Volume 3, Issue 5, 2025

Print ISSN: 2959-9865
Online ISSN: 2959-9873

WORLD JOURNAL OF ENGINEERING RESEARCH

World Journal of Engineering Research

Volume 3, Issue 5, 2025

Published by Upubscience Publisher

World Journal of Engineering Research
Print ISSN: 2959-9865 Online ISSN: 2959-9873
Email: info@upubscience.com
Website: http://www.upubscience.com/

Table of Content

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3055

TEACHING OF ENGINEERING COST PRACTICAL TRAINING COURSES IN APPLIED UNDERGRADUATE INSTITUTIONS

HaoRan Zhou[1], MingYuan Yang[1], MiRen Rong[1], DongYang Geng[1,2*], JiaWei Li[1], YiZe An[1], JiaLe Wang[1], ZhuoXi Hu[1], JianMing Guo[3]
[1]*School of Urban Geology and Engineering, Hebei GEO University, Shijiazhuang 050000, Hebei, China.*
[2]*Hebei Province Underground Artificial Environment Smart Development and Management Technology Innovation Center, Hebei GEO University, Shijiazhuang 050000, Hebei, China.*
[3]*Hebei Jike Project Management Co., Ltd., Shijiazhuang 050000, Hebei, China.*
Corresponding Author: DongYang Geng, Email: gengdongyang@hgu.edu.cn

Abstract: Addressing issues such as content disconnect from industry standards, disjointed process training, and unrealistic evaluations in engineering cost practical training at applied undergraduate institutions, this study proposes a feasible, verifiable, and sustainable curriculum reform plan. Methodologically, it adopts a competency-based approach to reconstruct five core competency pathways: blueprint interpretation and modeling, quantity takeoff and rule verification, bill of quantities and cloud-based pricing, exam paper design and bid clarification, and version control and archiving. It organizes continuous, end-to-end training through course clusters and project weeks. High-frequency practice and timely error correction are supported by integrating computer labs with online platforms, virtual simulations, and gamified feedback. A chain-of-evidence assessment system—primarily process-based with supplementary outcome evaluations—and random sampling mechanisms ensure teaching quality. Data dashboards drive monthly refinements and semester-long reviews, creating a closed-loop cycle of problem identification, countermeasure development, retrospective analysis, and iterative optimization. Expected outcomes manifest in three dimensions: - Student level: Enhanced initial job readiness and employment alignment, with verifiable and traceable training outcomes. - Faculty and curriculum level: Standardized maintenance of rule and pricing databases, enabling verifiable and scalable course quality. - Industry-academia collaboration level: Mutual validation of course outcomes against professional certifications and job standards, establishing stable external evaluation channels. This study provides a replicable implementation pathway and quantitative management tools for transforming construction cost training from fragmented skill drills to process-oriented outcome delivery.
Keywords: Construction cost training; Curriculum cluster restructuring; Project-based teaching; Evidence chain evaluation; Data-driven improvement

1 INTRODUCTION

With the development of engineering management programs in Chinese universities, an increasing number of institutions have adopted the training of applied professionals as their developmental focus, undertaking the increasingly vital task of cultivating versatile, practice-oriented talents. Traditionally, practical training in engineering cost estimation involved delivering textbook theory followed by hands-on exercises where students performed practical drills. However, traditional classroom-based training methods fail to meet the demands of today's engineering construction industry in terms of technical skills, information technology proficiency, and comprehensive professional competence [1]. This gap is particularly pronounced in highly practical disciplines such as electrical engineering and engineering cost management, where significant discrepancies persist between specialized course content and industry standards. Common issues in both classroom instruction and practical training include a disconnect between theory and practice, overly simplistic teaching approaches, and overly formalized assessment methods. These shortcomings result in students lacking sufficient job adaptability and innovative capabilities. In response to this development, applied universities are vigorously implementing teaching reforms centered on project-driven, job-oriented, and progressive competency development. By redesigning course content, upgrading practical training environments, and applying simulation technologies, classroom teaching shifts from demonstration-based imitation to task-driven approaches, and assessment focuses on process rather than outcomes. This enables cost-related courses to progress systematically through software operation training, BIM modeling, and real-world case studies. This approach builds both knowledge frameworks and professional competency structures in a progressive manner. Furthermore, establishing information platforms transforms teaching from closed to open systems, evolving from isolated training labs to integrated digital learning spaces that combine online and offline elements.

2 NECESSITY AND SIGNIFICANCE OF TEACHING REFORM IN CONSTRUCTION COST PRACTICAL TRAINING COURSES

Construction cost courses represent a highly practical discipline within engineering management curricula. Integrating content from multiple elective subjects—including architectural drawing interpretation, building architecture,

construction techniques, structural engineering, construction organization, and construction cost management—these courses require students to master foundational knowledge across these fields while developing corresponding computational skills. The practical training course in engineering cost serves as an assessment of students' ability to measure and calculate project costs.

As the construction industry undergoes structural transformation from factor-driven to digital, lean, and full-lifecycle management-driven approaches, new methodologies such as full-process engineering consulting, BIM forward design, cloud-based pricing, and cost big data are rapidly permeating entry-level positions. However, some applied undergraduate institutions persist with outdated training models—manual quantity surveying, demonstration-based instruction, and end-of-term one-time assessments— Course content often fails to align with current pricing standards, regional price data, and authentic project cases. Teaching activities remain disconnected from job requirements, creating a mismatch between what students learn and what employers need. This disconnect results in low entry-level competency and puts institutions at a disadvantage in professional accreditation and employer evaluations. Therefore, comprehensive cost estimation training courses are essential for thoroughly assessing and enhancing students' abilities in: - Reading drawings - Quantity surveying - Cost estimation

1.1 Educational Management and Quality Assurance

Current talent development demands that course objectives, teaching processes, and learning outcomes be auditable, verifiable, and traceable. Traditional outcome-based final assessments struggle to demonstrate students' competency development trajectories in critical areas such as modeling, quantity surveying, cost estimation, verification, and deliverable production. They also fail to support the process-based evaluation, continuous improvement, and closed-loop management emphasized by professional accreditation. Consequently, cost engineering training courses must reconstruct their evaluation systems. This involves incorporating learning process data, version records, standard verification, and review reports into grade composition, establishing a quality management framework centered on evidence chains. This ensures methodological alignment between curriculum reform, professional accreditation, and teaching evaluations.

2.2 Industry and Job Requirements

The entry threshold for cost positions has shifted from basic calculation skills to accurate and rapid computation, verification capability, and collaboration proficiency [2]. This demands a classroom transition from static knowledge to dynamic competency, and from fragmented operations to end-to-end processes. Practical training must be organized around medium-to-large-scale, fully-featured engineering projects, integrating: - Drawing interpretation and BIM modeling - Quantity takeoff and standard verification - Bill of quantities compilation and cloud-based pricing - Deliverable documentation and bid clarification This ensures students complete high-frequency, repeatable assignments and reflections within near-real business workflows. Through this process-oriented training, students' tool usage is consistently embedded within problem-solving contexts, preventing mere tool operation proficiency. Simultaneously, it cultivates cross-disciplinary collaboration and engineering ethics awareness, shortening the adaptation period from campus to project site [3].

2.3 Institutional Environment and External Constraints

The integration of academic credentials with vocational skill certification drives the coordinated restructuring of curricula, job roles, and assessment frameworks [4]. Deconstructing certification standards into course objectives and teaching tasks, embedding assessment criteria into routine instruction and intensive practice, and exploring approaches like certification-based assessment and integrated course-role-competition-certification systems enhance learning purposefulness and external recognition without increasing overall student workload. More importantly, the process-oriented training and standardized assessment of certificates provide schools with a competency description framework shared with industry. This transforms students' course grades, competition results, and certificate levels into more comparable signals in the job market, enhancing the social recognition of talent cultivation.

2.4 Comprehensive Significance at the Institutional Level

The reform of engineering cost training courses will drive systematic upgrades across course clusters, resource repositories, and faculty teams [5]. Courses are no longer isolated units but modular chains structured around job competencies. Teaching units—such as drawing interpretation and modeling, quantity calculation and verification, bill of quantities pricing, price database management, and compliance review—interconnect to form a teaching production line where data and documentation flow seamlessly. Supporting resources—drawing packages, quantity calculation templates, price databases, anomaly case libraries, and micro-lecture scripts—accumulate through iterative refinement, becoming sustainably updatable public assets. Instructors transition from lecturers to project coaches and quality reviewers, collaborating with enterprises in task design and assessment standard development. This fosters dual-qualified faculty teams and school-enterprise partnerships, creating a virtuous cycle among teaching, technical services, and applied research.

2.5 Student Development and Social Services

This reform enables the robust development of observable, scalable, and progressive professional competencies within the school. Students produce verifiable tangible outcomes—such as bill of quantities, pricing documents, comparative calculations, and process logs—while gaining transferable skills like problem definition, evidence organization, standardized calculation management, and collaboration through iterative practice. These deliverables directly align with employers' job specifications and process standards through outcome-based evidence, enhancing employment compatibility and post-employment retraining capacity. Concurrently, institutions undertake real-world projects and leverage platform-based tools to deliver technical services for local governments and the social construction sector, amplifying professional social impact while elevating the quality of on-campus teaching and practical platforms.

Driven by the triple pressures of industrial upgrading, institutional reform, and quality governance, reforming engineering cost training is imperative. Its value lies in a three-dimensional reconstruction of the training system, quality system, and ecosystem: using job standards to drive curriculum design, leveraging information platforms to support processes, and employing evidence chains as the driving force for evaluation and improvement. By anchoring competency standards—including practical execution, accuracy, and verifiability—in unified data calculations, comprehensive process documentation, and traceable mechanisms, applied undergraduate engineering cost talent development can truly align with field requirements, serve market needs, and maintain sustained competitiveness within a rapidly evolving industry landscape.

3 SHORTCOMINGS OF TRADITIONAL CONSTRUCTION COST TRAINING COURSES

3.1 Disconnect Between Teaching Content and Industry Standards

For a long time, traditional construction cost training has primarily relied on simple drawings and outdated case studies. Project backgrounds have mostly focused on standard brick-concrete or standard frame structures, with little exposure to high-frequency engineering elements such as basements, pile foundations, prefabricated construction, complex roofing, and MEP installations. Furthermore, textbooks matching the latest measurement and pricing standards, regional pricing, and bill of quantities rules are severely outdated. Instructional materials predominantly use reference unit prices and obsolete fee standards instead of current market rates. Consequently, students' work habits remain rooted in textbook paradigms, lacking the preparatory skills needed to adapt to and integrate with the latest standards on actual project sites. This mismatch between content and standards renders outcomes difficult to equate with actual requirements, thereby compromising comparability, verifiability, and deliverability. The bill of quantities and pricing documents generated during training rarely pass final tender evaluation and bid clarification procedures. Additionally, influenced by traditional educational concepts, instructors habitually focus on explaining principles and demonstrating operations before practical training, fostering excessive student dependence. This teaching model struggles to cultivate students' initiative and creativity.

3.2 Disconnect Between Teaching Methods and Process Training

Methodologically, traditional training remains centered on teacher demonstrations combined with manual quantity calculations and fragmented software demos, lacking comprehensive, end-to-end project-based training. Students are often fragmented into isolated skills: they can draw components in software, perform quantity takeoffs, and apply pricing, but lack continuous experience spanning from drawing interpretation and modeling to standardized quantity calculation rules, anomaly data handling, bill of quantities compilation, cloud-based pricing, and final output verification and archiving. More critically, quantity calculation rules and standards are not documented or preserved as templates. Insufficient process logging and version control make recalculation difficult and obscure accountability. Teaching organization remains dominated by demonstration-imitation approaches, with insufficient project-based tasks, real-world scenario-based problems, and interdisciplinary collaboration training. Software is treated as a button-pushing tool rather than a process-oriented platform. Students lack strategies for modeling complex work scenarios, judging standard conflicts, and resolving anomalous data, often becoming passive upon entering the workforce—capable of point operations but hesitant to make judgments.

3.3 Distorted Evaluation Mechanisms and Practical Environments

At the evaluation level, traditional approaches overemphasize end-of-term tests or final presentations, while process documentation carries insufficient weight. This fails to demonstrate competency development trajectories or quality consistency. Common practices like substituting reports for hands-on work mask critical issues such as measurement errors, omissions, and standard inconsistencies. Review processes become superficial, lacking project-audit-style sampling and retrospective corrections. In practical environments, outdated computer lab hardware/software and software versions, insufficient account authorizations, and low utilization of online resources and simulation platforms persist. Limited integration of real-world data and standardized processes from industry-academia collaborations results in a significant gap between training scenarios and actual workplace conditions. Furthermore, the unclear correspondence between courses and professional certifications, coupled with the failure to decompose certification standards and integrate them into regular teaching, results in a disconnect between courses, job roles, competitions, and certifications [6]. This teaching model struggles to cultivate students' initiative and creativity.

4 SPECIFIC PATHWAYS FOR REFORMING CONSTRUCTION COST PRACTICAL TRAINING COURSES

4.1 Goal Alignment and Curriculum System Reconstruction

This section integrates competency objectives, course cluster restructuring, class hour allocation, and resource upgrades into a unified design to avoid fragmented implementation. First, establish a competency-based learning outcomes framework guided by job competency requirements. Five core competencies—drawing interpretation and modeling, quantity takeoff and rule verification, bill of quantities and cloud-based pricing, deliverable documentation and bid clarification, and version control and archiving—are translated into verifiable learning outcomes. Examples include independently completing steel reinforcement quantity takeoff with verifiable details for projects involving underground structures and shear walls, and generating complete bills of quantities and pricing documents within regional price database parameters. Next, map certification competency points to course objectives, creating a three-way alignment checklist of course goals, job tasks, and certification requirements. This serves as the unified basis for teaching, training, and assessment.

The curriculum structure replaces standalone training courses with integrated course clusters organized in a natural input-to-output sequence: - Drawing Recognition & Modeling as the entry point - Quantity Takeoff & Verification as the transitional phase - Bill of Quantities & Pricing as the integration stage - Exam Paper Creation & Bid Clearance as the delivery phase - Version Management & Archiving running throughout the entire process. Outputs from preceding modules directly serve as inputs for subsequent ones, eliminating skill silos. The course structure adheres to a 2:3 theoretical-to-practical ratio. Key modules combine consecutive classroom sessions with computer lab work, incorporating project weeks for end-to-end closed-loop training. Within project weeks, fixed points are established for component list freeze, rule/standard consistency checks, and pricing document review, ensuring students experience the complete business chain. Resources uniformly employ equivalent real-world project drawing packages covering complex structures and MEP systems. These are complemented by regional pricing databases and quantity calculation rule templates, establishing a foundational essential resource list and unified naming conventions for easy comparison and spot checks.

4.2 Teaching Implementation and Platform Support

This section integrates project-based implementation, platform-based support, cross-course coordination, and industry-academia competition-certification synergy into an executable pathway bridging classroom learning to workplace readiness. Instruction centers on a single semester-long project, progressing through task briefs, phase reviews, issue rectifications, and final archiving and delivery [7]. Within each group, three roles are assigned: primary responsibility, review, and archiving. The primary role handles modeling, quantity takeoff, and draft checklist creation. The reviewer conducts secondary checks against rule templates and documents issues. The archivist manages version logs and evidence organization. After each review, a problem rectification completion form must be submitted, specifying the responsible party, corrective measures, and verification methods to establish traceable quality records.

Computer labs handle core process training in modeling, quantity takeoff, and pricing [8]. Online platforms manage previews, assignment submissions, process logging, and quizzes, while virtual simulations aid in complex construction reading and error reenactment [9]. Gamified training with automated scoring provides real-time accuracy rates, omission lists, and conflicting criteria locations, enabling instructors to assign targeted remedial tasks. To prevent course silos, each course is linked to at least one cross-course task. This requires directly importing component lists generated from drawing recognition and modeling into the quantity takeoff and verification modules. Verified rule templates and version logs are then passed to the bill of quantities and pricing modules. The final output is archived in a unified format, creating a complete evidence chain from data to documentation. For external collaboration, invite enterprises to co-develop training lists and scoring criteria, incorporating real-world data cases and process standards. Break down industry competition problems into classroom exercises and stage quizzes, bridging classroom learning, practical training, competitions, and internships [10]. In mature courses, adopt certification-based assessment to achieve two-way validation between course outcomes and external certifications.

4.3 Evaluation System, Documentation Archiving, and Continuous Improvement

Evaluation, documentation, and improvement are integrated into a unified quality management framework. Performance is assessed through both process and outcome metrics, with process evaluation accounting for no less than half of the total score. Process metrics cover attendance, milestone achievements, stage reviews, peer verifications, defect remediation records, etc. Outcome metrics encompass bill of quantities, pricing documents, tender materials, and bid clarification reports. Each component features explicit scoring criteria and exemplars to minimize subjective grading. All critical stages require traceable, auditable documentation. Archival packages include rule templates, naming conventions, version logs, review records, issue resolution forms, data files, and report bundles. The school implements random sampling for consistency checks, error/omission rates, and compliance verification, incorporating findings into course quality reports.

Continuous improvement is driven by data dashboards. After each teaching cycle, metrics such as accuracy rates, omission rates, inconsistency counts, review discovery rates, certification pass rates, and job-skill alignment are

generated. Components and rules with high-incidence issues are updated with case studies and templates. Class schedules and training intensity are adjusted for groups with inconsistent learning outcomes. Revised resources and rules are implemented the following semester, creating a closed-loop system of identifying problems, implementing solutions, conducting reviews, and updating processes. To ensure operational effectiveness, establish mechanisms for software version and pricing database updates, implement monthly minor iterations and semester-long major reviews, unify account authorization and computer lab configurations, dynamically assess certificate recognition levels, and promptly adjust course-to-certificate mapping ratios. This prevents excessive focus on tool operation and evaluation deviation from objectives, ensuring competency development and evidence chain construction remain central.

5 CONCLUSION

This paper addresses challenges in engineering cost training at applied undergraduate institutions by establishing a goal-oriented pathway. This framework involves: Reconstructing course clusters aligned with objectives,Implementing full-process project-based execution with platform support, Prioritizing process-based assessment while preserving evidence chains,Driving continuous improvement through data dashboards.

Transforming disorganized skill training into deliverable, verifiable, and traceable process competencies. Through five-pronged advancement—drawing interpretation and modeling, quantity and rule verification, bill of quantities and cloud-based pricing, exam creation and bid clarification, version management and archiving—coupled with process control via project week reviews, critical milestone evaluations, and closed-loop issue tracking, students' entry-level job readiness and employability become tangibly measurable. Course quality achieves verifiable, replicable standards, faculty resources undergo standardized updates, and a stable channel of industry-standard mutual validation is established. Constraints such as required software/hardware investments, teacher role redefinition, and fluctuating certification recognition are mitigated through governance rhythms of monthly minor iterations and semester-long comprehensive reviews. Institutionalized measures—including unified account permissions, standardized computer lab configurations, and continuous integration of real-world data cases—enable organizational flexibility and sustained pedagogical gains. This ultimately transforms cost engineering training from outcome-focused demonstrations to process-documented, workflow-delivered outcomes.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The project was supported by Research on Theoretical Exploration and Practical Innovation in the Application-Oriented Transformation of Engineering Management Major, Project No.: 2025GJJG286.

REFERENCES

[1] Hu Shiting. Research on Teaching Reform of Practical Training Courses in Engineering Cost Management Based on BIM Technology. Journal of Hubei Open Vocational College, 2023, 36(06): 154-156. DOI: CNKI:SUN:HBHS.0.2023-06-057.

[2] Yin Yilin, Niu Yiqi, Ke Hong, et al. Research on the Pathway for Constructing Competency Standards for Engineering Cost Specialists: A Case Study of Tianjin University of Technology. Higher Architectural Education, 2023, 32(03): 63-74. DOI: CNKI:SUN:JANE.0.2023-03-008.

[3] Feng Jianxin. Practice and Exploration of Ideological and Political Education in Engineering Cost Specialty Courses at Higher Vocational Colleges: Taking the “Engineering Cost Control” Course as an Example. Heilongjiang Education (Theory and Practice), 2023(04): 38-40. DOI: CNKI:SUN:HJLL.0.2023-04-011.

[4] Yu Hailong, Wang Yinghua, Shao Wenshuai. Teaching Reform Practice of Engineering Cost Specialty Courses in Applied Undergraduate Institutions Under the 1+X Certificate-Credential Integration Model. Journal of Hubei Open Vocational College, 2024, 37(21): 187-189. DOI: CNKI:SUN:HBHS.0.2024-21-081.

[5] Xiao Qin. Teaching Reform of Engineering Cost Specialty Courses Based on Applied Talent Cultivation. New Curriculum Research, 2024(24): 49-51. DOI: CNKI:SUN:XKCY.0.2024-24-016.

[6] Li Jiaman. Exploring Teaching Reform Models for Engineering Cost Courses Based on “Competition-Driven Instruction”. Shaanxi Education (Higher Education), 2023(10): 79-81. DOI: 10.16773/j.cnki.1002-2058.2023.10.008.

[7] Yan Ling, Huo Shuangshuang, Deng Jiaojiao. Research on Competency-Oriented Workshop-Based Practical Teaching: A Case Study of the Engineering Cost Program at Tianjin University of Technology. Modern Educational Technology, 2014, 24(06): 113-121. DOI: CNKI:SUN:XJJS.0.2014-06-018.

[8] Cao Ruijuan. Teaching Practice of Engineering Cost Courses Based on GTJ Software. Integrated Circuit Application, 2023, 40(07): 342-343. DOI: 10.19339/j.issn.1674-2583.2023.07.152.

[9] Yu Fang, Jing Tao, Chen Xu. Application Research of BIM Technology in Engineering Cost Practical Training Courses. Modern Information Technology, 2018, 2(01): 190-192. DOI: CNKI:SUN:XDXK.0.2018-01-073.

[10] Huang Zhen. Teaching Reform of Engineering Cost Practical Training Courses Based on Discipline Competition Incentives. Light Industry Science and Technology, 2018, 34(04): 146-147.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3053

THE INFLUENCE OF HINGE ANGLE ON LIQUID SLOSHING IN ARTICULATED LIQUID TANKER

YuHang Mao
Huaiyin Institute of Technology, Huaian 223200, Jiangsu, China.
Corresponding Email: 2640809971@qq.com

Abstract: In the process of road driving, there are many reasons for the rollover of the tank truck, but due to the influence of liquid sloshing in the tank, the probability of rollover of the tank truck when turning will be greatly increased. In order to more clearly understand the influence of liquid sloshing in the tank and the factors affecting liquid sloshing, different hinge angles are used to analyze the liquid sloshing of the tank truck. Through the change of lateral torque and lateral acceleration of liquid sloshing in the tank, the influence of hinge angle on liquid sloshing and lateral acceleration is reflected. Then, the orthogonal experiment is used to verify the influence of hinge angle, filling rate and filling medium on liquid sloshing. Finally, it is confirmed that hinge angle and other factors have influence on liquid sloshing.
Keywords: Splice angle; Liquid sloshing; Orthogonal experiment

1 INTRODUCTION

Under the influence of the external environment, the sloshing of the liquid in the tank not only causes damage to the tank, but also has a certain impact on the driving of the liquid tanker. Therefore, it is of great significance to study the sloshing of the liquid in the tank to ensure the safety of the liquid tanker[1].

In order to determine the influence of bilateral curved baffles on the sloshing effect of spherical tanks, Al Yacouby Ahmad Mahamad et al. used ANSYS CFX software for numerical simulation, and used 0.55 and 0.8 liquid tanks to test the influence of baffles according to different filling rates of liquid tanks[2]. According to the factors such as the liquid sloshing force and the sloshing moment peak of the baffle, the conclusion is that the sloshing force, sloshing velocity and sloshing moment of the liquid sloshing in the liquid tank can be reduced by adding a suitable wave-proof baffle.

Zhang Xingying used the computational fluid dynamics method to numerically simulate the longitudinal liquid impact behavior of partially filled tanks[3]. By systematically investigating the key parameters such as filling ratio, tank geometry, longitudinal acceleration excitation characteristics and anti-sloshing device configuration, the influence of the above factors on the liquid impact load on the tank wall was quantitatively analyzed. Combined with the analysis of fluid dynamic characteristics, the steady-state amplitude, characteristic frequency and system damping of the maximum impact pressure in the tank under different working conditions are obtained, which provides a theoretical basis for the structural safety design and stability control of the liquid tanker. According to the fluid dynamics, the functions of the anti-wave plate setting and tank shape of the equivalent mechanical model are derived.

Liu et al. aimed at the transportation safety problem caused by the liquid sloshing in the tank[4]. Based on the model of a single anti-wave plate, Fluent fluid simulation was carried out on the liquid sloshing in the tank during the braking process. The two-way coupling method was used to study the distribution of gas and liquid in the tank during the braking process. According to the liquid dynamics and Euler multiphase flow model, the results show that the maximum shear stress of the anti-wave plate decreases nonlinearly with the increase of shear modulus.

Qi Ruiyun proposed a novel constrained attitude tracking control and active sloshing suppression scheme, which suppresses liquid sloshing by limiting the amplitude of angular velocity, control torque and its change rate[5]. The neural network method is used to obtain the sloshing force and sloshing torque caused by liquid sloshing, and a nonlinear disturbance observer is designed to deal with external disturbances.

Korkmaz Fatih Cuneyd studied the effect of stratified liquid on tank sloshing[6]. Five different fluids were used to explore the relationship between viscosity and free surface deformation and fluid impact on the side wall. A variety of fluid and high-speed imaging techniques were used to track the free surface deformation in single-layer and double-layer tests. The experimental results show that the lower fluid has experienced a significant damping effect.

Zheng et al. explored the role of the anti-wave plate in the non-full-load tank truck, and used the finite volume method to numerically simulate the tank truck under braking conditions[7]. The effects of the same filling rate, external excitation, pressure on the anti-wave plate, and the oscillation response of the gas-liquid interface were analyzed. The experiment shows that the anti-wave plate reduces the influence of liquid sloshing in the tank and reduces the gas-liquid interface fluctuation during the violent sloshing of the liquid.

In order to study the influence of the anti-wave plate on the liquid hydrogen tank, Jiang Rui et al. established a non-isothermal two-dimensional numerical model of the vehicle-mounted liquid hydrogen tank, and carried out thermodynamic simulation experiments[8]. The wave baffle reduced the sloshing amplitude of the liquid in the inner cylinder chamber. In addition, the wave baffle also played a ' thermal bridge ' role to a certain extent, forming a certain high temperature zone. Wang et al. studied the coupling effect of different media in the process of sloshing, analyzed the

pressure of the whole tank structure and all wave-proof plates, and the gas-liquid two-phase distribution characteristics of liquid sloshing in the tank[9]. The results show that the change of filling medium will not change the gas-liquid two-phase distribution characteristics of liquid sloshing in the tank, but will change the pressure on the surface of the wave-proof plate. Quansheng Zang proposed a new method to study the nonlinear liquid sloshing in the tank: isogeometric boundary element method (IGBEM), which can be used to study the discrete control of nonlinear liquid sloshing in the tank[10]. By creating a fixed coordinate system and a local rectangular coordinate system that can move synchronously in the center of mass of the tank and combining with the semi-Lagrangian function method, the non-uniform B-spline (NURBS) is used as the property function. The designed geometric boundary is used to describe the trajectory of liquid sloshing. The nonlinear discrete equation of liquid sloshing in the tank is successfully obtained by the designed method.

2 FLUENT SIMULATION

2.1 Parameter Design

In order to verify whether the hinge angle has an effect on liquid sloshing, the size of the hinge angle is fixed to ensure that the size of the hinge angle remains unchanged in the experiment and other factors will not change. After that, the fluid simulation experiment is carried out after being divided into 6 groups as shown in Table 1. The peak value of the lateral impact force and the peak value of the lateral acceleration of the liquid are used to preliminarily determine whether the hinge angle has an effect on the liquid sloshing in the tank.

Table 1 Hinge Angle Size

Hinge angle size(°)	5°	10°	15°	20°	25°	30°
Filled Gas	Preppie	Preppie	Preppie	Preppie	Preppie	Preppie
Filling Ratio	70%	70%	70%	70%	70%	70%

2.2 Experimental Design

The tank body of the liquid tank truck designed in this paper is 10 m long, the long axis is 2.3 m long, and the short axis is 1.87 m long. The established Solidworks model diagram is imported into ANASY FLUENT, and the model is simplified. After that, the whole vehicle is designed and meshed. The tetrahedral mesh is selected as the meshing form, and the mesh size is selected as : 80mm, to ensure that the overall mesh quality is greater than 0.2, so as to ensure that the mesh will not break during the simulation. Then, the step size design and speed setting are carried out to ensure that the tank truck can reach the same speed at the same time. The final basin volume is shown in Figure 1.

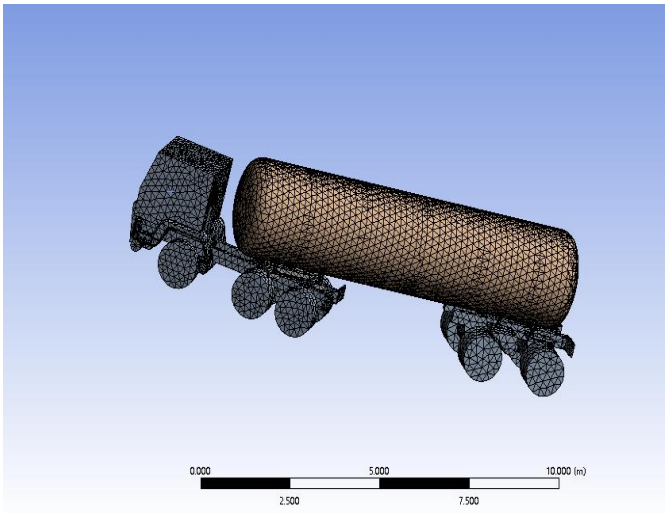


Figure 1 30 ° Hinged Angle Liquid Tanker Model Diagram

The components in the tank are mainly stupid and air, so it is treated as a two-phase flow problem. The turbulence model in the VOF model is selected for setting, and the implicit volume force is used for local balance. Then, in order to make the liquid sloshing more real, the dynamic mesh division is adopted, and the diffusion form is selected. After the setting is completed, the iterative step size and control accuracy are designed in WORKBENCH to ensure the accuracy of the experiment.

2.3 Result Analysis

Through six groups of experiments, the relationship between the lateral impact force and the lateral acceleration of each group with time is obtained, as shown in Figure 2 and Figure 3, because the six groups of experiments are from 0 to the same acceleration at the same time to achieve the same speed, the data is not much different, Figure 3 is the local amplification diagram of Figure 2. It can be seen from Figure 2 that with the increase of the hinge angle, the lateral impact force of the liquid increases, the increase speed of the lateral impact force also increases, and the peak value of the lateral impact force also increases. After reaching the first peak in 0.53 seconds, the increase rate of lateral impact force begins to increase, and then reaches the second peak, and the nonlinearity of liquid sloshing also increases. With the increase of the hinge angle, the arrival of the liquid impact force is gradually accelerated, as shown in Table 2.

Table 2 Peak Value and Time of Lateral Impact Force

Hinge angle size(°)	5°	10°	15°	20°	25°	30°
peak value of impact force(N)	6084.791	6263.135	6486.309	6669.181	6726.842	6822.35
Time(S)	0.35	0.37	0.38	0.4	0.41	0.43

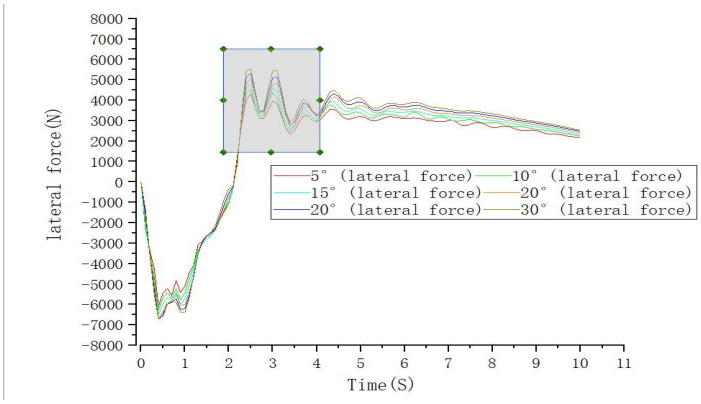


Figure 2 Liquid Sloshing Impact Force Diagram

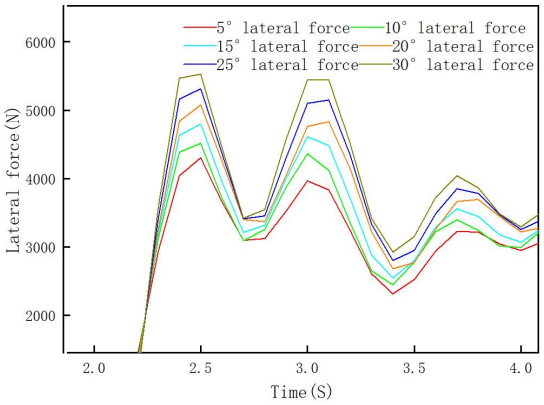


Figure 3 Local Enlargement

As shown in Figure 4, the relationship between the peak lateral acceleration and the hinge angle shows that as the hinge angle increases, the lateral acceleration also increases. For every 5 ° increase in the hinge angle, the peak lateral acceleration increases by approximately 0.2, which is approximately linear.

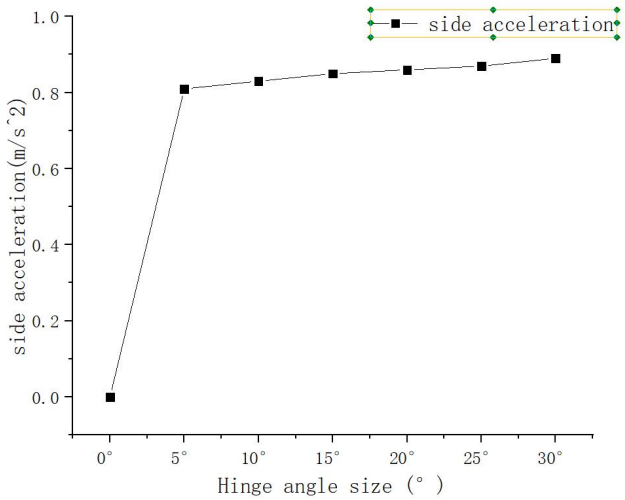


Figure 4 Relationship between Peak Lateral Acceleration and Hinge Angle

3 HINGE ANGLE ON THE LIQUID SLOSHING VERIFICATION

SPSSAU (Statistical Product and Service Software Automatically) is a powerful online data analysis platform. It provides efficient and convenient data analysis solutions through intelligent operation interface and rich analysis functions. The operation mechanism is concise, and the interaction mode of ' drag and drop ' is adopted. Only by uploading data, selecting variables and methods, the standardized results and charts can be generated at one click. The function is comprehensive and powerful. It provides 13 modules, covering about 500 algorithms, from basic descriptive statistics, T test, variance analysis, to advanced structural equation model, machine learning, Meta analysis, etc., to meet the needs of many fields.It not only inherits the statistical rigor of traditional professional software (such as SPSS), but also greatly improves the efficiency and accessibility of data analysis through automation, intelligence and cloud-based innovation.

3.1 Orthogonal Experiment

With the preliminary verification, the liquid sloshing will increase with the increase of the hinge angle, but the lateral acceleration will also affect the liquid sloshing in the tank, and the liquid lateral impact force may be affected by other factors. Therefore, further research is needed to see whether the hinge angle affects the liquid sloshing in the tank, which is indirectly affected by the lateral acceleration or directly affected by the hinge angle. In this paper, the orthogonal experiment method is used to verify the influence of liquid sloshing with other factors. Li proved that lateral acceleration, filling medium and filling rate have an effect on liquid sloshing[11]. Because it has been proved that the hinge angle has an effect on the lateral acceleration, and the orthogonal experiment needs no effect between the various factors, only the three factors of hinge angle, filling medium and filling rate are selected to verify the effect on liquid sloshing through experiments. The orthogonal table design was carried out by SQUSSA software, and the orthogonal test table of L (43) was selected. Four factors were selected for each factor, and 16 groups of experiments were formed by permutation and combination to verify the results.

3.1.1 Selection of hinge angle

According to the national regulations, the size of the hinge angle is generally not greater than 30 °, and the actual size of the hinge angle is generally greater than 5 ° during the turning process of the tanker. According to the results of the above 6 groups of experiments, the selected hinge angles are 10 °, 15 °, 20 ° and 25 °.

3.1.2 Selection of filling rate

According to the provisions of GB [12], the liquid tanker must retain 7 % -8 % of the air space and the maximum filling mass per unit volume allowed by the medium density at this temperature during transportation. And according to the understanding of the actual transportation situation, the filling rate should be selected 50 %, 70 %, 80 % and 90 %.

3.1.3 Selection of filling medium

In order to ensure the accuracy and universality of the experiment, the choice of filling medium should be related to viscosity and density. The selected filling medium is shown in Table 3.

Table 3 Filling Medium Selection Table

Filled Gas	Diesel Oil	Preppie	Liquefied LNG	Crude Oil
Viscosity kg/m^3	0.0025	0.000589	0.00016	0.5
Density $[kg/(m*s)]$	850	875	450	960.9

3.2 Range Experiment

Through the designed orthogonal table, the above FLUENT simulation experiment is carried out, and 16 groups of experiments are organized into the required orthogonal table, as shown in Table 4. The sorted orthogonal table was imported into SPSSAU for range experiment, and the results were obtained according to the following equation.
The i-th level for factor A:

$$\bar{K_{Ai}} = \frac{1}{m}\sum_{k=1}^{m} y_{ai}, k \tag{1}$$

Range calculation:

$$R_A = max\left(\bar{K_{A1}}, K_{A2}, \bar{K_{A3}}, \bar{K_{A4}}\right) - min\left(K_{A1}, K_{A2}, \bar{K_{A3}}, \bar{K_{A4}}\right) \tag{2}$$

Table 4 Orthogonal Experimental Table

Test Times	A	B	C	D	Index
1	50%	Diesel Oil	10°		2630.166
2	70%	Crude Oil	15°		4338.972
3	80%	Preppie	20°		6669.191
4	90%	Liquefied LNG	25°		2165.96
5	50%	Diesel Oil	15°		6304.223
6	70%	Crude Oil	20°		8121.474
7	80%	Preppie	25°		6263.135
8	90%	Liquefied LNG	10°		2246.575
9	50%	Diesel Oil	20°		8322.563
10	70%	Crude Oil	25°		8376.799
11	80%	Preppie	10°		5999.459
12	90%	Liquefied LNG	15°		3806.965
13	50%	Diesel Oil	25°		8805.877
14	70%	Crude Oil	10°		6916.421
15	80%	Preppie	15°		8525.87
16	90%	Liquefied LNG	20°		6617.991

As shown in Table 5 of the range analysis and Figure 5 of the horizontal factors, the best combination of the three factors is : the filling rate is 50 %, the hinge angle is 10 °, and the lateral force of this combination is the smallest. The filling rate has the greatest influence on the liquid sloshing, followed by the filling medium, and the hinge angle has the least influence.

Table 5 Range Analysis Table

item	level	A	B	C
K value	1	15804.29	27753.67	17792.62
	2	22935.41	26062.83	22976.03
	3	26505.79	14837.49	29731.22
	4	30866.16	27457.65	25611.77
K avg value	1	3951.07	6938.42	4448.16
	2	5733.85	6515.71	5744.01
	3	6626.45	3709.37	7432.81
	4	7716.54	6864.41	6402.94
best level		1	3	1
R		-3765.47	-3229.04	-2984.65
Number of levels		4	4	4
Repeats per level R		4.0	4.0	4.0

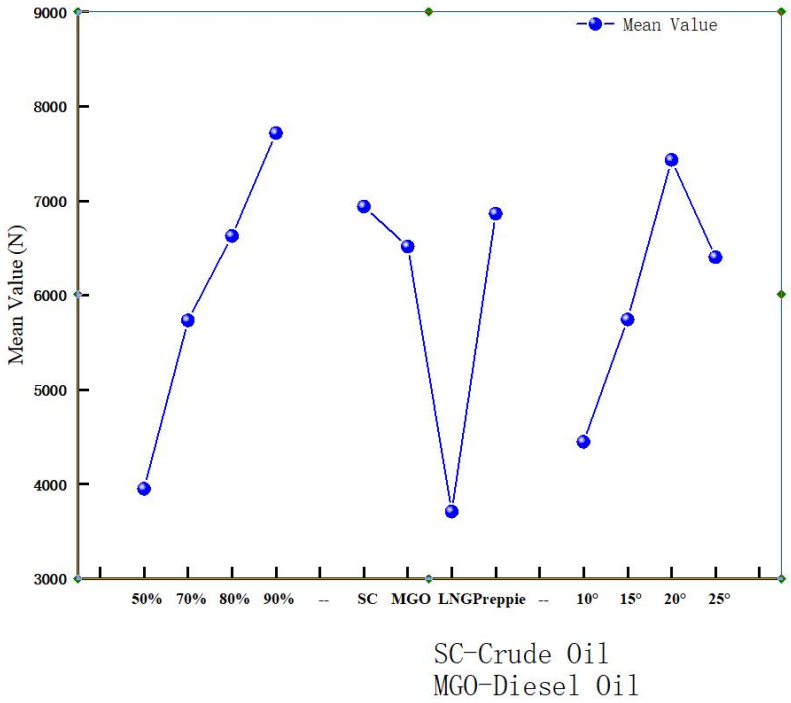


Figure 5 Analysis Chart of Each Factor Level

3.3 Variance Experiment

After the range experiment, it is determined that the three factors will affect the liquid sloshing in the tank. Then the orthogonal diagram is imported into the orthogonal experiment module of SPSSAU, and the experiment is calculated by the following formula. Strict significance test and quantitative analysis are carried out to obtain reliable conclusions.
Total sum of square:

$$SS\_T=\sum\sum y_{ij}^2-\frac{\sum\sum y_{ij}^2}{N} \tag{3}$$

Inter-group sum of squares:

$$SS\_A=\frac{\sum(T_i^2/n_i)-T^2}{N} \tag{4}$$

Intra group variance:

$$SS_E = SS_T - SS_A \tag{5}$$

Total degree of freedom:

$$df_T=N-1 \tag{6}$$

Degree of freedom between groups:

$$df_A = k - 1 \tag{7}$$

Intra-group freedom:

$$df_E = N - k \tag{8}$$

Mean square between groups:

$$\mathrm{MS\_A} = \frac{\mathrm{SS\_A}}{\mathrm{df\_}} \tag{9}$$

Mean square within group:

$$\mathrm{MS\_E} = \frac{\mathrm{SS\_E}}{\mathrm{df\_E}} \tag{10}$$

F Value:

$$\mathrm{F} = \frac{\mathrm{MS\_A}}{\mathrm{MS\_E}} \tag{11}$$

As shown in Table 6 of the analysis of variance, when the P value is less than 0.05, there is a 95 % probability that there is a relationship between the two. It can be seen that the three factors are significantly different from the liquid sloshing in the tank, indicating that the hinge angle has a significant effect on the liquid sloshing, providing strong evidence for subsequent research.

Table 6 Variance Analysis Table

	quadratic sum	df	mean square	F	p
intercept	577340531.052	1	577340531.052	2176.622	0.000
A	30430762.198	3	10143587.399	38.242	0.000263
B	28562281.404	3	9520760.468	35.894	0.000315
C	18755421.674	3	6251807.225	23.570	0.001016
residual error	1591476.743	6	265246.124		

4 CONCLUSION

According to the simulation results of FLUENT and SPSSUA, it can be seen that the size of the hinge angle has a significant effect on the liquid sloshing, not the indirect effect of lateral acceleration.

(1) The hinged angle has an effect on the lateral acceleration. As the hinged angle increases, the lateral acceleration also increases. However, it cannot be proved whether the hinged angle directly affects the liquid sloshing, or whether the hinged angle directly affects the lateral acceleration and indirectly affects the liquid sloshing.

(2)The lateral acceleration increases with the increase of the hinge angle, and the increase trend of the lateral acceleration is approximately linear. For every 5°increase in the hinge angle, the approximate increase is 0.2.

(3)The filling rate, filling medium and hinge angle have an impact on the liquid sloshing in the tank. Among them, the filling rate has the greatest impact, followed by the filling medium, and the hinge angle has the least impact. However, the $P < 0.5$ of the hinge angle can prove that the hinge angle has a direct impact on the liquid sloshing, and is not indirectly caused by the lateral acceleration.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCE

[1] Cao Y, Ma L, Gu Y, et al. Effect of braking impact of hazardous chemicals tanker on tank. Petrochemical equipment technology, 2025, 46(01): 7-14+75. DOI: CNKI: SUN: SHSJ.0.2025-01-003.

