia and internal features that could impact their interactions,

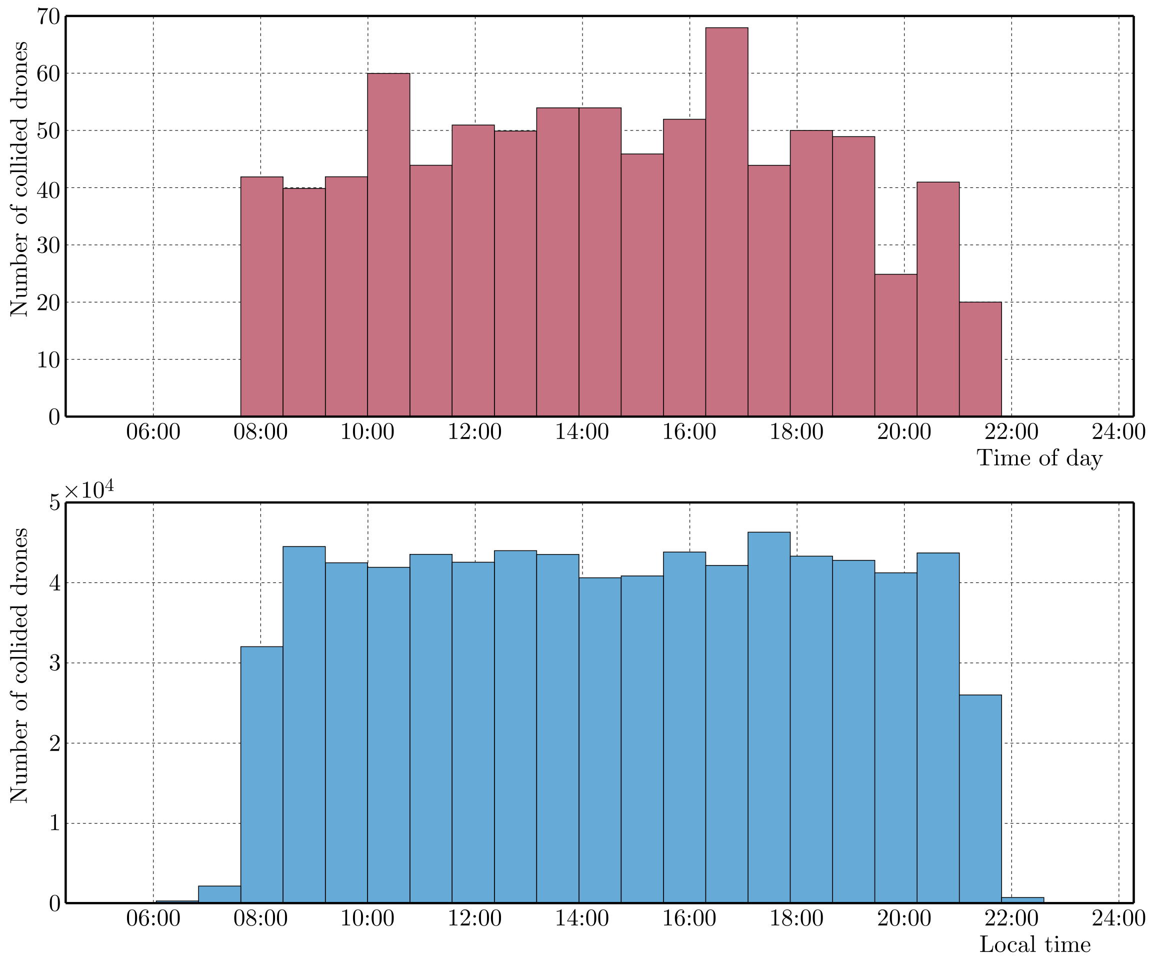


Fig. 6. Collision and collision avoidance manoeuvre histograms – case 3, no-fly zones and limited-fly zones

as well as to expand air traffic management such as to include optimisation of cruising altitude assignment. The presented system features provide a complete tool for simulation that could be used e.g. for national civil aviation authorities and U-Space users to define and estimate the capacity of the given airspace. Secondly, the tool could be used to define VTOLs requirements that have to be met so they are allowed to the traffic. Finally, various rules and methods for ATM could be tested and compared to provide the safest solution.

Acknowledgement

This work was initially presented at Delft International Conference on Urban Air Mobility – DICUAM Conference, 20-24 march 2024, Delft, Netherlands and further developed by the team under the Urban Air Mobility Project at Warsaw University of Technology, Warsaw, Poland.

References

1. Castro D.G., García E.V., 2021, Safety challenges for integrating U-space in urban environments, *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, Athens, Greece, 1258-1267
2. Ellis K., Koelling J., Davies M., Krois P., 2020, *In-Time System-Wide Safety Assurance (ISSA) Concept of Operations and Design Considerations for Urban Air Mobility (UAM)*, Technical Report, Hampton, USA

3. European Aviation Safety Agency, 2015, *Technical Opinion – Introduction of a regulatory framework for the operation of unmanned aircraft*, Technical Report
4. GIS Support, 2022, *Dane statystyczne GUS* (in Polish)
5. Hart P.E., Nilsson N.J., Raphael B., 1968, A formal basis for the heuristic determination of minimum cost paths, *IEEE Transactions on Systems Science and Cybernetics*, **4**, 2, 100-107
6. InPost S.A., 2023, *Annual Report 2022*, Technical Report
7. Jankowski Ł., Pisarski D., Konowrocki R., Popławski B., Faraj R., 2024, Efficient real-time positioning using Bayesian analysis and magnetic anomaly field, *Measurement*, **233**, 114738
8. Last Mile Experts, 2022, *Polish CEP Report 2022*, Technical Report
9. Meincke P., Duca G., Ciaburri M., Russo R., Enei R., *et al.*, 2022, Concept of operations for ATM service to passengers in intermodal transport system, *SESAR*, Technical Report
10. Polish Air Navigation Services Agency, 2019, *AIP VFR ENR 2.1.1 Prohibited Areas*
11. Polish Air Navigation Services Agency, 2023, *PansaUTM*
12. Stava, 2023, *Stava Report on Food Delivery Market in Poland 2023* (in Polish), Technical Report
13. Thipphavong D.P., Apaza R., Barmore B., Battiste V., Burian B., *et al.*, 2018, Urban air mobility airspace integration concepts and considerations, *2018 Aviation Technology, Integration, and Operations Conference*, AIAA AVIATION Forum, American Institute of Aeronautics and Astronautics
14. Warsaw Metropolitan Roads Administration (Zarząd Dróg Miejskich w Warszawie), 2023, *Results of Automatic Traffic Measurement System – Year 2022* (Wyniki pomiarów systemu automatycznych pomiarów ruchu APR ZDM – 2022 rok)
15. Żugaj M., Bibik P., Jacewicz M., 2016, UAV aircraft model for control system failures analysis, *Journal of Theoretical and Applied Mechanics*, **54**, 4, 1405-1415

Manuscript received July 15, 2024; accepted for publication November 8, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 115-129, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195720

DYNAMICS RESPONSE ANALYSIS AND MECHANICAL EXPERIMENTAL INVESTIGATION OF A DOUBLE-SPAN MULTI-SUPPORT SHAFT SYSTEM IN MICRO GAS TURBINES

Hao Lin, Leiming Song
Beijing Jiaotong University, School of Mechanical, Electronic and Control Engineering, Beijing, China
corresponding author L. Song, e-mail: lmsong@bjtu.edu.cn
Haipeng Geng, Sheng Feng
Xi'an Jiaotong University, School of Mechanical Engineering, Xi'an, China
Ling Li
Beijing Jiaotong University, School of Mechanical, Electronic and Control Engineering, Beijing, China

This study delves into analysis of dynamics and experimentation concerning a double-span multi-support shaft system in micro gas turbines. Critical speed analysis was conducted, accompanied by numerical solutions providing insights into the natural frequency. Additionally, an unbalanced response and vibration amplitude of the shaft were calculated, serving as crucial inputs for dynamic balancing considerations. The study also examined the influence of stiffness coefficient variations in bearings on the unbalance response. Dynamic balancing experiments of the motor rotor and the compressor-turbine rotor were conducted separately. Then a vibration experiment for the double-span multi-support shaft was executed at an experimental speed of 60000 r/min. The research revealed that rotors could be regarded as rigid bodies throughout the operational speed range. This study holds significant engineering implications and a practical application value.

Keywords: micro gas turbine, double-span multi-support shaft system, dynamics analysis, unbalance response, experiment

1. Introduction

1.1. Background

The industrialization process is increasingly constrained by resource and environmental limitations, as the high consumption and emission levels associated with industrial development are no longer sustainable. Consequently, environmental and energy concerns are gaining prominence. The utilization of reliable and stable high-end energy and power equipment emerges as a viable solution to address issues related to energy shortages and environmental pollution.

A significant portion of primary energy, approximately 66.4%, is directly dissipated into environment as thermal energy (Vishwanathan *et al.*, 2018). Notably, there exists a substantial amount of medium and low thermal energy that remains underutilized, highlighting the potential for energy recovery in micro gas turbines. Globally, the capacity of micro gas turbines ranges from 25 kW to 250 kW. Through combined generation of electricity and heat, these turbines achieve an impressive efficiency rate of 80%. Micro gas turbines offer various advantages, including a minimal number of rotating parts, compact design, small size and weight, as well as ease of assembly and maintenance (Wołowicz *et al.*, 2021; Villarroel-Schneider *et al.*, 2019). Particularly noteworthy are their low levels of pollutant emissions and noise, rendering micro gas turbines suitable for deployment in diverse facilities.

The adoption of advanced technologies, such as high-speed permanent magnet synchronous motors, flexible couplings, and elastic foil bearings, characterizes micro gas turbines. Leveraging the characteristics of high-speed and direct drive, these micro gas turbines efficiently recover energy and generate electricity (Bo *et al.*, 2018). Figure 1 illustrates the working principle of the system (Singh *et al.*, 2022).

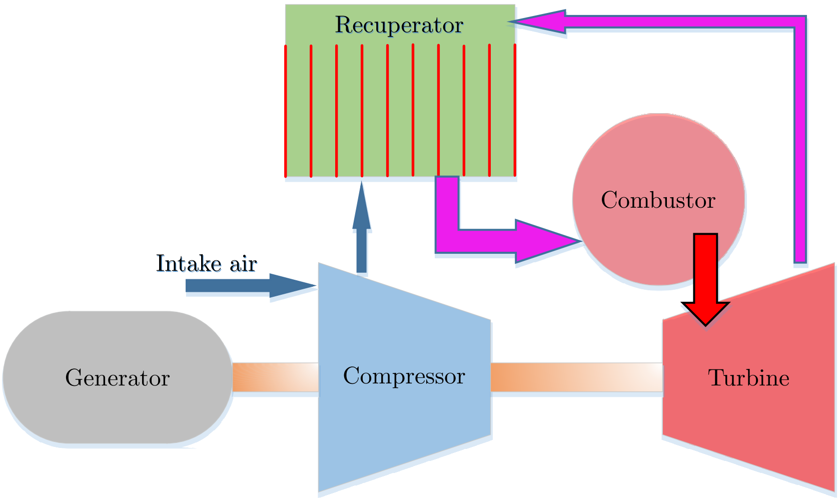


Fig. 1. The working principle of the micro gas turbine system

Ansaldo Energia (Ramaglia *et al.*, 2018) and Capstone (Othman and Boosroh, 2016) have produced micro gas turbines within the 30-200 kW power range, employing a double-span shaft with multi-supports. However, specific structural parameters and design schemes for the shaft system remain undisclosed. The pivotal technology of the high-speed shaft system, supported by oil-free lubrication bearings, significantly influences the performance of micro gas turbines, encompassing aspects like the high-speed coupling and compatibility of the bearing-shaft system. The high speed of the micro gas turbine, large pressure ratio, high efficiency, and high power density face constraints imposed by various factors. Ensuring the stable and reliable operation of the shaft system amid continuous speed improvements remains an ongoing and in-depth exploration.

The overarching objective of the industry is to maximize the operational speed of the high-speed double-span multi-support shaft system. However, achieving high rotation speed necessitates the shaft system to traverse multiple critical speeds during operation, rendering the amplitude limitation of the high-speed shaft system more challenging.

1.2. Research status

In the realm of high-speed, direct-drive energy machinery, the shaft system typically adopts a series structure. Power machinery with low-rated operating speeds, such as large steam turbine generators operating at 3000 r/min, commonly employs the typical series structure for the shaft system. Couplings are utilized to connect high and low-pressure rotors (permanent magnet rotors and generator rotors). The shaft system is supported by multiple bearings and encounters multiple critical speeds during operation. Under high-speed conditions, the personalized characteristics of shaft systems with varying speeds and powers render the structure of high-speed shaft systems non-universal. Notably, the amplitude limitation of the shaft system at critical speeds is particularly significant.

In the double-span shaft of the micro gas turbine, the motor rotor and turbine/compressor rotor are treated as rigid body rotors, respectively. A flexible coupling is employed to connect these two rotors. The critical speed of the flexible coupling plays a pivotal role in the overall performance of the shaft system. Koo *et al.* (2015) conducted analysis of the vibration of a shaft supported by elastic foil bearings with a diaphragm coupling. Schmid and Pradeto (1994) studied the rotor dynamic characteristics of a high-speed motor-compressor supported by electromagnetic bearings connected via a diaphragm coupling, highlighting the diaphragm coupling role in providing a significant margin between the maximum operating speed and the critical bending frequency. Lorenzen *et al.* (1989) compared the dynamic characteristics of shaft systems connected by different couplings and found that the diaphragm coupling offered a margin between the critical speed and operating speed. The structure of the diaphragm coupling is depicted in Fig. 2.

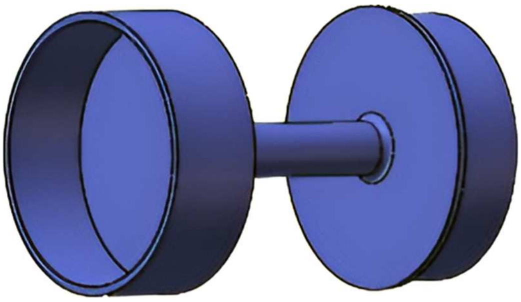


Fig. 2. Structure of the diaphragm coupling

While the aforementioned studies have delved into the dynamics of double-span multi--support shaft systems, they predominantly fall within the realm of the qualitative analysis. Emphasis has been placed on identifying reasonable critical speeds; however, there is a noticeable absence of detailed reports on the dynamic response of high-speed double-span multi-support shaft systems in micro gas turbines. Given the challenges posed by the high operating speed and the low stiffness-damping characteristics of gas bearings, successfully navigating the critical speed is a yet-to-be-resolved issue. Furthermore, insufficient attention has been devoted to elucidating strategies for traversing the first bending critical speed of the diaphragm coupling and judiciously establishing the margin between the critical speed and operational speed.

The selection of the micro gas turbine shaft system structure is multifaceted, contingent upon varying power levels and structural parameters. Following the proposal of the high-speed shaft system structure, uncertainties persist regarding the ability of elastic foil bearings to damp vibrations and the dynamic behavior of the shaft system under high-speed conditions. Beyond theoretical analysis, validation hinges on operational experiments involving the shaft system. Unfortunately, there is a paucity of related experiments for different diaphragm coupling and double-span shaft structures. Heshmat *et al.* (2014) presented an 8 kW turbine generator supported by elastic foil gas bearings connected via a diaphragm coupling. Walton *et al.* (2012) conducted operational experiments on two motors connected by a diaphragm coupling, and Heshmat and Walton (2016) experimented with a motor-flywheel supported by elastic foil gas bearings connected with a diaphragm coupling. Andrés *et al.* (2015) incorporated the diaphragm coupling in a rotor-bearing test device, yet instability was observed during the experiment at the speed range of 26000 r/min to 27000 r/min. These mixed experimental outcomes underscore the significance of experimental verification for the structure of high-speed double-span shaft systems. However, these experiments fall short of providing structure parameters and optimal dynamic characteristics for the flexible coupling, bearing, and shaft – essential for the effective operation of the shaft system. Due to the personalized structural characteristics and the scarcity of practical examples, there is a pressing need for both dynamic analysis and experimental verification of high-speed double-span shaft systems. The primary challenge lies in the complexity of the structure and the actual operation in high-speed shaft systems.

This study addresses this gap by conducting comprehensive dynamics analysis and an experiment of the double-span multi-support shaft system in a micro gas turbine. Firstly, critical speed analysis was performed in Section 2, accompanied by the provision of numerical solutions for natural frequencies. Secondly, the unbalance response and vibration amplitude of the shaft were calculated in Section 3, providing insights for dynamic balancing. The influence of stiffness coefficients of bearings on the unbalance response of the shaft system was analyzed, enabling the selection of appropriate bearing parameters for the shaft system based on the results. Thirdly, in Section 4, a balance experiment was conducted, then a vibration experiment of the double-span multi-support shaft was executed at an experimental speed of 60000 r/min, and the experimental results were employed to verify the analysis outcomes. The research methodology and results hold substantial engineering significance and a practical application value.

2. Analysis of critical speed and modal shape

2.1. Structure of the double-span multi-support shaft system

In the dynamics analysis of high-speed shaft systems, the conventional approach often involves considering two single-span rotors as rigid body modes, ensuring a sufficient margin across the operating speed range (Guan *et al.*, 2020). Adhering to these principles, the motor rotor in this study was designed with an approximately symmetrical structure, supported by two bearings. The mass distribution and moment of inertia of the motor rotor exhibit good uniformity and symmetry.

This study is grounded in practical engineering applications, specifically focusing on micro gas turbines. It provides a comprehensive understanding of the dynamics response of the shaft system, taking into account factors such as functional characteristics, power, rated speed, combination rotor structure, and load distribution of bearings (Kim *et al.*, 2023). The chosen shaft system structure is the double-span shaft, where the motor rotor is connected to the compressor-turbine rotor through a flexible coupling. Regarding support methods, there is flexibility in selecting either three or four bearings (Zywica and Baginski, 2019). The structure of the double-span multi-support shaft system for the micro gas turbine is illustrated in Fig. 3.

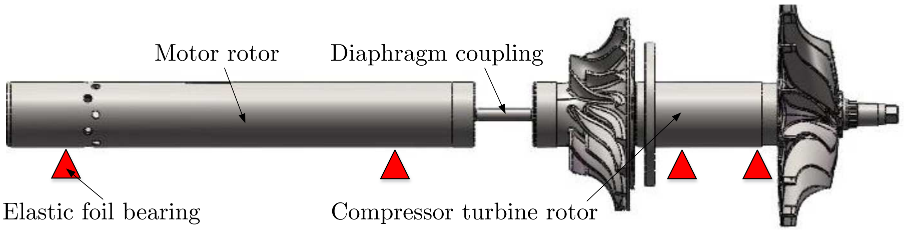


Fig. 3. Structure of the double-span multi-support shaft system

2.2. Theoretical model

In ensuring the stable operation of the shaft system under high-speed conditions, it is imperative to consider the vibration amplitude of the shaft, a parameter closely tied to both the natural frequency and the unbalance response of the shaft.

To delve into the vibration amplitude of the shaft system, the essential step involves determining the natural frequency. This necessitates analysis of the critical speed and modal characteristics of the shaft system (Zheng *et al.*, 2020).

Considering the diaphragms at the left and right ends of the diaphragm coupling as identical elastic components (Heshmat *et al.*, 2018), the displacement relationship between the left and right connection ends of the diaphragm coupling is illustrated in Fig. 4. In this depiction, the shafts represent rotors at both ends and the rod of the coupling. Taking the left end as an example, node i represents the node at the end of the rotor, and node j represents the node at the end of the coupling rod. The lateral displacement of these two nodes is identical, whereas the angular displacement differs. The radial direction is rigid, and the lateral displacements of the nodes are expressed as u_i and v_i.

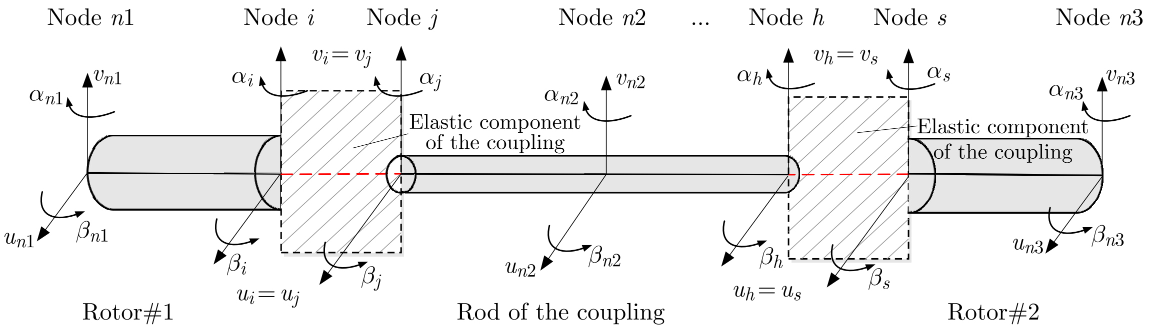


Fig. 4. Displacement relationship at the connection end of the diaphragm coupling

The column vector at the left end of the diaphragm coupling can be mathematically expressed as follows (Feng *et al.*, 2021)

$$\mathbf{q}_c = [u_{n1}, v_{n1}, \beta_{n1}, \alpha_{n1}, u_i, v_i, \beta_i, \alpha_i, \beta_j, \alpha_j, u_{n2}, v_{n2}, \beta_{n2}, \alpha_{n2}]^{\mathrm{T}} \tag{2.1}$$

The mechanical performance of the diaphragm coupling is characterized, as shown in Fig. 5.

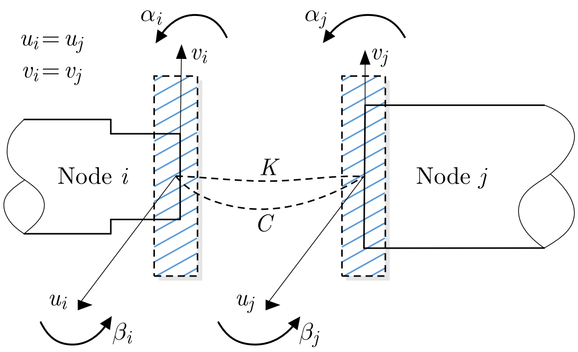


Fig. 5. Mechanical performance of the diaphragm coupling (Feng *et al.*, 2015)

The equation of motion of the shaft system was assembled by the matrix of the shaft and coupling $\mathbf{q}_c$ as follows

$$\mathbf{M}\ddot{\mathbf{q}} + \Omega\mathbf{G}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{Q}(t) \tag{2.2}$$

where the mass matrix $\mathbf{M}$ includes the shaft element matrix $\mathbf{M}_e$; and the disk mass and moment of inertia matrix $\mathbf{M}_d$; the gyroscopic matrix $\mathbf{G}$ includes the gyroscopic matrix of the disk $\mathbf{G}_d$ and the gyroscopic matrix of the shaft $\mathbf{G}_e$; the stiffness matrix $\mathbf{K}$ includes the stiffness matrix of the shaft element $\mathbf{K}_e$ and the stiffness matrix of the bearing $\mathbf{K}_b$; the column vector $\mathbf{q}$ is composed of variables from each node. Ω is the rotational speed.

2.3. Analysis of critical speed and modal shape

In the case of a 30 kW micro gas turbine, operating in the electric motor mode, the startup phase extends from the initial stage to the self-sustaining speed or ignition speed within the range of 0-60000 r/min. Beyond the self-sustaining speed, the micro gas turbine automatically transitions to the generator mode, where it outputs power externally, specifically at the rated operation speed of 96000 r/min. Structure parameters of the double-span multi-support shaft system are presented in Table 1.

Table 1. Structure parameters of the shaft system (30 kW, 96000 r/min)

Parameter	Value
Length of shaft system [mm]	440
Length of motor rotor [mm]	230
Diameter of motor rotor [mm]	32
Length of compressor-turbine rotor [mm]	180
Diameter of compressor-turbine rotor [mm]	32
Bearing support span of motor rotor [mm]	180
Bearing support span of compressor-turbine [mm]	40

To analyze the critical speed of this double-span multi-support shaft system, the calculation model, depicted in Fig. 6, was established by the finite element method. For the critical speed analysis, the right end term of Eq. (2.2) was set to zero, and the characteristic equation was solved. This process was executed using a MATLAB programme, yielding numerical solutions for the natural frequency, which are presented in Table 2.

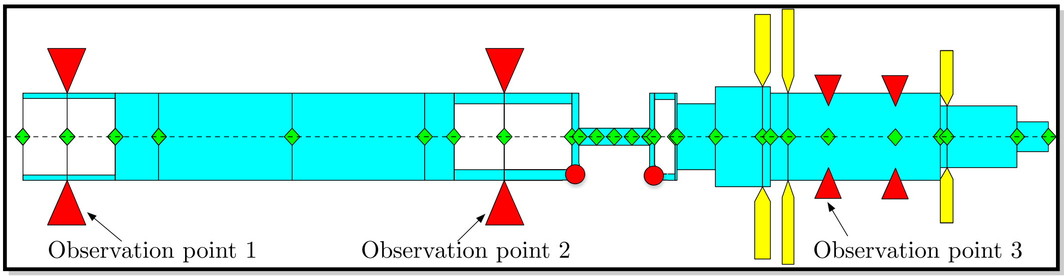
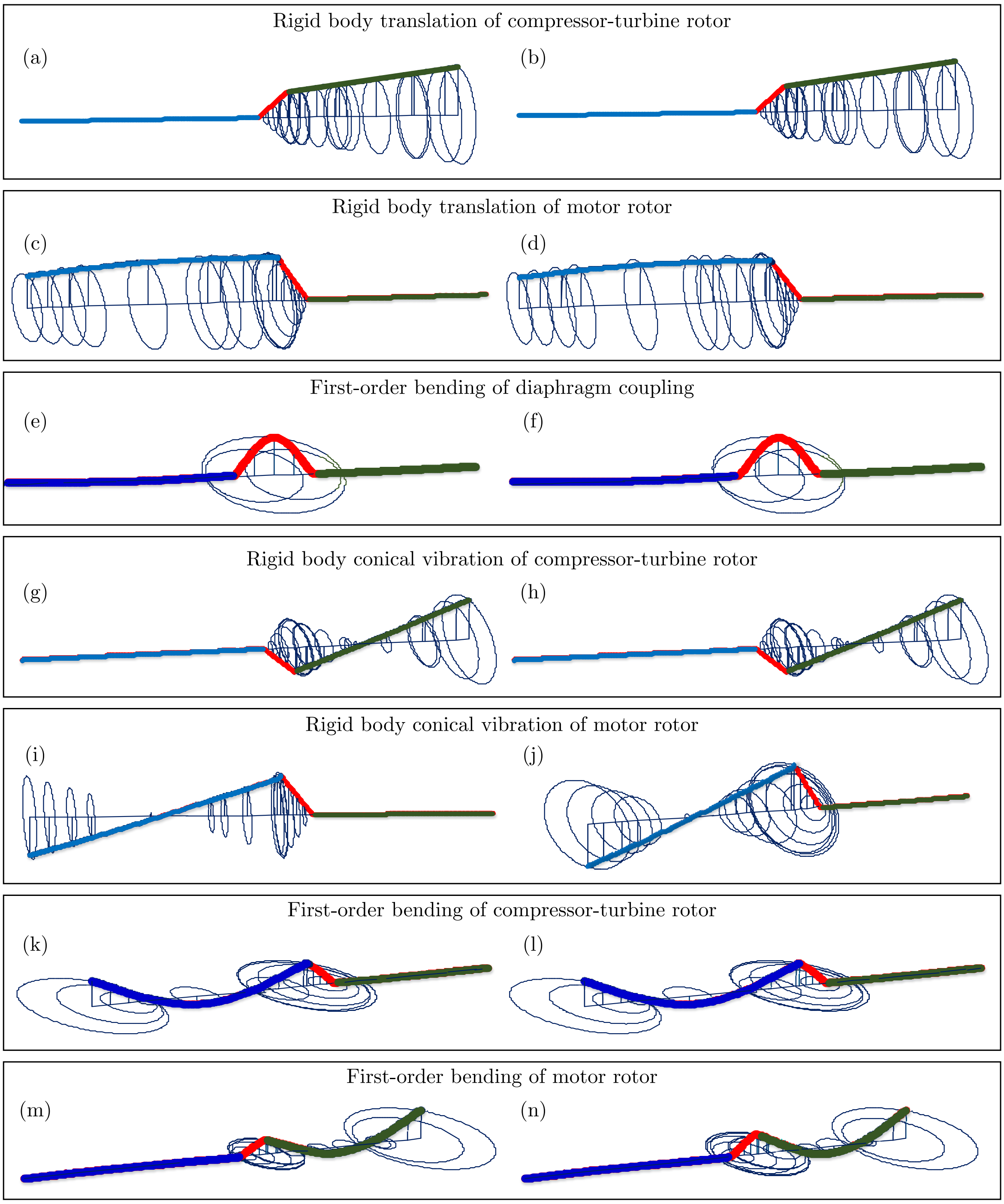


Fig. 6. Schematic diagram of the double-span multi-support shaft system model

Table 2. Natural frequencies of double-span multi-support shaft systems

No.	Natural frequency [Hz]	Rotation speed [r/min]	Vibration modes
a	319	19140	Rigid body translation of compressor-turbine rotor
b	337	20220	Rigid body translation of compressor-turbine rotor
c	392	23520	Rigid body translation of motor rotor
d	395	23700	Rigid body translation of motor rotor
e	427	25620	First-order bending of diaphragm coupling
f	496	29760	First-order bending of diaphragm coupling
g	768	46080	Conical vibration of compressor-turbine rotor
h	817	49020	Conical vibration of compressor-turbine rotor
i	1166	69960	Rigid body conical vibration of motor rotor
j	1217	73020	Rigid body conical vibration of motor rotor
k	3326	199560	First-order bending of compressor-turbine rotor
l	3796	227760	First-order bending of compressor-turbine rotor
m	3865	231900	First-order bending of motor rotor
n	4277	256620	First-order bending of motor rotor

The modal shapes of the shaft are illustrated in Fig. 7. The vibration modes corresponding to the rigid body are depicted in Fig. 7a to Fig. 7d and Fig. 7g to Fig. 7j. Figures 7e and 7f represent the first-order bending vibration mode of the diaphragm coupling. The first-order bending vibration modes of the motor rotor and the compressor-turbine rotor are captured in Fig. 7k to Fig. 7n. Notably, the diaphragm coupling effectively isolates bending vibrations, allowing the vibrations of the rotors to exhibit corresponding independence.


Rigid body translation of compressor-turbine rotor

Rigid body translation of motor rotor

First-order bending of diaphragm coupling

Rigid body conical vibration of compressor-turbine rotor

Rigid body conical vibration of motor rotor

First-order bending of compressor-turbine rotor

First-order bending of motor rotor

Fig. 7. Vibration mode of the double-span combined rotor shaft system

Figure 7 shows modal shapes of the double-span shaft system at the critical speed. The rigid body vibration mode of the shaft describes the rigid shape of both rotors and a flexible deformation of the diaphragm coupling. The flexible vibration mode of the shaft is that when the rotor bends.

In the rigid vibration mode, the critical speeds of forward precession were 20220, 23700, 49020, and 73020 r/min, while the critical speeds of backward precession were 19140, 23520, 46080 and 69960 r/min. Within the rated speed range of 96000 r/min, only rigid vibration modes of the two rotors existed. The first-order bending critical speeds of the diaphragm coupling were 25620 and 29760 r/min, significantly distant from the rated speed of 96000 r/min. The first-order bending critical speeds of the rotors were 199560, 227760, 231900, and 256620 r/min, exceeding one times the rated speed. The rated speed and the bending critical speed provided ample distance, ensuring a sufficient safety margin for the dynamical performance of the shaft.

As depicted in Fig. 7, it is evident that in the low-frequency rigid body mode of the shaft, one rotor exhibits significant vibration while the other rotor experiences minimal vibration. One rotor demonstrates a rigid body vibration mode, encompassing translation and conical vibration. All modes within the rated speed range are rigid body vibrations, indicating that the operation of the shaft is unaffected by rigid body vibration modes. However, caution should be exercised to avoid bending vibration modes. The high-frequency bending modes include the first-order bending modes of the two rotors, with the bending region primarily appearing at positions of the permanent magnet and bearings.

The first-order bending critical speed of the shaft was higher than the rated speed. Rotors were regarded as rigid bodies throughout the operation speed range. The shaft had the sufficient safety margin between the rated speed and the bending critical speed.

3. Analysis of unbalance responses

3.1. Unbalance response calculation

In addition to the resonant natural frequency derived from the preceding analysis, the primary cause of vibration amplitude is the inertial centrifugal force resulting from the unbalanced mass of the rotor. This centrifugal force is directly proportional to the square of the angular speed, making the unbalance response particularly prominent in high-speed systems. The analysis of unbalance response primarily concentrates on examining the sensitivity of the rotor to unbalance at different positions. By calculating the unbalanced response, predictions regarding the vibration amplitude and frequency of the rotor during operation were made. Measures were then implemented to restrict the excessive vibration amplitude.

Given the high operating speed of the system under study, there is a stringent requirement for dynamic balance accuracy. The permissible residual unbalance is calculated as

$$U_e = 1000\frac{mG}{\omega} \tag{3.1}$$

where U_e is the permissible residual unbalance [g·mm], m is the rotor mass [kg], G is the the numerical value of the selected balance quality grade [mm/s], ω is the angular velocity of the service speed [rad/s].

The selected balance quality grade is $G1$. The calculated maximum unbalance of the rotor is 0.263 g·mm. Leveraging the geometric and vibration characteristics of the double-span composite rotor system, unbalance values were applied at various locations, including the left end of the rotor, the center position of the permanent magnet, the thrust disk, and the center position of the disk, as depicted in Fig. 9. Observation points 1 to 3 were strategically chosen to align with the locations of three bearings in Fig. 6. Figures 8 to 11 display the amplitudes at observation points under different unbalance excitations, offering a comprehensive view of the vibration characteristics of the shaft system.

Figure 8 illustrates the response to unbalance excitation at the left end of the double-span composite rotor. Within the rated speed range of 1600 Hz, the peak frequencies of shaft vibration were approximately 506 Hz and 1222 Hz, aligning with the frequencies of rigid body vibration. These frequencies also corresponded to the natural frequencies of 496 Hz and 1217 Hz obtained in the modal analysis. This unbalance excitation mainly induced vibration at observation points 1 and 2, while observation point 3 exhibited almost no vibration response.

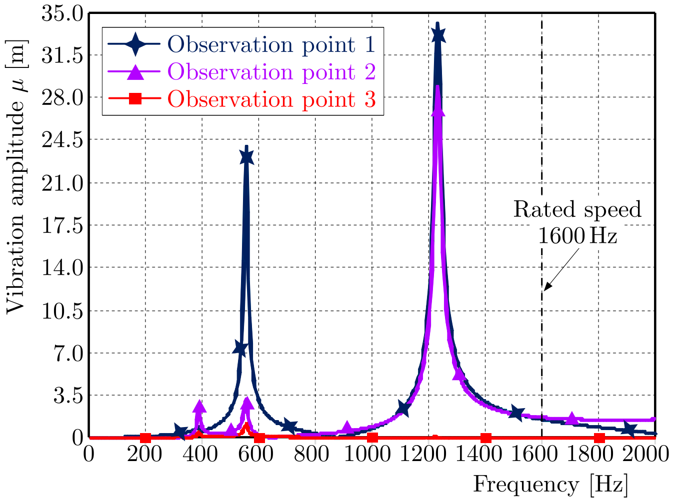


Fig. 8. Results of unbalance responses (unbalance excitation at the left end)

In Fig. 9, the unbalance excitation at the center of the permanent magnet is displayed. Within the rated speed range, the peak frequencies of rotor system vibration were approximately 400 Hz, 506 Hz, and 1222 Hz, corresponding to the natural frequencies of 395 Hz, 496 Hz, and 1217 Hz in the rotor modal analysis. Similar to Fig. 8, this unbalance excitation primarily caused vibration at observation points 1 and 2, with minimal impact on observation point 3. Comparing Fig. 8 and Fig. 9, it is evident that the unbalance excitation applied to the left end had little effect on the unbalance response at the right bearing. The amplitude of the unbalance vibration at the right bearing is considerably smaller than that at the left bearings, suggesting effective isolation by the diaphragm coupling.

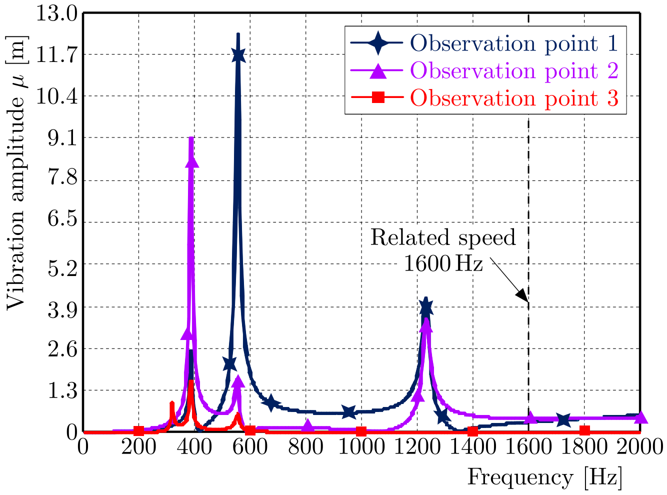


Fig. 9. Results of unbalance responses (unbalance excitation at the center of permanent magnet)

Figure 10 depicts the unbalance response to the excitation of the thrust plate. The peak vibration frequencies at observation points 1 and 2 correspond to the rigid body vibration of the rotor. The peak frequency of vibration at observation point 3 was 300 Hz, aligning with the natural frequency of the rigid body vibration of 319 Hz obtained from modal analysis. The amplitude at observation point 3 was the highest in Fig. 10.

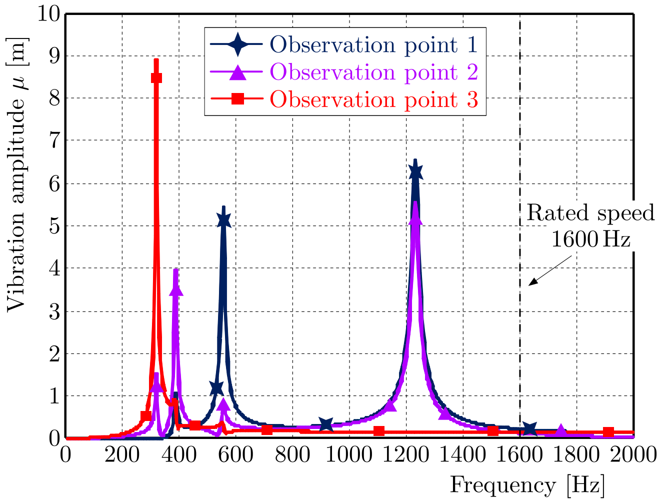


Fig. 10. Results of unbalance responses (unbalance excitation at the thrust disk)

In Fig. 11, the unbalance response to the excitation of the disk is shown, with the curve representing the response corresponding to the vibration frequency. The peak vibration frequencies at observation points 1 to 3 were consistent with those in Fig. 10. Comparing Fig. 10 and Fig. 11, it is apparent that when the unbalance was located at the right end, the vibration frequency and mode at each observation point were the same, and the amplitude at the right bearing was maximum. Thus, the vibration frequency and amplitude were related to the position of the unbalance excitation and the observation point.

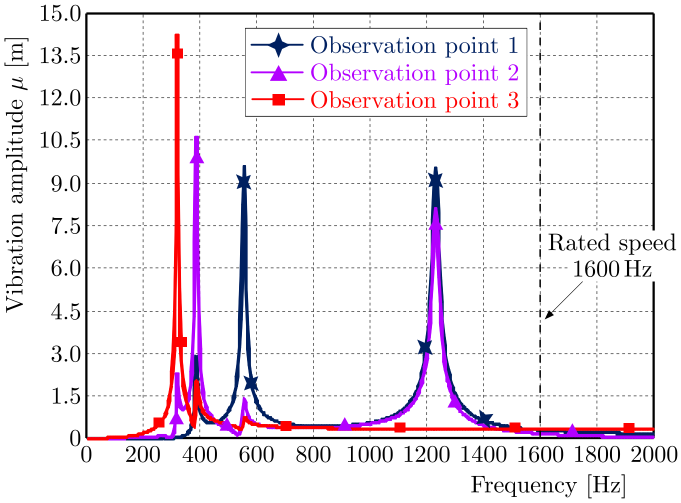


Fig. 11. Results of unbalance responses (unbalance excitation at the disk)

According to the results in Fig. 9, the unbalance response of the vibration amplitude was 35 μm at the left end of the rotor. In other cases, the maximum unbalance response did not exceed 15 μm, and most amplitudes were smaller than this value. Based on the theory of vibration analysis, the unbalance response at the same position is directly proportional to the unbalance excitation. Therefore, the unbalance response of the shaft was obtained by calculating the response at one position based on this proportional relationship.

Examining the trend of the above curves, it is apparent that within the rated speed range of 1600 Hz, the vibration amplitude near the peak frequency was larger than that at the rated speed. During the acceleration process, it is advisable to minimize prolonged stays at these frequency points.

3.2. Influence of bearing stiffness

In order to constrain the vibration amplitude of the shaft, various measures were implemented, including modifying the support span and adjusting the bearing stiffness. This study conducted analysis on the influence of both the bearing support span and the bearing stiffness coefficient on the system dynamics.

The bearing stiffness coefficient was varied within the range of 5e4 to 5e8 N/m, encompassing the stiffness coefficient values of various elastic foil bearings. The unbalance response of the double-span shaft was calculated under different bearing stiffness conditions, as shown in Fig. 12. The unbalance excitation was at the center of the permanent magnet. The bearing stiffness coefficients were chosen as 5e4 N/m, 5e6 N/m and 5e8 N/m.

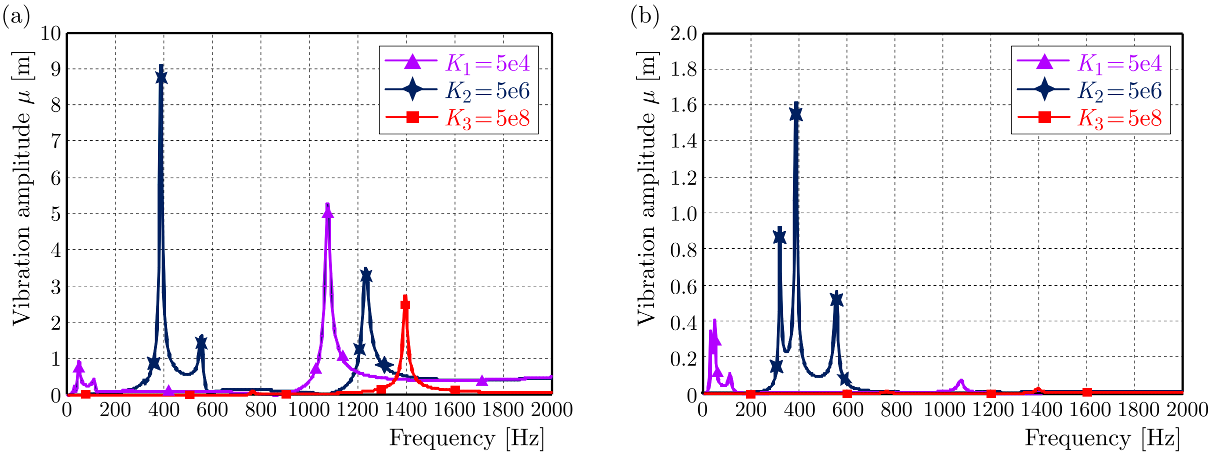


Fig. 12. Comparison of unbalance response under different bearing stiffness: (a) observation point 2, (b) observation point 3

Figure 12a depicts the unbalance response at observation point 2. The bearing stiffness had an impact on the unbalanced response. The maximum vibration amplitude approximately occured at the frequency of 400Hz, when the bearing stiffness was 5e6 N/m. The frequency range from 1000 Hz to 1600 Hz corresponded to the rigid body vibration. As the bearing stiffness increased, the vibration amplitude increased first and then decreased. In addition to the unbalance response, the selection of bearing stiffness needs to consider the stability of the shaft system.

