

ISSN 1854-6250

APEM
journal

Advances in Production Engineering & Management

Volume 19 | Number 3 | September 2024





University of Maribor

Published by CPE
apem-journal.org

Advances in Production Engineering & Management

Identification Statement

	ISSN 1854-6250 Abbreviated key title: Adv produc engineer manag Start year: 2006 ISSN 1855-6531 (on-line)
	Published quarterly by Chair of Production Engineering (CPE), University of Maribor Smetanova ulica 17, SI – 2000 Maribor, Slovenia, European Union (EU) Phone: 00386 2 2207522, Fax: 00386 2 2207990 Language of text: English APEM homepage: apem-journal.org University homepage: www.um.si

APEM Editorial

Editor-in-Chief

Miran Brezocnik

editor@apem-journal.org, info@apem-journal.org
University of Maribor, Faculty of Mechanical Engineering Smetanova ulica 17, SI – 2000 Maribor, Slovenia, EU

Desk Editor

Martina Meh

desk1@apem-journal.org

Janez Gotlih

desk2@apem-journal.org

Website Technical Editor

Lucija Brezocnik

desk3@apem-journal.org

Editorial Board Members

Eberhard Abele, Technical University of Darmstadt, Germany
Bojan Acko, University of Maribor, Slovenia
Joze Balic, University of Maribor, Slovenia
Agostino Bruzzone, University of Genoa, Italy
Borut Buchmeister, University of Maribor, Slovenia
Ludwig Cardon, Ghent University, Belgium
Nirupam Chakraborti, Indian Institute of Technology, Kharagpur, India
Edward Chlebus, Wroclaw University of Technology, Poland
Igor Drstvensek, University of Maribor, Slovenia
Illes Dudas, University of Miskolc, Hungary
Mirko Ficko, University of Maribor, Slovenia
Vlatka Hlupic, University of Westminster, UK
David Hui, University of New Orleans, USA
Pramod K. Jain, Indian Institute of Technology Roorkee, India
Isak Karabegović, University of Bihać, Bosnia and Herzegovina
Janez Kopac, University of Ljubljana, Slovenia

Changsong Ma, Geely University of China, China
Qingliang Meng, Jiangsu University of Science and Technology, China
Lanndon A. Ocampo, Cebu Technological University, Philippines
Iztok Palcic, University of Maribor, Slovenia
Krsto Pandza, University of Leeds, UK
Andrej Polajnar, University of Maribor, Slovenia
Antonio Pouzada, University of Minho, Portugal
R. Venkata Rao, Sardar Vallabhbhai National Inst. of Technology, India
Rajiv Kumar Sharma, National Institute of Technology, India
Katica Simunovic, J. J. Strossmayer University of Osijek, Croatia
Daizhong Su, Nottingham Trent University, UK
Soemon Takakuwa, Nagoya University, Japan
Nikos Tsourveloudis, Technical University of Crete, Greece
Tomo Udiljak, University of Zagreb, Croatia
Ivica Veza, University of Split, Croatia



Subsidizer: The journal is subsidized by Slovenian Research and Innovation Agency



Creative Commons Licence (CC): Content from published paper in the APEM journal may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Statements and opinions expressed in the articles and communications are those of the individual contributors and not necessarily those of the editors or the publisher. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. Chair of Production Engineering assumes no responsibility or liability for any damage or injury to persons or property arising from the use of any materials, instructions, methods or ideas contained herein.

Published by CPE, University of Maribor.

Advances in Production Engineering & Management is indexed and abstracted in the **WEB OF SCIENCE** (maintained by **Clarivate**): **Science Citation Index Expanded**, **Journal Citation Reports** – Science Edition, **Current Contents** – Engineering, Computing and Technology • **Scopus** (maintained by **Elsevier**) • **Inspec** • **EBSCO**: Academic Search Alumni Edition, Academic Search Complete, Academic Search Elite, Academic Search Premier, Engineering Source, Sales & Marketing Source, TOC Premier • **ProQuest**: CSA Engineering Research Database – Cambridge Scientific Abstracts, Materials Business File, Materials Research Database, Mechanical & Transportation Engineering Abstracts, ProQuest SciTech Collection • **TEMA (DOMA)** • The journal is listed in **Ulrich's** Periodicals Directory and **Cabell's** Directory



University of Maribor
Chair of Production Engineering (CPE)

Advances in Production Engineering & Management

Volume 19 | Number 3 | September 2024 | pp 297–410

Contents

Scope and topics	300
Dynamic scheduling for manufacturing workshops using Digital Twins, Competitive Particle Swarm Optimization, and Siamese Neural Networks	301
Weng, L.L.	
Optimization of cold chain multimodal transportation routes considering carbon emissions under hybrid uncertainties	315
Hou, D.N.; Liu, S.C.	
Enhancing workplace safety and ergonomics with motion capture systems: Present state and a case study	333
Vujica Herzog, N.; Buchmeister, B.; Jaehyun, P.; Kaya, Ö.	
Machine learning for enhancing manufacturing quality control in ultrasonic nondestructive testing: A Wavelet Neural Network and Genetic Algorithm approach	347
Song, W.T.; Huo, L.	
Optimizing obstacle avoidance path planning for intelligent mobile robots in multi-obstacle environments	358
Zhang, H.Q.	
Modelling and optimization of dimensional accuracy and surface roughness in dry turning of Inconel 625 alloy	371
Vukelic, D.; Milosevic, A.; Ivanov, V.; Kocovic, V.; Santosi, Z.; Sokac, M.; Simunovic, G.	
Enhancing calibration accuracy with laser interferometry for high-resolution measuring systems	386
Lipus, L.C.; Acko, B.; Klobucar, R.	
Improved Whale Optimization Algorithm for supply chain financial risk assessment of cloud warehouse platform	395
Zhang, H.; Guo, Y.W.; Hou, Y.; Tang, L.; Deveci, M.	
Calendar of events	408
Notes for contributors	409

Journal homepage: apem-journal.org

ISSN 1854-6250 (print)

ISSN 1855-6531 (on-line)

Published by CPE, University of Maribor.

Scope and topics

Advances in Production Engineering & Management (APEM journal) is an interdisciplinary refereed international academic journal published quarterly by the *Chair of Production Engineering* at the *University of Maribor*. The main goal of the *APEM journal* is to present original, high quality, theoretical and application-oriented research developments in all areas of production engineering and production management to a broad audience of academics and practitioners. In order to bridge the gap between theory and practice, applications based on advanced theory and case studies are particularly welcome. For theoretical papers, their originality and research contributions are the main factors in the evaluation process. General approaches, formalisms, algorithms or techniques should be illustrated with significant applications that demonstrate their applicability to real-world problems. Although the *APEM journal* main goal is to publish original research papers, review articles and professional papers are occasionally published.

Fields of interest include, but are not limited to:

Additive Manufacturing Processes	Machine Learning in Production
Advanced Production Technologies	Machine-to-Machine Economy
Artificial Intelligence in Production	Machine Tools
Assembly Systems	Machining Systems
Automation	Manufacturing Systems
Big Data in Production	Materials Science, Multidisciplinary
Block Chain in Manufacturing	Mechanical Engineering
Computer-Integrated Manufacturing	Mechatronics
Cutting and Forming Processes	Metrology in Production
Decision Support Systems	Modelling and Simulation
Deep Learning in Manufacturing	Numerical Techniques
Discrete Systems and Methodology	Operations Research
e-Manufacturing	Operations Planning, Scheduling and Control
Evolutionary Computation in Production	Optimisation Techniques
Fuzzy Systems	Project Management
Human Factor Engineering, Ergonomics	Quality Management
Industrial Engineering	Risk and Uncertainty
Industrial Processes	Self-Organizing Systems
Industrial Robotics	Smart Manufacturing
Intelligent Manufacturing Systems	Statistical Methods
Joining Processes	Supply Chain Management
Knowledge Management	Virtual Reality in Production
Logistics in Production	

Dynamic scheduling for manufacturing workshops using Digital Twins, Competitive Particle Swarm Optimization, and Siamese Neural Networks

Weng, L.L.^{a,*}

^aSouth China Business College, Guangdong University of Foreign Studies, Guangzhou, P.R. China

ABSTRACT

Flexible manufacturing workshops often encounter scheduling challenges due to complex processes and cumbersome procedures. To address these issues, a dynamic scheduling method is proposed. Initially, a discrete manufacturing workshop scheduling problem model is developed, considering the unique characteristics of the workshop. Digital Twin technology and a Competitive Particle Swarm Optimization algorithm are then integrated to create the scheduling model. Finally, Siamese Neural Networks are incorporated to form a dynamic scheduling mechanism that optimizes disturbance scheduling. The research model demonstrated a quick convergence, efficiently searching for the optimal fitness value using both the Sphere and Griewank functions. In the scheduling objective function test, the model achieved a maximum completion time of 244.8 minutes, the shortest time compared to similar technologies. In Siamese Neural Network experiments, the model successfully suppressed the influence of disturbances, maintaining optimal scheduling performance. Without adjustments for disturbances, the maximum completion time was 58.5 minutes. After optimization, it decreased to 54.2 minutes. These results demonstrate the effective application of the proposed technology in workshop scheduling. The findings provide valuable technical insights for the application of intelligent technologies in workshop scheduling optimization.

ARTICLE INFO

Keywords:
Manufacturing workshop;
Scheduling;
Digital Twin;
Siamese Network;
Competitive Particle Swarm
Optimization;
Siamese Neural Network

***Corresponding author:**
liliweng88@126.com
(Weng, L.L.)

Article history:
Received 14 October 2024
Revised 19 October 2024
Accepted 25 October 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

With the rapid development of digital technology, the global manufacturing industry has entered the era of Industry 4.0. The emerged digital twin technology has solved the shortcomings of traditional manufacturing in workshop control and decision-making [1]. It utilizes the mapping relationship between virtual space and physical entity space to achieve data fusion, and uses digital twin technology to achieve high-precision management, control, and decision-making in industrial manufacturing [2]. In industrial manufacturing, many manufacturing workshops commonly face workshop scheduling. Workshop scheduling arranges scheduling plans based on process conditions, material quantities, and machine usage to effectively reduce manufacturing cost and improve production efficiency. However, in actual manufacturing, a product will go through multiple processes, multiple machines, and other processing steps during production [3]. Due to the multiple influencing conditions and the susceptibility to external disturbances in the manufacturing process, many scheduling decisions in manufacturing workshops are not ide-

al. Therefore, to solve the scheduling problem in manufacturing workshops, a dynamic scheduling mechanism based on Siamese Network (SN) is proposed. There are two innovations in the research. One is the introduced digital twin technology and competitive particle swarm optimization to construct a data fusion scheduling model, achieving efficient scheduling in the workshop. Secondly, SN is introduced to optimize workshop disturbances, and the production efficiency of workshop manufacturing is improved by readjusting scheduling after disturbance. The research content will provide technical support for the improvement of workshop scheduling and the effective application of digital twin technology. The research has two contributions. Firstly, by combining digital twin technology and machine learning technology, a data fusion scheduling model was constructed, effectively improving the scheduling efficiency of manufacturing workshops and promoting the development of workshop intelligence. The second is to study the introduction of twin neural networks to optimize workshop disturbances, significantly improving the production efficiency of the workshop in the face of disturbances, and enhancing the safety and stability of the manufacturing workshop.

2. Literature review

Effective workshop scheduling is beneficial for resource integration and improving the overall operational efficiency of the production line. Relevant scholars have conducted extensive research on this. Luo *et al.* found that multi-resource constrained workshop scheduling faced many difficulties. To improve workshop efficiency and scheduling effectiveness, a problem decision model based on Markov was designed to represent the numerical state characteristics and related constraint relationships of the workshop through mathematical models. The results showed that this technology could significantly improve the work efficiency of the workshop. The technology was superior to similar models [4]. Li *et al.* found that automatic guided vehicles have become an indispensable tool in workshop production. In practical work, their scheduling system is unable to handle specific working conditions. A dynamic scheduling technique was proposed in this study, which utilized its characteristics and task requirements to establish a scheduling model. The proposed technology was reliable and solved the scheduling problem of automatic guided vehicles in complex environments [5]. Liu *et al.* found that the automatic tractor in current intelligent manufacturing plays a critical role in automated production lines, but it cannot be effectively integrated into the workshop scheduling system, which affects the efficiency of workshop operations. In response to this, a flexible automatic traction vehicle scheduling technology was proposed. Then, an optimization model was established using an optimized genetic algorithm, which was then used to complete scheduling optimization through learning. The results indicated that the improved technology could effectively improve the scheduling effect of the workshop, and its performance was superior to similar scheduling technologies [6]. Zhang *et al.* found that due to various factors such as environment and time constraints, it is difficult for carrier-based aircraft to complete maintenance work within an effective time. In response to this, an analysis was conducted on its maintenance scheduling. A carrier-based aircraft maintenance scheduling technology considering comprehensive factors was proposed. The improved genetic algorithm was used to establish a scheduling model, and the scheduling performance was improved through parameter optimization. Experimental tests showed that the proposed technology could significantly improve the maintenance effectiveness of carrier-based aircraft. The scheduling process was flexible and reliable [7].

With the development of deep learning technology, workshop scheduling technology based on twin technology has ushered in important development and plays an important role in workshop scheduling. Tliba *et al.* found that production lines faced problems such as low efficiency and unreasonable resource utilization in actual management due to factors such as resource and workshop allocation. To overcome the above problems, considering the constraints of workshop conditions and resource status, a twin driven dynamic scheduling technique based on twin technology was proposed. In specific scenarios, the proposed technology had good adaptability while meeting the constraints, improving the efficiency of workshop scheduling [8]. Huo *et al.* proposed a dynamic scheduling model based on digital twins and bacterial foraging algorithm to

improve the efficiency of production line scheduling. The model allocated and sorted schedules based on two sub problems in the scheduling process. Corresponding models were established. In the study, a certain mechanical grinding tool production line was selected as an experimental case. The technology was applied to the working environment. The results showed that the research technology could optimize the production line scheduling well, with higher efficiency and avoiding previous production line interruptions [9]. Peng *et al.* conducted research on automated industrial production lines and found that the contradiction between workshop system control management and execution layer affected industrial manufacturing efficiency. To improve the production efficiency of the production line, an intelligent scheduling architecture based on digital twin technology was proposed, which combined fuzzy control and deep learning, and reflected the impact relationship between different process scheduling through digital modeling. Finally, applying it to railway manufacturing workshops could significantly improve the workshop resource utilization and enhance the production efficiency of the production line [10]. In addition, Khadiri *et al.* found that solving the dynamic scheduling problem of production lines was necessary in intelligent manufacturing to effectively enhance the allocation efficiency and manufacturing efficiency of production lines. In this regard, a dynamic workshop scheduling system based on distributed holographic architecture and digital twin technology was proposed in the research. This technology solved the resource scheduling problem of manufacturing systems through bidirectional model design, and improved the controllability and control effect of the entire workshop [11].

According to the above research findings, twin technology can effectively optimize workshop scheduling and play a key role in intelligent control of the workshop. Therefore, to improve the production efficiency of manufacturing workshops, a smart workshop scheduling technology based on digital twin technology is proposed. Through workshop modelling and constraint analysis, workshop scheduling is improved to enhance the development effect of intelligent manufacturing in enterprises.

3. Research methodology

3.1 Modelling of discrete manufacturing workshop scheduling problem

To improve the workshop scheduling efficiency in manufacturing workshops, a SN-based manufacturing workshop scheduling model is built to promote the rapid development of the manufacturing industry [12]. The framework of workshop scheduling optimization technology based on SN is shown in Fig. 1.

In workshop scheduling based on twin networks, the first step is to clarify the workshop scheduling problem, followed by data fusion to determine the flow information mode. Finally, the scheduling model is established to solve the manufacturing workshop scheduling problem [13]. In specific discrete manufacturing workshops, Job-shop Scheduling Problem (JSP) is actually a production process priority decision-making process, in which multiple single projects are decomposed and scheduled [14]. The shortest allocation time is sought to complete production tasks in a limited resource environment [15]. Therefore, based on the scheduling characteristics of the manufacturing workshop, O_{ip} is the p -th process of workpiece i . The objective function is set to f , which will be represented as Eq. 1.

$$f = \min \left(\max(C_{ip}) \right) \quad (1)$$

In Eq. 1, C_{ip} represents the completion time of the p -th process of workpiece i in the production workshop. In the specific production scheduling process, there may be differences in the selection of optimization indicators based on enterprise plans, actual inventory materials, and process characteristics, but the priority goal is basically to complete tasks in the shortest possible time. At the same time, in manufacturing tasks, it is necessary to carry out according to the relevant requirements of the task, so corresponding workshop scheduling constraints need to be met [16]. One is that a certain process in manufacturing matches a unique production machine, and the machine can only operate on specific processes. Therefore, its constraint is shown in Eq. 2 [17].

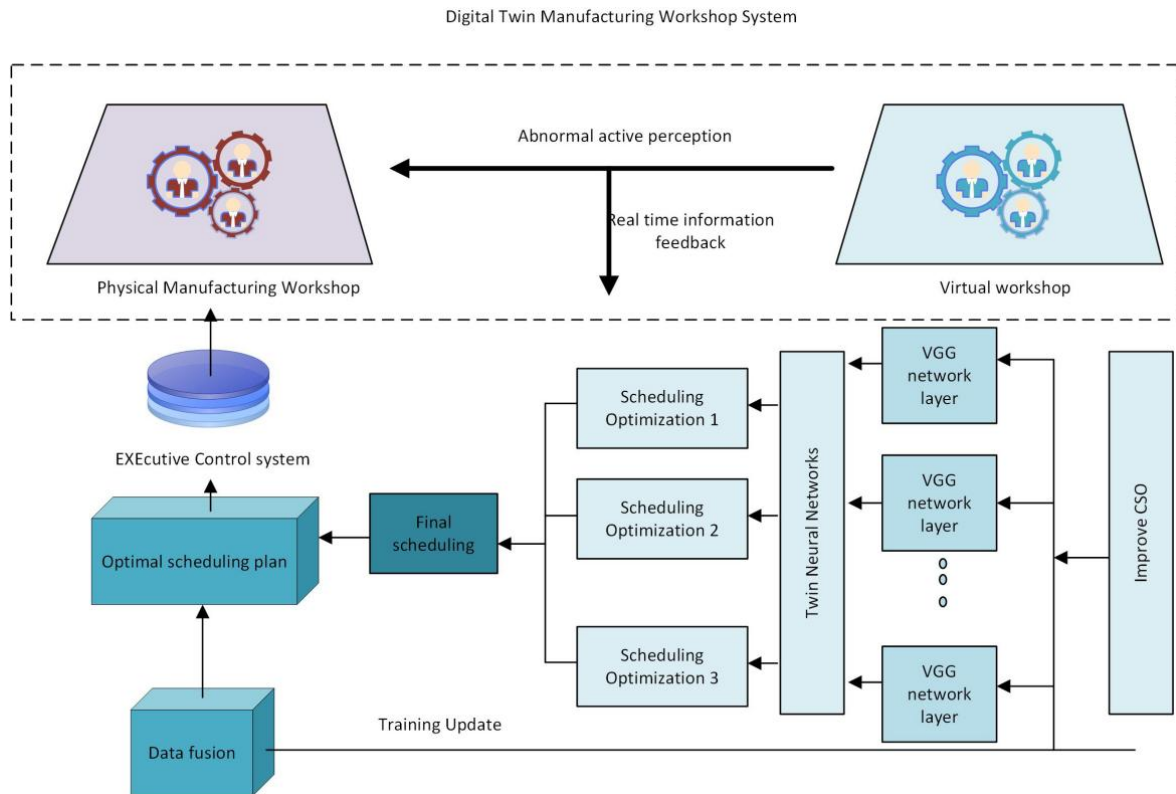


Fig. 1 Framework of workshop scheduling optimization technology based on SN

$$\sum_{k=1}^J X_{ipk} = 1, i = 1, 2, \dots, N, p = 1, 2, \dots, J \quad (2)$$

In Eq. 2, X_{ipk} represents the operational status of process O_{ip} on production machine k . N represents the number of workpieces. J represents the number of workpieces. In addition, in actual production, each machine needs to complete the previous task process before proceeding to the next task. The second constraint expression is shown in Eq. 3.

$$C_{ipk} \leq Q_{(i+1)pk}, i = 1, 2, \dots, N, p = 1, 2, \dots, J, k = 1, 2, \dots, m_p \quad (3)$$

In Eq. 3, $Q_{(i+1)pk}$ represents the next starting time of process O_{ip} on production machine k . C_{ipk} represents the completion time of process O_{ip} on production machine k , m_p represents the number of production machines. In addition, each workpiece in the manufacturing workshop also needs to complete the previous process before proceeding to the next task. The third constraint expression is shown in Eq. 4.

$$C_{ipk} \leq Q_{i(p+1)k}, i = 1, 2, \dots, N, p = 1, 2, \dots, J, k = 1, 2, \dots, m_p \quad (4)$$

In Eq. 4, $Q_{i(p+1)k}$ represents the start time of the next process in the production machine. The entire manufacturing workshop scheduling optimization process is shown in Fig. 2.

According to the workshop scheduling process in Fig. 2, the workshop scheduling problem needs to start from the manufacturing process path. Under the constraints of manufacturing, the manufacturing process is optimized to achieve the minimum task time.

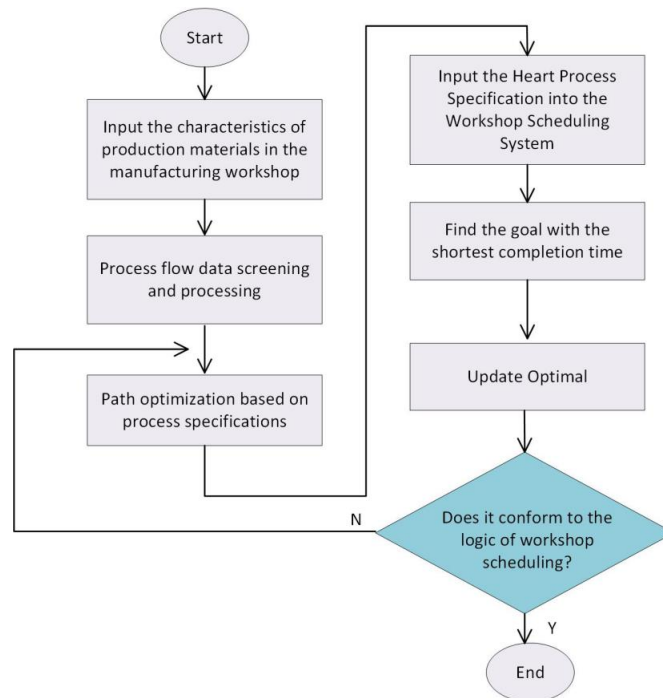


Fig. 2 Manufacturing workshop scheduling optimization process

3.2 Construction of data fusion scheduling model based on digital twins

The previous section analyses the scheduling problem in discrete workshops. In actual production in discrete workshops, due to the characteristics of product attributes, multiple process information needs to be analysed in manufacturing, resulting in heterogeneous data and difficulty in scheduling solution [18]. To address this issue, the research needs to describe complex production problems as similar feature problems. This study introduces digital twin technology for data fusion and uses an improved Competitive Swarm Optimizer (CSO) to solve it. The digital model of the workshop based on digital twin technology is shown in Fig. 3.