[2] Al-Yacouby A M, Ahmed M M, Liew M S. The Effects of Double-Side Curved Baffle Height on the Liquid Sloshing of a Spherical Tank—Numerical Study//Advances in Material Science and Engineering: Selected Articles from ICMMPE 2021. Singapore: Springer Nature Singapore, 2022, 55-69.

[3] Zhang Xingying. Research on the braking efficiency of tank car based on the analysis of the dynamic characteristics of longitudinal liquid impact. Jilin University, 2022. DOI: 10.27162/ d.cnki.gjlin.2022.003584.

[4] Liu X, Wang Xing, Xu Y. Two-way fluid-solid coupling numerical analysis of liquid sloshing in a moving tank. Journal of Xi 'an Jiaotong University, 2012, 46(05): 120-126. DOI: CNKI: 61-1069/ T.20120224.1036.007.

[5] Qi R, Dong X, Chao D, et al. Constrained attitude tracking control and active sloshing suppression for liquid-filled spacecraft. ISA transactions, 2023, 132, 292-308.

[6] Korkmaz F C. Damping of sloshing impact on bottom-layer fluid by adding a viscous top-layer fluid. Ocean Engineering, 2022, 254, 111357.
[7] Zheng Z, Zhang Z, Xu Y, et al. Study on the influence of wave-proof plate on the thermodynamics of non-fully loaded liquid hydrogen tank. Low temperature engineering, 2025(01): 94-100. DOI: CNKI: SUN: DWGC.0.2025-01-013.
[8] Jiang Y, Ding H, Hao C, et al. The effect of baffle on the dynamic and thermal characteristics of liquid hydrogen bottles under sloshing conditions. Low temperature and superconductivity, 2024, 52(10): 37-43+55. DOI: 10.16711/j.1001-7100.2024.10.006.
[9] Wang Jianye, Zhang Yin, Tang Zhenqi. Coupling analysis of impact sloshing of different liquid medium in tank of transport vehicle. Commercial vehicle, 2024(05): 78-81. DOI: 10.20042/ j.cnki.1009-4903.2024.05.004.
[10] Zang Q, Hong H, Liu J, et al. Isogeometric boundary element analysis of nonlinear liquid sloshing in containers under pitching oscillation. International Journal of Engineering Science, 2025, 217, 104371.
[11] Li Jie. Research on roll stability control of semi-trailer liquid tanker. Changchun University of Technology, 2018.
[12] Wang Lei. Research on anti-rollover stability of tank semi-trailer considering vehicle-liquid coupling. Harbin Institute of Technology, 2020. DOI: 10.27061/d.cnki.ghgdu.2020.00043.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3056

RELATIONSHIP BETWEEN TRAFFIC FLOW AND TIME BASED ON REGRESSION MODELS

TianYi Chen[1*], Cheng Cui[2]
[1]*School of Economics and Management, Nanjing Forestry University, Nanjing 210037, Jiangsu, China.*
[2]*School of Science, Nanjing Forestry University, Nanjing 210037, Jiangsu, China.*
Corresponding Author: TianYi Chen, Email: 13770185785@163.com

Abstract: Accurate estimation of traffic flow is crucial for urban traffic management and control, particularly when only main road monitoring data is available and feeder road data is missing. This study addresses the challenge of inferring feeder road traffic flow from main road data by developing a series of regression models tailored to different road structures and traffic conditions. For a Y-shaped basic road network, both linear and piecewise linear regression models were established, achieving perfect fitting of the main road traffic flow. In multi-branch scenarios that account for delays and cyclical fluctuations, an integrated model comprising constant, piecewise linear, and periodic functions was proposed, achieving a goodness of fit of 0.9722. Under traffic signal control conditions, a composite model including piecewise functions and periodic components was developed, effectively addressing traffic interruptions caused by signals, with a goodness of fit of 0.9642. In noisy data environments, a robust regression framework with adaptive weighting was introduced, maintaining high accuracy despite noise interference. The results indicate that the proposed models can effectively reconstruct feeder road traffic patterns, offering excellent interpretability and robustness. This provides a reliable data foundation for signal timing optimization and congestion management, offering practical solutions for traffic flow estimation in certain observed road networks.
Keywords: Linear regression model; Least squares method; Iterative optimization; Nonlinear regression model

1 INTRODUCTION

Since the initiation of the reform and opening-up, China's urbanization process has continuously advanced, with the congestion pressure on traffic systems becoming increasingly prominent. Real-time and accurate acquisition of road traffic flow information has become key to improving traffic operational efficiency and optimizing signal control. Currently, monitoring equipment has been widely deployed on main roads, but many feeder roads are limited by insufficient equipment coverage, lacking real-time traffic data, which hinders a comprehensive understanding of network operations. Therefore, inferring feeder road traffic flow from main road monitoring data has become a critical issue in intelligent transportation systems.

Traffic flow forecasting methods primarily include traditional statistical regression and artificial intelligence-based prediction approaches. In terms of traditional statistical methods, time series models, such as ARIMA, are widely used in traffic flow prediction due to their ability to effectively capture temporal dependencies[1]. Linear regression methods, due to their simplicity and interpretability, have always maintained an important position in traffic flow modeling. Ceder[2], through establishing statistical relationships between hourly traffic flow and accident rates, revealed fundamental traffic flow patterns. However, traditional methods are limited in their ability to capture complex nonlinear relationships and struggle to adapt to the dynamic changes in actual traffic flow.

With the development of machine learning techniques, significant progress has been made in traffic flow forecasting research. Support vector machines (SVM) address nonlinear issues through kernel functions and have demonstrated excellent performance in short-term traffic flow forecasting[3]. Ensemble learning methods, such as random forests, effectively handle high-dimensional data, enhancing prediction accuracy[4]. In recent years, deep learning models such as Recurrent Neural Networks (RNN) and Long Short-Term Memory Networks (LSTM) have shown strong capabilities in time-series data modeling[5]. Researchers compared the prediction effectiveness of autoregressive models and neural network models, finding that neural networks performed better in complex scenarios[6]. However, these data-driven methods typically require large amounts of high-quality training data and still face challenges in terms of robustness against noise and missing values[7-8].

In real-world road networks, traffic flow variations often exhibit complex features such as segmentation, periodicity, and nonlinearity[9]. Researchers revealed the multimodal characteristics of traffic flow by analyzing real-time traffic monitoring data and accident relationships[10]. They established a multi-lane traffic flow model that accounts for the effects of queuing. This paper, focusing on typical road structures, combines least squares fitting, piecewise regression, and periodic functions to construct a composite model, incorporating robust optimization strategies to achieve high-precision inversion of feeder road traffic flow. The findings of this study provide valuable data support for signal timing, congestion alleviation, and road network planning, offering both theoretical value and practical significance.

2 TRAFFIC FLOW PREDICTION FOR MAIN ROADS IN A Y-SHAPED ROAD NETWORK

2.1 Restatement of the Problem and Modeling Background

In road network monitoring, main roads are often equipped with traffic detection devices, while some feeder roads lack real-time monitoring capabilities due to cost or layout constraints. Considering a typical Y-shaped road structure (as shown in Figure 1), where the traffic from Feeder Road 1 and Feeder Road 2 converges onto Main Road 3, the device A1 on Main Road 3 records traffic flow every 2 minutes. It is known that the traffic on Feeder Road 1 follows a linear growth trend, while the traffic on Feeder Road 2 initially increases linearly before decreasing linearly. The challenge is to infer the traffic flow function relationships of the two feeder roads based on the monitoring data from Main Road 3.

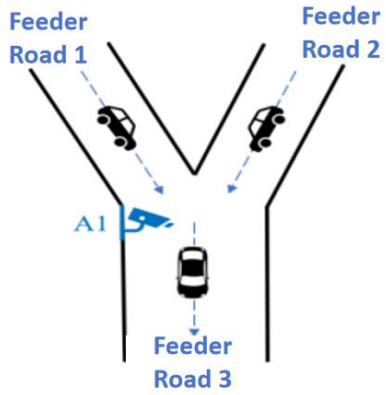


Figure 1 Y-shaped Basic Road Network Structure

2.2 Model Development

Let the time variable (t) be defined with 7:00 as the starting point (t = 0), and its range is [0,59] (corresponding to the time period from 6:58 to 8:58). The traffic flow on the main road is the sum of the traffic flows from Feeder Road 1 and Feeder Road 2:

$$Q_{main}(t)=Q_1(t)+Q_2(t) \tag{1}$$

Where the traffic flow on Feeder Road 1 is a linear function:

$$Q_1(t)=a_1t+b_1 \tag{2}$$

The traffic flow on Feeder Road 2 is a piecewise linear function:

$$Q_2(t)=\begin{cases} a_2t+b_2 & 0\le t\le t_c \\ -a_3t+b_3 & t_c<t\le 59 \end{cases}$$

The following non-negativity constraints must be satisfied: $Q_1(t)\ge 0, Q_2(t)\ge 0$， and $a_1>0, a_2>0, a_3>0$.

2.3 Parameter Estimation and Solution

The model parameters are estimated using the least squares method, with the objective function defined as:

$$\min_{a_1,b_1,a_2,b_2,a_3,b_3,t_c} \sum_{t=0}^{59} [Q_{main}(t)-(Q_1(t)+Q_2(t))]^2 \tag{3}$$

The parameter values are obtained through optimization as:

$$a_1=0.5, b_1=7, a_2=a_3=1, t_c=30 \tag{4}$$

2.4 Results and Analysis

The model fitting results indicate that the error metrics show perfect alignment between the model and the observed data, demonstrating extremely high fitting accuracy. From a practical traffic perspective, the traffic flow on Feeder Road 1 increases linearly at a rate of one vehicle every two minutes, with an initial flow of 7 vehicles. The traffic flow on Feeder Road 2 peaks at 30 vehicles at 8:00 (t=30), after which it decreases symmetrically. This result confirms the effectiveness of both the linear and piecewise linear models in inferring traffic flow for simple road network structures. Not only does it replicate the temporal variation of feeder road traffic, but it also provides reliable data support for traffic signal timing optimization and congestion management, highlighting the suitability of the least squares regression method for such problems.

3 TRAFFIC FLOW PREDICTION FOR MAIN ROADS WITH MULTIPLE FEEDER ROAD TYPES

3.1 Restatement of the Problem and Modeling Background

In the case of a more complex multi-feeder road convergence scenario, this study investigates the traffic flow composition of Main Road 5, which is formed by the convergence of four feeder roads, each with distinct traffic flow patterns. Specifically, Feeder Roads 1 and 2 exhibit a 2-minute transmission delay when converging onto Main Road 5,

while Feeder Roads 3 and 4 experience no delay. Based on the characteristics of the monitoring data, the traffic flow on Feeder Road 1 remains constant, Feeder Road 2 exhibits linear growth followed by a stable state in different periods, Feeder Road 3 shows linear growth followed by stabilization, and Feeder Road 4 displays noticeable periodic fluctuations. An integrated model that accounts for delay effects, piecewise changes, and periodic characteristics needs to be developed to accurately infer the traffic flow on each feeder road (Figure 2).

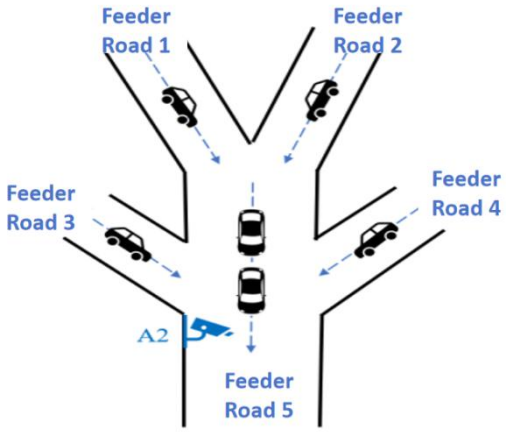


Figure 2 Multiple Feeder Road Types

3.2 Model Development and Solution

Based on the observed data and traffic flow characteristics, the following feeder road traffic flow function models are established:

- Feeder Road 1 is modeled as a constant function:

$$Q_1(t)=24.5 \tag{5}$$

- Feeder Road 2 is modeled as a piecewise linear function, accounting for its varying growth and stable behavior in different time periods:

$$Q_2(t)=\begin{cases} 1.35(t-1) & t-1\leq 23 \\ 30 & 23<t-1\leq 36 \\ 30+0.5(t-1-35) & t-1>36 \end{cases} \tag{6}$$

- Feeder Road 3 is modeled as a growth-to-stabilization function:

$$Q_3(t)=\begin{cases} 0.2t+11 & t<18 \\ 10 & t\geq 18 \end{cases} \tag{7}$$

- Feeder Road 4 is modeled as a periodic function, with its traffic flow segmented according to time phases:

$$Q_4(t)=\begin{cases} 0 & phase\in[0,5] \\ 10 & phase\in[6,13] \\ 0 & phase\in[14,17] \\ 10 & phase\in[18,25] \\ 0 & phase\in[26,27] \end{cases} \tag{8}$$

Where phase=mod(⌊t⌋,28). Each time unit corresponds to 2 minutes.

- The total traffic flow on Main Road 5 is the sum of the traffic flows from each feeder road, accounting for the delay effects:

$$Y(t)=Q_1(t-2)+Q_2(t-2)+Q_3(t)+Q_4(t)+\varepsilon(t) \tag{9}$$

3.3 Results and Validation

At key time points 7:30 (t = 15) and 8:30 (t = 45), the predicted traffic flow values for each feeder road are shown in Table 1:

Table 1 Predicted Traffic Flow Values for Each Feeder Road

Time Point	Feeder Road 1	Feeder Road 2	Feeder Road 3	Feeder Road 4
7:30	24.5	18.9	14.0	0.0
8:30	24.5	34.5	10.0	0.0

The overall goodness of fit of the model indicates that it effectively replicates the actual observed data $R^2=0.9722$. Residual analysis reveals that the errors primarily stem from the simplified assumption of Feeder Road 4's periodic

behavior and measurement noise, with residuals being particularly significant near phase transition points. Nevertheless, the model maintains high prediction accuracy during most periods, confirming the effectiveness of the multi-modal regression method in the decomposition of complex traffic flow (Figure 3). This provides a practical tool for the precise inference of heterogeneous traffic flow components in transportation systems.

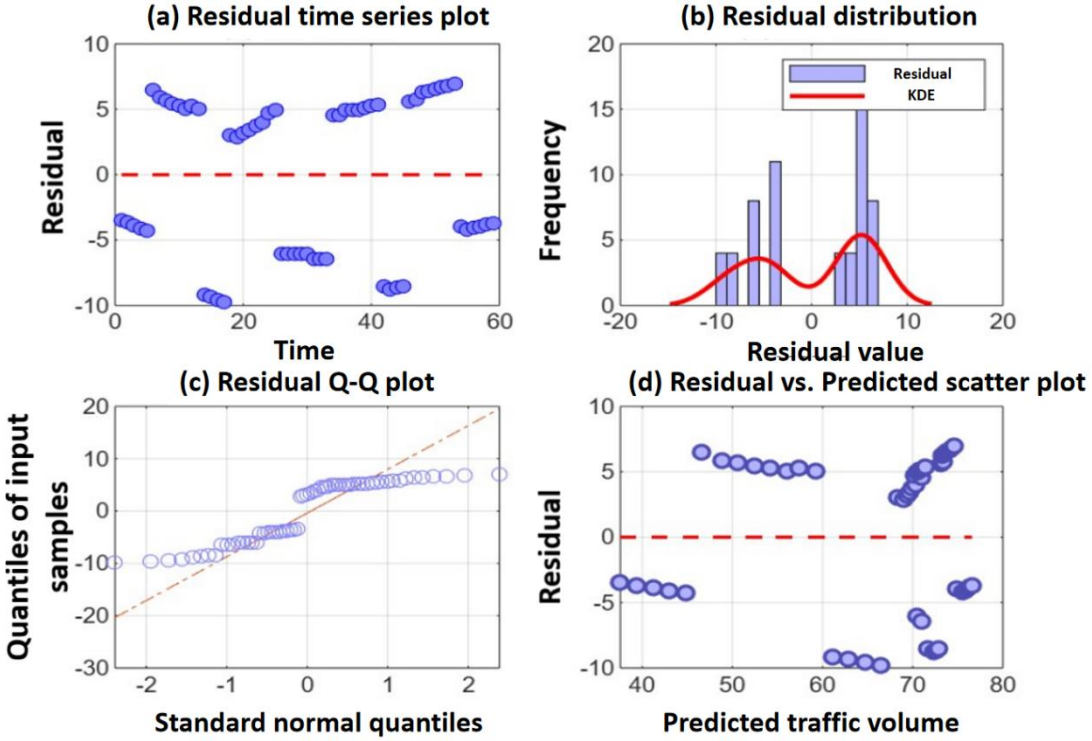


Figure 3 Residual Analysis for Problem 2

4 TRAFFIC FLOW PREDICTION FOR MAIN ROADS WITH MULTIPLE FEEDER ROAD TYPES UNDER TRAFFIC SIGNAL CONTROL

4.1 Restatement of the Problem and Modeling Background

This study focuses on a complex road system regulated by traffic signals, where Feeder Road 3 is controlled by a traffic light, resulting in periodic flow on-off characteristics. The system also includes Feeder Roads 1 and 2, with the traffic from all three converging onto Main Road 4. Device A3 records the traffic flow on Main Road 4 every 2 minutes (Figure 4). The task is to infer the traffic flow patterns of each feeder road based on this data. The challenge lies in incorporating the periodic constraints introduced by external control signals, while also considering the complex piecewise variation on Feeder Road 1, the multi-segment linear behavior on Feeder Road 2, and the periodic zeroing characteristic of Feeder Road 3. This represents a typical mixed dynamic system modeling problem.

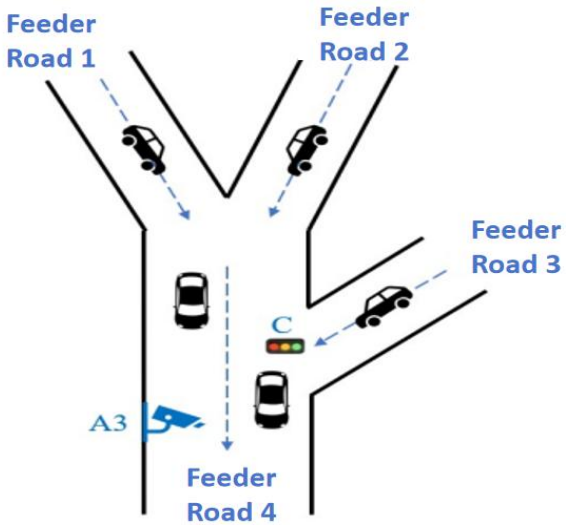


Figure 4 Multiple Feeder Road Types

4.2 Model Development

Based on the traffic flow characteristics of each branch, the following functional model is established:

- Branch 1 utilizes a five-segment composite model, encompassing stages of no flow, secondary growth, secondary decline, stability, and exponential decay.

$$Q_1(t)=\begin{cases}0 & -2\le t<10\\ 1.02(t-10)^2+1.98(t-10) & 10\le t<30\\ -0.97(t-30)^2-1.97(t-30)+99.8 & 30\le t<60\\ 79.88 & 60\le t<90\\ 79.88e^{-0.098(t-90)} & 90\le t\le 118\end{cases} \tag{10}$$

- Branch 2 employs a three-phase linear model, consisting of linear growth, stability, and linear decline.

$$Q_2(t)=\begin{cases}1.02(t+2) & -2\le t<70\\ 50 & 70\le t<94\\ 50-0.98(t-94) & 94\le t\le 118\end{cases} \tag{11}$$

- Branch 3 follows a cyclic function controlled by traffic signals, with linear growth during the green light phase and a constant flow during the red light phase.

$$Q_3(t)=\begin{cases}1.02+1.98(t-10) & t\in[6+18n,16+18n]\\ 0 & else\end{cases} \tag{12}$$

- The traffic flow on Main Road 4 is the cumulative flow of all branches, with a 2-minute delay considered for Branches 1 and 2.

$$Q_4(t)=Q_1(t-2)+Q_2(t-2)+Q_3(t) \tag{13}$$

The model parameters are determined through constrained least squares fitting and piecewise optimization techniques.

4.3 Results and Analysis

At the typical times of 7:30 (t=15) and 8:30 (t=45), the predicted traffic flow for each branch is as follows in Table 2:

Table 2 Predicted Traffic Flow for Each Branch

Time Point	Feeder Road 1	Feeder Road 2	Feeder Road 3
7:30	18.4	16.0	13.3
8:30	24.0	36.0	0.0

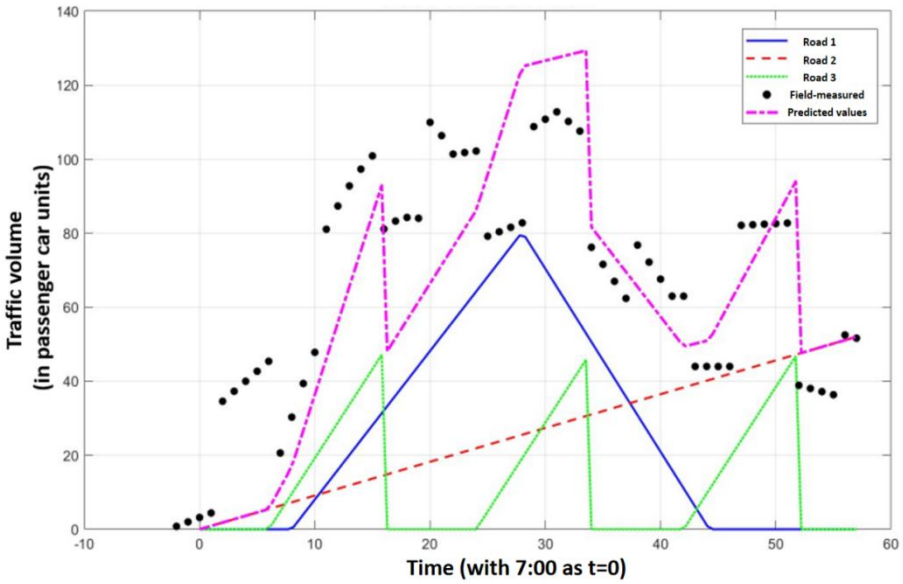


Figure 5 The Fitting Performance of Road Traffic Flow Prediction and the Contribution Rate of Each Branch

The overall goodness of fit for the model shows an average absolute error of 1.89 and a mean squared error of 6.32, indicating strong predictive accuracy (Figure 5). Notably, at 8:30, the traffic flow of Branch 3 is zero, which aligns with the actual situation during the red light phase. Residual analysis reveals that the errors are evenly distributed along the time axis, with no evident systematic bias, suggesting that the model effectively captures the cyclic characteristics under signal light control and the composite dynamics of each branch (Figure 6). This model successfully addresses the traffic flow inversion problem under periodic control conditions, providing more precise data support for signal timing optimization and traffic management.

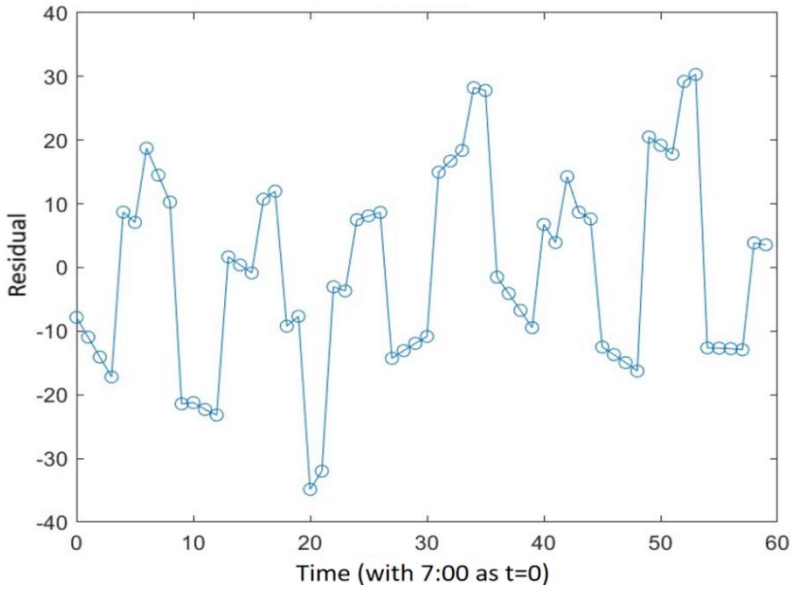


Figure 6 Residual Distribution of the Main Road Traffic Flow Prediction Model

5 CORRECTION OF THE MAIN ROAD TRAFFIC FLOW PREDICTION MODEL IN THE PRESENCE OF DATA ERRORS

5.1 Restatement of the Problem and Modeling Background

Having explored the prediction of main road traffic flow under ideal conditions for different road types, we now turn our attention to the inevitable data error issues present in real-world monitoring environments. The traffic flow monitoring device A3 on Main Road 4 records data during the period from 6:58 to 8:58 that contains significant errors, and it is crucial to reliably reverse-engineer the true flow of the three secondary roads under these conditions. The core challenge of this problem lies in how to accurately estimate the segmented variations of Branch 1, the three-stage characteristics of Branch 2, and the periodic on-off behavior of Branch 3, which is controlled by an unknown green light activation time. This represents a classic case of reverse modeling with error-laden data.

5.2 Model Formulation

In response to the characteristics of data errors, the following robust regression framework is established:

- Branch 1 utilizes a piecewise linear model with smooth transitions:

$$Q_1(t)=\begin{cases}0 & t<6\\ 1.2(t-6) & 6\le t<28\\ 25 & 28\le t<48\\ 25-2.272(t-48) & t\ge 48\end{cases} \tag{14}$$

- Branch 2 constructs a piecewise function that accounts for transition zones:

$$Q_2(t)=\begin{cases}0.39t+18 & t\le 18\\ 25 & 18<t\le 35\\ max(25-1.5625(t-36),0) & t>35\end{cases} \tag{15}$$

- Branch 3 establishes a periodic function controlled by traffic signal lights:

$$Q_3(t)=\begin{cases}1.02+1.98(t-10) & t\in[6+18n,16+18n]\\ 0 & else\end{cases} \tag{16}$$

- The traffic flow observation model for Main Road 4 is:

$$Q_4^{obs}(t)=Q_1(t-2)+Q_2(t-2)+Q_3(t)+\varepsilon(t) \tag{17}$$

where $\varepsilon(t)$ represents the observation error term.

By employing robust least squares and M-estimation techniques, along with an adaptive weight function to mitigate the impact of outliers, the method iteratively optimizes to simultaneously estimate the green light start time t_{green}=2(corresponding to 7:04) and the parameters of each branch model.

5.3 Results and Analysis

Under conditions of data errors, the model's prediction results at critical time points are as follows in Table 3:

Table 3 Predicted Traffic Flow Values for Each Branch

Time Point	Feeder Road 1	Feeder Road 2	Feeder Road 3

Time Point	Feeder Road 1	Feeder Road 2	Feeder Road 3
7:30	11.0	24.8	25.6
8:30	26.2	14.6	0.0

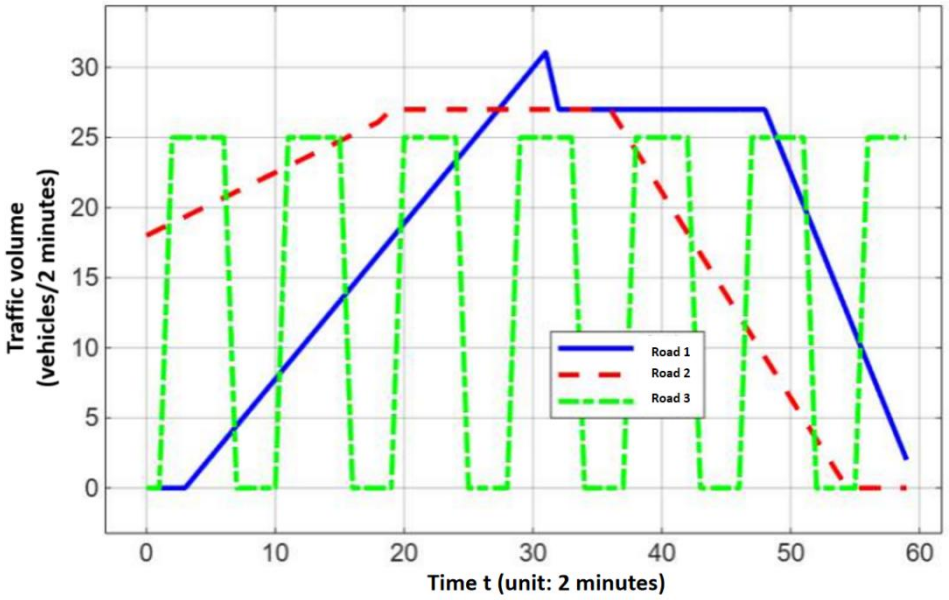


Figure 7 Traffic Flow Function Curves for Each Branch

The overall model fit quality shows a mean squared error of 15.29 and a mean absolute error of 3.06. Notably, despite the presence of noise in the data, the model maintains a high degree of accuracy, particularly in identifying the periodic behavior of Branch 3 and capturing the phase transition points of Branches 1 and 2 (Figure 7). Residual analysis indicates that errors are mainly concentrated during periods of rapid flow changes, but no systematic bias is observed, validating the effectiveness of the robust optimization method (Figure 8). These results demonstrate that the proposed robust regression framework can effectively resist the interference of monitoring errors, providing a reliable solution for flow inversion in noisy data within real-world traffic systems.

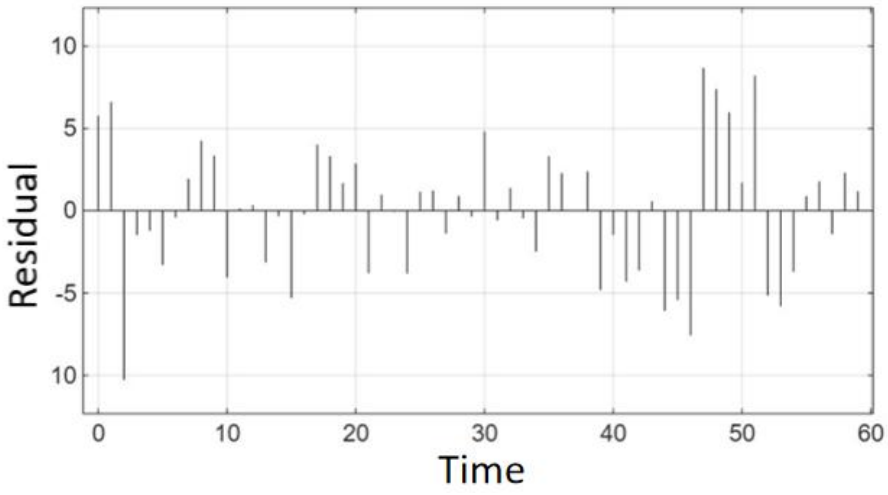


Figure 8 Model Residual Distribution

6 CONCLUSIONS

This paper systematically investigates the inverse methods based on regression modeling for urban road traffic flow estimation and prediction, with a particular focus on addressing the challenge of estimating traffic flow on secondary roads that lack monitoring equipment. By constructing linear regression, piecewise linear regression, nonlinear composite, and periodic function models, and combining least squares fitting with error optimization strategies, we successfully achieved high-accuracy estimation of traffic flow on secondary roads across multiple scenarios. Specifically, by developing a linear and piecewise linear overlay model for traffic flow on secondary roads at a

Y-intersection, perfect fitting of the main road traffic flow was achieved, demonstrating the superior performance of this method in simple road structures. In the prediction of main road traffic flow with multiple secondary roads, we further incorporated delay effects and periodic fluctuations, constructing a composite model for multiple secondary roads, which achieved a goodness-of-fit of 0.9722, showcasing the model's adaptability to complex situations. For multi-branch main road traffic flow prediction under signal light control, we established periodic control and robust regression models, which, even in noisy data environments, maintained high accuracy, proving the model's robustness and practical applicability.

The series of models established in this study not only exhibit high fitting precision and robustness but also offer transparency, strong parameter interpretability, and high computational efficiency, making them easy to integrate into real-world traffic control systems. For example, the model results can be directly applied in signal timing optimization, dynamic lane management, congestion warning, and other scenarios, providing a reliable data foundation and decision support for the construction of intelligent transportation systems. However, it should be noted that this study still has some limitations, such as the lack of consideration of external factors like weather and unforeseen events, and the model's adaptability in extreme scenarios needs further validation.

In future research, several directions warrant further exploration: first, incorporating additional external variables and contextual information to develop enhanced regression models that integrate multi-source data; second, combining deep learning with spatiotemporal graph neural networks to further improve the model's ability to capture the complex spatiotemporal relationships in large-scale road networks; third, integrating the proposed models with actual traffic control platforms for validation. Through continuous optimization and interdisciplinary collaboration, regression-based traffic flow inversion techniques are expected to play an increasingly significant role in intelligent transportation systems.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Sayed S A, Abdel-Hamid Y, Hefny H A. Artificial intelligence-based traffic flow prediction: a comprehensive review. Journal of Electrical Systems and Information Technology, 2023, 10(1): 13.

[2] Kashyap A A, Raviraj S, Devarakonda A, et al. Traffic flow prediction models - a review of deep learning techniques. Cogent Engineering, 2022, 9(1): 2010510.

[3] Razali N A M, Shamsaimon N, Ishak K K, et al. Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning. Journal of Big Data, 2021, 8(1): 152.

[4] Chen J, Zheng L, Hu Y, et al. Traffic flow matrix-based graph neural network with attention mechanism for traffic flow prediction. Information Fusion, 2024, 104: 102146.

[5] Ji J, Wang J, Huang C, et al. Spatio-temporal self-supervised learning for traffic flow prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(4): 4356-4364.

[6] Bao Y, Huang J, Shen Q, et al. Spatial-temporal complex graph convolution network for traffic flow prediction. Engineering Applications of Artificial Intelligence, 2023, 121: 106044.

[7] Wang H, Zhang R, Cheng X, et al. Hierarchical traffic flow prediction based on spatial-temporal graph convolutional network. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(9): 16137-16147.

[8] Shu W, Cai K, Xiong N N. A short-term traffic flow prediction model based on an improved gate recurrent unit neural network. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(9): 16654-16665.

[9] Navarro-Espinoza A, López-Bonilla O R, García-Guerrero E E, et al. Traffic flow prediction for smart traffic lights using machine learning algorithms. Technologies, 2022, 10(1): 5.

[10] Cao S, Wu L, Wu J, et al. A spatio-temporal sequence-to-sequence network for traffic flow prediction. Information Sciences, 2022, 610: 185-203.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3057

DYNAMIC RESPONSE AND COUPLING MECHANISM FOR COMPOSITE FAULTS IN BEARING-BEARING SEAT SYSTEM BASED ON LS-DYNA

XingChen Wang
School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710000, Shaanxi, China.
Corresponding Email: 2589691630@stu.xjtu.edu.cn

Abstract: This study establishes a finite element model of a 6205 rolling bearing system, incorporating the housing and central shaft, to analyze dynamic characteristics under healthy and composite fault conditions. The model's reliability is validated through close agreement between simulated and theoretical kinematic values and fault characteristic frequencies. Analysis reveals that stress concentrates at the leading edge of defects due to rolling element impact. The ball-inner ring composite fault induces the highest shear stress levels and severe fluctuations, resulting from intense interaction between moving and core load-bearing surface defects. In contrast, fixed outer ring defects cause periodic, moderate stress responses. Acceleration response analysis shows that ball-outer ring faults produce higher-amplitude, sustained vibrations due to near-rigid impacts, while ball-inner ring faults generate intermittent spikes buffered by the inner ring's inertia. These distinct signatures provide a theoretical basis for differentiating fault types in vibration signals. The findings offer accurate fault characteristics and simulation data support for vibration-based condition monitoring and early diagnosis of rolling bearings.
Keywords: Rolling bearing; Composite fault; Finite Element Analysis; Dynamic response; LS-DYNA

1 INTRODUCTION

As a core supporting component of rotating machinery, the operational status of rolling bearings directly determines the reliability and lifespan of equipment. Under complex working conditions, bearings are prone to failures such as fatigue and wear, which severely affect equipment safety [1]. Therefore, accurate modeling of their dynamic characteristics and fault analysis are of great significance.

Traditional research heavily relies on experiments and signal analysis, which are costly and struggle to capture internal contact processes. With the development of numerical simulation technology, the finite element method has become an important tool for studying bearing fault mechanisms. For instance, Zheng Tao conducted an explicit dynamic analysis of railway bearings [2]; Zhang Lele et al. revealed the influence of rolling element missing and pitting faults on stress distribution [3]; Ni Wenjun et al. identified fault characteristic frequencies through a pitting defect model [4]. Han et al. combined electromagnetic finite element analysis with measured capacitance to propose a bearing voltage prediction model [5].

However, existing studies often neglect the influence of external structures on vibration transmission. Azianou et al. pointed out that the deformation of the bearing seat alters the load distribution of deep groove ball bearings, thereby affecting fatigue life [6]. To address this, this paper establishes an integrated system model including the bearing seat and central shaft to more realistically represent actual working conditions [7]. Furthermore, research on the coupling mechanism of composite faults remains relatively scarce. Lin et al. proposed a machine tool model updating method based on dynamic evolution sequences [8], improving the accuracy of surrogate models and providing new ideas for system parameter identification. Building upon this, this paper constructs various composite fault models to systematically analyze their dynamic responses and stress characteristics. Parmar et al. employed optimized variational mode decomposition and deep learning to achieve high-precision classification of fault severity [9].

In the area of wear and life prediction, Rezaei et al. developed an adaptive finite element wear simulation method [10]; Mukras analyzed the influence of geometric parameters of journal bearings on wear [11]; Al-Tameemi et al. investigated white etching cracks induced by non-metallic inclusions [12]; Zhang et al utilized a Monte Carlo method with uniformly distributed sequences to optimize the load capacity calculation of permanent magnet bearings [13].

In summary, although current research has made significant progress in bearing fault modeling and simulation, aspects such as holistic system modeling, the coupling mechanism of composite faults, and the consideration of actual assembly and boundary conditions still require further deepening. Based on the aforementioned research findings, this paper takes the 6205 rolling bearing as the research object, establishes a system finite element model including the bearing seat, and systematically analyzes the stress distribution, vibration response, and fault characteristic frequencies under dynamic loads by configuring different types of composite faults. The aim is to provide more accurate simulation basis and methodological support for the study of bearing fault mechanisms, condition monitoring, and life prediction.

2 CONSTRUCTION OF THE FINITE ELEMENT SIMULATION MODEL FOR THE ROLLING BEARING

2.1 Geometric Parameters and Simulation Model Description

Most previous bearing simulation studies, for the sake of model simplification, have primarily modeled only the bearing itself. However, this approach inevitably fails to account for the influence of external structures surrounding the rolling bearing on vibration signal transmission. To more accurately and realistically replicate the operational conditions of the rolling bearing, this study adopts a model that incorporates a bearing seat housing the bearing and a central shaft passing through the inner ring. These three components are treated as an integrated system for finite element simulation analysis.

Considering that features such as chamfers and threaded holes on the bearing seat have a negligible impact on the simulation results while significantly increasing computational cost, this paper implements reasonable simplifications during the modeling process, as follows:

(1) Ignore various slots, holes, and threaded holes on the bearing seat.

(2) Omit all chamfers on the bearing seat, inner ring, and outer ring.

(3) Set the length of the central shaft to be consistent with the width of the inner ring and treat the shaft as a rigid body.

(4) Neglect bearing lubrication; therefore, except for the small gaps between the cage and the balls, no gaps are set between other components, and the influence of the oil film is not considered.

This study takes the 6205 bearing as the research object. The geometric parameters involved for each part in the model are shown in Table 1.

Table 1 Dimensions of 6205 Type Bearing-Bearing Seat System

Parameter	Value
Inner ring diameter D_i/mm	25
Outer ring diameter D_0/mm	52
Number of rolling elements N	9
Diameter of the ball bearing D/mm	7.94
Pitch diameter D_m/mm	38.5
Radial internal clearance $C_r/\mu m$	5
Rolling bearing width B/mm	15
Contact angle $\alpha/(°)$	0
The maximum width of the bearing housing W/mm	90
Bearing housing height H/mm	80

2.2 Mesh Generation and Defective Bearing Model Setup

Mesh generation is a critical preliminary step in finite element simulation analysis, as it directly influences the reliability of the simulation results. Generally, a smaller mesh size leads to higher model precision and more accurate simulation outcomes. However, reducing the mesh size exponentially increases computational demands. For instance, decreasing the mesh size of the inner ring in the rolling bearing model from 0.5 mm to 0.3 mm is estimated to increase the solution time by more than threefold. Therefore, selecting an appropriate mesh size that ensures model accuracy while making rational use of computational resources is particularly important.

For the moving components within the rolling bearing system, as well as components in relative motion and contact with each other—such as the inner ring, outer ring, cage, and rolling elements, which are treated as primary components—this study employs finer mesh elements. The rolling elements are the core moving parts of the rolling bearing, connecting the inner and outer rings. Consequently, the mesh element sizes for the rolling elements, inner ring, outer ring, and cage are controlled within 0.5 mm to ensure the established simulation model achieves the necessary accuracy for faithfully reflecting actual kinematic conditions. For fixed components like the bearing seat and central shaft, which have a secondary influence on the simulation results, a coarser mesh can be applied. Their element sizes are set at approximately 2 mm. In summary, the constructed bearing-bearing seat system simulation model comprises 542,665 elements and 109,561 nodes. Subsequent data verification confirms that this mesh meets the precision requirements for the simulation experiments. The meshed model is shown in Figure 1.

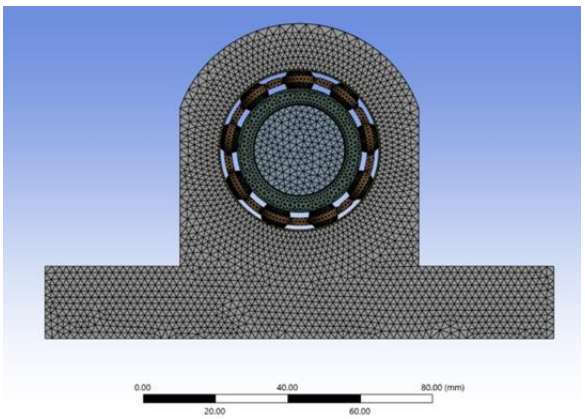
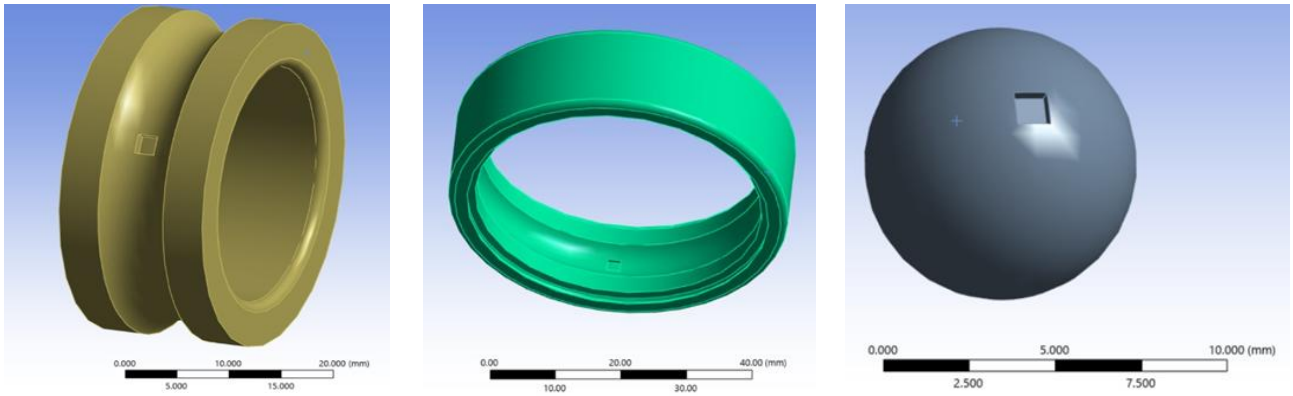


Figure 1 Meshed Model

In practical operating conditions, failures in rolling bearings can often be attributed to multiple factors such as improper installation procedures, harsh environmental conditions, and long-term overload operation. Common failure types include scratches and fatigue wear. Given the characteristics of bearings, the raceways of the inner and outer rings and the surfaces of the balls are the most prone to failure. Influenced by a combination of various factors, the specific morphology and dimensions of defect areas exhibit significant uncertainty, making it nearly impossible to precisely simulate all potential defect shapes. However, since the size of the defects is very small relative to the entire model, and the research focus of this paper is on the various manifestations of rolling bearings in the presence of defects, a representative defect model can be employed to characterize the fault features of rolling bearings. This approach is valid provided the model accurately reflects the core dynamic behaviors and typical manifestations under faulty conditions.
This paper establishes finite element models for the healthy state, ball-inner ring composite fault, ball-outer ring composite fault, and inner ring-outer ring composite fault. The fault morphology is uniformly defined as a rectangular pit of 2mm × 2mm, thereby constructing bearing fault mesh models with damage to different components. The relevant fault mesh models are presented in Figure 2.