In Fig. 12b, the unbalance response at observation point 3 is shown. At around 400 Hz, the maximum vibration amplitude corresponded to the frequency of 400 Hz with the bearing stiffness of 5e6 N/m. Comparing observation points 2 and 3, the unbalance response at observation point 3 was much smaller than that at observation point 2. Therefore, it cpuld be observed that the diaphragm coupling effectively isolated the impact of the unbalance excitation of the permanent magnet on the compressor-turbine rotor.

The insights gained from the analysis of rotor and bearing parameters hold significant importance, allowing the shaft system to select appropriate structural parameters based on the analysis results.

4. Mechanical experimental study of the shaft system

4.1. Dynamic balancing experiment

The dynamic balancing machine was depicted in Fig. 13. When the machine operates, rotation speed is measured by a photoelectric sensor. The driving power of the motor is 1.1 kW, and the maximum operation speed is 2800 r/min. The support parts are made of a non-magnetic material, which has no effect on the permanent magnet of the shaft.

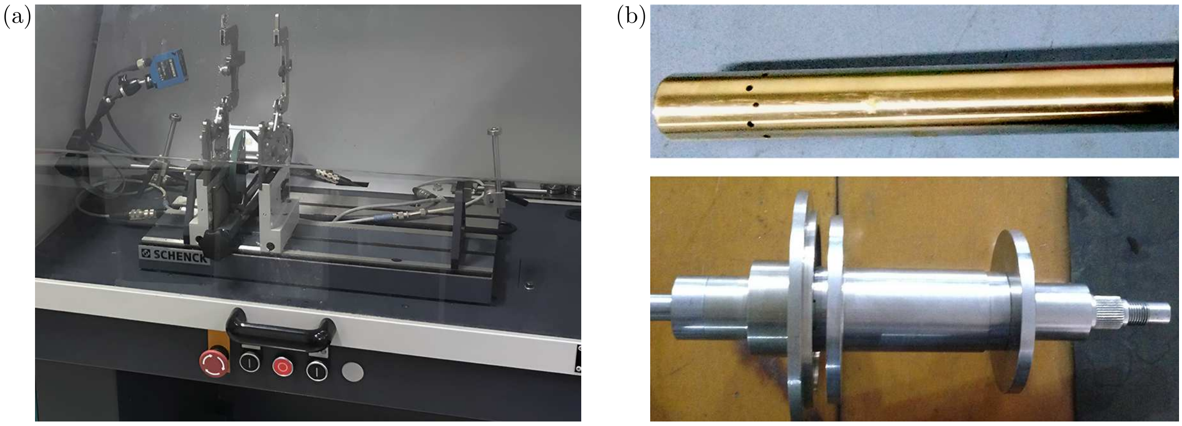


Fig. 13. Dynamic balancing experiment: (a) dynamic balancing machine, (b) motor rotor and compressor-turbine rotor

The dynamic balancing experiment of the shaft was conducted in this machine. The motor rotor and the compressor-turbine rotor were balanced separately in the rigid state because the diaphragm coupling could effectively isolate the unbalance response of the motor rotor and the compressor-turbine rotor.

According to the standard of ISO 1940 (2003), dynamic balancing results of the motor rotor and compressor-turbine rotor were obtained. Numerical values of the selected balance quality grade were $G1$.

4.2. High-speed dynamic experiment on the shaft system

The double-span multi-support shaft system in the micro gas turbine is illustrated in Fig. 14. The rated power of the micro gas turbine was 30 kW, and the rated speed was 96000 r/min. The shaft was propelled by a high-speed permanent magnet synchronous motor, and the operating speed was controlled within the range of 0-60000 r/min. This speed range encompassed the entire operational spectrum of the micro gas turbine, starting from 0 and reaching the self-sustaining speed or ignition speed. The experimental setup is depicted in Fig. 14.

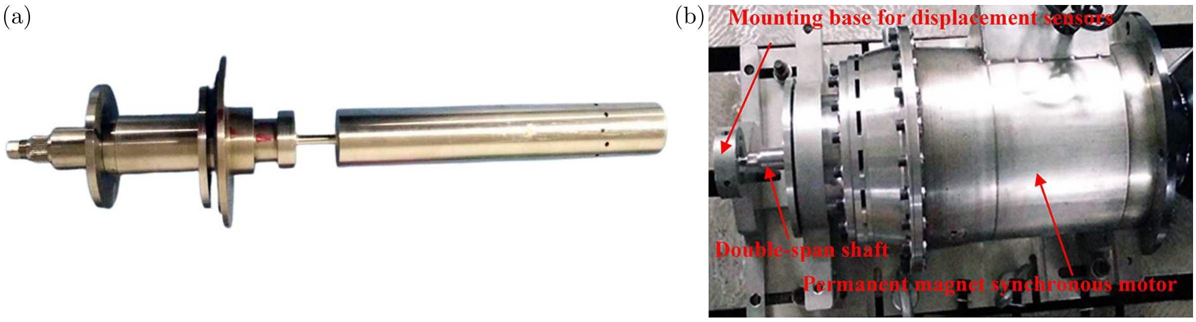


Fig. 14. Experimental device for the high-speed shaft with double spans and multiple supports: (a) double-span multi-support shaft, (b) experimental setup for the shaft

The double-span multi-support shaft was driven by the permanent magnet synchronous motor in the experimental device. The maximum experimental operating speed of the double-span multi-support shaft reached 60000 r/min. The vibration displacement of the shaft was measured by using eddy current displacement sensors.

Results of the vibration displacement for the double-span multi-support shaft at the experimental speed of 60000 r/min are presented in Fig. 15. In Fig. 15a, displacement amplitudes along the time history in vertical and horizontal directions are depicted, ranging from around 40 μm. The data of the vibration displacement was processed by the Fourier transform. Figure 15b illustrates the amplitude in the frequency domain, with the main frequency at 1000 Hz corresponding to the frequency caused by the rotor unbalance. The low-frequency at 248 Hz aligns with the rotor whirl frequency. Figure 15c shows the orbit of the shaft centerline at the speed of 60000 r/min. The experimental results validate capability of the shaft for vibration suppression and show its dynamic behavior at high speeds. The experimental findings indicate that the shaft operated effectively at the elevated speed of 60000 r/min.

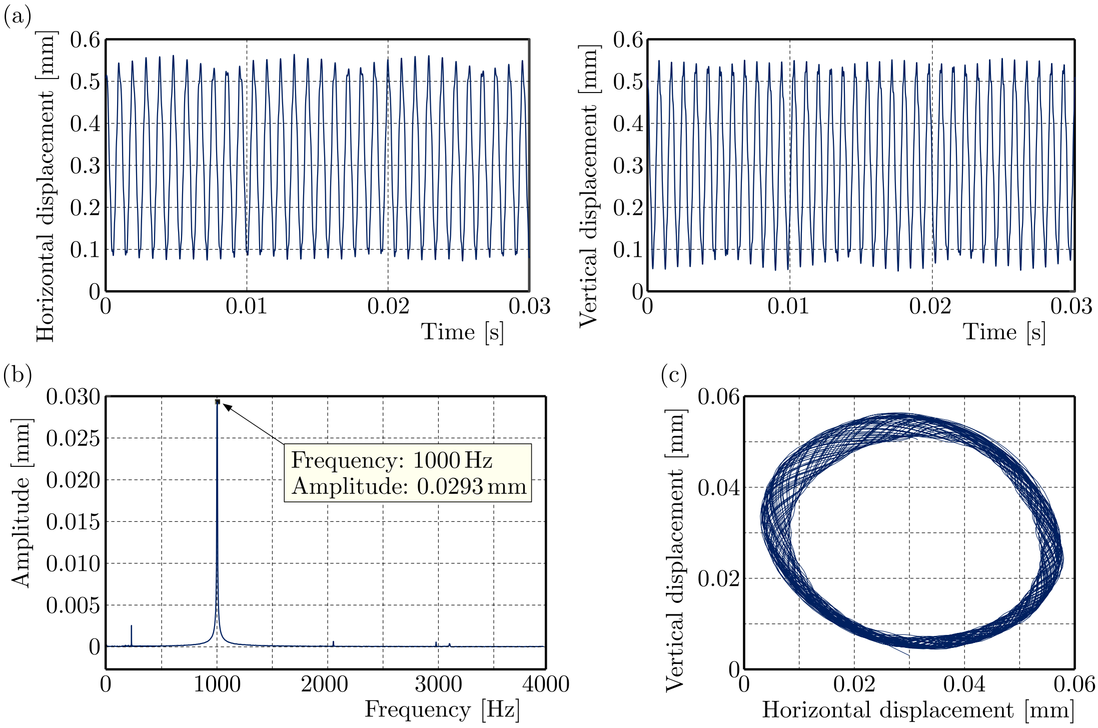


Fig. 15. Experimental vibration displacement of the shaft: (a) time domain diagram of the vibration displacement, (b) frequency of the vibration displacement, (c) shaft orbit diagram

5. Conclusion

This study conducted comprehensive dynamics analysis of the double-span multi-support shaft system in a micro gas turbine, leading to the following conclusions:

Critical Speed Analysis: The critical speed analysis was performed, providing numerical solutions for the natural frequency. The first-order bending critical speed of the shaft was found to be higher than the rated speed, with rotors considered as rigid bodies throughout the operational speed range. The shaft exhibited a sufficient safety margin between the rated speed and the bending critical speed.

Unbalanced Response Analysis: Unbalanced response calculations were conducted to predict vibration amplitudes of the shaft. The maximum unbalance response, located at the left end of the rotor, was approximately 35 μm. Within the rated speed range of 1600 Hz, vibration amplitudes near peak frequencies were larger than those at the rated speed, providing valuable guidance for dynamic balance considerations.

Effect of Bearing Stiffness: The variation range of the bearing stiffness coefficient, selected within the range of 5e4-5e8 N/m, had an effect on the unbalance response. The unbalance response increased first and then decreased with an increase in the bearing stiffness. The analysis provided insights for selecting appropriate bearing parameters for the shaft system.

Experimental Verification: Experimental results at the operational speed of 60000 r/min demonstrated a vibration displacement for the double-span multi-support shaft around 40 μm. The experimental findings closely aligned with the calculated results, confirming the effectiveness of the vibration suppression and dynamic behavior of the shaft at high speeds. In future experiments, the operational speed will be increased to 96000 r/min. The nonlinear dynamic stability of the double-span multi-support shaft will be analyzed. Furthermore, considering the unbalance response and bending vibration of the flexible coupling, nonlinear dynamics of the double-span multi-support shaft will be complicated.

In summary, this study contributes valuable insights into the dynamic characteristics of the double-span multi-support shaft system, offering practical guidance for design considerations, dynamic balance, and operational performance in micro gas turbines.

Acknowledgment

The research was supported by the Fundamental Research Funds for the Central Universities (No. 2022JBZY028), the Beijing Natural Science Foundation (No. 3244035), Hebei Natural Science Foundation (No. E2024105047), and Science and Technology Research and Development Plan of China National Railway Group Corporation Limited (No. L2022Z002).

References

1. Andrés L.S., Rohmer M., Park S., 2015, Failure of a test rig operating with pressurized gas bearings: a lesson on humility, *Proceedings of the ASME Turbo Expo 2015: Turbine Technical Conference and Exposition, Vol. 7A: Structures and Dynamics*, Montreal, Canada
2. Bo A., Giacomazzi E., Messina G., Di Nardo A., 2018, Analysis of a fuel flexible micro gas turbine combustor through numerical simulations, *Journal of Engineering for Gas Turbines and Power*, **140**, 12, 121504
3. Feng S., Geng H., Yu L., 2015, Rotordynamics analysis of a quill-shaft coupling-rotor-bearing system, *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **229**, 8, 1385-1398
4. Feng S., Yang B., Tian J., Yu L., 2021, A theoretical solution of the axial stiffness of a cylindrical shell with a circular plate, *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **235**, 18, 3563-3569
5. Guan H.Q., Feng K., Cao Y.L., Huang M., Wu Y.H., Guo Z.Y., 2020, Experimental and theoretical investigation of rotordynamic characteristics of a rigid rotor supported by an active bump-type foil bearing, *Journal of Sound and Vibration*, **466**, 115049
6. Heshmat H., Walton II J.F., 2016, Lubricant free foil bearings pave way to highly efficient and reliable flywheel energy storage system, *Proceedings of the ASME 2016 Tenth International Conference on Energy Sustainability*, 1-12
7. Heshmat H., Walton II J.F., Hunsberger A., 2014, Oil free 8 kW high-speed and high specific power turbogenerator, *Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, Vol. 1B: Marine; Microturbines, Turbochargers and Small Turbomachines; Steam Turbines*, Düsseldorf, Germany, 1-16

8. Heshmat H., Walton II J.F., Nicholson B.D., 2018, Ultra-high temperature compliant foil bearings: the journey to 870°C and application in gas turbine engines – experiment, *Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition, Vol. 7B: Structures and Dynamics*, Oslo, Norway
9. ISO 1940-1, 2003, Mechanical vibration-balance quality requirements for rotors in a constant (rigid) state, International Organization for Standardization
10. Kim H., Nutakor C., Singh S., Jaatinen-Värri A., Nerg J., *et al.*, 2023, Design process and simulations for the rotor system of a high-efficiency 22 kW micro-gas-turbine range extender for electric vehicles, *Mechanism and Machine Theory*, **183**, 105230
11. Koo B., Kim C.H., Lee J.S., Kim T.H., Sim K., 2015, Analysis and experimental test of a 200 kW oil-free micro gas turbine simulator, *Proceedings of the 9th IFToMM International Conference on Rotor Dynamics*, 307-319
12. Lorenzen H., Niedermann E., Wattinger W., 1989, Solid couplings with flexible intermediate shafts for high speed turbocompressor trains, *Proceedings of the Eighteenth Turbomachinery Symposium*, Texas A & M University, 101-110
13. Othman N.F., Boosroh M.H., 2016, Effect of H_2 and CO contents in syngas during combustion using Micro Gas Turbine, *Proceedings of IOP Conference Series: Earth and Environmental Science*, **32**, 012037
14. Ramaglia A.D., Ruedel U., Stefanis V., Florjancic S., 2018, *Ansaldo Energia Gas Turbine Technology Developments, Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition, Vol. 3: Coal, Biomass, and Alternative Fuels; Cycle Innovations; Electric Power; Industrial and Cogeneration; Organic Rankine Cycle Power Systems*, Oslo, Norway, American Society of Mechanical Engineers, New York
15. Schmied J., Pradetto J., 1994, Rotor dynamic behaviour of a high-speed oil-free motor compressor with a rigid coupling supported on four radial magnetic bearings, *Proceedings of the Second International Symposium on Magnetic Suspension Technology*, 557-573
16. Singh S., Petrov I., Pyrhönen J., Sergeant P., 2022, Conceptual design of high-speed permanent-magnet generator for a micro gas turbine, *2022 International Conference on Electrical Machines (ICEM)*, Valencia, Spain, 1696-1702
17. Villarroel-Schneider J., Malmquist A., Araoz J.A., Martí-Herrero J., Martin A., 2019, Performance analysis of a small-scale biogas-based trigeneration plant: An absorption refrigeration system integrated to an externally fired microturbine, *Energies*, **12**, 20, 3830
18. Vishwanathan G., Sculley J.P., Fischer A., Zhao J.C., 2018, Techno-economic analysis of high-efficiency natural-gas generators for residential combined heat and power, *Applied Energy*, **226**, 1064-1075
19. Walton II J.F., Heshmat H., Tomaszewski M.J., 2012, Power loss in high-speed micro turbomachinery: an experimental study, *Proceedings of the ASME Turbo Expo: Power for Land, Sea and Air*, 1-10
20. Wołowicz M., Kolasiński P., Badyda K., 2021, Modern small and microcogeneration systems-a review, *Energies*, **14**, 3, 785
21. Zheng M., Huang W., Gao C., 2020, Rotor stress and dynamics analysis of a high-speed permanent magnet machine for a micro gas turbine considering multiphysics factors, *IEEE Access*, **8**, 152523-152531
22. Żywica G., Bagiński P., 2019, Investigation of unconventional bearing systems for microturbines, *Advances in Mechanism and Machine Science, Proceedings of the 15th IFToMM World Congress on Mechanism and Machine Science*, Krakow, Poland, **73**, 1, 3439-3448

Manuscript received February 22, 2024; accepted for publication November 8, 2024

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 131-149, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195727

YAW FEEDBACK CONTROL OF ACTIVE STEERING VEHICLE BASED ON DIFFERENTIAL FLATNESS THEORY

Haiyan Qiang

Logistics Engineering College, Shanghai Maritime University, Shanghai, China and Key Laboratory of Railway Industry of Maglev Technology, Tongji University, Shanghai, China

Cheng Xiao, Hengyue Huang, Yichen Hai

Logistics Engineering College, Shanghai Maritime University, Shanghai, China

Yougang Sun

National Maglev Transportation Engineering R&D Center, Tongji University, Shanghai, China
e-mail: 1989yoga@tongji.edu.cn

In order to solve the problems of nonlinearity, underactuation and insufficient lateral stability of an active steering vehicle (ASV) in trajectory tracking tasks, a yaw feedback control strategy based on differential flatness theory is proposed in this paper. Firstly, the vehicle integrated monorail model is established, and the vehicle model is linearized by small angle approximation. Secondly, a suitable flat output is found to convert a complex vehicle model into a full drive system, and the flatness of the linear model is proved. Then, an equivalent form of the vehicle model is constructed based on the flat output and its derivatives, and a feedback controller based on the differential flat theory is designed to complete the trajectory tracking control through active steering and longitudinal motion. Finally, an ASV control simulation model is built in MATLAB/Simulink, and the simulation results show the effectiveness of the proposed control strategy under different maneuvering conditions.

Keywords: differential flatness; trajectory tracking; feedback control; underactuated system; lateral stability

1. Introduction

Active steering vehicles (ASVs) are poised to play a major role in the future of intelligent transportation. They have the potential to improve vehicle stability, enhance traffic efficiency and safety, and promote low-carbon transportation. With the integration of computer vision technology, sensor intelligent perception technology, Internet of Things (IoT) technology, and new control strategies (Ortiz *et al.*, 2023; Sun *et al.*, 2023; Luo *et al.*, 2022), the vehicle yaw control has made significant progress and has become a key research focus in modern land transportation. ASV yaw control can automatically track reference trajectories under varying speeds, complex road conditions, and uncertain disturbances, guaranteeing enhanced tracking accuracy and lateral stability.

Importantly, the differential flatness approach can be widely applied in the transportation sector, as shown in Fig. 1, which includes self-driving passenger cars, self-driving buses, intelligent express delivery cars, port container driverless transport vehicles, and engineering transport vehicles, among others.

In recent years, the intelligent vehicle trajectory tracking control has attracted attention of researchers worldwidely. Most of these studies focus on tracking accuracy, lateral stability, and control methods, yielding substantial results (Yu *et al.*, 2021; Sun *et al.*, 2024; Yang *et al.*, 2021;

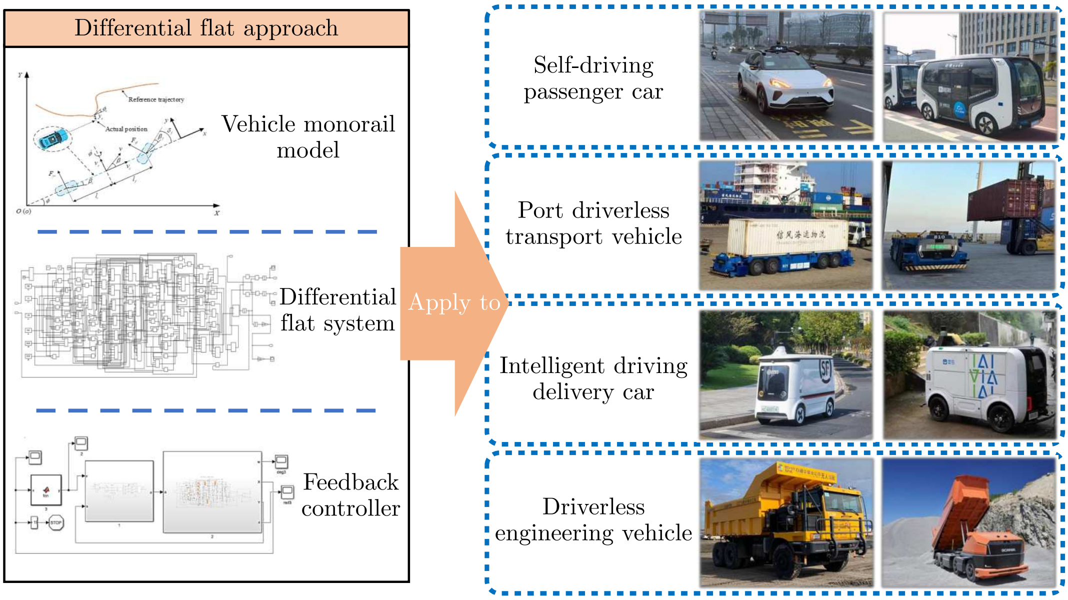


Fig. 1. Possible application areas for the differential flatness approach

Sun *et al.*, 2023). Wang and Sun (2023) designed a linear quadratic regulator combining feedforward and feedback, adopting a fuzzy control algorithm and cosine similarity updating mechanism to quickly adjust weight values, thereby improving vehicle tracking accuracy and computational efficiency. Rokonuzzaman (2021) developed a dynamic model predictive controller based on neural networks combined with the big data system of intelligent vehicles, enhancing the accuracy of vehicle models, and tracking performance. Bai (2019) introduced a nonlinear model predictive controller for mining vehicles, which effectively ensured stable and accurate tracking at high longitudinal speeds. Mata (2019) proposed a tube-based robust model predictive control method for vehicle path tracking, considering dynamic differences between the actual vehicle and the mathematical model and strict constraints on control signals and lateral errors to ensure tracking accuracy and driving comfort. Kang (2022) suggested an improved active disturbance rejection control method, applying a new continuous nonlinear function to the extended state observer, optimally allocating yaw moment to the four wheels to achieve differential control, enhancing tracking effect and anti-disturbance robustness. Guerrero (2023) designed a generalized super-twisting algorithm controller to address the disturbance problem of underwater vehicles, conducting stability analysis under different disturbances, with simulation results showing good performance.

It is evident that considering both the lateral displacement and yaw angle response in trajectory tracking research is crucial for improving vehicle tracking performance. However, many studies overlook the underactuated characters of vehicle systems, with few focusing on underactuated vehicle systems. Most research on underactuation are focused on the field of foot robots, robotic arms, aircraft, and surface vehicles, etc. To achieve more accurate control of the lateral displacement and yaw angle response, differential flatness theory and feedback control can be applied to address the underactuated vehicle system problem.

Fliess (1995) introduced a system equivalent to a linear system through special feedback, identified as a differentially flat system. Using crane motion planning as an example, it was demonstrated that the input of a differentially flat system could be expressed through a combination of the output and its derivatives, which significantly reduced complex integral operations. This illustrates those issues of uncontrollability, nonlinearity, and underactuation can be resolved using the properties of differential flatness theory. Huang (2019) proposed an active disturbance rejection control method based on differential flatness for nonlinear systems with periodic and aperiodic characteristics, identifying total system disturbances, selecting appropriate observers, and employing the fruit fly optimization algorithm to determine controller parameters. The effectiveness of the proposed method was validated through simulation. Elmi (2013) established a three-degree-of-freedom vehicle model, designed a robust linear quadratic regulator, and performed simulation analysis across a wide range of vehicle speeds and tire characteristics, demonstrating that the proposed controller ensured robust stability under varying conditions. Aschemann *et al.* (2008) combined differential flatness and sliding mode control to enhance error dynamics stability for unmodeled aerodynamic mechanisms. Differential flatness theory can also resolve system underactuation issues, as illustrated by Zhuang (2010) who applied differential flatness to achieve optimal motion planning for non-axisymmetrical spacecraft.

This paper focuses on the yaw control strategy of single-input, multiple-output underactuated vehicles in trajectory tracking tasks. The main contributions are summarized as follows:

- The differential flatness of an integrated monorail vehicle model is demonstrated, and the assumed differential flatness output is validated.
- Based on the differential flatness and flat output, the equivalent system form is derived, representing the system with the differential flat output and its derivatives, thus solving the underactuation problem.
- A feedback controller based on differential flatness theory is designed, providing high control accuracy and robust performance.

The organizational structure of this paper is as follows: In Section 2, an integrated monorail model of ASV is built, combining vehicle kinematics and dynamics. In Section 3, the validity of the assumed flat output and the differential flatness of the system are proved. In Section 4, the original system is transformed into an equivalent system represented by the differential flat output and its derivatives, and a feedback controller based on differential flatness theory is designed. In Section 5, simulations on ASV in Simulink demonstrate the superiority of the proposed strategy. Section 6 concludes the paper.

2. Vehicle integrated monorail model

ASV is an underactuated integrated system with high complexity, strong nonlinearity, and uncertainties. In trajectory tracking tasks, its dynamic model is characterized by numerous constraints and high nonlinearity, making accurate modeling challenging (Wang *et al.*, 2022). Therefore, in this Section, the vehicle kinematics and dynamics are combined to establish an integrated monorail model of an active steering vehicle.

To align with practical scenarios and facilitate subsequent research, the important characteristics of the lateral and yaw direction of the vehicle are the focus of this research. The following assumptions are made during the model establishment process:

- The influence of air resistance, tire changes, and the suspension system is ignored.
- Vertical, tumbling, and pitching movements are disregarded.
- The tire slip angle is small, and the tire lateral force is proportional to the slip angle.
- The road adhesion state and vehicle driving conditions are optimal.
- Active steering is applied to the front wheels of the vehicle.

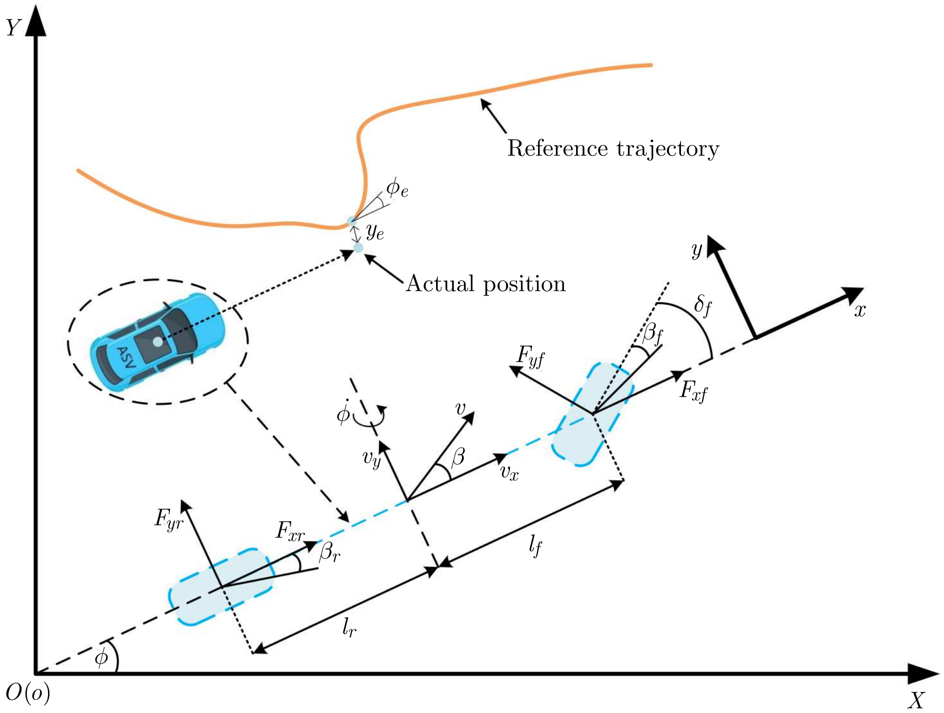


Fig. 2. Vehicle monorail model

Based on the assumptions above, a vehicle model, which is simplified to a monorail model, by considering the similar motion and dynamic characteristics exhibited for both sides of the vehicle, is established, as shown in Fig. 2.

In Fig. 2, XOY denotes the global coordinate system of the earth, xoy is the vehicle centroid coordinate system, y_e is the lateral deviation, ϕ_e is the deviation of yaw angle. The kinematic model considering the lateral and yaw of the vehicle can be expressed as follows

$$\Delta\dot{Y} = v_x \sin\phi + v_y \cos\phi \qquad \dot{\phi} = w \tag{2.1}$$

where v_x and v_y represent longitudinal velocity and transverse velocity, respectively, ϕ indicates the vehicle yaw angle, w indicates the yaw angular speed of the vehicle.

According to Fig. 2, the dynamic model considering the lateral and yaw of the vehicle can be expressed as follows

$$F_{yf}\cos\delta_f + F_{yr} = m(\dot{v}_y + v_x\dot{\phi}) \qquad l_f F_{yf}\cos\delta_f - l_r F_{yr} = I\ddot{\phi} \tag{2.2}$$

where F_{yf} and F_{yr}, respectively, represent the lateral force of the road on the front and rear wheels of the vehicle, l_f and l_r represent the distance from the center of the front and rear wheels to the center of mass, respectively. δ_f represents the active steering angle of the vehicle, m represents vehicle mass, I represents the yaw moment of inertia of the vehicle.

Considering the lateral force generated by the interactions among the front and rear wheels and the ground, combined with the assumptions made in the process of vehicle model building, it can be considered that the lateral force is proportional to the side yaw angle, so the lateral force and side yaw angle can be expressed as

$$\begin{aligned} &F_{yf} = -C_{kf}\beta_f \qquad && F_{yr} = -C_{kr}\beta_r \\ &\beta_f \approx (v_y + wl_f - \delta_f v_x)/v_x \qquad && \beta_r \approx (v_y - wl_r)/v_x \end{aligned} \tag{2.3}$$

where C_{kf} and C_{kr} represent the lateral stiffness of the front and rear wheels, respectively, β_f and β_r represent the front and rear wheel side angles.

Generally, a small angle approximation is carried out for the vehicle under normal driving conditions (Wang *et al.*, 2014), namely: $\cos\delta_f \approx 1$, $\sin\phi \approx \phi$, $\cos\phi \approx 1$. Equations (2.1)–(2.3) are combined to obtain the vehicle integrated monorail model equation

$$\begin{aligned}
&\Delta\dot{Y} = v_x\phi + v_y \qquad \dot{\phi} = w\\
&\dot{v}_y = -v_x\dot{\phi} + 2\big[C_{kf}(\delta_f v_x - wl_f - v_y) + C_{kr}(wl_r - v_y)\big]/v_x m\\
&\dot{w} = 2\big[l_f C_{kf}(\delta_f v_x - wl_f - v_y) - l_r C_{kr}(wl_r - v_y)\big]/v_x I
\end{aligned} \tag{2.4}$$

In order to facilitate subsequent derivation and research, the vehicle model equation is simplified, and the relevant parameter combinations are replaced by simple equivalent characters, and the simplified vehicle model equation can be obtained

$$\begin{aligned}
&\Delta\dot{Y} = g\phi + v_y \qquad \dot{\phi} = w\\
&\dot{v}_y = eh + (av_y + bw - g^2 w)/g \qquad \dot{w} = fh + (cv_y + dw)/g
\end{aligned} \tag{2.5}$$

where

$$\begin{aligned}
&a = -2(C_{kf} + C_{kr})/m \qquad e = 2C_{kf}/m\\
&b = 2(C_{kr}l_r - C_{kf}l_f)/m \qquad f = 2C_{kf}l_f/I\\
&c = 2(C_{kr}l_r - C_{kf}l_f)/I \qquad g = v_x\\
&d = -2(C_{kr}l_r + C_{kf}l_f)/I \qquad h = \delta_f
\end{aligned}$$

Assuming $x_1 = \Delta Y$, $x_2 = v_y$, $x_3 = \phi$, $x_4 = w$, and $x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^{\mathrm{T}}$ is regarded as the system state variable, $h = \delta_f$ as the control input, and y is defined as the output after system operation, then the vehicle model equation above can be converted into the state space equation

$$\dot{x} = \mathbf{A}_0 x + \mathbf{B}_0 h \qquad y = \mathbf{C}_0 x \qquad x \in R^4 \tag{2.6}$$

where

$$\mathbf{A}_0 = \begin{bmatrix} 0 & 1 & g & 0\\ 0 & a/g & 0 & (b-g^2)/g\\ 0 & 0 & 0 & 1\\ 0 & c/g & 0 & d/g \end{bmatrix} \qquad \mathbf{B}_0 = \begin{bmatrix} 0\\ e\\ 0\\ f \end{bmatrix} \qquad \mathbf{C}_0 = \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3. Differential flat system construction and output

3.1. Differential flatness theory

For a system $\dot{x} = \kappa_1(x, h)$, κ_1 represents a smooth continuous function, $x \in R^{n_1}$, $h \in R^{n_2}$. According to the property of differential flatness, a set of flat output $P \in R^{n_2}$ can be found to represent a combination of system state variables, control variables and their derivatives, and both the system state variables and control variables can be represented by the flat output and its derivatives (Menhour *et al.*, 2014). The system can be regarded as a system based on differential flatness theory.

The general form of a flat output is as follows

$$P = \boldsymbol{\xi}(x, h, \dot{h}, ..., h^{(i)}) \tag{3.1}$$

where $\boldsymbol{\xi}$ represents a smooth vector function.

The general form of the system state variable and control variable composed of the flat output and its derivatives is as follows

$$x = \boldsymbol{\psi}_x(P, \dot{P}, ..., P^{(i)}) \qquad h = \boldsymbol{\psi}_h(P, \dot{P}, ..., P^{(i)}) \quad x \in R^{n_1} \quad h \in R^{n_2} \quad P \in R^{n_2} \tag{3.2}$$

where $\boldsymbol{\psi}_x$, $\boldsymbol{\psi}_h$ both represent smooth vector functions.

3.2. Differential flatness output and flatness proof

If a linear system is controllable, it has differential flatness and can find the corresponding flat output. The combination of the flat output of a controllable linear system, the inverse of the Kalman controllability discriminant matrix and the system state variable can represent the corresponding flat output (Sira-Ramirez *et al.*, 2004; Xia *et al.*, 2016). It can be described more accurately by a linear system

$$\dot{x} = Ax + Bh \qquad y = Cx \qquad x \in R^n \tag{3.3}$$

Based on the above system, a general form of the flat output can be found

$$\mathbf{P}_0 = \begin{bmatrix} 0 & 0 & 0 & ... & 1 \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & ... & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}^{-1} x \tag{3.4}$$

where matrices $\mathbf{A}_0$ and $\mathbf{B}_0$ are relatively fixed values derived from the model, the system state variable x changes with the input value, so the flat output $\mathbf{P}_0$ mainly depends on the state variable of the system.

For active steering vehicle system (2.6), according to the controllability discrimination theory, its controllability matrix is shown in (3.5). By using data analysis software to calculate rank$(\mathbf{K}_0) = 4$, that is: matrix $\mathbf{K}_0$ is in a state of full rank. Therefore, system (2.6) is judged to be controllable. The matrix $\mathbf{K}_0$ looks like this

$$\begin{aligned} \mathbf{K}_0 &= \begin{bmatrix} \mathbf{B}_0 & \mathbf{A}_0\mathbf{B}_0 & \mathbf{A}_0^2\mathbf{B}_0 & \mathbf{A}_0^3\mathbf{B}_0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & e & (ae+bf)g^{-1} & (ea^2+ebc+fab+fbd)g^{-2}-fa \\ e & (ae+fb)g^{-1}-fg & (ea^2+eb+fab-bd)g^{-2}-ec-fa+d & \Delta_{24} \\ 0 & f & (ce+df)g^{-1} & (fd^2+fcb+eac+ecd)g^{-2}-cf \\ f & (ce+df)g^{-1} & (fd^2+bcf+ace+cde)g^{-2}-cf & \Delta_{44} \end{bmatrix} \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \Delta_{24} &= -e(ca+d)(g^{-1}-bg^{-3}) - a(a^2g^{-3} - c(g^{-1}-bg^{-3})) \\ &\quad - f((a^2g^{-2} - c + bcg^{-2})(g - bg^{-1}) + (a+1)dg^{-1} - (a+1)bdg^{-3}) \\ \Delta_{44} &= ce((bc+d^2)g^{-3} - cg^{-1}) + ae(a+d)cg^{-3} \\ &\quad - ef((a+d)cg^{-3}(1-bg^{-2}) - (bcd+d^3)g^{-4} + cdg^{-2}) \end{aligned}$$

When an active steering vehicle performs a trajectory tracking task, its system outputs are ΔY and ϕ, which are related to the value of matrix $\mathbf{C}_0$. It can be seen from (3.4) that the flat output is not affected by matrix $\mathbf{C}_0$, and the differential flat system is relatively independent, so the differential flat theory can be used here to analyze the system. In order to make the subsequent equations and derivations more concise, according to subsequent analysis and derivation, $ce^2 + f^2(g^2 - b) + (d - a)ef$ is set to G and $0.001(ce - af)$ is set to u in advance, and u is set as a constant which is not 0. Obviously, if $\mathbf{P}_0$ is valid, then $u\mathbf{P}_0$ can also be expressed as a set of flat outputs, so the assumed flat outputs are

$$P = u\mathbf{P}_0 = 0.001u\,[\,0 \;\; 0 \;\; 0 \;\; 1\,]\,\mathbf{K}_0^{-1}x \tag{3.6}$$

After expansion

$$P = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 \tag{3.7}$$

where

$$\begin{aligned}
&\alpha_1 = 0.001 \qquad \alpha_2 = 0.001fg(bf^2 - ce^2 + aef - def)/(afG - ceG)\\
&\alpha_3 = 0.001(bf - de)/(af - ce)\\
&\alpha_4 = 0.001eg(ce^2 - aef - bf^2 + def)/(afG - ceG)
\end{aligned}$$

To get the differential flat output of the state variable, take the derivative of P

$$\begin{aligned}
\dot{P} &= \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 h\\
\ddot{P} &= \eta_1 x_1 + \eta_2 x_2 + \eta_3 x_3 + \eta_4 x_4 + \eta_5 h\\
P^{(3)} &= \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4 + \lambda_5 h
\end{aligned} \tag{3.8}$$

where

$$\begin{aligned}
&\beta_1 = \beta_5 = 0 \qquad \beta_3 = 0.001g\\
&\beta_2 = 0.001f^2g^2/G \qquad \beta_4 = -0.001efg^2/G\\
&\eta_1 = \eta_3 = \eta_5 = 0 \qquad \lambda_1 = \lambda_3 = \lambda_5 = 0\\
&\eta_2 = 0.001fg(af - ce)/G \qquad \lambda_2 = 0.001(af - ce)^2/G\\
&\eta_4 = 0.001eg(ce - af)/G \qquad \lambda_4 = 0.001(ce - af)(fg^2 - bf + de)/G
\end{aligned}$$

From (3.7) we know that $\beta_5 = \eta_5 = \lambda_5 = 0$, so the state variable x can be represented by a combination of the flat output P and its derivatives $\dot{P}$, $\ddot{P}$, $P^{(3)}$, and by combining (3.6) and (3.7), we get

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \gamma_1 P + \gamma_2 \dot{P} + \gamma_3 \ddot{P} + \gamma_4 P^{(3)} \\ \mu_1 P + \mu_2 \dot{P} + \mu_3 \ddot{P} + \mu_4 P^{(3)} \\ \sigma_1 P + \sigma_2 \dot{P} + \sigma_3 \ddot{P} + \sigma_4 P^{(3)} \\ \tau_1 P + \tau_2 \dot{P} + \tau_3 \ddot{P} + \tau_4 P^{(3)} \end{bmatrix} \tag{3.9}$$

where

$$\begin{aligned}
&\gamma_1 = 1000 \qquad \gamma_3 = 1000e(ce - af)\\
&\gamma_2 = 1000(abf^2 - cde^2 + adef + bcef)/g \qquad \gamma_4 = 0\\
&\sigma_1 = \sigma_4 = 0 \qquad \sigma_2 = 1000/g \qquad \sigma_3 = 1000f/(ce - af)\\
&\mu_1 = \mu_2 = 0 \qquad \mu_3 = 1000(fg^2 - bf + de)/(afgG^2 - cegG^2)\\
&\mu_4 = 1000e(ce - af)/G^2\\
&\tau_1 = \tau_2 = 0 \qquad \tau_3 = 1000/(gG^2) \qquad \tau_4 = 1000f/(ceG^2 - afG^2)
\end{aligned}$$

Since the coefficients controlling the input amount h in $\dot{P}$, $\ddot{P}$, $P^{(3)}$ are all 0, the output expression of h cannot be flat, so it is necessary to obtain a higher derivative. From system (2.6), $\kappa_2(x, \dot{x}) = h$ can be obtained, and κ_2 represents a smooth continuous function. At this time, the fourth derivative of P is obtained

$$P^{(4)} = \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_3 x_3 + \varepsilon_4 x_4 + \varepsilon_5 h \tag{3.10}$$

where

$$\varepsilon_1=\varepsilon_3=0 \qquad \varepsilon_2=0.001(ce-af)(ace-fa^2+cfg^2-bcf+ced)/(gG)$$
$$\varepsilon_4=0.001(ce-af)(ed^2-abf+bce-bdf+afg^2-ceg^2+dfg^2)/(gG)$$
$$\varepsilon_5=0.001(ce-af)$$

From (3.10) we know that $\varepsilon_5 \neq 0$, so the control input h can be represented by the flat output and its derivatives, and by combining (3.7), (3.8), (3.9) and (3.10), we can get

$$h=\frac{(a+d)P^{(3)}}{gGu}-\frac{b\ddot{P}}{u}-\frac{(ad-bc)\ddot{P}}{g^2G^2u}-P^{(4)} \tag{3.11}$$

It is proved by (3.5), (3.9) and (3.11) that the differential flatness of system (2.6) is established and that the assumed output in (3.6) is indeed flat, that is, the state variables and control input variables of the system can be represented by the flat output and its finite derivatives, and the system has flatness.