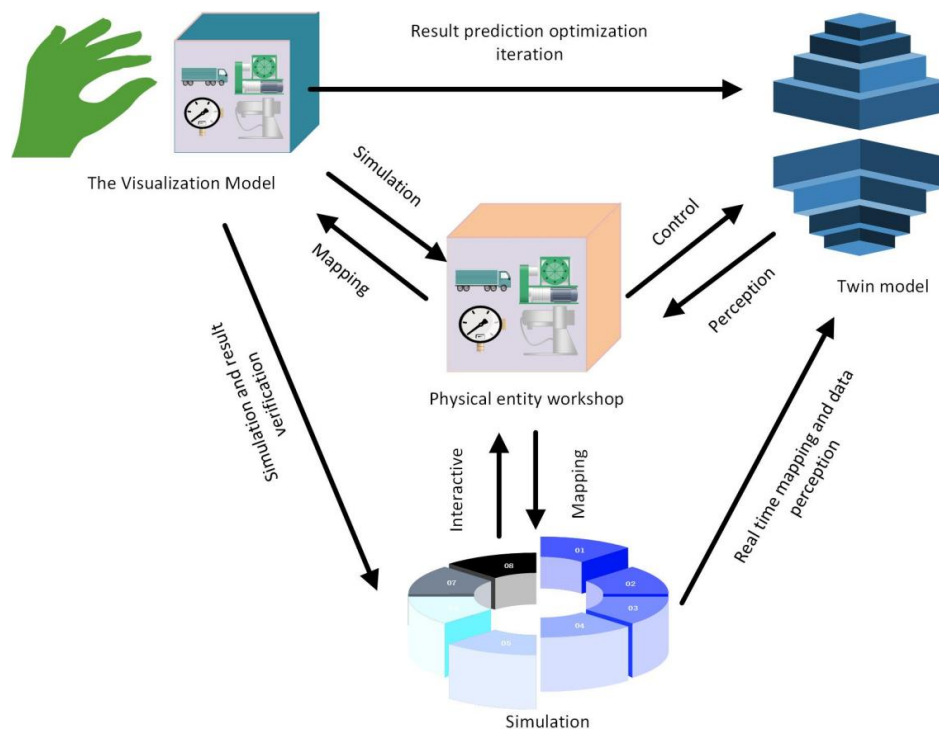


Fig. 3 Digital twin workshop digital model

From Fig. 3, digital twin technology has established a connection between virtual workshops and physical manufacturing workshops and achieved control and management of workshop production and manufacturing through interaction and mapping [19]. Therefore, it is used in the study to integrate multi-source information from manufacturing workshops, providing a foundation for subsequent scheduling solutions. In the digital twin manufacturing workshop, in order to describe the virtual real control relationship of the workshop, the initial three-dimensional coordinates and the adjusted three-dimensional coordinates are defined as P and p' respectively, as expressed in Eq. 5.

$$\begin{cases} P = (x, y, z) \\ P' = (x', y', z') \end{cases} \quad (5)$$

In Eq. 5, x , y and z are the initial coordinate parameters. x' , y' and z' represent the coordinate parameters after movement. According to the coordinate parameters. Multiplying P' with the four-dimensional coordinate $(x, y, z, 1)$ can calculate a three-dimensional model action change matrix, as shown in Eq. 6.

$$P' = PM = (x, y, z_1) \begin{bmatrix} b_{11} & b_{12} & b_{13} & p_x \\ b_{21} & b_{22} & b_{23} & p_y \\ b_{31} & b_{32} & b_{33} & p_z \\ t_x & t_y & t_z & s \end{bmatrix} \quad (6)$$

In Eq. 6, t_x , t_y and t_z represent the three-dimensional coordinate translation parameters. p_x , p_y and p_z represent the perspective projection transformation matrix in three-dimensional space, with three parameters serving as projection midpoints. The 9 parameters, including b_{11} and b_{12} , are coordinate transformation matrix parameters responsible for model scaling and rotation. s represents the integer scaling coefficient in the three-dimensional model which mainly controls the scaling size of the model in three-dimensional space. The twin manufacturing workshop information model is shown in Fig. 4.

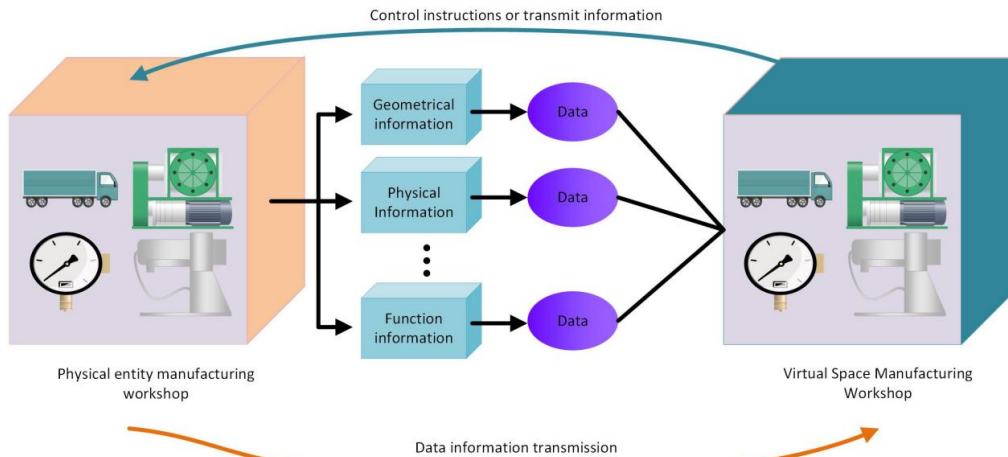


Fig. 4 Twin workshop information model

In the twin workshop, the information features that need to be extracted include labour, materials, equipment, production environment, and production knowledge. Based on this, key data for information fusion is constructed, as shown in Eq. 7 [20].

$$DT = DT_{cq} \cup DT_{pr} \cup DT_{pe} \cup DT_{cn} \cup DT_{kn} \quad (7)$$

In Eq. 7, DT_{eq} represents device information. DT_{pr} represents material information. DT_{pe} represents personnel information. DT_{en} represents environmental information. DT_{kn} represents production knowledge information. Among them, DT_{kn} information contains the entire workshop workpiece processing process information, which needs to be optimized starting from it. After obtaining the fused data, an improved CSO model is used to solve the scheduling problem.

CSO is a search algorithm derived from particle swarm optimization without memory, which has excellent optimization performance and strong convergence. CSO selects the optimal individual through competitive relationships. The CSO algorithm is updated, as shown in Eq. 8.

$$v_1(t+1) = R_1 v_1(t) + R_2 [x_w(t) - x_t(t)] + \varphi R_3 [\bar{x}(t) - x_t(t)] \quad (8)$$

In Eq. 8, $x_w(t)$ represents the dimensional vector of the winning individual after t rounds of competition. $x_t(t)$ is the dimension vector of the failed individual. $\bar{x}(t)$ represents the random position value of individuals competing t times. R_1 , R_2 and R_3 are all representative interval distributed random vectors. φ represents updated the influencing parameters. $v_1(t)$ represents the updated strategy. In CSO updates, population diversity is improved by half updating, but it can easily lead to a decrease in inter population searchability. Therefore, two encoding segments are introduced to improve CSO. Among them, circular and cross topology structures are added to the positions of the winning individuals, and the top 20 % of individuals are selected for neighborhood search to improve the overall optimization performance of the model. In addition, it is also considering that the circular topology structure allows encoding adjacent individuals to be connected to each other. But the expansion of the radius slows down information transmission and limit information sharing. Therefore, a Precedence Operation Crossover (POX) algorithm is adopted to integrate into the circular structure. Individuals and adjacent excellent individuals are selected as the subsequent crossover parents, and excellent individuals are selected from the iterative offspring to replace the current individual, achieving model improvement [21]. The crossover probability is shown in Eq. 9.

$$p_c = p_0 \times \frac{t_{rmax} - t_r}{t_{rmax}} \quad (9)$$

In Eq. 9, p_0 represent the initial values, while t_{rmax} denotes the maximum iteration, and t_r represents the current iteration count.

3.3 Modelling of workshop dynamic scheduling mechanism based on SN

In actual manufacturing workshops, different processes can lead to problems such as overlapping and scheduling errors in a certain link. To solve this problem, a workshop dynamic scheduling mechanism based on SN is proposed to address the scheduling disturbances in manufacturing workshops. SN is a coupled deep learning framework composed of multiple identical structures. The core of the SN is to use multiple sub networks to simultaneously train and compare parameter similarities and differences [22]. Based on this study, SN network is used to identify abnormal disturbances between physical and virtual workshops in digital twins. Cumulative disturbances are determined, and scheduling is adjusted by mining historical workshop data. At the same time, the impact of updated scheduling on actual workshop manufacturing is compared. Continuous iterations are repeated to approach the optimal workshop scheduling, achieving workshop scheduling disturbance optimization [23]. In the study, the loss function of the SN network is selected as Binary Cross Entropy, which has better stability, as displayed Eq. 10.

$$loss(y_o, t_r) = -[t_r \log(y_o) + (1 - t_r) \log(1 - y_o)] \quad (10)$$

In Eq. 10, t_r represents the label, while y_o denotes network output. In the sub networks of SN, Visual Geometry Group (VGG) has excellent application effects in the field of image feature processing. Therefore, with the help of VGG features, 3×1 convolution is introduced to design the sub network. At the same time, the Sigmoid function is selected as the network activation function, which has excellent performance in handling binary classification problems and is more advantageous for handling features with obvious differences, meeting the requirements of digital twin feature data processing. Therefore, the Sigmoid function is expressed as shown in Eq. 11.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

When processing data that is too large or too small, SN is prone to gradient vanishing problems, which can affect network training. Therefore, Batch Normalization (BN) is applied to address gradient vanishing problems. Increasing the BN layer after each convolution improves the

training speed of the network. The BN layer transformation is shown in Eq. 12.

$$\begin{cases} \hat{x}^{(i)} = \frac{x' - E[x^{(i)}]}{\sqrt{Var[x^{(i)}]}} \\ y^{(i)} = \gamma^{(i)} \hat{x}^{(i)} + \beta^{(i)} \end{cases} \quad (12)$$

In Eq. 12, $E[x^{(i)}]$ represents the mean, while $\sqrt{Var[x^{(i)}]}$ represents the standard deviation. $\gamma^{(i)}$ corresponds to the learning scaling parameters, and $\beta^{(i)}$ represents the offset parameter. $y^{(i)}$ is the output of neurons, and $x^{(i)}$ is the input vector. Finally, $\hat{x}^{(i)}$ represents transformed input vector. To enhance the ability of the SN model to extract workshop feature data, the study employs a double-layer VGG convolutional network structure. This includes various layers such as activation and pooling layers, designed for extracting features that capture differences in twin workshop information. Multiple similar structures are combined to form the SN model. The SN structure is shown in Fig. 5 [24].

After constructing the SN, it is necessary to select important feature data from the digital twin system as the output of the SN, such as processing time, production machine processes, process constraints, etc. Therefore, Table 1 data are taken as status characteristics.

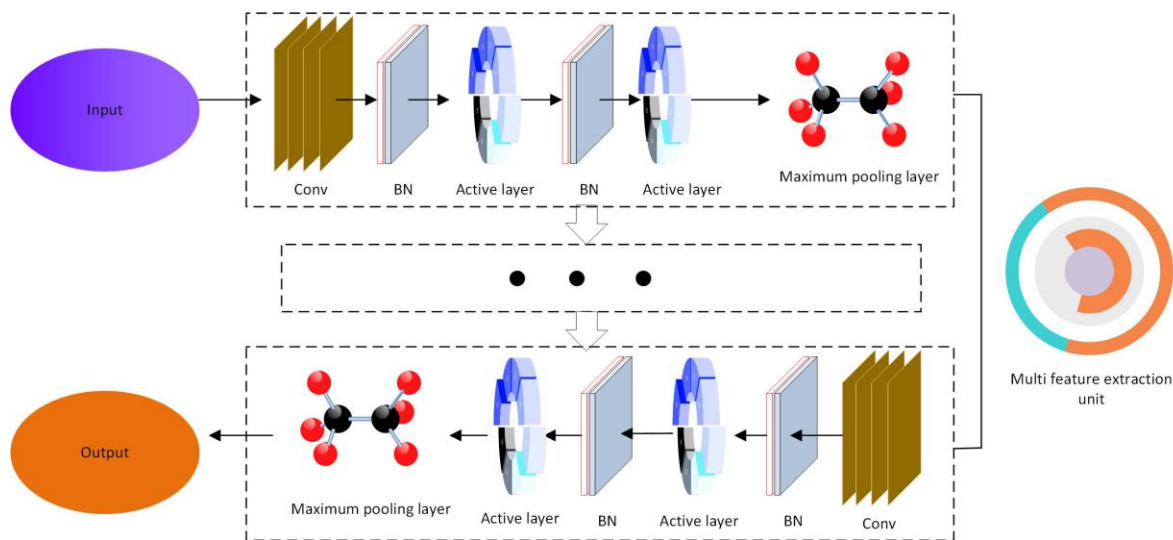


Fig. 5 SN sub network structure

Table 1 Status characteristics of digital twin manufacturing workshop

Signature sequence	Characterization
1	The total duration of workpiece processing
2	The number of processes in which the workpiece has been processed
3	The number of idle workpieces in the workshop
4	Number of idle machines in the workshop
5	The number of processes that the machine has not yet processed
6	The number of machines being processed in the workshop
7	The total duration of machine processing
8	Total idle time of the machine
9	The number of processes where the workpiece has not been processed yet
10	Number of processes for the workpiece
11	The number of processes to be processed by the machine
12	The machine number of the workpiece being processed
13	Total idle time of workpieces
14	The number of processes already processed by the machine

The min-max standard is used to preprocess selected data and improve the accuracy of network training. In actual training, if the SN output label is selected based on the accumulated disturbance time point, the output label of the data state before the accumulated disturbance is

determined to be 0. The output label of the data state after the accumulated disturbance point is determined to be 1. In addition, for traditional manufacturing workshops with a single scheduling rule, SN cannot provide good scheduling decisions. Therefore, a pseudo-SN is added to the SN to learn features, which simplifies the structural process, uses simplified convolutions as sub networks, and mines the mapping relationship between processes and machine scheduling. The connection layer of the network achieves feature fusion through concatenation, followed by the simplified double convolution to enhance feature extraction. Through the above research, the dynamic scheduling mechanism for manufacturing workshops is completed.

4. Results and discussion

4.1 Experimental analysis of workshop scheduling effect

To test the manufacturing workshop scheduling technology, corresponding experiments are conducted to verify the reliability of the technology. Firstly, the experimental environment is built. In the experiment, a flexible workshop for electrical manufacturing is taken as the research object. The workshop has a total of 26 production and manufacturing processes, several machines, and 6 workpieces. The workshop scheduling solution adopts an improved COS model and a SN scheduling mechanism to solve the dynamic scheduling disturbance problem caused by the accumulation of hidden disturbances in the workshop. In the digital twin workshop, the improved CSO algorithm is set to have 100 iterations, a population size of 60, and a crossover rate of 0.8. Simultaneously, the Genetic Optimization-Particle Swarm Optimization Algorithm (GA-PSO) and standard CSO algorithm are introduced for comparison. Sphere function and Griebank function are taken to test the performance of different scheduling models, as shown in Fig. 6.

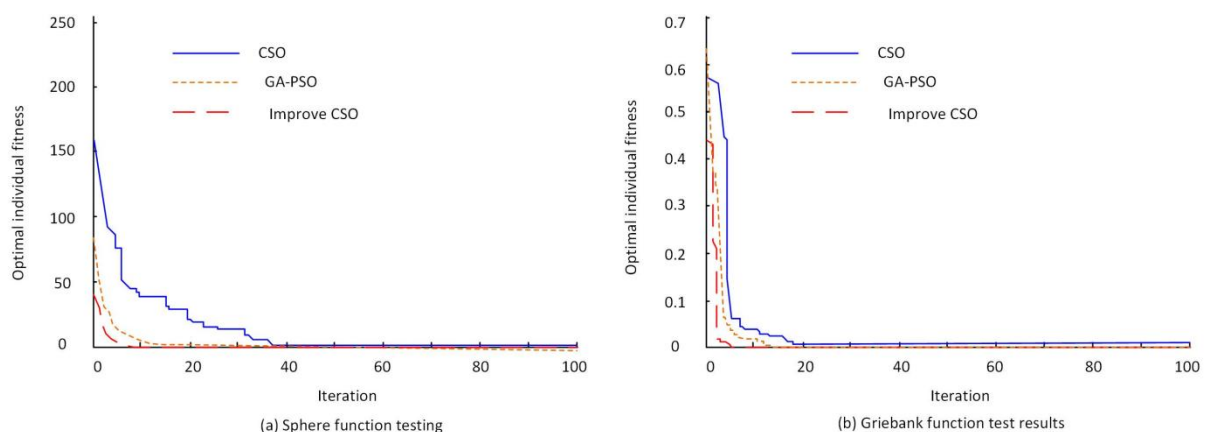


Fig. 6 Comparison of optimization performance of different scheduling models

Fig. 6(a) and 6(b) show the optimization test results of Sphere function and Griebank function, respectively. According to the test results, the improved CSO converged faster and was closer to the optimal value of the function compared with GA-PSO and CSO. In the Sphere function, the improved CSO converged 9 times, achieving an optimal fitness value of 0.025. In comparison, the GA-PSO converged 18 times, with a final fitness value of 0.156, while the standard CSO converged 38 times, reaching an optimal fitness value of 0.186. In the Griebank function, the improved CSO still performed better, with a 5.38 % and 11.25 % improvement in optimization ability compared with GA-PSO and CSO.

Next, the training results of the objective function for different scheduling models are shown in Fig. 7. This figure shows the results of solving scheduling objectives for different scheduling models. The standard CSO scheduling model performed the worst overall. According to the test curve, although the standard CSO model had an advantage in population diversity in the early stages of training, its search performance decreased in the later stages of training, and the model fell into local convergence. Its maximum completion time was ultimately 264.5 minutes. In contrast, GA-PSO had better optimization performance and convergence effect, but due to the influ-

ence of genetic algorithm parameters, premature convergence phenomenon occurred when processing high-dimensional data. Its maximum completion time was 249.2 minutes, which was better than the standard CSO model. The overall best performance is the improved CSO model, with a final target value of 244.8 minutes. It enhances the information communication between individuals through neighbourhood structure, avoids local optima, and has more advantages in convergence. Next, the scheduling Gantt chart results of different models are compared, as shown in Fig. 8.

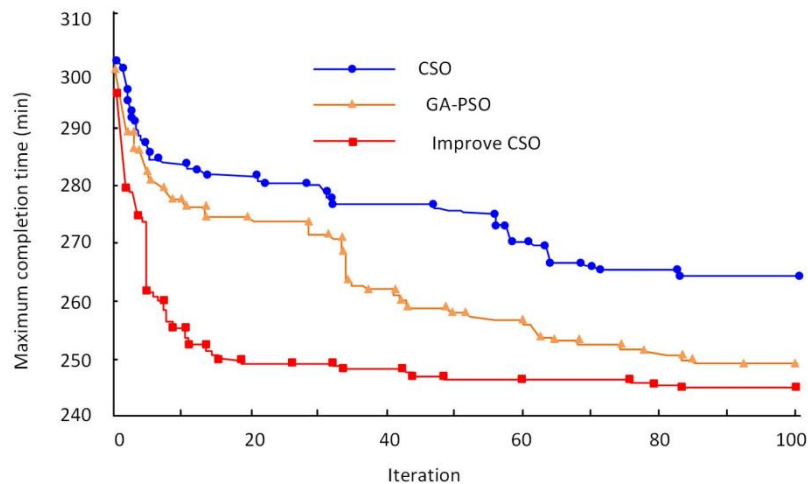


Fig. 7 Results of solving scheduling objectives for different scheduling models

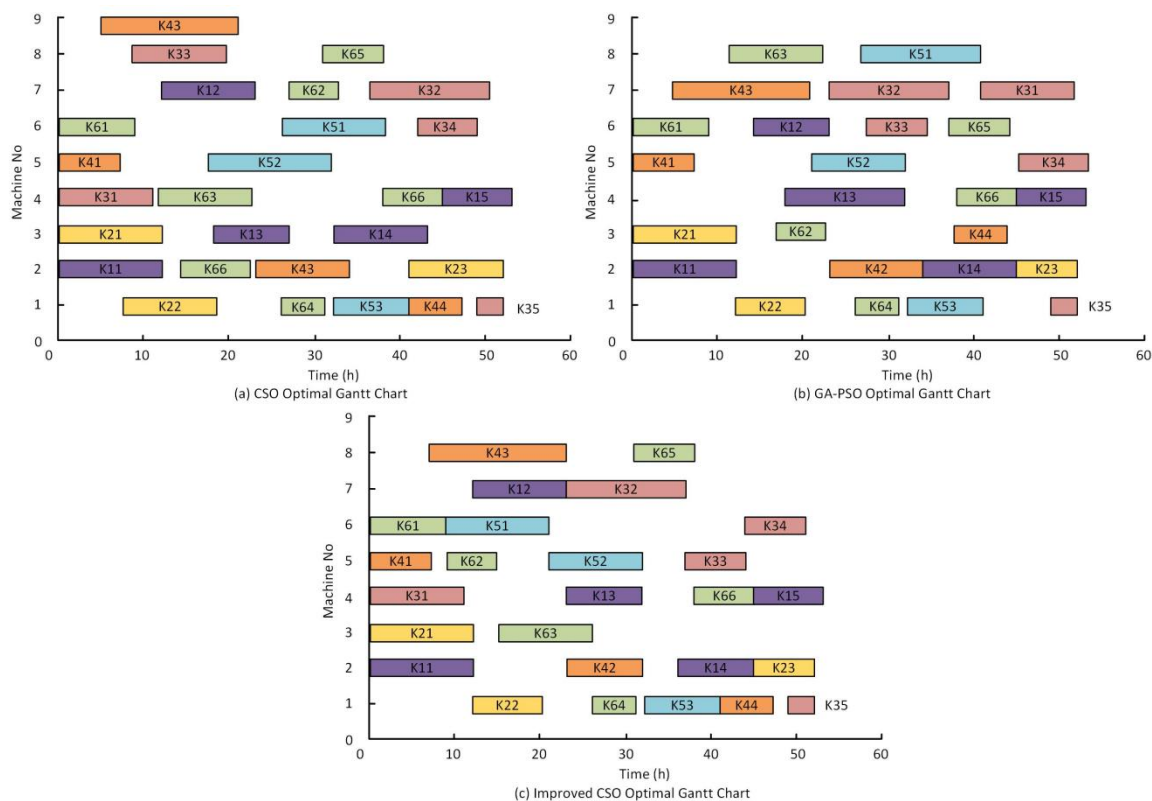


Fig. 8 Optimal Gantt chart results for different scheduling models

Fig. 8(a) shows the optimal result of the CSO Gantt chart. The number of processing machines under CSO training was 9, the number of processed workpieces was 6, and convergence was achieved after 84 iterations. The average processing time for a single workpiece was 59.2 minutes. Fig. 8 shows the optimal result of the GA-PSO Gantt chart. GA-PSO scheduling optimization was better. The processing machine consisted of 8 units and processed 6 workpieces. It

converged after 85 iterations, with an average processing time of 55.25 minutes for a single workpiece. Fig. 8(c) shows the optimal result of the improved CSO Gantt chart, which converged after 82 iterations. The number of processing machines was 8, the number of processed workpieces was 6, and the average processing time for a single workpiece was 52.2 minutes.

4.2 Experimental analysis of workshop scheduling based on SN

To test the application effect of SN in scheduling rules, the improved CSO model scheduling results are selected as a case study for analysis. Among them, there are 26 workshop processes, 8 machines, and 6 workpieces. In the experiment, the convolution kernel of the SN is set to 3×1 , with 5 feature extraction units, a sliding step size of 2, and an optimizer of Adam. In the experiment, the process code is set as K, the workpiece code is C, and the mechanical code is D. Table 2 shows the corresponding machining machine information for the manufacturing and processing process of this electrical appliance.

Table 2 Electrical manufacturing and processing information of a manufacturing workshop (Time/h)

Workpiece	C1				C2				C3				
Production processes	K11	K12	K13	K14	K15	K21	K22	K23	K31	K32	K33	K34	K35
D1	-	18	-	-	15	-	8	16	-	-	9	-	3
D2	12	-	-	11	-	-	-	7	-	-	-	12	-
D3	-	-	14	-	-	12	-	-	-	12	-	-	-
D4	-	19	9	-	8	19	9	-	11	18	15	-	4
D5	14	-	-	0.9	-	14	-	-	10	-0:	7	-	-
D6	-	-	17	-	-	-	-	9	-	-	-	5	8
D7	20	11	-	-	-	-	15	-	-	14	-	-	-
D8	-	-	-	12	18	-	-	-	13	-	-	9	-
Workpiece	C4				C5				C6				
Production processes	K41	K42	K43	K44	K51	K52	K53	K61	K62	K63	K64	K65	K66
D1	-	-	-	6	-	-	9	-	8	-	5	-	19
D2	19	-	11	-	-	18	-	11	-	-	-	11	-
D3	-	8	-	-	-	-	12	-	-	11	-	-	-
D4	-	-	8	-	22	-	-	-	-	-	15	-	7
D5	7	11	-	-	-	11	-	-	6	17	-	-	-
D6	-	-	-	-	12	-	-	9	-	-	-	8	15
D7	13	-	18	14	17	-	-	14	-	-	7	-	-

The scheduling results in Table 2 are obtained from the training of the improved CSO scheduling model. In the improved CSO model, the planned completion time was 52.2 min. The disturbance was added in the subsequent experiments to simulate the disturbance of different process abnormal states in scheduling. The scheduling data are input into the digital twin scheduling system and 21.2 min perturbations are added. Whether the system disturbance has been triggered is checked, as shown in Fig. 9. This figure shows the results of the disturbance effect test. According to the results, disturbance is added at 24.1 minutes of scheduling, as indicated by the red line in Fig. 9(b). Compared with the simulated output value of 0.523 at 24.1 minutes in Fig. 9(c), it exceeded the system's specified disturbance standard value of 0.5, indicating successful triggering. Next, the effect of different models on post disturbance scheduling adjustment is compared.