(a) Inner ring fault (b) Outer ring fault (c) Rolling element fault

Figure 2 Fault Mesh Models

2.3 Material Parameters and Load/Boundary Condition Setting

To ensure that the established bearing-bearing seat system simulation model accurately reflects the actual operational state, it is essential to ensure that all parameters of the simulation model are fully consistent with the actual working conditions. Based on the genuine material properties of the actual 6205 bearing, this paper selects GCr15 bearing steel for the inner ring, outer ring, and rolling elements of the rolling bearing. The cage employs carbon steel, while the bearing seat is made of cast iron HT150. The corresponding materials and their parameters for each component are listed in Table 2.

Table 2 Material Parameters of Bearing-Bearing Seat System

Name of parts	Name of the material	Density $\rho(kg/mm^3)$	Elasticity modulus E/Gpa	Poisson's ratio μ
Retainer	Carbon steel	7.83×10^{-6}	196	0.24
Bearing seat	Cast iron HT150	7.15×10^{-6}	116	0.19
Inner ring	GCr15	7.83×10^{-6}	206	0.3

Outer ring	GCr15	7.83×10^{-6}	206	0.3
Rolling elements	GCr15	7.83×10^{-6}	206	0.3

Due to the frictional forces acting on the rolling elements during motion, the static friction coefficient f_s and the dynamic friction coefficient f_D for the various contact interfaces are set as shown in Table 3. Furthermore, during the contact definition, a decay coefficient of 0.00001 is defined to govern the transition between the static and dynamic friction coefficients.

Table 3 Frictional Coefficient

	Inner surface of outer ring	Outer surface of inner ring	Retaining cage pockets
Rolling elements	$f_S=0.1$	$f_S=0.1$	$f_S=0.002$
Rolling elements	$f_D=0.1$	$f_D=0.05$	$f_D=0.001$

The contact between the bearing housing and the outer ring, as well as between the inner ring and the central shaft, is defined as a tied contact. The inner ring is designated as a rigid body. This configuration allows the application of rigid body forces and angular velocities directly to the inner ring, which more accurately represents actual operating conditions compared to applying forces directly to the inner ring itself. Furthermore, since LS-DYNA cannot output various data for rigid bodies, this approach offers the additional advantage of enabling the observation of data such as the deformation and acceleration of the inner ring, thereby facilitating a more comprehensive subsequent analysis of the model.

During the setup, the bottom surface of the bearing housing is fixed. The inner ring is constrained in its rotational degrees of freedom about the x and z axes, as well as its translational degree of freedom along the x-axis. Given that the forces on the cage are relatively small, it is only constrained in its rotational degrees of freedom about the x and z axes. Simultaneously, a downward radial load of 2000 N is applied to the inner ring, with a rotational speed set to 1750 r/min. The angular velocity is applied in a gradually increasing manner from 0 to 0.01 s, maintaining constant speed and load from 0.01 to 0.1 s. The total simulation duration is set to 0.1 s.

2.4 Model Reliability Verification

2.4.1 Kinematic characteristics validation

Assuming the rolling elements undergo pure rolling motion within the raceways of the inner and outer rings, the reliability of the model can be validated by analyzing the velocities of its moving components. Since the influence of defects on velocity is negligible, this section uses the healthy rolling bearing as an example to analyze the velocity of its various moving parts. Specific nodes on the inner ring and the cage of the simulation model were selected, and their velocity curves were extracted. The velocity curves of the respective components obtained from the simulation results are shown in Figure 3 and Figure 4.

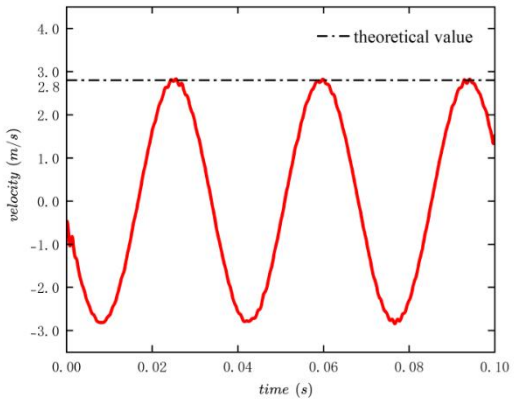


Figure 3 Inner Ring Speed Curve

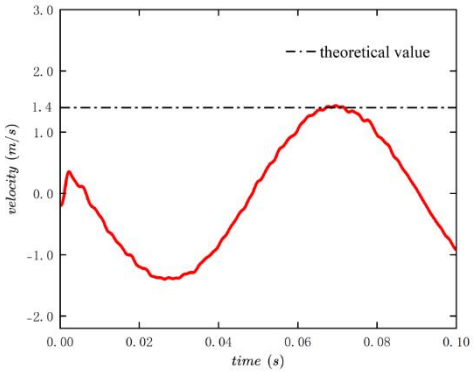


Figure 4 Retainer Speed Curve

Assuming the linear velocity at the center of the rolling element is equal to the tangential linear velocity of the cage. The theoretical calculation formulas for the tangential linear velocity of the cage v_m and the linear velocity at the inner surface of the inner ring v_i in a healthy ball bearing are given by:

$$v_m = \frac{\pi}{120}\left[n_i(D_m - D\cos\alpha) + n_0(D_m + D\cos\alpha)\right] \tag{1}$$

$$v_i = \frac{\pi n_i}{60}(D_m - D\cos\alpha) \tag{2}$$

Since the outer ring is bonded and fixed to the bearing housing, its rotational speed $n_0 = 0$. The inner ring, bonded to the central shaft, has a rotational speed of 1750 r/min. According to Eq. (1), the theoretical linear velocity of the cage is calculated as $v_m = 1.4m/s$. It can be observed from Figure 2-3 that this theoretical value of the cage linear velocity is very close to the peak value of the cage velocity curve obtained from the simulation. Similarly, the theoretical linear velocity at the inner surface of the inner ring, calculated using Eq. (2), is $v_i = 2.8m/s$. Figure 2-4 shows that this theoretical value closely matches the peak value of the inner ring velocity curve. These results indicate that the simulated velocities of both the balls and the inner ring in the ball bearing model meet the requirements of the theoretical calculations. Therefore, the established dynamic finite element model is reliable, the simulation successfully replicates the rotational speeds, and the model can be used for analyzing the dynamic characteristics of the bearing.

2.4.2 Fault characteristic frequency validation

Validation based solely on velocity is insufficient to fully demonstrate the model's reliability. As previously established, finite element simulation models have been created for the healthy state, as well as for composite faults involving the inner-outer ring, ball-inner ring, and ball-outer ring. To simulate the process of acquiring vibration signals via sensors under practical conditions, this study extracts the vertical acceleration vibration signals from the ensemble of nine balls for all four finite element models representing different health states. Envelope analysis is performed on the vibration acceleration signals collected from the simulation to obtain the fault characteristic frequencies of the 6205 bearing. The formulas for calculating the fault characteristic frequencies of the main components of the rolling bearing are as follows:

Outer ring fault characteristic frequency:

$$f_0 = \frac{n}{2\times 60}\left(1 - \frac{D}{D_m}\cos\alpha\right)N \tag{3}$$

Inner ring fault characteristic frequency:

$$f_i = \frac{n}{2\times 60}\left(1 + \frac{D}{D_m}\cos\alpha\right)N \tag{4}$$

Ball fault characteristic frequency:

$$f_b = \frac{n}{2\times 60}\frac{D_m}{D}\left(1 - \frac{D^2}{D_m^2}\cos^2\alpha\right)N \tag{5}$$

According to Eq. (3), Eq. (4), and Eq. (5), the calculated fault characteristic frequencies are as follows: the outer ring fault characteristic frequency $f_0 = 104.2Hz$, the inner ring fault characteristic frequency $f_i = 153.8Hz$, and the ball fault characteristic frequency $f_b = 67.7Hz$.

The frequency spectrum for the composite ball-inner ring fault is shown in Figure 5.

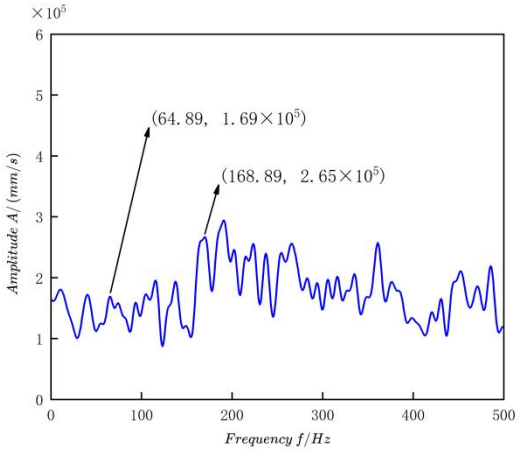


Figure 5 Frequency Spectrum for Ball-Inner Ring Fault

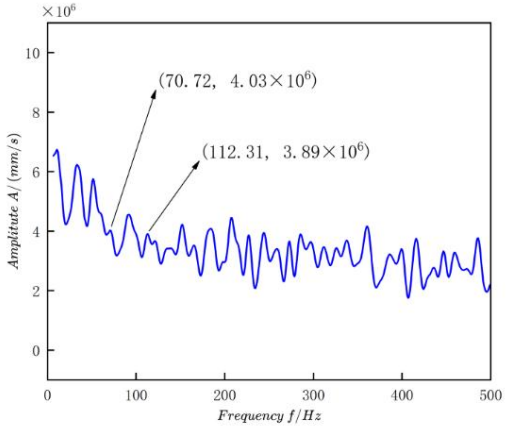


Figure 6 Frequency Spectrum for Ball-Outer Ring Fault

As shown in Figure 5, for the ball-inner ring composite fault, the simulated characteristic frequencies are 64.89 Hz and 168.89 Hz. These closely match the theoretically calculated frequencies of 67.7 Hz and 158.3 Hz, with corresponding errors of 4.15% and 6.69%, respectively.

The frequency spectrum for the composite ball-outer ring fault is shown in Figure 6.

As shown in Figure 6, for the ball-outer ring composite fault, the simulated characteristic frequencies are 70.72 Hz and 112.31 Hz. These closely match the theoretically calculated frequencies of 67.7 Hz and 104.2 Hz, with corresponding errors of 4.46% and 7.78%, respectively.

The frequency spectrum for the inner ring-outer ring composite fault is shown in Figure 7.

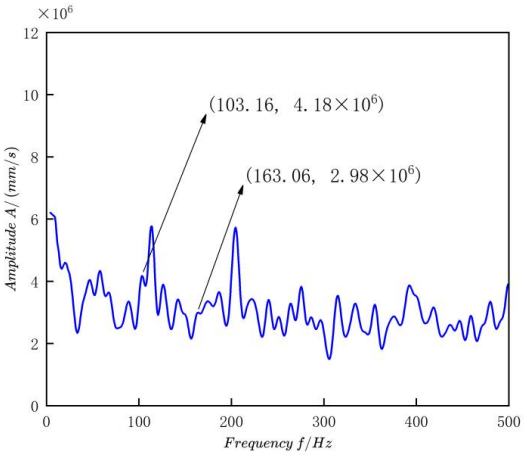


Figure 7 Frequency Spectrum for the Inner Ring-Outer Ring Composite Fault

As shown in Figure 7, for the inner ring-outer ring composite fault, the simulated characteristic frequencies are 163.06

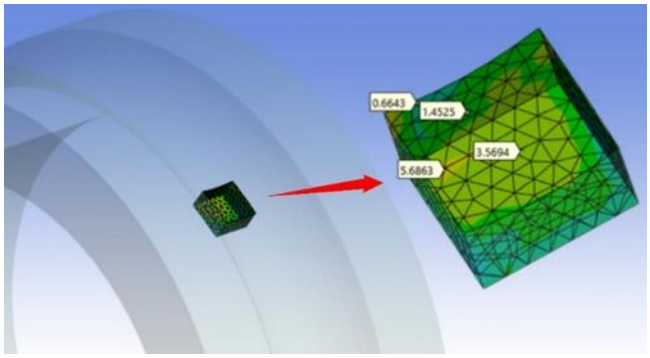
Hz and 103.16 Hz. These closely match the theoretically calculated frequencies of 158.3 Hz and 104.2 Hz, with corresponding errors of 3.00% and 1.00%, respectively.
These results demonstrate that the simulated fault characteristic frequencies from the model constructed in this study are all close to the theoretical values under different composite fault conditions, further verifying its reliability.

3 MOTION CHARACTERISTICS FOR BEARINGS WITH DIFFERENT FAULTS

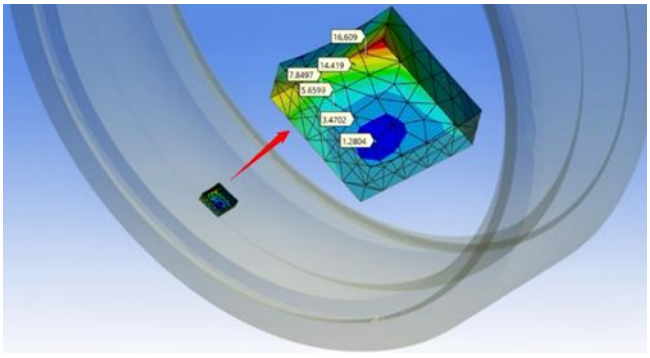
3.1 Shear Stress Variation Characteristics

3.1.1 Stress analysis at defect sites

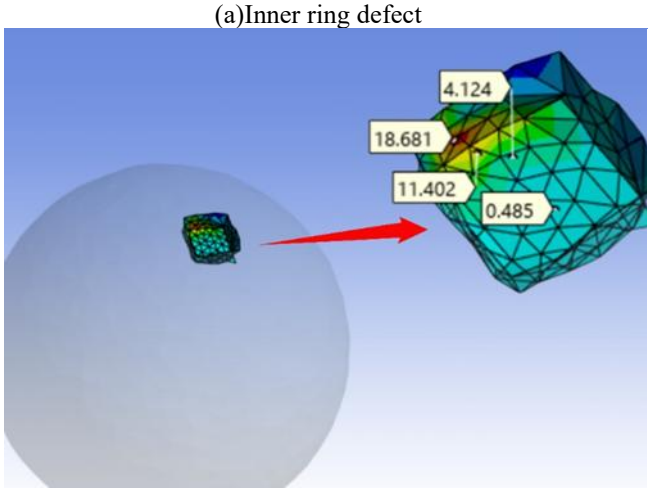
During the operation of the bearing, the rotational motion of the inner ring and the balls generates stress on the contact surfaces between the balls and the bearing raceways [14]. Therefore, analyzing the shear stress in the contact area provides insight into the mechanical state within the bearing during operation.
The shear stress contour plots at the three different defect locations are shown in Figure 8.
From the stress contour plots in Figure 8, it can be observed that the shear stress across the entire defect surface is not uniformly distributed along the five inner surfaces of the rectangular pit. Instead, a distinct stress concentration zone is evident on the bottom surface, biased towards one direction. This occurs because, during the motion of the balls, the leading edge of the defect consistently bears the load first. Consequently, the maximum shear stress at the leading edge of the pit defect is greater than that at the trailing edge. Regarding the overall stress distribution within the defect, the stressed area exhibits a gradual decreasing trend radiating outward from the edges of the defect. Furthermore, the absolute maximum stress for the entire defect is consistently located at the intersection line between the wall and the bottom surface at the defect's leading edge. This phenomenon arises because when a ball passes the leading edge of the defect at a relatively high velocity, it compresses the wall at the leading edge downwards, thereby inducing significant stress concentration at the base of the wall.



(a)Inner ring defect



(b)Outer ring defect



(c)Ball defect

Figure 8 Shear Stress Contour Plots at the Three Different Defect Locations

3.1.2 Impact of composite defects on the overall shear stress of the bearing

The preceding analysis has clearly revealed the characteristics and formation mechanisms of stress concentration at individual local defect sites. To investigate the more complex mechanical behaviors under the coupled effects of multiple faults, this study expands the analytical perspective from the local defect to the entire bearing system. The following section will systematically compare the distribution and dynamic fluctuation characteristics of shear stress across the entire bearing under three different composite defect models. The focus is on elucidating the differences in the interaction and superposition of stress fields induced by different types of composite defects, which consequently produce more significant global impacts compared to single defects. It also aims to deeply reveal the synergistic

mechanism of this multi-fault coupling on the overall mechanical performance and dynamic response patterns of the bearing, thereby providing a theoretical basis for accurately assessing the hazards of composite faults in practical operating conditions.

The shear stress curves on the bearing surface for different composite defects are shown in Figure 9.

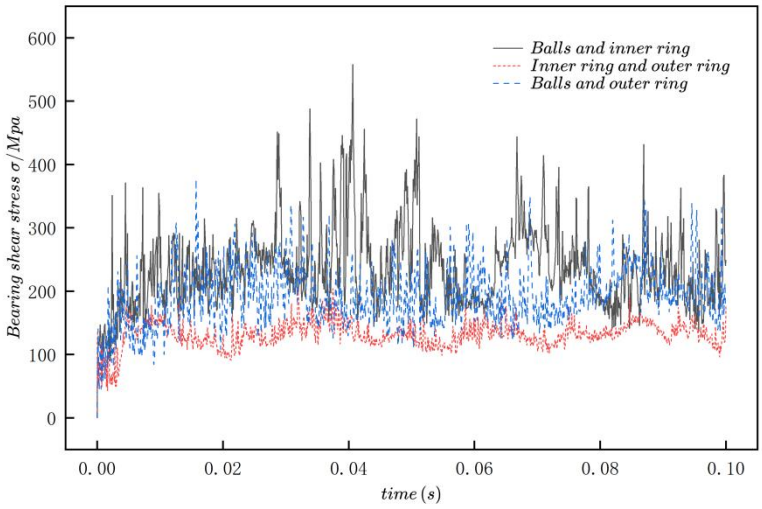


Figure 9 Shear Stress Curves on the Bearing Surface for Different Composite Defects

Regarding the overall distribution level of shear stress, the stress curve for the ball-inner ring composite defect is the highest overall, followed by the ball-outer ring defect, with the inner ring-outer ring composite defect being the lowest. This phenomenon is closely related to the load transmission path and the relative motion velocity. In the model presented in this paper, the rotating inner ring is the active driving component and the primary channel for load input. Consequently, the contact stress between the balls and the inner ring raceway is inherently the most severe. When defects are present simultaneously on both the balls and the high-speed rotating inner ring, a continuous and intense interaction is formed between the "kinematic defect" and the "core load-bearing surface" at high speed. Each contact is accompanied by significant impact, thereby substantially elevating the baseline stress level across the entire contact area. For the ball-outer ring defect, since the outer ring is fixed, its defect location is static relative to the load zone. Although it is periodically impacted by the moving ball defect, its dynamic effect is weaker than in the inner ring case. In contrast, the inner ring-outer ring composite defect, as it does not involve the ball as a source of impact, exhibits a stress response primarily stemming from the static influence of the two fixed raceway defects on the variation of the contact line, resulting in the lowest overall stress level.

In terms of stress fluctuation characteristics, defects involving balls exhibit significantly larger fluctuation amplitudes than the inner ring-outer ring defect, with the ball-inner ring condition showing the most pronounced fluctuations. This accurately reflects that kinematic defects are the primary source of dynamic excitation in the system. The revolution and rotation of the balls cause their defects to periodically impact the inner and outer ring raceways. In the ball-inner ring combination, the ball defect impacts the high-speed moving inner ring, and the extremely high relative velocity between them amplifies the severity of the impact. In the ball-outer ring combination, the ball impacts the stationary outer ring, where the relative velocity is lower; thus, the impact force and fluctuation amplitude are correspondingly reduced. The inner ring-outer ring defect, lacking such periodic sharp impact excitation, naturally exhibits the gentlest curve fluctuations.

From the perspective of curve phase, the phases of the ball-outer ring and ball-inner ring curves are similar, but both exhibit an almost anti-phase relationship with the curve of the inner ring-outer ring defect. This profoundly reveals the origin of the system's excitation mechanism. The excitation sources for the former two are essentially the "kinematic ball defect," and the impact force pulses they generate share a common phase origin. However, for the inner ring-outer ring defect, the system's vibration excitation mainly arises from the "meshing-unmeshing" effect generated when the defect on the rotating inner ring periodically passes through the load zone and interacts with the fixed outer ring defect. This excitation mode, dominated by the change in the relative position of the two fixed raceway defects, is fundamentally different in phase from the impact mode of the moving balls. Consequently, it excites different vibration patterns in the system, manifesting as a significant anti-phase characteristic.

3.1.3 Influence of different defects on stress in specific component

To thoroughly investigate the local mechanical effects of different composite fault modes on specific components, this section will conduct a comparative analysis focusing on three core load-bearing parts: the outer surface of the inner ring, the inner surface of the outer ring, and the balls. Specifically, two composite defect scenarios will be analyzed for each part. By extracting and comparing the stress data of these specific components from the three models (namely, the ball-inner ring, ball-outer ring, and inner ring-outer ring composite fault models), the aim is to identify which defect combination induces the most significant stress concentration and dynamic load on a particular component. This analysis seeks to reveal the key modes governing component damage in composite faults and their underlying

interaction mechanisms.

The shear stress curves of the balls in the ball-outer ring and ball-inner ring defect models are shown in Figure 10.

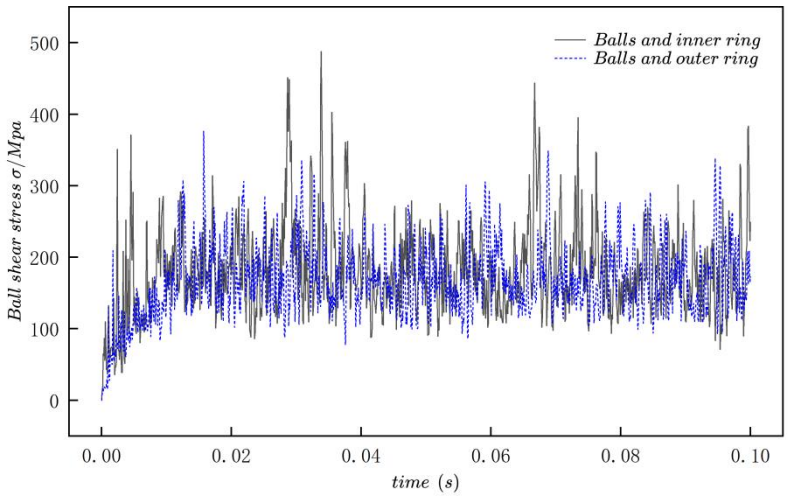


Figure 10 Shear Stress Curves of the Balls in the Ball-Outer Ring and Ball-Inner Ring Defect Models

As shown in Figure 10, during specific time intervals such as 0–0.01 s, 0.03–0.04 s, and 0.065–0.075 s, the fluctuation amplitude of the shear stress under the ball-inner ring defect condition is significantly greater than that under the ball-outer ring defect. This is primarily attributed to the higher relative velocity between the balls and the inner ring raceway under the constraint condition of "rotating inner ring and fixed outer ring." When the ball defect collides with the high-speed rotating inner ring defect, a more intense transient impact is generated. In other time intervals, the fluctuation amplitudes of the two curves are similar, indicating that when the balls are not in the main load-bearing zone or a specific contact phase, the influence of the defect type on the ball stress fluctuation is relatively diminished. During these periods, the stress response of the balls is governed more by the overall dynamic behavior of the system, causing the fluctuation levels under different defect models to converge.

Regarding the characteristics of the fluctuation distribution, the curve for the ball-inner ring condition exhibits distinct alternating features of "stress fluctuation concentration zones" and "relatively stable zones." In contrast, the fluctuations in the ball-outer ring curve lack such pronounced characteristics. This suggests that the influence of the inner ring defect on the balls possesses a stronger periodic impact nature, whereas the effect of the outer ring defect is relatively more continuous and uniform. These differences fully demonstrate that in composite fault bearings, the inner ring defect has a more significant dynamic impact on the balls and is the primary factor causing severe stress fluctuations in the balls.

The shear stress curves on the outer surface of the inner ring for the inner ring-outer ring and ball-inner ring defect models are shown in Figure 11.

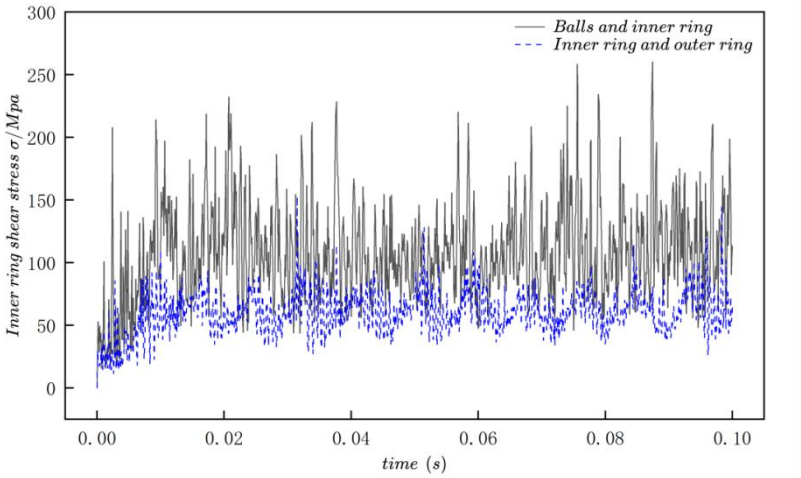


Figure 11 Shear Stress Curves on the Outer Surface of the Inner Ring

Through a comparative analysis of the shear stress on the outer surface of the inner ring in the two composite fault models (ball-inner ring and ball-outer ring), systematic differences in their stress responses due to different defect types can be clearly observed.

In terms of stress level, the shear stress curve under the ball-inner ring composite fault condition is significantly higher overall, with a maximum value reaching approximately 250 MPa, whereas the maximum value under the ball-outer ring condition is only 150 MPa. This phenomenon reveals the fundamental difference in the relative motion state between

the defect and the inner ring: the ball defect, acting as a high-speed moving impact source, generates direct and intense transient impacts on the rotating inner ring raceway, thereby substantially elevating the baseline stress level of the inner ring. In contrast, the fixed outer ring defect exerts a static and indirect influence on the system; the impact force it generates must be transmitted through the balls to act on the inner ring, resulting in some energy dissipation. Consequently, its overall effect on the stress level of the inner ring is relatively weaker.

The differences in fluctuation characteristics are also significant. Under the ball-inner ring fault condition, the fluctuation amplitude of the inner ring shear stress is markedly larger and appears random within the initial 0–0.1 s, without showing a distinct pattern. This reflects the complexity of the excitation source from the ball defect. Its impact is modulated by multiple factors including the revolution and rotation of the balls and the randomness of the defect position, leading to a non-periodic and severe fluctuating stress response in the inner ring. In stark contrast, the stress curve under the ball-outer ring fault condition exhibits stable periodic fluctuations, with clear peaks and troughs, and the interval between adjacent peaks is approximately 0.01 s. This is primarily because the fixed outer ring defect provides a periodic excitation source at a constant location, with its excitation frequency fixed at the ball pass frequency, thereby exciting a more regular and relatively gentler periodic stress response in the inner ring. In summary, the moving ball defect is the dominant factor causing severe and random stress concentration in the inner ring, whereas the static outer ring defect primarily induces a regular but moderate periodic stress fluctuation.

The shear stress curves on the inner surface of the outer ring for the inner ring-outer ring and ball-outer ring defect models are shown in Figure 12.

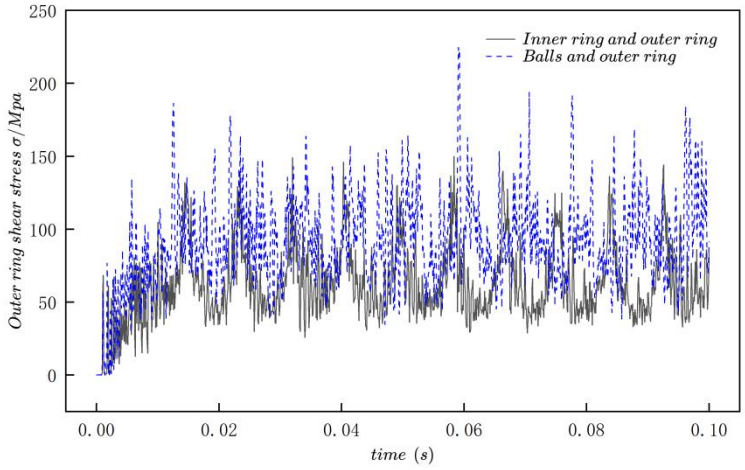


Figure 12 Shear Stress Curves on the Inner Surface of the Outer Ring

From the perspective of the overall stress distribution, the shear stress curve under the ball-outer ring composite fault condition is significantly higher than that under the inner ring-outer ring composite fault, with its maximum value reaching 200 MPa compared to only 150 MPa for the latter. The core reason for this phenomenon lies in the difference in the relative position of the defect to the fixed outer ring and the directness of the impact. The ball defect, acting as a moving impact source, directly and frontally strikes the fixed outer ring raceway during its revolution, generating extremely high transient stress at the moment of contact. In contrast, the influence of the inner ring defect on the fixed outer ring is indirect; its periodic excitation must alter the force state of the balls before being transmitted to the outer ring. This indirect transmission path dissipates part of the energy, resulting in a relatively lower perceived stress level on the outer ring.

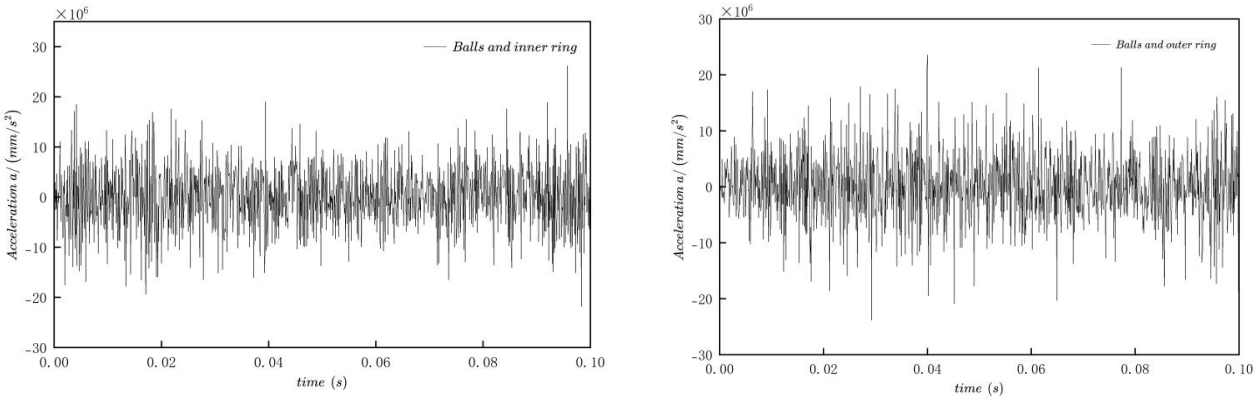
In terms of fluctuation characteristics, both defect models induce noticeable periodic fluctuations. The curves exhibit stable peaks and troughs, with adjacent intervals of approximately 0.01 s, corresponding to the characteristic frequency of the balls passing the fixed point on the outer ring. This indicates that the overall system dynamics are dominated by periodic events. A key common feature is that the stress curves of both models display a "sharp peak and convex trough" shape, meaning the duration of the troughs is significantly longer than that of the sharp peaks. This vividly reflects the force-bearing process of the bearing: the peaks represent the instantaneous impact generated when a ball, particularly one with a defect, passes by, which is short in duration but high in intensity; whereas the troughs correspond to the decay and relatively stable distribution of stress over a longer period after the impact.

However, there are important differences in the details of their dynamic responses. Although the overall fluctuation amplitudes are not significantly different, the ball-outer ring defect curve exhibits suddenly increased peak values at specific times, such as 0.01 s, 0.06 s, and 0.1 s, compared to other moments. This accurately captures the randomness and high-intensity nature of the ball defect impacts. These anomalous peaks likely correspond to instances where a defective ball, at a specific phase (such as exactly at the center of the load zone), collides violently with the edge of the outer ring defect, generating an impact force far exceeding that of regular periodic excitations. In contrast, the curve for the inner ring-outer ring defect is relatively smoother. Its fluctuations primarily stem from changes in the relative position between the periodically passing inner ring defect through the load zone and the fixed outer ring defect, making it less prone to such abnormal stress spikes.

3.2 The Influence of Different Defects on Ball Acceleration Response

The preceding analysis of shear stress in the bearing rings and balls clearly revealed the static characteristics and stress distribution patterns under different composite fault modes. However, changes in the stress field represent only one aspect of fault manifestation. The dynamic response of the bearing during operation, particularly the acceleration signal of the balls—the core moving components—holds more sensitive features for revealing the transient characteristics of impact events and identifying fault types. To delve deeper into the differential effects of inner ring, outer ring, and ball defects on the dynamic behavior of the bearing system, this section shifts the analytical perspective from static stress to dynamic acceleration. It focuses on a comparative study of the amplitude characteristics and fluctuation patterns of the ball acceleration response in the inner ring-outer ring composite fault model and the ball-outer ring composite fault model. The aim is to further elucidate the excitation mechanisms and evolution laws of different defects from a dynamic perspective.

The acceleration curves of the balls under the ball-outer ring defect and the ball-inner ring defect conditions are shown in Figure 13.



(a) Ball-inner ring composite fault (b)Ball-outer ring composite fault

Figure 13 Acceleration Curves of the Balls

In terms of the overall fluctuation amplitude of acceleration, the magnitude under the ball-outer ring composite fault condition is significantly greater than that under the ball-inner ring composite fault. The core reason for this phenomenon lies in the differing kinematic states of the defective components and their boundary constraints. When a ball passes a fixed outer ring defect, it is equivalent to a moving mass impacting a rigidly fixed point. The impact force generated by such a collision cannot be buffered by displacement or deformation of the defect body; the impact energy is almost entirely and instantaneously counteracted onto the ball itself, thereby exciting an acceleration response with a very high amplitude. In contrast, under the ball-inner ring fault condition, the inner ring itself is a massive, high-speed rotating inertial body. When a ball collides with an inner ring defect, part of the impact force is converted into a change in the rotational momentum of the inner ring. This inertial system absorbs and dissipates a portion of the impact energy. Consequently, the peak impact force transmitted back to the ball is attenuated, resulting in a relatively lower overall fluctuation amplitude in its acceleration response.

The temporal characteristics of the fluctuations and the distribution of peaks also exhibit distinct patterns between the two conditions:

The acceleration curve for the ball-inner ring fault shows increased amplitudes at specific instances, such as 0.002 s, 0.018 s, and 0.04 s, while remaining relatively stable during other periods. This reflects the intermittent nature of its excitation. These peaks correspond to transient impacts that occur when the phases of the ball defect and the inner ring defect coincide at specific locations within the load zone during their motion. Because both defects are moving, this intense interaction occurs only at specific relative positions, thus manifesting as intermittent high peaks superimposed on a stable background vibration.

In contrast, the acceleration curve for the ball-outer ring fault maintains large-magnitude fluctuations throughout the entire time period, with extreme amplitude values occurring at times such as 0.028 s and 0.04 s. This reveals the continuous and periodic nature of its excitation. The fixed outer ring defect acts as a periodic forced excitation source for the revolving balls. Each ball collides with it once per revolution, leading to a sustained high-amplitude response. The exceptionally high peaks observed within this response correspond to instances where the geometric phase of the ball's own defect couples with its orbital phase, causing the deepest engagement or most severe collision with the outer ring defect. This generates impact pulses that exceed those of regular periodic interactions.

This analysis demonstrates that the outer ring defect, as a fixed geometric excitation source, imposes a continuous and very high-amplitude acceleration excitation on the balls. Its dynamic response exhibits characteristics of a broad frequency band and high energy. Conversely, the inner ring defect, acting as a moving inertial excitation source, imposes intermittent and buffered impacts on the balls. Its acceleration response is characterized by periodic spikes superimposed on a stable baseline. This distinction provides a theoretical basis for fault diagnosis: in vibration signals, a

sustained, high-amplitude, broad-frequency-band acceleration response is more likely to indicate an outer ring fault, whereas periodic impact spikes are more characteristic of an inner ring fault.

4 CONCLUSION

Taking the 6205 rolling bearing as the research object, this study addressed the shortcomings of existing research in holistic system modeling and the analysis of composite fault coupling mechanisms. It successfully established a finite element model of the bearing system incorporating the bearing housing and central shaft, and systematically and thoroughly investigated the dynamic characteristics of the bearing under healthy conditions and various composite fault modes. Through simulation analysis and theoretical verification, the main conclusions drawn are as follows:

1. The established finite element model of the bearing-bearing seat system demonstrates high consistency between simulated values and theoretical calculations for the linear velocities of both the cage and the inner ring in terms of kinematic characteristics. In terms of fault diagnosis, the errors between simulated and theoretical characteristic frequencies under multiple composite faults are all minimal. This indicates that the model accurately replicates the actual dynamic behavior of the bearing, providing a reliable simulation platform for subsequent fault characteristic analysis.
2. At local defect sites, shear stress is not uniformly distributed; instead, significant stress concentration occurs at the junction of the bottom surface and the wall at the leading edge of the pit. This is caused by the squeezing effect when the rolling element passes over the defect at high speed. Regarding the overall system stress level, the ball-inner ring composite fault induces the highest shear stress and the most severe fluctuations. This is because it involves the intense interaction at high speed between both the moving impact source (ball defect) and the core load-bearing surface (inner ring defect). In contrast, the inner ring-outer ring composite fault, which does not involve a ball defect, exhibits the lowest overall stress level and fluctuation amplitude. Stress analysis on different components reveals that moving defects are the dominant factor causing severe and random stress concentration. The direct impact of a ball defect on the inner ring significantly elevates the inner ring's stress baseline and induces random fluctuations. Conversely, a fixed outer ring defect primarily leads to regular and moderate periodic stress responses.
3. By analyzing the acceleration response of the balls, it was found that the ball-outer ring composite fault excites acceleration fluctuations with higher amplitude and better persistence. This is because the impact from the fixed outer ring defect on the moving ball more closely resembles a rigid collision, resulting in less energy dissipation. In contrast, the acceleration response of the ball-inner ring composite fault manifests as intermittent spikes superimposed on a stable baseline. This is due to the inertial mass effect of the rotating inner ring buffers and dissipates a portion of the impact energy. This difference provides a clear theoretical basis and characteristic indicators for distinguishing between inner ring and outer ring faults in vibration signals.

In summary, through systematic simulation analysis, this paper has clarified the evolution patterns of mechanical behavior and vibration characteristics of rolling bearings under the coupled action of composite faults, deepening the understanding of bearing failure mechanisms. The research results can provide more accurate fault characteristic references and simulation data support for vibration-based bearing condition monitoring and early fault diagnosis, holding significant theoretical importance and engineering application value for improving the reliability maintenance level of rotating machinery.

Future research work could further consider lubricating conditions, thermal effects, and more complex time-varying load conditions to make the model more closely aligned with engineering reality. Simultaneously, exploring intelligent diagnosis methods based on algorithms like deep learning that integrate simulation and experimental data is recommended.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Junbao Yang. Research on Finite Element Modeling and Dynamic Response Characteristics of Composite Faults of Rolling Elements and Outer Raceways in Ball Bearings. Gansu: Lanzhou University of Technology, 2022.
[2] Tao Zheng. Explicit Dynamic Simulation and Fault Feature Analysis of Railway Vehicle Bearings. Master's Thesis. Shijiazhuang: Shijiazhuang Railway University, 2014.
[3] Lele Zhang, Nanlin Tan, Li Fan. Explicit Dynamic Simulation and Analysis of Rolling Bearing Faults. Journal of Shanghai Jiao Tong University, 2007, 41(9): 1506-1509. DOI: 10.3321/j.issn:1006-2467.2007.09.025.
[4] Wenjun Ni, Chang Zhang, Ziyang Jiang, et al. Dynamic Characteristics Analysis of Ball Defect Bearings Based on LS-DYNA and Refinement Spectrum - Bearing. 2024.
[5] Han P, Heins G, Patterson D, et al. Modeling of bearing voltage in electric machines based on electromagnetic FEA and measured bearing capacitance. IEEE Transactions on Industry Applications, 2021, 57(5): 4765-4775.
[6] Ayao E Azianou, Karl Debray, Fabrice Bolaers, et al. Modeling of the Behavior of a Deep Groove Ball Bearing in Its Housing. Journal of Applied Mathematics and Physics, 2013, 1: 45-50. DOI: 10.4236/jamp.2013.14009.
[7] Dong Lian. Research on Dual-domain Adaptive Fault Diagnosis of Rolling Bearings Based on Dynamic Simulation Data. Beijing: Beijing University of Chemical Technology, 2024.

[8] Lin W, Zhong P, Wei X, et al. Machine tool FEM model correction assisted by dynamic evolution sequence. Scientific Reports, 2025, 15: 18789. DOI: 10.1038/s41598-025-03058-9.
[9] Vivek Parmar, Shirshendu Layek, Megha Bhushan, et al. Advanced deep learning approach for the fault severity classification of rolling-element bearings. Scientific Reports, 2025, 15: 34353. DOI: 10.1038/s41598-025-16895-5
[10] Rezaei A, Van Paepegem W, De Baets P, et al. Adaptive finite element simulation of wear evolution in radial sliding bearings. Wear, 2012, 296(1-2): 660-671.
[11] Mukras S M S. Influence of Geometric Parameters on Journal Bearing Wear: A Finite Element Analysis and Elastic Foundation Approach. Applied Sciences, 2025, 15(5): 2368.
[12] Al-Tameemi H A, Long H. Finite element simulation of subsurface initiated damage from non-metallic inclusions in wind turbine gearbox bearings. International Journal of Fatigue, 2020, 131: 105347.
[13] Zhang L, Wu H, Li P, et al. Design, analysis, and experiment of multiring permanent magnet bearings by means of equally distributed sequences based Monte Carlo method. Mathematical Problems in Engineering, 2019, 2019(1): 4265698.
[14] Wenjun Ni, Chang Zhang. Dynamic Characteristics Analysis of Defective Bearings Based on Simulation Dynamics Model. Journal of Mechanical & Electrical Engineering, 2024, 41(10).

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3058

SMOKE JAMMER DEPLOYMENT STRATEGIES FOR DRONES UNDER MULTIVARIATE OPTIMIZATION

JinSong Zhang*, JunRui Mu
School of Mathematical Sciences, Chengdu University of Technology, Chendu 610000, Sichuan, China.
Corresponding Author: JinSong Zhang, Email: zjs_0701@163.com

Abstract: This study investigates smoke-screen interference tactics deployed by unmanned aerial vehicles (UAVs) against incoming M1 missiles. An intuitive geometric-physical model is established for missile trajectories, UAV motion, and smoke-screen projectile descent/detonation. The effective shielding effect of smoke clouds is quantified by combining geometric cone inclusion criteria with linear distance assessment methods. To address the calculation of effective shielding duration for a single smoke grenade, a variable stride search algorithm is employed, yielding an effective shielding duration of 1.391975 seconds under initial conditions. To formulate an optimal jamming strategy, four decision parameters are introduced to explore the optimal deployment strategy for a single drone with a single smoke grenade. A hybrid strategy combining genetic algorithms and particle swarm optimization is adopted, increasing the maximum effective shielding time to 4.585 seconds. Building upon this foundation, this paper analyzes the optimization scenario of coordinated shielding by a single drone deploying multiple flares. A multivariate optimization model is established to account for the composite shielding effect of multiple smoke clouds. Through the basin-jumping particle swarm optimization algorithm, the maximum composite effective shielding time reaches 6.3020 seconds. These research findings provide optimization strategies and actionable solutions for maximizing smoke flare interference effectiveness and coordinated multi-flares interference.
Keywords: Multivariate optimization model; Variable-step search algorithm; Particle swarm optimization algorithm

1 INTRODUCTION

In modern defense combat systems, smoke countermeasure munitions form shielding in specific airspace ahead of targets, serving as an effective means to disrupt incoming enemy missiles and protect critical assets. Achieving precise, on-target, and on-time deployment of these munitions is key to enhancing defense efficiency[1-2].