4. Yaw feedback control strategy based on differential flatness theory

The equivalent system of (2.6) is established based on the differential flatness theory, assuming $\chi=[\ \chi_1\ \ \chi_2\ \ \chi_3\ \ \chi_4\ ]^{\mathrm{T}}$ and $\chi_1=P$, $\chi_2=\dot{P}$, $\chi_3=\ddot{P}$, $\chi_4=P^{(3)}$, the equivalent system can be obtained as follows

$$\dot{\chi}=\mathbf{A}_1\chi+\mathbf{B}_1h_p \qquad y=\mathbf{C}_1\chi \tag{4.1}$$

where

$$\mathbf{A}_1=\begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & (bc-ad-bg^2G^2)/(g^2G^2) & (a+d)/(gG)\end{bmatrix}$$

$$\mathbf{B}_1=\begin{bmatrix} 0\\ 0\\ 0\\ (af-ce)/1000\end{bmatrix} \qquad \mathbf{C}_1=\begin{bmatrix} 1 & 0 & 0 & 0\end{bmatrix}$$

It can be assumed from (4.1) that $\kappa_3=\dot{\chi}_4-0.001(af-ce)h$, so that the objective function can be tracked by controlling the differential flat output function to solve the system underdrive problem of ASV in trajectory tracking control.

In the feedback controller on the differential flatness (FCDF) control strategy, according to the target state variable x_p and the target control input variable h_p obtained from the differential flatness theory, the deviation system with $z_1=x_1-x_{p1}$, $z_2=x_2-x_{p2}$, $z_3=x_3-x_{p3}$, $z_4=x_4-x_{p4}$ as the state variable and $h_z=h-h_p$ as the control input variable is constructed.

$$x_p=\begin{bmatrix} x_{p1}\\ x_{p2}\\ x_{p3}\\ x_{p4}\end{bmatrix}=\begin{bmatrix}\kappa_4(\chi_1,\chi_2,\chi_3,\chi_4)\\ \kappa_5(\chi_1,\chi_2,\chi_3,\chi_4)\\ \kappa_6(\chi_1,\chi_2,\chi_3,\chi_4)\\ \kappa_7(\chi_1,\chi_2,\chi_3,\chi_4)\end{bmatrix} \qquad h_p=\kappa_8(\chi_1,\chi_2,\chi_3,\chi_4,\dot{\chi}_4)$$

where κ_i $(i=4,5,6,7,8)$ is a smooth vector function.

Considering the deviation between the target value and the actual value of the state variable, the deviation of each state variable is substituted into the expression of the system state variable, and the deviation expression of the vehicle system state variable can be obtained by combining system (2.6)

$$\begin{aligned} &\dot{z}_1 = gz_3 + z_2 \qquad &&\dot{z}_2 = (a/g)z_2 + (b - g^2)z_4/g + eh_z \\ &\dot{z}_3 = z_4 \qquad &&\dot{z}_4 = (c/g)z_2 + (d/g)z_4 + fh_z \end{aligned} \tag{4.2}$$

The forward Euler method is used for discretization

$$\begin{aligned} z_1(k+1) &= z_1(k) + T(gz_3(k) + z_2(k)) \\ z_2(k+1) &= z_2(k) + T((a/g)z_2(k) + (b - g^2)z_4(k)/g + eh_z(k)) \\ z_3(k+1) &= z_3(k) + Tz_4(k) \\ z_4(k+1) &= z_4(k) + T((c/g)z_2(k) + (d/g)z_4(k) + fh_z(k)) \end{aligned} \tag{4.3}$$

where T indicates the sampling time, which is 0.05 seconds in this numerical study, and k indicates the sampling order.

Assuming $z(k) = \begin{bmatrix} z_1(k) & z_2(k) & z_3(k) & z_4(k) \end{bmatrix}^{\mathrm{T}}$, the new discrete time-varying deviation system obtained by combining (4.2) and (4.3)

$$z(k+1) = \mathbf{A}_2 z(k) + \mathbf{B}_2 h_z(k) \tag{4.4}$$

where

$$\mathbf{A}_2 = \begin{bmatrix} 1 & T & gT & 0 \\ 0 & 1 + (a/g)T & 0 & (b - g^2)T/g \\ 0 & 0 & 1 & T \\ 0 & (c/g)T & 0 & 1 + (d/g)T \end{bmatrix} \qquad \mathbf{B}_2 = \begin{bmatrix} 0 \\ eT \\ 0 \\ fT \end{bmatrix}$$

A state feedback controller $h_z(k) = R(k)z(k)$ can be designed to stabilize the control deviation system approaching zero, then the corresponding system can be expressed as

$$z(k+1) = [\mathbf{A}_2(k, k+1) + \mathbf{B}_2(k, k+1)R(k)]\, z(k) \tag{4.5}$$

In the infinite time domain, the performance index of feedback control based on differential flatness is designed as H, satisfying the following expression

$$H_\infty(k) = \sum_{k=0}^{\infty} z^{\mathrm{T}}(k)\mathbf{M}z(k) + \sum_{k=0}^{\infty} h_z^{\mathrm{T}}(k)\mathbf{N}h_z(k) \tag{4.6}$$

where $\mathbf{M} = \operatorname{diag}(m_1, m_2, m_3, m_4)$, $\mathbf{N} = [n]$, both are weight matrices.

According to the Riccati algebraic equation, where Q is a positive definite solution of the equation and satisfies the following expression

$$\mathbf{M} + \mathbf{A}_1^{\mathrm{T}}Q\mathbf{A}_1 - \mathbf{A}_1^{\mathrm{T}}Q\mathbf{B}_1(\mathbf{N} + \mathbf{B}_1^{\mathrm{T}}Q\mathbf{B}_1)^{-1}\mathbf{B}_1^{\mathrm{T}}Q\mathbf{A}_1 - Q = 0 \tag{4.7}$$

State feedback gain R satisfies the following expression

$$R = -(\mathbf{N} + \mathbf{B}_1^{\mathrm{T}}Q\mathbf{B}_1)\mathbf{B}_1^{\mathrm{T}}Q\mathbf{A}_1 \tag{4.8}$$

According to the vehicle monorail model, the reference active steering angle is $\delta_r = \arctan((l_f + l_r)(1 + \dot{y}^2)^{3/2} \big/ |\ddot{y}|)$, and the final control law of the vehicle actual active steering angle under the FCDF control strategy is

$$\delta_f = \delta_r + h_z(k) \tag{4.9}$$

Table 1. Vehicle parameters

Symbol	Meaning	Numerical value
m	Vehicle mass	1280 kg
l_f	Distance from center of mass to center of front wheel	1.2 m
l_r	Distance from center of mass to center of rear wheel	1.26 m
C_{kf}	Front wheel side stiffness	61000 N/rad
C_{kr}	Rear wheel side stiffness	61000 N/rad
I	Yaw moment of inertia	1630 kg·m^2

5. Simulation and analysis

To verify the effectiveness of the yaw feedback control strategy based on differential flatness, a vehicle model using differential flatness theory was established in MATLAB/Simulink. The parameters of the test vehicle are shown in Table 1.

During the path tracking process, it is assumed that the longitudinal speed of the vehicle is constant at $v_x = 50\,\text{km/h}$. In order to simulate the trajectory of a real road vehicle, the curvature of the curve is usually between 0.005 and 0.0001 according to highway engineering technical standards, so the smooth trajectory design helps the vehicle maintain a stable dynamic response when it is tested. Lane change and overtaking condition tests were conducted. The track reference value and yaw angle reference value for these conditions are shown in Figs. 3 and 4, respectively.

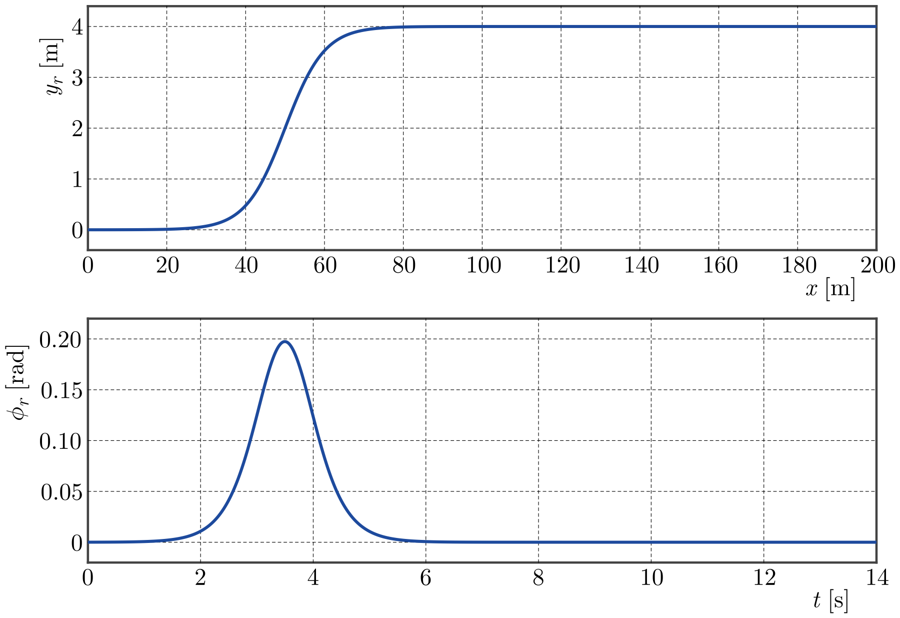


Fig. 3. Reference track and reference yaw angle under lane change conditions

The designed FCDF controls the vehicle for active steering, ensuring that the actual driving trajectory of the vehicle closely follows the reference trajectory. It also minimizes the deviation between the vehicle yaw angle and the reference value, ensuring good tracking accuracy and lateral stability. This Section demonstrates the superiority of the FCDF control strategy by comparing and analyzing its tracking effect against the PID control strategy.

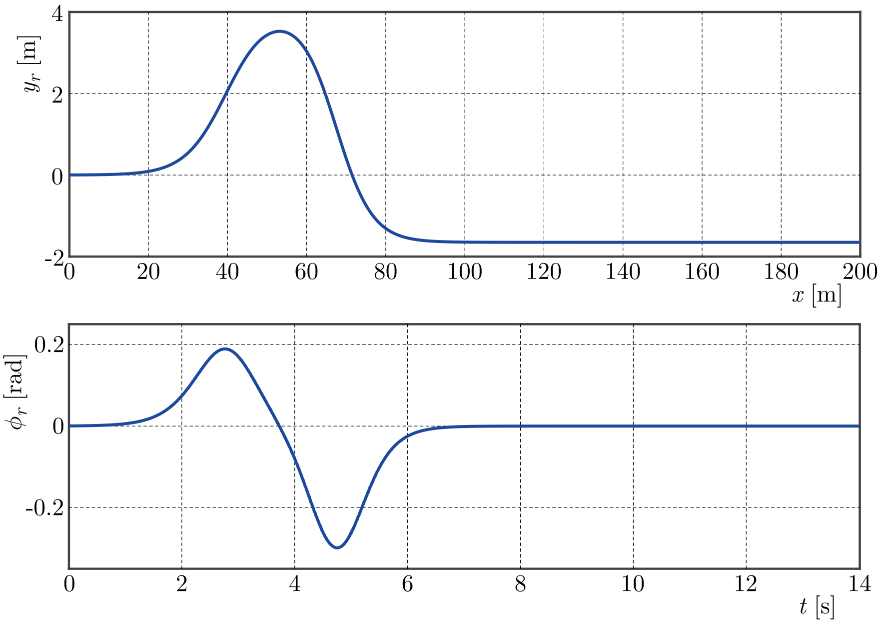


Fig. 4. Reference track and reference yaw angle under overtaking conditions

5.1. Lane change condition test

The lane change condition test verifies the track tracking effect and lateral performance of the vehicle under normal tracking and its response performance under external a random excitation.

Figure 5 shows the response curve of the trajectory tracking effect and lateral deviation. Simulation results indicate that the maximum absolute lateral deviation under PID and FCDF control is 0.1702 m and 0.1006 m, respectively. Compared to PID, the deviation amplitude under FCDF control is smaller, within the range of $(-0.1006, 0.0022)$ m. The linear section tracking accuracy is higher with minimal lateral deviation in the curved section. The final deviation convergence rate is 3.4 seconds faster than that for PID.

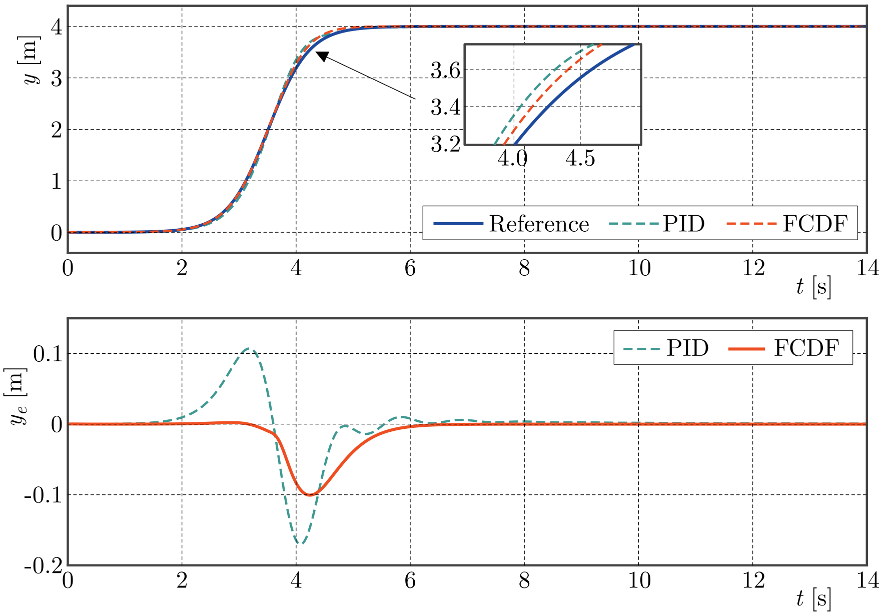


Fig. 5. Trajectory tracking effect and lateral deviation under lane change conditions

Figure 6 shows the yaw angle and its deviation response curve. The maximum yaw angle deviation under FCDF control is within 0.017 rad. The final deviation convergence rate is 1.3 seconds faster than that for PID. The maximum yaw angle deviation for PID control exceeds 0.036 rad, with noticeable oscillations. Figure 7 displays the active steering angle response. The maximum absolute active steering angle for PID and FCDF control is 3.527° and 2.740°, respectively. For PID control, the steering angle oscillation is significant at curve-straight line junctions, potentially causing vehicle instability, while for FCDF control, the steering angle oscillation is slight, ensuring vehicle stability.

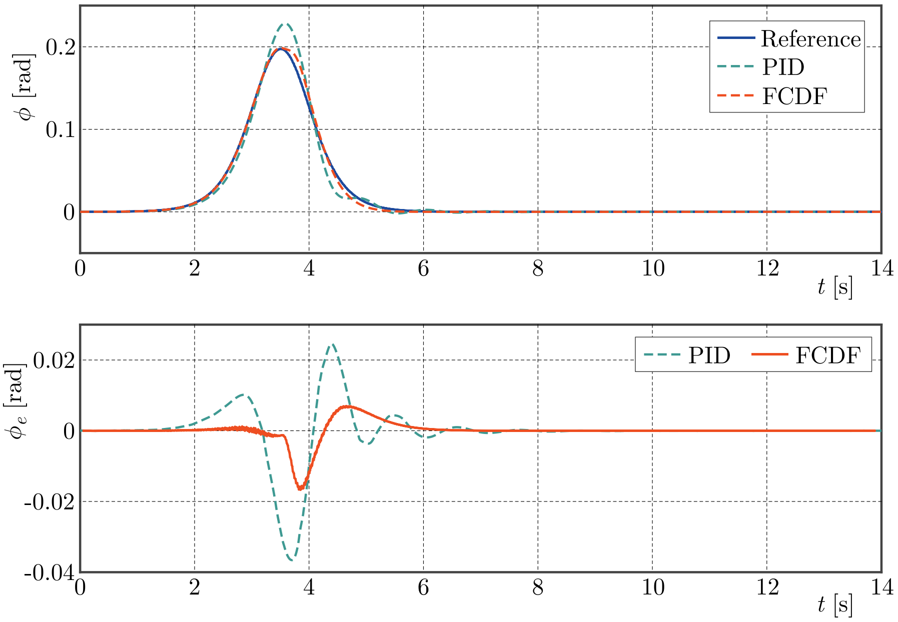


Fig. 6. Yaw angle response and deviation under lane changing conditions

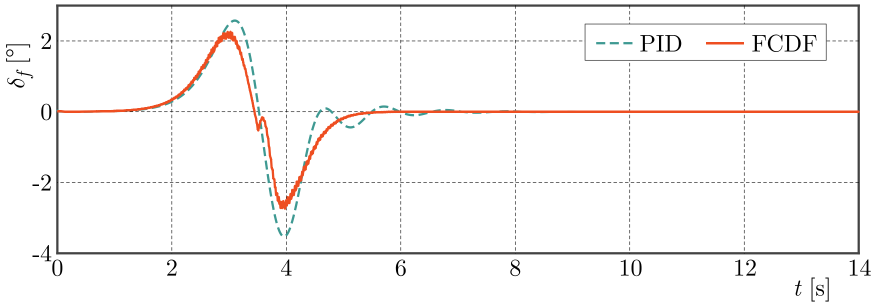


Fig. 7. Active steering angle under lane change condition

For quantitative analysis, the average absolute deviation (ABD), root mean square deviation (RSMDV), and maximum absolute deviation (MBD) were calculated and compared. Table 2 shows that ABD, RSMDV, and MBD are smaller for FCDF compared to PID, indicating better vehicle tracking for FCDF control.

To simulate continuous disturbances during lane changes, a random excitation R_1 is added to the process of obtaining the reference trajectory. The random excitation time in the whole simulation is 14 seconds

$$R_1 = \mu_{r1} + \sigma_{r1} Z_{r1} \tag{5.1}$$

where μ_{r1} represents the mean and its value is 0, σ_{r1} represents the standard deviation and its value is 1, Z_{r1} represents independent and equally distributed random numbers and fol-

Table 2. Comparison of vehicle tracking effects under lane change conditions

Evaluation parameter	Control strategy	ABD	RSMDV	MBD
Lateral displacement [m]	PID	0.0180	0.0411	0.1702
	FCDF	0.0084	0.0240	0.1006
Yaw angle [rad]	PID	0.0031	0.0076	0.0366
	FCDF	0.0010	0.0029	0.0169

lows the normal distribution $N_1(0,1)$, the unit of R_1 is meter, the sampling time is 0.01 seconds; the frequency is 100 Hz.

Figure 8 shows the trajectory tracking and yaw angle response under such conditions. It can be seen that the interference will reduce the vehicle tracking accuracy, increase the yaw angle response and reduce the lateral stability of the vehicle. Despite sustained oscillations and deviations, FCDF control shows better convergence and smaller deviation amplitude compared to PID, demonstrating better robustness against external disturbance.

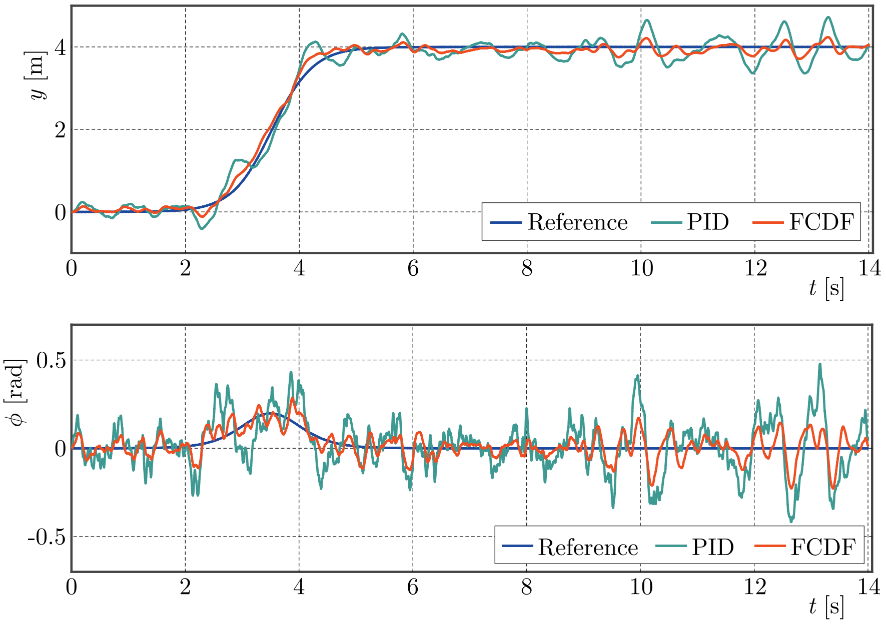


Fig. 8. Trajectory tracking effect and yaw angle response obtained by external random excitation under lane change condition

5.2. Overtaking condition test

This Section verifies the track tracking effect and lateral performance under normal tracking and response performance under external random excitation during overtaking.

Figure 9 shows the trajectory tracking effect and lateral deviation. The maximum lateral deviation for PID and FCDF control is 0.2835 m and 0.1946 m, respectively. FCDF control maintains smaller deviation amplitudes within $(-0.1946,\ 0.0052)$ m. The tracking accuracy is higher in linear sections, with a minimal lateral deviation in curved sections, and the final deviation convergence rate is 2.2 seconds faster than that for PID.

Figure 10 shows the yaw angle deviation response. The maximum yaw angle deviation for FCDF control is within 0.033 rad, with a 3.1 seconds faster convergence to 0 than PID. The maximum deviation for PID control exceeds 0.063 rad, with evident oscillations. Figure 11 shows the active steering angle response. The maximum steering angle for PID and FCDF control is 5.815° and 4.335°, respectively. For PID control, the oscillation amplitude is significant at curve-

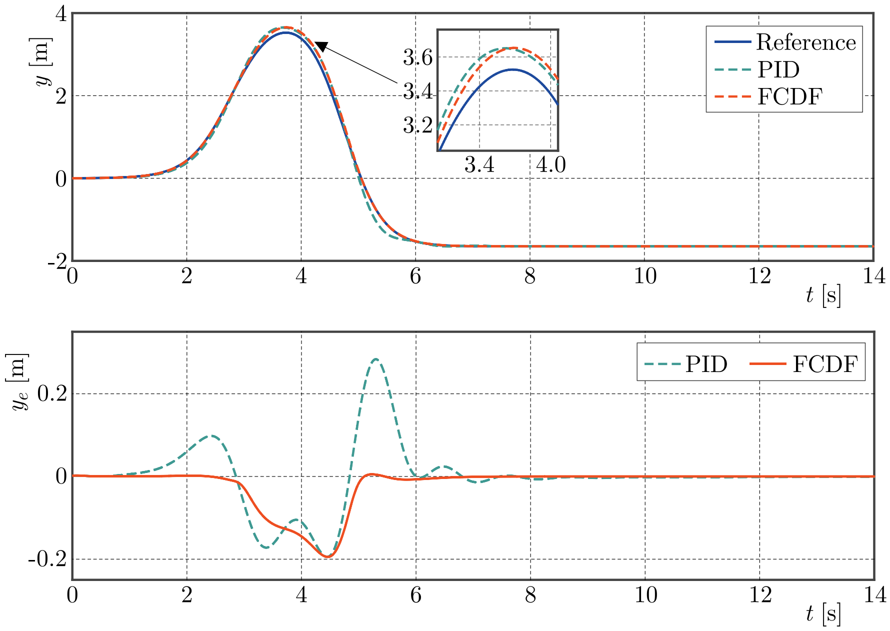


Fig. 9. Trajectory tracking effect and lateral deviation under overtaking conditions

-straight line junctions, risking instability. FCDF control maintains slight oscillations, ensuring stability.

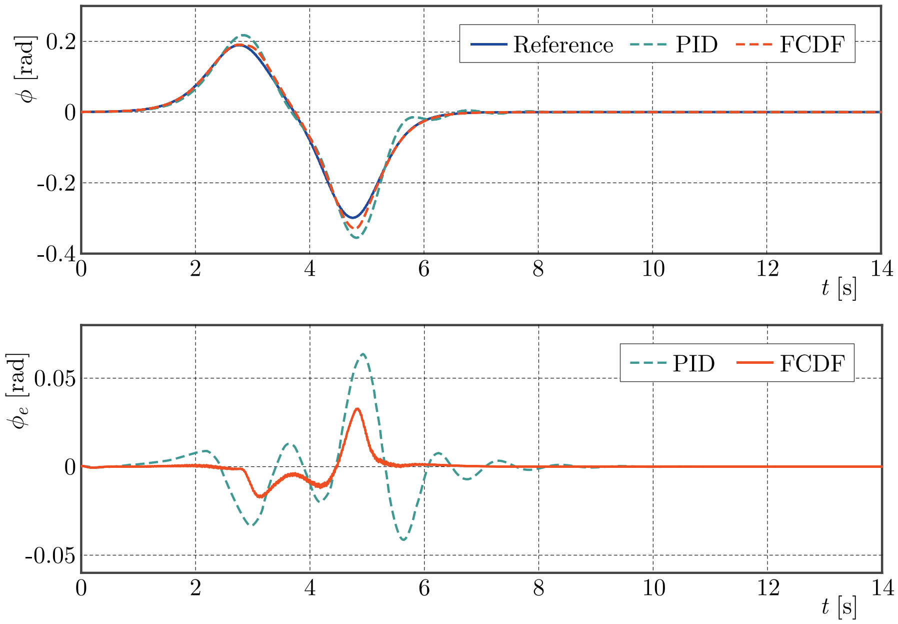


Fig. 10. Yaw angle response and deviation under overtaking conditions

Quantitative analysis (Table 3) of ABD, RSMDV, and MBD shows smaller values for FCDF, indicating better tracking performance than PID. Figure 12 shows that FCDF control has a better convergence and smaller deviation amplitude under external disturbances, demonstrating superior robustness.

Considering the interference of external uncertainties in the system, a random excitation R_2 is added to the process of obtaining the reference trajectory to simulate the change of the system due to continuous disturbance under overtaking conditions. The random excitation time

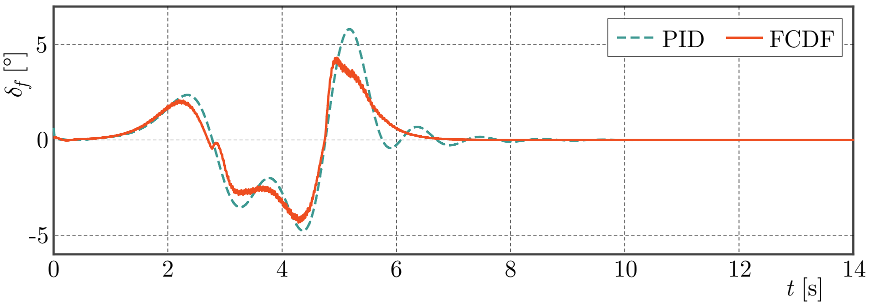


Fig. 11. Active steering angle under overtaking condition

Table 3. Comparison of vehicle tracking effects under overtaking conditions

Evaluation parameter	Control strategy	ABD	RSMDV	MBD
Lateral displacement [m]	PID	0.0427	0.0798	0.2835
	FCDF	0.0207	0.0525	0.1946
Yaw angle [rad]	PID	0.0071	0.0146	0.0636
	FCDF	0.0023	0.0058	0.0328

in the whole simulation is 14 seconds

$$R_2 = u_{r2} + \sigma_{r2} Z_{r2} \tag{5.2}$$

where μ_{r2} represents the mean and its value is 0, σ_{r2} represents the standard deviation and its value is 1, Z_{r2} represents independent and equally distributed random numbers and follows the normal distribution $N_2(0,1)$, the unit of R_2 is meter, the sampling time is 0.01 seconds, the frequency is 100 Hz.

The effect of trajectory tracking and yaw angle response are shown in Fig. 12. It can be seen that the interference will reduce the vehicle tracking accuracy, increase the yaw angle response and reduce the lateral stability of the vehicle. Through intuitive comparative analysis, it can be seen that although the trajectory tracking effect and yaw angle response under the two controllers have sustained oscillation and deviation, the trajectory tracking effect and yaw angle response under the control of FCDF have a better convergence, and the deviation change amplitude is small, and both of them are closer to the reference value. Therefore, compared with PID, FCDF has better robustness to external interference in the trajectory tracking effect and yaw angle response.

According to the comparison of the time-domain response and steady-state response in the simulation test of ASV under different working conditions, FCDF can solve the problem of underdrive in trajectory tracking control better than PID.

5.3. Robustness analysis of FCDF

ASV controller performance is particularly important during path tracking, and ASV is easy to lose stability at high speed. For this reason, the robustness of FCDF was analyzed for different vehicle speeds and road conditions. The reference values corresponding to longitudinal vehicle speeds of 30 km/h, 50 km/h and 90 km/h were Reference 1, Reference 2 and Reference 3, respectively.

Figure 13 shows the comparison of ASV driving conditions at different speeds under lane change conditions. Among them, Fig. 13a reflects the overall effect of ASV trajectory tracking and the lateral deviation generated in the tracking process. Figure 13b reflects the actual yaw response of the ASV and the deviation from the ideal yaw angle. It can be observed from the simulation test results that ASV can maintain a good tracking effect in general, and FCDF

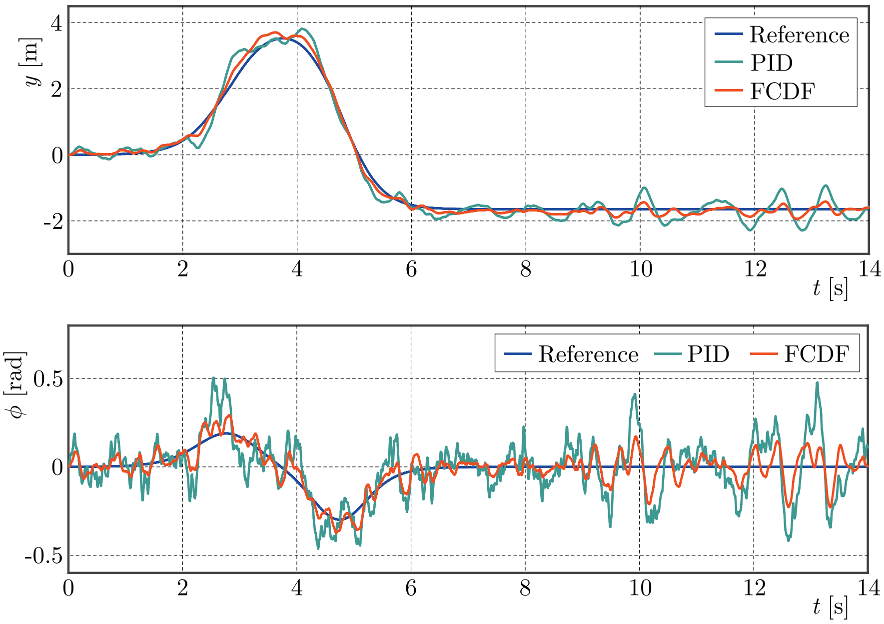


Fig. 12. Trajectory tracking effect and yaw angle response obtained by the external random excitation under the overtaking condition

has good control performance. For the 90 km/h speed test, the trajectory tracking effect and yaw angle response both oscillate slightly during 3 s to 5 s, but their amplitude is small and can converge to 0 quickly. Meanwhile, the lateral deviation and yaw angle deviation in the whole period are kept within a reasonable range.

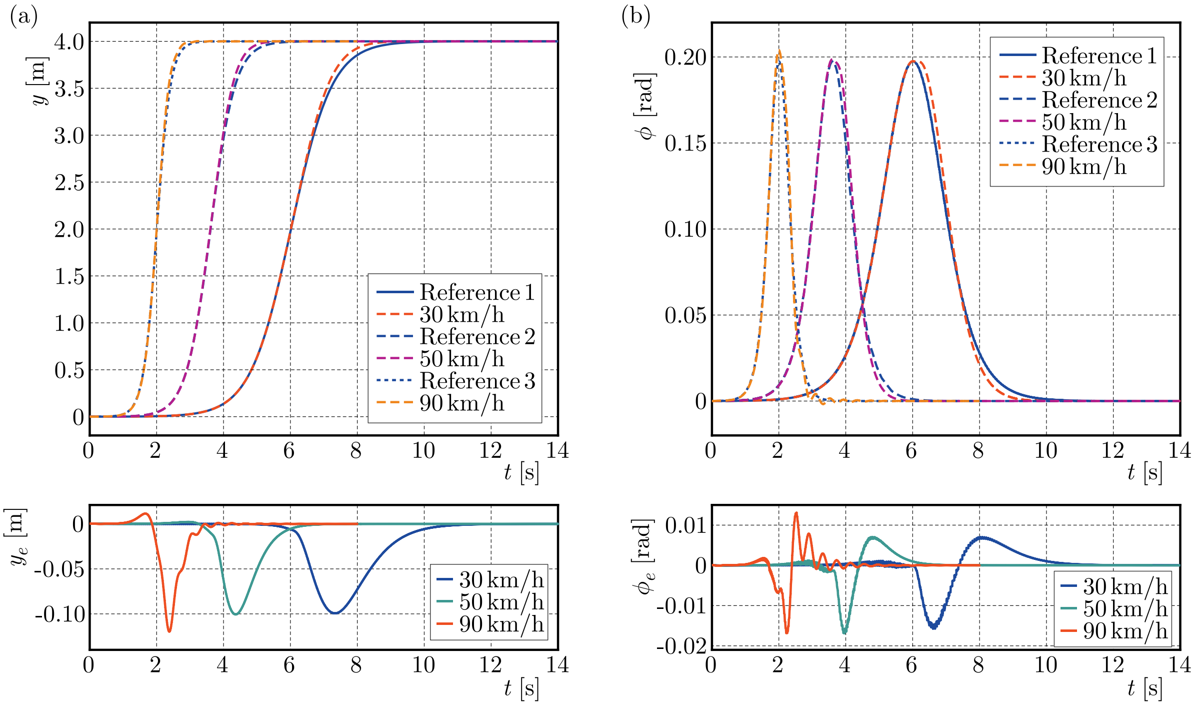


Fig. 13. The control effect of FCDF under different speed in the lane change condition: a) trajectory tracking effect and lateral deviation, b) yaw angle response and deviation

Figure 14 shows the comparison of ASV driving conditions at different speeds under overtaking condition. Among them, Fig. 14a reflects the overall effect of ASV trajectory tracking and the lateral deviation generated in the tracking process. Fig. 14b reflects the actual yaw response of the ASV and the deviation from the ideal yaw angle. According to the simulation test results, it can be found that a satisfactory tracking effect can be obtained at 30 km/h and 50 km/h, and the transverse deviation and yaw angle deviation values decrease with a decrease of speed. At 90 km/h, high speed and continuous sharp turns adversely affect the lateral stability of the ASV, and the control performance of the FCDF deteriorates, leading to a decline in tracking performance. Especially in the 2 s to 5 s period, the yaw angle oscillation is relatively sharp and the amplitude is large, which may be caused by the side of the vehicle tire, resulting in the tire state touching the nonlinear region.

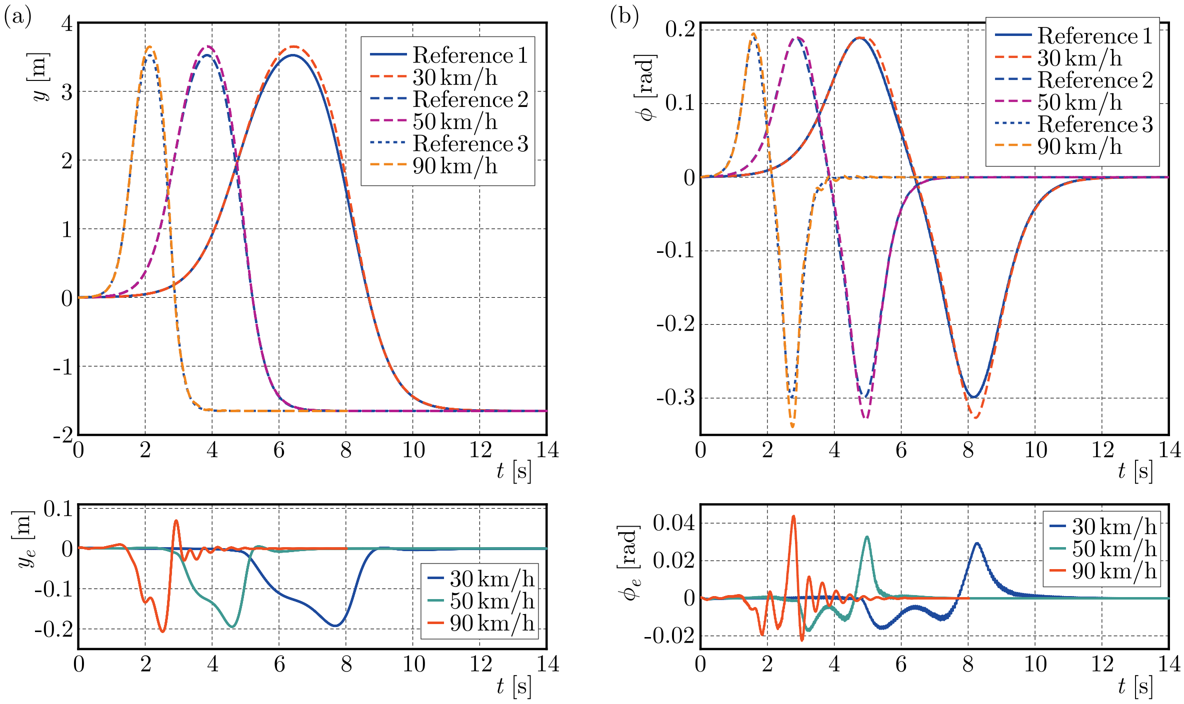


Fig. 14. The control effect of FCDF under different speed in the overtaking condition: a) trajectory tracking effect and lateral deviation, b) yaw angle response and deviation

By comparing the simulation results under different driving conditions and different speeds, the ASV can achieve accurate trajectory tracking and good lateral stability at 30 km/h and 50 km/h speeds, and the control performance of FCDF has good advantages and robustness. As the speed increases to more than 90 km/h, the ASV can achieve reasonable tracking effect in road conditions without continuous sharp turns, while the tracking performance will deteriorate to a certain extent in road conditions with continuous sharp turns. However, as long as the speed is guaranteed to change within the range of 90 km/h, even if there are external disturbances or internal uncertainties worse than the actual situation, the ASV can achieve a reasonable tracking effect in road conditions without continuous sharp turns. The simulation results show that the FCDF controller can still maintain good robustness and the control tracking lateral deviation and yaw angle deviation are bounded, which further shows that the ASV system based on differential flatness theory can effectively solve the underdrive problem.

6. Conclusion

In this paper, based on the active steering vehicle, the trajectory tracking problem of the single-input multi-output vehicle monorail model with an active steering angle input but lateral displacement and yaw angle output is studied. Firstly, through derivation and application of differential flatness theory, the original vehicle system is transformed into an equivalent form containing only the flat output and its derivatives. Then, the FCDF control strategy is applied to control the vehicle tracking and lateral stability, while dealing with external interference of the system. Finally, the vehicle system was built in Simulink to conduct simulation tests under lane change and overtaking conditions. The simulation results show that the proposed method has certain advantages and is conducive to solving nonlinear and underdrive problems of autonomous vehicles. But the differential flat control method is currently only validated in simulation, and its implementation can be quite complex, requiring an accurate system model and a deep understanding of the system dynamics, and may involve complex calculations. In addition, for systems that do not satisfy differential flatness conditions, other types of control strategies may need to be considered. Our future work will focus on applying the proposed control method to a practical experimental platform.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 52272374 & 52232013, in part by the open project of Key Laboratory of Railway Industry of Maglev Technology.

References

1. Aschemann H., Schindele D., 2008, Sliding-mode control of a high-speed linear axis driven by pneumatic muscle actuators, *IEEE Transactions on Industrial Electronics*, **55**, 11, 3855-3864
2. Bai G.X., Liu L., Meng Y., Luo W.D., Gu Q., Ma B.Q., 2019, Path tracking of mining vehicles based on nonlinear model predictive control, *Applied Sciences*, **9**, 7, 1372
3. Elmi N., Ohadi A., Samadi B., 2013, Active front-steering control of a sport utility vehicle using a robust linear quadratic regulator method, with emphasis on the roll dynamics, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **227**, 12, 1636-1649
4. Fliess M., Lévine J., Martin P., Rouchon P., 1995, Flatness and defect of non-linear systems: Introductory theory and examples, *International Journal of Control*, **61**, 6, 1327-1361
5. Guerrero J., Chemori A., Torres J., Creuze V., 2023, Time-delay high-order sliding mode control for trajectory tracking of autonomous underwater vehicles under disturbances, *Ocean Engineering*, **268**, 113375
6. Huang C.Z., Luo C.M., Li Y., Zhang T.Y., 2019, Differential flatness active disturbance rejection control approach for a class of nonlinear uncertain systems, *International Journal of Robotics and Automation*, **34**, 2, 146-155
7. Kang N., Han Y., Guan T., Wang S., 2022, Improved ADRC-based autonomous vehicle path-tracking control study considering lateral stability, *Applied Sciences*, **12**, 9, 4660
8. Li Z., Li J., Wang W., 2023, Path planning and obstacle avoidance control for autonomous multi-axis distributed vehicle based on dynamic constraints, *IEEE Transactions on Vehicular Technology*, **72**, 4, 4342-4356
9. Luo L., Cao S.Y., Sheng Z., Shen H.L., 2022, LiDAR-based global localization using histogram of orientations of principal normals, *IEEE Transactions on Intelligent Vehicles*, **7**, 3, 771-782
10. Mata S., Zubizarreta A., Pinto C., 2019, Robust tube-based model predictive control for lateral path tracking, *IEEE Transactions on Intelligent Vehicles*, **4**, 4, 569-577
11. Menhour L., d'Andréa-Novel B., Fliess M., Mounier H., 2014, Coupled nonlinear vehicle control: Flatness-based setting with algebraic estimation techniques, *Control Engineering Practice*, **22**, 135-146

12. Ortiz F.M., Sammarco M., Costa L.H.M., Detyniecki M., 2023, Applications and services using vehicular exteroceptive sensors: A survey, *IEEE Transactions on Intelligent Vehicles*, **8**, 1, 949-969

13. Rokonuzzaman M., Mohajer N., Nahavandi S., Mohamed S., 2021, Model predictive control with learned vehicle dynamics for autonomous vehicle path tracking, *IEEE Access*, **9**, 128233-128249

14. Sira-Ramirez H., Agrawal S.K., 2018, *Differentially Flat Systems*, CRC Press

15. Sun Y.G., He Z.Y., Xu J.Q., Sun W., Lin G.B., 2023, Dynamic analysis and vibration control for a maglev vehicle-guideway coupling system with experimental verification, *Mechanical Systems and Signal Processing*, **188**, 109954

16. Sun Y.G., Li F.X., Lin G.B., He Z.Y., 2023, Adaptive fault-tolerant control of high-speed maglev train suspension system with partial actuator failure: design and experiments, *Journal of Zhejiang University-SCIENCE A*, **24**, 3, 272-283

17. Sun Y.G., Qiang H.Y., Wang L., Ji W., Mardani A., 2023, A fuzzy-logic-system-based cooperative control for the multielectromagnets suspension system of maglev trains with experimental verification, *IEEE Transactions on Fuzzy Systems*, **31**, 10, 3411-3422

18. Sun Z.Y., Wang R.C., Meng X.P., Yang Y.Y., Wei Z.D., Ye Q., 2024, A novel path tracking system for autonomous vehicle based on model predictive control, *Journal of Mechanical Science and Technology*, **38**, 1, 365-378

19. Wang R.R., Zhang H., Wang J.M., 2014, Linear parameter-varying controller design for four-wheel independently actuated electric ground vehicles with active steering systems, *IEEE Transactions on Control Systems Technology*, **22**, 4, 1281-1296

20. Wang X., Sun W.C., 2023, Trajectory tracking of autonomous vehicle: A differential flatness approach with disturbance-observer-based control, *IEEE Transactions on Intelligent Vehicles*, **8**, 2, 1368-1379

21. Wang Z.Q., Sun K.Y., Ma S.Q., Sun L.T., Gao W., Dong Z.Z., 2022, Improved linear quadratic regulator lateral path tracking approach based on a real-time updated algorithm with fuzzy control and cosine similarity for autonomous vehicles, *Electronics*, **11**, 22, 3703

22. Xia Y.Q., Pu F., Li S.F., Gao Y., 2016, Lateral path tracking control of autonomous land vehicle based on ADRC and differential flatness, *IEEE Transactions on Industrial Electronics*, **63**, 5, 3091-3099

23. Yang T., Bai Z.W., Li Z.Q., Feng N.L., Chen L.Q., 2021, Intelligent vehicle lateral control method based on feedforward + predictive LQR algorithm, *Actuators*, **10**, 9, 228

24. Yu H.W., Zhang Z.Z., Xing J.F., 2021, Micro feed characteristic analysis of a new crawler guide rail dual drive servo system, *Science Progress*, **104**, 3, 1-24

25. Zhuang Y.F., Ma G.F., Huang H.B., 2010, Time-optimal motion planning of an underactuated rigid spacecraft, *Control and Decision* (in Chinese), **25**, 10, 1469-1473

Manuscript received July 14, 2024; accepted for publication November 8, 2024

**JOURNAL OF THEORETICAL
AND APPLIED MECHANICS**
63, 1, pp. 151-164, Warsaw 2025
https://doi.org/10.15632/jtam-pl/196984

LOAD TRANSFER LAW AND BEARING CAPACITY OF THE ANCHORAGE INTERFACE BASED ON A THREE-STAGE MODEL

Longfei Li, Xuebin Wang
College of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao, China
Tongbin Zhao, Minglu Xing, Weiyao Guo, Yubao Zhang
College of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao, China, and State Key Laboratory of Mining Disaster Prevention and Control Co-founded by Shandong Province and the Ministry of Science and Technology, Shandong University of Science and Technology, Qingdao, China
corresponding author Tongbin Zhao, e-mail: 2863898939@qq.com

The shear stress constitutive equation of an anchorage interface element is established based on a three-stage model, and the expressions for the interface shear stress and axial force distribution evolution during the whole process of anchor drawing are derived theoretically. The parameters of the interface model are given through anchor pull-out tests. Then, the load transfer law and the bearing performance of the anchorage interface under different anchoring agents are calculated and analyzed. The calculation method of the ultimate bearing capacity of the anchorage interface is verified by supplementary tests, which can provide theoretical guidance for the bolt support design.