The disturbance time point is 37.1 min, and the results are shown in Fig. 10. Fig. 10(a) to 10(d) show the original scheduled scheduling, perturbed unadjusted scheduling, pseudo-SN scheduling, and SN scheduling, respectively. According to the test results, the original workshop's planned completion time for the workpiece was 52.2 minutes. After disturbance, its maximum completion time in its unadjusted state became 58.5 minutes. After using pseudo-SN and SN for disturbance adjustment, the maximum completion time of both models was adjusted to 56.4 minutes and 54.2 minutes, respectively. Both models improved the scheduling deviation after disturbance, while the SN performed the best, reducing the unadjusted state time by 4.3 minutes, which was better than the pseudo-SN's 2.1 minutes. The SN dynamic scheduling mechanism can effectively suppress disturbances and meet the scheduling requirements of the manufacturing workshop.

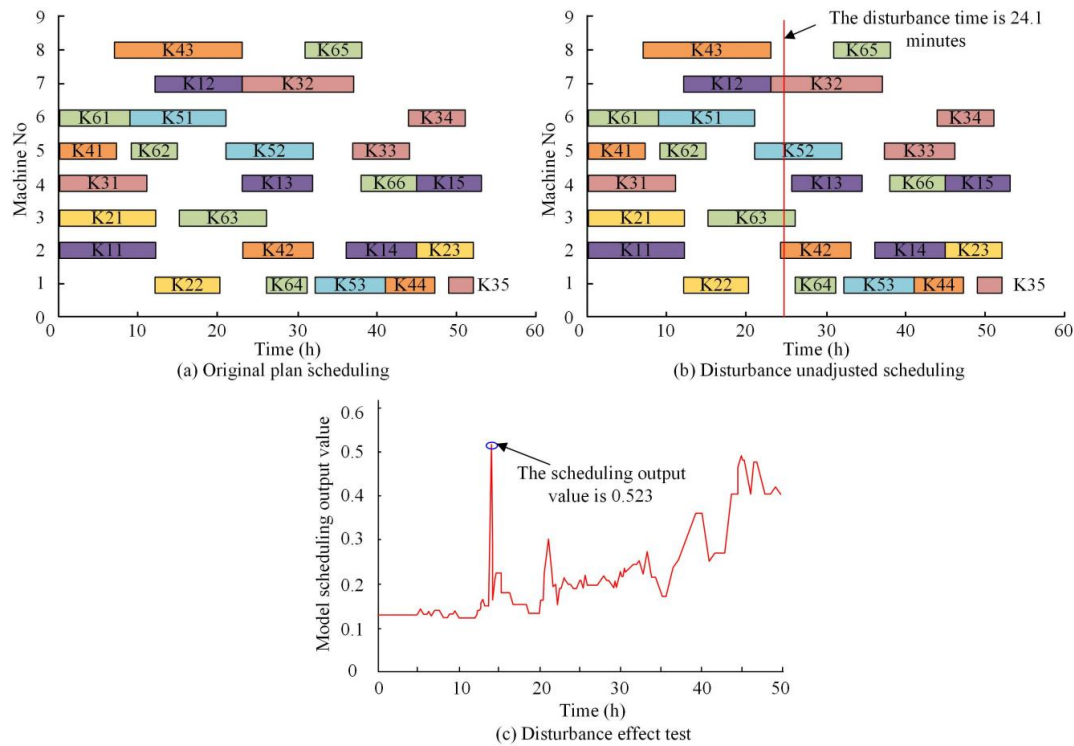


Fig. 9 Disturbance trigger test

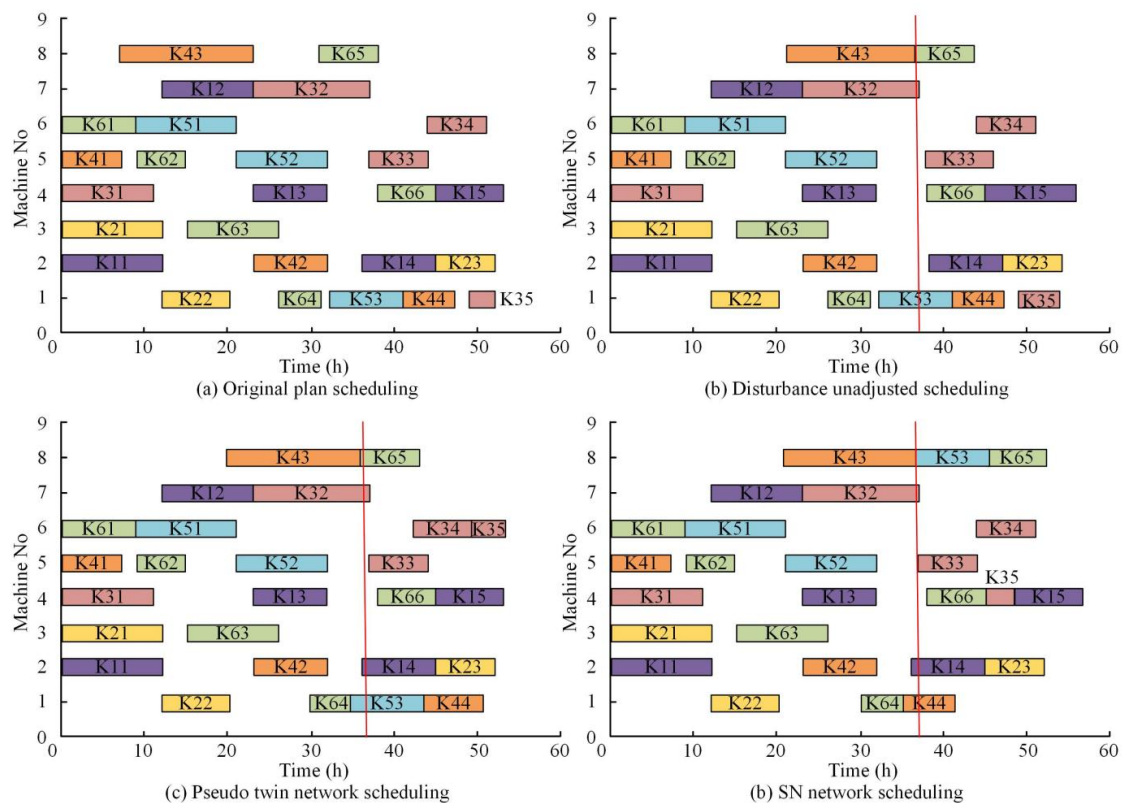


Fig. 10 Scheduling effects of different models after adding perturbations

In addition, studying the same process and increasing complexity to compare the effects of different technologies is shown in Table 3. This table shows the scheduling situation in complex scenarios, with each scenario repeated three times. According to the results, the overall performance of SN scheduling is better, with a typical mid-term disturbance time of 55.6 minutes,

pseudo-SN time of 57.5 minutes, and the original scheduling time of 60.5 minutes. In more complex repeated disturbances, the pseudo-SN has a large deviation from the original scheduling time, while the SN scheduling time is the shortest and the three times of time are close. From this, it can be seen that SN scheduling also has excellent performance and better stability in more complex scenarios.

Table 3 Time consumption for scheduling more complex scenarios

Scheduling scenario	Experiment number	Original scheduling (min)	Pseudo-SN scheduling (min)	SN scheduling (min)
General disturbance	1	60.5	57.5	55.6
	2	61.2	58.6	56.4
	3	60.8	57.8	55.4
Mild disturbance	4	71.3	66.5	61.2
	5	70.2	67.2	60.3
	6	72.5	66.9	61.2
Repeated disturbances multiple times	7	112.5	91.2	76.5
	8	109.5	95.2	78.5
	9	110.5	94.5	77.5

5. Conclusion

With the rapid development of digital industrialization, intelligent and automated manufacturing has become the key to the transformation of traditional manufacturing industry. However, in actual manufacturing workshops, many enterprises face insufficient workshop scheduling. To solve this problem, an intelligent scheduling mechanism was proposed based on a flexible production workshop. Firstly, the characteristics of workshop scheduling were analysed to construct a scheduling problem model. Considering the data scheduling and solving in discrete workshops, this study introduced an improved CSO algorithm based on digital twin technology to construct a scheduling model. Finally, the SN was introduced to construct a dynamic scheduling mechanism for manufacturing workshops, optimizing workshop scheduling disturbances. In the Griewank function scheduling performance test, the improved CSO showed a 5.38 % and 11.25 % improvement in optimization ability compared with GA-PSO and CSO, respectively. In the training of the objective function of the scheduling model, the maximum completion time of the improved CSO was 244.8 minutes, which was better than the 264.5 minutes and 249.2 minutes of CSO and GA-PSO. In the SN scheduling experiment, a disturbance was set at 37.1 minutes, and the maximum completion time of the SN after disturbance optimization was 54.2 minutes, while the pseudo-SN was 56.4 minutes. Compared with the unadjusted state of the disturbance, the maximum completion time of the SN was reduced by 4.3 minutes, and the optimization effect was the best. The workshop scheduling technology has good application effects. However, there are also shortcomings in the research. Due to the complexity of discrete workshop scheduling, only the shortest completion time is considered as the sole objective. In the future, cost, equipment load and other objectives can be considered to improve the practical application effect of the technology.

Acknowledgement

This project is funded by the 2024 Guangdong Provincial Education and Science Planning Project (Higher Education Special Project), specifically titled "Research on a New Industry-Education Integration Model Driven by the Integration of Digital and Practical Elements" (2024GXJK137).

Reference

- [1] Tliba, K., Diallo, T.M.L., Penas, O., Penas, O., Khalifa, R.B., Yahia, N.B., Choley, J.-Y. (2023). Digital twin-driven dynamic scheduling of a hybrid flow shop, *Journal of Intelligent Manufacturing*, Vol. 34, No. 5, 2281-2306, [doi: 10.1007/s10845-022-01922-3](https://doi.org/10.1007/s10845-022-01922-3).

- [2] Barnhart, C., Bertsimas, D., Delarue, A., Yan, J. (2021). Course scheduling under sudden scarcity: Applications to pandemic planning, *Manufacturing & Service Operations Management*, Vol. 24, No. 2, 727-745, doi: [10.1287/msom.2021.0996](https://doi.org/10.1287/msom.2021.0996).
- [3] Zhang, R., Lv, J., Bao, J., Zheng, Y. (2023). A digital twin-driven flexible scheduling method in a human-machine collaborative workshop based on hierarchical reinforcement learning, *Flexible Services and Manufacturing Journal*, Vol. 35, 1116-1138, doi: [10.1007/s10696-023-09498-7](https://doi.org/10.1007/s10696-023-09498-7).
- [4] Luo, P.C., Xiong, H.Q., Zhang, B.W., Peng, J.Y., Xiong, Z.F. (2022). Multi-resource constrained dynamic workshop scheduling based on proximal policy optimisation, *International Journal of Production Research*, Vol. 60, No. 19, 5937-5955, doi: [10.1080/00207543.2021.1975057](https://doi.org/10.1080/00207543.2021.1975057).
- [5] Li, Z., Sang, H., Pan, Q., Gao, K., Han, Y., Li, J. (2022). Dynamic AGV scheduling model with special cases in matrix production workshop, *IEEE Transactions on Industrial Informatics*, Vol. 19, No. 6, 7762-7770, doi: [10.1109/TII.2022.3211507](https://doi.org/10.1109/TII.2022.3211507).
- [6] Liu, Q., Wang, N., Li, J., Ma, T., Li, F., Gao, Z. (2023). Research on flexible job shop scheduling optimization based on segmented AGV, *Computer Modeling in Engineering & Sciences*, Vol. 134, No. 3, 2073-2091, doi: [10.32604/cmescs.2022.021433](https://doi.org/10.32604/cmescs.2022.021433).
- [7] Zhang, Y., Li, C., Su, X., Cui, R., Wan, B. (2023). A baseline-reactive scheduling method for carrier-based aircraft maintenance tasks, *Complex & Intelligent Systems*, Vol. 9, 367-397, doi: [10.1007/s40747-022-00784-9](https://doi.org/10.1007/s40747-022-00784-9).
- [8] Sun, Z.Y., Han, W.M., Gao, L.L. (2023). Real-time scheduling for dynamic workshops with random new job insertions by using deep reinforcement learning, *Advances in Production Engineering & Management*, Vol. 18, No. 2, 137-151, doi: [10.14743/apem2023.2.462](https://doi.org/10.14743/apem2023.2.462).
- [9] Tliba, K., Diallo, T.M.L., Penas, O., Khalifa, R.B., Yahia, N.B., Choley, J.-Y. (2023). Digital twin-driven dynamic scheduling of a hybrid flow shop, *Journal of Intelligent Manufacturing*, Vol. 34, 2281-2306, doi: [10.1007/s10845-022-01922-3](https://doi.org/10.1007/s10845-022-01922-3).
- [10] Huo, L., Wang, J.Y. (2022). Flexible job shop scheduling based on digital twin and improved bacterial foraging, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 525-536, doi: [10.2507/IJSIMM21-3-C014](https://doi.org/10.2507/IJSIMM21-3-C014).
- [11] Peng, F., Zheng, L. (2023). An improved multi-objective Wild Horse optimization for the dual-resource-constrained flexible job shop scheduling problem: A comparative analysis with NSGA-II and a real case study, *Advances in Production Engineering & Management*, Vol. 18, No. 3, 271-287, doi: [10.14743/apem2023.3.472](https://doi.org/10.14743/apem2023.3.472).
- [12] Peng, F., Zheng, L. (2023). Fuzzy rule-based neural network for high-speed train manufacturing system scheduling problem, *Neural Computing and Applications*, Vol. 35, 2077-2088, doi: [10.1007/s00521-022-07190-9](https://doi.org/10.1007/s00521-022-07190-9).
- [13] Khadiri, H., Sekkat, S., Herrou, B. (2023). Digital Twin Based SUDIHA architecture to smart shopfloor scheduling, *Journal of Manufacturing and Materials Processing*, Vol. 7, No. 3, Article No. 84, doi: [10.3390/jmmp7030084](https://doi.org/10.3390/jmmp7030084).
- [14] Yang, S.L., Wang, J.Y., Xin, L.M., Xu, Z.G. (2022). Verification of intelligent scheduling based on deep reinforcement learning for distributed workshops via discrete event simulation, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 401-412, doi: [10.14743/apem2022.4.444](https://doi.org/10.14743/apem2022.4.444).
- [15] Zhou, Z., Xu, L., Ling, X., Zhang, B. (2024). Digital-twin-based job shop multi-objective scheduling model and strategy, *International Journal of Computer Integrated Manufacturing*, Vol. 37, No. 1-2, 87-107, doi: [10.1080/0951192X.2023.2204475](https://doi.org/10.1080/0951192X.2023.2204475).
- [16] Nie, Q., Tang, D., Zhu, H., Sun, H. (2022). A multi-agent and internet of things framework of digital twin for optimized manufacturing control, *International Journal of Computer Integrated Manufacturing*, Vol. 35, No. 10-11, 1205-1226, doi: [10.1080/0951192X.2021.2004619](https://doi.org/10.1080/0951192X.2021.2004619).
- [17] Negri, E., Cattaneo, L., Pandhare, V., Macchi, M., Lee, J. (2022). Integrating PHM into production scheduling through a Digital Twin-based framework, *IFAC-PapersOnLine*, Vol. 55, No. 19, 31-36, doi: [10.1016/j.ifacol.2022.09.180](https://doi.org/10.1016/j.ifacol.2022.09.180).
- [18] Wan, J. (2022). Demand prediction and optimization of workshop manufacturing resources allocation: A new method and a case study, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 413-424, doi: [10.14743/apem2022.4.445](https://doi.org/10.14743/apem2022.4.445).
- [19] Yin, Y., Feng, J., Yi, B., Li, W., Yin, Q. (2024). Visual monitoring system of digital twin workshop for process manufacturing, *Journal of System Simulation*, Vol. 36, No. 1, 120-130, doi: [10.16182/j.issn1004731x.joss.22-1006](https://doi.org/10.16182/j.issn1004731x.joss.22-1006).
- [20] Tang, Q., Wu, B. (2024). A digital twin-driven multi-resource constrained location system for resource allocation, *The International Journal of Advanced Manufacturing Technology*, Vol. 130, No. 9, 4359-4385, doi: [10.1007/s00170-023-12886-w](https://doi.org/10.1007/s00170-023-12886-w).
- [21] Rožanec, J.M., Lu, J., Rupnik, J., Škrjanc, M., Mladenović, D., Fortuna, B., Zheng, X., Kiritsis, D. (2022). Actionable cognitive twins for decision making in manufacturing, *International Journal of Production Research*, Vol. 60, No. 2, 452-478, doi: [10.1080/00207543.2021.2002967](https://doi.org/10.1080/00207543.2021.2002967).
- [22] Wang, Y., Jiang, Z., Wu, Y. (2022). Model construction of material distribution system based on digital twin, *The International Journal of Advanced Manufacturing Technology*, Vol. 121, 4485-4501, doi: [10.1007/s00170-022-09636-9](https://doi.org/10.1007/s00170-022-09636-9).
- [23] Serrano-Ruiz, J.C., Mula, J., Poler, R. (2022). Development of a multidimensional conceptual model for job shop smart manufacturing scheduling from the Industry 4.0 perspective, *Journal of Manufacturing Systems*, Vol. 63, 185-202, doi: [10.1016/j.jmsy.2022.03.011](https://doi.org/10.1016/j.jmsy.2022.03.011).
- [24] Yang, J., Lin, F., Chakraborty, C., Yu, K., Guo, Z., Nguyen, A.-T., Rodrigues, J.J.P.C. (2023). A parallel intelligence-driven resource scheduling scheme for digital twins-based intelligent vehicular systems, *IEEE Transactions on Intelligent Vehicles*, Vol. 8, No. 4, 2770-2785, doi: [10.1109/TIV.2023.3237960](https://doi.org/10.1109/TIV.2023.3237960).

Optimization of cold chain multimodal transportation routes considering carbon emissions under hybrid uncertainties

Hou, D.N.^{a,b,*}, Liu, S.C.^c

^aDivision of Public Policy, Hong Kong University of Science and Technology, P.R. China

^bShanxi State-owned Capital Operation Co., Ltd, P.R. China

^cSchool of Business Administration, Fujian Business University, P.R. China

ABSTRACT

As the third largest greenhouse gas emission industry across the globe, the transportation industry dominates a major position in the total carbon emission. Multimodal transportation has incomparable advantages over single-modal transportation, and choosing transportation routes and modes under the background of carbon peaking and carbon neutrality is crucial. The uncertainties of demand and transportation time were described using the maximum regret value and Lyapunov central limit theorem, respectively, and a robust optimization model for cold chain multimodal transportation routes considering hybrid uncertainty of carbon emissions was established. Then, a dual-pheromone ant colony algorithm was designed, and the improved niche genetic algorithm was nested into the ant colony algorithm to solve the model. Results showed that the changes in transportation time and node transfer time lead to the changes in transportation cost and transportation scheme, and the robust optimization of the multimodal transportation route under hybrid uncertainties is affected by the regret value constraint and the fluctuation range of uncertain factors, resulting in the increase in transportation cost and carbon emission cost. Therefore, the decision-makers of cold chain multimodal transportation must predict the influence of uncertain factors, choose the appropriate maximum regret value, and pay attention to the mixed time window constraints of transportation time and node transfer time to reduce costs and improve efficiency.

ARTICLE INFO

Keywords:

Cold chain multimodal transportation;
Route optimization;
Hybrid uncertainties;
Niche genetic algorithm;
Dual-pheromone Ant Colony Algorithm;
Robust optimization;
Carbon emissions

**Corresponding author:*
denimhou@163.com
(Hou, D.N.)

Article history:

Received 27 September 2024

Revised 26 October 2024

Accepted 28 October 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

As an important constituent of the comprehensive transportation system, multimodal transportation can integrate various single modes of transportation and scattered transportation resources. Multimodal transportation has become an important way of cargo transportation and an inevitable development trend of the modern comprehensive transportation system by virtue of its efficient, safe, and economical characteristics. In addition, multimodal transportation plays an important role in integrating transportation resources, saving energy and reducing transportation costs [1], and serves as an effective way to realize the unification of transportation efficiency and cost [2]. According to statistics, the multimodal transportation volume in Europe and the United States has developed rapidly, and the overall proportion has increased year by year. Generally, the container multimodal transportation accounts for approximately 20 % interna-

tionally, 40 % in the United States, 35 % in France, 25 % in India, and 2 % in China [3]. Compared with single road transportation, multimodal transportation can reduce transport costs by 20 % and CO₂ emissions by 57 % [4]. The transportation efficiency and cost of cold chain container multimodal transportation are affected by multimodal transportation stations, transportation routes, means of transportation, transportation prices, transportation time and carbon emission restrictions, and different transportation organization schemes must be formulated by catering to the different needs of customers [5]. Cold chain container multimodal transportation involves many subjects, such as shippers, carriers, and consignees, and various modes and means of transportation, accompanied by problems such as the coordination of transportation organizations, the optimization of transportation routes, and the distribution of transportation benefits. In the complex and volatile external environment, multimodal transportation carriers should optimize transportation organization schemes considering multiple uncertain factors; for example, during the COVID-19 epidemic, some key nodes in the multimodal transportation network failed, the transportation demand and time were uncertain, while the goods were required to be delivered within the shortest time. Therefore, a modelling analysis under multifactor uncertainties was performed based on the study on container cold chain multimodal transportation route optimization under single-factor uncertainty. Many uncertain factors, such as transshipment time and uncertain transportation demand in the multimodal transportation network, interact with each other, and during cold chain multimodal transportation, various transportation equipment and refrigerated containers release carbon dioxide, which affects the choice of cold chain multimodal transportation routes. Hence, a cold chain multimodal transportation route optimization model considering carbon emissions under hybrid uncertainties was constructed, aiming to solve the problem of cold chain multimodal transportation route optimization under the hybrid uncertainties of transportation time and demand. Meanwhile, the model also considered the influence of carbon emissions on cold chain container multimodal transportation routes. Studying the intermodal route scheme under parameter changes through sensitivity analysis is of specific theoretical guiding importance for improving the operational capacity of the intermodal network in complicated situations.

2. Literature review

As for the research on the disturbance of uncertain demands, Qiu *et al.* established a robust optimization model of multimodal transportation routes through a scenario analysis among robust optimization methods, in which the objective function was the total cost composed of transportation cost, time value cost of goods and carbon emission cost, and the constraint conditions included the node operation time window [6]. On the basis of demand uncertainty and carbon trading policy, Deng *et al.* (2023) built a robust optimization model of multimodal transportation routes with the minimum economic cost as the objective and solved the model using a genetic algorithm to verify its rationality and effectiveness [7]. Li and Sun (2022) constructed a robust stochastic optimization model for container multimodal transportation considering the uncertainty of demand and the randomness of carbon transaction price and designed a hybrid firework algorithm based on gravity search operator [8]. During uncertainty model transformation, Wang (2024) used the Bertsimas robust optimization method to linearize the demand uncertainty parameter and then transformed the demand uncertainty model into a certainty model [9]. Tang *et al.* (2023) constructed a multi-objective model with the minimum total transportation cost and total transportation time, transformed the multiple objectives into a single objective by using the linear weighting method, and designed a genetic algorithm based on simulated annealing to solve the model [10]. Mainly considering the demand uncertainty in intermodal transportation, Li *et al.* (2017) designed the sequential decision on terminal location and route planning based on a two-stage stochastic programming model to realize an economical and efficient multimodal transportation network [11]. Considering the uncertain demand in multimodal transportation, Hu *et al.* (2017) proposed a two-stage stochastic programming model to determine the terminal location and transportation route under the uncertain customer demand [12]. Zhang *et al.* (2020) established a two-stage nonlinear stochastic programming model consider-

ing the uncertainty of iron ore demand and the economies of scale in port transshipment [13]. In accordance with the above research results, the transportation volume has been set to a specific value for optimization modelling in most of the existing literature while neglecting the uncertainty of customer transportation demands. Moreover, the design of transportation schemes is of certain foresightedness, and the capacity constraints of each transportation route and transshipment node have not been considered.