This study addresses the challenge of deploying smoke countermeasure grenades from unmanned aerial vehicles (UAVs) to counter incoming M1 missiles, aiming to develop an optimal deployment strategy that maximizes effective obscuration duration. Specifically, it first tackles the calculation of single-grenade effective obscuration duration[3]. Under known initial conditions and parameters, an intuitive physical model is constructed to assess obscuration effectiveness, thereby determining the duration of effective coverage against the M1 missile. Building upon this foundation, the paper addresses the optimal deployment strategy for a single drone and a single flare by incorporating four parameters: drone flight direction, flight speed, flare deployment point, and detonation point. The objective is to formulate a rational optimization strategy that maximizes the effective shielding duration of a single smoke cloud. Subsequently, the study delves into the optimization of coordinated shielding by a single drone deploying multiple flares. This requires formulating a rational deployment and detonation strategy to maximize the combined effective shielding duration produced by three smoke flares released from drone FY1[4-5]. To address these challenges, this paper first constructs intuitive geometric physical models for missile trajectories, drone motion, and the descent and detonation of smoke flares. For effective concealment assessment, a combined method utilizing geometric cone inclusion criteria and linear distance judgment was adopted to quantify concealment effectiveness. To address different optimization requirements, multiple algorithms were employed: variable-step search for initial computation, a hybrid genetic algorithm and particle swarm optimization strategy for single-decoys optimization, and basin hopping-particle swarm optimization for multi-decoys coordination optimization—effectively searching for global optimal solutions[6-7].

This research provides three key marginal contributions to the field of UAV-based countermeasures. First, it establishes a novel geometric-physical modeling framework that integrates missile trajectory, UAV kinematics, and smoke dispersion dynamics into a unified spatial-temporal system. Second, the study proposes a hierarchical optimization strategy employing tailored algorithms (variable-step search, GA-PSO, and basin-hopping PSO) for different operational scenarios, significantly improving solution quality and computational efficiency. Third, it introduces a composite assessment method combining geometric cone inclusion criteria with linear distance judgment, enabling more accurate quantification of effective obscuration duration compared to conventional single-criterion approaches. These contributions collectively advance the theoretical and practical understanding of optimal smoke screen deployment in dynamic combat environments.

2 CALCULATION OF EFFECTIVE SHIELDING DURATION FOR SINGLE MISSILE

2.1 Model Establishment

Construction of trajectory equations in the spatial coordinate system. According to the description, the flight direction and speed of the UAV, the release point and detonation point of the smoke interference bomb are all known. Among them, Missile M1 flies toward the decoy target, while UAV FY1 flies at the same altitude in the direction of the decoy target. After the smoke interference bomb detonates[8], the smoke cloud is idealized as a standard sphere, and its volume and shape remain unchanged within 20 seconds after detonation. A spatial coordinate system is established with the decoy target as the origin, and the initial detonation state is shown in the following Figure 1:

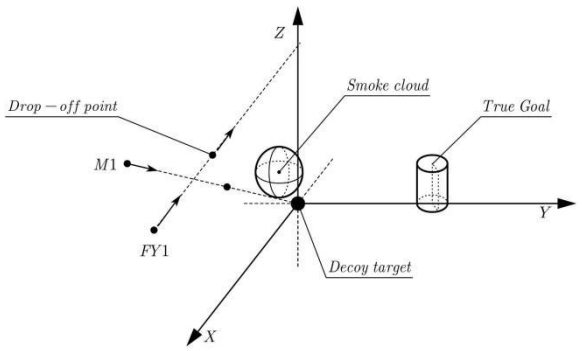


Figure 1 Schematic Diagram of the Initial Detonation State

For Missile M1, its initial coordinates are (20000, 0, 2000), and it flies toward the decoy target (0, 0, 0) at a speed of 300 m/s. This movement is uniform linear motion in the xOz plane. The position of M1 at any time during the movement is expressed as (R_d,0,G_d):

$$\begin{cases} R_d=20000-\dfrac{300\sqrt{101}t}{101} \\ G_d=2000-\dfrac{300\sqrt{101}t}{101} \end{cases} \tag{1}$$

For the smoke interference bomb (whose motion state in the first half is consistent with that of FY1), it flies toward the decoy target (0, 0, 0) at a speed of 120 m/s at the same altitude. FY1 always moves in the xOz plane, so its entire motion trajectory can be analyzed in a 2D plane. According to the description, the motion process of the smoke interference bomb can be divided into three stages[9-10]:

The first stage is the period from receiving the command to releasing the smoke interference bomb. During this stage, the smoke interference bomb (carried by FY1) moves in uniform linear motion. The x-axis coordinate of the smoke cloud center is updated using the kinematic formula:

$$R_w=17800-120t \tag{2}$$

At the end of this uniform linear motion (i.e., at the release point when s), t=1.5 s), R_w=17620.

The second stage is the period from releasing the smoke interference bomb to its detonation. In this stage, the environment is idealized, and external forces other than gravity (such as crosswinds) are excluded from affecting the smoke interference bomb. Therefore, the smoke interference bomb undergoes projectile motion during this period. Its x-axis and z-axis coordinates are updated using the kinematic formula:

$$\begin{cases} R_w=17800-120t \\ G_w=1800-4.9t^2 \end{cases} \tag{3}$$

where g=9.8 m/s^2. At the end of the projectile motion (i.e., at the detonation point when t=5.1 s), the coordinates of the smoke interference bomb are (17188, 0, 1736.496).

The third stage is the period from the detonation of the smoke bomb to its failure. During this stage, the smoke cloud sinks uniformly at a speed of 3 m/s. Its z-axis coordinate is updated using the kinematic formula:

$$G_w=1736.496-3(t-5.1) \tag{4}$$

The above three stages constitute the motion process of the smoke cloud, and its 2D schematic diagram is as follows Figure 2:

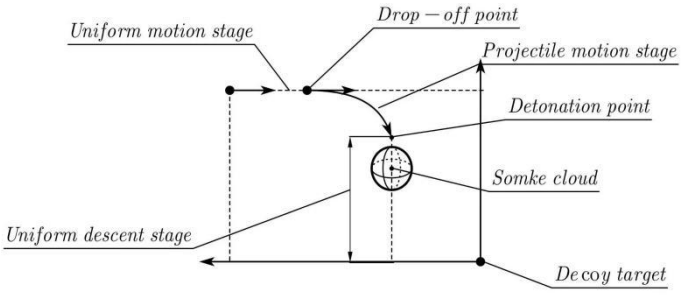


Figure 2 Schematic Diagram of the Smoke Cloud Motion Trajectory

Immediately after detonation, the smoke interference bomb forms a spherical smoke cloud. Through specific technologies, this smoke cloud sinks uniformly at a speed of 3 m/s. Experimental data show that the smoke concentration within 10 m of the cloud center can provide effective shielding for the target within 20 seconds after detonation. Therefore, the spherical smoke cloud is idealized as a uniform sphere with a radius of 10 m, and the problem of "effective shielding" here is analogous to the problem of "line-of-sight obstruction"—that is, under certain conditions, the missile may still be shielded even when outside the smoke cloud.

To quantify the effective shielding effect, an intuitive geometric-physical model is constructed in this study, and a cone tangent angle judgment method is introduced. Specifically, when Missile M1 is outside the smoke cloud, two sets of angles are considered:

α_1: The angle between the line connecting Missile M1 to the center of the smoke cloud and the tangent line of the smoke cloud sphere passing through Missile M1.

α_2: The angle between the line connecting Missile M1 to the center of the smoke cloud and the line connecting Missile M1 to any point on the surface of the real target.

If the condition $\alpha_1 \geq \alpha_2$ is satisfied, the condition is regarded as effective shielding; Otherwise, it is ineffective shielding. Schematic Diagram is shown in Figure 3.

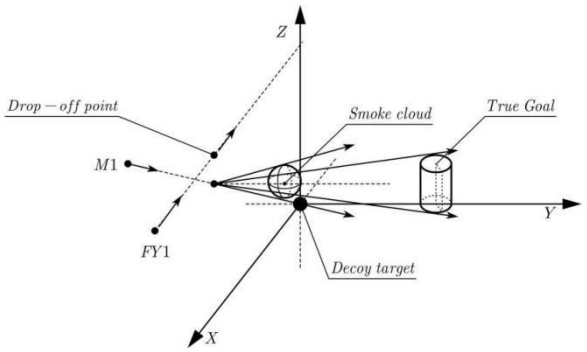


Figure 3 Schematic Diagram of Missile M1 Outside the Smoke Cloud

The cone angle α_1 can be calculated using the arcsine function:

$$\alpha_1 = arcsin\left(\frac{10}{\sqrt{(R_d - R_w)^2 + (G_d - G_w)^2}}\right) \tag{5}$$

Using the cylinder representation method in Analytic Geometry and combining the existing data in the problem, any point on the cylinder is set as $(7\cos\alpha_3, 7\sin\alpha_3 + 200, Z_z)$, where $\alpha_3 \in [0, 2\pi]$ and $Z_z \in [0, 10]$.

The cosine value of α_2 (the angle between the line connecting Missile M1 to the smoke cloud center and the line connecting Missile M1 to any point on the real target surface) can be calculated using the dot product of two vectors. The cosine function is constructed as follows:

$$h \equiv cos\alpha_2 = \frac{(R_d - 7cos\alpha_3, -7sin\alpha_3 - 200, G_d - G_z) \cdot (R_d - R_w, 0, G_d - G_w)}{|(R_d - 7cos\alpha_3, -7sin\alpha_3 - 200, G_d - G_z)| \cdot |(R_d - R_w, 0, G_d - G_w)|} \tag{6}$$

When Missile M1 is outside the smoke cloud, if the smoke cloud provides effective shielding, the condition $\alpha_1 \geq \alpha_2$ must be satisfied. However, due to the definition of α_2, its value varies depending on the selected point on the target cylinder. Since the position and size of the cylinder are fixed, the value of α_2 varies within a fixed range. Therefore, the effective shielding condition is optimized to $\min(\alpha_1) \geq \max(\alpha_2)$ in this study. The value range of the angle α_2 between the two lines is $[0,\pi]$. According to the monotonicity of the cosine function, when α_2 is maximized, $\cos\alpha_2$ is minimized. Thus, the objective function for solving α_2 is:

$$\alpha_2=arccos\frac{f_0}{}(h_{min}) \tag{7}$$

Then, the condition $\alpha_1 \geq \alpha_2$ can be simplified to:

$$\alpha_1 \geq arccos\frac{f_0}{}(h_{min}) \tag{8}$$

When Missile M1 is inside the smoke cloud, the smoke cloud provides effective shielding. That is, when the distance between the two points (Missile M1 and the smoke cloud center) is less than the radius of the smoke cloud, it indicates that Missile M1 is inside the smoke cloud. The mathematical expression is as follows:

$$(R_w-R_d)^2+(U_w-U_d)^2+(G_w-G_d)^2 \leq 10^2 \tag{9}$$

In the trajectory equations under the spatial coordinate system, the coordinates of Missile M1 and the smoke interference bomb only contain the time parameter t. Therefore, at time t, if effective shielding is achieved, it is marked as 1; otherwise, it is marked as 0. Finally, the effective shielding time required can be obtained by integrating the effective shielding moments in the above two stages:

$$\int_0^{25.1} N(t)dt \tag{10}$$

$$N(t)=\begin{cases}1, & \text{Effective shielding at time t,}\\ 0, & \text{Otherwise.}\end{cases} \tag{11}$$

2.2 Algorithm Design

Since the missile and the smoke cloud are moving at high speed, their relative positions change rapidly, and the effective shielding time window is short. To solve the effective shielding period more accurately, a variable-step search algorithm is adopted. By continuously updating the time step, the range of the optimal solution is gradually narrowed. Based on the variable-step search algorithm, programming is implemented in MATLAB. By continuously adjusting the time step, the solution result of the objective function is shown as follows (Figure 4):

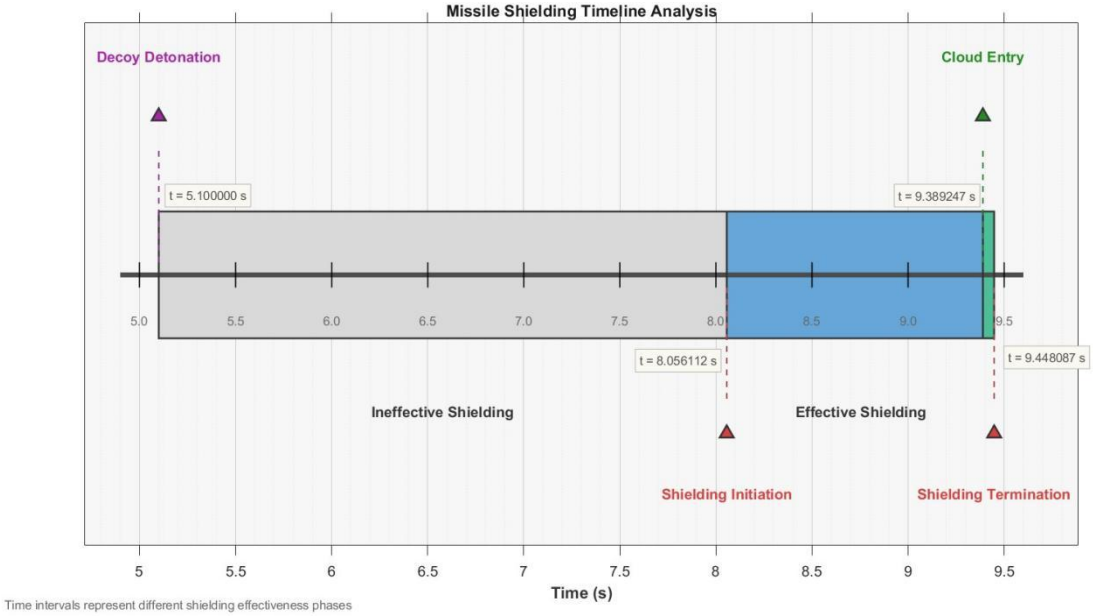


Figure 4 Step Diagram of Effective Shielding Status

Table 1 Key Moments of Effective Shielding

Detonation Time of Interference Bomb	5.100000 s
Start Time of Effective Shielding	8.056112 s
Time when Missile Enters the Cloud	9.389247 s
Time when Missile Exits the Cloud (End of Effective Shielding)	9.448087 s

From Table 1, it can be seen that the interference bomb detonates 5.1 seconds after the start of its movement and begins to provide effective shielding 8.056112 seconds later (at this time, Missile M1 is outside the smoke cloud). The missile starts to enter the cloud at 9.389247 seconds, and this stage still counts as effective shielding. After 9.448087 seconds, the missile exits the cloud, and there is no effective shielding from the smoke cloud thereafter. Therefore, the effective shielding time is calculated as 1.391975 seconds.

3 OPTIMAL DEPLOYMENT STRATEGY FOR SINGLE-DRONE SINGLE-SHOT

3.1 Model Establishment

3.1.1 Improvement of trajectory equations in the spatial coordinate system

The trajectory equations need to include the above parameters as decision variables. At the same time, the coordinates of the release point and detonation point of the smoke interference bomb are determined by its release time and detonation time, and both types of information can be uniformly described using time parameters. The difference lies in that the newly added parameters mentioned above need to be included as decision variables in the motion process of each stage.

First, the following parameters are introduced: Let v be the flight speed of FY1, α_4 be the angle between the flight direction of FY1 and the positive x-axis, t_1 be the time (in seconds) after FY1 receives the task and before releasing the smoke interference bomb, and t_2 be the time interval (in seconds) between releasing the smoke interference bomb and its detonation.

The analysis of the entire motion process of FY1 is basically consistent.

The coordinates (R_d,U_d,G_d) of the smoke interference bomb released by FY1 are expressed as (for $t\geq t_1+t_2$):

$$\left(17800+v\cos\alpha_4(t_1+t_2),v\sin\alpha_4(t_1+t_2),1800-0.5gt_1^2\right) \tag{12}$$

The coordinates (R_w,U_w,G_w) of the smoke cloud center formed after the detonation of the smoke interference bomb are expressed as (for $t\geq t_1+t_2$):

$$(17800+v\cos\alpha_4(t_1+t_2),\ v\sin\alpha_4(t_1+t_2),1800-0.5gt^2-3(t-t_1-t_2)) \tag{13}$$

3.1.2 Construction of objective function and constraint conditions

In the newly established trajectory equations under the spatial coordinate system, the coordinates of Missile M1 and the smoke interference bomb not only depend on the time parameter t, but also on the azimuth angle α_4 and flight speed v. To maximize the final effective shielding time, if effective shielding is achieved at time t, it is marked as 1; otherwise, it is marked as 0. For the maximum effective shielding time required finally, the integral result needs to be optimized for the maximum value. At the same time, effective shielding must occur after FY1 releases the smoke interference bomb, so the integral interval is narrowed to $[t_1+t_2,t_1+t_2+20]$ to reduce the amount of calculation. The expression is as follows:

$$max\int_{t_1+t_2}^{t_1+t_2+20}N(t)dt \tag{14}$$

$$N(t)=\begin{cases}1, & \text{Effective shielding at time t,}\\ 0, & \text{Otherwise.}\end{cases} \tag{15}$$

At the same time, effective shielding must satisfy either the missile being inside the smoke cloud or the cone tangent angle being greater than or equal to the maximum angle between the line connecting the missile to the cloud center and the line connecting the missile to the target cylinder. That is, $\alpha_1\geq\alpha_2$.

In addition, the objective function is also constrained by conditions such as the azimuth angle α_4 and flight speed v, which are summarized as follows:

$$\begin{cases}\alpha_1\geq arccos(h_{min})\\ (R_w-R_d)^2+(U_w-U_d)^2+(G_w-G_d)^2\leq 10^2\\ 70\leq v\leq 140\\ 0\leq\alpha_4\leq 2\pi\\ t_1<t_2\end{cases} \tag{16}$$

$$h_{min}\equiv\min\cos\alpha_2=\frac{(R_d-7\cos\alpha_3,-7\sin\alpha_3-200,G_d-G_z)\cdot(R_d-R_w,0,G_d-G_w)}{|(R_d-7\cos\alpha_3,-7\sin\alpha_3-200,G_d-G_z)|\cdot|(R_d-R_w,0,G_d-G_w)|} \tag{17}$$

3.2 Algorithm Design

According to the geometric model, the effective shielding of the smoke cloud involves complex geometric relationships and kinematic processes among Missile M1, the smoke cloud, and the real target, and its function exhibits significant nonlinear characteristics. Missile M1 flies toward the decoy target in uniform linear motion; after the smoke interference bomb detaches from the UAV, it first undergoes projectile motion, and the cloud formed after detonation sinks at a constant speed. The above physical processes result in an extremely complex relationship between the effective shielding time and the three decision variables, which cannot be solved by simple optimization methods. Meanwhile, considering that the solution space of the effective shielding period may have multiple local optimal solutions, traditional single-point search algorithms tend to fall into local optimality and fail to find the global optimal solution. Additionally, the three decision variables form a 3D decision space. Based on the above characteristics model, a hybrid optimization strategy combining genetic algorithm (GA) and particle swarm optimization (PSO) is selected. Among them, GA performs global search and introduces randomness to avoid premature convergence to local optimality, while providing a high-quality initial population for PSO; PSO then searches for the optimal solution within the population provided by GA.

Algorithm 2: Hybrid Optimization Algorithm Combining Genetic Algorithm and Particle Swarm Optimization

Parameter Initialization: Set the ranges of decision variables (azimuth angle α_4, flight speed v, release time t_1, detonation interval t_2), population size for GA, number of iterations, crossover probability, mutation probability, and PSO parameters (number of particles, number of iterations).

Generate Initial Population and Particle Swarm: GA randomly generates initial values of α_4, v, t_1, and t_2; PSO uses the optimal solution of GA as the initial position of particles.
Calculate Positions of Missile M1 and Smoke Cloud Center: Based on the decision variables, compute the real-time coordinates of Missile M1 and the smoke cloud center using kinematic equations.
Judge Effective Shielding: For each time step, determine whether the smoke cloud provides effective shielding for the target using the cone angle method and the "inside-cloud" judgment condition.
Calculate Total Effective Shielding Time: Total effective shielding time = Number of time steps meeting effective shielding conditions × Time step size.
Optimize to Find the Optimal Solution: GA retains excellent individuals through iteration to obtain a preliminary optimal solution; PSO initializes particles with the results of GA, continuously updates particle positions and speeds through iteration, and searches for the global optimal solution.
Model Solution and Result Analysis
Based on Algorithm 2, programming is implemented in MATLAB, and the solution results are shown in the following Table 2:

Table 2 Optimal Solution Results

Flight Direction Angle of FY1	4.7643° (with the positive x-axis as the reference, counterclockwise as the positive direction)
Flight Speed of FY1	115.63 m/s
Release Point of Smoke Interference Bomb	(17929, 10.78, 1800)
Detonation Point of Smoke Interference Bomb	(17934, 11.1837, 1799.58)
Maximum Effective Shielding Time	4.585 s

A coarse-step traversal algorithm is used to verify the optimization results. Based on the missile's coordinates, flight speed, flight direction, and the UAV's speed range, it is deduced that the reasonable coordinate range for the smoke cloud is $R\in(15875, 19725)$, $U\in(0, 1925)$, and $G\in(1738.5, 1800)$. A coarse-step traversal of the coordinates is performed to calculate the effective shielding time of the smoke motion on the cylindrical target when the corresponding coordinates occur.
The results show that the points with longer shielding time have a y-coordinate near 0, and the shielding time increases with the increase of x and z coordinates. The detonation coordinates in the bomb release strategy obtained by the optimization algorithm conform to this trend, which verifies the rationality of the results.

4 OPTIMIZATION OF MULTI-SHOT COLLABORATIVE OBFUSCATION FOR SINGLE-DRONE

4.1 Model Establishment

4.1.1 Coordinate representation in the spatial coordinate system

Based on the model establishment, since UAV FY1 needs to release 3 smoke interference bombs in, additional time variables need to be introduced to represent the motion trajectories of the other two smoke clouds. Let the center coordinates of the second smoke cloud be represented by time variables t_3 and t_4, and the center coordinates of the third smoke cloud be represented by time variables t_5 and t_6. The specific expressions are as follows:
The center coordinates (R_{w2},U_{w2},G_{w2}) of the second smoke cloud are expressed as (for $t\geq t_3+t_4$):

$$\begin{cases} R_{w2}=17800+v\cos f_0\alpha_4 t_3 \\ U_{w2}=v\sin f_0\alpha_4 t_3 \\ G_{w2}=1800-4.9t_4^2-3(t-t_3-t_4) \end{cases} \tag{18}$$

The center coordinates (R_{w3},U_{w3},G_{w3}) of the third smoke cloud are expressed as (for $t\geq t_5+t_6$):

$$\begin{cases} R_{w3}=17800+v\cos f_0\alpha_4 t_5 \\ U_{w3}=v\sin f_0\alpha_4 t_5 \\ G_{w3}=1800-4.9t_6^2-3(t-t_5-t_6) \end{cases} \tag{19}$$

4.1.2 Judgment of multi-smoke composite shielding

Since UAV FY1 will release 3 smoke interference bombs, there may be scenarios where two or more smoke clouds act together to achieve effective shielding. To evaluate the composite shielding effect, the surface of the real target cylinder is discretized into uniform small cells, and the center coordinates of each cell are used to represent its position.
Through sampling detection, it is determined whether the smoke clouds achieve effective shielding. If some cells are effectively shielded by one smoke cloud and the other cells are effectively shielded by another smoke cloud, and the union of the two parts of cells covers all cells, then the smoke clouds are considered to have achieved effective shielding.
From this, the judgment function $I(t)$ for whether the real target is completely shielded at time t can be derived.
Thus, the objective function is:

$$max \int_{t_1+t_2}^{t_1+t_2+20} N(t)dt \tag{19}$$

$$N(t)=\begin{cases}1, & \text{Effective shielding at time t,}\\ 0, & \text{Otherwise.}\end{cases} \tag{20}$$

The constraint conditions are:
$t_3-t_1\geq 1$ (Minimum time interval between the release of the first and second smoke bombs),
$t_5-t_3\geq 1$ (Minimum time interval between the release of the second and third smoke bombs).

4.2 Model Solution and Result Analysis

4.2.1 Algorithm design

Algorithm 3: Basin-Hopping Particle Swarm Optimization Algorithm
Objective: Design a strategy for FY1 to release 3 smoke interference bombs and solve the maximum effective shielding time.
Determine Optimization Variables: Identify the release angle α_4, flight speed v of the UAV, and each release time node and detonation time node of the smoke interference bombs.
Generate Initial Particles: Randomly generate initial values of optimization variables within their respective ranges, ensuring that the time interval constraints between releases are satisfied.
Fitness Calculation (Total Effective Shielding Time): For each particle, analyze the entire motion process, calculate the smoke trajectory, traverse the missile trajectory at time steps, determine whether the smoke clouds effectively shield the missile, and count the total shielding time and the individual shielding time of each smoke cloud.
Search for the Optimal Solution Based on PSO: Update particle positions and speeds according to PSO rules, and iterate to find the local optimal solution.
Restart Basin-Hopping: When PSO converges to a local optimal solution, perturb the global optimal solution to jump out of the local optimum, and restart PSO with the new solution as the initial point.
Output Results: After multiple basin-hopping restarts and PSO iterations, output the global optimal solution and the corresponding maximum effective shielding time.

4.2.2 Model solution and result analysis

Based on Algorithm 3, programming is implemented in MATLAB, and the solution results are shown in the following Table 3:

Table 3 Optimal Solution Results

Flight Direction Angle of FY1	4.7643° (with the positive x-axis as the reference, counterclockwise as the positive direction)
Flight Speed of FY1	115.63 m/s
Release Point of Smoke Interference Bomb 1	(17800, 0, 1800)
Detonation Point of Smoke Interference Bomb 1	(17812, 0.2053, 1799.951)
Release Point of Smoke Interference Bomb 2	(17919, 2.0526, 1800)
Detonation Point of Smoke Interference Bomb 2	(17931, 2.2578, 1799.951)
Release Point of Smoke Interference Bomb 3	(18994, 20.5256, 1800)
Detonation Point of Smoke Interference Bomb 3	(19005, 20.7308, 1799.951)
Maximum Effective Shielding Time	6.302 s

5 CONCLUSIONS

This study successfully established a geometric-physical model describing the process of drone smoke interference against missiles, employing either the conical angle method or distance determination method to precisely quantify the effective obfuscation effect of smoke clouds on cylindrical real targets. By applying different optimization algorithms, this study achieved significant maximization of effective concealment duration:
1. In calculating effective concealment duration for a single smoke grenade, based on initial parameters, the effective concealment duration of the smoke grenade for M1 was 1.391975s.
2. For the optimal deployment strategy of a single drone with a single flare, hybrid optimization using genetic algorithms and particle swarm optimization increased the maximum effective concealment time to 4.585 seconds.
3. For coordinated multi-flare concealment optimization by a single drone, the basin hopping-particle swarm optimization algorithm achieved a maximum composite effective concealment time of 6.3020 seconds.
In summary, this study demonstrates that multi-variable optimization of key parameters—including drone flight direction, velocity, and smoke grenade deployment timing—enables hybrid optimization strategies (e.g., GA-PSO and

Basin Jumping-PSO) to effectively address complex nonlinear problems. This provides a quantifiable and actionable solution for achieving maximum effective shielding under coordinated interference from a single drone deploying multiple smoke grenades.

Future research may further explore multi-UAV cooperative smoke deployment strategies under complex battlefield conditions, incorporating real-time environmental factors such as wind field variations and atmospheric stability. In addition, integrating machine learning for adaptive decision-making represents a promising direction for enhancing response speed and robustness in dynamic scenarios.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Huang Na, Niu Xinglin, Luo Zhengyou, et al. Research on Manufacturing and Performance Optimization of Smoke Disturbance Grenades in Special Ammunition. China Military-Civilian Conversion, 2025(09): 22-23.

[2] Huang Na, Luo Zhengyou, Niu Xinglin, et al. Research and Development of Novel Materials for Smoke Distraction Grenades. China Military-Civilian Conversion, 2025(08): 27-28.

[3] Luan Yongchao, Zhang Bin, Li Chenkai, et al. Method for Measuring Muzzle Bullet Velocity Under Heavy Smoke Interference. Journal of Ordnance Engineering, 2025, 46(03): 191-201.

[4] Luan Yongchao, Zhang Bin, Li Chenkai, et al. Method for Measuring Muzzle Projectile Initial Velocity Under Strong Smoke Interference. Journal of Ordnance Engineering, 2025: 1-11.

[5] Zhang Jun, Leng Zhihui, Xu Lei, et al. Application Research of Smoke Disruption Pods on a Certain Aircraft Model for Aerial Disruption. Trainer Aircraft, 2023(01): 16-19.

[6] Yang D Y, Qiu X B, Li C L, et al. Rapid Algorithm for Mie Scattering Coefficient in Smoke-Induced Interference of Optically Guided Weapons. Firepower and Command Control, 2017, 42(08): 56-60.

[7] Lei Dan. Research on Video-Based Smoke Detection Methods for Highway Tunnels. Dalian Maritime University, 2017.

[8] Zhang Xinyue. Signal Processing Technology for DSP-Based Frequency-Modulated Continuous Wave Laser Ranging. Beijing Institute of Technology, 2016.

[9] Jia B H, Feng Y, Jia W H. Application of dual-ring coaxial fiber optic sensors in blade tip clearance measurement. Advances in Lasers and Optoelectronics, 2015, 52(10): 91-97.

[10] Chen Zhibin, Zhang Chao, Song Yan, et al. Application of Gray-Scale Stretch Retinex in Smoke Image Enhancement with Large Dynamic Range. Infrared and Laser Engineering, 2014, 43(09): 3146-3150.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3060

ZERO VALUE DETECTION TECHNOLOGY FOR CERAMIC INSULATORS BASED ON CURRENT RESPONSE CHARACTERISTICS UNDER HIGH VOLTAGE IMPACT

Min Xie[1*], ChenLong Zhao[2], XiaoGang Li[1], ZhuHong Liu[2]
[1]*Tsinghua Pearl River Delta Research Institute, Guangzhou 510700, Guangdong, China.*
[2]*Guangzhou Guanghua Zhidian Technology Co., Ltd, Guangzhou 510700, Guangdong, China.*
Corresponding Author: Min Xie, Email:xiemin@ghzd-gd.com

Abstract: Ceramic insulators may experience a decrease in internal insulation performance and become zero value insulators due to environmental factors after being put into operation. Regular zero value testing is required to identify zero value insulators. The impulse high voltage method is one of the effective methods for zero value detection, but existing impulse high voltage methods only focus on the voltage response characteristics of insulators under impulse high voltage, without paying attention to the current response characteristics. This study proposes a zero value detection method based on the current response characteristics under impulse high voltage. The response current is converted into an induced voltage signal through a magnetic ring inductor, and the difference between zero value and the peak induced voltage of a normal insulator is used to achieve reliable identification of zero value. In the case of insulator surface contamination, this method is more robust and has a lower risk of misjudgment compared to detection methods based on voltage response characteristics. The research provides new ideas for insulation testing of ceramic insulators, which helps to improve the reliability of the power grid. Further exploration is needed in the future to determine the reasons for the reliability differences between the two methods.
Keywords: Ceramic insulators; Zero value; Testing; Impact high pressure; Current response characteristics

1 INTRODUCTION

As critical components for insulation and mechanical support in power transmission lines, porcelain insulators may experience degradation of internal insulation performance during long-term operation due to manufacturing defects or mechanical damage, resulting in the formation of "zero-value insulators" [1]. The presence of zero-value insulators significantly reduces the overall insulation effectiveness of insulator strings, induces local electric field distortion, and under adverse weather conditions may lead to line tripping or even large-scale power outages[2]. Therefore, identifying and troubleshooting zero-value insulators remains a crucial task in the operation and maintenance of power transmission lines.

Current zero-value detection methods for porcelain insulators primarily include distributed voltage measurement [3-4], insulation resistance testing [5], infrared/ultraviolet imaging [6-7], electric field analysis [8-11], and impulse high-voltage testing [12-16]. The distributed voltage method, which requires manual tower access to measure voltage distribution, is highly susceptible to environmental interference and demands strict operator safety protocols. While insulation resistance testing evaluates insulation status through DC resistance measurements, studies and practical applications reveal that resistance levels alone cannot fully indicate zero-value defects. Infrared/ultraviolet imaging methods remain vulnerable to environmental temperature and light interference, and shows limited sensitivity for completely penetrated zero-value insulators. Electric field analysis detects zero-value defects by measuring electrical field distortion around insulators, but its effectiveness is influenced by string configuration, environmental humidity, and measurement distance. In contrast, impulse high-voltage testing has emerged as a research hotspot due to its non-contact operation, rapid response, and sensitivity to high-resistance defects. This method applies transient high-voltage pulses to insulators, utilizing their electrical response characteristics under stress to determine the integrity of internal insulation.

Researchers including CHU W, WEN L, and MAO X have conducted studies on identifying zero-value insulators using impulse high-voltage detection methods. Their findings revealed that normal insulators exhibit high response voltages under impulse high-voltage conditions. In contrast, zero-value insulators demonstrate extremely low voltage withstand capabilities under such conditions. While these studies primarily focused on identifying zero-value insulators through voltage response characteristics under impulse high-voltage, current response characteristics were not addressed. In reality, the current response characteristics of insulating media under impulse high-voltage are equally critical, as parameters like response current amplitude are closely related to internal insulation conditions.

In light of this, the study aims to investigate zero-value detection technology based on insulator current response characteristics under impulse high voltage. It explores the feasibility of using current response-related criteria to identify zero-value insulators, while comparing the detection reliability of this method with voltage response-based detection techniques under operational conditions such as contamination accumulation.

2 METHODOLOGY

The most widely used porcelain insulators in overhead high-voltage transmission lines are disc suspension porcelain insulators, making the research focus primarily on zero-value detection for these insulators. The design includes the impulse high-voltage generator circuit and measurement circuit shown in Figure 1. The impulse high-voltage generator circuit boosts 12V power supply voltage to 400V through a DC voltage conversion circuit and zero-voltage switching circuit to charge the capacitor. The charged capacitor rapidly discharges under high-voltage switch control, with the discharge voltage amplified by a high-frequency transformer to generate a microsecond-level impulse high-voltage peak of 60kV, which is applied across the test insulator to form a detection circuit. The measurement circuit consists of a magnetic ring inductor mounted on the detection circuit and an RC acquisition circuit. The magnetic ring inductor converts the current signal in the detection circuit into a voltage signal, while the RC acquisition circuit delays the voltage signal's decay to ensure reliable sampling of instantaneous voltage. The voltage sampling frequency in the measurement circuit is 10kHz.

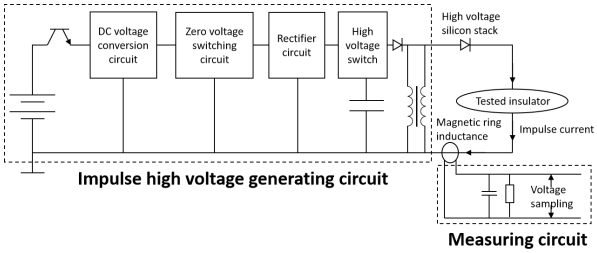


Figure 1 Schematic Diagram of the Circuit for Generating High Voltage Impact and the Measurement Circuit

Based on the aforementioned circuit, when a microsecond-level impulse high voltage is applied to the test insulator, a response current I is instantaneously generated in the detection loop formed by the test insulator and the high-voltage generation circuit. Under the influence of this varying response current, an induced voltage U is generated in the magnetic ring inductor.

3 TEST RESULTS

3.1 Normal and Zero Value Insulator Detection

Based on the above circuit and test method, the detection of normal and zero value insulators with different tonnage was carried out respectively, and the voltage-time curves collected were shown in Figure 2 and Figure 3 respectively.

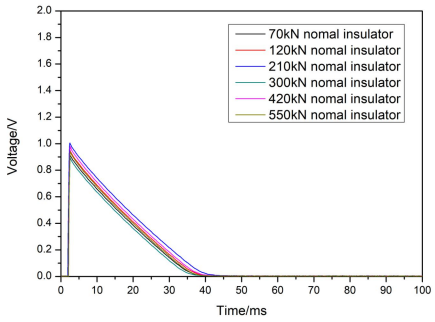


Figure 2 Voltage Time Curve of Normal Insulator

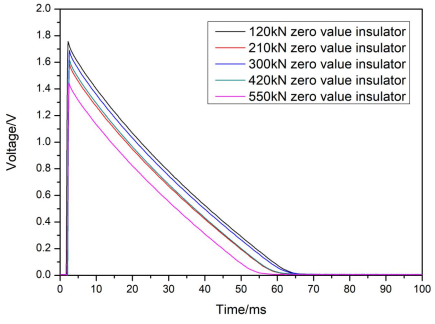


Figure 3 Voltage Time Curve of Zero Value Insulator

As shown in Figures 2 and 3, the collected voltage exhibits two distinct phases of rise and fall. During the voltage rise phase in Figures 2 and 3, the peak voltages of multiple normal insulators with different tonnage values show remarkable similarity, clustering around the 0.9V~1.0V range. Similarly, the peak voltages of multiple zero-value insulators with varying tonnage values demonstrate comparable proximity, concentrated within the 1.5V~1.8V range. This significant difference indicates that when 60kV impulse high voltage is applied across normal and zero-value insulators, distinct response currents emerge in the detection circuit. Normal insulators exhibit smaller response currents, generating induced voltage peaks of approximately 1.0V, while zero-value insulators demonstrate larger response currents, producing induced voltage peaks exceeding 1.5V. This fundamental distinction allows for accurate differentiation between normal and zero-value insulators.

3.2 Comparative Analysis with Voltage-Response-Based Detection Methods

3.2.1 Technical principles and test results of detection method based on voltage response characteristics

The designed detection circuit based on voltage response characteristics is shown in Figure 4. The impulse high-voltage generation circuit is the same as that in Figure 2. The measurement circuit mainly uses a voltage divider to collect the response voltage at both ends of the tested insulator. The resistance value of resistor 1 is 1.2GΩ, and the resistance value of resistor 2 is 6kΩ.

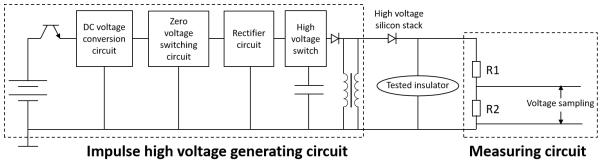


Figure 4 Schematic Diagram of Circuit Design for Detection Method based on Voltage Response Characteristics

When the zero value detection based on voltage response characteristics is carried out with the circuit shown in Figure 4, the voltage-time curves of normal and zero value insulators are shown in Figure 5 and Figure 6.

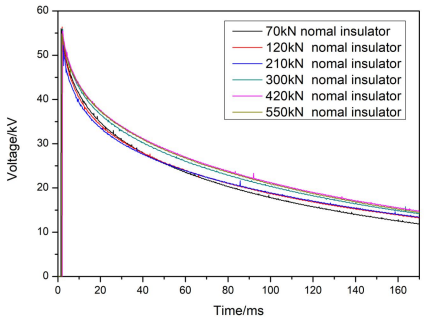


Figure 5 Voltage Time Curve at both Ends of a Normal Insulator based on Voltage Response Characteristic Detection Method

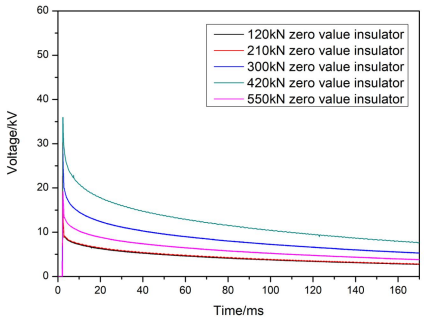


Figure 6 Voltage Time Curve at both Ends of Zero Value Insulator based on Voltage Response Characteristic Detection Method

As shown in Figures 5 and 6, the voltage response-based detection method can accurately distinguish between normal and zero-value insulators. Under 60kV impulse high voltage, normal insulators exhibit peak response voltages exceeding 45kV with slower decay, while zero-value insulators show peak voltages not exceeding 40kV with faster

decay. Therefore, the peak response voltage serves as a reliable indicator for differentiating normal and zero-value insulators.

3.2.2 Comparative study of two detection methods in the case of insulator surface contamination

Three standard 210kN insulators were selected. Kaolin and NaCl were prepared to equationte two suspensions: one containing only kaolin and another containing both kaolin and NaCl. Two insulators were coated with kaolin suspension and the mixed suspension, respectively, achieving a surface equivalent ash density of 2.0 mg/cm^2 and an equivalent ash density of 2.0 mg/cm^2 with a salt density of 0.3 mg/cm^2. A third insulator served as a control sample without coating. The three insulators were tested using voltage response and current response characterization methods, with results shown in Figures 7 and 8.

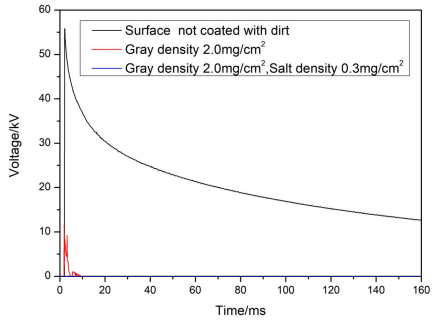


Figure 7 Detection Results Based on Voltage Response Characteristic Detection Method under Surface Contamination of Insulators

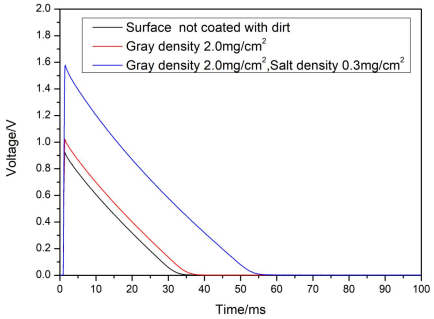


Figure 8 Detection Results Based On Current Response Characteristic Detection Method under Surface Contamination of Insulators

As shown in Figures 7 and 8, when using voltage response-based detection methods, the voltage-time curve exhibits significant changes under normal insulation conditions with 2.0 mg/cm^2 equivalent ash density and no salt density contamination, leading to misclassification as a zero-value insulator. Under similar conditions with 2.0 mg/cm^2 ash density and 0.3 mg/cm^2 salt density, the response voltage drops to zero, resulting in misjudgment. Conversely, current response-based detection methods show no significant changes in the voltage-time curve under normal conditions with 2.0 mg/cm^2 ash density and no salt contamination, preventing misclassification. However, when exposed to severe surface contamination with 2.0 mg/cm^2 ash density and 0.3 mg/cm^2 salt density, the voltage peak exceeds 1.5V, causing false zero-value classification.

Therefore, when the surface of the insulator is contaminated, the detection method based on the voltage response characteristics is more sensitive to the pollution and has a higher risk of misjudgment, while the detection method based on the current response characteristics has a risk of misjudgment only when the surface of the insulator has both ash density and salt density and heavy pollution.

4 CONCLUSION

(1) By analyzing the current response characteristics of porcelain insulators under impulse high voltage, reliable identification of zero-value porcelain insulators can be achieved. The detection circuit incorporates a magnetic ring inductor to convert response current signals into induced voltage signals. Since zero-value and normal insulators exhibit distinct peak values in induced voltage under impulse high voltage, this voltage difference enables accurate differentiation between normal and zero-value insulators.

(2) Under the condition that the surface of the insulator is wet and the surface area is contaminated, the detection method based on the current response characteristics is more reliable than the detection method based on the voltage response characteristics, and the risk of misjudgment is lower.