Keywords: constitutive model, anchor pull-out test, law of load transfer, interface bearing capacity

1. Introduction

The rockbolt support has been widely used in the roadway surrounding rock support engineering, because it can fully mobilize the strength and self-stabilizing ability of the surrounding rock (Zhao *et al.*, 2021; Yi *et al.*, 2020). In the engineering site, failure of the rockbolt support system will lead to serious instability and damage of the surrounding rock. Therefore, the analysis and evaluation of the bearing capacity of the rockbolt support system are of great significance for improving the support effect and reducing the failure risk of the rockbolt support system (Fan *et al.*, 2021).

In order to explore the failure mechanism of the rockbolt support system, researchers have conducted a large number of laboratory experiments and in situ tests. The research results show that failure of the rockbolt support system is mainly caused by interface debonding or rockbolt breaking (Li, 2010; Wang *et al.*, 2023; Kilic *et al.*, 2002). According to the results of rockbolt pull-out tests under different pull-out loads, Høien *et al.* (2021) classified failure mechanisms of the anchorage into three categories: debonding slip below the yield strength of the rockbolt, debonding slip between the yield load and ultimate load, and rockbolt breaking under the ultimate load; Liu *et al.* (2021) studied the failure mechanism of the anchorage interface through laboratory rockbolt pull-out tests, then divided the failure mode of the anchorage interface into the shear slip mode and shear expansion slip mode. Cao *et al.* (2014) found that the failure mode of the rockbolt system is closely related to the contact performance of the anchorage interface and the actual stress state of the material. Researches show that a high-quality anchorage interface bonding performance is an important part of the bolt support design. It is of great significance to improve the bearing capacity of the anchorage interface to ensure long-term effectiveness of the rockbolt support.

In order to further explore the relationship between mechanical parameters of the anchorage interface and anchorage bearing capacity, Kilic *et al.* (2002) conducted a series of laboratory tests on a grouting rockbolt, and put forward some empirical formulas for calculating the bearing capacity of the rockbolt. Holý (2017) used sandstone as an anchor matrix to carry out an anchor tensile test, and it was found that interfacial shear friction and surrounding rock properties had obvious influence on anchorage performance. Yahia *et al.* (1998) carried out a large number of tensile tests of rockbolts under different anchorage conditions, and found that the mechanical properties of anchoring agents affected the bearing capacity of the anchorage. A large number of experimental studies show that (Huang *et al.*, 2019; Wu *et al.*, 2019; Du *et al.*, 2021) the bond strength of the anchorage interface determines the critical shear stress when the anchorage interface is destroyed, and its deformation characteristics determine the load transfer mode and the distribution law of interfacial shear stress on the anchorage interface. Therefore, the shear stress-slip displacement relationship of the anchorage interface is the basis for studying deformation characteristics of the anchorage interface. A reasonable interface shear-slip model can effectively explain the interface load transfer and damage mechanism. The bond-slip models commonly used in the existing research to characterize the shear stress-shear displacement relationship of the anchorage interface are bilinear model (Cai *et al.*, 2004), trilinear model (Ma *et al.*, 2016; Nie *et al.*, 2019; Chen *et al.*, 2021) and nonlinear model (Nemcik *et al.*, 2014). According to the existing bond-slip model, many scholars have studied it by analytical and numerical methods (Chen *et al.*, 2020; Chen and Li, 2022; Yue *et al.*, 2022), which reveals the evolution law of shear stress distribution at the anchorage interface and provides a reference for the study of bearing capacity of the anchorage interface.

In summary, the bonding performance of the anchorage interface directly affects the bearing capacity of the anchorage, and the research on the influence rules of interface model parameters on the bonding performance of the anchorage interface still needs to be improved. Therefore, this paper deduces the load transfer law of the anchorage interface through theoretical derivation, carries out anchor rod pull-out tests under different anchoring conditions of anchoring agents, tests and analyzes the constitutive model parameters of the anchorage interface, and conducts in-depth research on the bearing capacity of the anchorage interface.

2. Analysis of load transfer laws at the anchorage interface

2.1. Three-stage interface element constitutive model and parameters

As shown in Fig. 1, the three-stage constitutive model divides the debonding failure process of the anchorage interface element into three stages: I – elastic stage: shear stress increases linearly with shear displacement. τ_s^d is the peak shear strength of the anchorage interface element in the elastic stage, and δ_s^d is the corresponding shear displacement; II – bond damage stage: shear stress decreases linearly with shear displacement. τ_s^s is the residual shear strength of the anchorage interface element, which is also the shear strength at the onset of slip; δ_s^s is the corresponding shear displacement at the onset of slip; III – frictional slip stage: in this stage, friction plays the dominant role, and the shear stress remains constant. In the figure, k_{ss} represents the elastic shear stiffness characterizing the evolution of shear stress with shear displacement in the elastic stage of the interface element $k_{ss} = \tau_s^d/\delta_s^d$.

As shown in Fig. 1, the constitutive equation for evolution of shear stress damage in the anchorage interface element characterized by the three-stage constitutive model is

$$\tau_s = \begin{cases} K_{OA}\delta_s & \delta_s \leqslant \delta_s^d \\ \tau_s^d + K_{AB}(\delta_s - \delta_s^d) & \delta_s^d \leqslant \delta_s \leqslant \delta_s^s \\ \tau_s^s & \delta_s^s \leqslant \delta_s \leqslant \delta_s^f \end{cases} \tag{2.1}$$

where $K_{OA} = k_{ss} = \tau_s^d/\delta_s^d$, $K_{AB} = (\tau_s^d - \tau_s^s)/(\delta_s^d - \delta_s^s)$.

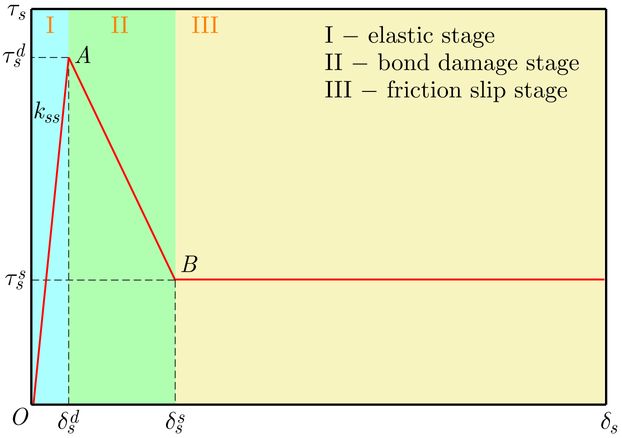


Fig. 1. Three-stage constitutive model and parameters

2.2. Load transfer equation for anchorage interface elements

As shown in Fig. 2a, to facilitate the analysis of load transfer laws at the anchorage interface, the shear interaction between the anchor rod and the rock mass is simplified into a spring contact model, with the side of the anchor rod connected to the rock mass through tangential line springs, the stiffness of which corresponds to the elastic shear stiffness of the anchorage interface.

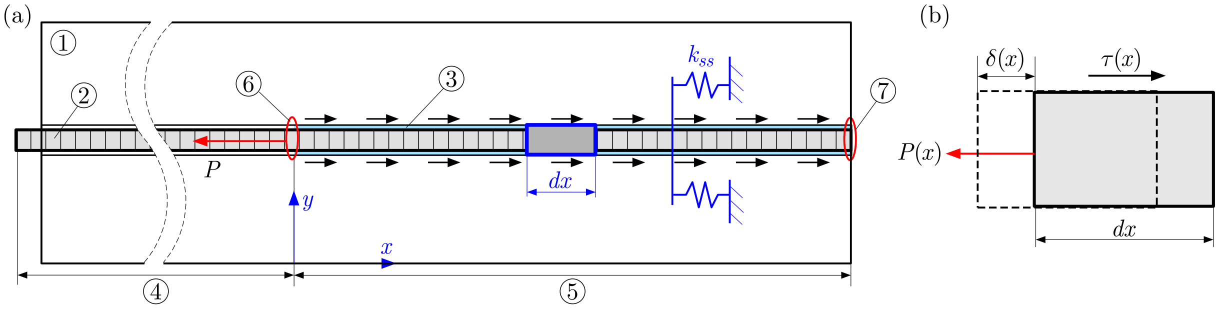


Fig. 2. Load transfer analysis diagram of the anchorage interface (1 – rock mass; 2 – anchor rod; 3 – anchor rod-rock mass coupling interface; 4 – free section of anchor rod; 5 – anchored section of anchor rod; 6 – anchored front end; 7 – anchored tail end): (a) simplified model of the anchorage interface, (b) anchorage interface element

Take an interface element at a distance of x from the anchored front end in the anchorage section. As shown in Fig. 2b, analyze the element using the load transfer method to determine the force equilibrium equations. The force equilibrium equations are

$$dP(x) = -2\pi r_b \tau(x)\,dx \qquad \frac{d\delta(x)}{dx} = -\frac{1}{\pi r_b^2 E_b} P(x) \tag{2.2}$$

where $P(x)$ is the axial force of the anchor rod at a distance x from the anchored front end, $\tau(x)$ is the shear stress, $\delta(x)$ is the shear displacement, r_b is the radius of the anchor rod, E_b is the elastic modulus of the anchor rod.

Substituting Eq. $(2.2)_1$ into Eq. $(2.2)_2$, the basic load transfer equation for the interface element is obtained

$$\frac{d^2\delta(x)}{dx^2} = \frac{2}{r_b E_b}\tau(x) \tag{2.3}$$

Substituting Eq. $(2.2)_2$ into Eq. (2.1), the differential equation for load transfer during the debonding failure process of the anchorage interface element is obtained

$$\frac{d^2\delta(x)}{dx^2}=\begin{cases}\alpha^2\delta(x) & 0\leqslant\delta(x)\leqslant\delta_s^d\\ -\beta^2\delta(x)+\dfrac{2\left(\tau_s^d-K_{AB}\delta_s^d\right)}{r_bE_b} & \delta_s^d\leqslant\delta(x)\leqslant\delta_s^s\\ \dfrac{2\tau_s^s}{r_bE_b} & \delta_s^s\leqslant\delta(x)\end{cases}\tag{2.4}$$

where $\alpha^2=(2K_{OA})/(r_bE_b)$, $\beta^2=-(2K_{AB})/(r_bE_b)$.

2.3. Calculation formulas for shear stress and axial force distribution at the anchorage interface

Under the action of axial tensile load, the anchorage interface element will sequentially experience elastic, bond damage, and frictional slip stages. As shown in Fig. 3, the anchorage interface is composed of anchorage interface elements. Therefore, the state of the anchorage interface is determined by the stages in which all interface elements are located. When the anchorage interface elements are in different stages, the anchorage interface will have a situation where parts of the elastic segment, bond damage segment, and frictional slip segment coexist.

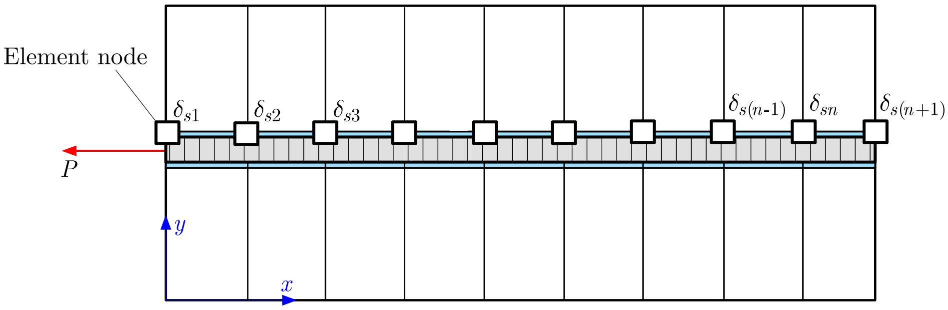


Fig. 3. Schematic diagram of the anchorage interface and interface element relationship

Depending on the state of the anchorage interface elements, the anchorage interface may exist in a full-length elastic stage, elastic-damage stage, full-length damage stage, elastic-damage-slip stage, damage-slip stage, or full-length slip stage. By using the boundary conditions of different stages, combined with Eqs. $(2.2)_1$ and $(2.2)_2$, by solving the load transfer differential equation as shown in Eq. (2.4), it is possible to theoretically calculate the evolution formulas for the distribution of shear stress and axial force along the anchorage interface at different anchorage lengths, as listed in Table 1.

3. Anchor rod pull-out tests

3.1. Preparation of anchorage specimens and test scheme

The evolution law of stress distribution at the anchorage interface is closely related to the parameters of the interface element constitutive model. To determine the model parameters of the anchorage interface and to explore the influence of different anchoring agents on the support effect of the anchor rod, three groups of anchor rod pull-out tests were designed. The tests used large cylindrical specimens made of C40 concrete with a diameter of $d_r = 300$ mm and a length of $l_r = 600$ mm to simulate the anchoring matrix. During preparation of the anchoring matrix, a pipe with an outer diameter of $d_a = 30$ mm was inserted into the center of the

Table 1. Formulas for shear stress and axial force distribution at the anchorage interface in different stages

Anchorage interface stage	Boundary conditions	Distribution formulas of interfacial shear stress and rockbolt axial force
Full-length elastic stage	$P(x)\vert_{x=0} = P_0$ $P(x)\vert_{x=L} = 0$	$\tau(x) = \frac{\alpha P_0 \cosh[\alpha(L-x)]}{2\pi r_b \sinh(\alpha L)} \qquad 0 \leqslant x \leqslant L$ $P(x) = \frac{P_0 \sinh[\alpha(L-x)]}{\sinh(\alpha L)} \qquad 0 \leqslant x \leqslant L$
Elastic- -damage stage	$P_e(x)\vert_{x=L} = 0$ $\tau_e(x)\vert_{x=l_d} = \tau_s^d$ $\tau_d(x)\vert_{x=l_d} = \tau_s^d$ $P_d(x)\vert_{x=l_d}$ $= P_e(x)\vert_{x=l_d}$	$\tau(x) = \begin{cases} \tau_s^d \cos[\beta(l_d - x)] \\ \quad -\frac{\beta\tau_s^d}{\alpha} \tanh[\alpha(L - l_d)] \sin[\beta(l_d - x)] \\ \quad 0 \leqslant x \leqslant l_d \\ \frac{\tau_s^d \cosh[\alpha(L-x)]}{\cosh[\alpha(L-l_d)]} \qquad l_d \leqslant x \leqslant L \end{cases}$ $P(x) = \begin{cases} 2\pi r_b \tau_s^d \{ \frac{\tanh[\alpha(L-l_d)] \cos[\beta(l_d-x)]}{\alpha} \\ \quad + \frac{\sin[\beta(l_d-x)]}{\beta} \} \qquad 0 \leqslant x \leqslant l_d \\ \frac{2\pi r_b \tau_s^d \sinh[\alpha(L-x)]}{\alpha \cosh[\alpha(L-l_d)]} \qquad l_d \leqslant x \leqslant L \end{cases}$
Full-length damage stage	$P(x)\vert_{x=0} = P_0$ $P(x)\vert_{x=L} = 0$	$\tau(x) = \frac{\beta P_0 \cos[\beta(L-x)]}{2\pi r_b \sin(\beta L)} \qquad 0 \leqslant x \leqslant L$ $P(x) = \frac{P_0 \sin[\beta(L-x)]}{\sin(\beta L)} \qquad 0 \leqslant x \leqslant L$
Elastic- -damage- -slip stage	$\tau_e(x)\vert_{x=l_d+l_f} = \tau_s^d$ $P_e(x)\vert_{x=L} = 0$ $\tau_d(x)\vert_{x=l_d+l_f} = \tau_s^d$ $\tau_d(x)\vert_{x=l_f} = \tau_s^s$ $P_d(x)\vert_{x=l_d+l_f}$ $= P_e(x)\vert_{x=l_d+l_f}$	$\tau(x) = \begin{cases} \tau_s^s \qquad 0 \leqslant x \leqslant l_f \\ \frac{\tau_s^s \sin[\beta(l_d+l_f-x)] - \tau_s^d \sin[\beta(l_f-x)]}{\sin(\beta l_d)} \\ \quad l_f \leqslant x \leqslant l_f + l_d \\ \frac{\tau_s^d \cosh[\alpha(L-x)]}{\cosh[\alpha(L-l_d-l_f)]} \qquad l_f + l_d \leqslant x \leqslant L \end{cases}$ $P(x) = \begin{cases} 2\pi r_b \tau_s^d \{ \frac{\sin(\beta l_d)}{\beta} + \frac{\tanh[\alpha(L-l_d-l_f)] \cos(\beta l_d)}{\alpha} \} \\ \quad + 2\pi r_b \tau_s^s (l_f - x) \qquad 0 \leqslant x \leqslant l_f \\ 2\pi r_b \tau_s^d \{ \frac{\tanh[\alpha(L-l_d-l_f)] \cos[\beta(l_d+l_f-x)]}{\alpha} \\ \quad + \frac{\sin[\beta(l_d+l_f-x)]}{\beta} \} \qquad l_f \leqslant x \leqslant l_f + l_d \\ \frac{2\pi r_b \tau_s^d \sinh[\alpha(L-x)]}{\alpha \cosh[\alpha(L-l_d-l_f)]} \qquad l_f + l_d \leqslant x \leqslant L \end{cases}$
Damage- -slip stage	$P_d(x)\vert_{x=L} = 0$ $\tau_d(x)\vert_{x=l_f} = \tau_s^s$ $P_f(x)\vert_{x=l_f}$ $= P_d(x)\vert_{x=l_f}$ $\delta_f(x)\vert_{x=l_f} = \delta_s^s$	$\tau(x) = \begin{cases} \tau_s^s \qquad 0 \leqslant x \leqslant l_f \\ \frac{\tau_s^s \cos[\beta(L-x)]}{\cos[\beta(L-l_f)]} \qquad l_f \leqslant x \leqslant L \end{cases}$ $P(x) = \begin{cases} 2\pi r_b \tau_s^s \{ l_f - x + \frac{\tan[\beta(L-l_f)]}{\beta} \} \qquad 0 \leqslant x \leqslant l_f \\ \frac{2\pi r_b \tau_s^s \sin[\beta(L-x)]}{\beta \cos[\beta(L-l_f)]} \qquad l_f \leqslant x \leqslant L \end{cases}$
Full-length slip stage	–	$\tau(x) = \tau_s^s \qquad 0 < x < L$ $P(x) = P_0 - 2\pi r_b \tau_s^s x \qquad 0 \leqslant x \leqslant L$

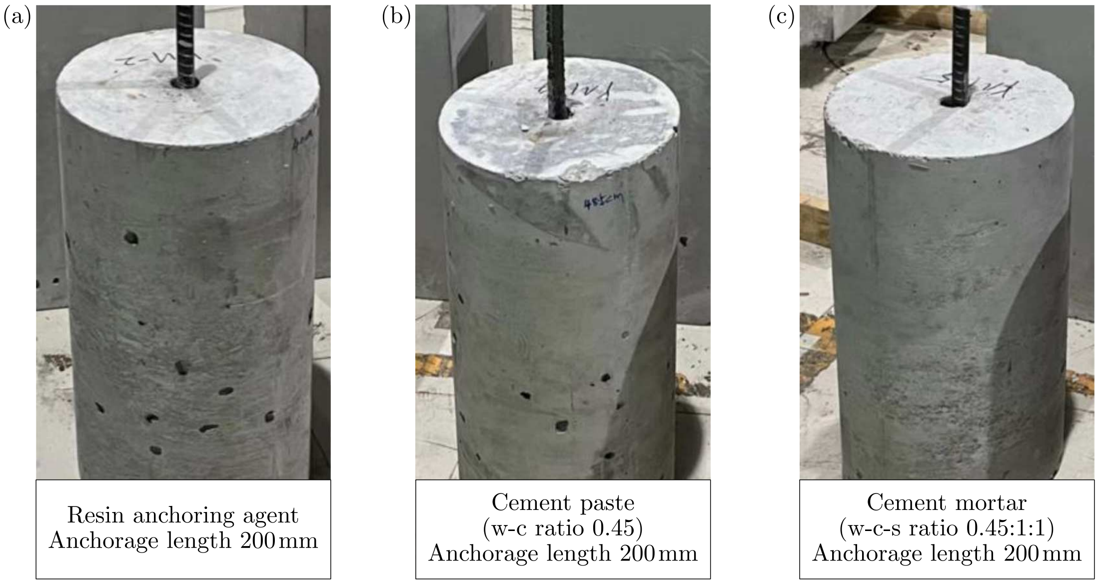
Note: l_d is the length of the bond damage segment, l_f is the length of the slip segment.

anchoring matrix preparation mold to prefabricate the anchoring borehole. The anchor rod used for anchoring was a left-handed thread steel anchor rod with a model number of MG400 and an equivalent diameter of $d_b = 20\,\mathrm{mm}$. The elastic modulus of the anchor rod was $E_b = 200\,\mathrm{GPa}$, the yield load $P_y = 150\,\mathrm{kN}$, and the breaking load $P_f = 210\,\mathrm{kN}$.

Three different types of anchoring agents were selected for the tests: resin anchoring agent, cement grout with a water-cement ratio of 0.45:1, and cement mortar with a water-cement-sand ratio of 0.45:1:1. The designed length of the anchorage section is $l_a = 200\,\mathrm{mm}$, and the calculation formula for the required volume of anchoring agent is

$$V = \frac{\pi}{4} l_a (d_a^2 - d_b^2) \tag{3.1}$$

According to Eq. (3.1), the volume of anchoring agent required for an anchorage length of 200 mm is calculated to be $V = 78.5\,\mathrm{cm}^3$. After injecting the anchoring material, the anchor rods are inserted for anchoring, and the tests are conducted after 7 days of curing. Figure 4 shows the three groups of anchoring samples prepared in the laboratory.


(a) Resin anchoring agent Anchorage length 200 mm

(b) Cement paste (w-c ratio 0.45) Anchorage length 200 mm

(c) Cement mortar (w-c-s ratio 0.45:1:1) Anchorage length 200 mm

Fig. 4. Diagram of three groups of anchorage specimens prepared in the laboratory


Fig. 5. The RTR-3122 full-scale anchorage performance tester

The anchor rod pull-out tests are conducted on the self-developed RTR-3122 model full-scale anchor rod anchorage performance testing machine (as shown in Fig. 5). Before the test, an anchorage pull-out device is installed on the anchorage performance testing machine. One end of the anchorage specimen is blocked by a rigid load-bearing plate and suspended behind the stop plate of the testing machine. The stop plate serves to fix the anchorage specimen and also prevents wedge failure of the rock mass during the pull-out process. The loading end applies a displacement loading by clamping the free end of the anchor rod with the wedge-shaped clamp of the testing machine. An LVDT (linear variable differential transformer) fixed frame is installed at a measuring point 200 mm away from the rock mass stop plate on the free section of the anchor rod. The spring probe of the LVDT displacement sensor is placed at the stop plate to record the relative displacement between the anchor rod measuring point and the rock mass, as shown in Fig. 6.

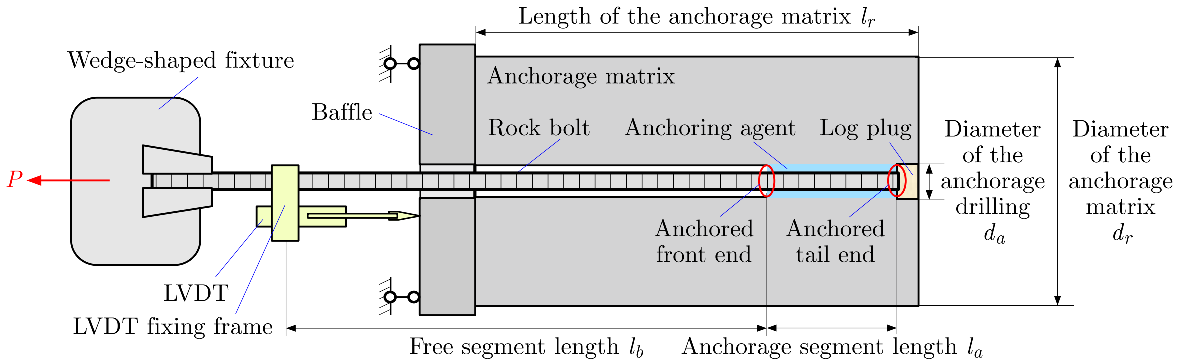


Fig. 6. Schematic diagram of the anchor rod pull-out tests

3.2. Mechanical characteristic curves of the anchorage interface and constitutive model parameters

The tests obtained the axial tensile load-displacement curves of short anchor rods under the action of different anchoring agents (P-δ_{s0} curves), as shown in Fig. 7. The peak tensile loads of the anchor rods under the three test conditions are all less than the yield load of the anchor rod, indicating that the anchor rods were pulled out in the elastic stage.

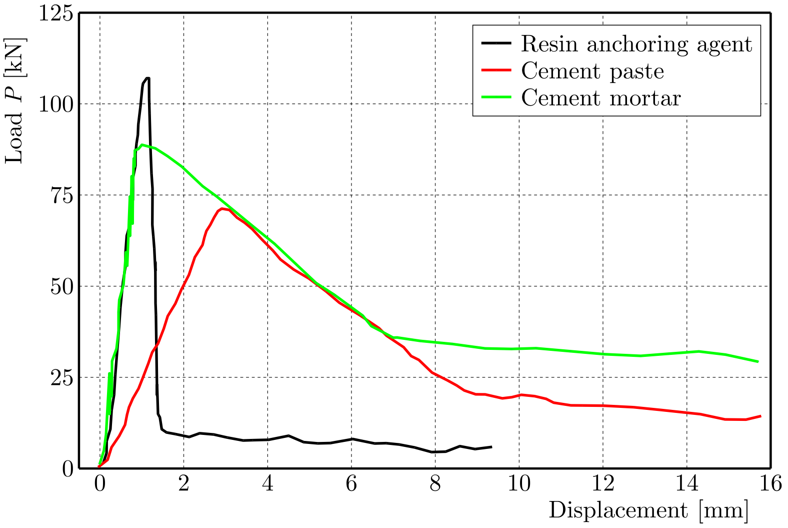


Fig. 7. Anchorage pull-out load-displacement curves

Since the LVDT displacement meter is fixed between the anchor rod and the testing machine stop plate, the measured displacement δ_{s0} is the displacement of the anchor rod measuring point relative to the stop plate, which includes both the elastic deformation of the free section of the anchor rod and the relative displacement between the anchor rod and the rock mass. Subtracting the elastic deformation of the free section of the anchor rod from δ_{s0} allows us to obtain the relative displacement between the anchor rod and the rock mass, i.e., the shear displacement of the anchor rod-rock mass coupling interface. The calculation formula is

$$\delta_s = \delta_{s0} - \frac{4Pl_b}{\pi d_b^2 E_b} \tag{3.2}$$

where l_b represents the length of the free section of the anchor rod.

Since the anchorage section length designed in the test is relatively short, it can be assumed that the shear stress at the anchorage interface is uniformly distributed. Based on the anchor rod pull-out load-displacement curve, the calculation formula for the shear stress τ_s at the anchor rod-rock mass coupling interface is

$$\tau_s = \frac{P}{\pi d_b l_a} \tag{3.3}$$

where l_a represents the length of the anchorage section.

Since the anchor rod does not yield during short anchorage pull-out, the test curve can be considered as the mechanical characteristic curve of a single anchorage interface element. Based on the shear stress-shear displacement relationship curve, the interface model parameters shown in Fig. 1 can be calibrated, where the calculation formulas for τ_s^d and τ_s^s are as follows

$$\tau_s^d = \frac{P_{max}}{\pi d_b l_a} \qquad \tau_s^s = \frac{P_{res}}{\pi d_b l_a} \tag{3.4}$$

where P_{max} is the peak tensile load in the elastic stage of axial stretching of the anchor rod, P_{res} is the residual tensile load in the axial stretching of the anchor rod.

The shear stress-shear displacement relationship curves of the anchor rod-rock mass coupling interface, calculated under different anchoring agent conditions according to Eqs. (3.2) to (3.3), are shown in Fig. 8. The constitutive model parameters of the anchorage interface element under the action of different anchoring agents are calculated and obtained, as listed in Table 2.

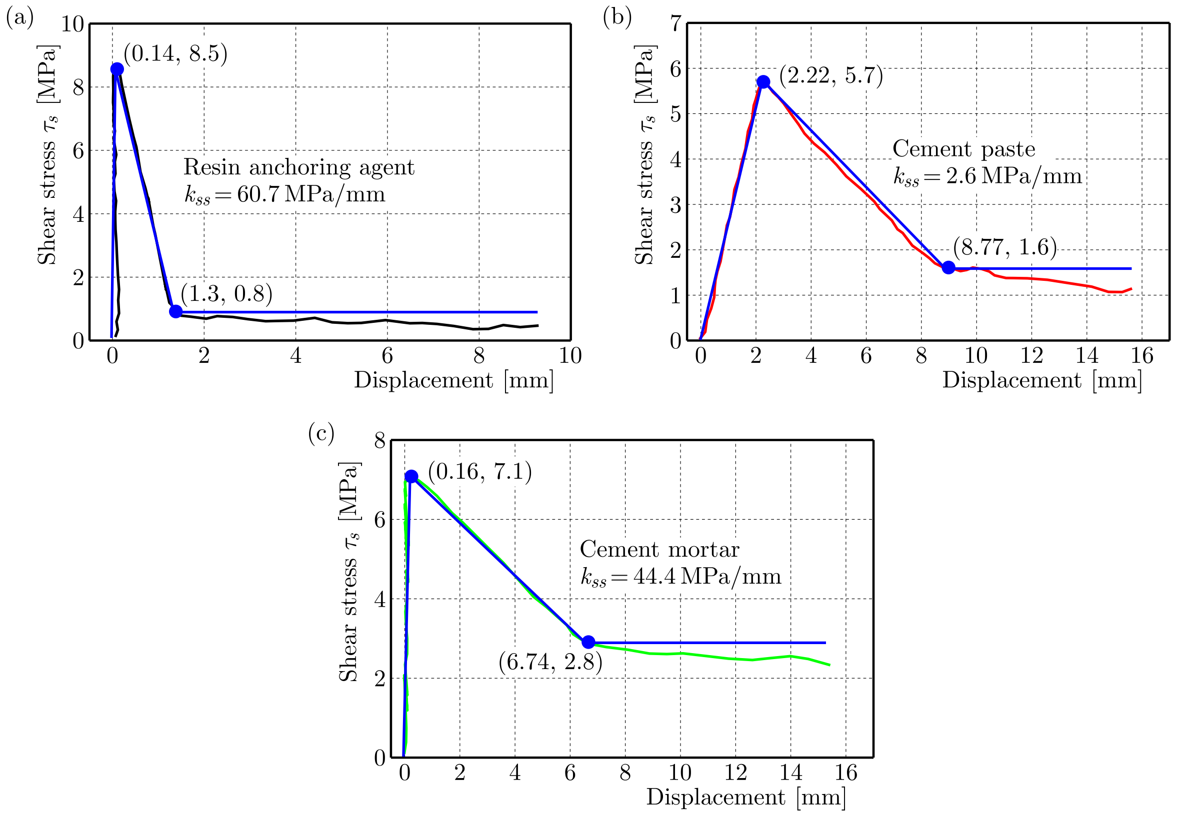


Fig. 8. Anchorage interface mechanical characteristic curves: (a) resin anchoring agent, (b) cement paste, (c) cement mortar

From Fig. 8, it can be seen that different anchoring agents have a significant impact on the mechanical parameters characterizing the bonding performance of the anchorage interface. Under the action of different anchoring agents, the ranking of the peak shear strength of the anchorage

Table 2. Trilinear bond-slip constitutive model parameters under different anchorage agent conditions

Anchoring agent types	τ_s^d [MPa]	δ_s^d [mm]	τ_s^s [MPa]	δ_s^s [mm]	k_{SS} [MPa/mm]
Resin anchoring agent	8.5	0.14	0.8	1.3	60.7
Cement paste	5.7	2.22	1.6	8.77	2.6
Cement mortar	7.1	0.16	2.8	6.74	44.4

interface is: resin anchoring agent > cement mortar > cement paste; the ranking of the residual shear strength of the anchorage interface is: cement mortar > cement paste > resin anchoring agent; the ranking of the residual shear displacement of the anchorage interface is: cement paste > cement mortar > resin anchoring agent; the elastic shear stiffness of the anchorage interface varies greatly, with the highest reaching 60.7 MPa/mm and the lowest being 2.6 MPa/mm, and the ranking is: resin anchoring agent > cement mortar > cement paste. Comparing Fig. 8b and Fig. 8c, it can be seen that when cement is used as the anchoring agent, adding sand particles to the anchoring agent can increase the peak shear strength, residual shear strength, and elastic shear stiffness of the anchorage interface, but at the same time, it reduces the initial sliding displacement of the anchorage interface.

3.3. Load transfer laws of the anchorage interface under the action of different anchoring agents

Assuming the anchorage length $L = 600$ mm, based on the constitutive model parameters of the anchorage interface element shown in Fig. 8, and using the evolution formulas for the distribution of shear stress and axial force at the anchorage interface listed in Table 1, theoretical calculations can be made to obtain the distribution and evolution process of the shear stress and axial force during the entire process of interface debonding failure under different anchoring agent conditions, as shown in Fig. 9.

Since the anchor rod itself has a certain bearing limit, when the theoretical calculated axial force exceeds the breaking load of the anchor rod, the anchor rod will break. The breaking load P_b of the anchor rod used in the test is 210 kN. From Fig. 9, it can be seen that when the anchorage length is 600 mm, the anchor body under the conditions of resin anchoring agent and cement mortar anchoring will experience anchor rod breakage failure, while under the condition of cement grout anchoring, the anchor body will experience interface debonding slip failure. Observing the distribution of interface shear stress when the axial force equals the breaking load of the anchor rod in Figs. 9a and 9e, it is shown that the anchor rod will break in the elastic-damage stage of the anchorage interface. At that moment, the degree of damage to the anchorage interface is relatively low, indicating that the anchorage support effect is greatly influenced by the strength limit of the anchor rod at this time. Observing Figs. 9a, 9c, and 9e, it can be seen from the distribution pattern of shear stress in different sections of the anchorage interface that the greater the elastic shear stiffness k_{ss} of the interface, the more obvious the decay of the shear stress and axial force in the elastic section of the interface along the anchorage length, and the smaller the axial force of the anchor rod when the interface begins to damage.

4. Analysis of the bearing performance of the anchorage interface

4.1. Theoretical analysis of the ultimate bearing capacity of the anchorage interface

Based on the theoretical analysis mentioned above, with the known breaking load of the anchor rod, the ultimate anchorage force of the anchorage interface can be calculated theoreti-

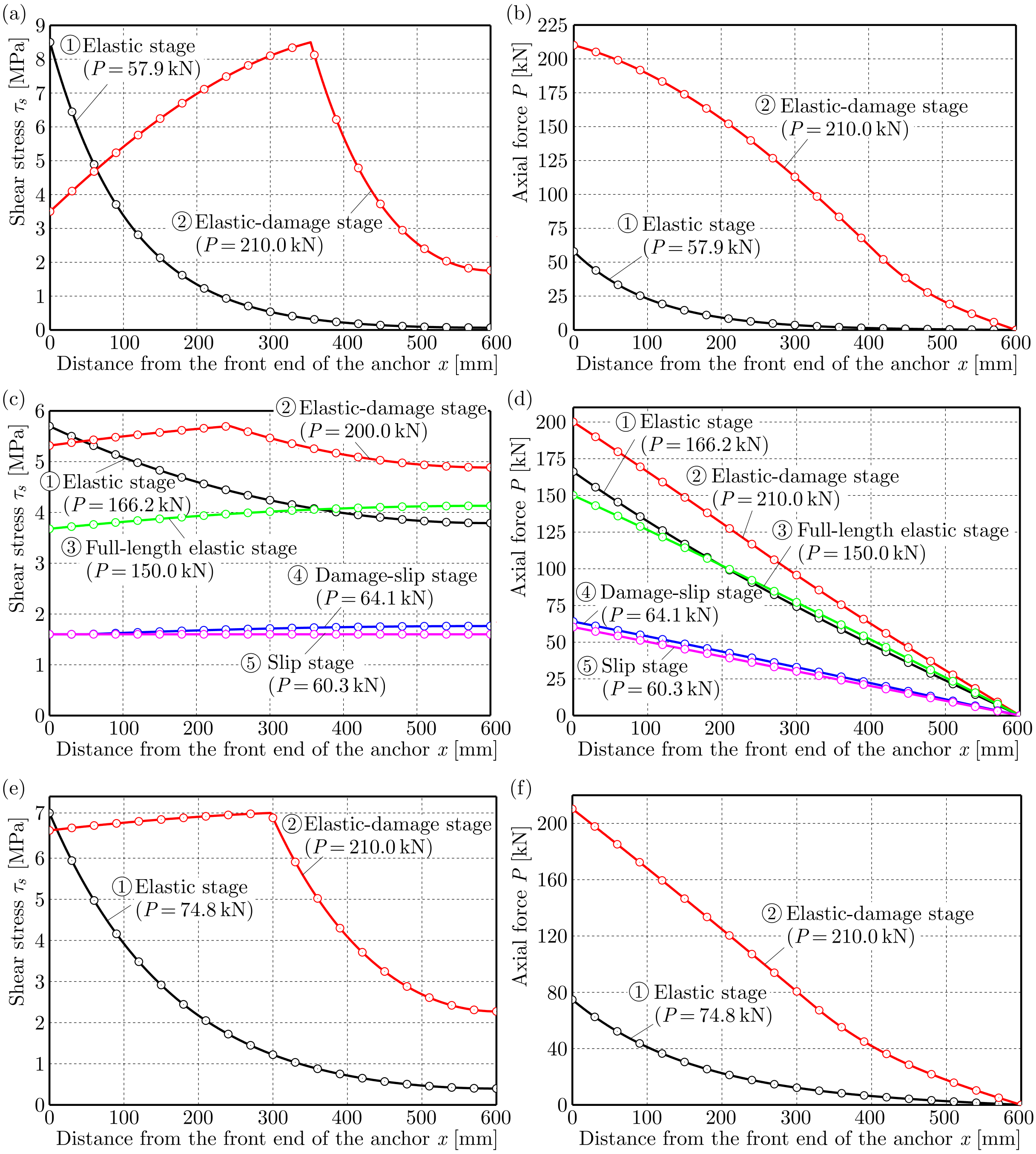


Fig. 9. Distribution curves of the shear stress and axial force under different anchorage agent conditions with anchorage length of 600 mm: (a)–(b) resin anchoring agent, (c)–(d) cement paste, (e)–(f) cement mortar

cally. According to the constitutive model parameters of the anchorage interface calibrated with different test data as shown in Fig. 8, using the aforementioned theoretical calculation method, the curve of the interface ultimate anchorage force changing with the anchorage length under different anchoring agent conditions can be calculated, as shown in Fig. 10. It can be seen from the figure that the relationship between the ultimate anchorage force of the anchorage interface and the anchorage length is related to the interface model parameters. Under the condition of resin anchoring agent, the ultimate anchorage force tends to approach a certain fixed value as the anchorage length increases. Under the conditions of cement grout and cement mortar anchoring, the ultimate anchorage force increases approximately linearly with the anchorage length.

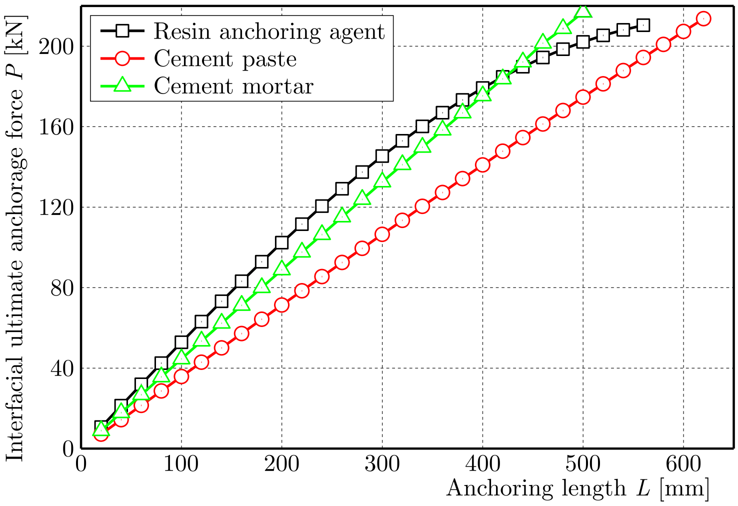


Fig. 10. Variation curves of interfacial ultimate anchorage force with anchorage length

At the same anchorage length, the interface ultimate anchorage force is generally smaller under the condition of cement grout anchoring; when the anchorage length is less than 420 mm, the interface ultimate anchorage force under the condition of resin anchoring agent is greater than that of cement mortar and cement grout anchoring at the same anchorage length. When the anchorage length is greater than 420 mm, the interface ultimate anchorage force with cement mortar is greater than that with resin anchoring agent at the same anchorage length. This indicates that when the anchorage length is relatively small, the magnitude of the interface ultimate anchorage force is determined by the bonding strength, and the greater the peak shear strength, the greater the interface ultimate anchorage force. However, when the anchorage length is relatively large, the contribution of residual stage friction to improving the interface bearing capacity increases, and at this time, the greater the residual shear strength, the greater the interface ultimate anchorage force.