Regarding the research on the uncertainty of transportation price, Chen (2020) mainly considered the uncertain disturbance relations, such as transportation process price, depicted the uncertainty of transportation price through the interval method, constructed a hybrid integer robust optimization model, and solved the model with an accurate algorithm [14]. Aiming at the uncertainty of transportation cost, Hu *et al.* (2018) built a facility planning model considering the uncertainty of fixed investment, logistics cost, inventory cost and recovery [15]. Zhang *et al.* (2021) established a hybrid robust stochastic optimization model and designed a catastrophic adaptive genetic algorithm based on Monte Carlo sampling [16]. Xie and Cao (2021) explored into the multimodal transportation of dangerous goods under uncertain conditions and proposed an interval 0–1 integer programming model with the transportation risk and transportation cost during the transportation of dangerous goods as the objectives [17]. Li *et al.* (2024) considered the influences of uncertain freight rates and navigation conditions of different route segments, constructed the Yangtze River container multimodal transportation route optimization model under uncertain conditions, and designed Monte Carlo simulation and improved wolf pack algorithm to solve the model [18]. Pian *et al.* (2022) constructed a two-level programming model with the objectives of maximizing the profit of carriers and minimizing the generalized cost of shippers, solved the model by using a sensitivity-based heuristic algorithm, and verified its effectiveness [19]. Aradi (2022) constructed a two-level programming model, a carrier revenue maximization model, and a shipper transportation mode selection model considering the influence of the pricing strategy on shippers [20]. Considering the generalized cost of customers and the revenue of railway operation enterprises, Aghabayk *et al.* (2020) established a pricing model with the minimum generalized cost as the objective and a two-level programming pricing model and concluded that the two-level programming model is better than the model with the minimum generalized cost [21]. The above studies mainly focus on the uncertainty of transportation price, which, however, is affected by market law, transportation cycle, and international environment and strongly influenced by the market policy with marked regional differences. Moreover, the maximum and minimum carbon trading prices vary greatly from region to region; thus, the effect is extremely limited if the minimum value of the objective function is utilized to describe the container multimodal route optimization model considering the uncertainty of transportation price.

Regarding the research on uncertain transportation time, the influence of traffic congestion on multimodal transportation routes has been analysed by some scholars. For instance, Norouzi *et al.* (2015) built a route optimization model with the objectives of maximum reliability and minimum time considering road dynamics [22]. Ehr Gott *et al.* (2012) constructed a hybrid integer optimization model based on fuzzy chance constraints aiming at two uncertain factors, namely, railway capacity and highway passage time, and designed an accurate solution to solve the model by the linearization method [23]. Considering road reliability, Gupta *et al.* (2018) constructed a multi-objective route optimization model with the objectives of maximum reliability, minimum time, and minimum material unsatisfaction degree [24]. Taking the time-dependent nature of disasters, Gupta *et al.* (2020) established an emergency logistics route planning model with maximum reliability and minimum transportation time as the objective functions [25]. Mardanya *et al.* (2021) studied the route selection model for emergency supplies considering the reconfiguration and reliability of routes [26]. Biswas *et al.* (2019) built a multimodal route optimization model carrying time windows in the multimodal network with full consideration of the uncertainties in transportation time and transshipment time [27]. Given the time-dependent nature of the transportation network, Liu (2021) established a cold chain container multimodal transportation route optimization model, designed two scenarios—static optimization and dynamic optimization, and adopted an improved genetic algorithm to solve the model [28]. When

studying multimodal route selection, Das *et al.* (2020) introduced the time window constraint to establish a hybrid integer fuzzy mathematical model with demand assumption as a fuzzy number, aiming to improve customer satisfaction to the greatest extent [29]. Maity *et al.* (2020) probed into the modelling and optimization of the public highway–railway multimodal transportation system based on the central radiation network under uncertain conditions, taking cost and time as objective functions and using the weighted sum method to deal with multi-objective problems [30]. Roy *et al.* (2017) established a multi-objective route decision model for multimodal transportation considering the uncertainties of highway transportation speed and transshipment time in the actual transportation process and solved the model by using the K-shortest route algorithm and nondominated sorting genetic algorithm [31]. All the parameters related to transportation time in most multimodal transportation models studied above are definite values, but in actual transportation, transportation time is also affected by time-varying transportation networks.

In summary, the multimodal route optimization problem that simultaneously considers the hybrid certainties of cold chain container transportation time and demand has been rarely involved in the existing literature, not to mention the research considering carbon emissions and the uncertainty of multimodal links. On this account, the carbon emission cost was imported into the container multimodal route optimization model under the hybrid uncertainties in transportation time and demand to construct a cold chain container multimodal low-carbon route optimization model under the hybrid uncertainties of time and demand. For the NP-hard characteristic of the problem to be solved, a solution strategy combining niche genetic algorithm and ant colony algorithm was designed, followed by the verification of model and algorithm effectiveness through an example. This study is of certain theoretical importance and practical application value.

The remainder of this paper is organized as follows: In Section III, the cold chain container multimodal route optimization problem considering the hybrid uncertainties of time and demand was modelled; in Section IV, a “dual-pheromone matrix” storage method was designed to realize the niche genetic algorithm–ant colony algorithm combined hybrid algorithm guided by the matching and search between target objects; in Section V, analysis was performed by taking the multimodal network in Chinese Circum-Bohai Sea Economic Zone as an example to verify the effectiveness of the proposed model and algorithm; in the final section, the relevant conclusions were drawn.

3. Modelling

3.1 Problem description and model assumptions

It is assumed that there are totally n transportation nodes in the multimodal transportation network in a region, with m modes of transportation between every two nodes. Thus, transportation is conducted using *teu* refrigerated containers. In consideration of the constraint conditions of the transportation network, the minimum cost and the minimum carbon emissions during the transportation process are achieved by selecting the combinations of different modes of transportation.

A batch of cold chain cargoes are transported from city O to city D , where O is the starting point of transportation ($O \in I$), D is the endpoint ($D \in I$) and I represents the transportation set. Several middle nodes are passed in the process of transportation; moreover, between any two nodes in the multimodal network are one or multiple selectable modes of transportation, such as highway, railway, and waterway, and the difference in the mode of transportation selected results in the differences in carbon dioxide emissions, transportation distance, and transportation cost, and transportation time and transshipment time are set to be random numbers. In this study, the minimization of total cost, including carbon emission cost, was solved, the decision on the multimodal transportation route and mode was made, and carriers were required to meet the time window limit of a certain confidence level when delivering goods to the endpoint. The following assumptions were made:

- The driving speed and the total transportation volume demanded under different modes of transportation on each road section follow a probability distribution.
- After transshipment is completed, emergency supplies can be transported to the next node only upon the nearest shift.
- The goods are no special requirements for the mode of transportation. All nodes except the starting point and endpoint can serve as transshipment nodes.
- At transshipment nodes, emergency supplies can only be transshipped to one node instead of split transshipment.

3.2 Model parameters and variables

Based on the problem description and model assumptions, the symbol descriptions are exhibited in Table 1.

Table 1 Symbol descriptions

Parameter	Parameter definition	Unit
I	Transportation node set, representing the locations passed during the transportation of goods, $i \in I$;	—
J	Set of modes of transportation, indicating the type of transportation mode, $j \in J, k \in J$;	—
teu	Number of cold chain transportation containers;	Ea
$x_{i,i+1}^j$	If the j -th mode of transportation is adopted from location i to location $i + 1$, the value is 1, or otherwise, it is 0;	—
$l_{i,i+1}^j$	Transportation distance under the j -th mode of transportation from location i to location $i + 1$;	—
$t_{i,i+1}^j$	Transportation time under the j -th mode of transportation from location i to location $i + 1$;	Min
$e_{i,i+1}^j$	Unit transportation cost under the j -th mode of transportation from location i to location $i + 1$;	yuan
$co_{i,i+1}^j$	Unit carbon emission under the j -th mode of transportation from location i to location $i + 1$;	kg
co_z	Unit carbon emission during the rehandling operation at nodes;	T
$y_i^{j,k}$	If the mode of transportation is changed from j to k at location i , the value is 1, or otherwise, it is 0;	—
$e_i^{j,k}$	Rehandling cost of unit goods if the mode of transportation is changed from j to k at location i . The value is zero if the mode of transportation at this location is unchanged;	—
$t_i^{j,k}$	Rehandling time of unit goods if the mode of transportation is changed from j to k at location i . The value is zero if the mode of transportation at this location is unchanged;	—
e_l	Refrigeration cost of cold chain containers within unit time;	yuan
co_l	Refrigeration-induced carbon emission of cold chain containers within unit time;	t
r_e	Carbon tax value;	yuan
γ_1	Storage management cost to be paid within unit time when the goods arrive at location i in advance;	yuan
γ_2	Penalty cost to be paid within unit time in case of delayed arrival of goods at location i ;	yuan
a_i	Earliest arrival time of goods allowed at location i ;	—
b_i	Latest arrival time of goods allowed at location i ;	—
v_j	The speed of the j -th mode of transportation.	—

3.3 Robust optimization model

Based on the actual operation process of cold chain container multimodal transportation and the above assumptions, the objective function of the model consists of the following parts:

- Transportation process cost C_1 , including internode transportation cost, cold chain equipment refrigeration cost, and rehandling cost at nodes, and each cost during the transportation process is expressed as follows:

$$\begin{aligned}
 C_1 = & \sum_{i=1}^n \sum_{j=1}^m e_{i,i+1}^j x_{i,i+1}^j l_{i,i+1}^j + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^m y_i^{j,k} e_i^{j,k} + e_l \left(\sum_{i=1}^n \sum_{j=1}^m t_{i,i+1}^j x_{i,i+1}^j \right. \\
 & \left. + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^m y_i^{j,k} t_i^{j,k} \right)
 \end{aligned} \quad (1)$$

- Carbon emission cost C_2 during transportation, including carbon emission cost during inter-node transportation, carbon emission cost induced by cold chain equipment refrigeration, and carbon emission cost induced by rehandling at nodes. The carbon emission cost during transportation is expressed as follows:

$$C_2 = \left(\sum_{i=1}^n \sum_{j=1}^m co_{i,i+1}^j x_{i,i+1}^j l_{i,i+1}^j + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^m y_i^{j,k} co_z \right. \\ \left. + co_l \left(\sum_{i=1}^n \sum_{j=1}^m t_{i,i+1}^j x_{i,i+1}^j \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^m y_i^{j,k} t_i^{j,k} \right) \right) r_e \quad (2)$$

- Time window cost C_3 , including the storage cost because of advanced arrival and the penalty cost because of delayed arrival, specifically as follows:

$$C_3 = \gamma_1 \sum_{i=1}^n \max(a_i - t_i, 0) + \gamma_2 \sum_{i=1}^n \max(t_i - b_i, 0) \quad (3)$$

- Cargo damage cost C_4 . Cold chain goods are sensitive to environmental changes, and time and temperature changes probably lead to the deterioration of goods. In this study, the cargo damage cost was calculated via Arrhenius equation as follows:

$$C_4 = P_{teu} [1 - e^{-\mu \frac{l_{i,i+1}^j}{v_j}}] \quad (4)$$

where μ represents the sensitivity coefficient.

The objective functions of the model are displayed as follows:

$$\min C = teu(C_1 + C_2 + C_3 + C_4) \quad (5)$$

s. t.

$$x_{i,i+1}^j + x_{i-1,i}^j \geq 2y_i^{j,k} \quad \forall i \in I, \forall j, k \in J \quad (6)$$

$$t_{i,i+1}^j \geq 0, t_i^{j,k} \geq 0, l_{i,i+1}^j \geq 0, teu \geq 0 \quad (7)$$

$$t_{i,i+1}^j = \frac{l_{i,i+1}^j}{v_j} \quad \forall i \in I, \forall j \in J \quad (8)$$

$$\sum_{j=1}^m \sum_{k=1}^m y_i^{j,k} = 1 \quad \forall i \in I, \forall j, k \in J \quad (9)$$

$$x_{i,i+1}^j (t_{i+1} - t_i) \geq 0 \quad \forall i \in I \quad (10)$$

$$\sum_{j=1}^m x_{i,i+1}^j = 1 \quad \forall i \in I, \forall j \in J \quad (11)$$

$$t_i^{j,k} \sim N(\mu_i^{j,k}, (\sigma_i^{j,k})^2) \quad (12)$$

$$t_{i,i+1}^j \sim N(\mu_{i,i+1}^j, (\sigma_{i,i+1}^j)^2) \quad (13)$$

$$x_{i,i+1}^j = \begin{cases} 1, & \text{transportation mode } j \text{ is selected between node } i \text{ and } i+1 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

$$y_i^{j,k} = \begin{cases} 1, & \text{transportation mode is changed from } j \text{ to } k \text{ at the node } i \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

Eq. 5 is the objective function, which consists of four parts: transportation cost, time window cost, carbon emission cost, and damage cost of cold chain products. Eq. 6 aims to ensure the continuity and accuracy of the conversion between various modes of transportation in the process of cold chain multimodal transportation. Eq. 7 ensures that the transportation time, transshipment time, transportation distance, and transportation volume are positive values. Eq. 8 is the transportation time of goods between two adjacent nodes (i and $i+1$), which is related to the speed. Eq. 9 means that the cold chain container can only be converted from the current mode of transportation to another mode of transportation at transshipment nodes. Eq. 10 indicates the

continuity of cargo transportation time, and the time when the cold chain container arrives at the back node ($i + 1$) is later than the time when the cargo arrives at the previous node (i). Eq. 11 indicates that only one mode of transportation can be selected between adjacent nodes. Eqs. 12 and 13 mean that transshipment time and transportation time obey a normal distribution. Eq. 14 is a value constraint of the decision variable, representing whether a certain mode of transportation is used between two nodes. Eq. 15 is also a value constraint for decision variables, representing whether to change the mode of transportation at a certain node.

4. Algorithm design

4.1 Description of hybrid algorithm

In accordance with the characteristics of the problem, the “dual-pheromone matrix” storage method was designed to realize the “interactive selection” of routes and modes of transportation. Next, a hybrid algorithm combining genetic algorithm and ant colony algorithm was designed. First, the similarity between schemes was reduced through the niche technology based on genetic algorithm, and multiple groups of feasible solutions that contained more representative solution spaces were obtained; second, each optimal route and mode of transportation were combined, and the dual-pheromone heuristic matrix in ant colony algorithm was updated so that the initial ant colony possessed an improved matching basis for the route and mode of transportation to guide and accelerate the route optimization process of ants and further enhance the solution accuracy and convergence rate of the hybrid algorithm.

4.2 Niche genetic algorithm

In this study, the niche genetic algorithm was introduced into the ant colony algorithm. To prevent the algorithm from sampling the fixed area of the target space [16], two-point mutation and mask crossover were designed as genetic operators, and the diversity of the population was guaranteed, as shown in Fig. 1.

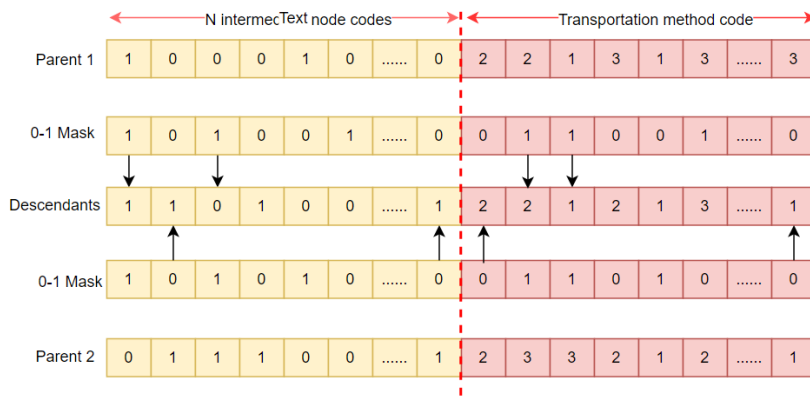


Fig. 1 Schematic of chromosome coding

4.3 Dynamic dual-pheromone ant colony algorithm

In this study, the “dual-pheromone matrix” storage method was designed, aiming to eliminate the blindness of route mode of transportation matching. As for the dual-pheromone storage structure, the heuristic information $\eta_{ij}(t)$ at time t indicates the probability of route selection, pheromone concentration $\tau_{ij}(t)$ denotes the probability for the selection of modes of transportation, $p_{ij}^a(t)$ is the probability for ant a to transfer from node i to node j at time t , and the probability for the selectable mode of transportation that matches with the selected arc segment (i, j) can be solved in a similar way.

$$p_{ij}^a(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta}{\sum_{r \in allowed_a} [\tau_{ir}(t)]^\alpha [\eta_{ir}(t)]^\beta} & j \in allowed_a \\ 0 & j \notin allowed_a \end{cases} \quad (16)$$

Where the set of nodes or modes of transportation that can be selected by ant a is expressed as $\tau_{ij}(t+1) = (1-\rho) \times \tau_{ij}(t) + \Delta\tau_{best}$; $\tau_{ij}(t)$ stands for the pheromone intensity of route (i, j) at time t ; $\tau_{ij}(t+1) = (1-\rho) \times \tau_{ij}(t) + \Delta\tau_{best}$ is the influencing factor, indicating the reliance on the pheromone matrix and heuristic matrix during node selection by ants; $\tau_{ij}(t+1) = (1-\rho) \times \tau_{ij}(t) + \Delta\tau_{best}$ is the influencing factor, reflecting the reliance of the mode of transportation on the pheromone matrix and heuristic matrix [26].

The max-min ant strategy is adopted to avoid the stagnation of algorithm search in the later stage, and the pheromone value of each arc and each mode of transportation must be within the interval of (τ_{min}, τ_{max}) , where $L(S_{best})$ is the favourable solution of each niche or the best solution of the ant colony algorithm in the current iteration. The range of initial pheromone and the solution obtained from the niche genetic algorithm are determined as per Eqs. 17 and 18 [29].

$$\tau_{max}(t) = \frac{1}{2(1-\rho)} \times \frac{1}{L(S_{best})} \quad (17)$$

$$\tau_{min}(t) = \frac{\tau_{max}(t)}{20} \quad (18)$$

After the niche genetic algorithm ends, τ_{min} is determined via Eq. 18 and τ_{max} via Eq. 19.

$$\tau_{max}(t) = \frac{1}{2(1-\rho)} \times \frac{1}{L(S_{best})} + \frac{1}{L(S_{best})} \quad (19)$$

4.4 Improved time-varying pheromone evaporation operator

By analysing the factors influencing the algorithm search space and algorithm convergence, the importance of the pheromone evaporation factor in the selection of route or mode of transportation is found to be directly proportional [32]. In the niche genetic algorithm and ant colony algorithm mechanism designed in this study, therefore, the pheromone evaporation factor ρ increases with the increase in the number of iterations during algorithm operation [31].

$$\rho = \rho_0(1 - \mu_0^{-\lambda}) \quad (20)$$

where μ_0 is the pheromone evaporation speed; λ denotes the number of iterations performed by the ant colony algorithm; ρ_0 is the initial pheromone evaporation factor. μ_0 controls the pheromone evaporation factor to finally converge to ρ_0 , and the value of μ_0 is directly proportional to the increase rate of the pheromone evaporation factor with the increase in the number of algorithm iterations [33].

4.5 Dual-pheromone hybrid algorithm flow

The dual-pheromone hybrid algorithm flow is displayed in Fig. 2.

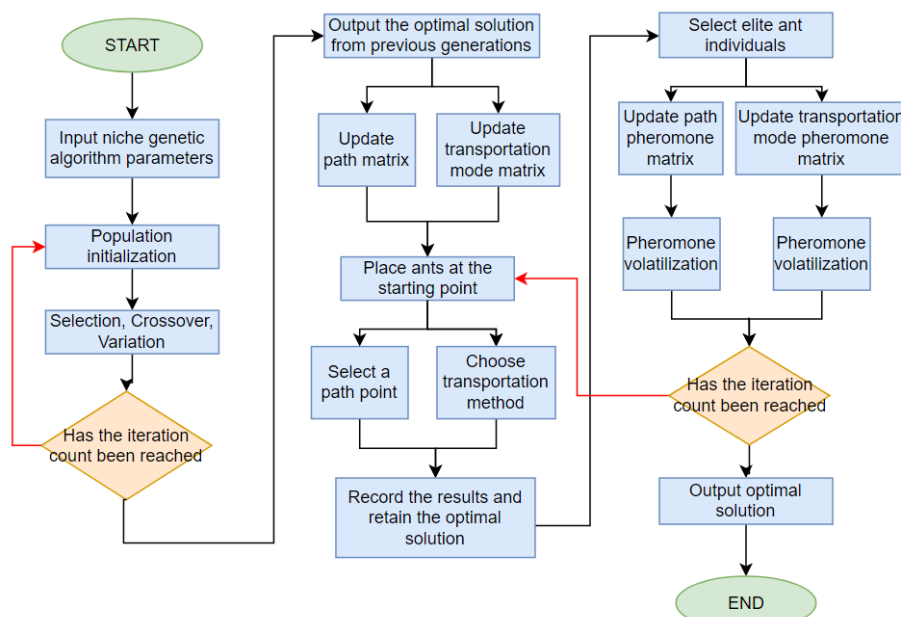


Fig. 2 Flowchart for solving the multimodal route optimization problem based on an improved genetic algorithm

5. Example analysis

5.1 Data description

With the Chinese Circum-Bohai Sea Economic Zone taken as the study object, the most important dense extravert and multifunctional urban agglomerations in Northern China were constituted with Beijing and Tianjin as the center, coastal cities such as Dalian, Qingdao, Yantai and Weihai as the sector, and provincial cities such as Shenyang, Jinan, and Shijiazhuang as the regional pivots. These urban agglomerations exerted the agglomerating, radiating, serving, and driving effects in nationwide and regional economy, powerfully facilitating the development of their characteristic economic zones. The location map of the Circum-Bohai Sea Economic Zone is shown in Fig. 3. The shortest route mileages among 35 main nodes under such modes of transportation as highway, railway, and waterway were acquired in accordance with the China railway information network, Baidu Map, and relevant electronic maps. Among the 35 cities, transportation could be realized between any two cities by means of highway and railway, and no waterway transportation existed between some cities. The multimodal network with 35 nodes was simulated to investigate the feasibility of the model and algorithm. The basic data tables were omitted due to the limitations on the length of the paper.

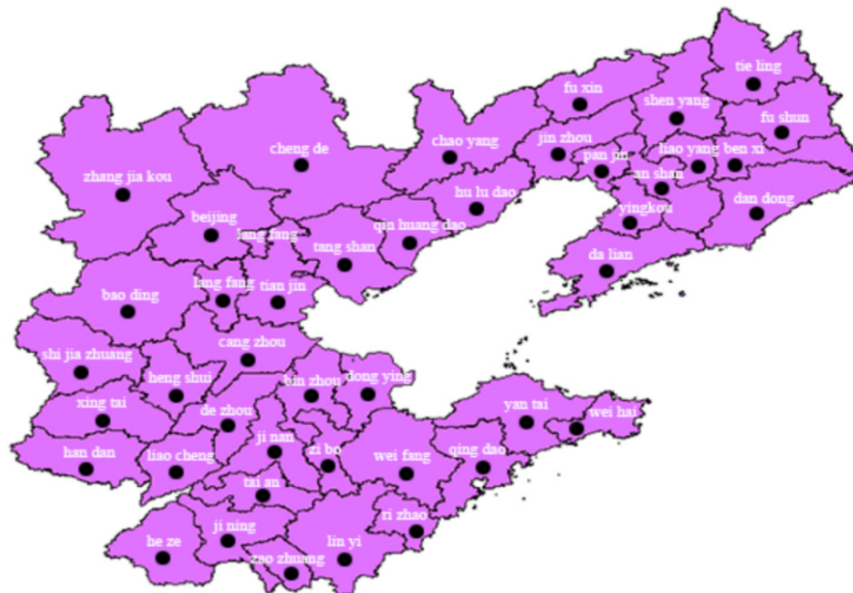


Fig. 3 Multimodal transportation network

5.2 Example design

Table 2 shows three transportation tasks undertaken by a transportation company. The unit warehousing cost and penalty cost induced by the advanced and delayed arrival of goods at the endpoint relative to the customer-expected arrival time are 60 and 100 yuan/teu*h, respectively. In reference to the actual transportation speed under each mode of transportation in China, the highway, railway, and waterway transportation speeds were respectively taken as 90, 65, and 35 km/h. The ratio of transportation time to transportation speed is the reference transportation time T , and the reference volume of goods transported is $Q = 28$.