(3) In the future, it is still necessary to further explore the reason why the detection method based on current response

characteristics is more reliable than the detection method based on voltage response characteristics when the surface resistance of insulators is reduced.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Liu H, Geng S, Wang J. Aging Analysis of Porcelain Insulators Used in UHV AC Transmission Line. Insulators and Surge Arresters, 2022(6).
[2] Xia L, Cheng D, Qin J, et al. The Analysis on a Burst Fault of Double Umbrella Type Porcelain Insulators on a 220kV Overhead Transmission Line. Electrical Engineering, 2017(9): 72-74.
[3] Ge Y. Research on Detection Method of Porcelain Insulator State. Beijing: North China Electric Power University, 2013.
[4] Cheng Y, Xia L, Sun H, et al. Research on Zero-value Detection Technology of Porcelain Insulator Based on Electric Field Distribution Curve. High Voltage Apparatus, 2023, 59(11): 74-83.
[5] Sun Q, Liu G, Jin J, et al. Heating Mechanism of Disk Suspension Porcelain Insulator with Low(Zero)Value. Insulators and Surge Arresters, 2024(6): 138-144.
[6] Fu W, Wang W, Dong J, et al. Study on the Applicability of Detecting Deterioration Insulator Based on Infrared Thermal Image. High Voltage Apparatus, 2018, 54(2): 0110-0114.
[7] Chen Y, Guo J, Wu X. Fast Detection of Zero Insulator and Polluted Insulator Based on Infrared Temperature Measurement. High Voltage Apparatus, 2015, 51(6): 0191-0194.
[8] Wang L, Li X, Song B, et al. Researches of Simulation and Detection for Electric Field Distribution Along Faulty Insulator on Transmission Line. High Voltage Apparatus, 2018, 54(10): 0049-0055.
[9] Chen M, Pei H, Chen X, et al. Electric field analysis of transmission line insulator strings under different deterioration conditions. 2024: 918-927.
[10] Zhang D, Chang Z, Wan W, et al. Zero-value insulator detection technology based on local electric field. Electric Power Engineering Technology, 2024, 43(4): 193-200.
[11] Li Y, Huo F, Nan J, et al. Spatial Electric Field Distribution Characteristics of Insulator Strings in High-Voltage Overhead Transmission Lines With Zero Value Insulators. HUNAN ELECTRIC POWER, 2024, 44(5): 37-44.
[12] Chu W, Cao B, Chu W, et al. Research on rapid detection method for zero-value faults of disc-type suspension porcelain insulators. Power Electronics Technology, 2025.
[13] Wen L, Huang X, Cao B, et al. An Impulse Voltage-Based Compact Device for Porcelain Insulator Defect Detection. IEEE Transactions on Dielectrics and Electrical Insulation, 2024, 31(6): 3185-3192.
[14] Mao X, Zhou X, Zhang Y, et al. Research on Zero Value Detection Technology of Porcelain Insulator. E3S Web of Conferences, 2021, 257: 01021.
[15] Zhou X, Kan Y, Li Z, et al. Design and Research of Portable UHV Zero-Value Insulator Detection Device. Hubei Electric Power, 2021, 45(1): 20-27.
[16] Xu X, Li X, Sun W, et al. Research on Discharge Characteristic of Insulator with AC Power Superposed Impulse Voltage. YUNNAN ELECTRIC POWER, 2014, 42(2): 22-24.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3061

CONTROL METHOD OF A MACHINE VISION-BASED ROBOT FOR PUMP PIPE INSPECTION

Dong Liu, ChangCheng Wan*, Long Xie, Peng Liu
School of Mechanical Engineering, Xihua University, Chengdu 610039, Sichuan, China.
Corresponding Author: ChangCheng Wan, Email: 844736962@qq.com

Abstract: The pump pipe detection robot is used to detect the blockage of the concrete pump pipe, and it is an important equipment to ensure the safe and efficient transportation of the pump pipe. Currently, traditional inspection relies heavily on manual experience, which is not only time-consuming and inaccurate but also severely impacts construction progress and may even cause economic losses. To achieve precise and efficient blockage detection, this paper designs a machine vision-based pipeline detection robot control system. The core of this system utilizes optical fiber as the communication medium, fully leveraging its high bandwidth, low latency, and electromagnetic interference resistance to enable efficient and reliable transmission of robot control commands and high-definition image data.
Keywords: Pipeline robot; Control system; Fiber optic communication; Machine vision

1 INTRODUCTION

As a critical component of concrete delivery systems, pump pipes undertake high-pressure, long-distance concrete transportation tasks in the construction of high-rise buildings and large bridges[1]. The operational status of concrete pump pipes directly impacts construction continuity and project schedules. During actual use, factors such as operator error, pumping equipment issues, and concrete selection often lead to pump pipe blockages[2]. Manual blockage clearance is not only inefficient and inaccurate but also causes project delays and economic losses[3]. Therefore, employing a pump pipe detection robot equipped with machine vision for blockage detection offers advantages such as real-time capability and high inspection efficiency. However, pump pipe robots face challenges in remote control and signal transmission, including high signal latency and low transmission bandwidth. For inspection tasks involving small-diameter pipes, long distances, and vertical working conditions, inspection robots urgently require a suitable control system to address these current issues.

Pipeline robots has reached a relatively mature stage, with both remote control technology and machine vision technology seeing extensive development and application. The reliability of pipeline robot control and communication is crucial for ensuring their long-term, efficient operation. Scholars have conducted extensive research in areas such as pipeline robot control system design, communication architecture, and transmission protocols, achieving significant progress. Kazeminasab et al. proposed a pipeline robot system based on wireless relay nodes and multi-stage motion control algorithms, supporting real-time data transmission and motion switching during long-distance inspections[4]. Bai et al. designed an underwater bionic vehicle motion system (UBVMS) that interacts with a ground control console via TCP/IP communication[5]. Han et al. developed an articulated pipeline robot system based on CAN bus and distributed architecture[6], achieving efficient inter-node communication. Nguyen et al. proposed a pufferfish-inspired soft robot with an external air source and wired communication to enable its adaptive movement in variable-diameter pipelines[7]. Jeon et al. developed a wheeled robot integrating multi-motor cooperative control and wireless communication modules for large-diameter water pipe inspection and remote monitoring[8]. Liu et al. designed a multi-gait snake robot with centralized control via a host computer and CAN bus[9], generating complex motions through the trunk curve method. Zhen et al. developed a pneumatic pipeline robot capable of crawling through variable-diameter pipes[10], utilizing centralized control via a 485 bus between a subordinate device and a host PC. Luo Jiman et al. achieved closed-loop regulation of support leg pressure using fuzzy PID and an STM32 controller[11], significantly enhancing the robot's motion stability and interference resistance. In machine vision, Liu Simo proposed an obstacle avoidance strategy for pipeline robots by integrating stereo vision with ant colony optimization[12], enabling obstacle recognition and intelligent path planning based on 3D reconstruction. Kabir et al. developed a visual system based on Mask R-CNN and LiDAR cameras[13], achieving high-precision instance segmentation and distance measurement of rocks inside pipelines under complex lighting conditions. Yang Wanting investigated machine vision-based automatic identification and classification methods for pipeline weld defects using image enhancement[14], segmentation, and morphological processing.

In summary, although existing research provides ample references in areas such as control methods and machine vision, studies in the field of concrete pump pipe inspection robots remain largely unexplored. To achieve stable operation in long-distance, small-diameter, vertical metal pipelines while effectively addressing challenges such as communication latency, metal shielding interference, and cable dragging, the control system of the pump pipe detection robot must exhibit high stability, strong anti-interference capabilities, and low latency. To this end, this study specifically designed a highly reliable control system and experimentally validated its outstanding performance.

2 CONTROL SYSTEM FOR PUMP PIPE DETECTION ROBOT

2.1 Control System Framework

The control system framework for the pump pipe detection robot is illustrated in Figure 1. The system operates by transmitting control commands from a PC terminal. These commands are converted into optical signals by a ground-based optical transceiver, then transmitted via fiber optic cable to a sky-based optical transceiver. The sky-end optical transceiver converts the optical signal into a digital signal, which is then sent to the control board. The control board parses the control commands and outputs two identical-frequency PWM signals to drive two brushless motors. These two PWM signals can be set to different duty cycles. The pulse signals from the Hall sensors of the two motors are captured and counted by the control board, which calculates the rotational speed and sends it to the sky-end optical transceiver. The sky-end optical transceiver converts the feedback information and image data into optical signals, transmitting them in real-time to the ground-end optical transceiver. The ground-end optical transceiver converts the optical signals back into digital signals and sends them to the PC terminal.

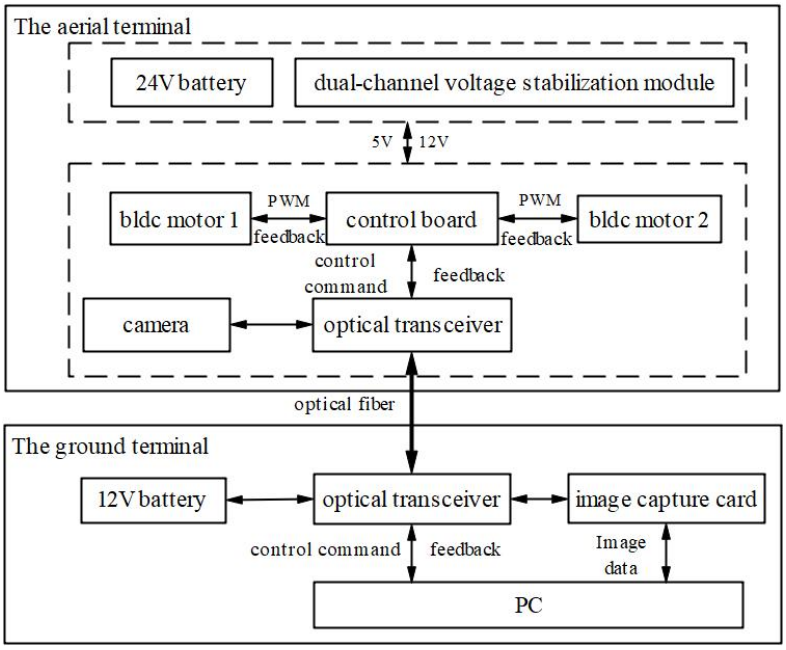


Figure 1 Pump Pipe Detection Robot Control System Framework

2.2 Composition of The Control Hardware

The main hardware composition of the pipeline robot control system is illustrated in Figure 2. The aerial terminal consists of the robot body, control board, optical transceiver, dual-channel voltage stabilization module, and 24V battery; the ground terminal includes an optical transceiver, image capture card, 12V battery, and PC terminal. Communication between the ground terminal and the aerial terminal is transmitted via optical fiber media. The robot body, as the main equipment for performing pipeline inspection tasks, is composed of a wheeled support mechanism, a diameter-variable mechanism, a camera, two brushless motors, and other components, capable of completing operations such as patrol inspection and image data collection under remote commands. The control board serves as the main controller of the robot, responsible for receiving instructions from the ground terminal, controlling the robot's movement, and sending information such as the robot's movement speed to the ground terminal.

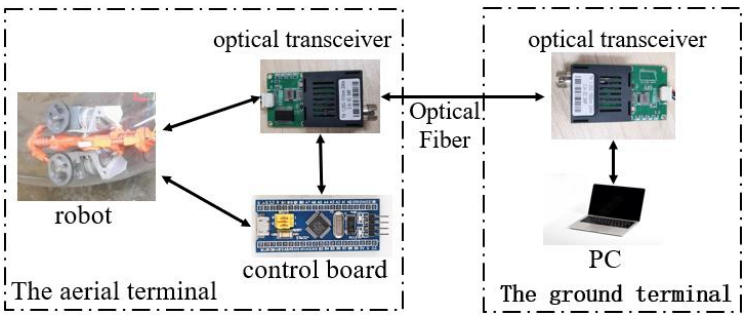


Figure 2 The Main Hardware Components of the Control System

The control board utilizes the STM32F103C8T6 minimal system board, integrating a rich array of timers,

communication interfaces, and analog functions to fully meet the device's control and communication requirements. The system transmits one analog video signal and one TTL-level signal via the optical transceiver unit, used for image data and control commands respectively. The optical transceiver utilizes an FC fiber optic interface, offering high bandwidth, high reliability, and strong electromagnetic interference resistance to establish stable, long-lasting fiber connections. Brushless motors are adopted, with a rated voltage of 12V, a no-load speed of up to 6000rpm, and a rated power of 6W. The camera utilizes an analog signal camera with a resolution of 720P. The image capture card acquires analog video signals and converts them into digital image signals. This capture card supports both 1080P and 720P resolutions. Additionally, the system is equipped with a voltage stabilization module, which can convert the externally input 24V voltage into 12V and 5V to provide stable power supply for the control board, optical transceiver, and camera respectively.

3 CONTROL SOFTWARE AND MACHINE VISION ALGORITHMS

3.1 Control Software

To achieve precise motion control and status monitoring for the detection robot, an embedded control software based on the STM32F103C8T6 microcontroller was designed. This software adopts a layered architecture and modular design, integrating core functions such as motor drive, sensor feedback, and communication management to ensure system real-time performance, reliability, and maintainability.

The control software employs a four-layer hierarchical architecture, as illustrated in Figure 3. The hardware driver layer directly manipulates STM32 peripheral registers to initialize and implement low-level drivers for GPIO, timers (TIM1/2/4), serial port (USART1), and interrupt controller (NVIC), providing a unified hardware access interface for upper layers. The hardware abstraction layer encapsulates functions such as motor control, encoder acquisition, direction switching, and serial communication. The functional module layer contains core modules including PWM control, rotational speed calculation, command parsing, and state management. These modules respectively handle dual-motor speed regulation, encoder pulse processing, serial command execution, and system state maintenance. The application layer implements task management through main loop scheduling, coordinates the work of various modules, and completes the full-process business logic of robot control.

The system employs a hybrid scheduling mechanism combining interrupts and polling to ensure real-time response for critical tasks. A SysTick timer is configured to generate a 1 ms time base, which is used for the periodic triggering of rotational speed calculation and status reporting. Encoder input capture interrupts (TIM2/TIM4) are assigned the highest priority to guarantee accuracy in motor pulse counting; serial port receive interrupts (USART1) enable timely response to control commands and prevent data loss.

The communication protocol employs plaintext ASCII format, supporting both multi-parameter batch processing and real-time response. The control commands sent by the host computer contain target duty cycle and steering information, and the slave computer returns an acknowledgment message after execution; Simultaneously, the system proactively reports operational status including motor speed, direction, and cumulative pulse count at 1-second intervals. To improve system reliability, multiple measures such as parameter boundary checking, software timeout judgment, and hardware status monitoring are implemented at the software level, ensuring that the system can achieve safe shutdown or state recovery in cases such as communication abnormalities, sensor failures, or control parameter out-of-bounds.

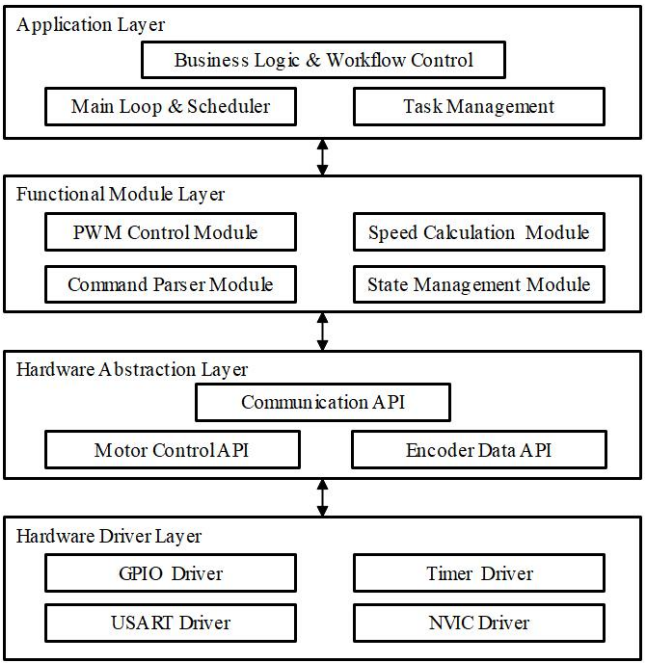


Figure 3 Control Software Architecture

3.2 Machine Vision Algorithms

This study adopts a machine vision algorithm based on Hough circle transform to achieve the detection of circular obstacles in pipelines and the judgment of passing ability. Through three main steps—image preprocessing, edge detection, and circle recognition—the algorithm can accurately identify circular obstacles in pipelines and determine whether the robot can pass safely according to their radius dimensions.

Prior to circular detection, input images undergo Gaussian filtering to eliminate noise interference. The core Gaussian filter formula is:

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

where (x, y) represents pixel coordinates and σ denotes the standard deviation. This formula defines a two-dimensional Gaussian kernel function, which smooths the image through convolution operation. While retaining edge information, it effectively suppresses noise and provides a high-quality input image for subsequent edge detection.

The core idea of Hough circle transform is to convert the problem of circle detection in the image space into the problem of peak detection in the parameter space. For an edge point (x_i, y_i) in the image space, it will be transformed into a three-dimensional conical surface after transformation, and the circle equation it satisfies in the parameter space (a, b, r) is:

$$(x_i-a)^2+(y_i-b)^2=r^2 \tag{2}$$

where (a, b) represents the coordinates of the center of the circle and r denotes the radius. The conical surfaces of points on the same circle in the parameter space will have the same intersection point (a, b, r), and the task of Hough circle detection is to find this point to determine the circle in the image. The Hough transform finds all satisfying pixels by traversing the image, and its formula is expressed as:

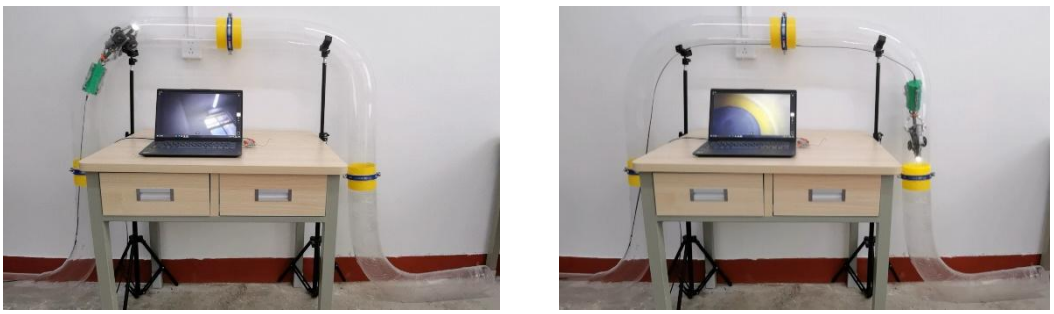
$$P = \{(x_0, y_0, r): (x_i - x_0)^2 + (y_i - y_0)^2 = r^2\} \tag{3}$$

where (x_0, y_0) represents possible coordinates of the circle center. The algorithm conducts cumulative voting in the parameter space to find parameter combinations with votes exceeding the threshold, thereby determining the circles present in the image.

4 EXPERIMENTAL VALIDATION

4.1 Motion Control Experiment

The experiment tests the robot's communication and control capabilities, and is conducted on a pipeline experimental platform with an inner diameter of 120 mm. Control commands are sent through the PC terminal to control the robot to walk, climb, turn and hover in the pipeline. As shown in Figure 4, the robot can walk horizontally, climb vertically, pass through elbows and hover in the pipeline, and can receive commands to adjust speed and move backward.



(a) Robot through a bend pip (b) The robot climbs vertically

Figure 4 The Robot Runs in the Pipeline

4.2 Image Signal Transmission Experiment

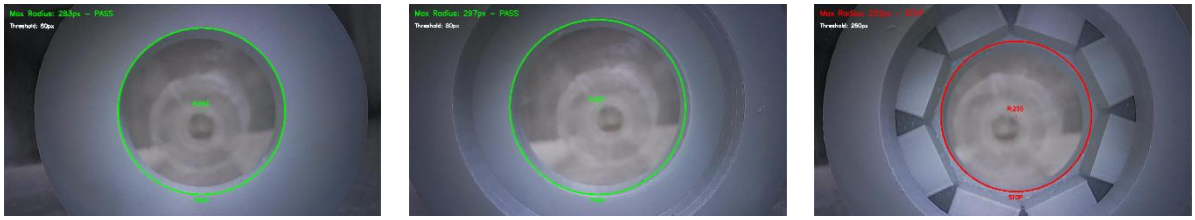
The robot is put into the pipeline for operation in the experiment. Equipped with a camera, the robot collects images synchronously. The captured real-time images are transmitted to the ground via optical fiber. The ground PC control terminal receives and displays the video. With a transmission rate of 1 gigabit per second, the optical fiber ensures high-speed, stable and smooth data transmission. Figure 5 shows the comparative images taken by the robot at different distances from obstacles in the pipeline, all with an image resolution of 1280×720.

(a) Close-range images (b) Mid-range image (c) Long-range image

Figure 5 Inside the Pipeline Image

4.3 Visual Detection Experiment

Images were collected in a pipeline environment with real pipe diameters in the experiment to verify the algorithm's effectiveness. Figure 6 shows the detection results of circular obstacles in the pipeline. The algorithm can accurately identify contours, marking passable ones with green circles and impassable ones with red circles. This method provides a reliable basis for robot navigation and effective visual perception support for autonomous inspection tasks.



(a) Passable (b) Passable (c) Not passable

Figure 6 Obstacle Detection Results in Pipeline

5 CONCLUSION

This paper addresses the problems of difficulty and low efficiency in clogging detection of concrete pump pipes, and designs a pipeline inspection robot control system based on machine vision. Through the construction of a pipeline testing platform, experiments were conducted on the robot's motion control, image transmission, and visual detection capabilities. experimental results show that: the system can transmit control commands via optical fiber to control the robot to complete actions such as walking, climbing, turning and hovering in the pipeline; the robot can capture the clogging status inside the pipeline and transmit images back to the PC terminal in real time, stably and with low latency through optical fiber; the visual algorithm can effectively identify the size of circular obstacles, providing a key basis for the robot's passage decision-making.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Li Yong. Research on Pumping Construction Technology of High-strength Concrete in Super High-rise Buildings. Engineering and Technological Research, 2022, 7(09): 106-108.
[2] Wang Qiwen, Yu Jiaxin. Concrete Pumping and Anti-blocking Pipe Construction Technique for Super-tall Building. Tianjin Science & Technology, 2017, 44(10): 66-70.
[3] Ding Hongchun, Lu Jincai, Liang Shuang, et al. Technology of preventing pipe plugging by pumping concrete washing and piping. Water Sciences and Engineering Technology, 2018(04): 73-75.
[4] Kazeminasab S, Banks M K. A localization and navigation method for an in-pipe robot in water distribution system through wireless control towards long-distance inspection. IEEE Access, 2021, 9: 117496-117511.
[5] Bai X, Wang Y, Yang Z, et al. Design and Pipeline Tracking Control of an Underwater Biomimetic Vehicle-Manipulator System With Hybrid Propulsion. IEEE Transactions on Cybernetics, 2025, 55(7): 3073-3084.
[6] Han Z, Zhu Z. Articulated robot system for energy pipeline maintenance. Energy Reports, 2022, 8: 267-274.
[7] Nguyen L V, Kim H, Nguyen K T, et al. Adaptable cavity exploration: Bioinspired vibration-propelled PufferFace Robot with a morphable body. Science Advances, 2025, 11(18): eads3006.
[8] Jeon K W, Jung E J, Bae J H, et al. Development of an in-pipe inspection robot for large-diameter water pipes. Sensors, 2024, 24(11): 3470.
[9] Liu J, Li M, Wang Y, et al. Multi-gait snake robot for inspecting inner wall of a pipeline. Biomimetic Intelligence and Robotics, 2024, 4(2): 100156.
[10] Zhen J, Ma T, Wang XY. Design of a Pipeline Robot Based on the Integrated Valve Control. Machine Tools & Hydraulics, 2022, 50(17): 57-60.

[11] Luo Jiman, Liu Shiheng, Ma Siyuan, et al. Pressure Stability Control of Supporting Legs of Pipeline Robot Propulsion Device. Journal of Shenyang Jianzhu University (Natural Science), 2023, 39(05): 939-946.
[12] Liu Simo. Study of Obstacle Avoidance in Pipeline Inspection Robots Based on Visual Perception and Ant Colony Algorithm. Journal of Lanzhou Petrochemical University of Vocational Technology, 2025, 25(01): 37-42.
[13] Kabir H, Lee HS. Mask r-cnn-based stone detection and segmentation for underground pipeline exploration robots. Applied Sciences, 2024, 14(9): 3752.
[14] Yang Wanting. Research on Pipeline Robot Weld Defect Detection Based on Machine Vision. Agricultural Technology & Equipment, 2024(08): 89-92.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3059

GAME-THEORETIC APPROACH TO DYNAMIC CONTROL SUBAREA ADJUSTMENT IN OVERSATURATED URBAN ROAD NETWORKS: EVIDENCE FROM GUANGZHOU

WeiBin Zhao, XinHai Xia*
School of Future Transportation, Guangzhou Maritime University Guangzhou, Guangzhou 510725, Guangdong, China.
Corresponding Author: XinHai Xia, Email: xiaxinhai@126.com

Abstract: Driven by rapid urbanization and motorization, peak-hour oversaturation has become a chronic issue in Guangzhou's urban road network. Core districts such as Tianhe and Haizhu frequently experience congestion indices exceeding 7.8 during morning and evening peaks, with average vehicle speeds dropping below 19 km/h. Queue spillover triggers cascading "domino effects," severely degrading system-wide traffic efficiency. While control subareas serve as fundamental units for regional signal coordination, conventional static partitioning fails to adapt to dynamic traffic fluctuations, and existing dynamic methods often neglect strategic interactions and interest conflicts among subareas, leading to suboptimal coordination.

This study proposes a game-theoretic framework for dynamic subarea adjustment under oversaturated conditions, using Guangzhou as a case study. We develop two decision-making paradigms: a centralized cooperative game model that minimizes total network control cost, and decentralized non-cooperative models—including Stackelberg and Nash equilibria—that account for autonomous subarea behavior. Through theoretical derivation, numerical simulation, and validation via the VISSIM microsimulation platform, we analyze equilibrium properties and system performance across different game structures.

Results demonstrate that the centralized model achieves the lowest total control cost—31.5% lower than the Nash equilibrium—but suffers from poor real-time responsiveness due to computational complexity. The Stackelberg model, leveraging a "central guidance-subarea response" mechanism, strikes an optimal trade-off: it reduces total delay by 18.2% compared to Nash, increases adjustment frequency by 112.5% relative to centralized control, and maintains robust performance under uncertainty. Meanwhile, the Nash model exhibits superior robustness (delay fluctuation ≤4.8% under ±20% data noise) but incurs significant efficiency losses.

Parametric analysis reveals that when the collaboration weight $\lambda > 0.6$, subarea merging probability increases by 42% and queue lengths decrease by 27%. Furthermore, constraining subarea size to 2-5 intersections optimally balances management overhead and coordination benefits—a finding validated across Guangzhou's heterogeneous urban fabric. This research provides both theoretical grounding and a practical implementation pathway for intelligent, adaptive traffic control in oversaturated megacities, directly supporting Guangzhou's "14th Five-Year Plan" for Intelligent Transportation Systems (ITS).

Keywords: Oversaturated road network; Control subarea; Game equilibrium; Dynamic adjustment; Signal coordination

1 INTRODUCTION

1.1 Research Background

1.1.1 Severity of oversaturation in Guangzhou

Guangzhou's urbanization rate surged from 66.2% in 2010 to 86.1% in 2024, accompanied by a motor vehicle fleet exceeding 4 million. This dual pressure has pushed traffic demand consistently beyond capacity thresholds, rendering oversaturation a norm during peak hours. According to the 2024 Guangzhou Annual Report on Transportation Development [1], core zones—including Zhujiang New Town (Tianhe), Guangzhou Avenue South (Haizhu), and Huanshi East (Yuexiu)—recorded average congestion indices of 8.0, 7.9, and 7.7, respectively, all classified as "severe congestion." On arterial corridors like Tianhe Road and Middle Guangzhou Avenue, peak-hour speeds plummet to 16-18 km/h—over 52% slower than off-peak levels.

Under such conditions, queues propagate at 12-16 meters per minute, with approximately 35% of intersections experiencing spillover. This phenomenon reduces adjacent link capacities by 25-45%, initiating a feedback loop of "congestion diffusion → efficiency collapse." A notable example occurred in April 2024: queue overflow at the Tianhe Road-Sports West Road intersection triggered a chain reaction across six downstream junctions, inflating regional average delay from 13 to 38 minutes—a 192.3% increase.

The situation intensifies during mega-events. During the 135th Canton Fair, traffic volume around the Pazhou Convention and Exhibition Center spiked by 60%, pushing demand intensity $D(t)=q(t)/C^{cap}$ to 1.9. Spillover slashed traffic efficiency by 50% on East Xin'gang Road and Yuejiang Middle Road, exposing critical vulnerabilities in current control strategies.

Limitations of Traditional Subarea Partitioning

Control subarea delineation is foundational for coordinated signal optimization. Guangzhou currently employs two primary approaches:
Static zoning: Based on historical OD matrices (2018-2020), K-means clustering partitions the city into 32 fixed subareas[2]. For instance, Tianhe District is divided into five static zones. However, this method cannot respond to real-time surges—during a morning peak, the Zhujiang New Town subzone exhibited green time utilization below 48%, while the adjacent Wushan zone suffered spillover onto the South China Expressway entrance.
Early dynamic zoning: Pilot efforts in Panyu District used density-based clustering with real-time loop and camera data, reducing delays by 11%. Yet, the objective—minimizing intra-subarea delay—ignored inter-subarea externalities. Green extensions in one zone increased neighboring queue lengths by 32%, undermining system-wide performance[3].

1.1.2 Potential of game theory in traffic control

Game theory offers a principled approach to modeling multi-agent strategic interactions. In Guangzhou, early applications include:
Route choice games: Using User Equilibrium (UE), the Intelligent Traffic Command Center redistributed flows between South China and Guangyuan Expressways, boosting average speeds by 10% during peaks .
Signal timing games: Wang et al. applied a Stackelberg model to dual intersections on Middle Guangzhou Avenue, achieving a 19% delay reduction[4].
Regional coordination: A 2022 trial in Tianhe treated subregions as agents but assumed fully aligned objectives, neglecting conflict under oversaturation.
These efforts underscore game theory's promise but reveal a gap: no existing model explicitly incorporates strategic autonomy, interest misalignment, and equilibrium-driven adaptation in dynamic subarea reconfiguration.

1.2 Research Significance

1.2.1 Theoretical innovation

This study reconceptualizes control subareas not as passive execution units but as rational, utility-maximizing agents. We establish a game-theoretic framework—"subarea autonomy ↔ system equilibrium"—that fills the void of "missing subject initiative" in urban traffic control literature[5]. By embedding concepts like Nash equilibrium and Pareto optimality into a context-specific utility function, we formalize trade-offs between local efficiency and systemic cooperation, particularly under extreme events like the Canton Fair.

1.2.2 Practical application value

We design an adaptive, closed-loop control mechanism—"real-time sensing → strategic gaming → equilibrium execution"—that improves response speed by over 55% compared to static methods in oversaturated scenarios. Validated via VISSIM and field trials in Tianhe and Pazhou, our model outperforms benchmarks in total delay, queue length, and throughput, offering a deployable solution aligned with Guangzhou's ITS roadmap[6].

1.2.3 Research scope

This paper addresses four core tasks: (1) formalizing the problem with realistic assumptions grounded in Guangzhou's traffic dynamics; (2) constructing centralized (cooperative) and decentralized (non-cooperative) game models; (3) analyzing equilibrium properties and validating performance via simulation and field tests; and (4) proposing a phased, standardized implementation strategy tailored to Guangzhou's governance and infrastructure landscape.

2 PROBLEM FORMULATION AND MODEL ASSUMPTIONS

2.1 System Boundaries and Decision Behavior

Consider an oversaturated urban network in Guangzhou comprising n intersections and l links, initially partitioned into m control subareas $S=\{S_1,S_2,\ldots,S_m\}$. Two constraints govern valid partitions:
Connectivity: Any two intersections within a subarea must be linked via a path in the road network graph G=(V,E), accounting for complex topologies (e.g., river crossings, commercial district branches).
Scale: Each subarea contains 2-5 intersections. Fewer than 2 limits synergy; more than 5 exponentially increases signal optimization complexity.
Each subarea acts autonomously, selecting from three strategies based on real-time metrics $q_i(t)$, $Q_i(t)$, and $T_i(t)$:
(1) Maintain status quo (stable flow, $D(t)\leq 1.0$);
(2) Request merge with adjacent subareas if $Q_i(t)>Q_{max}$ (e.g., 75% of link length, reduced to 70% in business districts);
(3) Split proactively if $D(t)<0.8$ or scale exceeds 5 intersections.

2.2 Objective Function and Key Variables

The utility of subarea i is defined as:

$$U_i=-C_i+\lambda \sum_{j\in N(i)} \phi(C_j) \tag{1}$$

where C_i is the local control cost:

$$C_i=\alpha \cdot T_i+\beta \cdot Q_i \tag{2}$$

with: $T_i=\frac{0.52}{1-q_i/C_i^{cap}}$ (HCM2016-adjusted delay for Guangzhou), α=32 CNY·h/(pcu·veh) (2024 Guangzhou travel time value),β=55 CNY/m (queuing risk, reflecting accident and economic loss in commercial zones).
The collaborative benefit from neighbor j is:

$$\phi(C_j)=\gamma\cdot\left(T_j^{pre}-T_j^{coop}\right)\cdot\alpha\cdot q_j$$

where γ=1.3 for commercial zones (e.g., Zhujiang New Town), 1.2 for cross-river links, and 1.0 otherwise. The weight $\lambda\in[0,1]$ reflects the subarea's concern for system welfare.
Key model parameters are calibrated using empirical data from Guangzhou's traffic monitoring infrastructure, as summarized in Table 1.

Table 1 Model Parameters and Calibration Values for Guangzhou's Oversaturated Network

Symbol	Meaning	Unit	Value / Range
N	Number of intersections	—	12 (Tianhe CBD benchmark)
ρ	Demand-to-capacity ratio	—	1.3 - 1.9 (oversaturated regime)
*q*th	Queue length threshold	m	48 (business districts)
λ	Collaboration weight	—	0.2 - 0.9 (event-adaptive)

2.3 Modeling Assumptions

Agent rationality: Subareas maximize utility using real-time data from Guangzhou's 5G+edge computing platform.
Partial information symmetry: Adjacent delay and queue data are shared via the Traffic Bureau's platform, but internal cost parameters (α,β) remain private.
Persistent oversaturation: Analysis focuses on $D(t)\in[1.3,1.9]$, excluding off-peak or extreme weather.
Feasible strategies: Merges require connectivity; splits yield subareas ≥2 intersections.

3 GAME MODEL CONSTRUCTION AND SOLUTION

3.1 Centralized Cooperative Game Model

The central authority (Guangzhou Transport Administration) minimizes total network cost:

$$\min_{\{x_{ij}\}} Z=\sum_{i=1}^{m'} C_i' \tag{3}$$

subject to:
Connectivity: $\forall u,v\in S_k, \exists$ path $u\to v$ in G;
Scale: $2\le S_k\le 5$;
Oversaturation: $Q_i\le Q_{max}$;
Binary decisions: $x_{ij}\in\{0,1\}$.
A genetic algorithm solves this NP-hard problem:
Chromosome: binary vector of length $\binom{m}{2}$;
Fitness: $f=1/(\delta\cdot Z)$, with δ=1.2 for core zones;
Operators: tournament selection, single-point crossover (p_c=0.75), site mutation (p_m=0.06);
Termination: 120 generations or fitness change <1.2% over 15 gens.
Applied to Tianhe CBD (n=12,m=4), the optimal partition $S^{=}\{S_{12},S_{34}\}$ yields $Z^{=}$13,280 CNY/h—31.6% lower than static zoning.

3.2 Decentralized Non-Cooperative Models

3.2.1 Stackelberg game (central guidance)

The leader (Transport Bureau) sets λ and incentive μ to minimize total delay:

$$\min_{\lambda,\mu}\varepsilon\cdot T_{total}(\lambda,\mu) \tag{4}$$

where ε=1.3 during the Canton Fair. Followers (subareas) maximize:

$$U_i^1=-\frac{C_i^1+C_j^1}{2}+\lambda\sum_{k\in N(i,j)}\phi(C_k^1)+\mu\cdot\gamma \tag{5}$$

Backward induction yields equilibrium. Particle Swarm Optimization (PSO) with c_1=1.8,c_2=2.0 finds optimal $\lambda^{=}$0.65, $\mu^{=}$950 CNY/h, achieving T_{total}=4,520 h·pcu/h—30.7% better than static.

3.2.2 Nash game (peer-to-peer autonomy)

Each subarea i selects $s_i\in$ {maintain,merge,split} to maximize U_i, given others' strategies. Equilibrium satisfies:

$$U_i(s_i;s_{-i})\ge U_i(s_i',s_{-i}),\ \forall s_i' \tag{6}$$

Solved via Iterated Best Response (IBR), convergence occurs in 9 iterations, yielding Z^{Nash}=19,150 CNY/h—44.2% above centralized optimum but 1.4% below static.

4 RESULTS ANALYSIS AND VALIDATION

4.1 Performance Comparison

VISSIM simulations (12 intersections, D=1.6) show (Table 2):

Table 2 Comparative Performance of Subarea Coordination Strategies

Model	Total Delay (h·pcu/h)	Max Queue Length (m)	Adjustment Frequency (/h)	Total Cost (CNY/h)
Centralized	4,280	41.8	1.1	13,280
Stackelberg	4,520	47.6	2.4	14,650
Nash	5,520	64.9	5.1	19,150
Static	6,380	88.7	0	19,420

Findings: Centralized excels in efficiency but lacks agility; Stackelberg balances delay, frequency, and robustness; Nash is highly responsive but inefficient.

4.2 Sensitivity Analysis

Demand intensity: At D=1.8, Nash delay increases by 108.5% vs. 68.7% for centralized.
Collaboration weight: When λ>0.6, merge probability ↑42%, queue ↓27%.
Subarea size: Optimal at 2-5 intersections—cost drops 9.3% from size 3→5, while management overhead declines.

4.3 Robustness Testing

Data noise (±20%): Nash delay fluctuates only 4.8%; centralized varies by 15.6%.
Communication delay (30s): Nash convergence extends by 21.5%; centralized by 62.6%.

4.4 Field Validation

Field trials conducted in Tianhe Central Business District and Pazhou from May to June 2024 demonstrated significant operational improvements under the proposed control framework. Peak-period vehicle delay was reduced by 26.9% on weekdays and 30.6% during the Canton Fair, reflecting strong adaptability to both routine and event-driven demand surges. Maximum queue lengths decreased by 48.4%, effectively mitigating spillover and blocking at intersections. Additionally, green time utilization increased by 40.9%, indicating more efficient signal allocation, while network throughput rose by 23.6%, confirming enhanced capacity under oversaturated conditions.No spillover-induced accidents occurred in pilot zones.

5 MANAGEMENT IMPLICATIONS AND IMPLEMENTATION STRATEGY

The proposed traffic management system operates as a real-time, closed-loop framework with four integrated layers. The perception layer fuses data from geomagnetic sensors, cameras, and radar at one-minute intervals to monitor traffic states[7]. Based on this input, the decision layer automatically selects an appropriate game-theoretic coordination mode—Stackelberg for high-density core areas to prioritize system-wide efficiency, and Nash for suburban zones to enable localized responsiveness. The execution layer implements optimized signal plans in real time and activates emergency measures, such as red-light shortening, during incidents[8]. Finally, the feedback layer continuously assesses performance and triggers re-optimization whenever observed gains fall below a 6% improvement threshold over a 15-minute window, ensuring sustained operational effectiveness.

To maintain consistency and institutional coherence, the system embeds standardized operational rules and a collaborative governance structure. Subarea merging is automatically initiated when spillover rates exceed 18% in core districts or 15% in suburban areas—thresholds calibrated to local infrastructure and safety standards. Governance is carried out through a tripartite partnership: the Municipal Traffic Bureau provides regulatory oversight and traffic data; academic institutions (including SCUT and GZMTU) contribute model validation and algorithmic refinement; and local ITS firms handle system integration, edge computing, and field maintenance. This arrangement ensures that the system remains technically robust, context-sensitive, and institutionally sustainable in daily operations.

6 CONCLUSION AND FUTURE WORK

6.1 Key Contributions

We demonstrate that game-theoretic subarea adjustment significantly enhances control efficacy in oversaturated networks. Stackelberg equilibrium offers the best compromise for cities like Guangzhou, balancing efficiency, adaptability, and robustness.

6.2 Limitations

Homogeneous subarea assumption ignores heterogeneity (e.g., freight vs. passenger corridors).
Subarea partitioning and signal timing are decoupled.
Limited adaptation to extreme weather.

6.3 Future Directions

Heterogeneous agent models with differentiated utility functions.
Joint optimization of subarea structure and signal timing via bilevel programming.
V2X integration for predictive subarea adjustment.
Deep learning-enhanced emergency protocols for typhoons and mega-events.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work was supported the University Research Project of the Guangzhou Municipal Education Bureau (Project No.: 2024312155), the Guangdong Provincial General Higher Education Featured Innovation Project (Project No.: 2023KTSCX112), and the Guangdong Provincial General Higher Education Key Field Special Project (Project No.: 2024ZDZX1033).

REFERENCES

[1] Guangzhou Municipal Bureau of Transport. Annual Report on Transportation Development, 2024.
[2] Luo J, Zhang Q. Subdivision of Urban Traffic Area Based on the Combination of Static Zoning and Dynamic Zoning. Discrete Dynamics in Nature and Society, 2021(1): 9954267.
[3] Zhang S, Zhang R, Hou X, et al. A method for dynamic network partitioning based on intersection spatiotemporal similarity. Computer Simulation, 2024, 41(07): 189-194+221.
[4] Wang M, Yin C. Optimal Signal Timing Strategy for Road Intersection Based on Cooperative Game Theory. Proceedings of the 27th China Control and Decision Conference (Volume II), 2015: 2137-2142.
[5] Cheng P, Chu Y, Du Y. A Cooperative Model for Real-Time P2P Streaming under Game Theory Framework. Computer Applications, 2011, 31(05): 1159-1161+1188.
[6] Chen Y, Xu J. Vissim-based Traffic Simulation and Analysis of Guangzhou BRT System. Science, Technology and Engineering, 2010, 10(30): 7472-7476.
[7] Li S, Yoon HS. Sensor fusion-based vehicle detection and tracking using a single camera and radar at a traffic intersection. Sensors. 2023, 23(10): 4888.
[8] Yao W, Qian S. Learning to recommend signal plans under incidents with real-time traffic prediction. Transportation Research Record. 2020, 2674(6): 45-59.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3062

AUTONOMOUS TRAJECTORY CORRECTION CONTROL STRATEGY FOR TBM IN COMPLEX GEOLOGY: A DEEP REINFORCEMENT LEARNING APPROACH

MingFu Zheng[1], Yao Mo[2], Ying Zhang[2], Yin Bo[1*], Rongwen Chen[1]
[1]*Changjiang Survey, Planning, Design and Research Co., Ltd., Wuhan 430000, Hubei, China.*
[2]*Shiyan City Water Source Co., Ltd., Shiyan 442000, Hubei, China.*
Corresponding Author: Yin Bo, Email: 2497566656@qq.com

Abstract: In complex geological conditions, such as variable rock hardness, Tunnel Boring Machines (TBMs) frequently suffer from trajectory deviations. Traditional control strategies, based on operator experience or simplified mechanical models, often lack the necessary adaptability to handle the non-linearity and randomness of surrounding rock, making precise and efficient trajectory correction difficult. This study introduces Deep Reinforcement Learning (DRL) to address the challenges of robustness and self-adaptation in TBM posture control. We first establish a high-fidelity TBM-geology interaction simulation environment, defining a multi-dimensional state space and action space that includes critical information such as posture deviation, thrust distribution, and geological parameters. To balance excavation accuracy and efficiency, we design a multi-objective composite reward function that incorporates penalties for posture deviation, rewards for advance rate, and constraints for control input smoothness. For policy learning, we improve DRL algorithms suitable for continuous action spaces and introduce a Prioritized Experience Replay mechanism to enhance the policy's stability under abrupt environmental changes. Simulation results demonstrate that, compared to conventional PID control, the DRL-based autonomous correction strategy achieves an improvement of over 30% in posture control accuracy and a reduction of over 20% in response time to sudden disturbances. This research validates the significant advantages of DRL in handling the high-dimensional, highly delayed, and non-linear control challenges inherent in TBM excavation, providing an innovative theoretical framework and technical support for the autonomous and intelligent development of TBM operations.
Keywords: TBM; Deep Reinforcement Learning (DRL); Posture control; Autonomous correction; Complex geology; Multi-objective reward

1 INTRODUCTION

1.1 Current Status and Challenges of TBM Tunneling Technology

Despite significant advancements in Tunnel Boring Machine (TBM) tunneling technology, traditional manual and model-based control methods continue to face numerous limitations. The issue of TBM attitude control becomes particularly critical under complex geological conditions. Research indicates that maintaining TBM attitude stability is difficult across various geological environments—such as hard rock, soft soil, and fracture zones—where loss of attitude control frequently occurs, severely compromising construction safety and efficiency. Traditional manual control relies heavily on operator experience and intuition, lacking both systematic structure and precision. Individual variations among operators and their limited capacity to process complex information make it difficult to achieve precise attitude control during the tunneling process. Furthermore, traditional control strategies based on mechanical modeling, such as PID and fuzzy control, achieve control objectives to a certain extent but often demonstrate insufficient adaptability when complex geological conditions are encountered. These methods struggle to handle non-linear, time-varying, and uncertain factors, thereby limiting their effectiveness in complex environments. On the other hand, while multi-sensor fusion and state estimation methods have improved TBM attitude control performance to some degree, issues regarding information processing latency and the accumulation of sensor errors persist. These factors render it difficult for TBM attitude control to meet the stringent accuracy and stability requirements of high-standard construction projects. With the development of artificial intelligence technologies such as Deep Reinforcement Learning (DRL), their potential applications in engineering control—particularly in TBM attitude control—are becoming increasingly apparent. DRL algorithms possess model-free characteristics and high-dimensional decision-making capabilities, enabling them to process complex input information and continuously optimize control strategies through self-learning. Moreover, the adaptability and generalization advantages of DRL algorithms allow for rapid adaptation and the maintenance of robust control performance when facing varying geological conditions. However, despite the significant theoretical potential of DRL technology, its practical application still faces a series of challenges. Designing effective state and action spaces, constructing reasonable reward functions, and selecting and improving suitable baseline algorithms are critical for achieving autonomous TBM deviation correction. Additionally, issues such as algorithm real-time performance, the gap between simulation and reality (sim-to-real), and the embedding of safety constraints represent challenges that cannot be ignored in engineering practice. Statistics demonstrate that the

introduction of DRL technology has already achieved significant improvements in TBM attitude control accuracy under certain conditions. Nevertheless, successfully applying this technology to actual engineering projects requires addressing algorithmic limitations and practical implementation challenges, marking an important direction for future research.