4.2. Experimental verification of the ultimate bearing capacity of the anchorage interface

To verify the accuracy of the theoretical calculation results of the ultimate anchorage force of the anchorage interface, cement grout with a water-cement ratio of 0.45:1m was used as the anchoring agent. The anchor rod pull-out test scheme shown in Fig. 6 was adopted to conduct anchor rod tensile tests on anchor bodies with anchorage lengths of 300 mm, 400 mm, 500 mm, and 600 mm. The tensile load-displacement curves obtained from the test are shown in Fig. 11. The experimental and theoretical values of the ultimate anchorage force at different anchorage lengths, as well as the theoretical deviation rates, are listed in Table 3. The deviation rates are relatively small, indicating a high degree of conformity between the experimental and theoretical values. When the anchorage length is 500 mm and 600 mm, the anchor rod yields during the tensile process, and the experimental value of the interface ultimate anchorage force is less than the theoretically calculated value, with a theoretical deviation rate reaching 3.0%. This indicates that the yield strengthening of the anchor rod will lead to reduction in the bearing capacity of the anchorage interface.

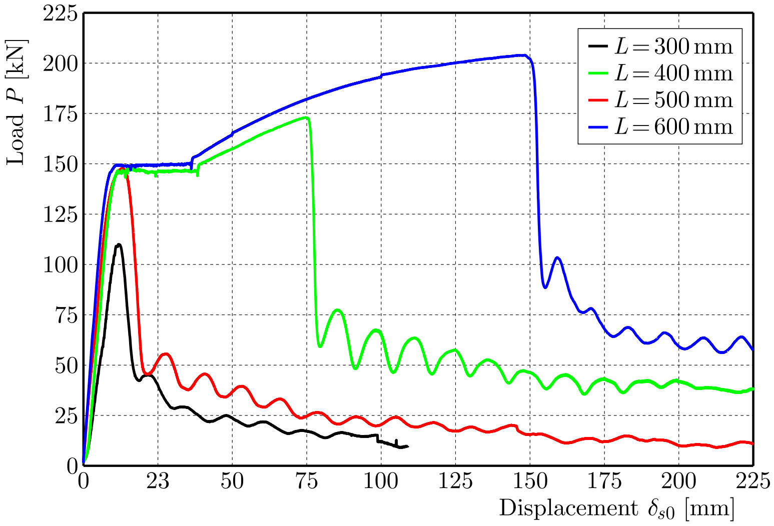


Fig. 11. Load-displacement curves with different anchorage lengths under the cement paste anchorage condition

Table 3. Experimental and theoretical values of the interfacial ultimate anchorage force

Anchorage length L [mm]	Experimental values [kN]	Theoretical values [kN]	Deviation rate of theoretical values [%]
300	108	107	−0.9
400	144	141	−2.1
500	170	175	2.9
600	201	207	3.0

5. Conclusions

(1) Based on the three-stage model, a constitutive equation for the evolution of shear stress damage in the anchorage interface element was established, and theoretical derivations were obtained for the expressions of shear stress and axial force distribution evolution at the anchorage interface during the entire process of anchor rod pull-out.

(2) Anchor rod pull-out tests were conducted to provide a method for calibrating the interface model parameters. Different anchoring agents have a significant impact on the interface model parameters. Compared with cement grout and cement mortar, applying the resin anchoring agent, the interface peak shear strength is larger, but the initial slip displacement and friction force during the slip stage are the smallest.

(3) The load transfer laws and bearing performance of the anchorage interface under the action of different anchoring agents were calculated and analyzed. Since the anchor rod has a certain bearing limit, when the theoretical calculated axial force exceeds the breaking load of the anchor rod, the anchor rod will fail by breaking. Therefore, the theoretical calculation and analysis method can predict the mode of anchorage failure. The greater the elastic shear stiffness of the anchorage interface, the more obvious the decay of shear stress and axial force in the elastic section of the interface along the anchorage length, and the smaller the axial force of the anchor rod when the interface begins to damage.

(4) The ultimate anchorage force of the interface at different anchorage lengths can be obtained through theoretical calculations. The test validation results show that the error between the experimental and theoretical value of the interface ultimate anchorage force is within 3%.

However, the yield strengthening of the anchor rod will reduce the anchorage bearing capacity, and the error of the theoretical calculation value will increase.

Acknowledgement

This research was supported by Taishan Scholar in Shandong Province of China (Grant No. tstp20221126), and the National Natural Science Foundation of China (Grant Nos. 52374097 and 52104137). The authors of this article are very grateful for the support provided.

References

1. Cai Y., Esaki T., Jiang Y.J., 2004, A rock bolt and rock mass interaction model, *International Journal of Rock Mechanics and Mining Sciences*, **41**, 7, 1055-1067
2. Cao C., Ren T., Cook C., 2014, Introducing aggregate into grouting material and its influence on load transfer of the rockbolting system, *International Journal of Mining Science and Technology*, **24**, 3, 325-328
3. Chen J.H., He F.L., Zhang S.B., 2020, A study of the load transfer behavior of fully grouted rock bolts with analytical modelling, *International Journal of Mining Science and Technology*, **30**, 1, 105-109
4. Chen J.H., Li D.Q., 2022, Numerical simulation of fully encapsulated rock bolts with a tri-linear constitutive relation, *Tunnelling and Underground Space Technology*, **120**, 104265.
5. Chen J.H., Liu P., Zhao H.B., Zhang C., Zhang J., 2021, Analytical studying the axial performance of fully encapsulated rock bolts, *Engineering Failure Analysis*, **128**, 105580
6. Du Y.L., Feng G.R., Kang H.P., Zhang Y.J., Zhang X.H., 2021, Effects of different pull-out loading rates on mechanical behaviors and acoustic emission responses of fully grouted bolts, *Journal of Central South University*, **28**, 7, 2052-2066
7. Fan J.Q., Shi X.Y., Kong F.L., Guo J.Q., 2021, Calculation method of the effective anchorage length of anchorage structures based on analysis of the ultimate bearing capacity, *Geotechnical and Geological Engineering*, **39**, 6, 4205-4212
8. Holý O., 2017, Evaluation of many load tests of passive rock bolts in the Czech Republic, *Geoscience Engineering*, **63**, 1, 1-7
9. Huang M.H., Zhao M.H., Chen C.F., 2019, Study of the mechanical features of the bolts under tension based on the nonlinear bond-slip model (in Chinese), *Journal of Safety and Environment*, **19**, 4, 1196-1203
10. Høien A.H., Li C.C., Zhang N., 2021, Pull-out and critical embedment length of grouted rebar rock bolts-mechanisms when approaching and reaching the ultimate load, *Rock Mechanics and Rock Engineering*, **54**, 1431-1447
11. Kilic A., Yasar E., Celik A.G., 2002, Effect of grout properties on the pull-out load capacity of fully-grouted rock bolt, *Tunnelling and Underground Space Technology*, **17**, 4, 355-362
12. Li C.C., 2010, Field observations of rock bolts in high stress rock masses, *Rock Mechanics and Rock Engineering*, **43**, 4, 491-496
13. Liu X.H., Yao Z.S., Xue W.P., Wang X.S., Huang X.W., 2021, Experimental study of the failure mechanism of the anchorage interface under different surrounding rock strengths and ambient temperatures, *Advances in Civil Engineering*, **2021**, 1, 622418
14. Ma S.Q., Zhao Z.Y., Nie W., Gui Y.L., 2016, A numerical model of fully grouted bolts considering the tri-linear shear bond–slip model, *Tunnelling and Underground Space Technology*, **54**, 73-80

15. Nemcik J., Ma S.Q., Aziz N., Ren T., Geng X., 2014, Numerical modelling of failure propagation in fully grouted rock bolts subjected to tensile load, *International Journal of Rock Mechanics and Mining Sciences*, **71**, 293-300
16. Nie W., Zhao Z.Y., Guo W., Shang J., Wu C., 2019, Bond-slip modeling of a CMC rock bolt element using 2D-DDA method, *Tunnelling and Underground Space Technology*, **85**, 340-353
17. Wang Y., Li C.D., Cai Z.L., Zhu G.Q., Zhou J.Q., Yao W.M., 2023, Mechanical behaviors of anchorage interfaces in layered rocks with fractures under axial loads, *Journal of Earth Science*, **34**, 2, 354-368
18. Wu Q.H., Chen L., Shen B.T., Dlamini B.G., Li S.Q., Zhu Y.J., 2019, Experimental investigation on rockbolt performance under the tension load, *Rock Mechanics and Rock Engineering*, **52**, 11, 4605-4618
19. Yahia A., Khayat K.H., Benmokrane B., 1998, Evaluation of cement grouts for embedding anchors under water, *Materials and Structures*, **31**, 4, 267-274
20. Yi K., Kang H.P., Ju W.J., Liu Y.D., Lu Z.G., 2020, Synergistic effect of strain softening and dilatancy in deep tunnel analysis, *Tunnelling and Underground Space Technology*, **97**, 103280
21. Yue Z.W., Li A., Wang P., Wang P., 2022, An analytical analysis for the mechanical performance of fully-grouted rockbolts based on the exponential softening model, *International Journal of Mining Science and Technology*, **32**, 5, 981-995
22. Zhao T.B., Xing M.L., Guo W.Y., Wang C.W., Wang B., 2021, Anchoring effect and energy-absorbing support mechanism of large deformation bolt, *Journal of Central South University*, **28**, 2, 572-581

Manuscript received October 2, 2024; accepted for publication December 6, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 165-175, Warsaw 2025
https://doi.org/10.15632/jtam-pl/196983

EFFECT OF MATERIAL PARAMETERS ON THE CUMULATIVE FATIGUE DAMAGE OF SONIC OSCILLATORS

Shengyu He, Changgen Bu, Xiuhua Zheng, Jing Xiao, Xiangzhou Kong
School of Engineering and Technology, China University of Geosciences, Beijing, China
corresponding author Changgen Bu, e-mail: bucg@cugb.edu.cn

The standing-wave vibration drill-string in a sonic drill is prone to fatigue damage. Commonly used alloy materials for drill-strings include S135, 4145H and G105, but the material sensitivity coefficient is large when the fatigue strength is high. To choose an appropriate drill-pipe material and improve the drill-string service life, this paper analyzes the cumulative fatigue damage of three types of material drill-strings with variable length. The results indicate that for short drill-strings with high-frequency vibrations, the material constant has a greater effect on fatigue damage. However, for long drill-strings with low-frequency vibrations, the fatigue limit has a greater impact on fatigue damage.

Keywords: sonic oscillators, sonic drilling, cumulative fatigue damage, standing wave vibration, material parameters

Nomenclature

c – damping coefficient
d_i, d_o – inner and outer diameter of drill string, respectively
$f(x,t)$ – distributed load
k – number of additional drill pipes
k_{max} – total number of additional drill pipes within fatigue damage zone
l – length of drill string
l_b – length of drill string without fatigue damage
l_0 – maximum drill string length before entering standing wave vibration
m – material constant
m_e – static moment of sonic vibrator
m_H, m_G, m_S – material constant of 4145H, G105, S135, respectively
n_k – number of cycles under stress level σ_{-1k}
$u(x,t)$ – displacement response of drill string
x_f – distance from any point on drill string to drill bit
D – cumulative fatigue damage
D_H, D_G, D_S – fatigue damage of 4145H, G105, S135, respectively
E – elastic modulus
N_C, N_k – total number of cycles under stress level σ_{-1} and σ_{-1k}, respectively
R_{HG}, R_{SG} – damage ratio at maximum damage point of 4145H and G105; S135 and G105, respectively
S – cross-sectional area of drill string
V – rate of penetration (ROP)
Δl – length of single drill pipe

ζ – damping ratio
ρ – density
$\varepsilon(x,t), \sigma(x,t)$ – dynamic strain and stress, respectively
σ_b – tensile strength
σ_{-1} – symmetrical cyclic fatigue limit stress
σ_{-1k} – stress level
φ_i – phase angle
ω_i – i-th order natural angular frequency

1. Introduction

In the context of geological and oil drilling operations, a drill tool is subjected to tension, bending, torsion, impact, and hydraulic pressure in the wellbore. These complex stress conditions are a primary cause of fatigue failure in drill strings (Albdiry and Almensory, 2016). Research has shown that over 50% of drill-string failures are caused by fatigue (Moradi and Ranjbar, 2009). Constant axial, torsional, and internal pressure loads do not cause fatigue, but they can exacerbate fatigue under the influence of cyclic stress (Asgharzadeh *et al.*, 2019). Drill strings fatigue damage is a major concern for researchers worldwide.

Recently, sonic (or ultrasonic) vibration technology has rapidly developed and been applied in the field of drilling. Sonic (ultrasonic) drilling is an efficient drilling method that utilises resonant waves in a drill string to generate higher speeds in the drill bit, thereby achieving high-speed drilling. It has the advantages of high drilling speed, high fidelity of rock samples, and low environmental pollution and is widely used in fields such as engineering exploration and infrastructure construction. However, the standing wave resonance in drill strings can cause extreme dynamic stresses at specific locations within the drill string, which can lead to fatigue failure or even fracture of the drill string, increasing drilling costs (Zamani *et al.*, 2016).

Fatigue failure is the process of fatigue damage accumulation reaching a critical value. Numerous theoretical models have been developed to describe the development of cumulative damage (Benkabouche *et al.*, 2015). In engineering applications, the Miner cumulative fatigue damage theory is notable for its simplicity and practical utility. It often aligns well with experimental results and is adept at predicting the average fatigue life of engineering structures under random loads (Sun *et al.*, 2014). This theory is widely applied in predicting the fatigue life of various mechanical components in engineering, such as aircraft engines, skin, and hydraulic tubes and other components (Baek *et al.*, 2008; Chen *et al.*, 2014; Jiao *et al.*, 2018). Rao *et al.* (2001) reported that the Palmgren–Miner law is the most used theory for predicting the fatigue life in blades subjected to variable stress amplitudes. Zhao *et al.* (2018) employed the rainflow-counting method to analyze continuous bending stress history in whirl responses, and evaluated the cumulative fatigue damage using Miner's rule. Additionally, Zhu *et al.* (2012) indicated that the Palmgren–Miner rule provides a better lifetime prediction under simple loading conditions, as compared to other methods. Bu *et al.* (2023) theoretically investigated the fatigue damage in sonic drilling caused by inertial excitation, and conducted a fatigue damage theory for deep-hole sonic drilling by using Miner's cumulative fatigue damage rule.

Research on fatigue experiments has mainly focused on fatigue damage analyzes of material samples. The metal magnetic memory method can be employed to quantitatively assess material damage based on the characteristics of metal magnetic memory changes during stress or fatigue of different materials. Many studies have been conducted on the processes of tensile fatigue and rotary-bending fatigue (Li *et al.*, 2013, 2016). Lin *et al.* (2015) conducted fatigue performance tests on commonly used API standard S135 and G105 materials by employing a PQ-6 rotary-bending fatigue-testing machine. By utilizing the Basquin model and mathematical statistics theory, Lin obtained the S-N curves for these materials. Wang *et al.* (2016) using the same method as Lin *et al.* (2015), obtained the S-N curves for 4145H, 4145H is international standard drilling tool steel certified by the American Iron and Steel Institute (AISI). Furthermore, it was found that the frequency of ultrasonic testing did not impact the fatigue life in tests of asymmetric cyclic loads with constant- and variable-amplitude loads (Mayer *et al.*, 2013; Fitzka and Mayer, 2016). Experimental studies on material fatigue damage have only focused on a specific load. However, during sonic drilling, as the borehole extends, the excitation frequency decreases, and the location of the maximum dynamic stress changes. Therefore, the dynamic stress experienced by each point in the drill string is different, and each point in the drill string is subjected to varying amplitude loads (Bu *et al.*, 2015). Therefore, fatigue damage analysis and research on resonant drill strings in sonic drilling should not only focus on the material itself but also be combined with the actual working conditions.

In our previous research, we developed a theory on fatigue damage of sonic drill strings, which theoretically proved the primary role of resonant order stress in causing fatigue damage, and our findings suggested that selecting materials with a high fatigue limit and low material constant (index of σ-N curve) can reduce the damage of drill strings (Bu *et al.*, 2023). However, the influence of materials on the fatigue life was not theoretically analyzed and the fatigue life of different materials under the same stress cycle was not evaluated; thus, whether a larger fatigue limit leads to a longer fatigue life is still unknown. Compared to G105 materials, S135 materials have a higher fatigue limit and a higher material constant, which is more sensitive to the stress cycle. The material constant and fatigue limit of 4145H material are both between S135 and G105. Therefore, during construction, both the fatigue limit and the material constant should be considered. For shallow sonic drills using piezoelectric ceramics or giant magnetostrictive materials as exciters (Bu *et al.*, 2022) or ultrasonic knives used in the medical field, the excitation frequency is high (up to 20–100 kHz). In such cases, considering the impact of material constants on the fatigue life is crucial.

This study compares the fatigue damage of S135, 4145H and G105 drill strings, which are commonly used in drilling projects, by applying the Miner cumulative fatigue damage rule combined with the standing-wave vibration characteristics of sonic drilling. Furthermore, the fatigue damage of three types of drill strings having the same length is analyzed, and the application range of the three materials is determined. This study is important for guiding engineering practices on choosing appropriate drill pipe materials, reducing drilling accidents and enhance the service life of sonic drill strings.

2. Methodology

2.1. Stress analysis of sonic drill string

In sonic drilling, a sonic vibrator is installed at the top of the drill to induce vibrations in the drill string. This vibrator is isolated using air springs, ensuring it does not participate to the resonance of the drill string (Xiao *et al.*, 2020). According to our previous work (Bu *et al.*, 2023), the differential equation to describe the longitudinal vibration of a sonic drill string is

$$\rho S\frac{\partial^2 u}{\partial t^2} + c\frac{\partial u}{\partial t} - ES\frac{\partial^2 u}{\partial x^2} = f(x,t) \tag{2.1}$$

where ρ is the material density, S is the cross-sectional area of the drill string, c is the damping coefficient, E is the elastic modulus, and $f(x,t) = m_e\omega^2\sin(\omega t)\delta(x)$, $\delta(x)$ is the Dirac delta function, $\int_0^l f(x,t)\,dx = m_e\omega^2\sin(\omega t)$, m_e is the static moment of the sonic vibrator.

The boundary condition is expressed as

$$\frac{\partial u(0,t)}{\partial x} = \frac{\partial u(l,t)}{\partial x} = 0 \tag{2.2}$$

The vibration differential equation and boundary conditions of the sonic drill string are obtained by applying Eqs. (2.1) and (2.2), and the steady-state solution of the forced vibration displacement response of the drill string is obtained by using the separated variable method (Sun *et al.*, 2017)

$$u(x,t) = \frac{2m_e\omega^2}{\rho Sl}\sum_{i=0}^{\infty}\frac{\sin\omega\, t - \varphi_i}{\sqrt{(\omega_i^2-\omega^2)^2 + (2\zeta_i\omega_i\omega)^2}}\cos\frac{i\pi x}{l} \tag{2.3}$$

where $\omega_i = (i\pi/l)\sqrt{(E/\rho)}$ $(i = 1,2,3,...)$ is the natural angular frequency, $\varphi_i = \arctan((2\zeta_i\omega_i\omega)/(\omega_i^2 - \omega^2))$ is the phase angle, and ζ_i is the mode shape damping ratio (it is usually assumed that the damping ratio of each mode is the same; therefore, $\zeta_i = \zeta$). When the excitation angular frequency ω approaches ω_i, the phase angle φ_i approaches $\pi/2$.

Then, the stress expression can be obtained as follows

$$\sigma(x,t) = E\varepsilon(x,t) = E\frac{\partial u(x,t)}{\partial x} = -\frac{2\pi E m_e\omega^2}{\rho S l^2}\sum_{i=0}^{\infty}\frac{i\sin\omega t - \varphi_i}{\sqrt{(\omega_i^2-\omega^2)^2 + (2\zeta_i\omega_i\omega)^2}}\sin\frac{i\pi x}{l} \tag{2.4}$$

The material parameters of S135, 4145H and G105 drill pipes are listed in Table 1.

Table 1. Material parameters of drill pipes (Lin *et al.*, 2015; Wang *et al.*, 2016)

Drill pipe material	Fatigue limit σ_{-1} [MPa]	Tensile strength σ_b [MPa]	Material constant m	Density ρ [kg/m^3]	Elastic modulus E [Pa]
S135	527.37	1000	16.53	7850	2.06e11
4145H	482.89	1100	12.81	7850	2.06e11
G105	435.62	793	9.77	7850	2.06e11

The technical parameters of the sonic vibrator and drill pipe are presented in Table 2.

Table 2. Technical parameters of the sonic vibrator and drill pipe (Sun *et al.*, 2017; Bu *et al.*, 2023)

Total static moment of sonic vibrator m_e	Drill-string inner/outer diameter d_i/d_o	Drill-string cross-sectional area S
0.126 kg · m	92.46/114.3 mm	3.547e-3 m^2

Table 1 indicates that the three materials have the same density and elastic modulus. Then, according to Eq. (2.4), the dynamic stress produced by standing wave vibration is the same for drill strings having the same length. We can determine the relationship between the maximum dynamic stress amplitude of the first-order resonant sonic drill string and its length by substituting the technical parameters listed in Table 2 into Eq. (2.4), as shown in Fig. 1.

Figure 1 illustrates that as the length of the drill string increases, its maximum dynamic stress amplitude gradually decreases, and the peak dynamic stress of the first-order resonance is always located in the middle of the drill string. Fatigue damage occurs when the dynamic stress amplitude of the drill string exceeds the fatigue limit. Due to the highest fatigue limit of the S135 material, the S135 drill pipe enters the non-damage zone earliest with the extension of the drill string.

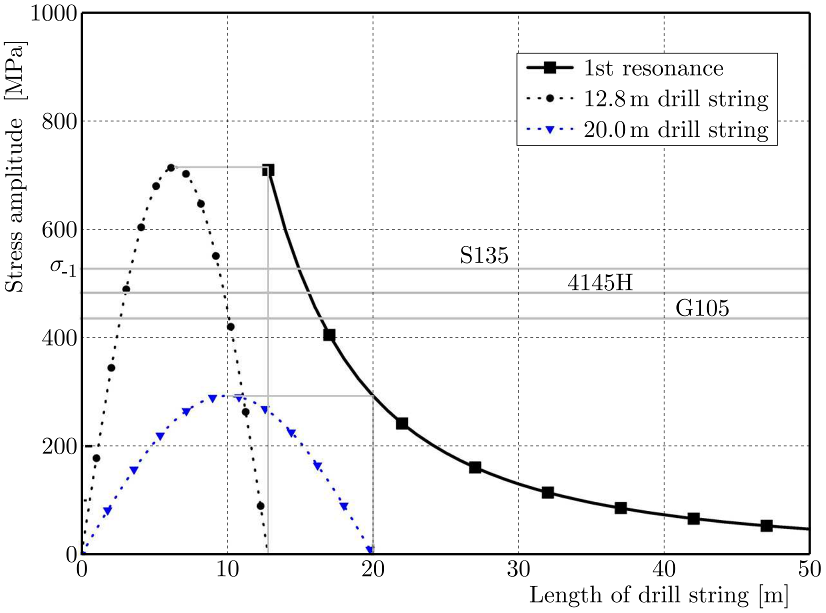


Fig. 1. Maximum stress amplitude of sonic drill strings having different lengths. The black dashed line represents the dynamic stress amplitude at each point of the 12.8 m drill string, and the blue dashed line represents the dynamic stress amplitude at each point of the 20 m drill string

2.2. Cumulative fatigue damage of the sonic drill string

When the influence of various dynamic stresses in the drill string on the fatigue life is analyzed, the fatigue damage of the drill string can be considered caused by the resonant dynamic stress. According to Miner's cumulative fatigue damage theory, $D = \sum(n_k/N_k)$, where n_k is the number of cycles under the stress level σ_{-1k} $(k = 1, 2, 3, ...)$, N_k is the total number of cycles that the test piece can withstand under the stress level σ_{-1k}, and the fatigue damage expression for the drill string (Bu *et al.*, 2023) can be obtained as

$$D = \frac{\pi^m r^{m+1}}{2N_C(\sigma_{-1})^m}\Big(\sqrt{\frac{E}{\rho}}\Big)^{2m+1}\Big(\frac{m_e}{S}\Big)^m \frac{1}{V\zeta^m}\sum_{k=1}^{k_{max}}\Big[\frac{\Delta l}{(l_0+k\Delta l)^{2m+1}}\sin^m\Big(\frac{\pi x_f}{l_0+k\Delta l}\Big)\Big] \tag{2.5}$$

where V is the rate of penetration (ROP), σ_{-1} is the symmetrical cyclic fatigue limit, N_C is the total number of cycles under the stress level σ_{-1}, m is the material constant, x_f is the distance from any point on the drill string to the drill bit, l_0 is the maximum length of the drill string before entering the standing wave vibration, Δl is the length of a single drill pipe, k is the number of additional drill pipes, and k_{max} is the total number of additional drill pipes within the fatigue damage zone, which is affected by the excitation frequency and the material fatigue limit.

As shown in Table 1, the S135, 4145H and G105 materials have the same elastic modulus and density. Therefore, based on Eq. (2.5), the factors affecting the damage between the two materials are the material fatigue limit and the material constant. The two parameters have opposite effects on the fatigue damage, with a high fatigue limit resulting in low fatigue damage, and a high material constant resulting in high fatigue damage. Table 1 reveals that the S135 material has the highest fatigue limit and highest material constant; so, the S135 material is most sensitive to stress cycles. Therefore, drill-string materials should be selected for more specific operating conditions to obtain a longer fatigue life.

3. Results and discussion

According to the foregoing analysis, the total cumulative fatigue damage during the drilling process was obtained based on the stress cycle times and fatigue life of the drill string at different stress levels. In this Section, the cumulative fatigue damage of the drill string for all materials listed in Table 1 are compared by setting continuous drilling conditions under different excitation frequencies. Additionally, the damage ratios of the fixed-length drill string oscillator are analyzed at different frequencies to explore the effects of the material fatigue limit and material constants on damage.

3.1. Fatigue damage comparison of variable-length sonic drilling

In sonic drilling, the extension of the drill string leads to a continuous change in force distribution. The static moment of the sonic vibrator and geometric parameters of the drill string are shown in Table 2, and the working conditions (damping ratio $\zeta = 0.025$ and ROP = 15 m/h) are set. Based on setting the length of a single drill pipe to $\Delta l = 0.1$ m (equivalent to the continuous extension of the drill string), the fatigue damages of S135, 4145H and G105 drill strings due to the standing wave resonance at initial excitation frequencies of 180 and 206 Hz are calculated.

When the initial excitation frequency is 180 Hz, the sonic standing wave begins at a drill string length of 14.2 m. Figure 2 shows the cumulative fatigue damages of the S135, 4145H and G105 drill strings as the drilling proceeded until no further fatigue damage occurred.

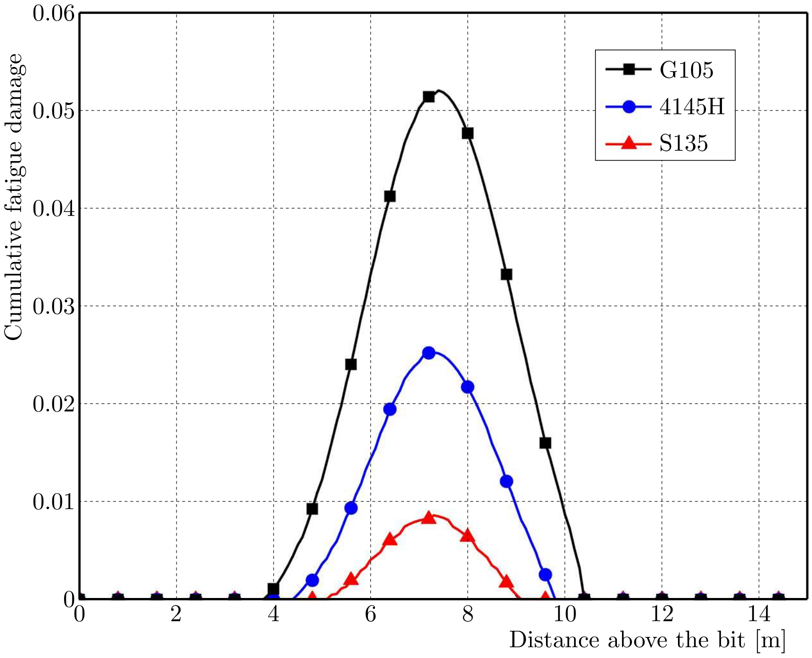


Fig. 2. At the initial excitation frequency of 180 Hz, the starting length of the sonic standing wave is 14.2 m; drilling is performed to the point where fatigue damage no longer occurs, and the cumulative fatigue damages at various points of S135, 4145H and G105 drill strings are shown (drilling conditions: $\zeta = 0.025$, $\Delta l = 0.1$ m, and ROP = 15 m/h)

Figure 2 shows that the S135 drill pipe, which has the highest fatigue limit, has lowest damage at all points; the G105 drill pipe, which has the lowest fatigue limit, has highest damage at all points; the fatigue limit and the fatigue damage of 4145H drill pipe are both between G105 and S135.

This result supports the idea that the higher the material fatigue limit, the lower the fatigue damage and the longer the lifespan. When the maximum dynamic stress amplitude of the drill string reaches the fatigue limits of S135, 4145H and G105, respectively, the specific excitation frequencies determined by Eq. (2.4) are 171.9 Hz, 164.5 Hz and 156.2 Hz, respectively. Notably, if the initial excitation frequency is lower than 156.2 Hz, the maximum dynamic stresses of these three materials are less than the fatigue limits, and none of them will experience fatigue damage. Therefore, these three materials can be selected.

In the case of long drill string at low frequencies, materials with a high fatigue limit exhibit low fatigue damage. However, for shallow sonic standing-wave drilling or drilling with a shorter drill string, the drilling string has a higher excitation frequency and undergoes more stress cycles per unit time. According to the data in Table 1, S135 material with the highest fatigue limit has the same highest material constant and is most sensitive to dynamic stress cycling. Therefore, the use of S135 for high-frequency short drill strings requires further investigation.

The starting length of the sonic standing wave is 12.4 m when the initial excitation frequency increases to 206 Hz. Drilling is performed to the point where fatigue damage no longer occurs, and the cumulative fatigue damages of the S135, 4145H and G105 drill strings are shown in Fig. 3.

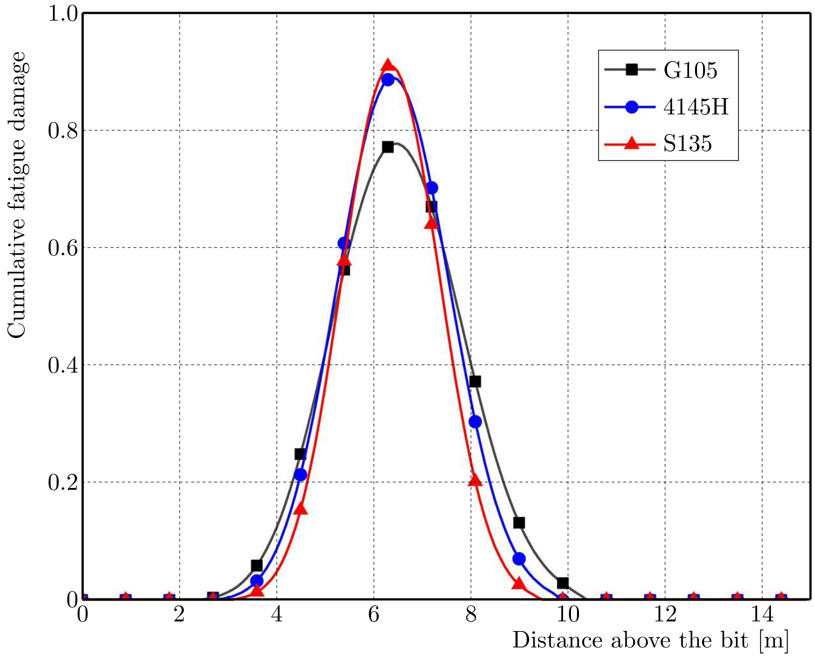


Fig. 3. At the initial excitation frequency of 206 Hz, the starting length of the sonic standing wave is 12.4 m; drilling is performed to the point where fatigue damage no longer occurs, and the cumulative fatigue damages at various points of S135, 4145H and G105 drill strings are shown (drilling conditions: $\zeta = 0.025$, $\Delta l = 0.1$ m, and ROP = 15 m/h)

As shown in Fig. 3, the maximum damage points of these three materials are near the distance from the drill bit of half the starting length of the sonic standing wave. The maximum cumulative fatigue damage of S135 is the highest, while the maximum cumulative fatigue damage of G105 is the lowest, and the maximum cumulative fatigue damage of 4145H is between G105 and S135. This result is inconsistent with the conclusion drawn from Fig. 2. If the starting length of the sonic standing wave decreases and the standing-wave excitation frequency increases, the fatigue damage to the S135 drill string becomes greater, potentially leading to its failure. This is because fatigue damage is not only related to the fatigue limit but also significantly influenced by the material constant. The higher-frequency excitation causes the dynamic stresses in the symmetric cycle in the drill string to change faster, and the S135 material has the highest constant and is most sensitive to stress changes; therefore, the maximum damage is highest in these three materials.

A comparison of the two types of special cases presented in Figs. 2 and 3 reveal that when drilling with a low-frequency and long drill string, the greater the fatigue limit, the smaller is the fatigue damage; when drilling with high-frequency short drill strings, the smaller the material coefficient, the smaller is the fatigue damage. This also indicates that under different excitation frequencies, the fatigue limit and material constant have different dominant effects on the fatigue life of the drill strings. Therefore, it is not feasible to evaluate the cumulative fatigue damage of the drill string based on a single parameter only, and a combination of fatigue limits and material constants needs to be considered.

3.2. Effect of standing wave oscillator length on the fatigue damage ratio of two materials

In contrast to the special case analysis of the standing-wave vibration drill string previously described, a more general case is discussed. The theory of cumulative fatigue damage means that when the dynamic stress is below the fatigue limit of the material, damage will no longer occur (Miner, 1945). According to our previous study, the excitation frequency and dynamic stress gradually decrease when the length of the drill string varies from short to long (Bu *et al.*, 2023). The formula for calculating the length of drill-string for which the standing wave resonance no longer causes fatigue damage is

$$l_b = \sqrt{\frac{\pi r m_e}{\zeta \sigma_{-1} S}} \sqrt{\frac{E}{\rho}} \tag{3.1}$$

The lengths of the three drill strings without fatigue damage are obtained as 14.9 m, 15.6 m and 16.4 m by substituting the material parameters of S135, 4145H and G105 presented in Table 1 into Eq. (3.1). Therefore, the analysis should be performed when the length of the drill-string oscillator is less than 14.9 m.

A fixed length of the drill string oscillator is set to analyze the fatigue damage of the three materials. To compare the extent of damage to the three materials more intuitively, the material parameters of S135, 4145H and G105 are introduced into Eq. (2.5), and the fatigue damage of the three materials can be obtained as D_S, D_H and D_G; therefore, taking D_G as the denominator, D_S and D_H as the numerators, the damage ratios can be obtained as

$$\begin{aligned} \frac{D_S}{D_G} &= \Big(\frac{\pi r m_e E}{\zeta \rho S l^2} \sin\Big(\frac{\pi x_f}{l}\Big)\Big)^{m_S - m_G} \frac{(\sigma_{-1G})^{m_G}}{(\sigma_{-1S})^{m_S}} \\ \frac{D_H}{D_G} &= \Big(\frac{\pi r m_e E}{\zeta \rho S l^2} \sin\Big(\frac{\pi x_f}{l}\Big)\Big)^{m_H - m_G} \frac{(\sigma_{-1G})^{m_G}}{(\sigma_{-1H})^{m_H}} \end{aligned} \tag{3.2}$$

where subscripts S, H and G represent the S135, 4145H and G105 materials, respectively. When $\sin(\pi x_f/l) = 1$, the damage ratios at the maximum damage point of drill strings are obtained as

$$\begin{aligned} R_{SG} &= \frac{D_{S\,max}}{D_{G\,max}} = \Big(\frac{\pi r m_e E}{\zeta \rho S l^2}\Big)^{m_S - m_G} \frac{(\sigma_{-1G})^{m_G}}{(\sigma_{-1S})^{m_S}} \\ R_{HG} &= \frac{D_{H\,max}}{D_{G\,max}} = \Big(\frac{\pi r m_e E}{\zeta \rho S l^2}\Big)^{m_H - m_G} \frac{(\sigma_{-1G})^{m_G}}{(\sigma_{-1H})^{m_H}} \end{aligned} \tag{3.3}$$

For the three materials under the same working conditions, all the parameters in Eqs. $(3.3)_1$ and $(3.3)_2$ are the same, except for the material constant, m, and the fatigue limit, σ_{-1}, with only one variable, l. Therefore, the damage ratios are a function of the drill-string oscillator length, l.

The functional relationship between the fatigue damage ratio at the maximum damage points of the three drill strings and the oscillator length can be obtained by substituting the parameters in Table 2 into Eqs. $(3.3)_1$ and $(3.3)_2$, as shown in Fig. 4.

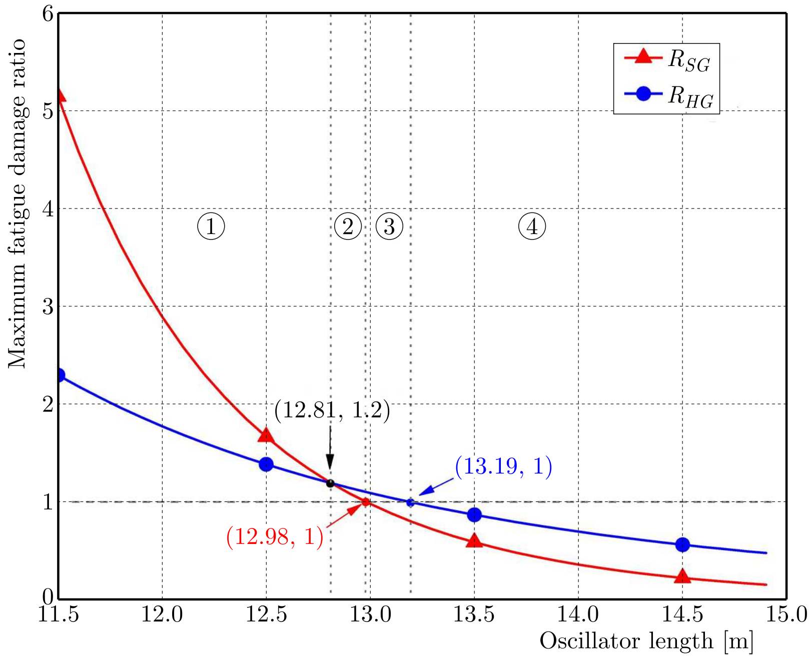


Fig. 4. Relationship between the length of the drill-string standing-wave oscillator and the maximum fatigue ratios; the damping ratio is set to 0.025

Figure 4 depicts that as the oscillator length increases, the damage ratio R_{SG} (red curve) decreases more sharply than R_{HG} (blue curve). Because the materials fatigue limit and material constant are fixed values, when the oscillator length l increases, $1/l^2$ becomes smaller, and the damage ratios R_{SG} and R_{HG} are proportional to $(1/l^2)^{m_S-m_G}$ and $(1/l^2)^{m_H-m_G}$, respectively. Meanwhile $m_S - m_G > m_H - m_G > 1$, thus R_{SG} decreases more sharply than R_{HG}.

As the oscillator length varies, the relationship between the maximum fatigue damage of these three materials also changes. According to Eqs. $(3.3)_1$ and $(3.3)_2$, when $D_{S\,max} = D_{H\,max}$, the damage ratio $R_{SG} = R_{HG} = 1.2$, $l = 12.81\,\mathrm{m}$. Define $l \leqslant 12.81\,\mathrm{m}$ as the region ① (shown in Fig. 4) where $R_{SG} \geqslant R_{HG} > 1$ (i.e. $D_{G\,max} < D_{H\,max} \leqslant D_{S\,max}$). When $R_{SG} = 1$ or $R_{HG} = 1$, $l = 12.98\,\mathrm{m}$ or $l = 13.19\,\mathrm{m}$; with $l = 12.98\,\mathrm{m}$ and $l = 13.19\,\mathrm{m}$ as regions boundaries, define regions ② $12.81\,\mathrm{m} < l \leqslant 12.98\,\mathrm{m}$, ③ $12.98\,\mathrm{m} < l \leqslant 13.19\,\mathrm{m}$, and ④ $13.19\,\mathrm{m} < l < 14.90\,\mathrm{m}$. The relationships between the maximum fatigue damage of materials in all four regions ①, ②, ③ and ④ can be summarized as

$$\begin{array}{ll} \text{①} : D_{G\,max} < D_{H\,max} \leqslant D_{S\,max} & l \leqslant 12.81\,\mathrm{m} \\ \text{②} : D_{G\,max} \leqslant D_{S\,max} < D_{H\,max} & 12.81\,\mathrm{m} < l \leqslant 12.98\,\mathrm{m} \\ \text{②} : D_{S\,max} < D_{G\,max} \leqslant D_{H\,max} & 12.98\,\mathrm{m} < l \leqslant 13.19\,\mathrm{m} \\ \text{④} : D_{S\,max} < D_{H\,max} < D_{G\,max} & 13.19\,\mathrm{m} < l < 14.90\,\mathrm{m} \end{array} \tag{3.4}$$

From Eq. (3.4), it can be seen that the G105 oscillator experiences the lowest damage when the oscillator length is less than 12.98 m. Therefore, for a shallow high-frequency sonic drilling under piezoelectric ceramic or giant magnetostrictive excitation, materials with low material constants should be selected to reduce their sensitivity to stress changes and fatigue damage. However, the S135 oscillator experiences the lowest damage when the oscillator length is greater than 12.98 m. Therefore, for long drill strings with a low-frequency excitation, materials with a high fatigue limit should be selected to extend the fatigue life of the drill string.

The best case for selecting drill string materials is when the material has both a high fatigue limit and a small material constant, such that the standing-wave vibration drill string can have a high fatigue life in both high- and low-frequency operating ranges. However, the cost of these materials must be considered in engineering applications. In the case of commonly used alloy materials, simultaneously satisfying these two conditions is challenging. Therefore, an appropriate drill pipe material must be selected depending on different construction conditions.

4. Conclusions

Based on the Miner cumulative fatigue damage theory combined with the characteristics of a standing-wave vibration sonic drill string under variable frequency and amplitude dynamic stress, the fatigue damage of S135, 4145H and G105 materials commonly used in drilling construction is compared, and the following conclusions are drawn.