Table 2 Table of transportation tasks

Transportation task	Starting point	Endpoint	Quanty (teu)	Cargo value
Task 1	Qingdao	Chengde	5	150,000/teu
Task 2	Qingdao	Beijing	5	150,000/teu
Task 3	Qingdao	Shenyang	8	150,000/teu

Four possible demand scenarios for the transportation volume in the current transportation task were assumed, and the probability for their occurrence was 0.1, 0.2, 0.3, and 0.4, respectively. The robust optimization model based on scenario sets was proposed by Mulvey *et al.* [34],

which divided scenario sets according to data characteristics so that they could cover all the possible occurrence conditions of events. The customer demand under each scenario is listed in Table 3.

Table 3 Customer demand under each scenario

Transportation task	Scenario I	Scenario II	Scenario III	Scenario IV	Cargo value	
Task 1	5	6	4	4	150,000/teu	5.3
Task 2	8	4	5	4	150,000/teu	5.5
Task 3	7	10	7	6	150,000/teu	8.4

The total value of each teu of cold chain goods is 150000 yuan, with a weight of 20 tons. A total of 33 node cities were chosen as the possible transshipment nodes and numbered as 1, 2...33 for the convenience of expression. The average data of transportation time between two nodes under different modes of transportation are difficult to obtain; thus, the ratio of transportation distance to transportation speed was assumed to be the average value of transportation time, and the variance of highway transportation, railway transportation, and waterway transportation was taken as 0.52, 0.22, and 0.752, respectively. Table 4 shows the relevant parameters of three transportation modes.

Table 4 Profiles of three modes of transportation

Mode of transportation	Highway	Railway	Waterway
Unit freight charge (yuan/teu*km)	8.5	2.025	0.5
Transportation speed (km/h)	(100, 0.52)	(60, 0.22)	(30, 0.752)
Carbon emission (teu/tkm)	1.4	0.412	0.332

For the railway line existing between nodes, a nonstop train can be started between some nodes, and this train does not perform unloading or transshipment operation at the middle node, and it is not affected by the departure time of other trains at the transshipment node. In this case, the transportation time of the nonstop train is the sum of transportation time on each section of the railway line, and the unit cargo transportation cost is the average value of the unit cargo transportation cost on each section of the railway line. The transshipment cost between different modes of transportation, the carbon emission, and the transshipment time are exhibited in Table 5.

Table 5 Rehandling circumstances between different modes of transportation

Mode of transportation	Highway			Railway			Waterway		
	Cost (yuan)	Carbon emission (kg)	Time (h)	Cost (yuan)	Carbon emission (kg)	Time (h)	Cost (yuan)	Carbon emission (kg)	Time (h)
Highway	0	0	0	30	1.4	(0.1, 12)	50	3.5	(0.3, 12)
Railway	30	1.4	(0.1, 12)	0	0	0	60	4.2	(0.4, 12)
Waterway	50	3.5	(0.3, 12)	60	4.2	(0.4, 12)	0	0	0

5.3 Model solving

Certainty model solving

For the certainty model, the influence of transportation volume on the cold chain container multimodal route is not considered. Under the certainty model, it takes 11.10 s to solve three tasks via MATLAB2014b, and the optimal multimodal route is displayed in Fig. 4. The convergence curve of the multimodal transportation cost for the three tasks under the certainty model is exhibited in Fig. 5, and the minimum total cost for the three tasks is 79,789 yuan.

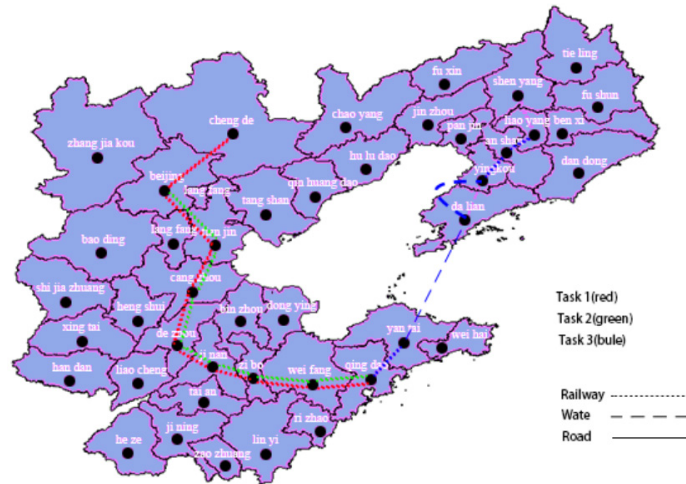


Fig. 4 Multimodal route with the minimum total cost under certainty model

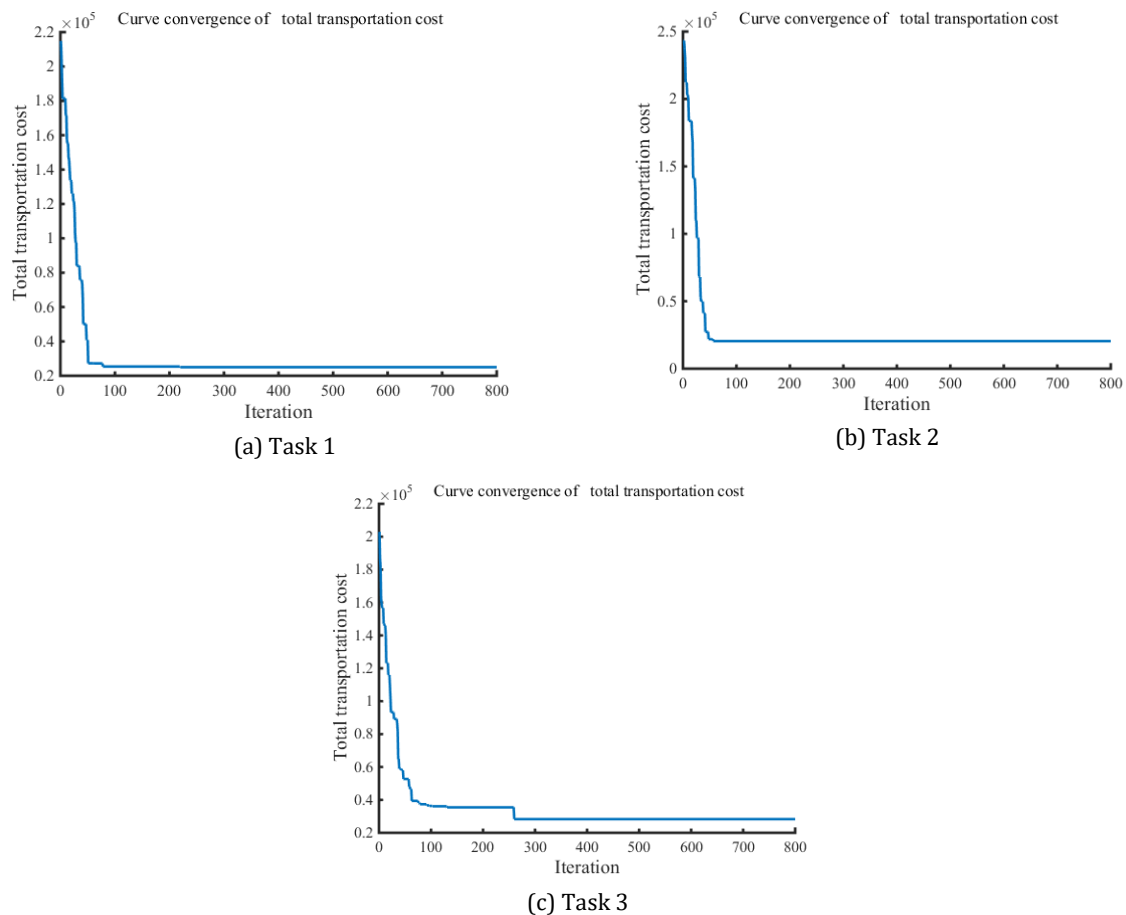


Fig. 5 Convergence curve of the multimodal transportation cost for three tasks under the static model

Time uncertainty model solving

The influence of different transportation task characteristics on the transportation scheme was compared by designing three different types of examples according to the different transportation tasks and transportation scales, that is, according to the number of transportation tasks and the size of transportation batches. To simulate the uncertainty of transportation time and re-handling time, transportation time and rehandling time were processed using fuzzy time, the average value of transportation time was the ratio of transportation distance to transportation speed, and the variance of highway transportation, railway transportation, and waterway transportation was 0.52, 0.22, and 0.752, respectively. The average value of transportation time was

changed by adjusting the average value of transportation speed according to the transportation time. The average transshipment time and variance were acquired by analysing the statistical data on container multimodal transshipment time at one transshipment node. At the confidence level of 90 %, the transportation time and transshipment time were lengthened to solve the time uncertainty mode, and the total cost of the three tasks was 82,672 yuan, as shown in Fig. 6. The convergence curve of each task is displayed in Fig. 7.

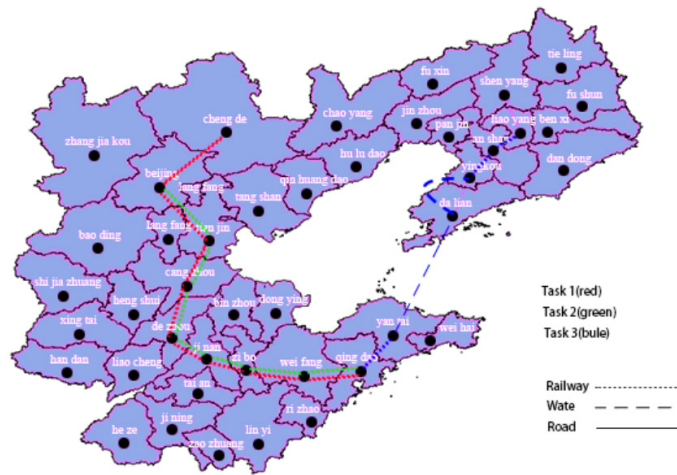


Fig. 6 Multimodal route with the minimum total cost under fuzzy time model

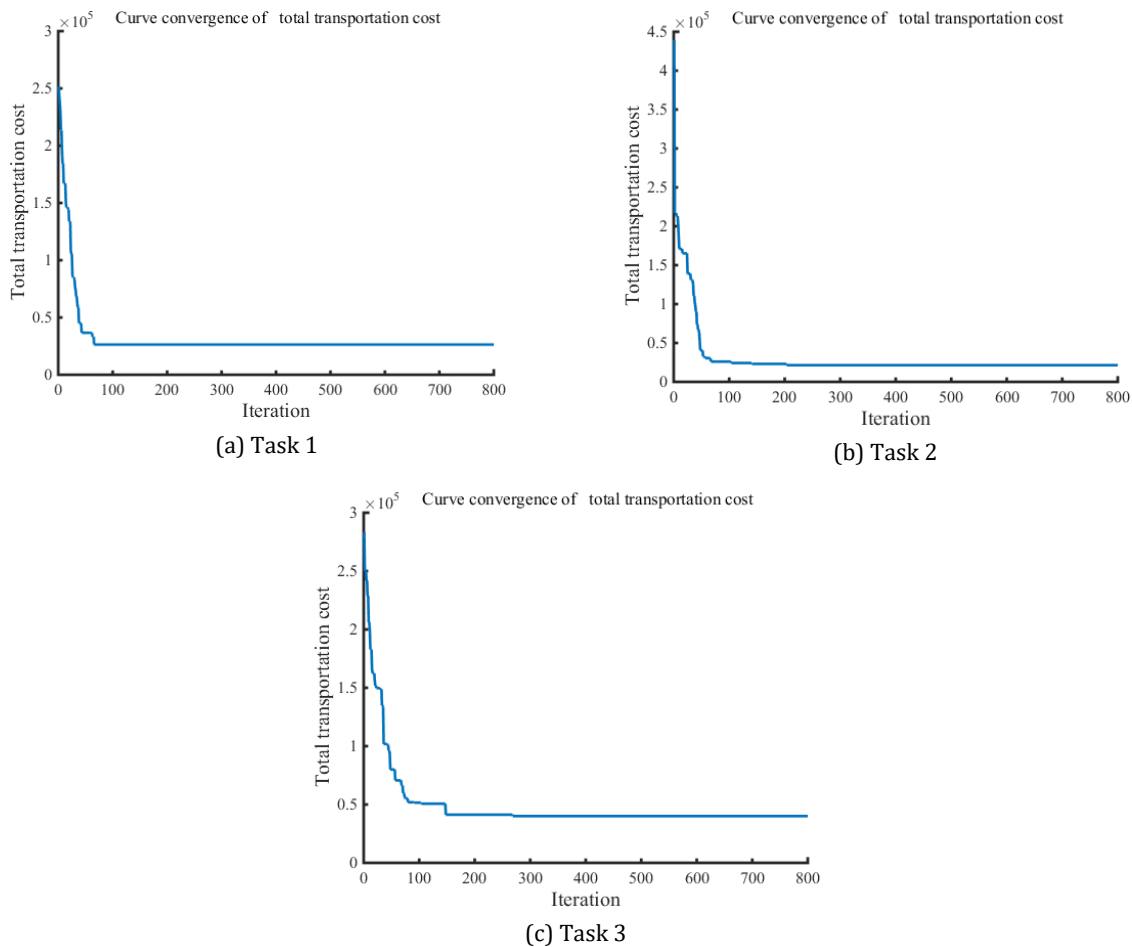


Fig. 7 Convergence curve of multimodal transportation cost for three tasks under the static model

Demand uncertainty model solving

In case of any change in the transportation volume, the optimal multimodal route changes if the optimization scheme under static environments is continuously executed. Under the dynamic model, it takes 10.094210 s to solve the three tasks via MATLAB2014b. The optimal multimodal route is as shown in Fig. 8, and the minimum total cost of the three tasks is 88,663 yuan. The convergence curve of the total multimodal transportation cost for the three tasks under the dynamic model is exhibited in Fig. 9.

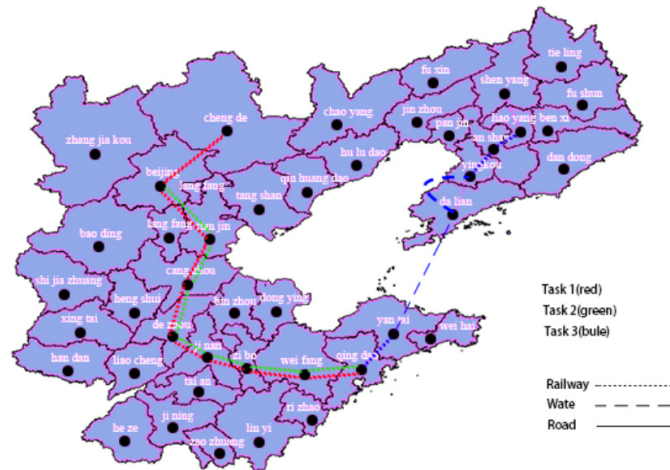


Fig. 8 Cold chain multimodal route considering demand uncertainty

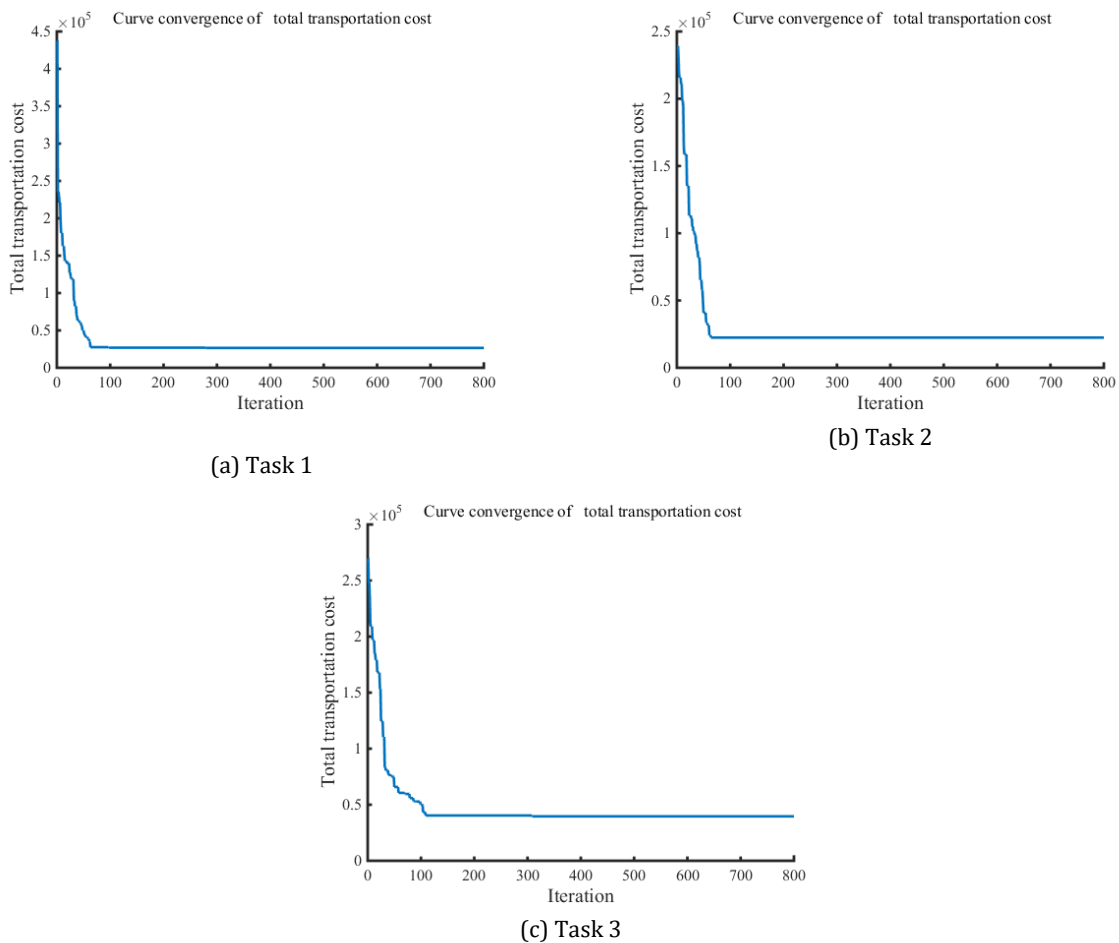


Fig. 9 Convergence curve of multimodal transportation cost for three tasks under the dynamic model

The solution that is feasible in all transportation scenarios and satisfies the regret coefficient constraint is referred to as a robust feasible solution, while the robust optimal solution is that with the lowest expected cost in all transportation scenarios. The robust feasible solution and the robust optimal solution are related to the regret coefficient, and the robust solution changes under different regret value coefficients but the same transportation scenario. In case of a too small regret value coefficient, if the regret constraint for the solution is too strict, then the model has no solution. As the regret value coefficient gradually increases, the robustness constraint for the solution is loosened, the number of robust feasible solutions grows, and the robust optimal solution changes accordingly. Four transportation scenario sets were respectively substituted into the certainty model to calculate the corresponding optimal value and compared with the robust optimization model to determine the advantages of robust models; the results are exhibited in Fig. 10.

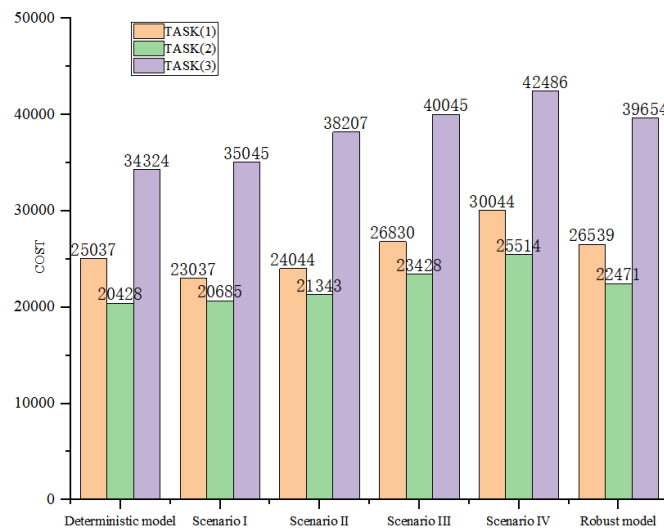


Fig. 10 Comparison of robust optimization results

Hybrid uncertainty model solving

To explore the influences of demand uncertainty and time uncertainty on the cold chain container multimodal route, it takes 13.50 s to solve the three tasks in MATLAB2014b through the hybrid algorithm designed in the previous part, the optimal multimodal route for the three tasks under the hybrid uncertainty model is as shown in Fig. 11, and the minimum total cost of the three tasks is 88,664 yuan. The convergence curve of the multimodal transportation cost for the three tasks under the hybrid uncertainty model is displayed in Fig. 12.

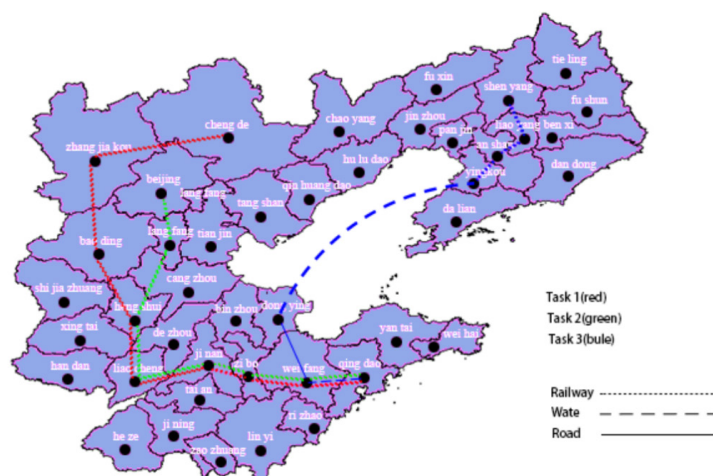


Fig. 11 Cold chain multimodal route considering hybrid uncertainties

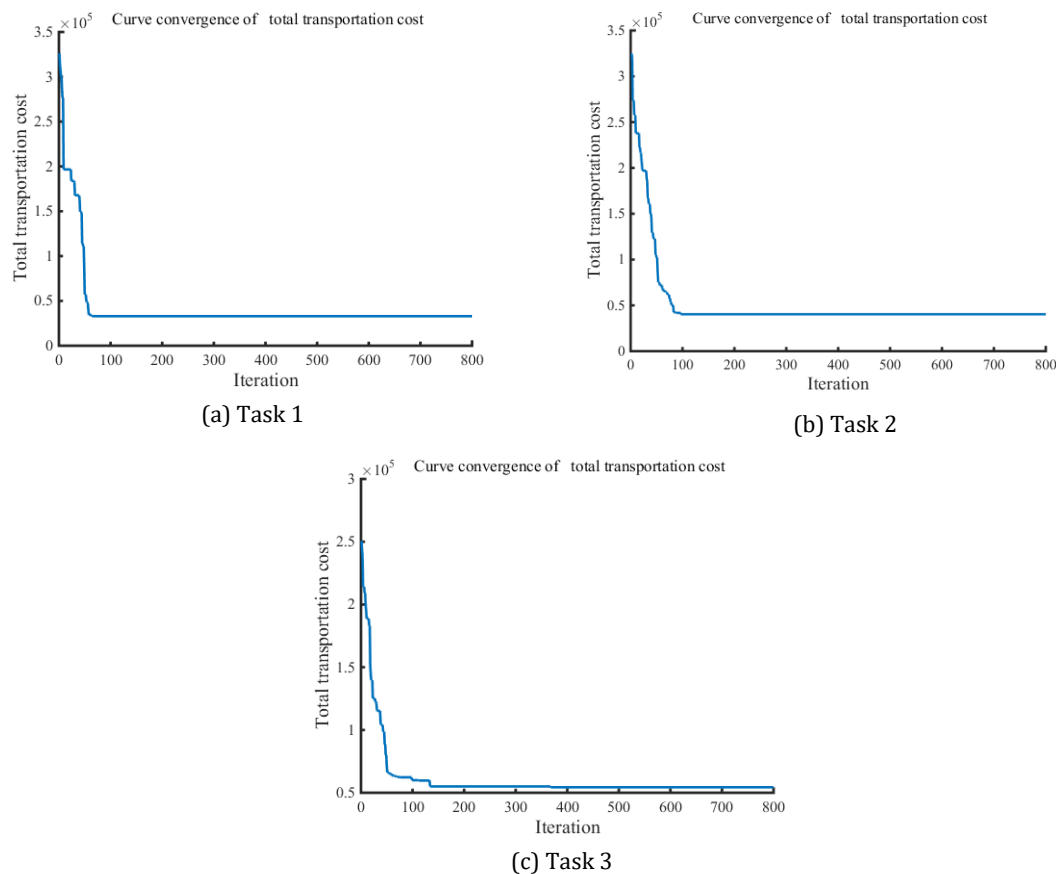


Fig. 12 Convergence curve of multimodal transportation cost for three tasks under hybrid uncertainties

Result analysis and comparison of examples

In this study, various costs involved in the multimodal network were analysed when dual hybrid uncertain factors followed a random distribution, and the different results acquired are displayed in Fig. 13. Three transportation tasks were compared by comparatively analysing the changes in the multimodal route programming with time uncertainty and demand uncertainty under four scenarios. When the transportation time and transshipment time changes, the total multimodal transportation cost under the time uncertainty model increased slightly. In case of demand changes, the total multimodal transportation cost under the demand uncertainty model also grew slightly. When time uncertainty or demand uncertainty was independently considered, the transportation route and the corresponding mode of transportation did not change mainly for two reasons: the transportation volume under each mode of transportation in the multimodal network was not constrained during model design; the transportation time and transshipment time under each mode of transportation changed inconsequentially. When time uncertainty and demand uncertainty were simultaneously considered, the total multimodal transportation cost grew remarkably under the hybrid uncertainty model, and the transportation route and the corresponding mode of transportation also changed evidently under the hybrid uncertainty scenario. The reasons were explained as follows: the comprehensive transportation volume for cold chain container multimodal transportation tasks was acquired by fully considering all scenarios of the robust model; the increasing demand led to the increase in the total cost of cold chain container multimodal transportation. Therefore, the gap in the total cost of cold chain container multimodal transportation under all scenarios in the robust model is stable, being always acceptable, and this verifies the interference immunity and high-cost characteristic of the robust model.