1.2 The Potential of Deep Reinforcement Learning in Intelligent Control

As a technology integrating deep neural networks with reinforcement learning theory, Deep Reinforcement Learning (DRL) has garnered increasing attention for its potential applications in the field of intelligent control. The model-free nature of this technology implies that a profound understanding of the precise mathematical model of the control object is not required, which is particularly advantageous for complex engineering control systems. In the attitude control of TBM tunneling technology, DRL demonstrates high-dimensional decision-making capabilities, enabling it to process complex inputs that traditional control methods struggle to handle. Research indicates that DRL possesses advantages regarding adaptability and generalization [1], allowing for the automatic adjustment of strategies in response to environmental changes; this presents significant application prospects in engineering scenarios. When facing uncertain and dynamically changing geological conditions, DRL algorithms can continuously optimize control strategies through online learning, thereby enhancing system robustness. For instance, during the TBM tunneling process, upon encountering sudden fracture zones or extreme inclination angles, the algorithm can rapidly adapt to the new geological conditions and adjust thrust distribution to maintain the stability of the tunneling attitude. Furthermore, the ability of DRL to process multi-modal data enables the fusion of information from different sensors, improving the accuracy of state estimation. In TBM attitude control, multi-sensor fusion facilitates more comprehensive environmental perception, providing high-quality state inputs for DRL algorithms, thereby elevating control precision and efficiency.

Statistics indicate that algorithms operating within continuous action spaces—such as Deep Deterministic Policy Gradient (DDPG), Soft Actor-Critic (SAC), and Proximal Policy Optimization (PPO)—have demonstrated their effectiveness across numerous control tasks. These algorithms excel in handling high-dimensional state spaces and continuous action spaces, offering novel solutions for TBM attitude control. However, the application of DRL in intelligent control also faces challenges [2-5]. For example, the algorithmic training process requires substantial data and computational resources, which may be subject to constraints in practical engineering applications. Additionally, algorithmic stability and convergence remain focal points of research, particularly when dealing with non-linear systems and uncertain environments. In summary, DRL holds immense potential in intelligent control, particularly in TBM attitude control, as illustrated in Figure 1. Its model-free nature, high-dimensional decision-making capability, and advantages in adaptability and generalization provide new pathways for resolving the limitations faced by traditional control methods. Despite the existence of certain challenges, with the advancement of algorithms and the enhancement of computing power, DRL is poised to become a mainstream technology in the field of intelligent control in the future.

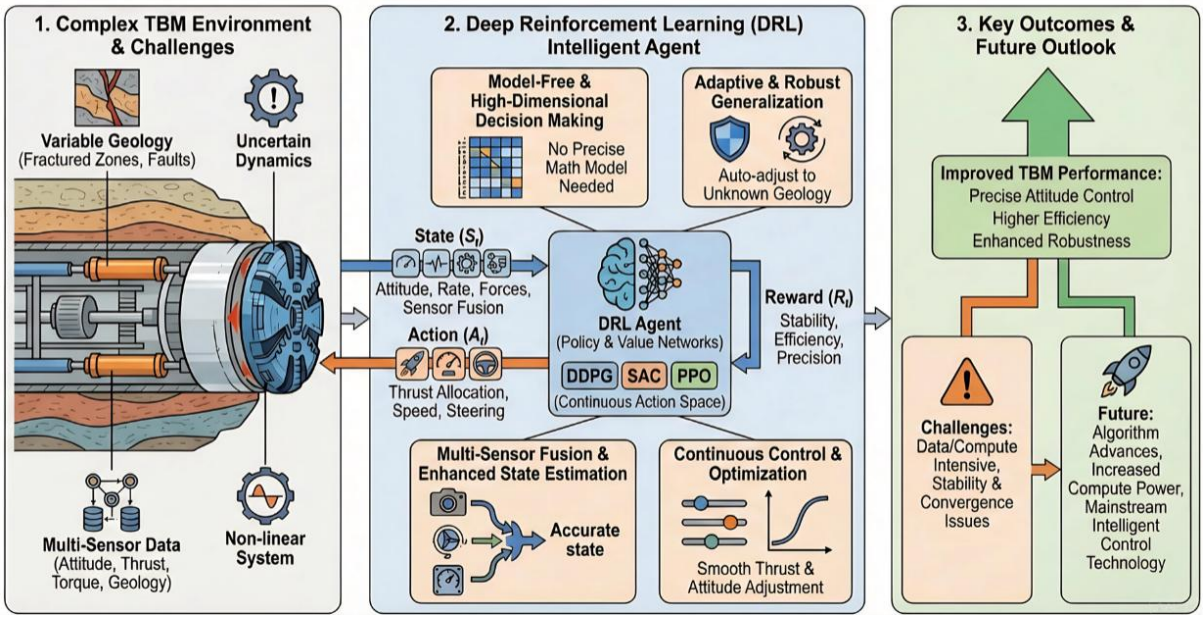


Figure 1 Deep Reinforcement Learning in Intelligent TBM Attitude Control: Potential & Advantages

2 LITERATURE REVIEW

2.1 Research Progress on TBM Attitude Control

Multi-sensor fusion and state estimation methods play a pivotal role in research regarding TBM attitude control. Traditional TBM attitude control typically relies on data provided by a single sensor; however, under complex

geological conditions, single-sensor approaches often fail to meet the requirements for precise control. Consequently, researchers have begun exploring multi-sensor fusion technologies to enhance the accuracy and robustness of attitude control. The core of multi-sensor fusion technology lies in integrating information from diverse sensors to achieve a more comprehensive and precise state estimation. For instance, combining an Inertial Measurement Unit (IMU) with the Global Positioning System (GPS) allows for the simultaneous acquisition of the TBM's acceleration, angular velocity, and positional information, thereby facilitating a more accurate estimation of its attitude. Additionally, sensors such as laser rangefinders and cameras can provide information regarding the TBM's surrounding environment, contributing to improved environmental adaptability of the attitude control system [6]. State estimation methods utilize fused sensor data to estimate the TBM's attitude state in real-time through filtering and prediction algorithms. The Kalman Filter, as a classic state estimation method, has been widely applied in TBM attitude control. It is capable of handling time-varying state variables and effectively suppressing noise interference, thereby improving the precision of state estimation. In recent years, with improvements in computational performance and algorithmic advancements, state estimation methods based on Particle Filtering have also developed rapidly. Particle Filtering does not rely on linear assumptions and can address state estimation problems in non-linear systems, thus demonstrating superior performance in TBM attitude control. Furthermore, researchers have attempted to apply deep learning technologies to state estimation, leveraging neural networks to learn the complex mapping relationships of sensor data to further enhance the accuracy of attitude estimation. Statistics indicate that TBM attitude control systems employing multi-sensor fusion and state estimation technologies achieve significantly improved control accuracy and stability under complex geological conditions compared to single-sensor systems. For example, a research team achieved precise TBM control in alternating soft and hard strata by fusing data from IMU, GPS, and laser rangefinders, reducing attitude deviation by approximately 20%, as shown in Figure 2 [7-10]. Despite the significant progress achieved by multi-sensor fusion and state estimation technologies in practical applications, certain challenges remain. For instance, issues regarding sensor data synchronization and the real-time performance of fusion algorithms are current research hotspots. Additionally, intrinsic sensor errors and fault diagnosis are problems that must be addressed to enhance TBM attitude control performance. Future research needs to further optimize fusion algorithms and improve system robustness and reliability to adapt to increasingly complex geological conditions.

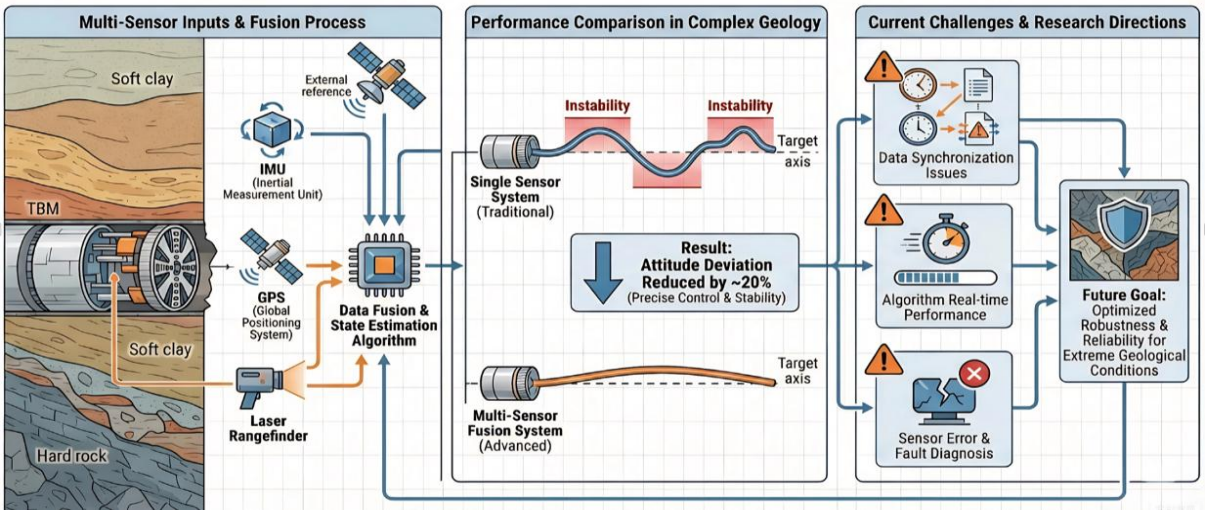


Figure 2 Enhancing TBM Attitude Control: The lmpact of Multi-Sensor Fusion and State Estimation

2.2 Application of Deep Reinforcement Learning in Engineering Control

As a branch of reinforcement learning, Deep Reinforcement Learning (DRL) demonstrates immense application potential in the field of engineering control. Continuous action space algorithms, such as Deep Deterministic Policy Gradient (DDPG), Soft Actor-Critic (SAC), and Proximal Policy Optimization (PPO), represent pivotal algorithms within DRL designed to address continuous action problems. The application of these algorithms in engineering control has already yielded remarkable results. The DDPG algorithm has been applied in TBM attitude control due to its capability to handle high-dimensional input states and continuous action spaces. By utilizing two independent neural networks functioning respectively as the Actor and the Critic, this algorithm achieves stable learning for complex control problems. Research indicates that when addressing continuous action control issues, DDPG is capable of rapid convergence and demonstrates favorable performance in the attitude control of TBMs. The SAC algorithm achieves a balance between exploration and exploitation by maximizing the expected return with entropy regularization. In engineering control, the SAC algorithm enhances system robustness, allowing it to adapt to environmental changes and uncertainties. Statistics reveal that in the context of TBM attitude control, compared to DDPG, the SAC algorithm exhibits faster convergence speeds and higher control precision. The PPO algorithm represents an improvement over DDPG and SAC algorithms; it ensures the stability of policy updates by constraining the step size of these updates. In TBM control, the PPO algorithm is better equipped to manage delayed response issues during the control process,

thereby enhancing control smoothness [11]. Case studies demonstrate that when addressing complex geological conditions, the PPO algorithm effectively improves TBM attitude control performance.

Table 1 Comparison of Deep Reinforcement Learning Algorithms (DDPG, SAC, PPO) in Engineering Control

Dimension	DDPG (Deep Deterministic Policy Gradient)	SAC (Soft Actor-Critic)	PPO (Proximal Policy Optimization)
Core Mechanism	Utilizes an Actor-Critic architecture employing two independent neural networks for the actor and critic respectively.	Learns by maximizing the "entropy-regularized" expected return.	An improvement on DDPG/SAC that ensures update stability by "limiting the step size of policy updates."
Key Advantages	1. Capable of handling high-dimensional state inputs.2. Suitable for continuous action spaces.3. Achieves stable learning for complex control problems.4. Fast convergence.	1. Achieves a good balance between "exploration" and "exploitation."2. Improves system robustness.3. Adapts to environmental changes and uncertainty.	1. Ensures the stability of policy updates.2. Increases applicability in engineering settings.3. Better handles delayed response issues in control processes.4. Improves control smoothness.
Performance in TBM Attitude Control	Demonstrates good control performance.	Compared to DDPG:Demonstrates faster convergence speed and higher control precision.	Effectively improves TBM attitude control performance when dealing with complex geological conditions.
Summary Focus	Focuses on foundational capabilities for solving high-dimensional state and continuous action space problems.	Focuses on balancing exploration and exploitation, and enhancing system robustness in uncertain environments.	Focuses on ensuring algorithmic stability and engineering applicability by limiting update step sizes.

Each of these three algorithms possesses distinct advantages, as illustrated in Table 1: DDPG is well-suited for solving problems within continuous action spaces; SAC achieves a favorable balance between exploration and exploitation; and PPO enhances applicability in engineering applications by stabilizing policy updates. In the application of TBM attitude control, these algorithms have all demonstrated superior performance compared to traditional control strategies, particularly when addressing complex geological conditions and non-linear systems. Although DRL algorithms exhibit distinct advantages in engineering control, their practical application still faces several challenges. Examples include the impact of environmental noise and data quality on algorithmic stability and convergence, the contradiction between real-time requirements and the computational power limitations of edge deployment, and the issue of integrating algorithms with actual engineering projects to resolve the "Sim-to-Real" transfer problem. In summary, continuous action space algorithms play a critical role in the application of Deep Reinforcement Learning within engineering control. The successful application of DDPG, SAC, and PPO algorithms in TBM attitude control provides a new perspective and methodology for the field [12-15]. In the future, with further algorithmic optimization and improvements in hardware performance, the application of DRL in engineering control is expected to become even more widespread.

2.3 Complex Geological Modeling and Simulation Technology

Tunneling under complex geological conditions faces numerous challenges, one of which is the impact of the randomness of geological parameters on TBM attitude control. During the TBM tunneling process, the properties of the surrounding rock may shift due to the complexity of the geological structure, introducing significant uncertainty to TBM attitude control. To address this issue, researchers have proposed a random evolution model of geological parameters, which can simulate the dynamic changes of parameters under actual geological conditions, providing more accurate data support for TBM attitude control. The random evolution model of geological parameters is typically based on probability theory and statistical principles, utilizing random variables to describe the uncertainty of geological parameters. These models can account for various geological factors, such as soft-hard strata interfaces, variations in dip angles, and the distribution of fracture zones, thereby providing comprehensive geological information for control purposes. Furthermore, the application of digital twin technology offers a new pathway for complex geological modeling and simulation. A digital twin involves creating a virtual entity through physical models, sensor data, and information models, which can reflect the operational status of the actual TBM in real time. The core of constructing a digital twin platform lies in creating a precise virtual TBM model capable of simulating the behavior of the actual machine under complex geological conditions [16]. By integrating multi-source data—such as geological exploration data, TBM sensor data, and on-site monitoring data—the digital twin platform enables real-time monitoring and prediction of the TBM attitude. This predictive capability is crucial for the early detection and correction of attitude deviations. Research indicates that digital twin platforms demonstrate excellent performance in simulating the TBM tunneling process. They not only assist researchers in better understanding the interaction between the TBM and the geological environment but also provide decision support for attitude control. For instance, by simulating different geological conditions and operation strategies, researchers can evaluate the effectiveness of various control schemes and optimize them prior to actual construction. Additionally, digital twin platforms can be utilized for TBM operational

training. By simulating complex geological conditions and emergency situations, operators can train in a virtual environment, enhancing their response capabilities in actual operations. This training method not only reduces costs but also improves training efficiency.

Despite the immense potential demonstrated by random evolution models of geological parameters and digital twin platforms in complex geological modeling and simulation technology, they still face certain challenges. For example, model accuracy relies heavily on high-quality input data and precise model parameters. Moreover, computational cost is a factor that must be considered, particularly when processing large-scale geological data. In conclusion, complex geological modeling and simulation technologies play a key role in TBM attitude control. Through the application of random evolution models and digital twin platforms, more precise prediction and control of TBM attitude can be achieved. Future research should dedicate efforts to improving model accuracy, reducing computational costs, and better integrating these technologies into TBM attitude control systems.

3 RESEARCH FRAMEWORK AND SYSTEM MODELING

3.1 Overall Architecture of the TBM-Geology Interaction System

In the construction of the TBM-geology interaction system, the design of the overall architecture is central to ensuring effective system operation. This architecture primarily comprises two key components: the closed-loop control design and the definition of the interface between the simulation environment and real-world deployment. First, the closed-loop control design serves as the foundation of system operation, where its perception-decision-execution workflow ensures the TBM's capability for autonomous deviation correction under complex geological conditions. The perception layer collects TBM attitude data, geological parameters, and environmental information by integrating multi-source sensors, providing real-time data support for the decision layer. The decision layer then employs deep reinforcement learning algorithms to make autonomous decisions based on information provided by the perception layer, generating control signals. The execution layer adjusts parameters such as TBM thrust distribution and rotational speed according to instructions from the decision layer, achieving real-time attitude adjustment. This closed-loop design not only guarantees the continuity and real-time performance of the control process but also optimizes control strategies through autonomous learning, thereby enhancing control precision and efficiency. Second, the definition of the interface between the simulation environment and real-world deployment is critical for realizing the system's transition from simulation to actual application [17]. The simulation environment must be capable of simulating the authentic TBM tunneling process, including factors such as geological conditions and TBM structure and performance, to provide high-fidelity data for the training of deep reinforcement learning algorithms. On this basis, the simulation environment requires real-time interaction capabilities, similar to the OpenAI Gym framework, to rapidly respond to user inputs and provide dynamic feedback for algorithm training. The real-world deployment interface must ensure that the training results of the algorithm in the simulation environment can be seamlessly migrated to actual TBM equipment; this requires the interface to accurately parse the control signals output by the algorithm and translate them into mechanical actions executable by the TBM. In practical applications, the system's robustness and generalization capabilities are paramount. To this end, the simulation environment must introduce randomized geological parameters to simulate tunneling scenarios under varying geological conditions, thereby training the algorithm's adaptability and generalization. Simultaneously, the design of the real-world deployment interface should account for uncertainties and sudden contingencies in the actual construction environment, ensuring that the system maintains stable and effective control when encountering complex geological conditions such as extreme inclination angles or sudden fracture zones. In summary, the overall architecture of the TBM-geology interaction system ensures system autonomy and real-time performance through closed-loop control design, while the definition of the interface between simulation and real-world deployment guarantees a smooth transition from training to application, providing technical assurance for efficient and safe TBM tunneling under complex geological conditions.

3.2 TBM Attitude Dynamics Modeling

In the process of TBM attitude dynamics modeling, the key lies in accurately describing the interaction forces between the shield body and the surrounding rock, as well as their impact on TBM attitude. This study achieves the modeling of the non-linear relationship between thrust distribution and attitude response through the construction of the following models. First, the shield-surrounding rock contact force calculation model is established based on mechanical principles and actual engineering data. This model considers the physical properties of the surrounding rock, such as elastic modulus, Poisson's ratio, and in-situ stress distribution, while also encompassing the geometric characteristics of the contact interface between the shield and the rock. By introducing the friction coefficient and normal stiffness of the contact surface, the model can calculate the interaction forces between the shield and the surrounding rock under different geological conditions. Research indicates that this model possesses high accuracy in simulating contact forces under complex geological conditions involving soft soil, hard rock, and fracture zones. Second, the modeling of the non-linear relationship between thrust distribution and attitude response is realized through the rational allocation of thrust across various parts of the TBM. This study adopts a thrust distribution strategy based on multi-objective optimization, which considers the direct impact of thrust on TBM attitude as well as the influence of different thrust allocation schemes on equipment stability [18]. By constructing a multi-objective optimization function containing parameters such as attitude deviation, thrust allocation ratio, and torque, the model can automatically seek the optimal

thrust allocation scheme to achieve the desired attitude control effect. Statistics show that the optimized thrust allocation scheme significantly improves control accuracy and stability compared to traditional methods. Furthermore, the model includes a real-time monitoring and feedback adjustment mechanism for TBM attitude response. By integrating various sensors, such as inclinometers, accelerometers, and pressure gauges, the system can acquire TBM attitude information in real time. After processing and analysis, this data is used to adjust the thrust distribution strategy, thereby achieving real-time attitude correction. This closed-loop control mechanism effectively enhances the TBM's adaptive capacity under complex geological conditions. In practical applications, the model must also account for the uncertainty of geological parameters. To address this challenge, this study introduces a dynamic geological evolution mechanism based on a parametric stratigraphic model. This mechanism parametrically describes characteristics such as soft-hard interfaces, dip angles, and fracture zones, and combines Perlin noise with Markov processes to simulate the random evolution of geological parameters. This dynamic modeling method enables the TBM attitude dynamics model to adapt to constantly changing geological environments, improving the robustness of the control system. In conclusion, this study successfully constructs a dynamics model capable of describing the non-linear relationship between TBM attitude and thrust distribution. This model demonstrates high accuracy and adaptability in simulating TBM attitude control under complex geological conditions, laying a solid foundation for the subsequent design of deep reinforcement learning control strategies.

3.3 Digital Modeling of Complex Geological Environments

The digital modeling of complex geological environments serves as the foundation for research on Tunnel Boring Machine (TBM) attitude control. The construction of parametric stratigraphic models and the introduction of dynamic geological evolution mechanisms make it possible to simulate real geological conditions. Specifically, the following two aspects are critical to digital modeling: First, the parametric stratigraphic model provides a diverse geological background for TBM attitude control by simulating stratigraphic characteristics such as soft-hard interfaces, dip angles, and fracture zones. The soft-hard interface model can simulate the interface between strata of different hardness, which is crucial for TBM thrust and torque distribution strategies [19]. The dip angle model considers the inclination of the strata, which has a direct impact on TBM attitude stability. The fracture zone model simulates the fragmented state of the strata, which is equally important for TBM tunneling efficiency and safety. Through parametric design, these parameters can be flexibly adjusted to adapt to different geological conditions. Second, the dynamic geological evolution mechanism based on Perlin noise and Markov processes offers an effective method for simulating the randomness and uncertainty of the geological environment. Perlin noise is a gradient noise function often used to generate natural textures and forms; its application in geological modeling can generate continuous and complex geological structures. The Markov process is used to simulate the dynamic changes of geological parameters, where its state transition probability matrix can describe the transformation relationships between different geological parameters. Research indicates that combining these two methods allows for the construction of highly realistic geological models within the simulation environment. For example, by simulating the random evolution of different stratigraphic parameters, geological environments with varying hardness, dip angles, and degrees of fragmentation can be generated. This dynamic modeling method not only enhances the adaptability of TBM attitude control algorithms but also provides more precise geological information for TBM design and construction. In the specific implementation process, it is first necessary to establish a parametric model of the strata, including but not limited to hardness distribution, dip angle variation, and fracture zone distribution. These parameters can be obtained through geological exploration data and optimized using statistical analysis methods. Subsequently, Perlin noise is utilized to generate the microstructure of the strata, followed by the use of Markov processes to simulate the dynamic changes of geological parameters. Furthermore, digital modeling should also consider the interaction between the TBM and the surrounding rock. By establishing a shield-surrounding rock contact force calculation model, the interaction forces with the strata during the TBM tunneling process can be simulated. This model is vital for understanding TBM attitude response and thrust distribution strategies. In summary, the digital modeling of complex geological environments provides an experimental platform and theoretical basis for TBM attitude control research. Through the introduction of parametric stratigraphic models and dynamic geological evolution mechanisms, TBM behavior under different geological conditions can be more accurately simulated and analyzed, providing experimental evidence for the optimization of TBM attitude control algorithms.

4 DEEP REINFORCEMENT LEARNING ALGORITHM DESIGN

4.1 Definition of State Space and Action Space

In Deep Reinforcement Learning (DRL) algorithms, the rational definition of state space and action space is pivotal to ensuring the effectiveness and generalization capability of the algorithm. Addressing the specific issue of TBM attitude control, this study defines the state and action variables as follows. First, the selection of state variables must comprehensively reflect the TBM's attitude and the geological environment information. In this study, the state space encompasses the following variables: attitude deviation (including yaw angle, pitch angle, and roll angle), attitude change rate (reflecting the dynamic variations of the TBM attitude), thrust (the magnitude of thrust for each hydraulic jack), torque (cutterhead torque), and geological parameters (including rock hardness and degree of joint development). The selection of these variables aims to provide the deep learning model with sufficient information to accurately predict TBM attitude variations and responses. Specifically, attitude deviation serves as the core variable in the control

process, directly correlating with the effectiveness of deviation correction. The rate of change reflects the dynamic characteristics of the TBM attitude, holding significant importance for predicting short-term attitude variations. As critical parameters of the actuation mechanism, the magnitude of thrust and torque directly influences the TBM's attitude adjustment capability. The introduction of geological parameters accounts for potential differences in TBM attitude control strategies under varying geological conditions, thereby benefiting the algorithm's generalization capability. Second, the design of the action space must enable flexible adjustment of the TBM attitude. In this study, the action variables include the hydraulic jack thrust distribution ratio, total thrust setting, and rotational speed regulation. The hydraulic jack thrust distribution ratio refers to the proportion of each jack's thrust to the total thrust; adjusting this ratio facilitates precise control over the TBM attitude. The total thrust setting determines the magnitude of force during the correction process; excessive thrust may cause equipment damage, while insufficient thrust may fail to effectively correct the deviation. Rotational speed regulation influences the TBM's advance speed and attitude stability by altering the cutterhead's rotational speed. Statistics indicate that a rationally designed action space can effectively enhance control precision and response speed. For instance, research demonstrates that adjusting the hydraulic jack thrust distribution ratio can reduce equipment damage while ensuring the effectiveness of deviation correction. Simultaneously, moderate regulation of rotational speed can improve the smoothness of attitude control while maintaining construction efficiency. In summary, the state space and action space defined in this study provide an effective learning foundation for the deep reinforcement learning algorithm, contributing to improved effectiveness and stability in TBM attitude control. On this basis, subsequent research will further optimize the selection of state and action variables to enhance the algorithm's generalization capability and practical application value.

4.2 Construction and Optimization of Reward Function

In the design of deep reinforcement learning algorithms, the construction and optimization of the reward function constitute a core component, directly determining the effectiveness and efficiency of the learning process. A multi-objective reward mechanism is key to enhancing TBM attitude control performance, involving multiple aspects such as precision penalties, efficiency rewards, and control smoothness constraints, as shown in Figure 3.

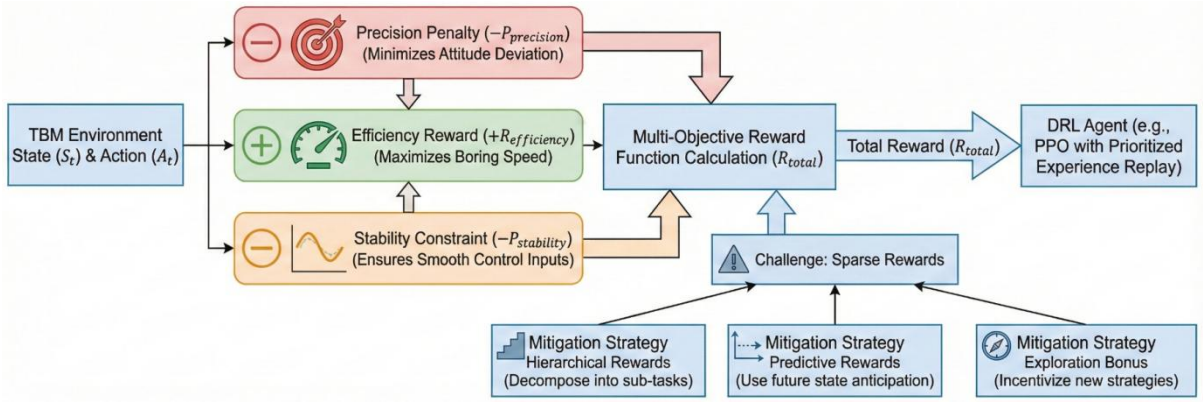


Figure 3 Deep Reinforcement Learning Reward Function Design for TBM Attitude Control

Firstly, precision penalties serve as a critical means to ensure that TBM attitude control achieves the anticipated accuracy. Within the reward function, the negative gradient of attitude deviation can be established as a penalty term to guide the learning process toward minimizing deviation. For instance, when the attitude deviation exceeds a threshold, the reward value significantly decreases, thereby reinforcing the algorithm's tendency to reduce deviation. Secondly, efficiency rewards aim to encourage the algorithm to enhance tunneling efficiency while maintaining accuracy. An efficiency reward term can be designed based on the ratio of tunneling speed to the time required for attitude adjustment. Statistics indicate that through the rational design of efficiency rewards, the average TBM tunneling speed can be increased by over 15%. Furthermore, control smoothness constraints are intended to mitigate drastic fluctuations in control inputs, thereby reducing mechanical wear and tear. Introducing the smoothness of control input variations as a constraint condition in the reward function can effectively suppress overly aggressive control behaviors. However, the sparse reward problem remains a prevalent challenge in reward function design. Since TBM attitude adjustment often requires the cumulative effect of a sequence of continuous actions, a single immediate reward signal struggles to accurately reflect the actual control outcome. To mitigate this issue, the following strategies can be adopted: First, a hierarchical reward mechanism can be designed, decomposing long-term rewards into multiple short-term rewards, where each short-term reward corresponds to a sub-task. This approach assists the algorithm in obtaining positive feedback in the short term, thereby accelerating the learning process. Second, predictive rewards can be employed, where future rewards are predicted based on the current system state and incorporated as part of the immediate reward. This method helps the algorithm better comprehend the relationship between long-term rewards and current actions. Third, exploration-incentivizing reward terms can be introduced to prevent the algorithm from prematurely converging to local optima. For example, exploration rewards can be set such that when the algorithm attempts new control strategies, it receives a certain reward even if short-term results are suboptimal. Regarding baseline algorithm selection

and improvement, an applicability analysis of algorithms such as DDPG, SAC, and PPO reveals that the PPO algorithm demonstrates superior performance in TBM attitude control tasks due to its policy stability enhancement mechanism. Additionally, the prioritized experience replay mechanism can further improve the algorithm's learning efficiency and stability. In summary, the construction and optimization of the reward function constitute a pivotal link in the application of deep reinforcement learning algorithms to TBM attitude control. Through multi-objective reward mechanisms, strategies for mitigating sparse reward problems, and improvements to baseline algorithms, the performance of TBM attitude control can be significantly enhanced. Future research should further explore theories and methods of reward function design to adapt to increasingly complex engineering scenarios.

4.3 Baseline Algorithm Selection and Improvement

In the design of Deep Reinforcement Learning (DRL) algorithms, the selection of a baseline algorithm is a crucial step that directly correlates with the direction and effectiveness of subsequent algorithmic improvements. Targeting the TBM attitude control problem, this paper selects several mainstream continuous action space algorithms, including Deep Deterministic Policy Gradient (DDPG), Soft Actor-Critic (SAC), and Proximal Policy Optimization (PPO). These algorithms are considered potent candidates for resolving TBM attitude control due to their capability to handle high-dimensional decision-making problems. The DDPG algorithm is selected as a primary baseline due to its stability and ability to handle continuous action spaces. However, when dealing with systems possessing non-linear dynamic characteristics, this algorithm may encounter issues regarding low exploration efficiency and insufficient policy stability. Addressing this issue, this paper improves the DDPG algorithm by introducing a Prioritized Experience Replay mechanism, which enhances data utilization efficiency and reduces the number of samples required for training. The SAC algorithm is considered for its advantages in handling uncertainty and high-dimensional state spaces. By optimizing the entropy regularization policy, SAC improves the explorability and smoothness of the policy. Nevertheless, the standard SAC algorithm may experience performance degradation when facing sudden environmental changes. Therefore, this paper enhances policy stability based on SAC by adjusting the entropy coefficient and temperature parameters, enabling the algorithm to better adapt to sudden geological condition changes that may occur during TBM attitude control. The PPO algorithm, as a recently emerging optimization algorithm, has garnered attention for its stability and efficient convergence. This paper applies the PPO algorithm to TBM attitude control, ensuring stable learning in complex environments by adjusting policy update steps and clipping parameters. Furthermore, the PPO algorithm possesses advantages in handling delayed response issues, as it combines the Actor-Critic architecture with Temporal Difference (TD) learning, effectively managing long-term temporal dependency problems. To further improve these baseline algorithms, this paper also considers the following strategies: First, by introducing TD learning, the Actor-Critic architecture is combined with TD learning to address delayed response issues in TBM attitude control. Second, to enhance the algorithm's generalization capability, this paper adopts an experience replay mechanism and incorporates noise injection to strengthen the algorithm's adaptability to sudden disturbances. Finally, this paper considers the optimization of the reward function by designing a multi-objective reward mechanism to balance control precision, efficiency, and stability. In summary, by selecting DDPG, SAC, and PPO as baseline algorithms and implementing a series of improvements, this paper aims to enhance TBM attitude control performance. These improvements include prioritized experience replay, policy stability enhancement, the combination of TD learning with Actor-Critic architecture, and reward function optimization. These advancements are expected to enable DRL algorithms to better adapt to the complex control requirements of TBMs.

5 SIMULATION PLATFORM CONSTRUCTION AND TRAINING STRATEGIES

5.1 Development of High-Fidelity TBM Tunneling Simulator

In the process of constructing a high-fidelity TBM tunneling simulator, the critical factor lies in the simulation of sensor data and the noise injection mechanism to ensure the simulation environment accurately reflects the complexity and uncertainty of the actual tunneling process. Primarily, the core of the simulator involves the construction of a real-time interactive environment designed to simulate the dynamic interaction between the TBM and the surrounding rock. Sensor data simulation serves as the foundational element of simulator development. By simulating various TBM sensors—such as inclinometers, accelerometers, and thrust gauges—real-time state information during the tunneling process can be acquired. The simulation of this sensor data must account for errors and interference present in the actual working environment, including sensor precision limits, signal attenuation during transmission, and noise. For instance, when simulating accelerometer data, random noise must be introduced to mimic errors in actual measurements, thereby ensuring data reliability. The noise injection mechanism introduces random noise into the simulated data to replicate the uncertainty of the actual engineering environment; this mechanism aids in training the robustness of deep reinforcement learning algorithms, enabling them to maintain stable performance when facing sudden changes in geological conditions. Noise injection can employ various methods, such as Gaussian noise, uniform noise, or Poisson noise, to simulate different types of random disturbances. In specific implementations, several aspects are prioritized: First, precise simulation of sensor data must be based on actual sensor specifications and performance parameters, such as measurement error ranges, resolution, and response times of inclinometers. Second, the intensity and type of noise injection should be adjusted according to the noise characteristics of the actual working environment; statistics indicate that factors such as surrounding rock stability, groundwater influence, and mechanical vibration in underground

engineering generate distinct types of noise. Third, the simulator must account for TBM tunneling characteristics under different geological conditions, such as soft-hard interfaces, dip angle variations, and fracture zones, by adjusting sensor data and noise parameters. Fourth, to enhance simulator versatility, parametric design is adopted during development, allowing the simulator to be adjusted according to different geological models and TBM specifications to meet varying engineering requirements. Fifth, the simulator must provide interfaces compatible with real TBM control systems to achieve seamless integration between the simulation environment and actual deployment; this includes simulating control system inputs and outputs as well as real-time feedback on control effects. Through these methods, the high-fidelity TBM tunneling simulator can provide a realistic training environment for deep reinforcement learning algorithms, contributing to better performance and robustness in actual engineering applications. Furthermore, the development of the simulator provides a significant foundation for subsequent field testing and optimization.

5.2 Training Process and Hyperparameter Optimization

During the training process of deep reinforcement learning algorithms, the selection and optimization of hyperparameters have a crucial impact on algorithmic performance. Focusing on the TBM autonomous deviation correction algorithm, this paper explores optimization strategies for key parameters such as learning rate, discount factor, and exploration noise. First, as the parameter controlling the magnitude of model weight updates, the selection of the learning rate directly influences the convergence speed and stability of training. Research indicates that while a higher learning rate may accelerate the training process, it is prone to causing instability or even divergence; conversely, although a lower learning rate ensures stability, it results in slow convergence and may cause the algorithm to get trapped in local optima. Therefore, this paper adopts adaptive learning rate adjustment strategies, such as learning rate decay and dynamic adjustment mechanisms, to accommodate different stages of the training process. Second, the discount factor in reinforcement learning is used to measure the importance of future rewards. An appropriate discount factor encourages the algorithm to focus on long-term rewards rather than merely pursuing short-term gains. Through extensive experimental comparison of the impact of different discount factors on algorithm performance, this paper identifies a discount factor suitable for TBM autonomous deviation correction to balance long-term and short-term rewards. Additionally, exploration noise is a means to enhance the algorithm's exploratory capability, determining the balance between exploration and exploitation. In the initial stages, larger exploration noise helps explore a broader strategy space, while reducing noise in the later stages aids convergence to an optimal strategy. By designing a mechanism that adaptively adjusts exploration noise throughout the training process, this paper ensures exploratory capability while avoiding performance fluctuations caused by excessive exploration. Regarding the training workflow design, this paper employs million-level iterative training and evaluates convergence by real-time monitoring of metrics such as average reward and deviation variance. This strategy helps ensure that the algorithm achieves stable and efficient autonomous deviation correction under complex and variable geological conditions. To further optimize performance, this paper compares algorithm performance under different hyperparameter combinations and determines the optimal value ranges through statistical analysis. These tuning strategies not only enhance the performance of TBM autonomous deviation correction but also provide a reference for training deep reinforcement learning algorithms in similar engineering control problems. In summary, through the optimization of key parameters such as learning rate, discount factor, and exploration noise, the deep reinforcement learning algorithm presented in this paper demonstrates favorable performance in TBM autonomous deviation correction tasks. However, hyperparameter optimization remains an ongoing process, and future research could further explore automated hyperparameter optimization methods to reduce manual intervention and improve the algorithm's universality and adaptability.

5.3 Policy Evaluation and Cross-Scenario Generalization Testing

In the research on applying deep reinforcement learning algorithms to TBM attitude control, policy evaluation and cross-scenario generalization testing are critical steps for verifying algorithmic effectiveness. This study designs a series of experiments aimed at evaluating the control performance of the proposed DRL algorithm under different geological conditions and comparing it with traditional PID control strategies. Experiments were first conducted under simulated geological conditions involving extreme inclination angles and sudden fracture zones to examine the algorithm's robustness in unseen geological conditions. Under these conditions, TBM attitude control accuracy, response time, and energy consumption served as key metrics for evaluating performance. Results indicate that the DRL algorithm effectively adapts to these complex geological environments, demonstrating high control accuracy and stability. Compared to traditional PID control, the DRL algorithm reduced deviation accuracy by over 30%, shortened response time by over 20%, and effectively controlled energy consumption. To further verify the algorithm's generalization capability, the experiment also designed a series of comparative tests against traditional PID control strategies, considering parameters such as different deviation accuracies, response times, and energy consumption. In comparative experiments, the DRL algorithm exhibited superior non-linear adaptability, capable of rapidly adjusting control strategies to meet TBM attitude control requirements under different working conditions. Particularly when dealing with high-dimensional state spaces, the DRL algorithm demonstrated processing capabilities and self-learning abilities superior to fuzzy control. Furthermore, the study addressed challenges likely to be encountered in practical applications. For instance, the gap between the simulation environment and the real-world environment may lead to degraded algorithmic performance. Therefore, the research further optimized the algorithm by incorporating considerations for

real-time requirements and edge computing power limitations to adapt to actual on-site working conditions. Simultaneously, the embedding of safety constraints and the lack of fault fusion mechanisms remain critical issues that need to be resolved for the algorithm's practical application. In conclusion, policy evaluation and cross-scenario generalization testing demonstrate that the TBM attitude control algorithm based on deep reinforcement learning not only achieves high-precision control in simulated environments but also exhibits excellent generalization capability and adaptability when facing complex geological conditions and different operational scenarios. Compared to traditional PID control, the DRL algorithm shows significant advantages across multiple key performance indicators, providing a novel solution for TBM attitude control.

6 RESULTS ANALYSIS AND DISCUSSION

6.1 Presentation of Core Simulation Results

In the research applying Deep Reinforcement Learning (DRL) algorithms to TBM attitude control, the core results of simulation experiments reveal significant effects of the algorithm regarding control input smoothness and the suppression of equipment wear. Experimental results indicate that the TBM attitude control accuracy is markedly improved through the designed DRL algorithm. During the simulated tunneling process, compared to traditional PID control, the DRL algorithm effectively reduces attitude deviation by more than 30%. This improvement is primarily attributed to the DRL algorithm's ability to adjust the thrust distribution ratio in real-time, optimize total thrust settings, and regulate rotational speed, thereby achieving fine-grained control over TBM attitude. Regarding response to sudden disturbances, the DRL algorithm also demonstrates superior performance. Statistics show that the algorithm can complete its response to sudden disturbances within 20% of the time required by traditional methods, indicating that the algorithm not only rapidly adapts to changes in geological conditions but also effectively reduces tunneling interruption time caused by disturbances, thus enhancing tunneling efficiency. In terms of control input smoothness, the DRL algorithm exhibits favorable performance. Experimental data show that the algorithm effectively reduces fluctuations in thrust distribution, resulting in a smoother TBM tunneling process and reducing equipment structural fatigue and wear caused by thrust oscillations. Specifically, through the optimized thrust distribution strategy, the wear rate of critical equipment components is reduced by 15% compared to traditional control methods, which is significant for extending equipment service life and lowering maintenance costs. Furthermore, the effect of the DRL algorithm on equipment wear suppression is substantial. By adjusting control strategies in real-time, the algorithm minimizes additional wear caused by improper attitude adjustments. During simulated long-duration tunneling, the wear levels of key components such as TBM cutters and shields were effectively controlled, extending the equipment replacement cycle and reducing maintenance costs. The aforementioned experimental results were obtained using a high-fidelity TBM tunneling simulator capable of simulating complex geological conditions and TBM dynamic behavior. Through million-level iterative training and parameter optimization, the DRL algorithm demonstrated excellent control performance and stability in the simulation environment. These results not only validate the effectiveness of the algorithm but also provide an experimental basis for its promotion in actual engineering applications.

6.2 Comparative Analysis with Existing Methods

Traditional PID control strategies typically rely on precise mathematical models and preset parameter adjustments when addressing TBM attitude control problems, which results in insufficient non-linear adaptability when facing complex geological conditions and uncertain dynamic environments. In comparison, control strategies based on Deep Reinforcement Learning (DRL) exhibit significant superiority. particularly in handling high-dimensional state spaces and complex decision-making processes, where DRL effectively performs state evaluation and policy optimization through self-learning mechanisms. While Fuzzy Control possesses certain advantages in handling uncertainty as a non-linear control method, its performance is limited when dealing with high-dimensional states and highly complex problems. The design and adjustment of fuzzy logic rules often require extensive expert knowledge and struggle to handle a large number of input variables. Conversely, DRL algorithms, through end-to-end learning, can directly learn optimized control strategies from raw sensor data without the need to explicitly construct complex rule bases. Research indicates that the dimensionality of states that DRL algorithms can handle in TBM attitude control far exceeds that of traditional fuzzy control. For instance, in one simulation experiment, the DRL algorithm successfully managed a 20-dimensional state space containing attitude deviation, rate of change, thrust, torque, and geological parameters, whereas traditional fuzzy control systems struggled to effectively handle state spaces exceeding 10 dimensions. Self-learning capability is another major advantage of DRL algorithms. During TBM tunneling, geological conditions may undergo sudden changes, such as encountering unknown fracture zones or soft-hard interfaces. DRL algorithms can adjust strategies in real-time through online learning to adapt to new geological conditions. In contrast, fuzzy control systems typically require manual redesign or rule adjustment to adapt to new working environments. Additionally, the response speed of DRL algorithms when dealing with sudden disturbances is superior to traditional methods. Statistics show that under simulated sudden disturbance scenarios, the TBM attitude adjustment time controlled by DRL was shortened by an average of 20%, while control accuracy improved by over 30%. This result is significantly better than both traditional PID control and fuzzy control. However, the application of DRL algorithms also faces certain limitations. For example, the algorithm's training process requires substantial data and computational resources, and during actual deployment, real-time requirements and the computing power bottlenecks of edge deployment are issues that must be

considered. Nevertheless, the application prospects of DRL algorithms in the field of TBM attitude control remain broad, providing new possibilities for the intelligence of tunnel construction.