When the impact of materials on damage is examined, analyzing only a single material parameter is not sufficient. The fatigue limit and material constant must be comprehensively considered, and the length of the drill string and the excitation frequency influence the material selection. As the drill-string length increases, the frequency of the standing wave excitation decreases, the effect of the material constant on the fatigue damage gradually diminishes, and the impact of the fatigue limit on fatigue damage increases. For high-frequency short drill strings (oscillators), drill pipes with low material constants should be selected to reduce drill-string fatigue damage. For low-frequency long drill strings (oscillators), drill pipes with a high fatigue limit should be selected to reduce fatigue damage. This conclusion provides a practical value and theoretical guidance for selection of drill pipe materials in production.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 51775523), Science and Technology Innovation Special Project of Xiongan New Area, Ministry of Science and Technology of the People's Republic of China (No. 2022XAGG0500).

References

1. Albdiry M.T., Almensory M.F., 2016, Failure analysis of drillstring in petroleum industry: A review, *Engineering Failure Analysis*, **65**, 74-85
2. Asgharzadeh A., Meza O.G., Bannenberg M., Paz C., Holzmann J., Oppelt J., 2019, New concept for multiaxial fatigue rigs in reconstructing combined loads to investigate fatigue life of casing connections in casing drilling applications, *Offshore Mediterranean Conference and Exhibition*, OMC-2019-0856
3. Baek S.H., Cho S.S., Joo W.S., 2008, Fatigue life prediction based on the rainflow cycle counting method for the end beam of a freight car bogie, *International Journal of Automotive Technology*, **9**, 1, 95-101
4. Benkabouche S., Guechichi H., Amrouche A., Benkhettab M., 2015, A modified nonlinear fatigue damage accumulation model under multiaxial variable amplitude loading, *International Journal of Mechanical Sciences*, 100, 180-194
5. Bu C.G., Sun L., Hu Y.B., Xia B.R., 2015, Research on flexible drill string vibration induced by sonic harmonic excitation, *Transactions of the Canadian Society for Mechanical Engineering*, **39**, 2, 281-291
6. Bu C.G., Tang X.Z., Wu W.G., 2022, A shallow subsurface magnetostrictive-driven sonic wave drill (in Chinese), Chinese Patent: CN202210343686.9

7. Bu C.G., Xiao J., He S.Y., Wiercigroch M., 2023, Theoretical study on fatigue damage of sonic standing wave resonant drill-string, *Journal of Petroleum Science and Engineering*, **220**, Part A, 111160
8. Chen H.X., Chen Y.X., Yang Y., 2014, A fatigue and low-energy shock-based approach to predict fatigue life, *Journal of Mechanical Science and Technology*, **28**, 10, 3977-3984
9. Fitzka M., Mayer H., 2016, Constant and variable amplitude fatigue testing of aluminum alloy 2024-T351 with ultrasonic and servo-hydraulic equipment, *International Journal of Fatigue*, **91**, Part 2, 363-372
10. Jiao R., He X.F., Li Y.H., 2018, Individual aircraft life monitoring: An engineering approach for fatigue damage evaluation, *Chinese Journal of Aeronautics*, **31**, 4, 727-739
11. Li C.C., Dong L.H., Wang H.D., Li G.L., Xu B.S., 2016, Metal magnetic memory technique used to predict the fatigue crack propagation behavior of 0.45%C steel, *Journal of Magnetism and Magnetic Materials*, **405**, 150-157
12. Li J.W., Zhong S., Lv G.P., Wang Z.G., Tian C., 2013, The variation of surface magnetic field induced by fatigue stress, *Journal of Nondestructive Evaluation*, **32**, 3, 238-241
13. Lin Y.H., Li G.H., Hu Q., Liu W.Y., 2015, Experimental study on drill pipe stress-fatigue life curve (in Chinese), *Petroleum Drilling Techniques*, **43**, 4, 124-128
14. Mayer H., Schuller R., Fitzka M., 2013, Fatigue of 2024-T351 aluminium alloy at different load ratios up to 10^{10} cycles, *International Journal of Fatigue*, **57**, 113-119
15. Miner M.A., 1945, Cumulative damage in fatigue, *Journal of Applied Mechanics*, **12**, 3, 159-164
16. Moradi S., Ranjbar K., 2009, Experimental and computational failure analysis of drillstrings, *Engineering Failure Analysis*, **16**, 3, 923-933
17. Rao J.S., Pathak A., Chawla A., 2001, Blade life: A comparison by cumulative damage theories, *Journal of Engineering for Gas Turbines and Power*, **123**, 4, 886-892
18. Sun L., Bu C.G., Hu P.D., Xia B.R., 2017, The transient impact of the resonant flexible drill string of a sonic drill on rock, *International Journal of Mechanical Sciences*, **122**, 29-36
19. Sun Q., Dui H.N., Fan X.L., 2014, A statistically consistent fatigue damage model based on Miner's rule, *International Journal of Fatigue*, **69**, 16-21
20. Wang Q., Lin Y.H., 2016, Effect of nitriding temperature on fatigue life of 4145 steel in QPQ salt bath composite heat treatment (in Chinese), *Technol. Innov. Appl.*, **33**, 23-24
21. Xiao J., Bu C.G., Hu Y.B., He S.Y., 2020, Influence of sonic vibrator mass on the modal frequency of the drill string, *Transactions of the Canadian Society for Mechanical Engineering*, **44**, 1, 65-71
22. Zamani S.M., Hassanzadeh-Tabrizi S.A., Sharifi H., 2016, Failure analysis of drill pipe: A review, *Engineering Failure Analysis*, **59**, 605-623
23. Zhao D., Hovda S., Sangesland S., 2018, Whirl simulation of drill collar and estimation of cumulative fatigue damage on drill-collar connection, *SPE Journal*, **23**, 2, 286-300
24. Zhu S.P., Huang H.Z., Liu Y., He L.P., Liao Q., 2012, A practical method for determining the Corten-Dolan exponent and its application to fatigue life prediction, *International Journal of Turbo and Jet Engines*, **29**, 2, 79-87

Manuscript received July 4, 2024; accepted for publication December 6, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 177-195, Warsaw 2025
https://doi.org/10.15632/jtam-pl/194881

RESPONSE ANALYSIS AND CONTROL OF COMPOSITE THIN PLATE WITH PIEZOELECTRIC ACTUATORS BY FINITE ELEMENT METHOD

Trieu Long Vo, Dai Duc Mai, Long Nhut-Phi Nguyen
HCMC University of Technology and Education (HCMUTE), Ho Chi Minh City, Vietnam
Nhan Nguyen Dang
Cao Thang Technical College (CTCT), Ho Chi Minh City, Vietnam
Son Hoai Nguyen
Nguyen Tat Thanh University (NTTU), Ho Chi Minh City, Vietnam
corresponding author Son Hoai Nguyen, email: sonnh@ntt.edu.vn

This paper presents the shape deformation of a thin plate coupled to piezoelectric actuators and sensors analyzed using the finite element method (FEM). The coupling effects between electric and mechanical properties of the piezoelectric material draw attention to potential applications, such as actuators, sensors, etc. The proposed method is analyzed and evaluated, and its effectiveness is proven. Firstly, a rectangular piezoelectric actuator with three symmetrically bonded sensors is used. Secondly, it is applied to control the swimming pool diving board. Combination of FEM and LQR active control algorithms through numerical simulation results shows changing shape and position of the piezoelectric patch which makes premises for an experiment and production.

Keywords: piezoelectric actuator, piezoelectric sensor, thin plate theory, coupling effect, Finite Element Method (FEM), LQR

1. Introduction

In recent years, many studies about the behavior of piezoelectric structures have been researched and investigated. It is a key standout among them that piezo materials with some advantages such as the quick response, low energy consumption, and high linearity, have been studied and developed for a decade (Uchino, 2010; Lumentut and Howard, 2014; Bhalla *et al.*, 2017; Wei *et al.*, 2018; Chen *et al.*, 2020). The piezoelectric device is interesting in structure engineering, applying shape control, reducing noise and controuing stability of structures (Uchino, 1986; Wang and Shen, 1998; Adriaens *et al.*, 2000; Chen etal, 2016; Chen *et al.*, 2018; Reddy, 1999). For example, Saravanos and Heyliger (1999) covered important information about theories, analytical approaches, computational models, and numerical solutions for analyzing laminates and structures in piezoelectric actuator or sensor systems. Jafferis *et al.* (2016) used multilayer laminated piezoelectric bending actuators in the design and manufacturing to achieve optimum efficiency and power density. Moreover, Bailey and Hubbard (1985) presented that vibration of a cantilever beam was controlled by an adaptative law using a PFDV film as the actuator. A two-dimensional piezoelectric material bonded on the surface with a simply supported plate were revealed by Dimitriadis *et al.* (1991). Benjeddou *et al.* (2000) described the shear actuation mechanism, which offerred several promising features for the use of piezoelectric ceramics. Luo and Tong developed a finite element model to analyze and simulate twisting and bending shape control using an orthotropic piezoelectric actuator (Luo and Tong, 2006). Huang and Sun (2006) used piezoelectricity in an actuator to control dynamic adaption with an anisotropic elastic structure. Additionally, the effect of piezoelectric arrays symmetrically attached to opposite plate surfaces in continuous operation of a composite structure was evaluated by Dimitriadis *et al.* (1991), Crawley and Luis (1987). Phung-Van *et al.* (2013) proposed the cell-based smoothed discrete shear gap approach (CS-FEM-DSG3) to improve static vibration of a fan and introduced a dynamic control model for composite plates integrated into piezoelectric sensors and actuators. Furthermore, Hoa *et al.* (2018) used the cell-based smoothed discrete shear gap method for evaluating free and static vibration of laminated composite shells. Moretti and Silva (2019) presented the use of the Topology optimization method (TOM) in designing a bi-material piezoelectric actuator (BPEA), which was capable of eliminating vibration with the active velocity feedback control (AVFC). By reading velocity state variables in the time domain, the AVFC shows its influence on the system structural damping. This study uses Newmark's time-integration method to yield dynamic response results for the rectangular four-noded finite element (FE) analysis (this method uses physical and adjoint systems because the coupling formula is extremely important in sensitivity analysis). The authors use the gradient-based optimization method when applying a mechanical load instantaneously to minimize the displacement energy output, which is determined at a predefined BPA's DOF (degree of freedom). Cao *et al.* (2020) used the classical laminated plate theory and the Fourier transformation to analyze and control actively vibration of thin-constrained composite plates and damping characteristics with dual piezoelectric layers. The study results about natural frequencies and losses of a damped composite plate bound with two piezoelectric layers were found through 3D electric potential equations. Trojanowski and Wiciak (2020) presented numerical simulation results (using ANSYS software) of the influence of sensor-actuator size on the plate performance. Two piezoelectric actuators were placed on a steel plate: one to stimulate the plate and one (a standard actuator or a sensor-actuator with different shapes and sizes) to reduce plate vibrations. The results of numerical analysis show agreement with the objective function: the minimum value of the sum of the displacement vectors of n nodes. Karegar *et al.* (2021) presented the dynamic analysis of a concrete frame with a smart layer under earthquake load conditions. The article evaluated the influence of external voltage, plate thickness, boundary conditions, geometric parameters of the frame, and damping of the structure on the seismic displacement of the frame. The authors used the Grey Wolf (GW) optimization algorithm and hyperbolic shear deformation (HSD) theory to model the flat frame and Hamilton's principle to derive the governing equation of the frame. Based on that relationship, the numerical methods of differential quadrature and Newark's one were used to study this concrete frame response. Gohery *et al.* (2022) developed a Levi-type analytical solution procedure to describe the static and dynamic deformation response of smart laminated rectangular composite plates supported under the influence of inclined piezoelectric actuators with some excitation frequency. Latrache and Menasri (2022) used piezoelectric actuators and sensor pads to actively control vibration of the classical laminated plate with embedded piezoelectric patches. The authors built a coupled finite element (FE) model with mechanical and electrical degrees of freedom based on the first-order shear deformation (FSD) theory and Hamilton's principle. At the same time, based on the independent mode space control techniques, the authors designed the Linear Quadratic Regulator (LQR) controller to limit the system vibration. Her and Chen (2022) proposed a theoretical model predicting the vibration response of a laminate composite (LC) plate to control the shape and suppress vibrations through piezoelectric actuators. The analytical solution to that vibration response was derived using the composite mechanics and plate theory. Numerical simulation results using the finite element method (FEM) through ANSYS software were very well compared with the proposed model.

In this article, the shape deformation of the plate bonded to piezoelectric actuators and sensors is investigated by the FEM. Based on the Kirchhoff plate model, a finite element analysis has been evolved for the analysis of smart composite structures with a piezoelectric material. Two cases are implemented to prove the effectiveness of the proposed method in improving the reliability of the model algorithm. The tasks are done as follows. Firstly, a rectangular piezoelectric actuator with symmetrically bonded three sensors is considered. In the second case, the piezoactuator of the injector is implemented and controlled by the diving board in diving sports. The simulated results verify the behavior of plate drive modules with changes in the piezo patch position and composite plate fiber direction. The article also analyzes and compares classical control strategies (constant amplitude and constant velocity feedback). Classical techniques have the advantages of avoiding the necessity of digital control, reducing time delays, and providing stability. A major limitation of the LQR is that all states must be measured when generating control.

2. Structure modeling

As mentioned in (Reddy, 1999), the assumptions are considered as follows: (1) the piezoelectric layers are perfectly bonded to each other, (2) the behavior of linear elastic materials is a disadvantage of the presented formula, (3) based on the Kirchhoff hypothesis (thin plate), the horizontal normal remains straight after deforming and rotating, ensuring that it is always perpendicular to the mid-surface.

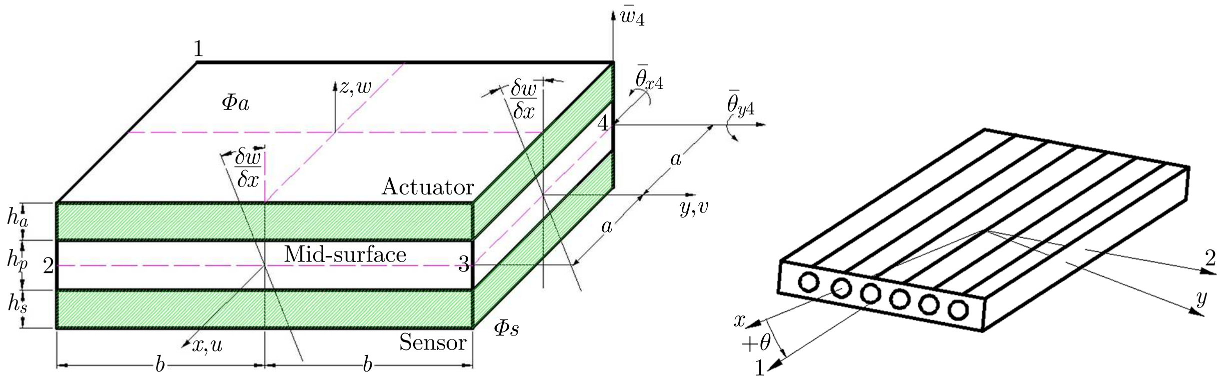


Fig. 1. A laminated finite element coordinate system with the integrated piezoelectric material (left) and the fiber direction in local and global coordinate systems (right)

Based on the Kirchhoff hypothesis, the displacement fields in u, v and w variables can be obtained as follows (Reddy, 1999)

$$u = -z\frac{\partial w}{\partial x} \qquad v = -z\frac{\partial w}{\partial y} \qquad w = w(x,y) \tag{2.1}$$

where $Oxyz$ is the Descartes coordinate system, located at the mid-surface. Additionally, u and v are displacements of the x and y-axes, while the transverse displacement w follows (or aka deflection) in the z-axis

$$\boldsymbol{\varepsilon} = [\varepsilon_x, \varepsilon_y, \gamma_{xy}] = -z\left[\frac{\partial^2 w}{\partial x^2}, \frac{\partial^2 w}{\partial y^2}, \frac{\partial^2 w}{\partial x \partial y}\right]^{\mathrm{T}} \tag{2.2}$$

The relation between the plane stress σ and strain ε of the isotropic material is denoted by

$$\boldsymbol{\sigma} = \mathbf{D}\boldsymbol{\varepsilon} \tag{2.3}$$

where

$$\boldsymbol{\sigma} = [\sigma_x, \sigma_x, \tau_{xy}]^{\mathrm{T}} \qquad \mathbf{D} = \frac{E_p}{1-v^2}\begin{bmatrix} 1 & v & 0 \\ v & 1 & 0 \\ 0 & 0 & (1-v)/2 \end{bmatrix} \tag{2.4}$$

and σ, ε, ν, E_p are stress, strain field, Poisson's ratio, and Young's modulus, respectively.

The composite material panels are made of many consecutive layers in which the direction of the fiber or basic direction is different. The relationship between stress and strain in the global system is as follows

$$\begin{bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{bmatrix} = \mathbf{Q}' \begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{bmatrix} \begin{bmatrix}Q'_{11} & Q'_{12} & Q'_{16}\\ Q'_{12} & Q'_{22} & Q'_{26}\\ Q'_{16} & Q'_{26} & Q'_{66}\end{bmatrix} \begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{bmatrix} \tag{2.5}$$

where

$$\begin{aligned}
Q;_{11} &= Q_{11}\cos^4\theta + 2(Q_{12}+2Q_{66})\sin^2\theta\cos^2\theta + Q_{22}\sin^4\theta\\
Q'_{12} &= (Q_{11}+Q_{22}-4Q_{66})\sin^2\theta\cos^2\theta + Q_{12}\sin^4\theta\cos^4\theta\\
Q'_{22} &= Q_{11}\sin^4\theta + 2(Q_{12}+2Q_{66})\sin^2\theta\cos^2\theta + Q_{22}\cos^4\theta\\
Q'_{16} &= (Q_{11}-Q_{12}-2Q_{66})\sin\theta\cos^3\theta + (Q_{12}-Q_{22}+2Q_{66})\sin^3\theta\cos\theta\\
Q'_{26} &= (Q_{11}-Q_{12}-2Q_{66})\sin^3\theta\cos\theta + (Q_{12}-Q_{22}+2Q_{66})\sin\theta\cos^3\theta\\
Q'_{66} &= (Q_{11}+Q_{12}-2Q_{12}-2Q_{66})\sin^2\theta\cos^2\theta + Q_{66}(\sin^4\theta+\cos^4\theta)
\end{aligned} \tag{2.6}$$

In fact, $\mathbf{Q}'$ is a complete matrix indicating that the shear strain γ_{xy} in the (x, y) plane is associated with the normal strain ε_x, ε_y. This behavior is called the shear strain and shear strain long form.

Similarly

$$\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \tau_{xy}\end{bmatrix} = \mathbf{S}' \begin{bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{bmatrix} \begin{bmatrix}S'_{11} & S'_{12} & S'_{16}\\ S'_{12} & S'_{22} & S'_{26}\\ S'_{16} & S'_{26} & S'_{66}\end{bmatrix} \begin{bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{bmatrix} \tag{2.7}$$

Layer softness composition $\mathbf{S}' = \{S'_{ij}\}$ with $i, j = 1, 2$

$$\mathbf{S}' = \mathbf{T}_\varepsilon^{-1}\mathbf{S}\mathbf{T} \tag{2.8}$$

where

$$\begin{aligned}
S'_{11} &= S_{11}\cos^4\theta + S_{22}\sin^4\theta + 2(S_{12}+2S_{66})\cos^2\theta\sin^2\theta\\
S'_{12} &= S_{12}(\cos^4\theta+\sin^4\theta) + (S_{11}+S_{22}-S_{66})\cos^2\theta\sin^2\theta\\
S'_{16} &= (S_{11}-S_{12}-2S_{66})\sin\theta\cos^3\theta + (S_{12}-S_{22}+2S_{66})\sin^3\theta\cos\theta\\
S'_{22} &= S_{11}\sin^4\theta + S_{22}\cos^4\theta + (2S_{12}+S_{66})\sin^2\theta\cos^2\theta\\
S'_{26} &= (2S_{11}-2S_{12}-S_{66})\cos\theta\sin^3\theta + (2S_{12}+S_{66}-2S_{22})\cos^3\theta\sin\theta\\
S'_{66} &= S_{66}(\cos^4\theta+\sin^4\theta) + 2(S_{11}+S_{22}-4S_{12}-S_{66})\sin^2\theta\cos^2\theta
\end{aligned} \tag{2.9}$$

with

$$\begin{aligned}
&S_{11} = \frac{1}{E_1} \qquad && S_{12} = -\frac{\nu_{12}}{E_2} \qquad && S_{21} = -\frac{\nu_{21}}{E_2}\\
&S_{22} = \frac{1}{E_2} \qquad && S_{66} = \frac{1}{G_{12}} \qquad && S_{16} = S_{26} = S_{61} = S_{62} = 0
\end{aligned} \tag{2.10}$$

where E_1, E_2, ν_{12}, ν_{21}, G_{12} are Young's elastic modulus, Poisson's coefficients, and the corresponding in-plane shear elastic modulus (Fig. 1, right).

Based on the original plate theory in the finite element analysis, considering a four-node rectangular plate bending element is developed (Bailey and Hubbard, 1985). Figure 2 shows the

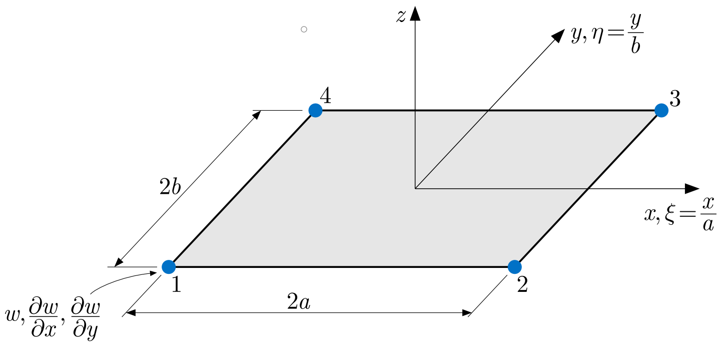


Fig. 2. 3-DOF (degrees of freedom) of the rectangular element per node

DOF (degrees of freedom) number of each node of the element: w displacement in the z direction, $\theta_x = \partial w/\partial x$ rotation around the x-axis, and $\theta_y = \partial w/\partial y$ rotation around the y-axis.

Based on the Pascal triangle law, the interpolation function is selected. The displacement w at an arbitrary point in the element is as follows

$$\begin{aligned} w(x_i,y_i) &= c_1 + c_2x_i + c_3y_i + c_4x_i^2 + c_5y_ix_i + c_6y_i^2 + c_7x_i^3 \\ &\quad + c_8x_i^2y_i + c_8x_iy_i^2 + c_{10}y_i^3 + c_{11}x_i^3y_i + c_{12}x_iy_i^3 \end{aligned} \tag{2.11}$$

where $(i = 1,2,3,4)$

$$x_1 = x_4 = -a \qquad x_2 = x_3 = a \qquad y_1 = y_2 = -b \qquad y_3 = y_4 = b \tag{2.12}$$

The transverse displacement field w is expressed by $w = \mathbf{P}^T\mathbf{c}$, where

$$\begin{aligned} \mathbf{P} &= [1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^3y, xy^3]^{\mathrm{T}} \\ \mathbf{c} &= [c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}] \end{aligned} \tag{2.13}$$

In the rectangular element, the node displacement field vector $\mathbf{d}_i$ is expressed by

$$\mathbf{d}_i = [\overline{w}_1, \overline{\theta}_{x1}, \overline{\theta}_{y1}, \overline{w}_2, \overline{\theta}_{x2}, \overline{\theta}_{y2}, \overline{w}_3, \overline{\theta}_{x3}, \overline{\theta}_{y3}, \overline{w}_4, \overline{\theta}_{x4}, \overline{\theta}_{y4}]^{\mathrm{T}} \tag{2.14}$$

where

$$\overline{w}_i = w\Big|_{x_i,y_i} \qquad \overline{\theta}_{xi} = \frac{\partial w}{\partial y}\bigg|_{x_i,y_i} \qquad \overline{\theta}_{yi} = -\frac{\partial w}{\partial x}\bigg|_{x_i,y_i} \tag{2.15}$$

And the displacement field can be expressed as follows

$$\mathbf{d} = \mathbf{H}\mathbf{L}^{\mathrm{T}}\mathbf{X}^{-1}\mathbf{d}_i \qquad i = 1,\dots,n \tag{2.16}$$

where

$$\mathbf{H}=\begin{bmatrix}1 & 0 & 0\\ 0 & -z & 0\\ 0 & 0 & -z\end{bmatrix} \qquad \mathbf{L}=\begin{bmatrix}0 & 0 & 0 & 2 & 0 & 0 & 6x & 2y & 0 & 0 & 6xy & 0\\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2x & 6y & 0 & 6xy\\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 4x & 4y & 0 & 6x^2 & 6y^2\end{bmatrix}$$

$$\mathbf{X}=\begin{bmatrix}1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 & x_1^3 & x_1^2y_1 & x_1y_1^2 & y_1^3 & x_1^3y_1 & x_1y_1^3\\ 1 & 0 & 1 & 0 & x_1 & 2y_1 & 0 & x_1^2 & 2x_1y_1 & 3y_1^2 & x_1^3 & 3x_1y_1^2\\ 0 & -1 & 0 & -2x_1 & -y_1 & 0 & -3x_1^2 & -2x_1y_1 & -y_1^2 & 0 & -3x_1^2y_1 & -y_1^3\\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 & x_2^3 & x_2^2y_2 & x_2y_2^2 & y_2^3 & x_2^3y_2 & x_2y_2^3\\ 1 & 0 & 1 & 0 & x_2 & 2y_2 & 0 & x_2^2 & 2x_2y_2 & 3y_2^2 & x_2^3 & 3x_2y_2^2\\ 0 & -1 & 0 & -2x_2 & -y_2 & 0 & -3x_2^2 & -2x_2y_2 & -y_2^2 & 0 & -3x_2^2y_2 & -y_2^3\\ 1 & x_3 & y_3 & x_3^2 & x_3y_3 & y_3^2 & x_3^3 & x_3^2y_3 & x_3y_3^2 & y_3^3 & x_3^3y_3 & x_3y_3^3\\ 1 & 0 & 1 & 0 & x_3 & 2y_3 & 0 & x_3^2 & 2x_3y_3 & 3y_3^2 & x_3^3 & 3x_3y_3^2\\ 0 & -1 & 0 & -2x_3 & -y_3 & 0 & -3x_3^2 & -2x_3y_3 & -y_3^2 & 0 & -3x_3^2y_3 & -y_3^3\\ 1 & x_4 & y_4 & x_4^2 & x_4y_4 & y_4^2 & x_4^3 & x_4^2y_4 & x_4y_4^2 & y_4^3 & x_4^3y_4 & x_4y_4^3\\ 1 & 0 & 1 & 0 & x_4 & 2y_4 & 0 & x_4^2 & 2x_4y_4 & 3y_4^2 & x_4^3 & 3x_4y_4^2\\ 0 & -1 & 0 & -2x_4 & -y_4 & 0 & -3x_4^2 & -2x_4y_4 & -y_4^2 & 0 & -3x_4^2y_4 & -y_4^3\end{bmatrix} \tag{2.17}$$

3. Piezoelectric constitutive equations

The linear constitutive relations of piezoelectric materials are given as follows

$$\boldsymbol{\sigma}=\mathbf{C}^E\boldsymbol{\varepsilon}-\mathbf{e}^{\mathrm{T}}\mathbf{E} \qquad \mathbf{D}=\mathbf{e}\boldsymbol{\varepsilon}+\boldsymbol{\xi}^s\mathbf{E} \tag{3.1}$$

In the matrix form

$$\begin{bmatrix}\boldsymbol{\sigma}\\ \mathbf{D}\end{bmatrix}=\begin{bmatrix}\mathbf{C}^E & -\mathbf{e}^{\mathrm{T}}\\ \mathbf{e} & \varsigma^s\end{bmatrix}\begin{bmatrix}\boldsymbol{\varepsilon}\\ \mathbf{E}\end{bmatrix} \tag{3.2}$$

where $\boldsymbol{\sigma}$, $\mathbf{D}$, $\boldsymbol{\varepsilon}$, $\mathbf{E}$, $\mathbf{C}^E$, $\mathbf{e}$, ς^S are the stress field, electric flux vector, strain field, electric field vector, elastic constant matrix, piezoelectric coupling constant matrix, and electric coefficient matrix, respectively.

Applying voltage to the element seems similar to applying heat to a bimetallic strip. The voltage Φ_a across the bender element forces the bottom layer to expand, as illustrated in Fig. 3a, while the top layer is contracted.

As a result of these physical phenomena, there is a significant curvature, implying a substantial deflection at the tip while the other end is clamped. Due to the reciprocity effect, the sensor deformation generates a charge across the electrode, collected as a voltage Φ_s through another sensor surface. The equation expresses the applied/perceived electric potential via the actuator or sensor element (Lopes *et al.*, 2000)

$$\phi_z=\Big(\frac{z-0.5h_p}{h}\Big)\phi \tag{3.3}$$

where h, ϕ and $z(z_a,z_s)$ are the thicknesses, maximum electric potentials at the external surfaces of the corresponding piezoelectric elements. The variables $z(z_a,z_s)$ can be obtained as follows

$$\frac{h_p}{2}\leqslant z_a\leqslant\frac{h_p}{2}+h_a \qquad -\frac{h_p}{2}\geqslant z_s\geqslant-\frac{h_p}{2}-h_s \tag{3.4}$$

The assumption for the electric field E remains constant regardless of the thickness of actuator and sensor parts, and the gradient operators are recast as follows

$$E=-\operatorname{grad}\phi=-\frac{\partial\phi_z}{dz}=-B_z \qquad \varphi=-\frac{\phi}{h} \tag{3.5}$$

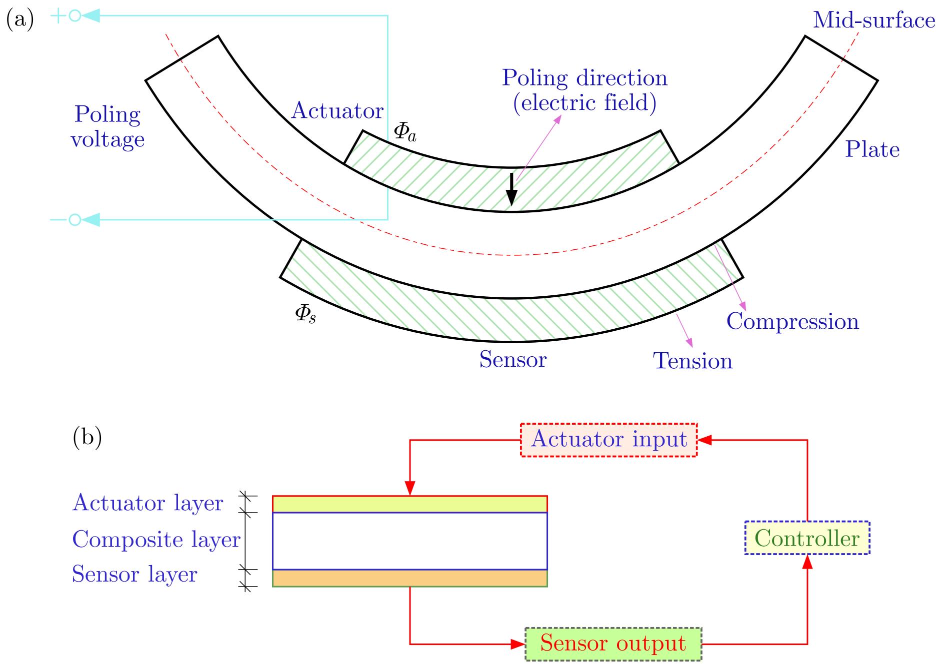


Fig. 3. (a) Curvature of the plate caused by mid-layer and contraction of both sides. (b) A control diagram for the laminate plate with integrated piezoelectric sensors and actuators

However, for implementation of an experiment in this article, the electric field $\mathbf{E}$ is assumed

$$\mathbf{E} = \begin{bmatrix} 0 \\ 0 \\ E_3 \end{bmatrix} \qquad \text{with } \mathbf{B}_z = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{h} \end{bmatrix} \tag{3.6}$$

where h is the thickness of the piezoelectricity material.

With linear piezoelectric materials, the elastic coefficient matrix, piezoelectric stress coefficient matrix, and dielectric coefficient matrix are given by

$$\mathbf{C}^E = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & C_{33} \end{bmatrix} \qquad \mathbf{e} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ e_{31} & e_{31} & 0 \end{bmatrix} \qquad \boldsymbol{\xi}^s = \begin{bmatrix} \xi_{11} & 0 & 0 \\ 0 & \xi_{11} & 0 \\ 0 & 0 & \xi_{11} \end{bmatrix} \tag{3.7}$$

4. The governing equations and finite element formulation

A two-dimensional piezoelectric problem in the domain Ω bounded by Γ is considered to be carried out. The governing equations and boundary conditions for linear piezoelectric materials are introduced as follows (Saravanos and Keyliger, 1999)

$$\sigma_{ij,i} + f_j = \rho \ddot{u}_j \qquad \varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}) \qquad D_{i,j} = 0 \qquad E_i = -\phi_{,i} \tag{4.1}$$

By integrating with the boundary conditions

$$\begin{array}{llll} \sigma_{ij} n_j = \overline{t}_i \quad \text{on} \quad \Gamma_\sigma & & u_i = \overline{u}_i \quad \text{on} \quad \Gamma_u \\ \phi = \overline{\phi} \quad \text{on} \quad \Gamma_\phi & & D_i n_i = -\overline{q} \quad \text{on} \quad \Gamma_q \end{array} \tag{4.2}$$

where σ_{ij} and ε_{ij} denote the stress and strain tensors, respectively. f_j is the body force density, u_j is the mechanical displacement vector, and ρ is the mass density. The electric displacement vector is D_i, the electric field vector is E_i, and the scalar electric potential field is φ. Summing the kinetic energy, strain energy, dielectric energy, and potential energy of external fields yields the general functional L. The disadvantage of the governing equations of piezoelectric structures is to use the Hamilton principle, written as follows (Lopes etal, 2000; Phung-Van *et al.*, 2013)

$$I = \int_{t1}^{t2} L\, dt = \int_{t1}^{t2} [\delta(T - U + W_e - W_m) + \delta W]\, dt \tag{4.3}$$

where t_1 and t_2 are two random instants. L, T, U and W_e are the general energy functional, kinetic energy, potential energy, and the work done by electrical forces, respectively. The variable W is the mechanical force, being negligible for piezoelectric material. The total potential energy U and kinetic energy T of the composite structure are described as follows

$$U = \frac{1}{2}\int \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\sigma}\, dV \qquad T = \frac{1}{2}\int \rho \dot{\mathbf{d}}^{\mathrm{T}}\dot{\mathbf{d}}\, dV \tag{4.4}$$

where $\dot{\mathbf{d}}$ is the differentiation of $\mathbf{d}$ with respect to time, $\mathbf{d}$ is the nodal displacement field, and dV is defined by

$$dV = dV_a + dV_p + dV_s \tag{4.5}$$

Herein, the subscripts a, p and s denote the actuator, plate, and sensor parts, respectively, and dV_a, dV_p and dV_s are given by

$$\begin{aligned} dV_p &= \int_{-h_p/2}^{h_p/2}\int_{-b}^{b}\int_{-a}^{a} dx\, dy\, dz \qquad dV_a = \int_{h_p/2}^{h_p/2+h_a}\int_{-b}^{b}\int_{-a}^{a} dx\, dy\, dz \\ dV_s &= \int_{-h_p/2-h_s}^{-h_p/2}\int_{-b}^{b}\int_{-a}^{a} dx\, dy\, dz \end{aligned} \tag{4.6}$$

The W_e work of electrical forces and the variable W for mechanical forces is illustrated by

$$W_e = \frac{1}{2}\int_V \mathbf{E}^{\mathrm{T}}\mathbf{D}\, dV \qquad W = \int_V \mathbf{q}^{\mathrm{T}}\mathbf{f}_b\, dV + \int_A \mathbf{q}^{\mathrm{T}}\mathbf{f}_A\, dA + \int_A \phi\sigma_q\, dA \tag{4.7}$$

Substituting equation (3.1) into equation (4.4) and equation $(4.7)_1$, we obtain the following

$$U = \frac{1}{2}\int_V \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{C}^E\boldsymbol{\varepsilon}\, dV - \frac{1}{2}\int_V \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{e}^{\mathrm{T}}\mathbf{E}\, dV \qquad W_e = \frac{1}{2}\int_V \mathbf{E}^{\mathrm{T}}\mathbf{e}\boldsymbol{\varepsilon}\, dV + \frac{1}{2}\int_V \mathbf{E}^{\mathrm{T}}\boldsymbol{\xi}^S\mathbf{E}\, dV \tag{4.8}$$

From equations (4.7) substituted to equation (4.3) one arrives at

$$\int_{t1}^{t2} \delta\mathbf{q}_k^{\mathrm{T}}[\mathbf{M}_{qq}^e\ddot{\mathbf{q}}_k + \mathbf{K}_{qq}^e\mathbf{q}_k + \mathbf{K}_{q\varphi}^e\boldsymbol{\varphi} - \overline{\mathbf{f}}] + \delta\boldsymbol{\varphi}[\mathbf{K}_{\varphi q}^e\mathbf{q}_k + \mathbf{K}_{\varphi\varphi}^e]\boldsymbol{\varphi} + \mathbf{Q}_a]\, dt = 0 \tag{4.9}$$

where

$$\begin{aligned} &\mathbf{M}_{qq}^e = \rho\int_V \mathbf{X}^{-\mathrm{T}}\mathbf{L}_M\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{L}_M^{\mathrm{T}}\mathbf{X}^{-1}\, dV \qquad \mathbf{K}_{qq}^e = \mathbf{X}^{-\mathrm{T}}\int_V z^2\mathbf{L}_K^{\mathrm{T}}\mathbf{D}\mathbf{L}_K\mathbf{X}^{-1}\, dV \\ &\mathbf{K}_{q\varphi}^e = [\mathbf{K}_{\varphi q}^e]^{\mathrm{T}} = -\mathbf{X}^{-\mathrm{T}}\int_V z\mathbf{L}_K^{\mathrm{T}}\mathbf{e}^{\mathrm{T}}B_z\, dV \qquad \mathbf{K}_{\varphi\varphi}^e = -\int_V B_z\boldsymbol{\xi}^S B_z\, dV \\ &\overline{\mathbf{f}} = \int_V \mathbf{f}_b\, dV + \int_A \mathbf{f}_A\, dA \qquad \mathbf{Q}_a = \int_A \sigma_q\, dA \end{aligned} \tag{4.10}$$

Allowing arbitrary variations of dk and φ, from (4.9), we have two equilibrium equations for k-th element in generalized coordinates that are now obtained as follows

$$\mathbf{M}^e_{dd}\ddot{\mathbf{d}}_k + \mathbf{K}^e_{dd}\mathbf{d}_k + \mathbf{K}^e_{\phi\phi}\boldsymbol{\phi} - \overline{\mathbf{f}} = \mathbf{0} \qquad \mathbf{K}^e_{dd}\mathbf{d}_k + \mathbf{K}^e_{\phi\phi}\boldsymbol{\phi} + \mathbf{Q} = \mathbf{0} \tag{4.11}$$

5. Dynamic control

From equation (4.11), the structure system can be rearranged as follows

$$\begin{aligned} &\mathbf{M}_{dd}\ddot{\mathbf{d}} + \mathbf{K}_{dd}\mathbf{d} + \mathbf{K}_{d\phi}\boldsymbol{\phi} - \overline{\mathbf{f}} = \mathbf{0} \qquad \mathbf{K}_{\phi d}\mathbf{d} + \mathbf{K}_{\phi\phi}\boldsymbol{\phi} + \mathbf{Q} = \mathbf{0} \\ &\mathbf{M}_{dd}\ddot{\mathbf{d}} + (\mathbf{K}_{dd} + \mathbf{K}_{d\phi} + \mathbf{K}^{-1}_{\phi\phi} + \mathbf{K}_{\phi d})\mathbf{d} + \mathbf{K}_{d\phi}\boldsymbol{\phi} = \overline{\mathbf{f}} + (\mathbf{K}^{-1}_{\phi\phi} + \mathbf{K}_{\phi d})\mathbf{Q} \end{aligned} \tag{5.1}$$

The resultant control gain is

$$\varphi_a = G_d\varphi_s + G_\nu\varphi'_s \tag{5.2}$$

where G_v and G_d are the velocity and displacement feedback control gains.

The modeling equaton of the system is found as follows

$$M\dot{d} + (C + C_R)\dot{d} + K^*d = F \tag{5.3}$$

where

$$\begin{aligned} &K^* = K_{uu} + G_d[\mathbf{K}_{u\phi}]_s[\mathbf{K}^{-1}_{\phi\phi}]_s[\mathbf{K}_{\phi u}]_s \qquad C = G_v[\mathbf{K}_{u\phi}]_a[\mathbf{K}^{-1}_{\phi\phi}]_s[\mathbf{K}_{\phi u}]_s \\ &C_R = \alpha M + \beta K_{uu} \end{aligned} \tag{5.4}$$

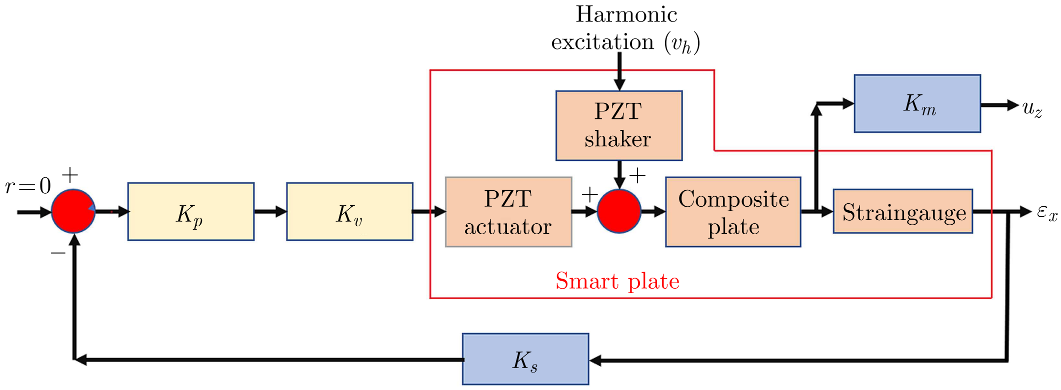


Fig. 4. The control structure of the composite plate by PZT

6. Numerical results

6.1. The first case

Using numerical techniques, the displacement comparison in the first case verifies the proposed method effectiveness. To do this work, the characteristics of piezoelectric material and simply supported (SSSS) of rectangular plates are presented in Table 1 and Fig. 5, respectively. The SSSS boundary conditions are symmetrically bonded with the matching architecture of the piezoelectric actuators and sensors. Furthermore, the effect of piezoelectric patches on the structure static behavior has been investigated.