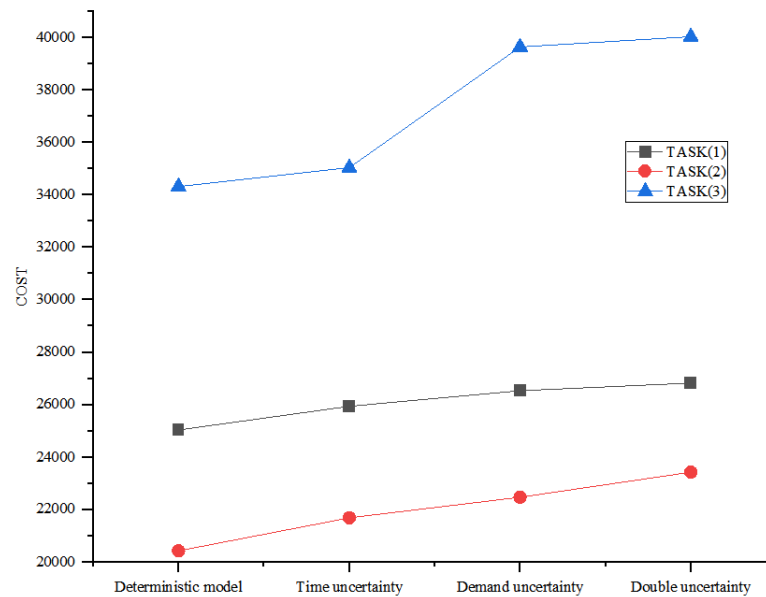


Fig. 13 Comparison of costs under four scenarios

6. Conclusions

The problem of cold chain container multimodal route optimization considering carbon emissions under the hybrid uncertainties of transportation time and demand was explored in this study. Then, a robust optimization model for the low-carbon cold chain container multimodal route under hybrid uncertainties was constructed on the basis of the robust optimization theory and solved by designing a hybrid algorithm based on the niche genetic algorithm and the dynamic dual-phormone ant colony algorithm. Finally, the following conclusions were drawn:

- The uncertainties in transportation time and transshipment time result in the changes in transportation cost, carbon emission cost, and transportation scheme.
- When the problem of cold chain container multimodal route optimization under demand uncertainty is solved using a robust model, the decision-maker must predict the fluctuation range of demand uncertainty: the smaller the fluctuation range, the easier it is to obtain the robust optimal solution under a small regret coefficient value, thus enhancing the quality of the robust optimal solution.
- Studying the problem of cold chain container multimodal route optimization considering carbon emissions under the hybrid uncertainties of time and demand reveals that the carbon emission cost and total cost of cold chain multimodal transportation under hybrid uncertainties increase remarkably, accompanied by the evident changes in the transportation route and the corresponding mode of transportation.

Therefore, when formulating the intermodal transportation scheme, the decision-maker must consider the hybrid time windows at nodes and the endpoint according to the actual situation. Given the different influences of different factors on the robustness of the solution, the decision-maker should respectively determine the fluctuation range of different uncertain factors when using a robust optimization model. The study results can be used as reference to solve the problem of multimodal route optimization under uncertain environments. However, in a cold chain multimodal network, the cold chain multimodal route is affected by such factors as risk uncertainty, transportation price uncertainty, and customer satisfaction in addition to time and demand uncertainties. In a follow-up study, this problem remains to be further analysed and studied by combining the queuing, disassembly, and marshalling situations of trains at nodes.

References

- [1] Lu, J., Lin, W., Dabić, M. (2024). Coevolution analysis of the sustainable development systems of energy firms from a strategic philanthropy perspective, *IEEE Transactions on Engineering Management*, Vol. 71, 6887-6902, [doi: 10.1109/TEM.2023.3244500](https://doi.org/10.1109/TEM.2023.3244500).
- [2] Wang, X., Ma, J., Yang, K. (2024). Domestic and foreign research progress on inland ports and research prospect for chinese inland ports: Based on CiteSpace bibliometric graphing analysis, *Journal of Beijing Jiaotong University(Social Sciences Edition)*, Vol. 23, No. 2, 80-88, [doi: 10.16797/j.cnki.11-5224/c.20240513.014](https://doi.org/10.16797/j.cnki.11-5224/c.20240513.014).
- [3] Yuan, R., Wang, W., Li, J., Zhao, Q. (2023). Multi-stage emergency materials scheduling based on multimodal transportation under uncertainty, *Operations Research and Management Science*, Vol. 32, No. 6, 33-39, [doi: 10.12005/orms.2023.0179](https://doi.org/10.12005/orms.2023.0179).
- [4] Yi, S. (2022). Development mode and countermeasure system of multimodal transport under the new development pattern, *Hebei Academic Journal*, Vol. 42, No. 5, 146-154.
- [5] Xu, Y., Jia, J. (2021). Sino-Russia joint rail and water transportation legislation in the B&R initiative, *Academic Exchange*, No. 6, 45-55, [doi: 10.3969/j.issn.1000-8284.2021.06.005](https://doi.org/10.3969/j.issn.1000-8284.2021.06.005).
- [6] Qiu, Y., Hu, J., Guo, J., Liu, K. (2024). Multimodal transport path selection method considering time value of goods under uncertain demand, *Journal of Railway Science and Engineering*, Vol. 21, No. 3, 994-1003, [doi: 10.19713/j.cnki.43-1423/u.T20230828](https://doi.org/10.19713/j.cnki.43-1423/u.T20230828).
- [7] Deng, M., Dai, Y., Li, X. (2023). Robust optimization of multi-modal transportation routing with low-carbon under demand uncertainty, *Industrial Engineering Journal*, Vol. 26, No. 4, 104-113, [doi: 10.3969/j.issn.1007-7375.2023.04.013](https://doi.org/10.3969/j.issn.1007-7375.2023.04.013).
- [8] Li, M., Sun, X. (2022). Path optimization of low-carbon container multimodal transport under uncertain conditions, *Sustainability*, Vol. 14, No. 21, Article No. 14098, [doi: 10.3390/su142114098](https://doi.org/10.3390/su142114098).
- [9] Wang, J. (2023). *Optimization of low-carbon multimodal transport routes for containers considering uncertain demand*, Master's thesis, Lanzhou Jiaotong University, Lan Zhou, China.
- [10] Tang, R., Hu, J., Wen, C., Zhang, W. (2023). Multimodal transport route selection considering dynamic transit time under uncertain demand, *Journal of Nanjing Tech University (Natural Science Edition)*, Vol. 45, No. 5, 554-562, [doi: 10.3969/j.issn.1671-7627.2023.05.010](https://doi.org/10.3969/j.issn.1671-7627.2023.05.010).
- [11] Li, S., Chen, Z., Liu, L., Wu, Y., Sun, S. (2018). Intermodal logistic network design under uncertain demand, *Journal of Donghua University (Natural Science Edition)*, Vol. 44, No. 4, 550-554, [doi: 10.3969/j.issn.1671-0444.2018.04.008](https://doi.org/10.3969/j.issn.1671-0444.2018.04.008).
- [12] Li, S.-X., Sun, S.-F., Wang, Y.-Q., Wu, Y.F., Liu, L.-P. (2017). A two-stage stochastic programming model for rail-truck intermodal network design with uncertain customer demand, *Journal of Interdisciplinary Mathematics*, Vol. 20, No. 3, 611-621, [doi: 10.1080/09720502.2016.1258831](https://doi.org/10.1080/09720502.2016.1258831).
- [13] Zhang, D., Ni, N., Lai, X., Liu, Y. (2020). Multimodal iron ore inbound logistics network design under demand uncertainty, *Maritime Policy & Management*, Vol. 48, No. 7, 941-965, [doi: 10.1080/03088839.2020.1791991](https://doi.org/10.1080/03088839.2020.1791991).
- [14] Chen, D. (2020). *Robust optimisation of container multimodal transportation route under uncertainty*, Doctoral dissertation, Southeast University, Nan Jing, China, [doi: 10.27014/d.cnki.gdnau.2020.002963](https://doi.org/10.27014/d.cnki.gdnau.2020.002963).
- [15] Hu, Y., Luo, Y., Li, X. (2018). Research on reverse logistics network design in uncertain environment based on hybrid algorithm, *Industrial Engineering and Management*, Vol. 23, No. 1, 90-95, [doi: 10.19495/j.cnki.1007-5429.2018.01.014](https://doi.org/10.19495/j.cnki.1007-5429.2018.01.014).
- [16] Zhang, X., Jin, F.-Y., Yuan, X.-M., Zhang, H.-Y. (2021). Low-carbon multimodal transportation path optimization under dual uncertainty of demand and time, *Sustainability*, Vol. 13, No. 15, Article No. 8180, [doi: 10.3390/su13158180](https://doi.org/10.3390/su13158180).
- [17] Xie, L., Cao, C. (2021). Multi-modal and multi-route transportation problem for hazardous materials under uncertainty, *Engineering Optimization*, Vol. 53, No. 12, 2180-2200, [doi: 10.1080/0305215X.2020.1850708](https://doi.org/10.1080/0305215X.2020.1850708).
- [18] Li, J., Wen, X., Liang, X. (2024). Route optimization of multimodal transportation for containers in the Yangtze river with uncertain freight rates, *Industrial Engineering Journal*, Vol. 27, No. 2, 138-146, [doi: 10.3969/j.issn.1007-7375.220245](https://doi.org/10.3969/j.issn.1007-7375.220245).
- [19] Pian, F., Chen, Y., Pang, S., Su, M. (2022). Game pricing of container road and multimodal transport considering railway discount, *Journal of Transportation Systems Engineering and Information Technology*, Vol. 22, No. 4, 1-10, [doi: 10.16097/j.cnki.1009-6744.2022.04.001](https://doi.org/10.16097/j.cnki.1009-6744.2022.04.001).
- [20] Aradi, S. (2022). Survey of deep reinforcement learning for motion planning of autonomous vehicles, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 2, 740-759, [doi: 10.1109/TITS.2020.3024655](https://doi.org/10.1109/TITS.2020.3024655).
- [21] Aghabayk, K., Radmehr, K., Shiwakoti, N. (2020). Effect of intersecting angle on pedestrian crowd flow under normal and evacuation conditions, *Sustainability*, Vol. 12, No. 4, Article No. 1301, [doi: 10.3390/su12041301](https://doi.org/10.3390/su12041301).
- [22] Norouzi, N., Sadegh-Amalnick, M., Alinaghiyan, M. (2015). Evaluating of the particle swarm optimization in a periodic vehicle routing problem, *Measurement*, Vol. 62, 162-169, [doi: 10.1016/j.measurement.2014.10.024](https://doi.org/10.1016/j.measurement.2014.10.024).
- [23] Ehrgott, M., Wang, J.Y.T., Raith, A., van Houtte, C. (2012). A bi-objective cyclist route choice model, *Transportation Research Part A: Policy and Practice*, Vol. 46, No. 4, 652-663, [doi: 10.1016/j.tra.2011.11.015](https://doi.org/10.1016/j.tra.2011.11.015).
- [24] Gupta, S., Ali, I., Ahmed, A. (2018). Multi-objective capacitated transportation problem with mixed constraint: A case study of certain and uncertain environment, *Opsearch*, Vol. 55, 447-477, [doi: 10.1007/S12597-018-0330-4](https://doi.org/10.1007/S12597-018-0330-4).
- [25] Gupta, S., Garg, H., Chaudhary, S. (2020). Parameter estimation and optimization of multi-objective capacitated stochastic transportation problem for gamma distribution, *Complex & Intelligent Systems*, Vol. 6, No. 3, 651-667, [doi: 10.1007/s40747-020-00156-1](https://doi.org/10.1007/s40747-020-00156-1).

- [26] Mardanya, D., Maity, G., Roy, S.K., (2021). Solving bi-level multi-objective transportation problem under fuzziness, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 29, No. 3, 411-433, [doi: 10.1142/S0218488521500185](https://doi.org/10.1142/S0218488521500185).
- [27] Biswas, A., Shaikh, A.A., Niaki, S.T.A. (2019). Multi-objective non-linear fixed charge transportation problem with multiple modes of transportation in crisp and interval environments, *Applied Soft Computing*, Vol. 80, 628-649, [doi: 10.1016/j.asoc.2019.04.011](https://doi.org/10.1016/j.asoc.2019.04.011).
- [28] Liu, S. (2023). Multimodal transportation route optimization of cold chain container in time-varying network considering carbon emissions, *Sustainability*, Vol. 15, No. 5, Article No. 4435, [doi: 10.3390/su15054435](https://doi.org/10.3390/su15054435).
- [29] Das, S.K., Roy, S.K., Weber, G.-W. (2020). Application of type-2 fuzzy logic to a multiobjective green solid transportation-location problem with dwell time under carbon tax, cap, and offset policy: Fuzzy versus nonfuzzy techniques, *IEEE Transactions on Fuzzy Systems*, Vol. 28, No. 11, 2711-2725, [doi: 10.1109/TFUZZ.2020.3011745](https://doi.org/10.1109/TFUZZ.2020.3011745).
- [30] Maity, G., Roy, S.K., Verdegay, J.L. (2020). Analyzing multimodal transportation problem and its application to artificial intelligence, *Neural Computing & Applications*, Vol. 32, No. 4, 2243-2256, [doi: 10.1007/s00521-019-04393-5](https://doi.org/10.1007/s00521-019-04393-5).
- [31] Roy, S.K., Maity, G., Weber, G.W., Gök, S.Z.A. (2017). Conic scalarization approach to solve multi-choice multi-objective transportation problem with interval goal, *Annals of Operations Research*, Vol. 253, No. 1, 599-620, [doi: 10.1007/s10479-016-2283-4](https://doi.org/10.1007/s10479-016-2283-4).
- [32] Han, X., Zhao, P.X., Kong, D.X. (2022). A bi-objective optimization of airport ferry vehicle scheduling based on heuristic algorithm: A real data case study, *Advances in Production Engineering & Management*, Vol. 17, No. 2, 183-192, [doi: 10.14743/apem2022.2.429](https://doi.org/10.14743/apem2022.2.429).
- [33] Guzman, E., Poler, R., Andres, B. (2023). A matheuristic approach combining genetic algorithm and mixed integer linear programming model for production and distribution planning in the supply chain, *Advances in Production Engineering & Management*, Vol. 18, No. 1, 19-31, [doi: 10.14743/apem2023.1.454](https://doi.org/10.14743/apem2023.1.454).
- [34] Mulvey, J.M., Vanderbei, R.J., Zenios, S.A. (1995). Robust optimization of large-scale systems, *Operations Research*, Vol. 43, No. 2, 264-281, [doi: 10.1287/opre.43.2.264](https://doi.org/10.1287/opre.43.2.264).

Enhancing workplace safety and ergonomics with motion capture systems: Present state and a case study

Vujica Herzog, N.^{a,*}, Buchmeister, B.^a, Jaehyun, P.^b, Kaya, Ö.^c

^aUniversity of Maribor, Faculty of Mechanical Engineering, Maribor, Slovenia

^bKonkuk University, Department of Industrial Engineering, Seoul, Republic of Korea

^cUniversity of Uşak, Faculty of Fine Arts, Uşak, Turkey

ABSTRACT

A motion capture system (MoCap) records and tracks the real-time movements of objects or people, generating data that can also be used for ergonomic analysis. Using specialised cameras and sensors, these systems translate movements into digital data that can animate figures or objects in a digital environment. While they are widely used in industries such as film, video games, sports, virtual reality and biomechanics, MoCap systems can also be used in the ergonomic design of workplaces. These systems allow researchers to analyse human movements, identify risks such as repetitive movements, awkward postures or excessive forces and develop solutions to improve safety, comfort and performance. The collected data helps to optimise workplace design, products and processes, reduce musculoskeletal disorders and improve well-being. Recent advancements have made motion capture systems highly sophisticated and capable of recognising subtle movements and expressions with exceptional precision. This paper consists of two parts: a concise literature review and a case study. The literature review explores the application of advanced human motion capture technologies in manufacturing and highlights their potential for ergonomic improvements. The case study illustrates these applications and provides practical insights into how motion capture can transform the design and functionality of workplaces.

ARTICLE INFO

Keywords:
Ergonomics;
Ergonomic risks;
Biomechanical research;
Motion capture;
Xsens suit;
Industry 4.0;
Workplace analysis;
Virtual reality

***Corresponding author:**
natasa.vujica@um.si
(Vujica Herzog, N.)

Article history:
Received 20 August 2024
Revised 27 September 2024
Accepted 29 September 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

The ergonomic design of the workplace, taking into account all the characteristics of the worker, is still a challenge for experienced and even more so for untrained ergonomists. A suitable approach should consider a system of various elements such as space, equipment, tools and devices as well as environmental influences such as noise, lighting, temperature, humidity and air velocity. Since the efficient placement of workplace elements and equipment influences productivity and employee satisfaction, our approach to workplace design is of great importance. Another important aspect that should be considered when designing the workplace is employee movements and their evaluation. Modern approaches that use motion capture systems are a major advance in the evaluation of workers' movements.

Motion capture technology is used today in industry and various other working environments. It is a technology first used in culture (movies, games and animations), but later also in other professions, in various fields of industry, medicine, sports and beyond. Several review articles have been published in recent years, indicating the topicality of the subject [1, 2]. Fig. 1 shows the results of the review by Rybníkář *et al.* [1] among the articles published in 2010-2022 with the keywords: ergonomics, motion capture, Industry 4.0, manufacturing, human factors, occupational health, safety, ergonomic assessment and work-related musculoskeletal disorders. According to Salisu *et al.*, tracking the movements of the human body is currently one of the most expanding areas of research [2].

The ongoing trend of Industry 4.0 (I4.0), characterized by "smart" technologies like sensors, communication systems, simulation, and data-driven modeling, enables timely, accessible, and secure information flow, e.g. [3-8]. This trend is driving companies to gradually automate traditional manufacturing processes and update existing monitoring systems and ergonomic practices [9]. In response, concepts such as Ergonomics 4.0 and intelligent ergonomic processes have been developed to prevent risks and promote physically healthy workplaces [10, 11]. The principles of Industry 4.0 are now being integrated with the emerging Industry 5.0 paradigm to adopt a more human-centered approach to industrial work system design. According to the European Commission's 2022 guidance, Industry 5.0 enhances Industry 4.0 by focusing on research and innovation that support the transition to a sustainable, human-centered, and resilient industry.

Workers remain a vital part of production systems, performing the majority of tasks. They bring diverse skills and knowledge and vary in how they approach tasks, in terms of speed, motivation, and diligence. This diversity presents a significant challenge for companies, especially those with high employee turnover and labor-intensive processes involving substantial physical demands.

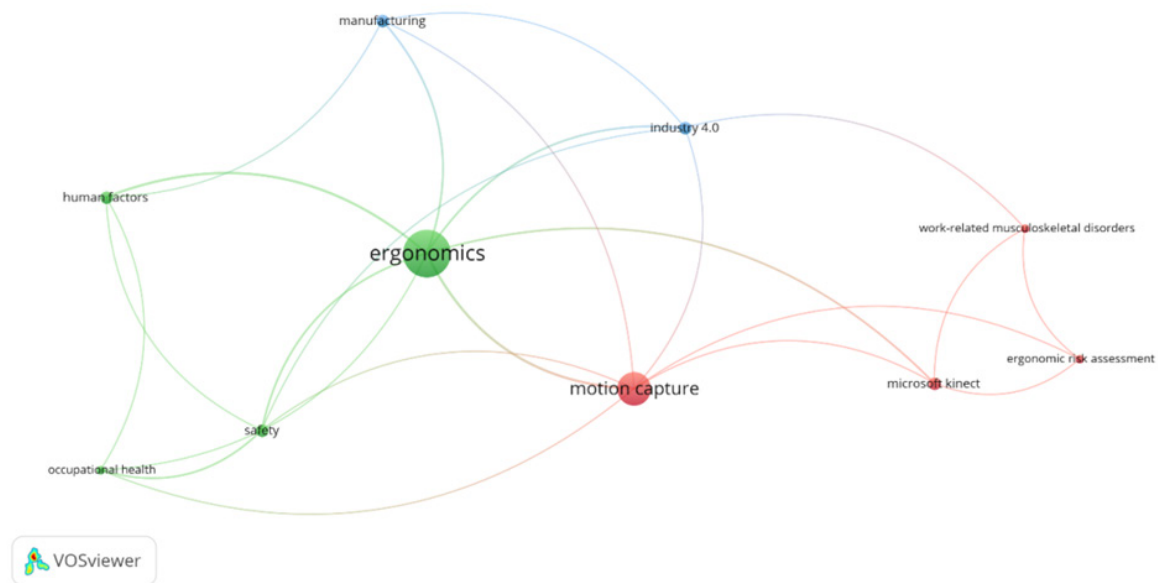


Fig. 1 Occurance of motion capture systems and ergonomics in scientific publications [1]

2. Literature review

2.1 Motion capture systems

A motion capture system is a technology that can be used to track and record the movement of individuals and objects in real time. It can be very useful for the ergonomic assessment of workplaces as it can speed up the assessment process and it is believed that the data can be analyzed more accurately than manually by an ergonomist. In the past, various approaches using different technologies have been developed to transfer the movements of workers to software that enables ergonomic analysis.

The first motion capture systems were developed in the 1960s and 1970s for military and aerospace applications. These systems used mechanical sensors and digitizers to capture the movements of pilots and astronauts. In the 1980s, the first optical motion capture systems with cameras and reflective markers were developed to track the movements of actors and performers. At the same time, pioneers in the field of biomechanics began to study the movement of humans and animals, and this knowledge was later shared. The first motion capture system that resembled modern systems was developed in the 1970s by a group of researchers at the University of Pennsylvania. This early system used reflective markers attached to the body of the person being captured, which were tracked by a series of cameras positioned around the person. The data captured by the cameras was then processed by a computer to create an animated representation of the person's movements.

An example of this group of technologies is the Vicon Motion Capture System, developed in Oxford, UK in 1979 and launched in 1984 by Oxford Medical Systems (later Oxford Dynamics) [12]. The Vicon system uses high-quality cameras and markers to track motion and is primarily used for professional motion capture applications in areas such as sports, animation and robotics. It is known for its high accuracy in motion tracking.

Kinect is a motion detection input device developed by Microsoft for its Xbox gaming console [12]. It was launched in 2010 and quickly gained popularity due to its ability to track human movement without the need for controllers or other input devices. The Kinect sensor uses a combination of cameras and depth sensors to track the movements of a person in front of it. It can detect the position and movement of individual body parts and also recognize facial expressions and voice commands. Originally developed for gaming, the Kinect is now used in a variety of other applications, such as education, healthcare and research. Its ability to track human movement in real time has made it a valuable tool for studying biomechanics, assessing progress in physiotherapy and creating interactive exhibitions. Despite its popularity, Kinect is no longer being actively developed. However, its influence on the motion capture industry and its innovative use of depth sensor technology has helped pave the way for other motion capture systems.