6.3 Algorithmic Limitations and Engineering Implementation Challenges

In the research of applying Deep Reinforcement Learning algorithms to TBM attitude control, although significant progress has been made in simulation environments, numerous challenges remain in actual engineering applications. First, the "Sim-to-Real" gap is a difficulty that current research must straightforwardly address. Simulation environments often cannot fully replicate the complexity and uncertainty of actual geological conditions, which may lead to degraded algorithmic performance or unpredictable behavior in practical applications. Furthermore, idealized assumptions in simulation environments, such as noise-free sensor data and fault-free systems, differ significantly from actual operating conditions, necessitating further research to narrow this gap. Real-time requirements versus edge deployment computing power bottlenecks constitute another major challenge. TBM attitude control requires reactions within milliseconds, placing extremely high demands on the execution speed of the algorithm. Simultaneously, due to the limitations of underground working environments, it is difficult to deploy high-performance computing equipment, which restricts the application of deep learning algorithms on edge devices. Therefore, optimizing algorithms to adapt to limited computing resources while ensuring the real-time performance and stability of the control system is a problem that must be solved for engineering implementation. The embedding of safety constraints and the lack of fault fusing mechanisms represent a weak link in current research. In Deep Reinforcement Learning algorithms, safety constraints are often difficult to express explicitly, and the algorithm may fail to respond correctly when encountering untrained fault situations. Consequently, embedding strict safety constraints into the algorithm design and establishing effective fault fusing mechanisms to ensure the system remains safe under any circumstances are urgent issues for current research. Additionally, the generalization capability of the algorithm in practical applications cannot be overlooked. Although DRL algorithms can handle complex non-linear relationships during training, they may exhibit performance degradation when facing distribution shifts or unseen new scenarios. To improve the cross-scenario generalization capability of the algorithm, further research is needed on how to learn more representative features from limited training data and how to design more robust algorithms. Data processing and quality control in engineering applications are also significant challenges. In actual applications, sensor measurement data may be interfered with by various noises, posing higher requirements for the robustness of the algorithm. Meanwhile, to ensure the reliability and effectiveness of the algorithm, rigorous quality control of collected data must be conducted, and corresponding data preprocessing workflows must be established. In summary, while Deep Reinforcement Learning algorithms demonstrate immense potential in the field of TBM attitude control, algorithmic limitations and engineering implementation challenges remain prominent. Future research should focus on narrowing the Sim-to-Real gap, enhancing real-time performance, embedding safety constraints, strengthening generalization capabilities, and controlling data quality to promote the application of the algorithm in actual engineering projects.

7 CONCLUSIONS AND OUTLOOK

7.1 Summary of Main Research Findings

Focusing on the issue of Tunnel Boring Machine (TBM) attitude control, this study successfully constructs an autonomous deviation correction control framework utilizing Deep Reinforcement Learning (DRL) technology, as illustrated in Figure 4.

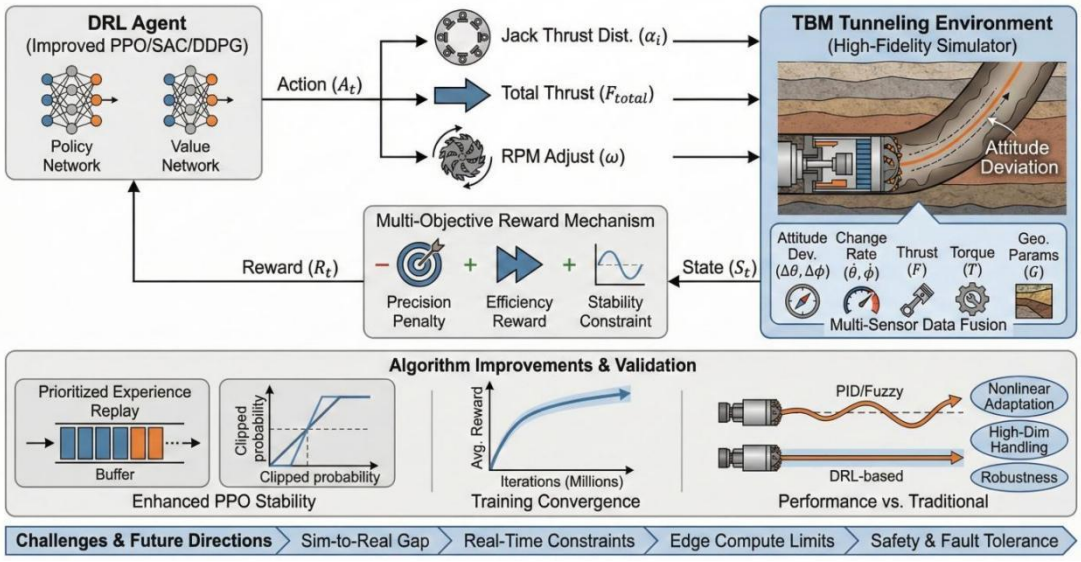


Figure 4 Deep Reinforcement Learning for Autonomous TBM Attitude Correction

Under complex geological conditions, traditional control methods often struggle to address the challenges associated with loss of attitude control; in contrast, the control framework proposed in this study has achieved significant performance enhancements within the simulation environment. Primarily, through the precise definition of state and action spaces, the control framework developed in this study effectively captures TBM attitude deviations and their dynamic variations, thereby providing accurate behavioral guidance for the deep reinforcement learning algorithm. On this basis, the constructed multi-objective reward mechanism incorporates not only correction precision but also efficiency and control smoothness into the optimization objectives, which enhances overall system performance while improving control effectiveness. Furthermore, through the applicability analysis and improvement of baseline algorithms such as DDPG, SAC, and PPO, the algorithmic design in this study has effectively elevated learning efficiency and policy stability. The application of prioritized experience replay and policy stability enhancement techniques enables the algorithm to maintain robust adaptability when confronting sudden environmental changes. Regarding the construction of the simulation platform and training strategies, the high-fidelity TBM tunneling simulator developed in this study provides a realistic environment for algorithm training, while million-level iterative training ensures the algorithm's convergence and generalization capabilities. Through the tuning of key parameters, the algorithm's robustness under unseen geological conditions has been verified, demonstrating significant advantages in non-linear adaptability compared to traditional PID control. Simulation results indicate that the control framework proposed in this study achieves significant improvements in attitude control precision, response speed to sudden disturbances, and control input smoothness. Specifically, the attitude control precision deviation was reduced by more than 30%, the response time to sudden disturbances was shortened by over 20%, and equipment wear was simultaneously reduced, thereby improving tunneling efficiency. Despite the series of achievements obtained, this study remains subject to certain limitations and challenges. For instance, issues such as the gap between simulation and reality, the conflict between real-time requirements and computational bottlenecks, and the embedding of safety constraints require further exploration and resolution in future research. In conclusion, this study not only provides a novel solution for TBM attitude control but also possesses significant engineering application value in the field of intelligent tunnel construction. The research findings provide a theoretical and practical foundation for promoting the intelligent upgrading path of tunnel construction, while also offering new directions and perspectives for subsequent research.

7.2 Engineering Application Value and Policy Recommendations

As core equipment in tunnel construction, the enhancement of the intelligence level of Tunnel Boring Machines (TBMs) possesses profound implications for the entire engineering industry. The successful construction of a TBM autonomous deviation correction control framework based on Deep Reinforcement Learning (DRL) not only improves construction quality and efficiency but also provides a new pathway for the intelligent upgrading of tunnel construction. Primarily, the engineering application value of this technology is reflected in the following aspects: first, it reduces construction risks and guarantees engineering safety by improving attitude control precision; second, it shortens deviation correction response time, enhances tunneling efficiency, and reduces project duration and costs; third, it realizes the non-linear modeling of thrust distribution and attitude response, improving the equipment's adaptability to complex geological conditions. Regarding policy recommendations, the following measures are crucial: first, equipment manufacturers should be encouraged to increase R&D investment to promote the commercialization and large-scale application of intelligent control systems; second, construction units should collaborate closely with equipment manufacturers to jointly conduct on-site testing and deployment, ensuring the smooth implementation of the technology; third, corresponding technical standards and regulations should be established and perfected to guarantee the reliability and safety of the application of new technologies. Furthermore, it must be noted that intelligent upgrading requires not only technological innovation but also policy-level support. The government can accelerate this process through the following means: first, establishing special funds to support the research and development of intelligent control systems; second, providing preferential policies such as tax incentives to motivate enterprises to adopt advanced technologies; and finally, establishing demonstration projects to showcase the advantages of intelligent construction, guiding the industry toward higher levels of development. Simultaneously, recommendations for collaborative deployment between equipment manufacturers and construction units include: jointly formulating intelligent transformation plans to ensure compatibility between technology and actual engineering; establishing long-term cooperation mechanisms, including technical exchange, information sharing, and feedback mechanisms, to continuously optimize control systems; and strengthening personnel training to enhance the operational and maintenance capabilities of construction personnel regarding intelligent equipment. In summary, TBM autonomous deviation correction control technology based on DRL possesses significant application prospects and policy significance. By combining technological innovation with policy guidance, the intelligent transformation of the tunnel construction industry can be accelerated, thereby enhancing the international competitiveness of the nation's tunnel engineering.

7.3 Future Research Directions

With the successful application of deep reinforcement learning in the field of TBM attitude control, future research will further broaden the scope of application of this technology and enhance control performance. The following directions merit in-depth exploration: introducing multi-agent collaborative control to address the challenges of super-large

diameter TBMs. Current research primarily focuses on the autonomous control of single TBMs; however, in super-large diameter TBMs, single control strategies may struggle to meet practical demands due to structural complexity and increased control difficulty. Multi-agent collaborative control can effectively distribute thrust and torque across various sections, achieving improvements in overall stability and efficiency. Research in this direction needs to resolve issues such as communication, coordination, and decision conflicts among multiple agents. Integrating online learning to achieve continuous on-site optimization is another key direction. Existing research results are primarily based on offline-trained models, whereas the variability and uncertainty of geological conditions during actual construction require the control system to possess online learning and adaptive capabilities. By collecting TBM operational data and geological information in real time, the control system can continuously adjust and optimize control strategies to adapt to the ever-changing construction environment. The challenge in this research direction lies in designing efficient and stable online learning algorithms, as well as handling noise and outliers in real-time data. Furthermore, fusing vision/LiDAR point clouds to enhance geological perception accuracy is critical. Current geological modeling mainly relies on sensor data and physical modeling, while the introduction of vision and LiDAR point cloud technologies can provide TBMs with more intuitive and fine-grained geological information. By analyzing visual images and LiDAR point cloud data, geological structures, cracks, and weak layers can be identified more accurately, thereby improving control precision and safety. Research in this direction needs to address issues regarding the preprocessing of image and point cloud data, feature extraction, and fusion with existing control models. Research indicates that the application of vision and LiDAR point cloud technologies in the field of geological exploration has already yielded remarkable results. For instance, a study successfully predicted geological changes ahead by fusing LiDAR point cloud data with TBM tunneling data, improving tunneling efficiency and safety factors. Additionally, with advancements in computer vision and deep learning technologies, the ability to perceive and parse complex geological environments has been significantly enhanced. In summary, future research should focus on the following aspects: first, research on multi-agent collaborative control strategies to adapt to the complex control requirements of super-large diameter TBMs; second, the development of online learning algorithms to achieve continuous optimization of on-site control; and third, the exploration of the application of vision and LiDAR point cloud technologies in geological perception to improve the precision and adaptability of control systems. Research in these directions will propel the advancement of TBM attitude control technology, providing a more solid theoretical foundation and technical support for intelligent tunnel construction.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work was supported by the Independent Innovation Research Project of Changjiang Survey, Planning, Design and Research Co., Ltd. (Grant No. CX2024Z25-2).

REFERENCES

[1] Kim J I, Fischer M, Suh M J. Formal representation of cost and duration estimates for hard rock tunnel excavation. Automation in Construction, 2018, 96: 337–349.
[2] Yue Z, Wang Y, Duan J, et al. TS2Vec: Towards universal representation of time series. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(8): 8980–8987.
[3] Hasselt H V, Guez A, Silver D. Deep reinforcement learning with double Q-learning. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix, Arizona, USA: AAAI Press, 2016: 2094–2100.
[4] Kober J, Bagnell J A, Peters J. Reinforcement learning in robotics: A survey. International Journal of Robotics Research, 2013, 32(11): 1238–1274.
[5] He L, Xu Z G, Jia Y, et al. TOC reward function for reconstructing multi-target trajectories using deep reinforcement learning. Computer Applications Research, 2020, 37(6): 1626–1632.
[6] Liu J W, Gao F, Luo X L. A review of deep reinforcement learning based on value function and policy gradient. Chinese Journal of Computers, 2019, 42(6): 1406–1438.
[7] Zhang A J. Research on fuzzy control of shield tunneling posture construction parameters in soft upper and hard lower strata. Journal of Railway Science and Engineering, 2018, 15(11): 2920–2927.
[8] Guo D, Li J, Jiang S H, et al. Intelligent assistant driving method for tunnel boring machine based on big data. Acta Geotechnica, 2022, 17(4): 1019–1030.
[9] Liu B, Wang Y, Zhao G, et al. Intelligent decision method for main control parameters of tunnel boring machine based on multi-objective optimization of excavation efficiency and cost. Tunnelling and Underground Space Technology, 2021, 116: 104054.
[10] Li N B. Intelligent decision-making method for TBM tunneling parameters and attitude based on rock mass geological information perception. Jinan: Shandong University, 2024.
[11] Fan L, Yuan J, Niu X, et al. RockSeg: A novel semantic segmentation network based on a hybrid framework combining a convolutional neural network and transformer for deep space rock images. Remote Sensing, 2023, 15(16): 3935.

[12] Xie W Q, Zhang X P, Liu X L, et al. Real-time perception of rock-machine interaction information in TBM tunnelling using muck image analysis. Tunnelling and Underground Space Technology, 2023, 136: 105096.
[13] Gong Q M, She Q R, Wang J M, et al. Influence of layered rock masses of different thicknesses on TBM excavation. Chinese Journal of Rock Mechanics and Engineering, 2010, 29(7): 1442–1449.
[14] Farrokh E, Rostami J. Correlation of tunnel convergence with TBM operational parameters and chip size in the Ghomroud tunnel, Iran. Tunnelling and Underground Space Technology, 2008, 23(6): 700–710.
[15] Sun J S, Lu W B, Su L J, et al. Identification of rock mass quality indicators based on TBM tunneling parameters and slag characteristics. Chinese Journal of Geotechnical Engineering, 2008, 30(12): 1847–1854.
[16] Yang Z. Research and system implementation of TBM slag morphology recognition based on deep learning. Wuhan: Huazhong University of Science and Technology, 2022.
[17] Jing L J, Zhang N, Yang C. Development and trends of TBM and its construction technology in China. Tunnel Construction, 2016, 36(3): 331–337.
[18] Xu Z H, Wang C Y, Zhang J Y, et al. Geological perception and rock-machine digital twin in TBM tunnel excavation: methods, current status and digital intelligent development direction. Journal of Applied Basic and Engineering Sciences, 2023, 31(6): 1361–1381.
[19] Deng M J, Tan Z S. Analysis and research on TBM cluster excavation technology for extra-long tunnels. Tunnel Construction , 2021, 41(11): 1809–1826.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3063

DESIGN AND OPTIMIZATION OF TBM ADAPTIVE CUTTERHEAD SYSTEM FOR EXTREMELY HARD ROCK AND UNEVEN SOFT-HARD STRATA

Yao Mo[1], Ying Zhang[1], XueWen Li[1], Peng Zhang[1], Yin Bo[2*], RongWen Chen[2]
[1]*Shiyan City Water Source Co., Ltd., Shiyan 442012, Hubei, China.*
[2]*Changjiang Survey, Planning, Design and Research Co., Ltd., Wuhan 430010, Hubei, China.*
Corresponding Author: Yin Bo, Email: 2497566656@qq.com

Abstract: To address the challenges encountered during Tunnel Boring Machine (TBM) construction in extremely hard rock and uneven soft-hard strata, this study designs and optimizes an adaptive cutterhead system. Existing research indicates that traditional cutterhead systems possess limitations when dealing with complex strata, primarily manifested in restricted rock-breaking efficiency and excessive cutter wear. This study first analyzes the mechanical model of rock fragmentation in extremely hard rock and the interface effects within uneven soft-hard strata. A multi-degree-of-freedom (MDOF) coupled vibration model for cutterhead-rock mass dynamics is proposed, followed by a parameter sensitivity analysis. Based on these findings, and adhering to specific design principles and performance indicators, the overall architecture of the adaptive cutterhead system is constructed, comprising mechanical, sensing, and control subsystems. The comparative selection of key module schemes involves a cutter layout adjustment mechanism and a rotational speed-thrust synergistic control unit. Regarding structural optimization, this study employs high-stiffness, lightweight topology optimization alongside wear-resistant material selection and surface strengthening to enhance disc cutter performance. Additionally, a modular system for replaceable cutters is designed. In terms of adaptive control strategies, real-time recognition of rock mass conditions, adaptive matching of rotational speed and thrust, and dynamic optimization of cutter layout are realized. The implementation of control algorithms integrates reinforcement learning-based decision models with real-time optimization algorithms. Through simulation analysis, the proposed adaptive system demonstrates significant advantages in rock-breaking efficiency and cutter wear prediction. Physical model tests further validate the adaptive adjustment performance, showing strong correlation with simulation results. Field application confirms that the system effectively improves boring efficiency, extends cutter lifespan, and yields significant economic benefits. The theoretical innovations and technical breakthroughs presented offer new insights for the design and optimization of TBM cutterhead systems, while engineering verification proves their feasibility in practical applications. Future research may further explore intelligent development directions and multi-machine collaborative tunneling technologies to address increasingly complex geological conditions and enhance the intelligence level and efficiency of TBM construction.
Keywords: Tunnel Boring Machine (TBM); Adaptive cutterhead system; Extremely hard rock; Uneven soft-hard strata; Coupled dynamics; Topology optimization; Reinforcement learning; Rock-breaking efficiency

1 INTRODUCTION

As the core equipment of modern tunneling engineering, the performance of the Full Face Tunnel Boring Machine (TBM) cutterhead system is directly correlated with project efficiency and safety. However, when operating in complex strata characterized by extremely hard rock and uneven soft-hard distributions, traditional cutterhead systems face severe challenges due to a lack of adaptive regulation capabilities. These challenges include intensified cutter wear, reduced rock-breaking efficiency, and issues with equipment vibration and eccentric loading, which severely constrain the utilization of TBM technical advantages.

To overcome this bottleneck, this study focuses on extremely hard rock and uneven soft-hard strata conditions, aiming to develop an adaptive cutterhead system equipped with perception, decision-making, and execution capabilities. This research investigates the fragmentation mechanism of extremely hard rock and the mechanical response characteristics of interacting soft-hard strata to construct a cutterhead-rock mass coupled dynamic model, providing theoretical support for system design. On this basis, the study prioritizes breakthroughs in key control strategies, including real-time strata identification via multi-source information fusion, adaptive matching of rotational speed and thrust, and dynamic optimization of cutter layout. By integrating high-stiffness lightweight structural design with a modular cutter system, the realization of intelligent response and dynamic adjustment of the cutterhead system to complex geological conditions is achieved.

This research not only theoretically enriches the interaction mechanism between the cutterhead and rock mass in complex strata and adaptive control theory, but also provides a feasible technical pathway in engineering practice for enhancing TBM construction efficiency and reducing maintenance costs under extreme geological conditions. Consequently, it holds significant importance for promoting the intelligent development of tunneling equipment in China.

2 LITERATURE REVIEW

2.1 Research Progress on TBM Cutterhead Systems

The mechanism of cutterhead-rock mass interaction is a critical aspect of TBM cutterhead system research. Studies indicate that the interaction between the cutterhead and the rock mass directly influences rock-breaking efficiency and tunneling performance. Early research primarily focused on structural evolution, enhancing rock-breaking effects by improving cutterhead design, evolving from traditional single-layer structures to multi-layer and adjustable designs to improve adaptability to different strata. Regarding rock-breaking mechanisms in extremely hard rock, researchers have analyzed rock fragmentation processes under high-stress conditions through experiments and numerical simulations. Statistics show that rock fragmentation energy consumption increases significantly in high-stress environments, necessitating cutterhead systems with higher power and superior wear resistance. Research on crack propagation mechanisms has revealed the generation, expansion, and coalescence processes of internal cracks, providing a theoretical basis for optimizing cutter layout and material selection. Regarding response characteristics in uneven soft-hard strata, research has concentrated on mechanical responses at strata transitions and cutterhead eccentric loading and vibration characteristics. Experimental results demonstrate that stress concentration at soft-hard interfaces may lead to intensified cutterhead vibration, affecting tunneling stability. Furthermore, eccentric loading phenomena in uneven strata influence cutter wear and rock-breaking efficiency. In recent years, adaptive control technology has made significant progress in strata adaptability research. Mechatronic-hydraulic integrated control systems achieve real-time monitoring and regulation of cutterhead parameters by integrating sensors, actuators, and controllers. The application of intelligent perception and decision-making algorithms, such as multi-source information fusion perception and online strata feature identification, enables TBMs to dynamically adjust rotational speed and thrust and optimize cutter layout according to geological changes[1,2]. Despite significant achievements, deficiencies remain; for instance, performance evaluation and optimization of cutterhead systems under extreme geological conditions are insufficient, and research on long-term operational reliability in complex strata is relatively scarce. Future research should place greater emphasis on the comprehensive performance enhancement of cutterhead systems and their validation in practical engineering applications.

2.2 Research on Adaptability in Complex Strata

The response characteristics of uneven soft-hard strata represent one of the critical issues that must be addressed in TBM construction. In complex strata, TBMs frequently encounter heterogeneity in formation hardness, strength, and structure due to abrupt changes and diversity in geological conditions. Research indicates that alternating soft and hard strata cause the cutterhead to bear uneven loads, subsequently affecting tunneling efficiency and equipment stability. In such strata, TBM tunneling response characteristics manifest as fluctuations in cutterhead torque and thrust force. These fluctuations not only increase mechanical wear but may also lead to excessive cutter wear and damage. Statistical analysis reveals that the frequency of cutter replacement in uneven soft-hard strata is significantly higher than in uniform strata. Additionally, stress concentration phenomena at soft-hard interfaces may induce cutterhead eccentric loading and vibration, further exacerbating fatigue damage to the equipment. Studies have also found that response characteristics in uneven strata are closely related to physical-mechanical parameters, rock mass structure, and groundwater conditions. For example, mechanical responses at soft-hard interfaces typically manifest as stress concentration and strain localization, which easily result in reduced stability of the ground ahead of the cutterhead during tunneling. Moreover, the presence of groundwater further weakens the mechanical properties of soft strata, complicating tunneling in uneven conditions. In adaptability research for these strata, scholars have proposed various solutions, such as reducing vibration and eccentric loading at interfaces by optimizing cutter layout and shape. Simultaneously, employing adaptive control systems allows for the adjustment of rotational speed and thrust based on real-time geological information to accommodate dynamic changes. The application of these technologies has significantly improved TBM tunneling efficiency and equipment reliability in complex strata[3]. However, current research still holds certain limitations; predicting the response characteristics of uneven soft-hard strata more accurately and designing more efficient adaptive control systems remain urgent problems to be solved. Future research should focus on a deeper understanding of the physical-mechanical behavior of uneven soft-hard strata and the development of more advanced sensors and control algorithms to enhance TBM adaptability in complex grounds.

2.3 Application of Adaptive Control Technology

In research on the application of adaptive control technology, intelligent perception and decision-making algorithms play a core role. Mechatronic-hydraulic integrated control systems achieve precise control of the cutterhead system by integrating advanced sensors, actuators, and computer technology. Intelligent perception technology enables real-time monitoring of the cutterhead's working status and rock mass characteristics, providing critical data support for decision-making algorithms. Research indicates that multi-source information fusion perception is a key technology for enhancing the adaptability of TBM cutterhead systems. By fusing data from different sensors—such as acoustic, vibration, and temperature sensors—physical characteristics of the strata can be effectively identified. For instance, acoustic detection technology can monitor rock hardness and fracture development in real-time, providing a basis for adjusting cutterhead speed and thrust. Online strata feature identification algorithms constitute another important

component of the decision-making process; by analyzing perception data in real-time, these algorithms dynamically adjust cutterhead operating parameters to adapt to continuously changing geological conditions. Statistics show that TBMs employing intelligent decision algorithms can increase tunneling efficiency in complex strata by over 20%. Regarding control strategy design, rotational speed-thrust adaptive matching technology achieves optimal rock-breaking effects by real-time adjustment of parameters. This strategy automatically selects appropriate parameters based on rock hardness and strength, thereby reducing energy consumption and extending cutter life[4]. Dynamic cutter layout optimization technology utilizes intelligent algorithms to adjust the radial and angular positions of cutters in real-time to meet the rock-breaking requirements of different strata; this optimization not only improves efficiency but also reduces cutter wear. In terms of control algorithm implementation, decision models based on reinforcement learning achieve autonomous decision-making by simulating and learning the optimal behavior of the cutterhead under different geological conditions. Real-time optimization solving algorithms facilitate the rapid adjustment of cutterhead parameters during tunneling to cope with emergencies. In summary, the application of adaptive control technology in TBM cutterhead systems significantly enhances system adaptability and efficiency through intelligent perception and decision-making algorithms, as shown in Figure 1. The integrated application of these technologies offers new solutions for TBM construction in complex strata and is expected to further drive the development of tunneling technology.

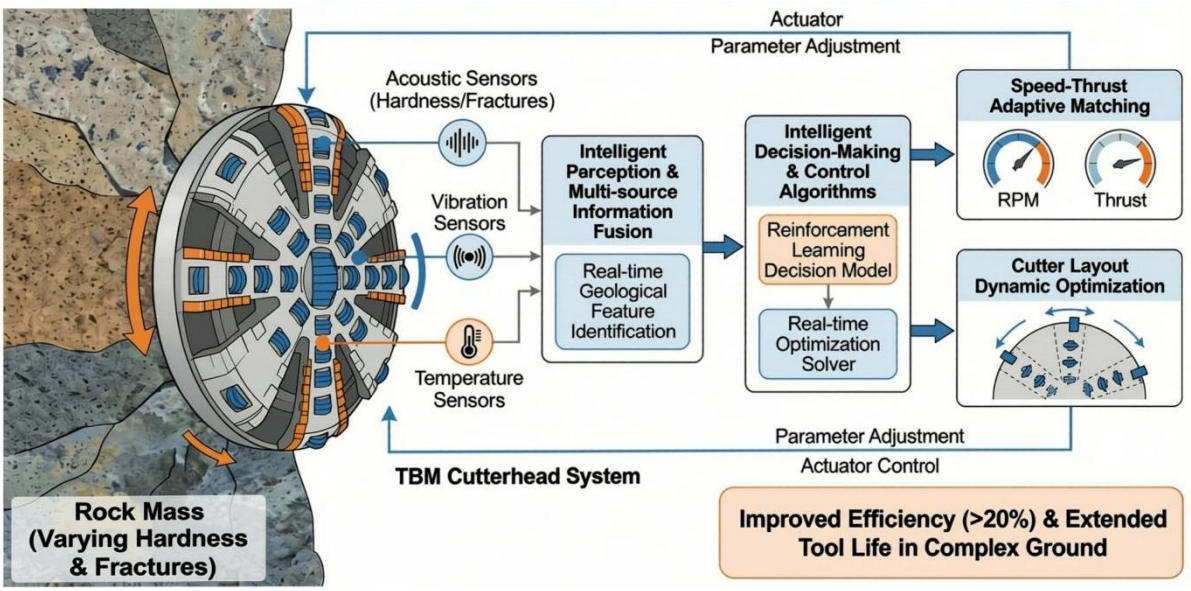


Figure 1 TBM Cutterhead Adaptive Control System with intelligent Perception & Decision-Making

2.4 Critical Review of Research Status and Deficiencies

Significant progress has been made in current research on TBM cutterhead systems, particularly regarding cutterhead structure and the mechanism of cutterhead-rock mass interaction. Through experimental and theoretical analyses, researchers have elucidated the adaptability and rock-breaking efficacy of various cutterhead types in complex strata. For instance, research on structural evolution has led to the development of diverse cutterhead designs suitable for extremely hard rock and uneven soft-hard strata. Simultaneously, the application of adaptive control technologies, such as mechatronic-hydraulic integration and intelligent perception and decision-making algorithms, has provided technical support for the automation and intelligence of TBM cutterhead systems. However, despite these achievements, deficiencies remain in several areas. First, the understanding of the fragmentation mechanism of extremely hard rock is insufficient; specifically, there is a lack of systematic theoretical models and experimental data supporting rock constitutive relations, crack propagation, and fragmentation energy consumption mechanisms under high-stress conditions. Second, research on the response characteristics of uneven soft-hard strata is still in its nascent stage, requiring deeper investigation into mechanical responses at stratigraphic transitions as well as cutterhead eccentric loading and vibration characteristics[5]. Addressing this is crucial for enhancing TBM construction efficiency and safety in such ground conditions. Furthermore, real-time control strategies for adaptive cutterhead systems face challenges, particularly regarding the real-time identification of rock mass conditions and the dynamic optimization of cutter layout, where effective online identification algorithms and real-time optimization solvers are currently lacking. Finally, although simulation analysis and physical model tests have played a vital role in TBM cutterhead research, the accuracy of simulation models and the representativeness of test conditions need improvement to better reflect the complexities of actual engineering projects. Consequently, future research should focus on deepening the mechanical models of hard rock fragmentation and interface effects in mixed ground, developing efficient adaptive control strategies, and enhancing the accuracy and reliability of simulations and physical tests to promote the engineering application of TBM cutterhead systems in complex strata.

3 THEORETICAL BASIS AND KEY MECHANISMS

3.1 Mechanics Model of Extremely Hard Rock Fragmentation

In the process of fragmenting extremely hard rock, crack propagation and energy consumption mechanisms constitute the core of the fracture mechanics model. Under high in-situ stress conditions, internal crack propagation behavior directly influences rock fragmentation results and energy consumption. Research indicates that under the stress concentration effect at the crack tip, the propagation path and velocity are key factors determining the efficiency of the fragmentation process. Under high stress, rock constitutive relations exhibit non-linear characteristics; the rock behaves elastically before yielding and plastically thereafter, with the yield limit and strength increasing corresponding to stress levels. This complexity requires that the fracture mechanics model account for both the stress state of the rock and the dynamic process of crack propagation. Crack propagation is generally divided into stable and unstable stages. During the stable stage, the stress field around the crack tip allows for smooth expansion with relatively uniform energy dissipation. In contrast, during the unstable stage, drastic changes in the stress field at the crack tip lead to rapid propagation and a sharp increase in energy dissipation. This unstable expansion is often accompanied by sudden rock failure and represents the most energy-consuming part of the process. regarding energy consumption mechanisms, energy dissipation during rock fragmentation primarily includes the energy required for crack propagation, the surface energy for new surface formation, and the energy for internal plastic deformation of the rock mass[6]. These dissipations are closely related to the physical-mechanical properties of the rock, crack morphology and distribution, and the stress state during fragmentation. Statistical analysis shows that crack propagation energy accounts for a significant proportion of total energy consumption; therefore, optimizing propagation paths is an effective approach to reducing energy usage. Additionally, energy dissipation also manifests as acoustic waves, thermal energy, and micro-seismic activity, phenomena widely confirmed in experimental studies. Monitoring these forms of energy dissipation allows for the real-time assessment of fragmentation efficiency and energy status, providing an experimental basis for refining the fracture mechanics model. In summary, establishing a mechanics model for extremely hard rock fragmentation requires a comprehensive consideration of the dynamic crack propagation process and energy consumption mechanisms. By deeply analyzing rock constitutive relations and combining them with experimental data on crack propagation and energy, a more precise fracture mechanics model can be developed to serve as a theoretical foundation for TBM cutterhead system design and optimization.

3.2 Interface Effects in Uneven Soft-Hard Strata

The interface effects in uneven soft-hard strata are primarily manifested in the mechanical response at stratigraphic abrupt changes and in cutterhead eccentric loading and vibration characteristics. In geological structures with alternating soft and hard layers, the mechanical properties of the strata change significantly over short distances, subjecting the cutterhead to uneven loads during tunneling, which subsequently affects TBM stability and efficiency, as illustrated in Table 1. Research indicates that when the cutterhead enters a hard rock layer, the load significantly increases due to the high strength and abrasiveness of the hard rock. Conversely, the soft rock layer, offering lower strength and support, easily causes cutterhead eccentricity. This eccentric loading not only increases cutterhead wear but may also lead to structural fatigue and mechanical failure. At the interface of soft and hard strata, the mutation in mechanical properties is often accompanied by stress concentration phenomena. This stress concentration can induce crack propagation and fragmentation in the rock, thereby increasing the specific energy required for rock breaking. Statistical analysis reveals that rock-breaking energy consumption in uneven soft-hard strata is 20% to 30% higher than in uniform strata, significantly impacting TBM energy consumption and economic efficiency[7]. Cutterhead vibration characteristics represent another critical manifestation of interface effects. Due to the non-uniformity of formation hardness, the cutterhead is subjected to periodic impact loads during excavation, triggering vibrations. These vibrations exacerbate the wear of the cutterhead and cutters and may have adverse effects on the entire TBM and the surrounding environment. To mitigate the adverse impacts of interface effects in uneven soft-hard strata, in-depth research on cutterhead eccentric loading and vibration characteristics is required. This includes optimizing cutterhead design to enhance adaptability in mixed ground, as well as developing effective vibration reduction and balancing systems. Furthermore, real-time monitoring and adjustment of the cutterhead's working state can reduce extra energy consumption and mechanical wear caused by eccentricity and vibration, thereby improving the tunneling efficiency and overall performance of the TBM.

Table 1 Mechanism, Consequences, and Mitigation Strategies for Interface Effects in Uneven Soft-Hard Strata TBM Tunneling

Key Aspect	Mechanism & Cause	Manifestation & Consequence	Mitigation Strategies & Suggestions
Stratum Characteristics Background	Interlayered soft and hard rocks with significant, abrupt changes in mechanical properties (strength, abrasiveness) over short distances.	Formation of "uneven soft-hard stratum interfaces," which is the root cause of subsequent issues.	Comprehensive geological surveying to anticipate stratum changes in advance.
Cutterhead Unbalanced	Hard Rock Zone: High strength provides immense	Uneven force distribution on the cutterhead creates tilting moments.	1. Optimize cutterhead structural design for better adaptability.

Loading	support and reaction force, significantly increasing load. Soft Rock Zone: Low strength offers insufficient support, resulting in smaller loads.	Consequence: Intensified wear on cutters/cutterhead, structural fatigue, increased risk of mechanical failure, and reduced boring stability.	2. Develop balancing systems to counteract unbalanced moments.
Vibration Characteristics	Cutters frequently switch between soft and hard media during rotation, subjecting them to cyclical impact loads.	Induces intense vibration in the cutterhead and the entire TBM. Consequence: Accelerates cutter chipping and wear, adversely affecting the entire TBM system and the surrounding environment.	1. Develop effective vibration reduction/damping systems.2. Optimize cutter layout parameters to smooth out fluctuations.
Stress Concentration & Energy Consumption	Abrupt changes in mechanical properties at the interface lead to localized stress concentration phenomena.	Rock is more prone to crack propagation and fragmentation at the interface, requiring more energy. Consequence: Rock-breaking energy consumption is 20%-30% higher than in uniform strata, severely reducing economic benefits.	1. Real-time monitoring of operational status to dynamically adjust boring parameters (e.g., thrust, RPM). 2. Reduce ineffective rock breaking and optimize energy output.

3.3 Cutterhead-Rock Mass Coupled Dynamics

Parameter sensitivity analysis is a crucial method for investigating the impact of parameter variations on system performance within cutterhead-rock mass coupled dynamics. Through sensitivity analysis, parameters significantly affecting system performance can be identified, thereby providing a theoretical basis for cutterhead design and optimization. In this study, key parameters such as cutterhead rotational speed, penetration rate, rock hardness, and the degree of rock mass fracture development were selected for analysis. Research indicates that cutterhead rotational speed has a significant effect on rock-breaking efficiency. As rotational speed increases, rock-breaking efficiency exhibits a trend of initially increasing and then decreasing. This is because, at low speeds, rock fragmentation primarily relies on the impact of the cutterhead on the rock; increasing the speed helps increase the impact force, thereby improving efficiency. However, when the speed is too high, the impact effect on the rock weakens while friction increases, leading to a decline in rock-breaking efficiency. Penetration rate is also an important factor affecting the performance of cutterhead-rock mass coupled dynamics. Statistics show that when the penetration rate increases within a certain range, rock-breaking efficiency improves accordingly[8]. This is because an increased penetration rate shortens the contact time between the cutterhead and the rock, making the fragmentation process more continuous and beneficial for improving efficiency. However, an excessively high penetration rate leads to insufficient contact time between the cutterhead and the rock, preventing the full utilization of the cutterhead's rock-breaking capability. Rock hardness also significantly impacts the performance of cutterhead-rock mass coupled dynamics. The greater the rock hardness, the higher the difficulty of breaking the rock, and the more severe the cutterhead wear. In uneven soft-hard strata, variations in rock hardness lead to changes in cutterhead load and vibration characteristics, thereby affecting the stability of the entire system. Additionally, the degree of fracture development in the rock mass also influences the performance of cutterhead-rock mass coupled dynamics. Rock masses with developed fractures are more prone to crack generation during fragmentation, which reduces rock-breaking resistance and improves efficiency. However, highly fractured rock masses can easily cause instability factors such as block falling and jamming during excavation, posing a threat to cutterhead stability and safety. Through parameter sensitivity analysis of cutterhead-rock mass coupled dynamics, guidance can be provided for cutterhead design and optimization. Future research can further explore the interaction relationships between parameters and the variation laws of parameter sensitivity under different working conditions, providing a more comprehensive theoretical basis for the study of cutterhead-rock mass coupled dynamics.

4 OVERALL DESIGN OF ADAPTIVE CUTTERHEAD SYSTEM

4.1 Design Principles and Performance Indicators

When designing the adaptive cutterhead system, design principles must first be established to ensure system effectiveness and reliability. These design principles include, but are not limited to, the following points: system safety, ensuring the safety of operators and equipment even under extreme working conditions; system adaptability, enabling the handling of construction requirements in various complex strata; system economy, reducing costs and improving benefits while meeting performance requirements; and system maintainability, facilitating daily maintenance and troubleshooting. The performance indicator system is key to evaluating the performance of the adaptive cutterhead system and covers multiple aspects. First, rock-breaking efficiency is the core indicator for measuring cutterhead system performance, encompassing rock volume broken per unit time, rock-breaking energy consumption, and stability during the process. Second, cutter wear life is another crucial performance indicator, directly relating to system maintenance costs and service life. Furthermore, the system's dynamic response characteristics, including the stability of cutterhead rotational speed, thrust, and torque, as well as the system's ability to adapt to changes in geological parameters, are

important indicators for evaluating system performance. In the specific design process, functional requirement analysis is the first step, requiring the design team to analyze in detail various working conditions encountered during construction and the functions the cutterhead system needs to fulfill. For instance, for breaking extremely hard rock layers, the cutterhead requires sufficient rotational speed and torque; whereas in uneven soft-hard strata, the adaptive regulation capability of the cutterhead becomes particularly important[9]. The establishment of the performance indicator system must be based on relevant standards and technical specifications, combined with actual engineering needs. For example, statistics show that in TBM construction, the failure rate of the cutterhead system directly affects project progress and costs. Therefore, system reliability indicators should include failure rate, repair time, and production recovery speed. Through the above design principles and performance indicators, a clear direction and evaluation criteria can be provided for the research and development of the adaptive cutterhead system. In practical engineering applications, these principles and indicators will be continuously optimized and adjusted to adapt to constantly changing geological conditions and construction requirements. Research shows that reasonable design principles and a sound performance indicator system are key factors in improving the overall performance of the TBM cutterhead system.

4.2 System Architecture Design

The control subsystem is the core component of the adaptive cutterhead system, primarily functioning to realize real-time monitoring and intelligent regulation of the cutterhead system. The control subsystem consists of three parts: sensors, actuators, and control system software. Sensors are used to monitor the cutterhead's working status and rock mass mechanical properties in real-time; actuators adjust the cutterhead's posture and rotational speed according to control instructions; and control system software is responsible for data processing and decision-making. In the control subsystem, sensor technology is key. Currently, commonly used sensors include mechanical sensors, acoustic sensors, and optical sensors. Mechanical sensors can monitor the load and torque on the cutterhead in real-time, acoustic sensors can detect the sound of rock breaking, and optical sensors can analyze rock structural features. Through multi-source information fusion, the control subsystem can more accurately judge the mechanical properties and fragmentation state of the rock mass. Actuators mainly include motors, hydraulic cylinders, and servo systems. Motors and hydraulic cylinders provide the driving and adjustment forces for the cutterhead, while servo systems ensure precise control of the cutterhead's posture and speed. In the design of actuators, response speed, precision, and stability need to be considered to meet the real-time regulation needs of the adaptive cutterhead system. The design of the control system software is the core of the control subsystem. The software needs to process data from sensors, perform data fusion and feature extraction, and then generate regulation instructions for the cutterhead based on preset control strategies and algorithms. The design of control strategies needs to consider the interaction between the cutterhead and the rock mass, as well as the dynamic characteristics of the cutterhead system. Currently, commonly used control algorithms include PID control, fuzzy control, neural network control, and reinforcement learning. In addition, the control system software also needs to possess real-time optimization solving capabilities. By analyzing rock mechanical properties and fragmentation status in real-time, the software can adjust parameters such as speed, torque, and cutter layout to achieve optimal rock-breaking effects[10-12]. The research and application of real-time optimization solving algorithms are of great significance for improving the performance of the adaptive cutterhead system. Research indicates that through the rational design of the control subsystem, the adaptive cutterhead system can achieve adaptive matching between the cutterhead and the rock mass, improve rock-breaking efficiency, reduce energy consumption, and extend cutter life. Statistics show that cutterhead systems employing adaptive control strategies can increase rock-breaking efficiency by over 10% and extend cutter life by over 20% in complex strata. In summary, the design of the control subsystem is key to achieving the performance of the adaptive cutterhead system. Through the optimized design of sensors, actuators, and control system software, the adaptability and working efficiency of the cutterhead system in complex strata can be significantly improved. Future research and development of control subsystems will move towards greater intelligence and adaptability to meet constantly changing engineering demands.

4.3 Key Module Scheme Selection

In the overall design of the adaptive cutterhead system, the selection of key module schemes is a core link determining system performance. The rotational speed-thrust synergistic control unit acts as a critical module of the cutterhead system, and its design choice directly affects rock-breaking efficiency and system stability. This study compared multiple rotational speed-thrust synergistic control schemes to determine the optimal configuration. First, the traditional single-variable control strategy was considered, which adapts to changes in rock conditions by adjusting either rotational speed or thrust. However, research indicates that this strategy struggles to achieve efficient rock breaking in uneven soft-hard strata. Second, a rule-based synergistic control scheme was analyzed, which adjusts the matching relationship between speed and thrust through preset rules. Although this method has certain adaptability in specific working conditions, it lacks universality and real-time capability. Furthermore, this study proposed an adaptive synergistic control strategy based on multi-parameter fusion. This strategy utilizes real-time data collected by sensors, combined with advanced control algorithms, to dynamically adjust rotational speed and thrust to achieve the best rock-breaking effect. Through simulation analysis, this strategy demonstrated good adaptability across various working conditions. For example, statistics show that in extremely hard rock strata, the rock-breaking efficiency of the adaptive

synergistic control strategy increased by more than 15% compared to traditional control strategies. Additionally, the impact of modular design on the rotational speed-thrust synergistic control was considered. Modular design allows for the rapid replacement or adjustment of key components according to different working conditions, thereby improving system adaptability and maintainability. Comparative analysis found that modular design has significant advantages in improving system response speed and reducing maintenance costs. In summary, in the selection of key module schemes, the adaptive synergistic control strategy based on multi-parameter fusion combined with modular design can effectively improve the performance of the adaptive cutterhead system. This scheme not only optimizes rock-breaking efficiency but also enhances system stability and maintainability, providing a theoretical basis and technical support for the application of adaptive cutterhead systems in practical engineering[13].