6.1.1. With steel plate

Table 1. Properties of the steel plate and piezoelectric material

Properties	Piezoelectric material		Steel plate
	sensor	actuator	
Young's modulus E [GPa]	2	69	207
Density ρ [kg/m^3]	1780	7700	7870
Poisson ratio ν	0.3	0.3	0.29
Thickness h [m]	$0.205 \cdot 10^{-3}$	$0.254 \cdot 10^{-3}$	$1 \cdot 10^{-3}$
Piezodielectric ζ^s [F/m]	$1.06 \cdot 10^10$	$1.6 \cdot 10^8$	–
Piezoelectric strain e [C/m^2]	0.046	-12.5	–
Capacitance C [F]	$5.2 \cdot 10^9$	$6.3 \cdot 10^7$	–
Geometry $L_x \times L_y$ [m]	0.1×0.1	0.1×0.1	0.6×0.4

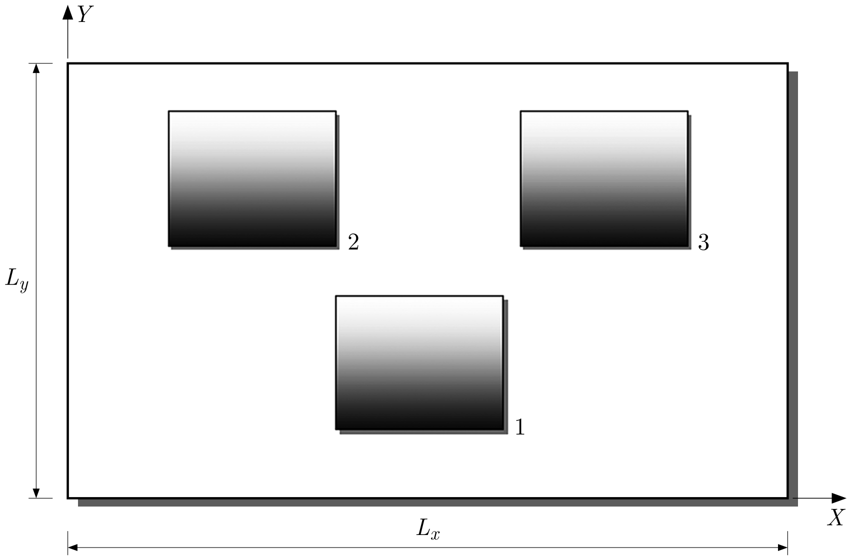


Fig. 5. Piezoelectric actuators and sensors test configuration

The mesh grid (24×16) in the first test is shown in Fig. 6.

Table 2 presents the position of sensors and actuators.

Table 2. Piezoelectric element positions in the x and y dimensions

Properties	Actuators and sensors		
	1	2	3
x	0.25	0.15	0.35
y	0.05	0.25	0.25

Considering the first section, where the position is the plane $y = 0.2$, the displacement graph is shown in Fig. 7.

The total plate displacement amplitude results using the proposed method when a static voltage is applied to the actuator with the magnitude of $\boldsymbol{\Phi}_a = [-1, 1, 1]$, are shown in Fig. 8.

Table 3 also compares the sensor-generated electric potential results obtained by Abreu *et al.* (2004) and from the proposed method.

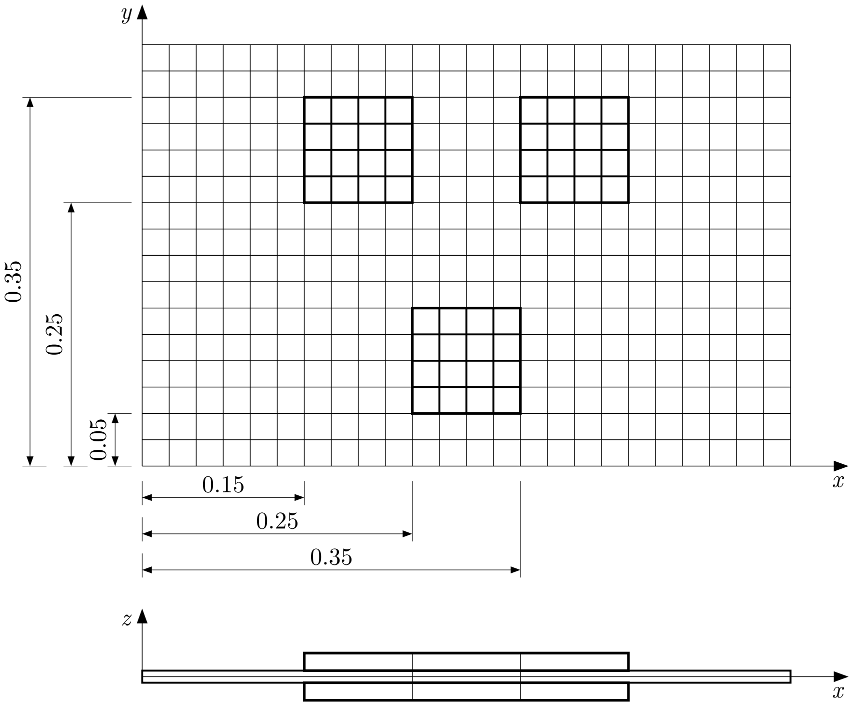


Fig. 6. Mesh grid (24 × 16) and position of actuators and sensors

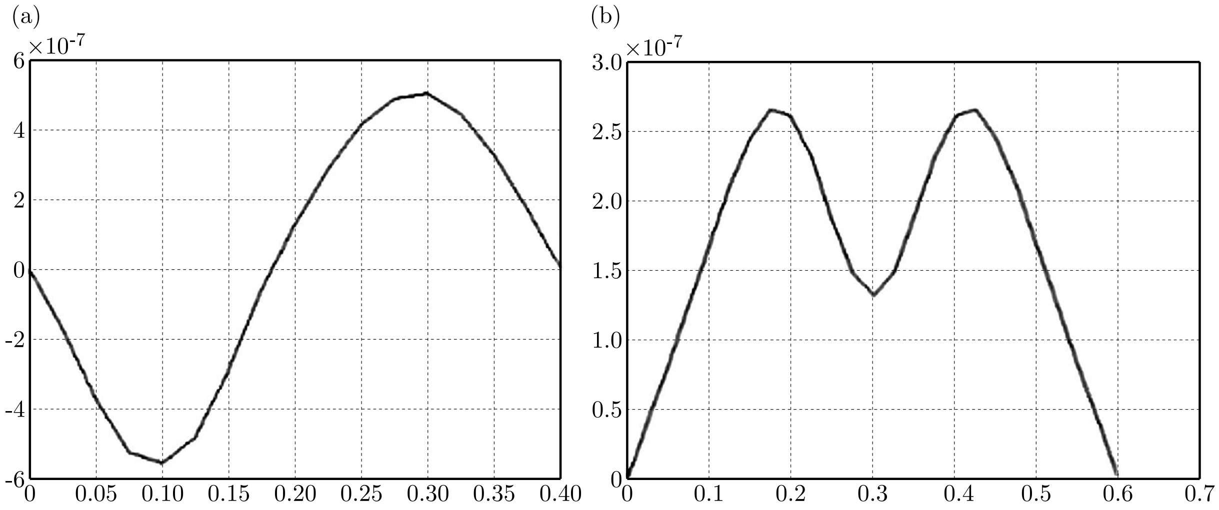


Fig. 7. The displacement graph at the plane $x = x/2$ (left) and $y = y/2$ (right)

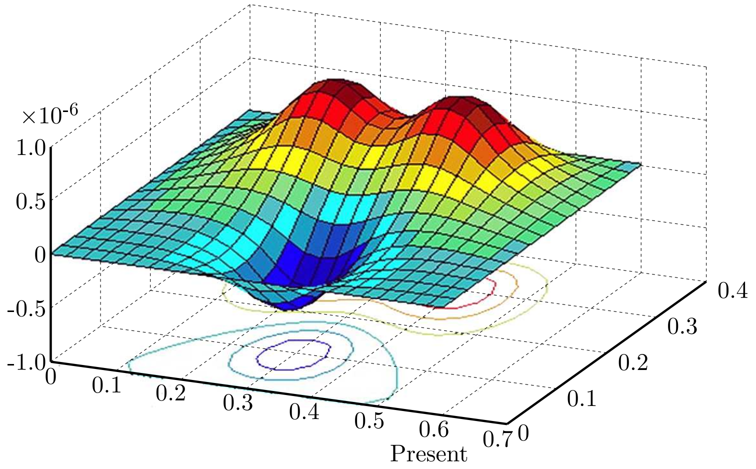


Fig. 8. The total displacement amplitude of the plate

Table 3. The electric potential from sensors

Actuators and sensors	Electric potential V			Err [V]
	Exact solution Abreu *et al.* (2004)	Present	Abreu *et al.* (2004)	
1	+0.0139	+0.0154	+0.0162	0.0015
2	−0.0139	−0.0154	−0.0162	0.0015
3	−0.0139	−0.0154	−0.0162	0.0015

Compared to the exact solution, the errors in Table 3 are small: Err = 0.0015 V, the relative error is about 10% for actuators and sensors. It is declared that the proposed method can be accurately applied to simulate the bending effect (displacement field) in the plate.

As mentioned above, the displacement of the plate is a result of supplying the static voltage to the transmission system. The plate is controlled using the closed loop algorithm to reduce the displacement. The results are illustrated in Fig. 9. The Z-direction displacement in the noncontrollable sensor is bigger than in the controllable one.

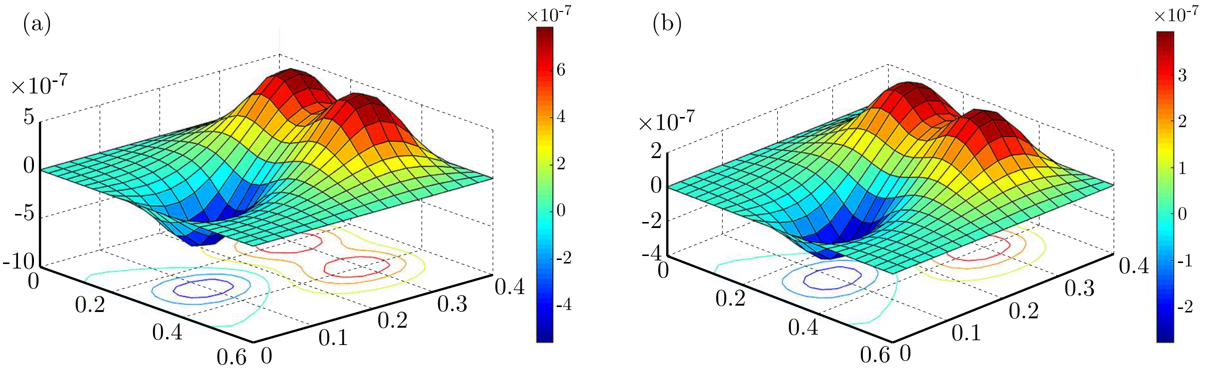


Fig. 9. The total displacement amplitude of the plate: (a) without control, (b) with control

Let us continue to develop the problem with the above material parameters. Apply an evenly distributed load $P = -10$ N, place the piezoelectric plate at the center, and control with the input voltages to 10 V, 20 V, 50 V, 100 V and 200 V, and get the results shown in Fig. 10.

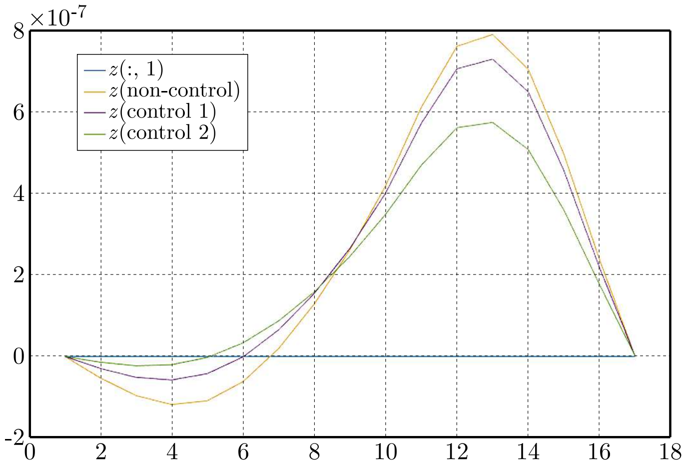


Fig. 10. The displacement graph at the plane $y/2$

The plate displacement decreases over time when other values are kept constant and the voltage is gradually increased. Combined with active control of the plate, the amplitudes of vibrations decrease progressively over time, and the speed of extinguishing the vibrations is faster than without control. The obtained results are shown in Figs. 11 and 12.

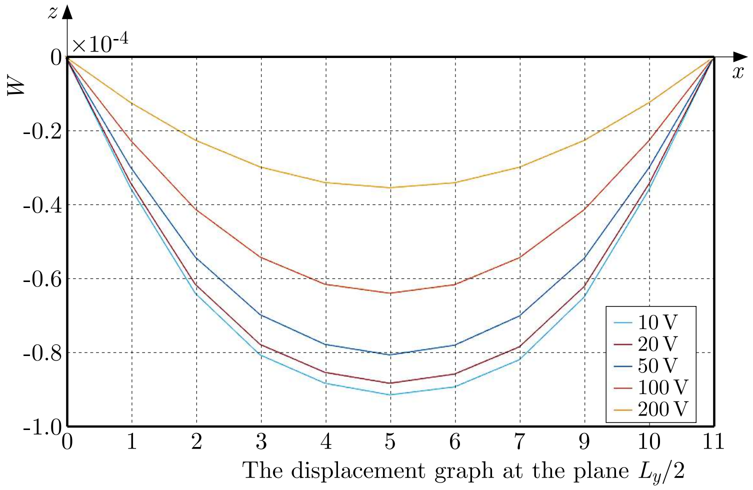


Fig. 11. The displacement graph at the plane $y/2$

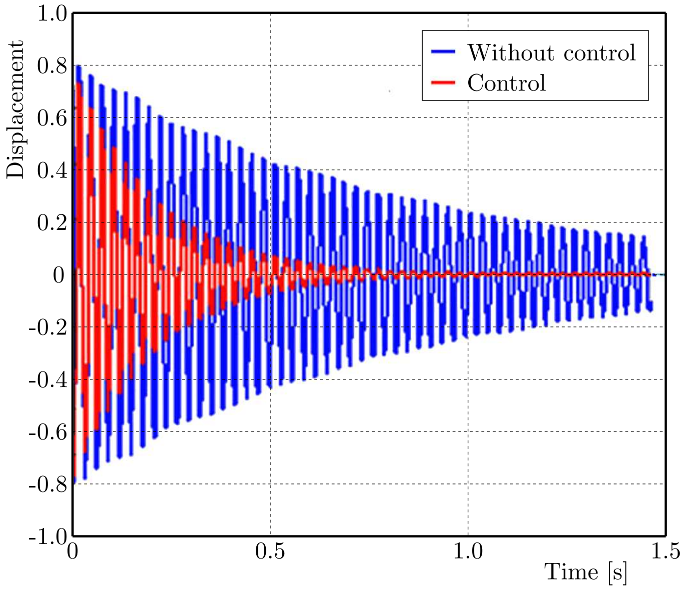


Fig. 12. Comparison of the displacement of the composite plate with control and without control

6.1.2. Changing the composite layers for the plate

The material is a carbon/epoxy composite with a stacking sequence [0/90] s. The properties are listed in Table 4. This composite laminate plate has length $a = 0.38\,\mathrm{m}$, width $b = 0.3\,\mathrm{m}$, thickness $t_p = 1.5876\,\mathrm{mm}$. The PZT G-1195 piezoelectric actuator has Young's modulus $E_{pe} = 63\,\mathrm{GPa}$, Poisson's ratio $\upsilon_{pe} = 0.3$, density $\rho_{pe} = 7600\,\mathrm{kg/m^3}$, piezoelectric constant $\zeta^s = 1.9 \cdot 10^{-10}\,\mathrm{V/m}$, and thickness $t_{pe} = 0.15876\,\mathrm{mm}$ (Dileep *et al.*, 2017). Through parametric research, the influence of the actuator size and position on the composite plate deflection and deformation shape when activated by the surface-bonded piezoelectric actuator is evaluated.

As shown in Fig. 13, the plate displacement generated relative to PZT decreases linearly as the number of layers increases.

Table 4. Properties of carbon/epoxy material

Longitudinal module E_1 [GPa]	Transverse modulus E_2 [GPa]	Shear modulus G_{12} [GPa]	Poisson's ratio	
			ν_{12}	ν_{21}
108	10.3	7.13	0.28	0.28

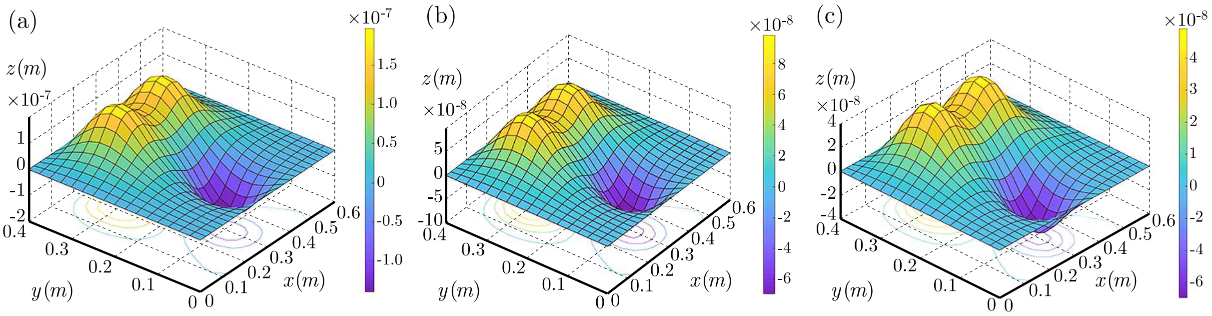


Fig. 13. The total displacement amplitude of the plate with 2-layer (left), 4-layer (between), and 8-layer (right)

When changing the fiber direction of an 8-layer plate, in the case of the [75, −75] s plate arrangement, the best ability to resist displacement caused by the PZT plates is shown in Fig. 14.

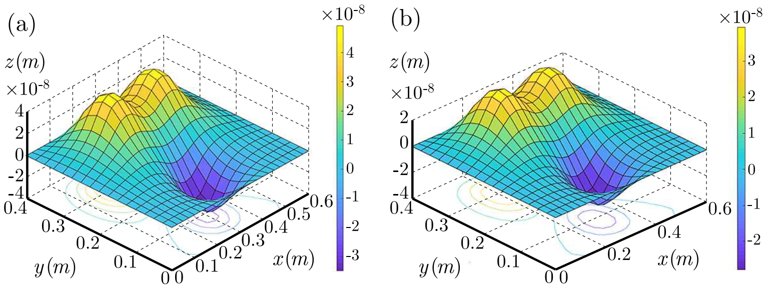


Fig. 14. The total displacement amplitude of the 8-layer composite plate with the fiber direction [30, 60] s (left) and [75, −75] s (right)

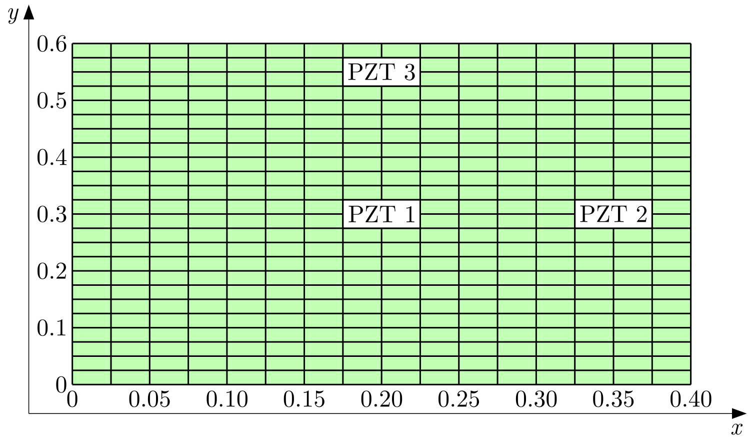


Fig. 15. Mesh model of the composite plate and the positions of 3 PZT patches

The actuators are placed at various locations to study their capability of controlling the plate deflection shape. In this case, the piezoelectric actuators are surface bonded at the plate central, right, and top regions, respectively, as shown in Fig. 15. Those are three typical positions chosen arbitrarily to demonstrate the effect of actuator position on the deflection.

Table 5. Piezoelectric element central positions in x and y dimensions

Properties	Actuators and sensors		
	1	2	3
x	0.2	0.35	0.2
y	0.3	0.3	0.55

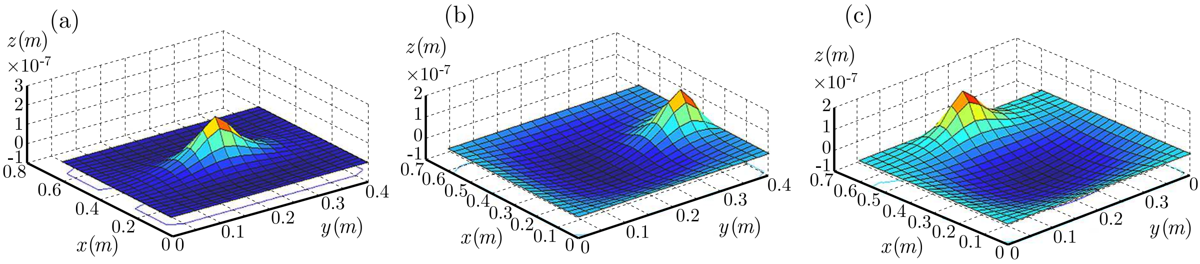


Fig. 16. Displacement of the composite plate with the PZT patch positioned at the center (left), x-axis (between), and y-axis (right)

Table 6. Comparison of the maximum displacement of the composite plate induced by PZT patches at three different positions

Positions	W [mm]
1	-0.2358
2	-0.0849
3	-0.0691

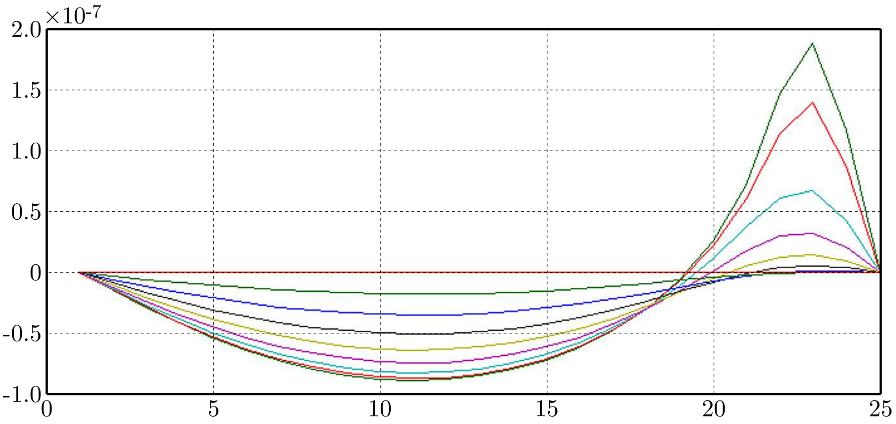


Fig. 17. The displacement graph at the plane $x/2$

When the position of the PZT piece on the composite plate changes, the displacement also changes. The PZT pieces placed closer to the boundary edges will gradually decrease the displacements.

If the composite panel is rectangular, PZT panels placed near the edge along the width direction will have smaller displacements than PZT panels placed at the edge along the length of the panel.

Based on Fig. 17, the torque generated deforms the composite panel where the PZT panel is placed and in the opposite direction in the surrounding area.

6.2. The second case

In this simulation, the proposed algorithm is investigated to control the diving board in diving sports. This article aims to show that the diving board can be safe in the case of overload and help improve athletes' performance before touching the swimming pool surface. The physical parameters of the diving board are $2.5\,\mathrm{m}\times 0.4\times 0.01\,\mathrm{m}$ for the composite 4-layers $[75°/-75°/-75°/75°]$ and $[30°/60°/60°/30°]$ (shown as Fig. 18). The control diagram is shown in Fig. 3b.

Fig. 18. The model of the diving board (desired scenario)

With diving board composite 4-layers $[75°/-75°/-75°/75°]$ and $[30°/60°/60°/30°]$, the piezoelectric ceramic (PZT 1V) is stuck in the diving board, as seen in Fig. 19.

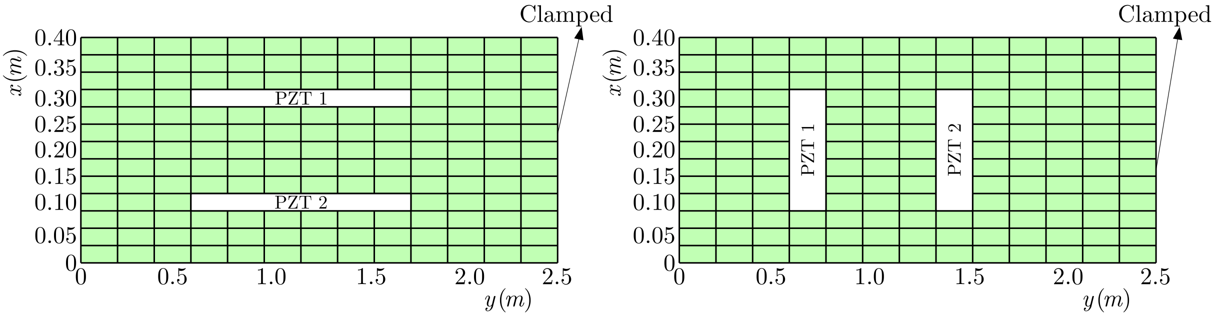


Fig. 19. The position of PZT: scenario 1 (left) and scenario 2 (right)

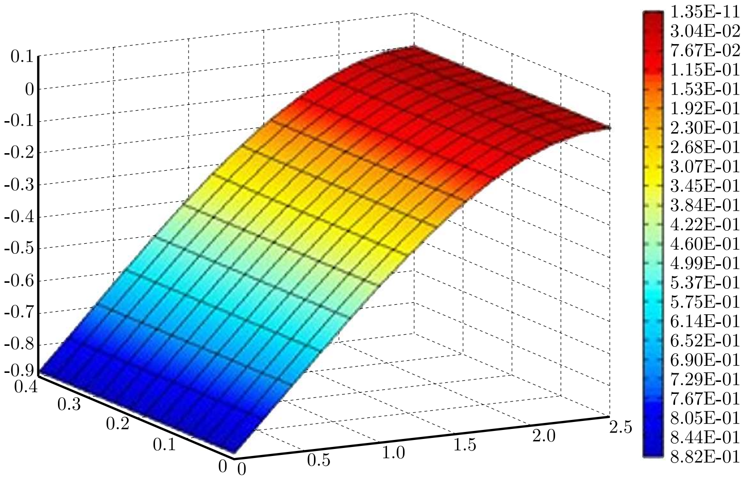


Fig. 20. The displacement (with m unit) of the diving board in non-controllable

The simulation results are found from two scenarios displayed in Figs. 21 and 22 with the composite 4-layers $[75°/-75°/-75°/75°]$ and $[30°/60°/60°/30°]$.

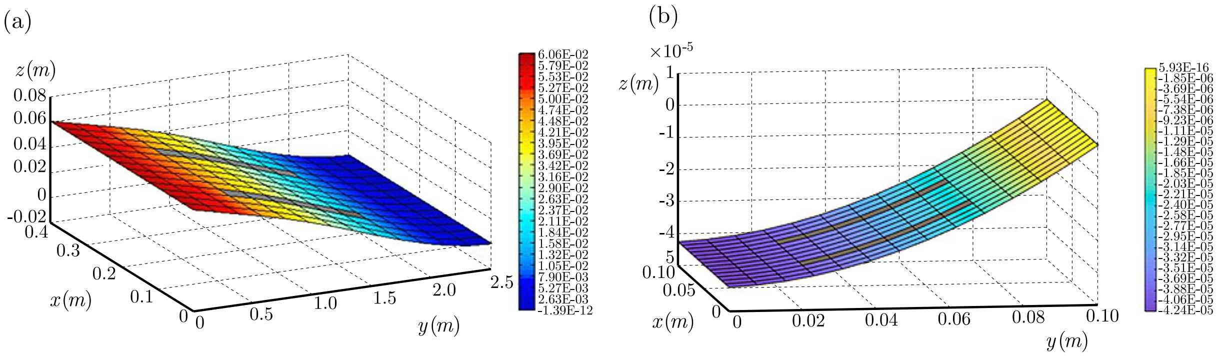


Fig. 21. The displacement of a controlled diving board for scenario 1: fiber direction $[75, -75]$s (left), fiber direction $[30, 60]$s (right)

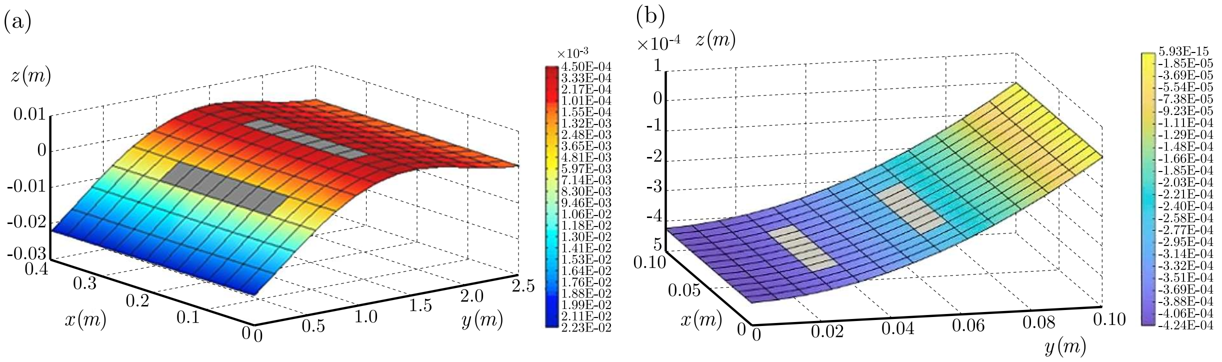


Fig. 22. The displacement of a controlled diving board in scenario 2: fiber direction $[75, -75]$s (left), fiber direction $[30, 60]$s (right)

Table 7. Piezoelectric element positions in x and y dimensions

Direction of fibers	Maximum displacement [m]		
	Non-controllable	Scenario 1	Scenario 2
$[75°/-75°/-75°/75°]$	0.8	0.0606	0.0223
$[30°/60°/60°/30°]$	0.8	0.0424	0.0424

The control method backs the diving board approximately to the initial position. Therefore, the displacement of the z-axis of the diving board is narrowed down. Although the displacement of scenario 2 is less than that of scenario 1, the shape of the diving board after controlling scenario 1 is nearly close to the desired scenario. The input voltage controls the desired displacement; therefore, an athlete performs well when turning on the thrust assist bar.

7. Conclusion

- In this article, the deformed shape of the plate bonded to piezoelectric actuators and sensors is investigated by the finite element method (FEM). Based on the Kirchhoff plate model, the FE (finite element) formula was developed for smart composite structures with piezoelectric materials. The modeling technique combining FEM and LQR algorithms was proposed. By analyzing and evaluating the results obtained from the two cases, the effectiveness of computation and application was presented.

- In the first case, a rectangular piezoelectric actuator with symmetrically bonded three sensors is considered. The results are analyzed, errors estimated and compared with other methods and reliable analytical solutions.
- In the second case, the piezoactuator of the injector is considered. From the obtained results, it can be said that the deformed shape of piezoactuators can be computed accurately by FEM. Furthermore, the obtained results also showed that the displacement at the center of the piezoactuator in the injector has a linear relation with the voltage level. These results help the understanding the coupling effects of mechanical and electric properties of piezoelectric actuators and sensors for applications. Besides, the results present the performance of active control LQR and behavior of the actuator modules of the injector in a common system. The convergence rate of the energy norm versus the degree of freedom is reliable.

References

1. Abreu G.L.C.M. de, Ribeiro J.F., Steffen V. Jr., 2004, Finite element modeling of a plate with localized piezoelectric sensors and actuators, *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, **26**, 2, 117-128

2. Adriaens H.J.M.T.S., De Koning W.L., Banning R., 2000, Modeling piezoelectric actuators, *IEEE/ASME Transactions on Mechatronics*, **5**, 4, 331-341

3. Bailey T., Hubbard J.E., 1985, Distributed piezoelectric-polymer active vibration control of a cantilever beam, *Journal of Guidance, Control, and Dynamics*, **8**, 5, 605-611

4. Benjeddou A., Trindade M.A., Ohayon R., 2000, Piezoelectric actuation mechanisms for intelligent sandwich structures, *Smart Materials and Structures*, **9**, 3, 328-335

5. Bhalla S., Moharana S., Talakokula V., Kaur N., 2017, *Piezoelectric Materials: Applications in SHM, Energy Harvesting and Biomechanic*, Wiley

6. Cao X., Tanner G., Chronopoulos D., 2020, Active vibration control of thin constrained composite damping plates with double piezoelectric layers, *Wave Motion*, **92**, 102423

7. Chen L., Xue J., Pan S., Chang L., 2020, Study on cantilever piezoelectric energy harvester with tunable function, *Smart Materials and Structures*, **29**, 7, 075001

8. Chen W., Liu Y., Liu Y., Tian X., Shan X., Wang L., 2018, Design and experimental evaluation of a novel stepping linear piezoelectric actuator, *Sensors and Actuators A: Physical*, **276**, 259-266

9. Chen X., Su C.-Y., Li Z., Yang F., 2016, Design of implementable adaptive control for micro/nano positioning system driven by piezoelectric actuator, *IEEE Transactions on Industrial Electronics*, **63**, 10, 6471-6481

10. Crawley E.F., de Luis J., 1987, Use of piezoelectric actuators as elements of intelligent structures, *AIAA Journal*, **25**, 10, 1373-1385

11. Dileep K.K., Subba Rao V.V., 2017, Vibration control of rectangular cross-ply FRP plates using PZT materials, *Journal of Engineering Science and Technology*, **12**, 12, 3398-3411

12. Dimitriadis E.K., Fuller C.R., Rogers C.A., 1991, Piezoelectric actuators for distributed vibration excitation of thin plates, *Journal of Vibration and Acoustics*, **113**, 1, 100-107

13. Gohery S., Mozafari F., Moslemi N., Mouloodi S., Alebrahim R., *et al.*, 2022, Static and dynamic deformation response of smart laminated composite plates induced by inclined piezoelectric actuators, *Journal of Composite Materials*, **56**, 21, 3269-3293

14. Her S.-C., Chen H.-Y., 2022, Vibration excitation and suppression of a composite laminate plate using piezoelectric actuators, *Materials*, **15**, 6, 2027

15. Hoa P.Q., Van T.T., Dat P.T., Hau D.T., Ha N.V., *et al.*, 2018, Static and free vibration analyses of laminated composite shells by cell-based smoothed discrete shear gap method (CS--DSG3) using three-node triangular elements, *Vietnam Journal of Mechanics*, **40**, 1, 89-103
16. Huang G.L., Sun C.T., 2006, The dynamic behaviour of a piezoelectric actuator bonded to an anisotropic elastic medium, *International Journal of Solids and Structures*, **43**, 5, 1291-1307
17. Jafferis N.T., Lok M., Winey N., Wei G.-Y., Wood R.J., 2016, Multilayer laminated piezoelectric bending actuators: design and manufacturing for optimum power density and efficiency, *Smart Materials and Structures*, **25**, 5, 055033
18. Karegar M., Bidgoli M.R., Mazaheri H., 2021, Smart control and seismic analysis of concrete frames with piezoelectric layer based on mathematical modelling and numerical method, *Structures*, **32**, 1171-1179
19. Latrache M., Menasri N., 2022, Active control vibration of a smart composite plate for various boundary conditions, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, **13**, 2, 671-688
20. Lopes V. Jr., Pereira J.A., Inman D.J., 2000, Structural FRF acquisition via electric impedance measurement applied to damage location, [In:] *IMAC XVIII. Proceedings of IMAC--XVIII: A Conference on Structural Dynamics*, San Antonio, USA, 41, 1549-1555
21. Lumentut M.F., Howard I.M., 2014, Electromechanical finite element modelling for dynamic analysis of a cantilevered piezoelectric energy harvester with tip mass offset under base excitations, *Smart Materials and Structures*, **23**, 9, 095037
22. Luo Q., Tong L., 2006, High-precision shape control of plates using orthotropic piezoelectric actuators, *Finite Elements in Analysis and Design*, **42**, 11, 1009-1020
23. Moretti M., Silva E.C.N., 2019, Topology optimization of piezoelectric bi-material actuators with velocity feedback control, *Frontiers of Mechanical Engineering*, **14**, 2, 190-200
24. Phung-Van P., Nguyen-Thoi T., Le-Dinh T., Nguyen-Xuan H., 2013, Static and free vibration analyses and dynamic control of composite plates integrated with piezoelectric sensors and actuators by the cell-based smoothed discrete shear gap method (CS-FEM-DSG3), *Smart Materials and Structures*, **22**, 9, 095026
25. Reddy J.N., 1999, On laminated composite plates with integrated sensors and actuators, *Engineering Structures*, **21**, 7, 568-593
26. Saravanos D.A., Heyliger P.R., 1999, Mechanics and computational models for laminated piezoelectric beams, plates, and shells, *Applied Mechanics Reviews*, **52**, 10, 305-320
27. 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
28. Uchino K., 1986, Electrostrictive actuators: materials and application, *American Ceramic Society Bulletin*, **65**, 4, 647-652
29. Uchino K. (Edit.), 2010, *Advanced Piezoelectric Materials: Science and Technology*, Woodhead Publishing Limited
30. Wang X., Shen Y., 1998, On the characterization of piezoelectric actuators attached to structures, *Smart Materials and Structures*, **7**, 3, 389-395
31. Wei H., Wang H., Xia Y., Cui D., Shi Y., *et al.*, 2018, An overview of lead-free piezoelectric materials and devices, *Journal of Materials Chemistry C*, **6**, 46, 12446-12467

Manuscript received April 8, 2024; accepted for publication October 18, 2024

JOURNAL OF THEORETICAL
AND APPLIED MECHANICS
63, 1, pp. 197-209, Warsaw 2025
https://doi.org/10.15632/jtam-pl/195836

EXPLORATION OF FREQUENCY CHARACTERISTICS OF BALL BEARING WITH UNEVEN LOAD BEARING EFFECT CAUSED BY WEAR FAILURE OF MULTIPLE ROLLING BODIES UNDER STARVED LUBRICATION

Zhong-Tang Huo, Jian-Qi Chen, Ling-Juan Hao, Jian-Song Gao

Handan University, Mechanical and Electrical College, Handan, China

corresponding author Zhong-Tang Huo, e-mail: huozhongtang123@hdc.edu.cn

Pointed at the problem of increased wear of rolling bodies (RBs) in a bearing under starved lubrication as well as abnormal vibration and increased noise of the bearing after occurence of wear fault, a fault dynamics model was proposed to simulate the interaction between components by quantifying the degree of rolling bodies wear. The results indicated that the bearing exhibited an uneven load bearing effect after the wear fault. The fractional multiple of the rotation frequency of the cage could be used as the basis for monitoring wear fault of rolling bodies. The research provides a reference for the diagnosis of wear fault in rolling bodies.

Keywords: starved lubrication, dynamic model, wear, bearing, frequency

1. Introduction

Ball bearings play an important role in modern industrial fields and are also a key link in the health monitoring and diagnosis of mechanical systems. The importance of their health status to the entire system cannot be ignored (Cao *et al.*, 2019; Visnadi and de Castro, 2019). Lubrication is an important factor affecting the performance of ball bearings during operation. In many cases, such as deterioration or reduction of the lubricant caused by extreme working conditions such as aircraft engines, the lubrication state of the bearing will transition to starved lubrication. The contact roughness between RBs and raceways will increase, and friction between internal components of will increase, thereby accelerating bearing wear. The levels of vibration and noise of the system are greatly grown, shortening the service life of the equipment (Wang *et al.*, 2018; Wen *et al.*, 2023). Therefore, for ensuring healthy and stable operation of the system, it is crucial to obtain accurate wear characteristics of RBs for the bearing condition monitoring.

The dynamic model is an important method to obtain various fault characteristics of bearings, it clarifies the fault mechanism, and establishes the mapping relationship between the fault excitation and characteristic frequency. Jiang *et al.* (2019) explored vibration mutation caused by contact force changes, Bai *et al.* (2021) explored the trend of outer ring fault frequency changes using temperature field thermal deformation theory, and for the study of RB fault, Zhang *et al.* (2021) established a dynamic model of the outer ring – RB composite faults to obtain fault characteristics. Cheng *et al.* (2019) set a local defect on the surface of RB and obtained the relationship between the geometric shape of the defect and deformation, the occurrence of local defect on RB significantly increased vibration of the bearing system. However, a wear fault cannot directly trigger impact forces like surface defects, so the local fault dynamics modeling method cannot be directly applied, but should be studied based on changes in size or spherical roundness. The above research clearly has low applicability. In addition, when conducting research on bearing dynamics, in order to reduce computational complexity, the bearings are generally considered to be in dry friction or a fully lubricated state (Liu *et al.*, 2020).

However, in actual use of bearings, it is almost impossible to achieve ideal lubrication conditions, and due to the influence of extreme working conditions, most bearings are in a state of starved lubrication (Maruyama and Saitoh, 2015). Bian *et al.* (2021) considered the influence of insufficient lubrication conditions on the bearing performance, and showed that when the lubricant oil supply was insufficient, the bearing was in the starved lubrication state, and the bearing bearing capacity of the oil film was reduced, which easily led to bearing wear. Regarding friction and wear behavior of RBs, Shi *et al.* (2022) developed a friction dynamics analysis model, which can be used to evaluate the lubrication and wear status of bearings, becoming the theoretical basis for establishing a friction lubrication model in this paper. Tu (2021) developed a dynamic model for sliding of cylindrical roller bearings. The most direct consequence of RB wear is a decrease in ball diameter. Regarding the difference in ball diameter, Bai *et al.* (2019) established a dynamic model suitable for full ceramic ball bearing considering RB size errors. Shi and Bai (2020) pointed out that the non-uniform loading situation varies with the diameter tolerance and arrangement of the RB, and the characteristic frequency of the RB raceway contact shifts from f_c to f_c', $2f_c'$ and $3f_c'$. By relying on the idea of ball diameter difference, this paper evaluates the phenomenon of uniform wear of RBs, and provides valuable materials for the study of dynamic models of RB wear.

From the above analysis, it can be seen that it is necessary to establish a dynamic model with RBs wear fault considering starved lubrication and obtain the corresponding vibration response of the fault. Therefore, the difference in ball diameter to simulate RBs wear fault was used in this paper, and based on the theory of friction-lubrication, the frequency characteristics of RBs wear fault under starved lubrication was obtained. The research results provide a sufficient reference for bearing wear fault monitoring in practical situations.

2. Establishment of the dynamic model for RB wear fault under starved lubrication

As shown in Fig. 1, $O_iY_iZ_i$ represents the center of mass coordinate system of the inner ring, and $O_jY_jZ_j$ represents the center of mass coordinate system of the j-th RB. The RB revolves around the outer ring axis at an angular velocity of ω_{mj}, while the cage rotates around the outer ring axis at ω_c. ϕ_j is the azimuth angle of the j-th RB in the coordinate system $O_iY_iZ_i$, and ψ_j is the azimuth angle of the j-th RB in the coordinate system OYZ.