In the 1990s, magnetic motion capture systems were developed. These systems used magnetic fields and sensors to track movements and later, in the early 2000s, the first generation of inertial motion capture systems were developed using accelerometers and gyroscopes.

The Mocup system from Synertial [13] has been commercially available since 2004. It was developed to accurately capture and record the movement of the body of a person wearing a full body suit. It uses wireless sensors (15 on the body and 7 on each hand) attached to various parts of the body, including hands, feet, head and torso. These sensors are small, lightweight and unobtrusive, allowing a natural range of motion without interfering with the performer's movements. The sensors use an inertial measurement unit (IMU) to measure the acceleration, angular velocity and magnetic field of the body part (to improve the estimation of segment orientation) to which they are attached. This data is transmitted wirelessly to a computer where it is processed and used to create a real-time representation of the performer's movements.

Unity is a cross-platform game engine developed by Unity Technologies, first released in June 2005 [14]. Since its launch, the engine has been continuously expanded to support a wide range of platforms, including desktop, mobile, console, and virtual reality. Unity is used to create both three-dimensional (3D) and two-dimensional (2D) games, as well as interactive simulations and various other experiences. Beyond gaming, Unity has been adopted in industries such as film, automotive, architecture, engineering, construction, and even by the United States Armed Forces. Originally launched for Mac OS X, it later extended support to Microsoft Windows and web browsers.

In 2007, another interesting solution was launched, initially called Moven and then renamed Xsens's MVN in 2009 [15]. The company was founded in 2000 by two graduates of the College of Twente. The Xsens motion sensing technology, also known as motion tracking or inertial measurement units (IMUs), uses small, lightweight and highly accurate sensors that are attached to the body or objects to capture and record motion data [14]. The Xsens sensors contain accelerometers, gyroscopes and magnetometers that work together to provide highly accurate and reliable measurements of movement. The data from the sensors is processed by sophisticated algorithms

to calculate the exact position, orientation and movement of the body or object in real time. This allows users to capture and analyze motion data with high precision and accuracy. This enables them to make informed decisions and create realistic and engaging digital content.

Noraxon has introduced an innovative IMU-based motion capture system that provides a versatile solution for measuring and analyzing human motion [16]. Like Xsens, it uses a series of small, lightweight IMUs attached to the body to capture joint angles, orientation and linear acceleration. Noraxon's system is characterized by its portability, which allows it to be used both in laboratory environments and in the real world. The data processed by advanced algorithms is visualized in real time via graphs and a skeletal avatar, providing immediate feedback for on-site adjustments. This makes the device an invaluable tool for researchers, clinicians and sports professionals who need accurate movement data in different environments.

The VR ErgoLog system is another type of motion capture system designed to analyze and improve human movement in virtual reality (VR) environments [17]. It is often used in sports training and rehabilitation. The system uses motion capture sensors attached to the user's body, a VR headset that immerses the user in a virtual environment, and combined with heart rate monitoring. The user can then perform various exercises or movements and the system tracks their movements in real time. The data collected by the system is analyzed by software that provides the user with feedback on their performance. This feedback can include information on posture, balance and alignment, as well as suggestions for improvement. The VR ErgoLog system is designed to be highly customizable and can adapt the level of difficulty and type of exercises to the user's specific needs. It can also be used for remote training and analysis, allowing trainers or therapists to monitor the user's progress remotely.

The HTC Vive is a virtual reality (VR) headset developed by HTC and Valve Corporation [18]. It was launched in 2016 and is known for its high-quality VR experiences. It offers room-based tracking that allows the user to move freely in a physical space while interacting with a virtual environment. The system includes a headset, two handheld controllers and two base states that enable 360-degree motion tracking. Unity and Unreal Engine are popular game engines that support the HTC Vive and allow developers to create VR experiences and integrate motion capture data. The motion capture capabilities of the HTC Vive, provided by the base stations and controllers, are highly accurate and offer low latency. This makes it suitable for various applications beyond gaming, such as training simulations and virtual prototyping.

Table 1 Motion capture systems and ergonomics

Motion capture systems	Used technology	Quality	Calibration process	Costs	Software used to process data	Ergonomics assessment
Vicon Motion capture system	Uses high-quality cameras and markers to track motion	High	Complex setup and calibration process with multiple cameras and markers	Very expensive	✓	✓
Kinect	A hardware device that uses depth-sensing cameras and microphones to track user movements and voice commands.	Medium	Relatively simple calibration and setup process	Expensive	✓	✓
Mocup system by Synertial	Based on inertial measurement units (IMUs)	High	Not particularly complex calibration process	Expensive	✓	✓
Unity	A software platform that provides developers with tools to create and deploy interactive content		Does not have a specific calibration process	Free	✓	

Table 1 (Continuation)

Xsens	Uses small motion sensors (IMU) attached to the body which communicate with a central hub to capture and analyze motion data	High	Simple calibration process	Very expensive	✓	✓
Noraxon	Uses small motion sensors (IMU) attached to the body which communicate with a central hub to capture and analyze motion data	High	Simple calibration process	Very expensive	✓	✓
VR – ErgoLog System	An inertial motion capture system integrated with immersive reality and combined with a heart rate monitoring	High	Not particularly complex calibration process	NA	✓	✓
Leap motion	Combination of optical and infrared sensors to track hand and finger movements	High	Some initial setup and calibration is needed	Cheap	✓	-
HTC Vive	Combination of advanced display technology, positional tracking, and wireless controllers to create VR experience	High	Simple calibration process	Not very expensive	✓	-

2.2 Motion capture systems in different working environments

Motion capture technology has a wide range of potential applications in different working environments, and its versatility makes it a valuable tool for improving efficiency, safety and productivity [19, 20]. The literature review by Rybníkář [1] shows the percentage of use in different sectors as follows:

- Manufacturing, 37 %,
- Logistics 23 %,
- Healthcare 17 %,
- Sports 10 %,
- Construction 7 %,
- Entertainment 7 %.

In healthcare, motion capture systems can be used to analyze patients' movements and gait to identify potential problems with balance, stability or other motor skills [21]. This information can be used to develop tailored rehabilitation programs.

It is also widely used in sports to analyze athletes' movements and improve their performance [22, 23]. It can help prevent injuries by identifying potential weak points or imbalances.

In education and training, interactive simulations can be created for training and educational purposes, e.g. for surgical training or virtual reality training for hazardous environments.

The use of motion capture systems and simulations is also important in manufacturing and on assembly lines [13, 24, 25]. It can be used to track the movements of workers on a factory floor, which can increase efficiency and reduce the risk of accidents. It can also be used to create virtual training simulations for workers to practice complex procedures in a safe and controlled environment.

Recently, it has also been used extensively when working with robots and collaborative robots [26]. Motion capture systems can be used to program and control robots so that they can mimic human movements and interact seamlessly with their environment.

2.3 Ergonomic workplace design

An effective ergonomic workplace design can lead to reduced lead and cycle times, increased productivity, lower production costs, improved return on investment, enhanced product quality and flexibility, fewer human and system errors, reduced idle time during work hours, and lower injury-related expenses. There are several ergonomic methods for evaluating the posture of workers at the workplace: Ovako Working Posture Assessment System (OWAS), Rapid Upper Limb Assessment (RULA), Rapid Entire Body Assessment (REBA), National Institute of Occupational Safety and Health (NIOSH), Occupational Repetitive Actions (OCRA), – Evaluacion del Riesgo Individual (ERIN), Potential Ergonomics Issue List (PEIL), Strain Index (SI) and many others [27-32].

Some methods focus on evaluating strain on specific body parts, while others offer a more comprehensive assessment. Certain approaches are tailored to different types of tasks, such as repetitive versus non-repetitive tasks, and some are designed to evaluate static loads, which involve maintaining the same posture for extended periods. Special attention should also be given to tasks involving manual handling. For example, an integrated approach for determining working times has been developed for assembly lines using the classical Methods Time Measurement (MTM) approach and the EAWS method for ergonomics [33]. In addition to manual assessment tools, which are time-consuming, there are also several computerized tools that reduce assessment time and usually offer several methods for body assessment (e.g. Jack, Process Simulate, Ergomas).

2.4 Future ergonomic challenges

Aging of the population

Demographic data shows an aging world population with an increasing dominance of older workers in the labor market. Rising life expectancy and falling birth rates are leading to slower population growth and a higher proportion of older people. Although the changes associated with aging are normal, they affect the physiological and psychological functioning of individuals and impact on their ability to work and quality of life. According to Cammen et al [34], aging is a multidimensional process of change in the physical, mental and social domains that leads to functional decline.

Increased life expectancy has substantial socio-economic implications for both companies and society. As people continue to work later in life and this demographic grows, addressing workplace conditions and ensuring effective management with a focus on health and safety has become a top priority [35].

Ergonomics is fundamentally concerned with the design of the workplace according to the needs and abilities of workers. In this context, an ergonomic approach considers adequate lighting, avoiding excessive heat, taking adequate breaks and redesigning unnecessarily complex tasks. From an ergonomic perspective, maintaining age-related functionality is a priority issue for the quality of life of older people [36].

In addition, motion capture technology can be used to develop customized prostheses or assistive devices that mimic natural human movement patterns and enable older individuals to maintain their independence and quality of life. Overall, motion capture technology can play a valuable role in promoting healthy aging by helping to identify potential problems early and develop personalized interventions to address them.

Workers with disabilities

Approximately 23.5% of workers in the EU-27 region suffer from some form of chronic illness, while 19% experience long-term illnesses [37]. The highest rates of disability are found among individuals over the age of 55, with no significant difference between men and women [38]. The most commonly reported disabilities are related to back and neck issues, followed by heart conditions, high blood pressure, and problems affecting the legs and feet. Half of the disabilities were acquired outside of work, 17% were congenital, and 18% were the result of work-related injuries or conditions [38]. The severity of a disability also impacts employment, as individuals with severe disabilities are more likely to be unemployed [39].

Sex and gender in ergonomics

To explore this emerging research area, the IEA established the Technical Committee (TC) on Gender and Work in 2006 [40]. Since then, numerous committees have been formed across various countries to examine the connections between sex, gender, health, and work in different occupational sectors. Collectively, these efforts highlight ergonomics as both a theoretical and practical discipline, developing innovative methods to incorporate sex and gender analysis into policy development and evaluation. Gender considerations were presented for various evaluation methods, working conditions and impact/process benefits in ergonomic interventions.

3. Materials and methods

To gain deeper insights into the topic, a case study was conducted using Xsens sensors [15] to capture human movements, which were later analyzed from an ergonomic perspective using the Process Simulate software (Siemens PLM Software). The sensors were directly attached to the body, allowing measurements to be taken in a natural environment and in real time. The ergonomic assessment of the worker's postures and movements was performed using the OWAS and RULA method [27, 28].

3.1 Process Simulate

Process Simulate is one of the many computer programs for the simulation and optimization of manufacturing processes developed by Siemens Digital Industries. It is a comprehensive platform for the development, simulation, validation and thus optimization of production in manufacturing environments. In our research, we used it to create a digital model of the workplace under investigation and to carry out ergonomic validation. The program enables the input of human movements and positions recorded with the Xsens suit. The Task Simulation Builder (TSB) in Process Simulate can also be used to create/edit the simulation of human movements. Minor corrections were made to the movements captured with the Xsens suit via the TSB interface.

3.2 Xsens suit

Founded in 2000 and now part of Movella, Xsens specializes in 3D motion sensing technology, offering wearable sensors and inertial devices that leverage miniature MEMS technology. The company provides innovative and precise solutions for real-time tracking and analysis of human and object movements. Its sensors, which include accelerometers, gyroscopes, and magnetometers, are placed on different body parts to measure acceleration, angular velocity, and spatial orientation. Advanced algorithms process this data to accurately reconstruct 3D movements.

In our research, we used the Xsens MVN Awinda system, which includes 17 wireless motion sensors (Fig. 2). These sensors are securely attached to a special Lycra suit and straps to ensure proper fixation. Fig. 2 shows the body parts and the position of sensors.



Fig. 2 Xsens MVN Awinda system for body movement data gathering and Xsens suit with sensors

The biomechanical model used in this system consists of 23 segments: pelvis, vertebrae (L5*, L3*, T12*, T8), neck*, head, shoulders, upper arms, lower arms, hands, thighs, lower legs, feet, and toes*. Segments marked with * do not have dedicated sensors; their movement is estimated by integrating data from the adjacent segments and the biomechanical model. By combining sensor data and correcting for errors—since the sensors are not rigidly connected to the body segments—a precise and drift-free estimate of the relative position and orientation of each segment is achieved.

To ensure accuracy, sensor calibration is necessary to account for factors that may affect measurement precision. This sensor-to-segment calibration typically involves standing in a known pose (e.g., N or T pose) and estimating sensor orientation by processing the sensor readings (Fig. 3). Gyroscopes detect short-term orientation changes, while accelerometers and magnetometers maintain long-term stability.

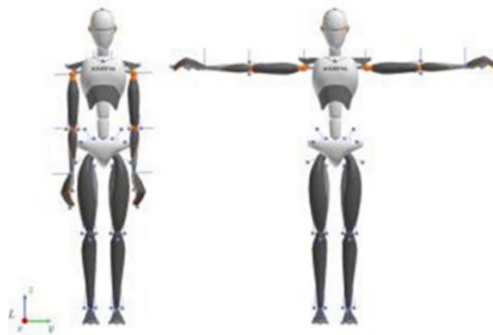


Fig. 3 Xsens MVN avatar in N-pose (left) and T-pose (right)

3.3 OWAS and RULA method

The OWAS method [27] was used for the ergonomic assessment of human movements, evaluating 28 postures. These include the back (four positions), upper limbs (four), hands (three), lower limbs (nine), head and neck (five), and the load or force applied (three). Each posture is classified into one of the following risk categories:

- Green: No changes needed,
- Yellow: Changes needed in the near future,
- Orange: Immediate changes required,
- Red: Requires intensive observation.

The Process simulate software enables various ergonomic analyses; we therefore also carried out an analysis using the RULA method to compare the results. The RULA method (Rapid Upper Limb Assessment) is an ergonomic tool for assessing the risk of musculoskeletal injuries in the workplace, particularly in relation to the upper limbs [28]. The aim is to observe the positions of certain body segments individually in order to determine the degree of deviation from neutral posture. The greater the deviation, the higher the score assigned to the respective body part.

3.4 Research environment and measuring procedures

The research was conducted in the laboratory of the Faculty of Mechanical Engineering at the University of Maribor, which is fully equipped to transfer human movements into a virtual environment. A specialized Xsens training suit, fitted with 17 sensors, was used to capture human movements, positions, and postures. During the process, a receiver collects the data from the suit and records it on a PC. The data is processed using the Xsens MVN software to generate a simulation. Simultaneously, the human movements were recorded in real time via video, which was later used to recreate the movements using Task Simulation Builder (TSB).

For the case study demonstration with the Xsens suit, we designed a simple polygon created with Process Simulate software (Fig. 4) that contained the following tasks:

- Walk to the table,
- Get a small box with both hands,
- Bring the box to the locker,
- Put the box on the locker,
- Walk to the chair,
- Get the chair,
- Place the chair to the table.

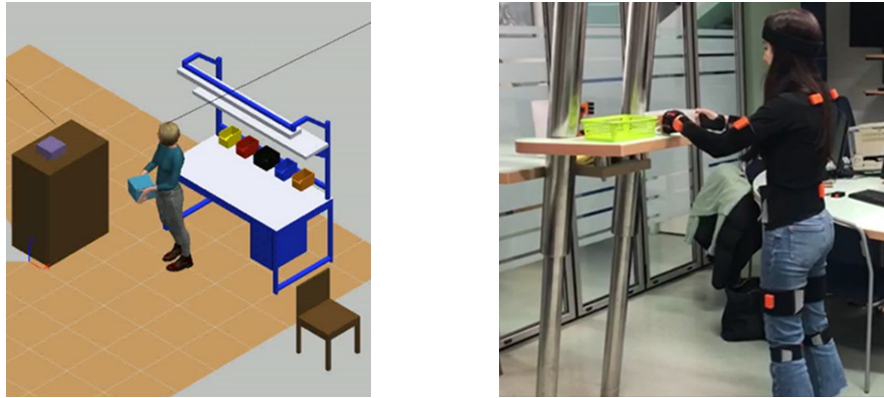


Fig. 4 Workplace created with Process Simulate software and Xsens suit with 17 sensors

5. Results and discussion

The movements of workers simulated using the Process Simulate software package were assessed with the OWAS ergonomic method first. We examined these movements with two human models representing different ages: 25 and 60 years old. Fig. 5 shows some of the positions of the human model that were performed during the simulation for the selected postures of a 25-year-old woman on the left and a 60-year-old woman on the right.

Figs. 6 and 7 display the cumulative OWAS analysis results for both age groups, with slight differences noted in the movements highlighted in orange and red.

Overall, the analysis results are similar for most simulated movements. However, a significant difference was observed in one specific posture—when turning to walk to another point (see Figs. 5, 6, and 7). For the 25-year-old model, the actions was marked in orange, indicating a moderately high risk and recommending prompt changes. Conversely, for the 60-year-old model, the same actions was marked in red, signifying a very high risk and recommending immediate changes. For further details on the research, please refer to the chapter published by DAAAM International Scientific Book [41].

During the research, we also compared ergonomic analyses performed manually with TSB interface with those based on data from the Xsens suit. The postures selected and designed by the TSB programme in Process Simulate are more ergonomic as they are selected by the programme itself, resulting in lower risk scores. An example of this is the posture in which the woman lifts the chair and places it at the desk (Fig. 8). In the simulation carried out with the suit, this posture is regarded as unergonomic and categorised as risk level 4. However, in the simulation performed by the programme, this posture represents a lower risk and is classified as level 2. This is because the simulation performed by the programme ensures a correct grip posture of the chair, in contrast to the other simulation.

Table 2 shows some of the results of the RULA method for the most important postures. As we can see, there is one posture that refers to Jill bending over to push the chair. According to RULA, this posture should be analysed and changed immediately.

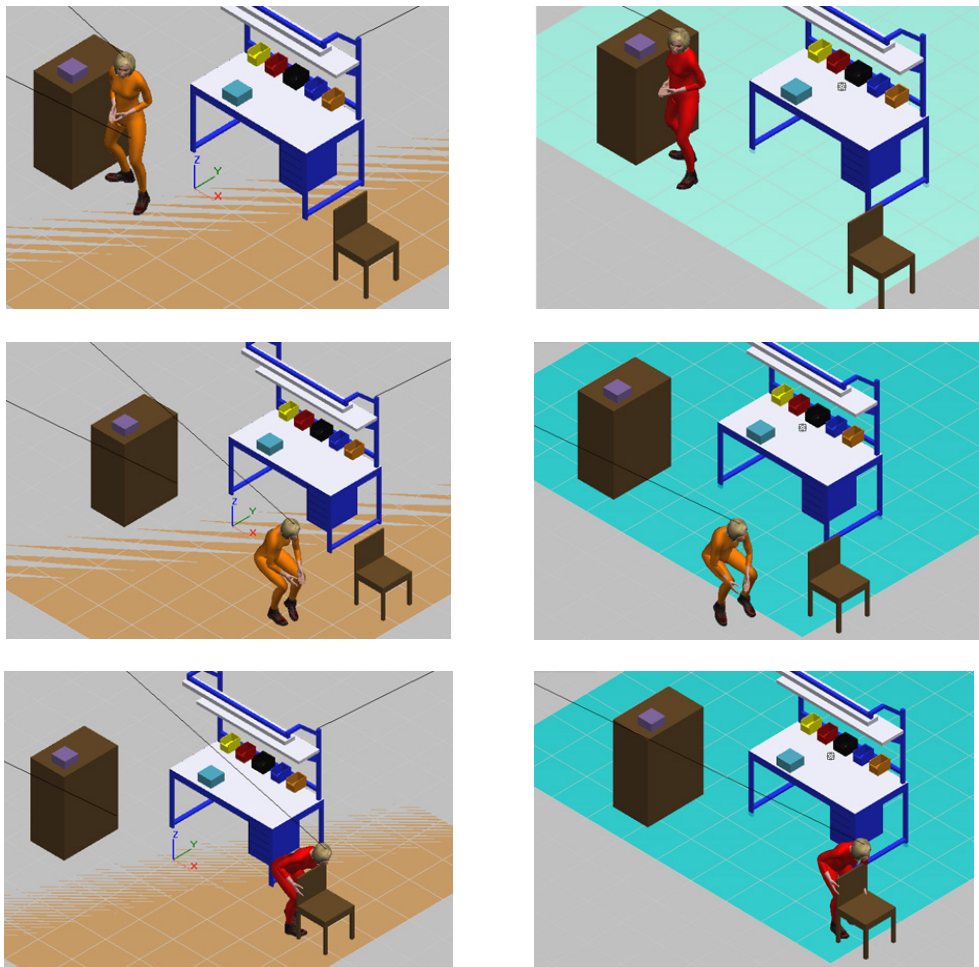


Fig. 5 Working environment with 25 and 60 year-old person

Time (Sec)	Operation	Object Weight (kg)	Action Category	Code Posture Combination					
				Back	Arms	Legs	Load	-	Head
0	Frame 4	0	1	1	1	2	1	-	1
0.2		0	2	1	1	4	1	-	1
8.3	Frame 88	0	3	2	1	4	1	-	1
8.7	Frame 94	0	2	1	1	4	1	-	1
16.25	Frame 175	0	3	2	1	4	1	-	1
16.9	Frame 181	0	2	1	1	4	1	-	1
19.2	Frame 205	0	3	2	1	4	1	-	1
22.45	Frame 241	0	4	4	1	4	1	-	1
23.1	Frame 247	0	3	2	1	4	1	-	1
23.8	Frame 253	0	2	1	1	4	1	-	1
24.2	Frame 259	0	3	2	1	4	1	-	1
24.3		0	2	1	1	4	1	-	1
24.5	Frame 262	0	3	2	1	4	1	-	1
25.3	Frame 271	0	4	4	1	4	1	-	1
26.5	Frame 283	0	3	2	1	4	1	-	1
27.2	Frame 289	0	2	1	1	4	1	-	1
27.5	Frame 292	0	3	2	1	4	1	-	1
27.6	Frame 295	0	2	1	1	4	1	-	1
28	Frame 298	0	3	2	1	4	1	-	1
29	Frame 310	0	2	1	1	4	1	-	1
29.1		0	3	2	1	4	1	-	1
29.3	Frame 313	0	2	1	1	4	1	-	1
30.9	Frame 331	0	3	2	1	4	1	-	1
32.2	Frame 343	0	3	2	1	4	1	-	4
33.5	Frame 358	0	3	2	1	4	1	-	1
36.8	Frame 394	0	2	1	1	4	1	-	1

Fig. 6 OWAS results: 25-year-old model

Time (Sec)	Operation	Object Weight (kg)	Action Category	Code Posture Combination					
				Back	Arms	Legs	Load	-	Head
22.2	Frame 406	0	2	1	1	4	1	-	1
22.25		0	3	3	1	4	1	-	1
22.5	Frame 412	0	2	1	1	4	1	-	1
22.59	Frame 415	0	3	2	1	4	1	-	1
23.38	Frame 430	0	3	2	1	4	1	-	4
23.5		0	4	4	1	4	1	-	4
23.6	Frame 433	0	3	2	1	4	1	-	4
23.85	Frame 436	0	4	4	1	4	1	-	4
24.15	Frame 442	0	3	2	1	4	1	-	4
24.34	Frame 448	0	3	2	1	4	1	-	1
24.93	Frame 460	0	2	1	1	4	1	-	1
29.05	Frame 535	0	3	2	1	4	1	-	1
29.55	Frame 544	0	3	2	1	4	1	-	4
31.65	Frame 583	0	3	2	1	4	1	-	1
34.55	Frame 637	0	2	1	1	4	1	-	1

Fig. 7 OWAS results: 60-year-old model

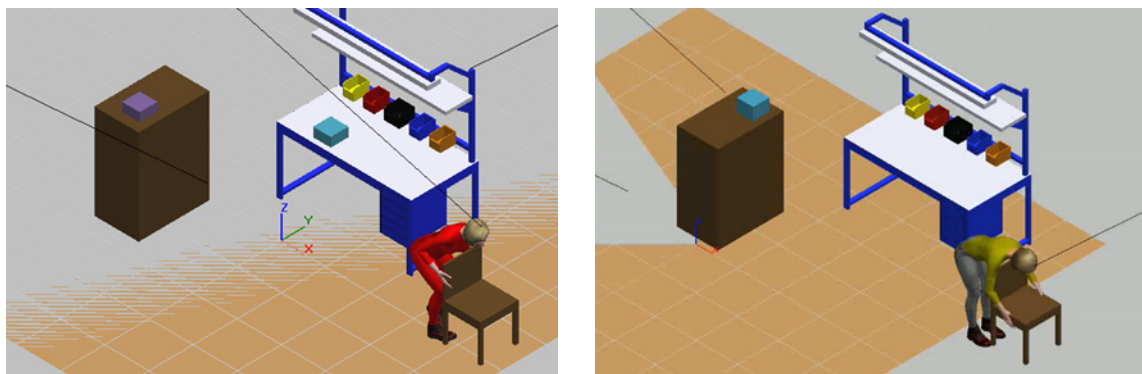


Fig. 8 Taking chair from Recording on the left and Taking chair from Simulation on the right

Table 2 Compilation of results of the RULA method for the most relevant postures

Position	25 year-old person	60 year-old person
Put_chair_Jill_1_Release	2	2
Go_to_target_Jill_Arrise_From_Bend	4	4
Put_box_Jill_Release	3	3
Put_part1_Jill_Bend_and_Reach	7	7
Go_to_target_Jill_Arrise_From_Bend	6	7
Go_to_target_Jill_Walk	4	4
Put_chair_Jill_Bend_and_Reach	7	7
Put_chair_Jill_Reach	6	6

6. Conclusion

The study of workplace ergonomics is essential for promoting worker health, well-being, and productivity. Ergonomic principles aim to tailor the work environment to individual needs, optimizing the interaction between workers, tasks, and their surroundings. This approach helps prevent injuries, reduce fatigue and stress, and improve performance and job satisfaction, benefiting both employees and companies. Recent advancements in motion capture systems have contributed significantly to ergonomic workplace design. These systems are highly valuable for virtual reality applications, detailed biomechanical analyses, and exploring ways to integrate motion capture into human-computer interactions. Additionally, they facilitate access for individuals with disabilities, promoting more inclusive participation across various environments.