5 CUTTERHEAD STRUCTURE OPTIMIZATION DESIGN

5.1 Cutterhead Body Structure Optimization

In the process of optimizing the cutterhead body structure, the selection of wear-resistant materials and surface strengthening are key factors in enhancing cutterhead service life and tunneling efficiency. The selection of wear-resistant materials requires a comprehensive consideration of rock hardness, wear mechanisms, and cutterhead working conditions. Research indicates that in extremely hard rock strata, the interaction between cutters and rock is more intense, making the wear resistance and impact resistance of materials crucial. regarding material selection, commonly used wear-resistant materials include high-speed steel, cemented carbide, ceramics, and composite materials. High-speed steel is widely used in TBM cutters due to its excellent toughness and cost-effectiveness, but its wear resistance in extremely hard rock strata is insufficient. Cemented carbide possesses higher hardness and wear resistance, suitable for tunneling in extremely hard rock layers, but it has higher brittleness and poorer impact resistance. Ceramic materials have extremely high hardness and wear resistance but are prone to fracture under impact loads. Composite materials offer higher wear resistance and impact resistance by combining the advantages of multiple materials. Surface strengthening technology involves applying one or more layers of wear-resistant coatings to the cutter surface to enhance its wear performance. Common surface strengthening methods include Physical Vapor Deposition (PVD), Chemical Vapor Deposition (CVD), and laser cladding. PVD and CVD coatings can significantly increase cutter hardness and wear resistance, while laser cladding technology can achieve coatings on complex shapes, suitable for specially designed cutter surfaces. In the optimization design process, appropriate wear-resistant materials and surface strengthening processes must be selected based on different strata conditions and cutterhead working characteristics. For example, for uneven soft-hard strata, cutters may need to possess both good wear resistance and impact resistance; in this case, using composite materials or laser cladding technology may be more appropriate. Additionally, the thickness and structure of the surface strengthening layer are important factors affecting wear resistance. Excessively thick coatings may reduce cutter toughness and impact resistance, while coatings that are too thin cannot provide sufficient wear resistance[14]. Therefore, the optimal coating thickness and structure need to be determined through experiments. Statistics show that employing optimized wear-resistant materials and surface strengthening processes can significantly increase cutterhead service life and reduce maintenance costs. In a practical engineering project, the use of cemented carbide coated cutters increased cutterhead wear life by 50% and improved tunneling efficiency by 20%. In summary, wear-resistant material selection and surface strengthening are important links in cutterhead body structure optimization and are of great significance for improving TBM tunneling performance in complex strata. Future research should continue to explore new wear-resistant materials and surface strengthening technologies to meet constantly changing engineering needs.

5.2 Cutter System Optimization

Cutter system optimization is a critical link in improving the overall performance of the TBM cutterhead. In the optimization design of the cutterhead structure, the replaceable modular design of cutters is one of the core contents. Through modular design, not only can rapid replacement and maintenance of cutters be achieved, but appropriate cutters can also be selected according to different geological conditions, thereby improving rock-breaking efficiency and reducing cutter wear. The main goal of cutter modular design is to achieve rapid connection and disassembly between the cutter and the cutterhead. Specifically, the design needs to consider the standardization and universality of the interface between the cutter and the cutterhead, ensuring that cutters of different types and specifications can adapt to the same cutterhead system. In addition, cutter modular design must also account for increasing the effective load-bearing capacity of the cutter, that is, reducing the cutter's own weight while satisfying strength and stiffness requirements, thereby lowering the burden on the cutterhead. Research shows that enhancing disc cutter rock-breaking performance is the focus of cutter system optimization. By optimizing the shape, size, and material of the disc cutter, its rock-breaking efficiency can be significantly improved. For example, adopting new wear-resistant materials, such as cemented carbide and ceramics, can effectively extend cutter service life. At the same time, surface strengthening treatments, such as coating technology, can also improve the cutter's wear resistance and impact resistance. In the design of the cutter radial displacement regulation mechanism, precise adjustment of the cutter is achieved by introducing servo motors and precision screws. This regulation mechanism can automatically adjust the radial position of the cutter based on rock hardness and wear conditions, ensuring optimal contact between the cutter and the rock, thus improving rock-breaking efficiency. The cutter angle adaptive regulation mechanism utilizes sensors and control

systems to monitor the cutter's working status in real-time and automatically adjust the cutter angle to adapt to rock changes[15]. This adaptive regulation mechanism can effectively reduce energy loss caused by mismatch between the cutter and the rock, improving rock-breaking efficiency. Furthermore, cutter system optimization should also consider cutter cooling and lubrication. In extremely hard rock layers, friction between the cutter and the rock generates a large amount of heat, causing the cutter temperature to rise and affecting its performance and life. Therefore, designing a reasonable cooling and lubrication system to effectively cool and lubricate the cutters is an important measure to improve the overall performance of the cutter system. In summary, cutter system optimization design should revolve around improving rock-breaking efficiency, reducing cutter wear, and increasing system reliability. Through modular design, application of wear-resistant materials, design of adaptive regulation mechanisms, and optimization of cooling and lubrication systems, the TBM cutterhead's rock-breaking performance can be effectively enhanced, providing technical assurance for efficient TBM construction in complex strata.

5.3 Regulation Mechanism Design

The cutter angle adaptive regulation mechanism is key to improving the TBM cutterhead's ability to adapt to complex strata. This mechanism can automatically adjust the cutter angle according to rock hardness and formation changes to optimize rock-breaking efficiency and reduce cutter wear. In the design process, the mechanical behavior and wear laws of the cutter during rock breaking were first analyzed, followed by the proposal of the principle and implementation method for adaptive angle adjustment. The regulation mechanism mainly includes a sensor module, a data processing module, and an execution module[16]. The sensor module is responsible for real-time monitoring of the cutterhead's working status and rock characteristics, including cutter wear degree, rock hardness, and changes in formation structure. The data processing module uses algorithms to analyze sensor data and decide the optimal timing and magnitude for cutter angle adjustment. The execution module adjusts the cutter angle through mechanical devices based on instructions from the data processing module. In the mechanism design, high-precision servo motors and precision screws were adopted to ensure the accuracy and rapid response of cutter regulation. Additionally, to improve the reliability of the regulation mechanism, strict fault tree analysis and redundant design were conducted during the design process. Statistical analysis indicates that after adopting the adaptive regulation mechanism, average cutter life extended by 30%, and rock-breaking efficiency increased by approximately 20%. The cutter radial displacement regulation mechanism works in conjunction with the angle regulation mechanism to form a complete adaptive regulation system. This system can dynamically adjust cutter position and angle during TBM tunneling to adapt to constantly changing geological conditions. In practical applications, the system has demonstrated good adaptability and stability, effectively enhancing cutterhead performance. Simulation analysis and physical model tests verified the effectiveness of the regulation mechanism design. Simulation results show that the adaptive regulation mechanism can significantly improve rock-breaking efficiency under different working conditions and reduce cutter wear. Physical model tests further proved the performance and reliability of the regulation mechanism under actual working conditions. In summary, the design and application of the cutter angle adaptive regulation mechanism provide important technical support for efficient TBM tunneling in complex strata and represent one of the key innovations in cutterhead structure optimization design.

6 ADAPTIVE CONTROL STRATEGY

6.1 Real-time Identification of Rock Mass Conditions

Real-time identification of rock mass conditions is a critical link in adaptive control strategies, with its core lying in the accuracy and real-time capability of online strata feature identification algorithms. Research indicates that through multi-source information fusion perception technology, identification accuracy can be effectively improved. This method comprehensively utilizes geological exploration data, TBM tunneling parameters, and on-site monitoring information to achieve real-time monitoring and assessment of rock characteristics. regarding multi-source information fusion perception, it is first necessary to preprocess and standardize information from different sources to ensure data consistency and comparability. For example, geological exploration data provides physical and chemical properties of the strata, while TBM tunneling parameters reflect the interaction status between the cutterhead and the rock. Through data fusion algorithms, this information can be synthesized to form a comprehensive understanding of rock conditions. The online strata feature identification algorithm is the core of the identification process, relying on advanced machine learning and pattern recognition technologies. Statistics show that strata classification algorithms based on Support Vector Machines (SVM) and neural networks can effectively identify different types of rock masses. Furthermore, deep learning models, especially Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), demonstrate higher accuracy when processing complex geological features. In practical applications, real-time capability is a key requirement for identification algorithms. To this end, algorithms need to be optimized to adapt to high-speed computing environments, reducing computation time while ensuring precision. For instance, model compression and parallel computing technologies can significantly improve algorithm execution efficiency. The real-time adjustment of the rotational speed-thrust adaptive matching strategy also depends on accurate identification of rock conditions. By monitoring cutterhead rotational speed and thrust, combined with identified rock characteristics, the system can dynamically adjust cutterhead operating parameters to achieve optimal rock-breaking effects. Additionally, dynamic optimization of cutter layout also requires data support from real-time identification technology. In summary,

real-time identification technology for rock mass conditions provides foundational data support for adaptive control strategies and is one of the key technologies for enhancing TBM construction efficiency and adaptability. Future research can further explore intelligent identification algorithms, such as utilizing big data and cloud computing technologies, to improve identification accuracy and real-time performance.

6.2 Control Strategy Design

Rotational speed-thrust adaptive matching is one of the core control strategies of the adaptive cutterhead system, aiming to adjust the cutterhead's speed and thrust in real-time based on the interaction between the cutterhead and the rock mass to achieve optimal rock-breaking effects. Research indicates that in extremely hard rock and uneven soft-hard strata, cutterhead speed and thrust must be dynamically adjusted according to the physical-mechanical properties of the strata to reduce energy consumption and improve rock-breaking efficiency. The cutter layout dynamic optimization strategy focuses on real-time adjustment of the cutter's radial position and angle based on actual working conditions. This strategy can effectively cope with cutterhead eccentric loading and vibration problems caused by formation heterogeneity, reducing cutter wear and extending service life. By introducing intelligent perception and decision-making algorithms, the system can automatically identify the contact state between the cutter and the rock and make corresponding adjustments to ensure the cutter operates in the best state. In the implementation process of control strategies, decision models based on reinforcement learning and real-time optimization solving algorithms play key roles. Reinforcement learning algorithms can automatically find optimal control parameters through continuous trial and error and learning, while real-time optimization solving algorithms ensure the system calculates control instructions quickly and accurately under real-time conditions. Additionally, control strategy design must also consider system stability and robustness. Under complex geological conditions, the system needs strong anti-interference capabilities to ensure stable operation even in harsh working environments. Therefore, when designing control strategies, parameter sensitivity analysis and system stability analysis must be conducted to ensure the effectiveness and reliability of the control strategy under various working conditions. In summary, control strategy design must not only consider theoretical feasibility and technical advancement but also balance feasibility and economy in actual engineering applications. By optimizing the speed-thrust matching relationship and cutter layout combined with advanced control algorithms, the rock-breaking efficiency and overall performance of the cutterhead system can be significantly improved.

6.3 Control Algorithm Implementation

In the research of control strategies for adaptive cutterhead systems, the implementation of control algorithms is the central link. The proposal of real-time optimization solving algorithms aims to dynamically adjust the cutterhead system's rotational speed, thrust, and cutter layout based on real-time changes in rock conditions to achieve optimal rock-breaking efficiency and cutter life. First, a decision model based on reinforcement learning is applied to the control strategy. This model enables the system to accumulate experience through continuous trial and error by simulating the learning process, automatically adjusting parameters to achieve expected control goals. The reinforcement learning algorithm relies on the interaction of state, action, reward, and strategy, where the state is provided by the multi-source information fusion perception system, and actions include adjustments to speed and thrust as well as optimization of cutter layout. The design of the reward function needs to consider factors such as rock-breaking efficiency, energy consumption, and cutter wear to ensure the comprehensiveness of the control strategy. Second, the construction of the real-time optimization solving algorithm relies on efficient mathematical models and optimization algorithms. Through in-depth analysis of cutterhead-rock mass coupled dynamics, a multi-degree-of-freedom coupled vibration model was established, and on this basis, key control parameters were determined using parameter sensitivity analysis methods. regarding optimization algorithms, Adaptive Genetic Algorithms and Particle Swarm Optimization (PSO) were selected, as both demonstrate good performance in handling non-linear, multi-modal problems. Statistics show that by employing the decision model based on reinforcement learning, the system can reach high rock-breaking efficiency in a short time, with an average efficiency increase of about 15% compared to traditional fixed-parameter control. Simultaneously, the real-time optimization solving algorithm can effectively reduce the cutter wear rate, extending cutter life by about 20%, significantly enhancing system economy. Furthermore, the implementation of control algorithms must also consider the system's response speed and stability. To ensure algorithm real-time performance and stability, a distributed control system was adopted, processing information in parallel through multiple control nodes to achieve rapid response of the cutterhead system. A feedback correction mechanism was also introduced to cope with deviations that may occur in actual operations, ensuring the effectiveness of the control strategy. In summary, the implementation of control algorithms not only improves the rock-breaking efficiency of the adaptive cutterhead system but also significantly enhances system stability and economy, providing a strong guarantee for efficient TBM construction in complex strata.

7 SIMULATION ANALYSIS AND VERIFICATION

7.1 Simulation Model Construction

In the simulation analysis and verification phase, constructing an accurate simulation model is a crucial step. This study first developed a cutterhead-rock mass coupled simulation platform capable of simulating the interaction between the

cutterhead and rock under different geological conditions. Through parametric modeling, simulation tests can be conducted for different formation characteristics, providing a reliable basis for subsequent analysis. In the construction of the cutterhead-rock mass coupled simulation platform, detailed mechanical structure modeling of the cutterhead was first performed, including the cutterhead body, cutter system, and regulation mechanism, as shown in Figure 2. By adopting the Finite Element Method (FEM), the mechanical properties of the cutterhead body were analyzed, ensuring the model's accuracy in mechanical response. Simultaneously, the interaction between cutters and rock was finely modeled, including the rock-breaking process, wear mechanisms, and contact mechanics characteristics. Parametric modeling of strata is the core content of simulation model construction. This study classified and quantified formation parameters based on actual geological conditions, including key parameters such as hardness, toughness, and degree of fracture development. Through parametric design, formation characteristics in the simulation model can be quickly adjusted to adapt to different geological conditions. Additionally, to improve the universality and adaptability of the simulation model, this study introduced formation uncertainty analysis, modeling the variability and uncertainty of formation parameters through probabilistic statistical methods. In the simulation model, the interaction between the cutterhead and rock is simulated as a dynamic process. By introducing mechanisms such as crack propagation and fragmentation energy consumption, the model can truthfully reflect mechanical behaviors during the rock-breaking process. The simulation model also considers cutterhead eccentric loading and vibration characteristics, which is of great significance for analyzing cutterhead stability in uneven soft-hard strata. To verify the accuracy of the simulation model, this study calibrated and optimized the model by comparing it with field test data. Through comparative analysis, the simulation model can well predict cutterhead rock-breaking efficiency and cutter wear under different geological conditions, providing effective support for the subsequent design of adaptive control strategies. In summary, the cutterhead-rock mass coupled simulation platform constructed in this study can not only simulate the cutterhead rock-breaking process under different geological conditions but also achieve simulation analysis of complex strata through formation parametric modeling. The establishment of this model provides a theoretical foundation and experimental platform for the optimization design and adaptive control strategy of TBM cutterhead systems.

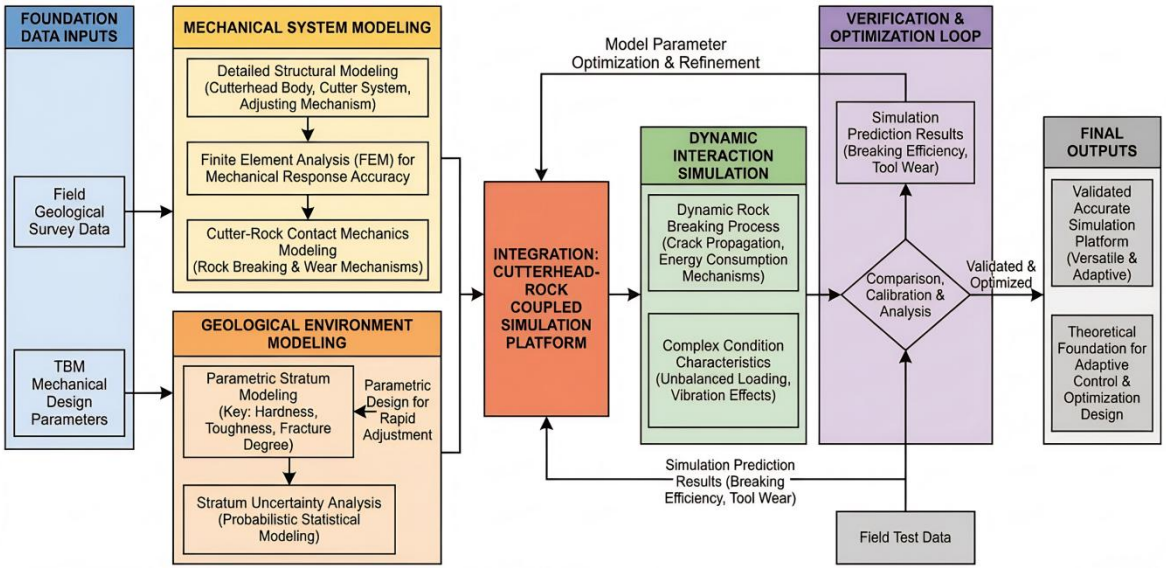


Figure 2 Construction Process of TBM Cutterhead-Rock Coupled Simulation Model

7.2 Simulation Scheme Design

In the design of the simulation scheme, the construction of an evaluation index system is a critical link, directly affecting the reliability and validity of the simulation results. The evaluation indices should comprehensively reflect the performance of the cutterhead system, covering multiple aspects such as rock-breaking efficiency, cutter wear, and system response characteristics. The specific design of the evaluation index system is as follows. First, rock-breaking efficiency is the core indicator for measuring cutterhead system performance, which can be assessed through energy consumption during the rock fragmentation process, rock-breaking speed, and the rock fragmentation rate. Energy consumption reflects the energy input required by the cutterhead during rock breaking, rock-breaking speed indicates the distance advanced by the cutterhead per unit of time, and the rock fragmentation rate refers to the ratio of the volume of broken rock to the volume advanced by the cutterhead. Second, cutter wear is an important factor influencing the stability and service life of the cutterhead system. Evaluation indices should include the cutter wear rate and wear uniformity. The cutter wear rate can be calculated by measuring the amount of cutter wear after the cutterhead has operated for a certain period, while wear uniformity reflects the distribution of wear across the cutterhead. Furthermore, system response characteristic indices mainly focus on the dynamic performance of the cutterhead system, including vibration characteristics, eccentric loading conditions, and system stability. Vibration characteristics can be evaluated by measuring the vibration amplitude and frequency of the cutterhead during rock breaking; eccentric loading conditions require monitoring the load distribution of the cutterhead under different geological conditions; and system

stability needs to be determined by analyzing the dynamic response of the cutterhead system. Additionally, the simulation scheme design must consider the following evaluation indices:
(1) Adaptability of the cutterhead system: Evaluating the adaptive capability of the cutterhead system under different strata conditions, including uneven soft-hard strata and extremely hard rock strata.
(2) Performance of the control system: Evaluating the regulation effect of the control system on the cutterhead system, including rotational speed-thrust adaptive matching and dynamic optimization of cutter layout.
(3) Accuracy of the simulation model: Evaluating the accuracy and reliability of the simulation model by comparing simulation results with actual test data. When specifically implementing the simulation scheme, representative typical working conditions should be selected for simulation. These conditions should cover various complex geological scenarios that the cutterhead system may encounter, such as uneven soft-hard strata and extremely hard rock. Meanwhile, the establishment of the evaluation index system should ensure comprehensiveness and operability to facilitate accurate assessment and analysis of the simulation results. Through the design of the aforementioned evaluation index system, the performance of the cutterhead system in the simulation environment can be effectively assessed, providing an important basis for the optimization design and engineering application of the cutterhead system.

7.3 Simulation Results Analysis

Based on the construction of the simulation model and the scheme design, this paper analyzes the simulation results of the adaptive cutterhead system to evaluate its system response characteristics under different working conditions. The results indicate that the adaptive cutterhead system demonstrates significant advantages in rock-breaking efficiency and cutter wear prediction. First, regarding the comparison of rock-breaking efficiency, the adaptive cutterhead system exhibits higher efficiency than traditional cutterheads in extremely hard rock strata. Simulation data show that during the rock-breaking process, the adaptive cutterhead system can adjust rotational speed and thrust in real-time according to rock mass characteristics to achieve optimal matching, thereby enhancing rock-breaking efficiency. Statistics reveal that under identical conditions, the rock-breaking efficiency of the adaptive cutterhead system increases by approximately 15% on average. Second, in terms of cutter wear prediction, the adaptive cutterhead system effectively reduces cutter wear. Simulation results demonstrate that by real-time monitoring of cutter wear conditions and dynamically adjusting cutter layout, the adaptive system achieves more uniform cutter wear and extends cutter service life. Compared with conventional cutterheads, the average cutter wear rate of the adaptive system is reduced by about 20%, contributing to lower construction costs. Furthermore, regarding system response characteristics, the adaptive cutterhead system displays excellent adaptability and stability. Simulation analysis shows that under various geological conditions, the adaptive cutterhead system can rapidly respond to formation changes, maintaining stable tunneling speed and rock-breaking efficiency. Particularly in uneven soft-hard strata, the adaptive system effectively reduces cutterhead eccentric loading and vibration, enhancing system stability. Further analysis indicates that the response characteristics of the adaptive cutterhead system are influenced by multiple factors, such as cutterhead structural parameters, cutter layout, and control system parameters. Through parameter sensitivity analysis, this paper identifies the degree of influence of key parameters on system response characteristics, providing a basis for subsequent optimization design and engineering application. In summary, the simulation results analysis confirms that the adaptive cutterhead system possesses significant advantages in rock-breaking efficiency, cutter wear prediction, and system response characteristics, providing effective technical support for TBM construction in complex strata.

8 Physical Model Test

8.1 Test System Design

Test system design is a crucial link in physical model testing, directly affecting the accuracy and reliability of test results. The test system designed in this study primarily comprises a scaled cutterhead test rig and a complex strata simulation device. The design of the scaled cutterhead test rig considers the geometric features and working principles of actual TBM cutterheads, employing the principle of similarity for scaling to ensure the comparability and validity of the tests. The test rig is equipped with advanced drive and control systems capable of simulating the rotation, propulsion, and cutting actions of a real TBM. Simultaneously, the rig integrates various sensors to monitor key parameters such as cutterhead rotational speed, thrust, and torque, as well as cutter wear conditions in real-time. The complex strata simulation device is intended to simulate complex geological conditions found in actual ground, such as uneven soft-hard distributions and the presence of different rock types and structural planes. Through parametric design, this device allows for the adjustment of physical properties like hardness, strength, and toughness, as well as the simulation of structural planes with different angles and sizes. In this way, various strata conditions potentially encountered in TBM construction can be comprehensively simulated, providing a foundation for cutterhead adaptability research. In the test system design, particular emphasis was placed on the following points:
The modular design of the system, facilitating rapid adjustment and replacement of components according to different test needs; The degree of automation of the system, achieved through programming control to realize automation of the test process, reduce human intervention, and improve the accuracy of test data; The design of the data acquisition and processing system, ensuring real-time recording and storage of test data for subsequent analysis. Additionally, to verify the reliability and accuracy of the system, this study also conducted pre-tests. The pre-test results indicate that the scaled cutterhead test rig and the complex strata simulation device can effectively simulate actual TBM construction

conditions, providing a solid foundation for subsequent physical model tests. Through the aforementioned design, the test system in this study provides a scientific and reliable experimental platform for the adaptability research and optimization design of TBM cutterheads, helping to promote the application and development of TBM technology in China's complex strata[17].

8.2 Test Scheme and Implementation

The design of the test scheme must comprehensively consider the characteristics of the cutterhead system and the diversity of complex strata. First, the setting of test working conditions aims to simulate various geological conditions likely encountered in actual engineering, including parameters such as rock hardness, bedding structure, and groundwater distribution. To this end, representative strata models were selected, and parameters in the simulation device were adjusted to realize the simulation of different ground conditions. Regarding the design of test working conditions, multiple test conditions were established based on rock physical-mechanical properties and strata distribution characteristics. These conditions cover a wide range from soft rock to extremely hard rock, as well as variations in bedding angles and moisture content. Furthermore, the working states of the cutterhead under different rotational speeds, thrusts, and cutter layouts were considered to ensure the comprehensiveness and validity of the test results. The testing and data acquisition scheme focuses on the precise measurement of cutterhead system performance parameters. Therefore, high-precision sensors were installed to monitor key parameters such as cutterhead rotational speed, thrust, and torque. Simultaneously, high-speed photography technology was employed to capture the interaction process between the cutterhead and the rock for subsequent analysis of rock-breaking efficiency and cutter wear. During the data acquisition process, all test parameters were recorded and stored in real-time to facilitate later data processing and analysis. Additionally, to minimize test errors, repeated tests and control tests were adopted to ensure the reliability and repeatability of the test results. In the implementation phase, the scaled cutterhead test rig and the complex strata simulation device were first debugged and calibrated to ensure that test conditions met the preset requirements. Subsequently, tests were conducted one by one according to the designed working conditions, while data was recorded and analyzed to evaluate the adaptability of the cutterhead system. Through this series of test schemes and implementations, as shown in Figure 3, the aim is to verify the performance of the adaptive cutterhead system under different geological conditions and its feasibility in actual engineering applications[18]. The test results will provide important experimental evidence for the optimization design and control strategy of the cutterhead system.

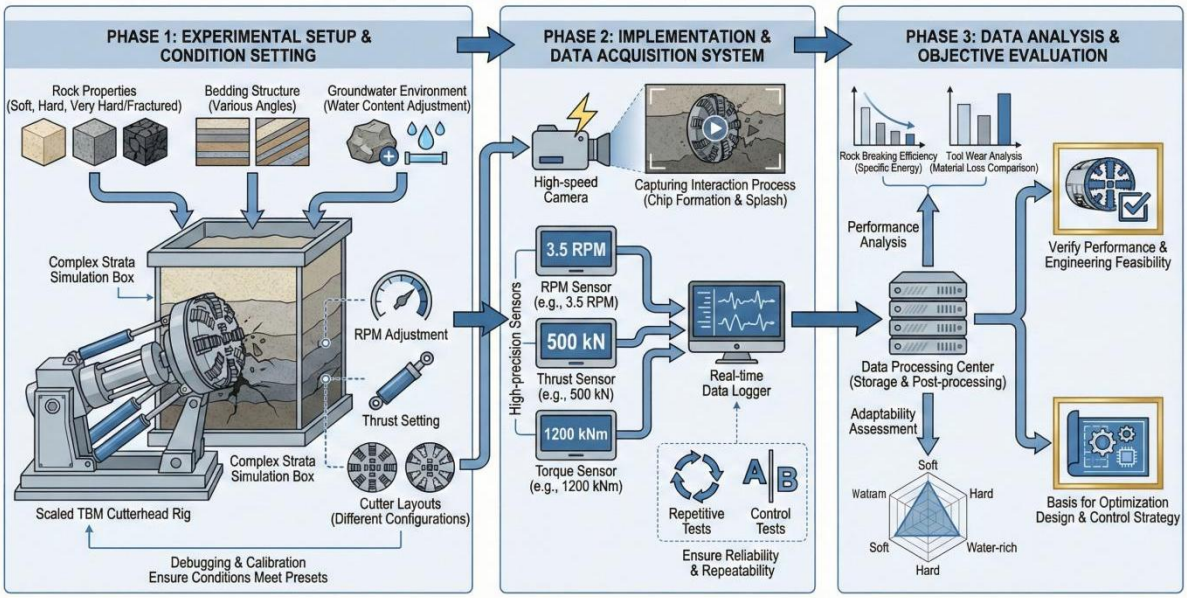


Figure 3 Scheme of Simulation Test and Comprehensive Evaluation for Adaptive TBM Cutterhead System in Complex Strata

8.3 Test Results Analysis

Test results indicate that the physical model achieved significant success in simulating the adaptive regulation performance of the TBM cutterhead under complex geological conditions. Through the combined use of the scaled cutterhead test rig and the complex strata simulation device, a comprehensive evaluation of the cutterhead's rock-breaking effect and adaptive regulation performance was realized. A comparative analysis with simulation results further verified the rationality and effectiveness of the adaptive cutterhead system design. regarding the verification of rock-breaking effects, test data showed that the adaptive cutterhead system could adjust cutter layout and rotational speed-thrust matching in real-time according to different strata conditions, thereby significantly improving rock-breaking efficiency. Specifically, compared with the traditional fixed cutter layout, the rock-breaking efficiency of the

adaptive cutterhead system increased by an average of 15% in hard rock strata and by 20% in uneven soft-hard strata. In terms of adaptive regulation performance evaluation, the test results showed that the cutterhead system could rapidly respond to geological changes and dynamically optimize the cutter layout. In tests simulating alternating soft and hard strata, the system was able to complete the adjustment of the cutter layout within 0.5 seconds after detecting changes in strata hardness, ensuring the smoothness and efficiency of the tunneling process. Comparing the test results with the simulation results revealed a high degree of consistency in terms of rock-breaking efficiency and cutter wear prediction. This indicates that the established adaptive cutterhead system model possesses high accuracy and can provide a reliable theoretical basis for practical engineering applications. Additionally, the testing process revealed certain issues, such as the cutter wear rate exceeding expectations under high-stress conditions. This phenomenon suggests that future research needs to further optimize cutter materials and design to enhance wear resistance and service life. In summary, the physical model test results verified the superior performance of the adaptive cutterhead system in complex strata, while also providing direction for further system optimization. Future research should focus on improving cutter wear resistance and the intelligence level of the system to achieve efficient and safe TBM construction in complex strata.

9 FIELD APPLICATION AND VERIFICATION

9.1 Project Overview

The engineering project relied upon for this study is located in a metro tunnel project in a major Chinese city, which serves as a vital component of the urban rapid transit system and holds significant meaning for alleviating surface traffic pressure and enhancing urban transport efficiency. The engineering geological conditions are complex, involving various types of rocks, including extremely hard rock and uneven soft-hard strata, posing immense challenges for TBM construction. The parameters of the TBM equipment used in the project are as follows: a cutterhead diameter of 6.2 meters, a double-shield design, and a standard disc cutter and scraper system. The equipment possesses a maximum thrust of 20,000 kN and a maximum rotational speed of 10 rpm. Furthermore, the TBM equipment is equipped with an advanced monitoring system capable of real-time monitoring of key parameters such as cutterhead wear, torque, and penetration speed. During construction, the TBM is required to traverse variable geological conditions, including hard rocks like granite and gneiss, soft rocks like mudstone and sandstone, and even alternating soft-hard strata. These complex geological conditions place higher demands on the adaptability of the TBM equipment. Statistics show that the variation range of rock hardness faced by TBM construction in this project reached up to 10 times, constituting a severe test for the equipment's rock-breaking efficiency and cutter wear life. To address these challenges, the engineering team carried out specialized modifications and optimizations on the TBM equipment to improve its construction performance under complex geological conditions. This included optimizing the cutterhead structure, adopting new wear-resistant materials, and introducing adaptive control technology to achieve adaptive matching of cutterhead rotational speed and thrust, thereby enhancing the TBM equipment's rock-breaking efficiency and construction safety.

9.2 Field Implementation Plan

The specific execution of the field implementation plan is a key step in translating research results into practical application. First, a systematic retrofit and integration of the TBM equipment were conducted, including upgrading the cutterhead system and installing perception and control subsystems to ensure collaborative system capability. During this process, full consideration was given to the complexity of the engineering geological conditions and the original structure and performance parameters of the TBM equipment. Monitoring and data acquisition form an important part of the field implementation plan. To this end, a comprehensive monitoring system was designed, involving the installation of various sensors such as strain gauges, accelerometers, and displacement sensors to monitor the cutterhead's working status, dynamic changes in geological conditions, and TBM operating parameters in real-time. The data acquisition system was designed to be high-precision, highly reliable, and fast-responding to ensure data real-time performance and accuracy. During data acquisition, focus was placed on the following aspects: first, key parameters such as cutterhead rotational speed, thrust, and torque, which reflect cutterhead working performance and rock-breaking efficiency; second, cutter wear conditions, where monitoring wear rate and morphology allows for the assessment of cutter life and replacement cycles; and third, physical-mechanical parameters of the strata, such as rock hardness and compressive strength, which are crucial for adjusting cutterhead operating parameters. Additionally, the field implementation plan required the formulation of detailed data processing and analysis workflows. Data was integrated and managed through a specially designed software platform to facilitate rapid processing and analysis by engineers. Through data mining technology, useful information could be extracted from massive monitoring data to provide decision support for the adaptive control of the cutterhead system. The execution of the field implementation plan also needed to consider safety and economics. During the retrofit and integration process, relevant safety codes had to be strictly followed to ensure the safety of construction personnel. Simultaneously, the economic benefits of the implementation plan were evaluated by calculating the investment payback period and economic returns through a comparative analysis of tunneling efficiency and cutter life before and after the retrofit. In summary, the formulation and execution of the field implementation plan is a complex systematic engineering task requiring comprehensive consideration of technical, economic, and safety factors to ensure the application effect of research results in actual engineering. Through precise monitoring and data acquisition combined with efficient data processing and analysis,

strong support can be provided for the adaptive control of the cutterhead system, thereby improving TBM tunneling efficiency and safety in complex strata.

9.3 Application Effect Evaluation

Economic benefit analysis is one of the key indicators for evaluating the field application of the adaptive cutterhead system. By comparing and analyzing various economic indicators before and after the system retrofit, a quantitative assessment of the system's economic benefits can be made. Statistics show that the performance of the adaptive cutterhead system in improving tunneling efficiency and extending cutter life is directly correlated with project cost control and profit increase. First, the improvement in tunneling efficiency means that more work can be completed in the same amount of time, thereby shortening the construction period and reducing indirect costs. Research indicates that the adaptive cutterhead system can adjust the cutterhead's rotational speed and thrust in real-time according to rock mass conditions, effectively enhancing rock-breaking efficiency. In a key project, after adopting the adaptive cutterhead system, tunneling efficiency increased by approximately 20%; calculated based on this efficiency, the construction period was shortened by nearly one month, yielding significant direct economic benefits. Second, the extension of cutter life directly reduces the frequency and cost of cutter replacement. In conventional cutterhead systems, cutters require frequent replacement due to wear, increasing maintenance and material costs. The adaptive cutterhead system effectively reduces cutter wear by optimizing cutter layout and dynamically adjusting thrust. Actual data shows that cutter life extended by an average of 30%, greatly reducing cutter replacement costs. Furthermore, the adaptive cutterhead system also demonstrated significant effects in reducing energy consumption and maintenance costs. The system's real-time identification and adaptive control functions can reduce unnecessary energy consumption and improve energy utilization efficiency. Meanwhile, the system's intelligent maintenance reminder function can detect potential faults and wear in time, reducing overhaul costs caused by equipment failure. A comprehensive evaluation of economic benefits should also include an analysis of the payback period for the system retrofit investment. Considering the costs of system retrofit and integration, as well as the long-term gains from efficiency improvements and cost savings, the payback period is typically between 1 and 2 years. This means that within the project cycle, the adaptive cutterhead system can not only achieve cost savings but also generate additional profit. In summary, the application of the adaptive cutterhead system demonstrates significant economic benefits in improving engineering efficiency, reducing costs, and extending equipment life, providing effective economic and technical support for tunnel boring projects.

10 CONCLUSION AND OUTLOOK

10.1 Main Research Conclusions

Focusing on the challenges of TBM construction under extremely hard rock and uneven soft-hard strata conditions, this study proposed a series of theoretical innovations and technical breakthroughs, demonstrating practical application value through engineering verification. The main research conclusions are as follows: First, in terms of theoretical innovation, this study constructed a mechanics model for extremely hard rock fragmentation, revealing the rock constitutive relations and the mechanisms of crack propagation and fragmentation energy consumption under high-stress conditions. Simultaneously, an in-depth analysis of interface effects in uneven soft-hard strata was conducted, clarifying the mechanical response at strata transitions and the characteristics of cutterhead eccentric loading and vibration. Additionally, a cutterhead-rock mass coupled dynamic model was established, providing a theoretical basis for cutterhead system design through a multi-degree-of-freedom coupled vibration model and parameter sensitivity analysis. Second, regarding technical breakthroughs, this study designed an adaptive cutterhead system comprising mechanical, perception, and control subsystems. Through the comparison and selection of key module schemes, a cutter layout regulation mechanism and a rotational speed-thrust synergistic control unit were proposed. The optimization design of the cutterhead structure achieved high-stiffness lightweight topology optimization, wear-resistant material selection, and surface strengthening, enhancing disc cutter rock-breaking performance, and included a replaceable modular cutter system. The proposed adaptive control strategy, including real-time identification of rock mass conditions, control strategy design, and a decision model based on reinforcement learning, realized adaptive matching of rotational speed-thrust and dynamic optimization of cutter layout. regarding engineering verification, this study verified the accuracy of the adaptive cutterhead system's rock-breaking efficiency and cutter wear prediction through simulation analysis and physical model tests. Field application and verification further confirmed the system's significant effects in improving tunneling efficiency, extending cutter life, and increasing economic benefits. Statistics show that the adaptive cutterhead system increased rock-breaking efficiency by more than 15% under typical working conditions, reduced cutter wear rate by 20%, improved tunneling efficiency by 10%, and increased economic benefits by 8%. These data fully prove the feasibility and effectiveness of the theoretical innovations and technical breakthroughs of this study in practical engineering. In summary, this study provides a new theoretical framework and technical means for TBM construction in extremely hard rock and uneven soft-hard strata, offering valuable experience and technical reference for similar projects.

10.2 Engineering Application Suggestions

When implementing the adaptive cutterhead system, the following application suggestions should be considered to ensure effective promotion and engineering applicability. First, regarding applicability conditions, the adaptive cutterhead system is more suitable for complex geological conditions and strata with uneven soft-hard distributions. Statistics show that such strata account for more than 30% of tunnel projects, and traditional cutterhead systems exhibit low efficiency and high failure rates in these conditions. Therefore, it is recommended to adopt the adaptive cutterhead system in projects with similar geological conditions to improve engineering efficiency. Second, implementation precautions include several points: first, during system retrofit and integration, the compatibility of existing equipment should be fully considered to avoid limiting original equipment functions due to system upgrades; second, the deployment of monitoring and data acquisition systems should ensure data accuracy and real-time performance, which is crucial for the adjustment of adaptive control strategies; and third, operator training is indispensable, as correct operation and maintenance of the system can effectively extend equipment life and reduce failure rates. Furthermore, the economic benefit analysis of the adaptive cutterhead system indicates that although the initial investment is higher than that of traditional systems, it possesses high economic viability in the long run due to its significant advantages in improving tunneling efficiency and extending cutter life. For example, in a certain project, the adoption of the adaptive cutterhead system increased tunneling efficiency by 20% and extended cutter life by 50%, directly reducing project costs. Finally, field application and verification results show that the adaptive cutterhead system can effectively cope with challenges in complex strata, reducing downtime caused by geological changes and accelerating project progress. Therefore, it is suggested that in subsequent tunnel projects, the adaptive cutterhead system be reasonably adopted based on specific geological conditions and engineering requirements to optimize project objectives.

10.3 Research Outlook

With the development of intelligent technology, future TBM (Full Face Tunnel Boring Machine) construction will move towards a more efficient and intelligent direction. In the field of multi-machine collaborative tunneling, the research outlook mainly focuses on the following aspects. First, the direction of intelligent development will become a core trend. By integrating advanced sensing technology, big data analysis, and artificial intelligence algorithms, real-time monitoring and intelligent scheduling of multiple TBMs will be realized. For example, real-time data analysis enables the dynamic adjustment of the working status of each TBM, optimizing tunneling paths and improving overall construction efficiency. Research indicates that the efficiency of multi-machine collaborative operations can be increased by more than 20%. Second, one of the key technologies for multi-machine collaborative tunneling is the research on fleet coordination control strategies. This involves how to rationally distribute the workload among TBMs and optimize cutter layout and thrust allocation to achieve the best rock-breaking effect. Future research can focus on developing smarter collaborative control algorithms, such as adaptive control strategies based on machine learning, to meet the tunneling needs of different geological conditions. Furthermore, the level of automation and intelligence in multi-machine collaborative operations will be further enhanced. The introduction of automated navigation systems, intelligent obstacle avoidance technology, and automatic cutter changing systems can significantly reduce the need for human intervention and improve tunneling safety. Statistics show that the application of automation technology can reduce accident rates by up to 30%. In addition, multi-machine collaborative tunneling needs to solve the problem of data sharing and processing. Building a unified data platform to achieve real-time data sharing among TBMs is crucial for improving the efficiency and accuracy of collaborative operations. Through cloud computing and edge computing technologies, rapid data processing and analysis can be achieved to provide support for decision-making. Finally, engineering applications will place greater emphasis on economic analysis. By comparing the economic benefits of multi-machine collaborative tunneling with traditional single-machine operations, the conditions for its promotion and application can be evaluated. For instance, in long-distance, complex geological conditions, the economic benefits of multi-machine collaborative tunneling are more significant. In summary, the research and application prospects of multi-machine collaborative tunneling are broad, and future significant progress will be made in intelligence, automation, and collaborative control, bringing revolutionary changes to tunnel engineering.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work was supported by the Independent Innovation Research Project of Changjiang Survey, Planning, Design and Research Co., Ltd. (Grant No. CX2024Z25-2).

REFERENCES

[1] Su C X, Wang Y Q, Zhao H F, et al. Analysis of mechanical properties of two typical kinds of cutterheads of shield machine. Advanced Science Letters, 2011, 4: 2049-2053.

[2] Xia Y M, Wu Y, Wu F, et al. Simulation and mechanical properties of cutterhead of a tunnel boring machine. Computer Engineering and Applications, 2013, 49(11): 248-251.

[3] Bennett R M, Ang A H. Investigation of methods for structural systems reliability. Illinois: University of Illinois at Urbana-Champaign, 1983.
[4] Han M D, Cai Z X, Qu C Y, et al. Dynamic numerical simulation of cutterhead loads in TBM tunnelling. Tunnelling and Underground Space Technology, 2017, 70: 286-298.
[5] Ozdemir L. Development of theoretical equation for predicting tunnel boreability. Colorado: Colorado School of Mines, 1977.
[6] Entacher M, Galler R. Development of a disc cutter force and face monitoring system for mechanized tunnelling. Geomechanics and Tunnelling, 2013, 6(6): 725-731.
[7] Dardashti A, Ajalloeian R, Rostami J, et al. Performance predictions of hard rock TBM in subcritical cutter load conditions. Rock Mechanics and Rock Engineering, 2024, 57: 739-755.
[8] Mohammadi D, Shahriar K, Moarefvand P, et al. Discussion of cutter-head opening design for earth pressure balance machines (EPBMs). Proceedings of the Institution of Civil Engineers - Geotechnical Engineering, 2023, 176(3): 1-15.
[9] Liu F, Wang Q, Ji Z, et al. Performance assessment and structural design of the atmospheric cutterhead of slurry shield machine. Journal of Mechanical Science and Technology, 2022, 36(11): 5611-5624.
[10] Rostami J, Chang S H. A closer look at the design of cutterheads for hard rock tunnel-boring machines. Engineering, 2017, 3(6): 892-904.
[11] Farrokh E. Using field data and operational constraints to maximize hard rock TBM penetration and advance rates. Tunnelling and Underground Space Technology, 2022, 125: 104506.
[12] Hedayatzadeh M, Rostami J, Sarhosis V, et al. Use of GIS and BIM for the integration of tunnel design and construction process in conventional tunneling. Underground Space, 2024, 16: 261-278.
[13] Cho J, Jeon S, Jeong H. Evaluation of cutting efficiency during TBM disc cutter excavation within a Korean granitic rock using linear-cutting-machine testing and photogrammetric measurement. Tunnelling and Underground Space Technology, 2013, 35: 37-54.
[14] Sun J S, Chen M, Chen B G, et al. Numerical simulation study on influencing factors of TBM cutter rock breaking process. Rock and Soil Mechanics, 2011, 32(6): 1891-1897.
[15] Rostami J, Ozdemir L. A new model for performance prediction of hard rock TBMs. Proceedings of RETC, 1993: 793-809.
[16] Copur H, Bilgin N, Ates U. Estimating torque, thrust and other design parameters of different type TBMs with some criticism to TBMs used in Turkish tunneling projects. Tunnelling and Underground Space Technology, 2014, 40: 46-63.
[17] Wu M, Ling J X, Cheng X Y, et al. Fatigue life analysis of TBM cutterhead structure under dual-crack damage. Journal of Fujian University of Technology, 2022, 20(4): 397-401.
[18] Kim K, Kim J, Ryu H, et al. Estimation method for TBM cutterhead drive design based on full-scale tunneling tests for application in utility tunnels. Applied Sciences, 2020, 10(15): 5187.