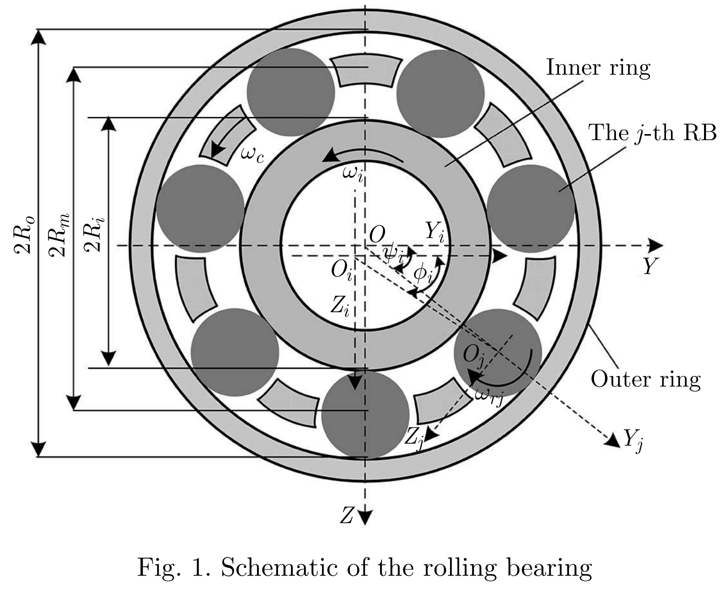


Fig. 1. Schematic of the rolling bearing

The azimuth ψ_j of the j-th RB can be expressed as

$$\psi_j = \omega_{mj}t + \theta_j + \frac{2\pi(j-1)}{N_b} \tag{2.1}$$

The diameter of the RB decreases after wear, and its load-bearing performance changes. D_j is the diameter of the RB, therefore, the relationship between ϕ_j and ψ_j is

$$\psi_j = \varphi_j + \arccos\frac{R_o^2 + R_i^2 - D_j(R_o - R_i) + 0.5D_j^2 - e^2}{2R_oR_i + D_j(R_o - R_i) - 0.5D_j^2} \tag{2.2}$$

where e is the eccentricity.

The deformation and force relationship between the RB and the raceway are as follows

$$\delta_j = \Big[y\cos\varphi_j + z\sin\varphi_j - \frac{1}{2}\lambda - (D_j - D_b)\Big]_+ \tag{2.3}$$

where "+" indicates that the equation takes only positive values. The contact force can be expressed as

$$Q_{ij} = K\beta\sqrt{\delta_j^3} \qquad Q_{oj} = K\beta\sqrt{\delta_j^3} + m_rR_m\omega_{mj}^2 \tag{2.4}$$

where K is the equivalent contact stiffness (Wang *et al.*, 2022); β is the switching function, it takes 1 when $\delta_j \geqslant 0$, otherwise it is 0.

Set the bearing area to cover the lower 1/3 symmetrical part of the bearing, as shown in Fig. 2. Due to wear on the RB, there is a difference between the actual ball diameter and the nominal diameter of the RB. Therefore, when the RB comes into contact with both the inner and outer rings, it is considered a load-bearing RB.

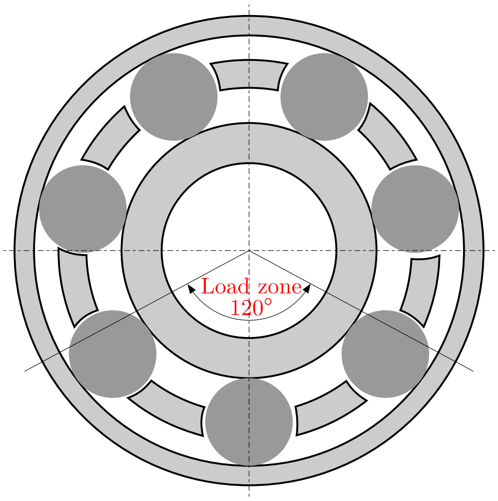


Fig. 2. Schematic of the bearing bearing area

When there is wear on the RB, that is, when there is a difference in the ball diameter of the RB, the ball diameter of the RB satisfies

$$D_j = D_n + A_j\delta \tag{2.5}$$

where A_j is the diameter coefficient corresponding to the j-th RB, and δ is the amplitude of the diameter difference.

When the j-th RB passes through the bearing area and D_{jlim} is given, the critical value for determining whether the RB is loaded can be obtained

$$D_{j\,\mathrm{lim}} = \frac{e^2 + R_o^2 - R_i^2 - 2eR_o(\psi_j - \psi_e)}{R_o + R_i - e(\psi_j - \psi_e)} \tag{2.6}$$

where ψ_e is the eccentricity angle of the inner ring, when $D_j \geqslant D_{j\,\mathrm{lim}}$, the j-th RB bears the load.

2.1. The traction force between RB and raceway under starved lubrication

When the RB and raceway are in the starved condition, due to the existence of micro-asperity contact friction between the two surfaces, the force relationship between the two contact surfaces will be affected, and one will need to calculate the oil film thickness and traction force under the starved condition, as shown in Fig. 3.

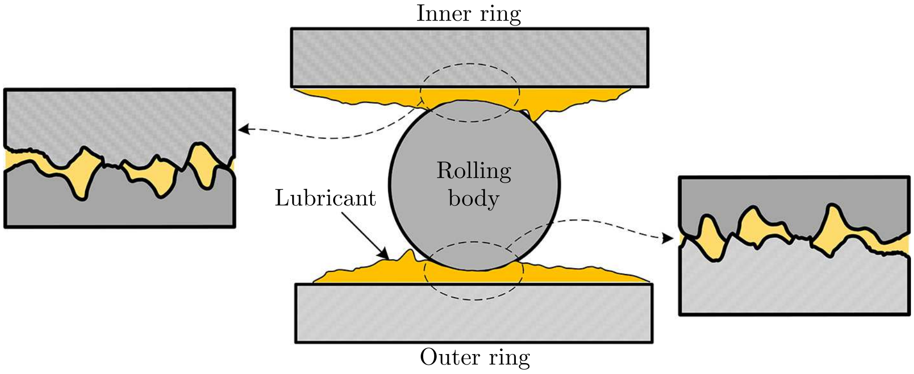


Fig. 3. Schematic of the contact surface between the RB and the raceway

The oil film thickness and traction force under the starved condition are calculated as

$$\begin{aligned} h_c &= 3.672 R_y W^{-0.045\kappa^{0.18}} U^{0.663\kappa^{0.025}} G^{0.502\kappa^{0.064}} \\ &\quad \cdot (1 + 0.025\overline{\sigma}^{1.248} V^{0.119} W^{-0.133} U^{-0.884} G^{-0.977} \kappa^{0.081})(1 - 0.573e^{-0.74\kappa}) \end{aligned} \tag{2.7}$$

and

$$(h_c)_{starved} = (1 - 1.561\chi^{0.849}\kappa^{-0.214})h_c \qquad \chi = 1 - \frac{\dot{m}_s}{\dot{m}_f} \tag{2.8}$$

where h_c is the film thickness in fully EHL, $(h_c)_{starved}$ is the film thickness in the state of starved lubrication, W is the dimensionless load, U is the dimensionless speed, G is the dimensionless material number, and V denotes the dimensionless hardness number, κ is the ellipticity coefficient, χ denotes the degree of starvation, $\dot{m}_s$ and $\dot{m}_f$ is the lubricant mass flow rate under starved conditions and fully EHL conditions, respectively.

The asperity contact ratio $L_{a-starved}$ of the micro-convex body between the RB and the raceway is calculated as

$$L_{a-starved} = 10W^{-0.083}U^{0.143}G^{0.314} \ln\Big(1 + (1-\chi)^{-7.326}\overline{\sigma}^{4.689}V^{0.509}W^{-0.501}U^{-2.9}G^{-2.87}\Big) \tag{2.9}$$

After quantifying the asperity contact ratio, the pressure distribution between the RB-raceway can be further determined

$$P_a = P_{max}\frac{L_{a-starved}}{100} \qquad P_h = P_{max}\Big(1 - \frac{L_{a-starved}}{100}\Big) \tag{2.10}$$

where P_{max} is the maximum contact pressure between the RB and the raceway contact surface, P_a is the contact pressure borne by the micro-asperity body, P_h is the contact pressure borne by the oil film.

Therefore, the traction force between the RB and the raceway can be calculated as

$$\begin{aligned}&\tau(x',y')=\min\Big[\frac{\eta(x',y')\Delta V}{(h_c)_{starved'}},\Lambda P_h\Big]+f_cP_a\\&F_{i/oj}=\iint\tau(x',y')\,dx'\,dy'\end{aligned}\tag{2.11}$$

where $\tau(x',y')$ is the oil film shear stress, $\eta(x',y')$ is the lubricant viscosity, ΔV is the sliding velocity between the RB and the raceway, Λ is the ultimate shear stress coefficient, and f_c denotes the roughness friction coefficient.

2.1.1. Establishment of bearing motion model

According to the Hertz theory and the geometric model of the bearing in Fig. 1, motion of the inner ring can be described as

$$\begin{aligned}&W_y-\sum_{i=1}^{N_b}(Q_{ij}\cos\varphi_j-F_{ij}\sin\varphi_j)-m_i\frac{d^2y}{dt^2}=c\frac{dy}{dt}\\&W_z-\sum_{i=1}^{N_b}(Q_{ij}\sin\varphi_j+F_{ij}\cos\varphi_j)-m_i\frac{d^2z}{dt^2}=c\frac{dz}{dt}\end{aligned}\tag{2.12}$$

where c is damping.

Motion of the RB can be described as

$$\begin{aligned}&F_{ij}-F_{oj}-Q_{cj}-m_rg\cos\psi_j-m_rR_m\frac{d\omega_{mj}}{dt}=0\\&(F_{ij}+F_{oj})R_r-m_rR_m\omega_{mj}^2R_r-m_rR_r^2\frac{d\omega_{rj}}{dt}=0\end{aligned}\tag{2.13}$$

where ω_{mj} is the circumferential angular velocity of the j-th RB

$$\omega_{mj}=\frac{d\varphi_j}{dt}\tag{2.14}$$

Motion of the cage can be described as

$$R_m\sum_{i=1}^{N_b}Q_{cj}-J_c\frac{d\omega_c}{dt}=0\tag{2.15}$$

where J_c is the moment of inertia of the cage.

3. Numerical simulation

The 6304 deep groove ball bearing is taken as an example for the study, and the parameters of the bearing and the performance parameters of the lubricant are shown in Table 1. The operating temperature is 40°C, radial load $W=100\,\mathrm{N}$, inner ring speed $\omega_i=4000\,\mathrm{r/min}$.

When the bearing area is uniformly loaded, the number of loaded RBs $N=N_b/3$, and the number of RBs in the bearing studied in this paper is $N_b=7$. Based on Fig. 2, it can be seen that the number of loaded RBs in the bearing area varies between 2 and 3 without considering the difference in ball diameter. This paper introduces wear in three RBs, assuming the radii of

Table 1. Main parameters of rolling bearing

Parameters	Value	Parameters	Value
Outer raceway radius R_o [mm]	22.77	Raceway surface roughness σ_1 [μm]	0.028
Inner raceway radius R_i [mm]	13.24	Surface roughness of ball σ_2 [μm]	0.042
Bearing pitch circle radius R_m [mm]	18	Material cypress pine ratio υ	0.3
RB radius R_r [mm]	4.76	Density of lubricant ρ [Kg/m^3]	826
Number of RBs N_b	7	Lubricant in viscosity η_0 [Pa·s]	0.04
Radial clearance ε [mm]	10e-5	Ultimate shear stress coefficient $\varLambda$	0.0434

the three worn RBs are 4.71 mm, 4.71 mm and 4.70 mm, respectively. The analysis of the sliding effect and force is based on the worn RB. Using a fixed time step ($\varDelta t = 5 \cdot 1\text{e-}4$), the fourth-order Runge-Kutta algorithm in Matlab is incorporated to solve differential equation systems.

The above analysis indicates that there may be 1-3 load-bearing RBs in the bearing area, as shown in Fig. 4. Figure 4 shows the situation of different numbers of load-bearing RBs or worn RBs in the bearing area, where the worn RBs are represented by dashed lines. The solid line represents healthy RBs.

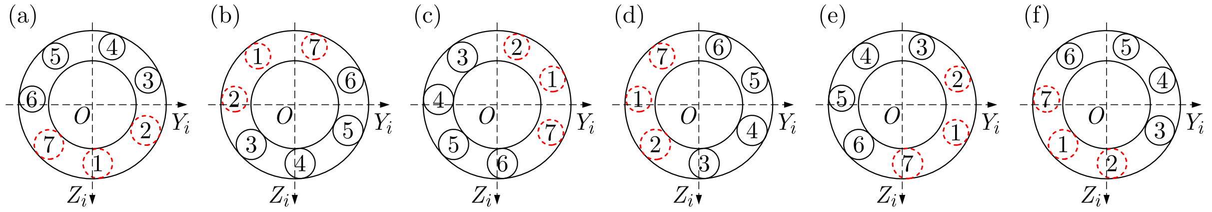


Fig. 4. Different bearing conditions

Numerical simulation methods are used to solve the model, as shown in Fig. 5. In Fig. 5a, the RB enters the contact zone when the azimuth angle is around $\phi_j = 300°$, and it can be observed that the oil film thickness between the RB and the raceway decreases with an increase of the starved lubrication degree. The reason for this phenomenon is that the supply of the lubricant is insufficient, and the rotating pair cannot stimulate enough lubricant to form an oil film. In the load area, due to the effect of gravity and load, the oil film is compressed, and thickness of the oil film is much lower than that in the non loaded area.

An increase in the degree of starved lubrication leads to an increase in friction between the RB-raceway, which drives the RB to rotate around the center of the bearing at an increased rate, i.e., the rotational speed of the RB and cage increases with the increase in the degree of starved lubrication, as shown in Figs. 5b and 5c. In Fig. 5b, the dashed line represents the rotational speed of the RB experiencing wear failure, which is lower than the healthy RB under the same lubrication conditions. Figure 5d shows the RB-raceway slip velocity, in the contact zone the RB is subjected to the radial load, so the slip velocity decreases rapidly. In addition it can be clearly observed that the slip rate decreases rapidly with an increase of the starved degree within one week of RB rotation.

Figure 6a shows the contact forces between the RBs and the raceway of the ball bearing without considering wear of the RBs, and the contact forces of each RB in the load zone are almost equal. Figures 6b-6d show the bearing capacity of each RB under the initial condition $\chi = 0$. Figure 6b corresponds to Fig. 4a, where the contact force of RBs 1, 2 and 7 is significantly lower than that of other RBs, exhibiting a significant uneven bearing effect. The 6th and 3rd RBs adjacent to the 7th and 2nd RBs have the highest contact force. This is because the diameter of the 1st, 2nd and 7th RBs is small, and according to the Hertz theory, the Hertz force generated is small. Therefore, the contact force is distributed to the 6th and 3rd RBs. Figures 6c and 6d are consistent with this description principle, where the RB with smaller diameter bears less

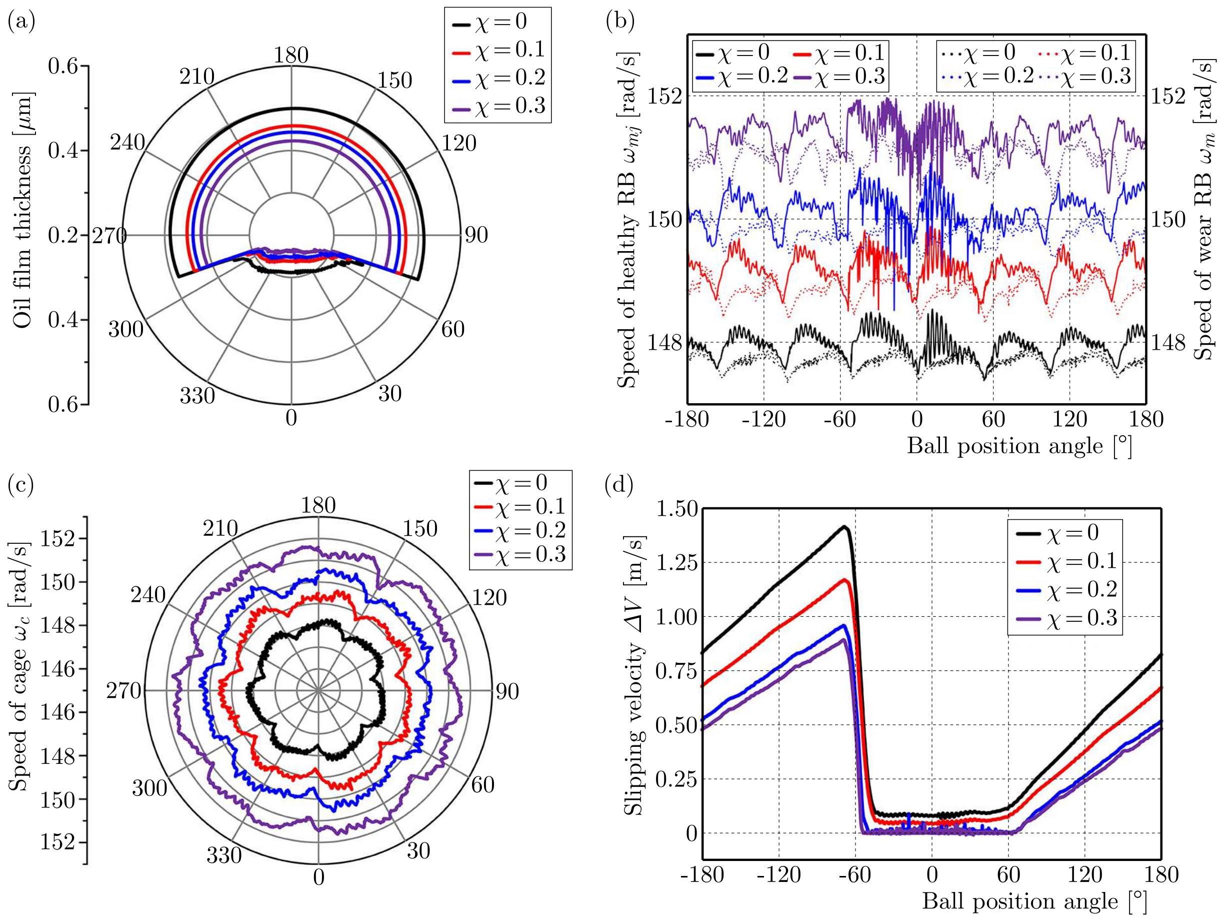


Fig. 5. Related calculation results: (a) the influence of starved lubrication on oil film thickness, (b) revolution speed of RB, (c) cage speed, (d) sliding speed of RB to raceway

pressure, while the adjacent RBs have a greater contact force, resulting in a significant uneven bearing effect.

Figure 7 shows variation of rhe contact force between each RB and the raceway at different χ values of 0, 0.1, 0.2 and 0.3. Figure 7 more intuitively reflects that the contact force of RBs 1, 2 and 7 is significantly lower than the others. And the 1st RB is adjacent to both the 7th and 2nd RBs, so a portion of the contact force from the 7th and 2nd RBs is shared on the 1st RB. In addition, as the degree of starved lubrication increases, the thickness of the oil film inside the bearing decreases, the contact stiffness K increases, resulting in an overall upward trend in the contact force (Wang *et al.*, 2022).

Equation (3.1) provides the calculation method for the friction torque between the RB and the raceway. Figure 8 shows variation of the friction torque for different degrees of starved lubrication. The curves A, B, C and D represent the depleted oil conditions at $\chi = 0$, 0.1, 0.2 and 0.3, respectively. After the wear of the RBs, as the degree of starved lubrication increases, the friction torque increases. In addition, changes in the friction torque can affect the degree of bearing sliding. According to Tu *et al.* (2021), the more sufficient friction between the RB and the raceway, the greater the driving force provided for the RB and the cage, thereby suppressing bearing sliding. With an increase of χ, that is an increase of the friction torque, the slip effect is reduced, verifying a decrease in slip velocity, which is consistent with Tu's research results. Meanwhile, this also confirms the conclusion in Fig. 5 that the sliding effect is suppressed as the degree of starved lubrication increases after the wear of the RB

$$M_{f,average} = \mathrm{AVERAGE}\Big[\sum_{j=1}^{N_b}(F_jR_j + F_oR_o)\Big] \tag{3.1}$$

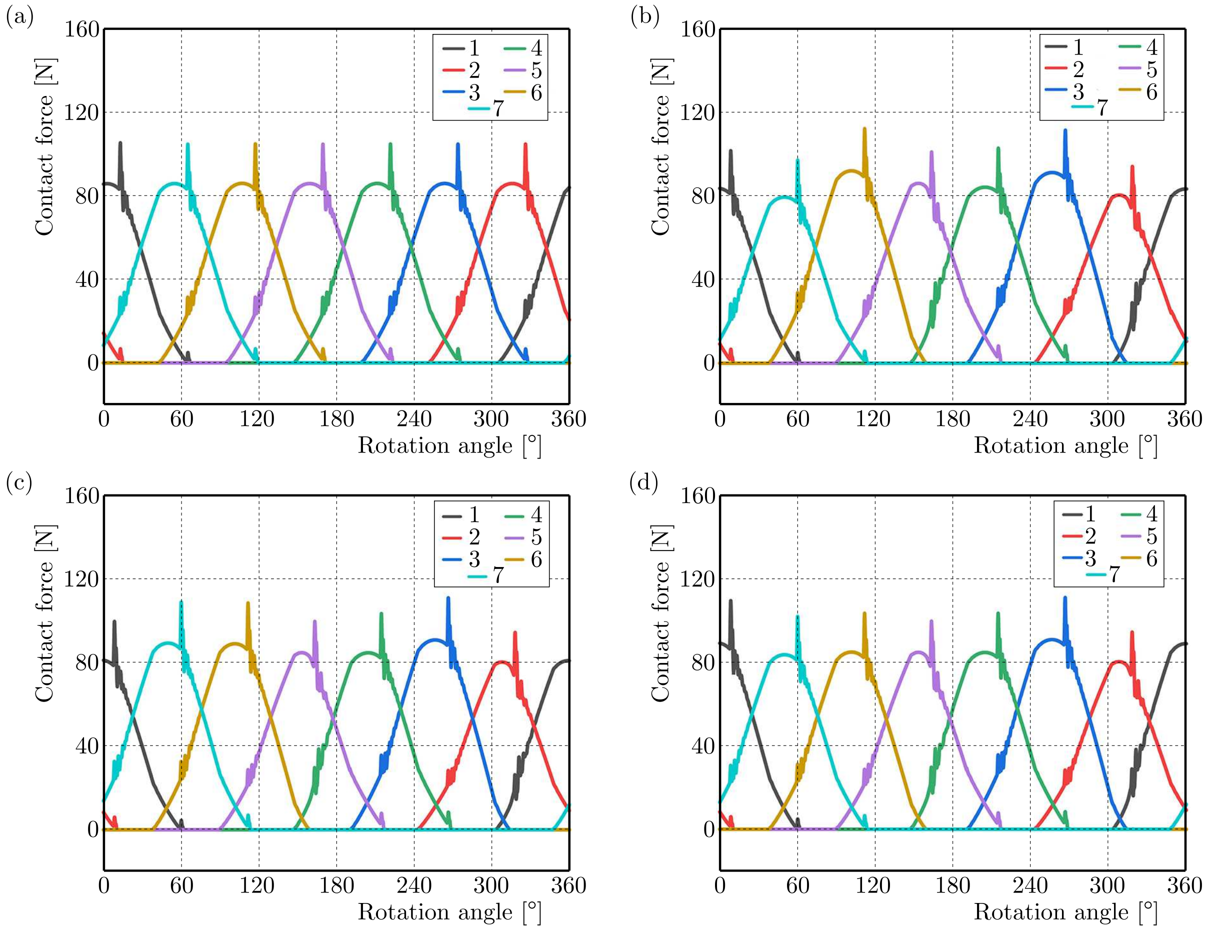


Fig. 6. The bearing capacity of each RB under the initial lubrication state

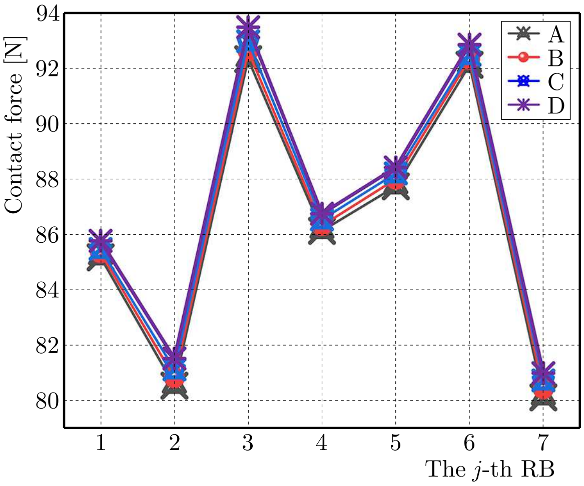


Fig. 7. The bearing capacity of each RB under different degrees of starved lubrication

The spectrum of Fig. 9 was obtained by solving the differential equation. Nf_r is the rotational frequency, where $nf_{cA,B,C,D}$ correspond to the rotational frequency and doubling frequency of the cage under four lubrication states, while $f_{bwA,B,C,D}$ correspond to the characteristic frequencies of RB wear faults. According to Shi's research (Shi *et al.*, 2019), when uneven load bearing effects occur inside the bearing, a fractional multiple of the cage frequency f_c, $f_{cw} = 3f_c/N$, appears in the spectrum. The above formula provides the characteristic frequency when there is

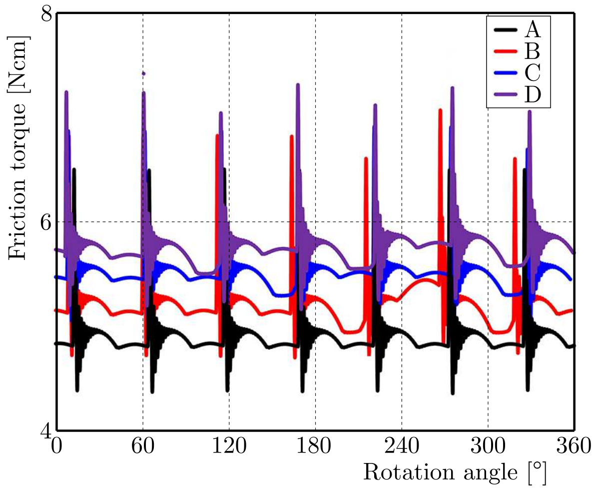


Fig. 8. Friction torque between the RB and raceways under different degrees of starved lubrication

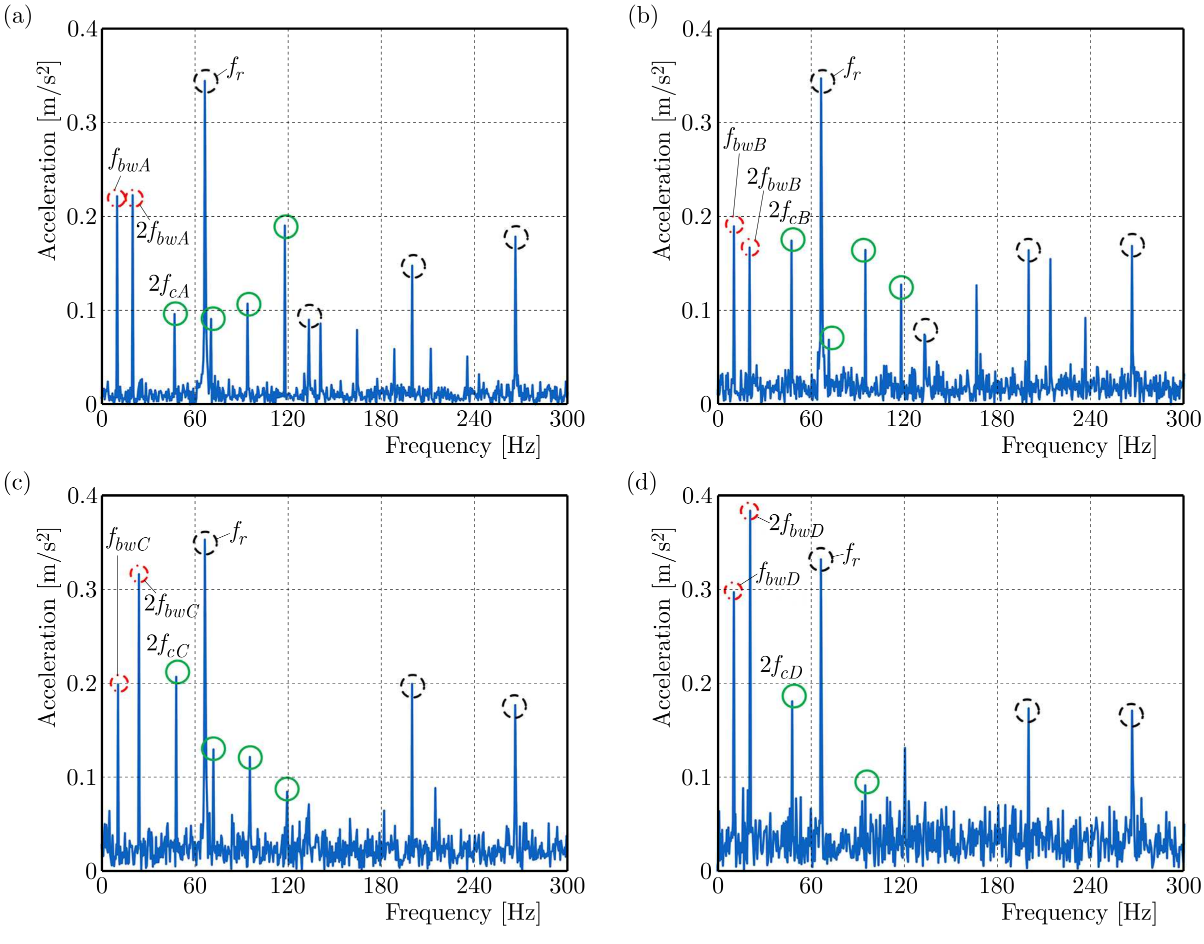


Fig. 9. Vibration spectrum of bearing dynamics model simulation: (a) A, (b) B, (c) C, (d) D

only one loaded RB in the bearing area, and when there is $2, 3, \dots, N/3$ in the bearing area, the corresponding characteristic frequencies are $2f_{cw}, 3f_{cw}, \dots, Nf_{cw/3}$. For the bearings selected in this article, there are 7 RBs, and there may be 1,2,3 loaded RBs simultaneously in the bearing area. When the worn RBs enter the bearing area, there are 2 loaded RBs in that area, and the corresponding characteristic frequency in the spectrum is $2f_{cw}$. After the worn RB leaves the bearing area, there are three loaded RBs in that area. At this time, $3f_{cw} > f_c$, and the load-bearing characteristics of bearings after RB wear will appear as f_c and its harmonic components.

As the degree of starved lubrication, the friction torque between the RB and the raceway increases, and the sliding effect is suppressed. This is manifested by an increase in the rotational frequency f_c and its harmonic component nf_c of the cage, but still lower than the theoretical value in ideal conditions. The spectrum shows an increase in f_{cA}, f_{cB}, f_{cC} and f_{cD}, corresponding to an increase in the RB wear frequency f_{bwA}, f_{bwB}, f_{bwC} and f_{bwD}, as shown in Fig. 9 and Table 2.

Table 2. Characteristic frequencies at different degrees of starved lubrication

Item	$\chi = 0$	$\chi = 0.1$	$\chi = 0.2$	$\chi = 0.3$	Change rate
f_c [Hz]	23.83	23.95	24.19	24.32	1.9%
f_{cw} [Hz]	20.19	20.38	20.80	20.92	3.5%

4. Experimental study

For verifying the correctness of the model proposed in this paper and detect the bearing vibration characteristics of RB wear fault under starved lubrication, an experimental setup as shown in Fig. 10 was established in this Section. In Fig. 10, the speed was controlled by a speed controller, and the motor output speed was from 0 to 5000 r/min. The vibration signal was obtained by a contact type acceleration sensor arranged on the bearing seat and further processed by a data collector, with a sampling frequency of 5000 Hz. The selected bearings were consistent with those in the simulation in the previous Section. Through special customization and procurement, the RB had a radius of 4.70 mm, the radius of the two RBs is 4.71, while the other RBs had the radius of 4.76 mm. This paper designed three sets of experiments P, Q and R. By manually adding or removing lubricants, the lubricants in the bearing cavity during the three sets of experiments were 0%, 33% and 75% of the cavity volume, respectively, to simulate changes in the degrees of starved lubrication.

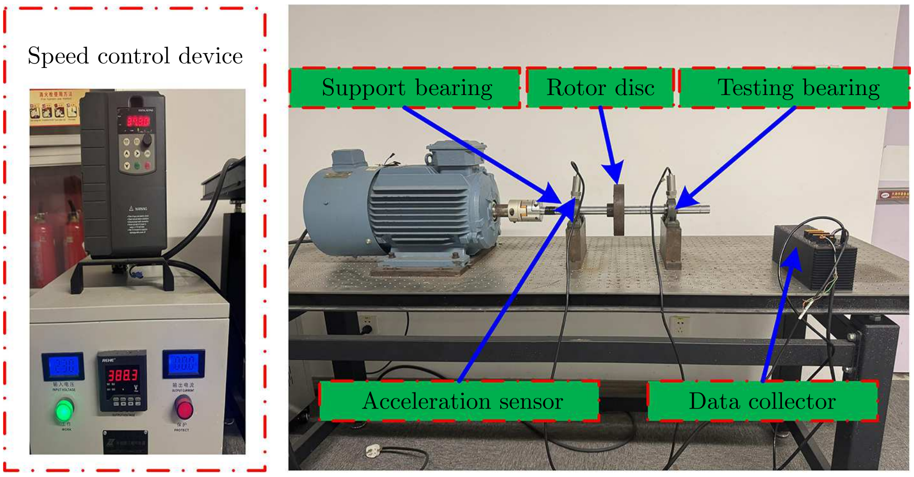


Fig. 10. Experimental testing system

Set the motor output speed to 4000 r/min, collect vibration data after the speed stabilizes, and perform Fourier transform on the vibration signal to obtain the spectrum of the right bearing as shown in Fig. 11. In the spectrum, it is clear to see changes in the characteristic frequencies $f_{cwP,Q,R}$ of RB wear caused by changes in lubricant supply, while the rotational frequency f_r and its harmonic components related to input speed remain unchanged. In the P, Q and R tests, the corresponding RB wear characteristic frequencies were 10.02 Hz, 10.24 Hz and 10.78 Hz, respectively. The corresponding double wear characteristic frequencies are approximately in the range of 20 Hz-22 Hz, which is consistent with the simulation results. The corresponding cage rotation frequency and its harmonic components were also excited. However, due to frequency modulation and mutual influence, the amplitude of the first harmonic frequency $f_{cP,Q,R}$ of the cage rotation was relatively low, which may be submerged in background noise in practice. However, the RB wear characteristic frequencies still vary significantly.

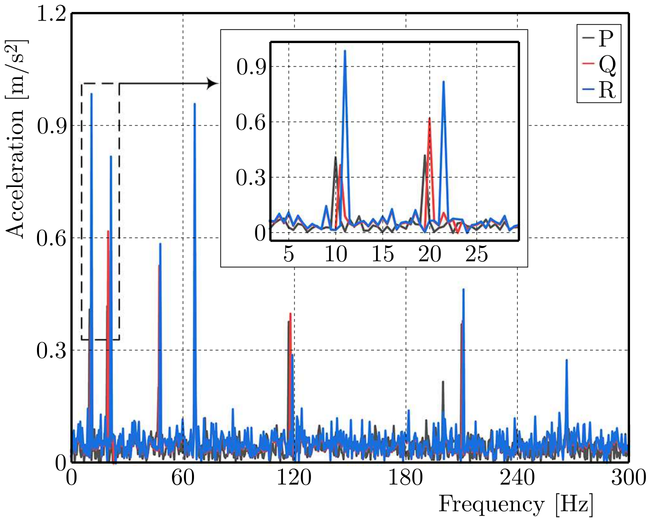


Fig. 11. Frequency curve of experimental data

Compared with simulation, it is difficult to accurately simulate the increase and decrease of lubricants. However, this experiment still verified the hypothesis that the characteristic frequency changed with the degrees of starved lubrication, and quantified the characteristics of RB wear fault under different degrees of starved lubrication, proving the correctness of the theoretical model and calculation.

5. Conclusion

This paper quantified the wear failure of ball bearings under starved lubrication based on the idea of ball diameter difference of RBs, and the mechanism of fault characteristics was analyzed by contact-friction theory. The main conclusions are as follows:

(1) RB wear failure is the cause of an uneven bearing load. Specifically, the contact force between the worn RB and the ring is small, while the adjacent healthy RBs experience a greater contact force.

(2) Deterioration of lubrication conditions will exacerbate the wear failure of RB. Specifically, as the degrees of starved lubrication increases, the same RB will experience greater contact and friction forces, leading to wear faults quickly radiating to healthy RBs.

(3) At the same lubrication level, the worn RB speed is slightly lower than healthy RBs. Specifically, due to wear of RB, it experiences a lower contact force and lower frictional driving force than healthy RB, resulting in a decrease in orbital speed.

(4) The fractional multiple of the rotation frequency of the cage can be used as an indicator for monitoring the RB wear fault. Specifically, considering the total number of RBs, the number of RBs loaded within the load zone, and the number of RBs experiencing wear and tear fault, relevant frequency components can be found in the spectrum.

(5) A change in the lubrication state will have an impact on the characteristic frequency of wear fault. In the actual monitoring of the ball bearing wear fault, lubrication conditions should be fully considered to effectively distinguish characteristic frequencies and ensure the accuracy of fault monitoring.

In the future, we will focus on the temperature rise phenomenon caused by bearing operation heating, fully consider the influence of temperature rise on lubrication conditions, radial clearance, and component dimensions, and incorporate the wear phenomenon of raceways into our research. We will establish a more detailed, comprehensive, and applicable bearing wear dynamic model to improve the theory of bearing fault dynamics.

References

1. Bai X.T., Wu Y.H., Rosca I.C., Zhang K., Shi H.T., 2019, Investigation on the effects of the ball diameter difference in the sound radiation of full ceramic bearings, *Journal of Sound and Vibration*, **450**, 231-250
2. Bai X., Zheng H., Wang Z., Wang Z., 2021, Raceway defect frequency deviation of full-ceramic ball bearing induced by fit clearance in wide temperature ranges, *Shock and Vibration*, **2021**, 1, 6650798
3. Bian Z., Song C., Meng L.H., Zhou Q.H., 2021, Response of the interface between the ball and the raceway of rolling bearing under starved lubrication, *IOP Conference Series: Materials Science and Engineering*, **1043**, 5, 052005
4. Cao H., Shi F., Li Y., Li B., Chen X., 2019, Vibration and stability analysis of rotor-bearing--pedestal system due to clearance fit, *Mechanical Systems and Signal Processing*, **133**, 106275
5. Cheng H., Zhang Y., Lu W., Yang Z., 2019, Research on ball bearing model based on local defects, *SN Applied Sciences*, **1**, 10, 1219-1229
6. Jiang Y., Huang W., Luo J., Wang W., 2019, An improved dynamic model of defective bearings considering the three-dimensional geometric relationship between the rolling element and defect area, *Mechanical Systems and Signal Processing*, **129**, 694-716
7. Liu Y., Wang W., Qing T., Zhang Y., Liang H., Zhang S., 2020, The effect of lubricant temperature on dynamic behavior in angular contact ball bearings, *Mechanism and Machine Theory*, **149**, 103832
8. Maruyama T., Saitoh T., 2015, Relationship between supplied oil flow rates and oil film thicknesses under starved elastohydrodynamic lubrication, *Lubricants*, **3**, 2, 365-380
9. Shi H.T., Bai X.T., 2020, Model-based uneven loading condition monitoring of full ceramic ball bearings in starved lubrication, *Mechanical Systems and Signal Processing*, **139**, 106583
10. Shi H.T., Bai X.T., Zhang K., Wu Y.H., Yue G.D., 2019, Influence of uneven loading condition on the sound radiation of starved lubricated full ceramic ball bearings, *Journal of Sound and Vibration*, **461**, 114910
11. Shi X., Lu X., Feng Y., Qiu Z., 2022, Tribo-dynamic analysis for aero ball bearing with 3D measured surface roughness, *Engineering Failure Analysis*, **131**, 105848
12. Tu W., Yu W., Shao Y., Yu Y., 2021, A nonlinear dynamic vibration model of cylindrical roller bearing considering skidding, *Nonlinear Dynamics*, **103**, 3, 2299-2313
13. Visnadi L.B., de Castro H.F., 2019, Influence of bearing clearance and oil temperature uncertainties on the stability threshold of cylindrical journal bearings, *Mechanism and Machine Theory*, **134**, 57-73
14. Wang Y., Yan C., Lu Z., Liu Y., Wu L., 2022, Effect of thermal elastohydrodynamic lubrication on vibration characteristics of ball bearing with local defect, *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-Body Dynamics*, **236**, 3, 488-500

15. Wang Z., Yu Q., Shen X., Chen X., 2018, A simple model for scuffing risk evaluation of point contact under mixed lubrication, *Journal of Tribology*, **140**, 3, 031502
16. Wen C., Meng X., Gu J., Xiao L., Jiang S., Bi H., 2023, Starved lubrication analysis of angular contact ball bearing based on a multi-degree-of-freedom tribo-dynamic model, *Friction*, **11**, 8, 1395-1418
17. Wu X., Qin Y., Luo J., Wang S., Chen B., 2022, Fault dynamic model of high-speed rolling bearing by a compound displacement excitation function considering the effect of defect roughness, *Mechanism and Machine Theory*, **177**, 105061
18. Zhang X., Yan C., Liu Y., Yan P., Wang Y., Wu L., 2020, Dynamic modeling and analysis of rolling bearing with compound fault on raceway and rolling element, *Shock and Vibration*, **2020**, 1, 8861899

Manuscript received July 18, 2024; accepted for publication November 11, 2024

INFORMATION FOR AUTHORS

Journal of Theoretical and Applied Mechanics (***JTAM***) is devoted to all aspects of solid mechanics, fluid mechanics, thermodynamics and applied problems of structural mechanics, mechatronics, biomechanics and robotics. Both theoretical and experimental papers as well as survey papers can be proposed.

JTAM accepts full-text articles as well as short communications.

We accept articles in English only. The contents of a paper submitted to *JTAM* should not exceed 12 pages (short communication: 4 pages) of standard A4 format (11-point type size, standard margins – 2.5 cm, single line spacing) including abstract, figures, tables and references.

The material for publication should be sent to the Editorial Office via electronic journal system: http://www.editorialsystem.com/jtam

Papers are accepted for publication after the review process. Blind review model is applied, which means that the reviewers' names are kept confidential to the authors. Reviewer(s) declare that there is no interpersonal relation with the author(s) that would affect the opinion and recommendation of the article for publication in *JTAM*. The final decision on paper acceptance belongs to the Editorial Board.

Publication of a paper in the Journal is subject to fee. The fee is EUR 700 for a full length paper and EUR 350 for a short communication. The fee is due upon acceptance of the paper for publication by the Editorial Board.

After qualifying your paper for publication we will require LaTeX or TeX or Word document file and figures. The best preferred form of figures are files obtained by making use of editorial environments employing vector graphics.

All papers in *JTAM* are published online under the CC BY open access license.

Requirements for paper preparation

Contents of the manuscripts should appear in the following order:

- Title of the paper.
- Authors' full name, affiliation and e-mail.
- Short abstract (maximum 100 words) and 3-5 key words (1 line).
- Article text (equations should be numbered separately in each section; each reference should be cited in the text by the last name(s) of the author(s) and the publication year).
- References (maximum 25) in alphabetical order.
- Titles of references originally published not in English, should be translated into English.

All the data should be reported in SI units.

Contents