The main contribution of the paper is an overview of motion capture technology and the possible integration with ergonomic assessment methods. It presents various possibilities and discusses future challenges in ergonomics.

A case study is included to demonstrate new technologies, such as the Xsens suit and Process Simulate software, in ergonomic workplace design, with a focus on older workers. This study has three limitations. First, while the findings offer valuable insights through a case study, they should be viewed as preliminary. The primary value of this research lies in its role as a pilot study. Since it is based on a single case, the results cannot be generalized. Future studies need to include a larger sample to validate the results in a broader population. Second, there may be discrepancies between the OWAS measurements obtained using the Xsens motion capture system and the data obtained by direct observation. Although we have made considerable efforts to minimize these differences by performing a frame-by-frame analysis, discrepancies may still occur. Finally, although OWAS and RULA were used for ergonomic assessment in this study, there are numerous other methods to assess physical strain. It is important to investigate whether these alternative methods can also be used reliably with motion capture data. Furthermore, given the real-time analysis capability offered by systems such as Xsens, this study points to the potential need to develop new methods of ergonomic assessment that go beyond traditional approaches and are tailored to take full advantage of modern motion capture technology.

In conclusion, motion capture technology offers significant advantages for ergonomic workplace design, providing efficient and accurate solutions. The conducted research has demonstrated great potential; however, future studies will need to be expanded, include a larger number of repetitions to verify the reliability of the equipment.

Acknowledgement

The authors thank the reviewers and the editor for their comments. This work was supported by the Slovenian Research Agency under Grant P2-0190. We also acknowledge the use of research equipment system for development and testing cognitive production approaches in Industry 4.0, procured within the project RIUM, which was co-financed by the Republic of Slovenia, the Ministry of Education, Science and Sport and the European Union from the European Regional Development Fund.

References

- [1] Rybníkář, F., Kačerová, I., Hořejší, P., Šimon M. (2023). Ergonomics evaluation using motion capture technology - Literature review, *Applied Sciences*, Vol. 13, No. 1, Article No. 162, doi: [10.3390/app13010162](https://doi.org/10.3390/app13010162).
- [2] Salisu, S., Ruhaiyem, N.I.R., Eisa, T.A.E., Nasser, M., Saeed, F., Younis, H.A. (2023). Motion capture technologies for ergonomics: A systematic literature review, *Diagnostics*, Vol. 13, Article No. 2593, doi: [10.3390/diagnostics13152593](https://doi.org/10.3390/diagnostics13152593).
- [3] Rosin, F., Forget, P., Lamouri, S., Pellerin, R. (2021). Impact of Industry 4.0 on decision-making in an operational context, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 500-514, doi: [10.14743/apem2021.4.416](https://doi.org/10.14743/apem2021.4.416).
- [4] Bun, P., Grajewski, D., Gorski, F. (2021). Using augmented reality devices for remote support in manufacturing: A case study and analysis, *Advances in Production Engineering & Management*, Vol. 16, No. 4, 418-430, doi: [10.14743/apem2021.4.410](https://doi.org/10.14743/apem2021.4.410).
- [5] Wawak, S., Sutoova, A., Vykydal, D., Halfarova, P. (2023). Factors affecting Quality 4.0 implementation in Czech, Slovak and Polish organizations: Preliminary research, *Advances in Production Engineering & Management*, Vol. 18, No. 3, 345-356, doi: [10.14743/apem2023.3.477](https://doi.org/10.14743/apem2023.3.477).
- [6] Stefanovska, E., Pepelnjak, T. (2022). Development of a flexible tooling system for sheet metal bending, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 311-325, doi: [10.14743/apem2022.3.438](https://doi.org/10.14743/apem2022.3.438).
- [7] Lasota, A.M. (2020). A new approach to ergonomic physical risk evaluation in multi-purpose workplaces, *Tehnički Vjesnik – Technical Gazette*, Vol. 27, No. 2, 467-474, doi: [10.17559/TV-20180312131319](https://doi.org/10.17559/TV-20180312131319).
- [8] Lyu, J., Peng, Q., Chen, M. (2022). Influence of body mass index on comfort and parametric optimization design of seat, *Tehnički Vjesnik – Technical Gazette*, Vol. 29, No. 4, 1262-1269, doi: [10.17559/TV-20210306023605](https://doi.org/10.17559/TV-20210306023605).
- [9] Diego-Mas, J.-A., Poveda-Bautista, R., Garzon-Leal, D.-C. (2015). Influences on the use of observational methods by practitioners when identifying risk factors in physical work, *Ergonomics*, Vol. 58, No. 10, 1660-1670, doi: [10.1080/00140139.2015.1023851](https://doi.org/10.1080/00140139.2015.1023851).
- [10] Kadir, B.A., Broberg, O., da Conceição, C.S. (2019). Current research and future perspectives on human factors and ergonomics in Industry 4.0, *Computers & Industrial Engineering*, Vol. 137, Article No. 106004, doi: [10.1016/j.cie.2019.106004](https://doi.org/10.1016/j.cie.2019.106004).
- [11] Lanzotti, A., Vanacore, A., Tarallo, A., Nathan-Roberts, D., Coccorese, D., Minopoli, V., Carbone, F., d'Angelo, R., Grasso, C., Di Gironimo, G., Papa, S. (2019). Interactive tools for safety 4.0: Virtual ergonomics and serious games in real working contexts, *Ergonomics*, Vol. 63, No. 3, 324-333, doi: [10.1080/00140139.2019.1683603](https://doi.org/10.1080/00140139.2019.1683603).

- [12] Dutta, T. (2012). Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace, *Applied Ergonomics*, Vol. 43, No. 4, 645-649, doi: [10.1016/j.apergo.2011.09.011](https://doi.org/10.1016/j.apergo.2011.09.011).
- [13] Simonetto, M., Arena, S., Peron, M. (2022). A methodological framework to integrate motion capture system and virtual reality for assembly system 4.0 workplace design, *Safety Science*, Vol. 146, Article No. 105561, doi: [10.1016/j.ssci.2021.105561](https://doi.org/10.1016/j.ssci.2021.105561).
- [14] Peruzzini, M., Grandi, F., Cavallaro, S., Pellicciari, M. (2021). Using virtual manufacturing to design human-centric factories: An industrial case, *The International Journal of Advanced Manufacturing Technology*, Vol. 115, 873-887, doi: [10.1007/s00170-020-06229-2](https://doi.org/10.1007/s00170-020-06229-2).
- [15] Roetenberg, D., Luinge, H., Slycke, P. (2013). *Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors*, Xsens technologies – Version April 3, from https://www.researchgate.net/publication/239920367_Xsens_MVN_Full_6DOF_human_motion_tracking_using_miniature_inertial_sensors, accessed August 13, 2024.
- [16] Cottam, D.S., Campbell, A.C., Davey, M.P.C., Kent, P., Elliott, B.C., Alderson, J.A. (2022). Measurement of uni-planar and sport specific trunk motion using magneto-inertial measurement units: The concurrent validity of Noraxon and Xsens systems relative to a retro-reflective system, *Gait & Posture*, Vol. 92, 129-134, doi: [10.1016/j.gaitpost.2021.11.012](https://doi.org/10.1016/j.gaitpost.2021.11.012).
- [17] Battini, D., Calzavara, M., Persona, A., Sgarbossa, F., Visentin, V., Zennaro, I. (2018). Integrating mocap system and immersive reality for efficient human-centred workstation design, *IFAC – PapersOnLine*, Vol. 51, No. 11, 188-193, doi: [10.1016/j.ifacol.2018.08.256](https://doi.org/10.1016/j.ifacol.2018.08.256).
- [18] Khan, S., Park, J.-W., Ham, S. (2023). Noncontact structural displacement measurements using a low-cost motion tracking system, *IEEE Sensors Journal*, Vol. 23, No. 11, 11695-11703, doi: [10.1109/JSEN.2023.3267450](https://doi.org/10.1109/JSEN.2023.3267450).
- [19] Kačerová, I., Kubr, J., Hořejší, P., Kleinová, J. (2022). Ergonomic design of a workplace using virtual reality and a motion capture suit, *Applied Science*, Vol. 12, No. 4, Article No. 2150, doi: [10.3390/app12042150](https://doi.org/10.3390/app12042150).
- [20] Dolgui, A., Sgarbossa, F., Simonetto, M. (2021). Design and management of assembly systems 4.0: Systematic literature review and research agenda, *International Journal of Production Research*, Vol. 60, No. 1, 184-210, doi: [10.1080/00207543.2021.1990433](https://doi.org/10.1080/00207543.2021.1990433).
- [21] Rudigkeit, N., Gebhard, M. (2020). AMiCUS 2.0 – System presentation and demonstration of adaptability to personal needs by the example of an individual with progressed multiple sclerosis, *Sensors*, Vol. 20, No. 4, Article No. 1194, doi: [10.3390/s20041194](https://doi.org/10.3390/s20041194).
- [22] Van der Kruk, E., Reijne, M.M. (2018). Accuracy of human motion capture systems for sport applications; State-of-the-art review, *European Journal of Sport Science*, Vol. 18, No. 6, 806-819, doi: [10.1080/17461391.2018.1463397](https://doi.org/10.1080/17461391.2018.1463397).
- [23] Ghattas, J., Jarvis, D.N. (2021). Validity of inertial measurement units for tracking human motion: A systematic review, *Sports Biomechanics*, 1-14, doi: [10.1080/14763141.2021.1990383](https://doi.org/10.1080/14763141.2021.1990383).
- [24] Bortolini, M., Faccio, M., Gamberi, M., Pilati, F. (2020). Motion Analysis System (MAS) for production and ergonomics assessment in the manufacturing processes, *Computers & Industrial Engineering*, Vol. 139, Article No. 105485, doi: [10.1016/j.cie.2018.10.046](https://doi.org/10.1016/j.cie.2018.10.046).
- [25] Grznar, P., Gregor, M., Gola, A., Nielsen, I., Mozol, S., Seliga, V. (2022). Quick workplace analysis using simulation, *International Journal of Simulation Modelling*, Vol. 2, No. 3, 465-476, doi: [10.2507/IJSIMM21-3-612](https://doi.org/10.2507/IJSIMM21-3-612).
- [26] Ojstersek, R., Javernik, A., Buchmeister, B. (2022). Importance of sustainable collaborative workplaces – Simulation modelling approach, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 627-638, doi: [10.2507/IJSIMM21-4-623](https://doi.org/10.2507/IJSIMM21-4-623).
- [27] Karhu, O., Härkönen, R., Sorvali, P., Vepsäläinen, P. (1981). Observing working postures in industry: Examples of OWAS application, *Applied Ergonomics*, Vol. 12, No. 1, 13-17, doi: [10.1016/0003-6870\(81\)90088-0](https://doi.org/10.1016/0003-6870(81)90088-0).
- [28] McAtamney, L., Corlett, E.N. (1993). RULA: A survey method for the investigation of work-related upper limb disorders, *Applied Ergonomics*, Vol. 24, No. 2, 91-99, doi: [10.1016/0003-6870\(93\)90080-s](https://doi.org/10.1016/0003-6870(93)90080-s).
- [29] Hignett, S., McAtamney, L. (2000). Rapid entire body assessment (REBA), *Applied Ergonomics*, Vol. 31, No. 2, 201-205, doi: [10.1016/s0003-6870\(99\)00039-3](https://doi.org/10.1016/s0003-6870(99)00039-3).
- [30] Occhipinti, E., Colombini, D. (2002). Updating reference values and predictive models of the OCRA method in the risk assessment of work related musculoskeletal disorders of the upper limbs, *Ergonomics*, Vol. 50, No. 11, 1727-1739, doi: [10.1080/00140130701674331](https://doi.org/10.1080/00140130701674331).
- [31] Kee, D., Karwowski, W. (2007). A comparison of three observational techniques for assessing postural loads in industry, *International Journal of Occupational Safety and Ergonomics*, Vol. 13, No. 1, 3-14, doi: [10.1080/10803548.2007.11076704](https://doi.org/10.1080/10803548.2007.11076704).
- [32] Vujica Herzog, N., Buchmeister, B. (2015). The review of ergonomics analysis for body postures assessment, In: Katalinić, B. (ed.), *DAAAM International scientific book 2015*, DAAAM International, Vienna, Austria, 153-164, DAAAM International scientific book, doi: [10.2507/daaam.scibook.2015.14](https://doi.org/10.2507/daaam.scibook.2015.14).
- [33] Schaub, K., Caragnano, G., Britzke, B., Bruder, R. (2013). The European assembly worksheet, *Theoretical Issues in Ergonomics Science*, Vol. 14, No. 6, 616-639, doi: [10.1080/1463922X.2012.678283](https://doi.org/10.1080/1463922X.2012.678283).
- [34] van der Cammen, T.J.M., Wang, G., Albayrak, A. (2019). Where ergonomics meets geriatrics: The connection between comprehensive geriatric assessment and design for ageing, *European Geriatric Medicine*, Vol. 10, 333-335, doi: [10.1007/s41999-019-00171-7](https://doi.org/10.1007/s41999-019-00171-7).
- [35] Araújo-Vila, N., Toubes, D.R., Fraiz-Brea, J.A. (2022). The age factor in the analysis of occupational risks in the wood industry, *Healthcare*, Vol. 10, No. 7, Article No. 1355, doi: [10.3390/healthcare10071355](https://doi.org/10.3390/healthcare10071355).
- [36] Pamuk, D., Faezi, S.A., Başibüyük, G.Ö. (2022). Ergonomics and aging: A bibliometric analysis, *Work*, Vol. 72, No. 3, 853-864, doi: [10.3233/WOR-210112](https://doi.org/10.3233/WOR-210112).
- [37] Vujica-Herzog, N., Harih, G. (2019). Decision support system for designing and assigning ergonomic workplaces to workers with disabilities, *Ergonomics*, Vol. 63, No. 2, 1-13, 225-236, doi: [10.1080/00140139.2019.1686658](https://doi.org/10.1080/00140139.2019.1686658).

- [38] Dupré, D., Karjalainen, A. (2003). *Eurostat, statistics in focus: Employment of disabled people in Europe in 2002, Eurostat, Theme 3: Population and Social Conditions*, Eurostat, Luxembourg, 1-8, from <https://aei.pitt.edu/85492/1/2003.26.pdf>, accessed August 13, 2024.
- [39] Greve, B. (2009). *The labour market situation of disabled people in european countries and implementation of employment policies: A summary of evidence from country reports and research studies*, Academic Network of European Disability Experts (ANED), from https://www.academia.edu/1902383/The_labour_market_situation_of_disabled_people_in_European_countries_and_implementation_of_employment_policies_A_summary_of_evidence_from_country_reports_, accessed August 13, 2024.
- [40] Laberge, M., Caroly, S., Riel, J. Messing, K. (2020). Considering sex and gender in ergonomics: Exploring the hows and whys, *Applied Ergonomics*, Vol. 85, Article No. 103039, doi: [10.1016/j.apergo.2019.103039](https://doi.org/10.1016/j.apergo.2019.103039).
- [41] Vujica Herzog, N., Kaya, Ö; Buchmeister, B. (2023). Ergonomics and ageing – Case study, In: Katalinić, B. (ed.), *DAAAM International scientific book 2023*, DAAAM International, Vienna, Austria, 17-32, doi: [10.2507/daaam.scibook.2023.02](https://doi.org/10.2507/daaam.scibook.2023.02).

Machine learning for enhancing manufacturing quality control in ultrasonic nondestructive testing: A Wavelet Neural Network and Genetic Algorithm approach

Song, W.T.^a, Huo, L.^{b,*}

^aSchool of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang, P.R. China

^bSchool of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, P.R. China

ABSTRACT

With the rapid development of the global manufacturing industry, an efficient and accurate quality control system has become key to enhancing competitiveness. Ultrasonic Nondestructive Testing (NDT), as an efficient means of quality inspection, plays a crucial role in improving manufacturing quality through the precision of its data analysis. This study aims to explore the application of ultrasonic NDT data in manufacturing quality control by integrating machine learning technologies, with a specific focus on the Wavelet Neural Network optimized by Genetic Algorithms (GA-WNN). This study achieved significant prediction and evaluation results by applying a GA-WNN to quality control in manufacturing. Compared to traditional Wavelet Neural Network (WNN) models, the GA-WNN more effectively identifies and predicts potential quality issues, especially in noisy data and complex production environments, demonstrating higher accuracy and stability. When predicting possible defect types in the manufacturing process, the GA-WNN showed a notable improvement in accuracy over other models. Additionally, in quality stability evaluation, GA-WNN was able to capture production fluctuations more accurately, providing more valuable results for decision-making. The methodologies and discoveries of this study offer new perspectives and tools for quality control in manufacturing and the analysis of ultrasonic NDT data, presenting broad application prospects.

ARTICLE INFO

Keywords:

Ultrasonic nondestructive testing (NDT);
Machine learning;
Genetic Algorithm (GA);
Wavelet Neural Network (WNN);
Quality prediction;
Quality stability assessment;
Quality control optimization

*Corresponding author:

hl@stdu.edu.cn
(Huo, L.)

Article history:

Received 9 April 2024
Revised 22 September 2024
Accepted 29 September 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

In the context of rapid global manufacturing industry development, ensuring manufacturing quality is not only the cornerstone of enterprise survival but also the key to enhancing competitiveness [1, 2]. With technological advancements, ultrasonic NDT, known for its efficiency and non-destructiveness, has been widely applied in quality control within the manufacturing sector [3-5]. However, traditional analysis of ultrasonic testing data largely relies on experienced operators, which is not only inefficient but also struggles to detect complex and minute defects [6, 7]. Therefore, efficiently and accurately analysing ultrasonic NDT data becomes a crucial issue for improving the level of manufacturing quality control.

The rapid development of machine learning technologies in recent years has offered new solutions to the aforementioned problems. By utilizing machine learning algorithms for in-depth

analysis of ultrasonic NDT data, it's possible to significantly improve the accuracy and efficiency of defect detection [8-11]. More importantly, this approach allows for automated detection under complex conditions, greatly enhancing the quality control capabilities of the manufacturing process. However, existing research primarily focuses on the improvement and optimization of algorithms, with less attention given to their application and effect evaluation in the actual manufacturing process, limiting the widespread application of machine learning in manufacturing quality control [12-14].

Although existing research methods have improved detection efficiency and accuracy to some extent, there are still some flaws and deficiencies [15-18]. Firstly, the algorithm models used in many studies are complex and require substantial computational resources, which is not conducive to application in resource-limited production environments. Secondly, some research lacks consideration of the variability in manufacturing processes, making it difficult to adapt to changing production conditions [19, 20]. Moreover, existing methods often overlook the interpretability of models, which is disadvantageous for problem diagnosis and decision support in practical applications [21].

This paper aims to introduce machine learning technology, specifically the GA-WNN, to build a comprehensive system suitable for manufacturing that includes quality prediction, stability assessment, and control optimization. Firstly, the study focuses on a manufacturing quality prediction model based on GA-WNN, aimed at providing a basis for early intervention through accurate prediction of potential quality issues in the manufacturing process. Secondly, a quality stability assessment system for the manufacturing process is established, evaluating production stability by analysing data fluctuations during the production process to provide references for quality control. Finally, an optimization model for manufacturing process quality control is researched, aiming to improve product quality and production efficiency by optimizing production parameters through algorithms. This study not only improves the level of quality control in manufacturing but also provides a new methodology for the efficient utilization of ultrasonic NDT data, having significant theoretical and practical value.

2. Construction of a manufacturing process quality prediction model based on GA-WNN

This paper proposes a model based on the GA-WNN (Fig. 1), designed to overcome the limitations of traditional quality prediction methods when faced with complex manufacturing stages, achieving accurate prediction of quality characteristic values throughout the entire manufacturing process. Based on its excellent generalization ability and noise resistance, the WNN's initial parameters are optimized using GA, not only speeding up the model's convergence but also significantly enhancing its predictive accuracy. Unlike models specifically for other assembly processes, the GA-WNN model considered in this study encompasses the entire chain from raw material processing to final product assembly. The model's input vector $A = \{a_1, a_2, \dots, a_l\}$ includes the actual values of quality characteristics from raw materials through each manufacturing step and their deviation from process standards, while the model output $B = \{b_1\}$ aims to predict the quality characteristic values of a specific manufacturing stage or the final product. This approach enables the model to provide real-time, accurate predictions of the quality condition of downstream processes or final product quality based on actual quality data and its fluctuations from upstream processes, offering strong support for early detection of potential quality issues and timely adjustment of production strategies.

WNN is a composite model that combines wavelet transform and neural networks. Its primary feature is the ability to decompose input signals into different frequency components through wavelet transforms, capturing both time and frequency characteristics of the signal. Compared to traditional Artificial Neural Networks (ANN), WNN not only handles nonlinear relationships but also effectively processes time-varying and noisy signals. For example, while ANN often requires signal preprocessing or manual feature extraction, WNN can automatically extract key features through adaptive multi-scale analysis, overcoming the limitations of traditional meth-

ods. In ultrasonic non-destructive testing data analysis, WNN can accurately capture subtle fluctuations in the signal and predict potential quality issues, while ANN may fail in such applications due to its insufficient ability to process time-frequency features. When constructing the manufacturing process quality prediction model based on GA-WNN, the design of the model's structure and the selection of parameters decisively impact its performance. To adapt to the complexity of quality control in manufacturing, this paper adopts a three-layer structure of the WNN, designed to fully capture and analyse the complex quality information contained in ultrasonic NDT data. In determining the optimal number of nodes in the hidden layer, a crucial parameter in the network structure, this study employs a trial-and-error method due to the absence of a unified standard. By continuously adjusting the number of nodes and evaluating model performance, the aim is to achieve optimal quality characteristic prediction effects. Assuming the number of input layer nodes is represented by l , and the number of output layer nodes is represented by j , with a constant ranging from 1 to 10 represented by s , the formula is:

$$v = \sqrt{l + j} + s \quad (1)$$

Moreover, selecting an activation function is another crucial parameter setting. This paper chooses the Morlet wavelet function as the activation function $\psi(a)$ for the hidden layer, due to its good anti-interference capability and computational stability when processing ultrasonic NDT data. This is crucial for improving the model's adaptability and predictive accuracy in complex manufacturing environments. The expression for $\psi(a)$ is:

$$\psi(a) = \cos(1.75a)e^{\frac{-a^2}{2}} \quad (2)$$

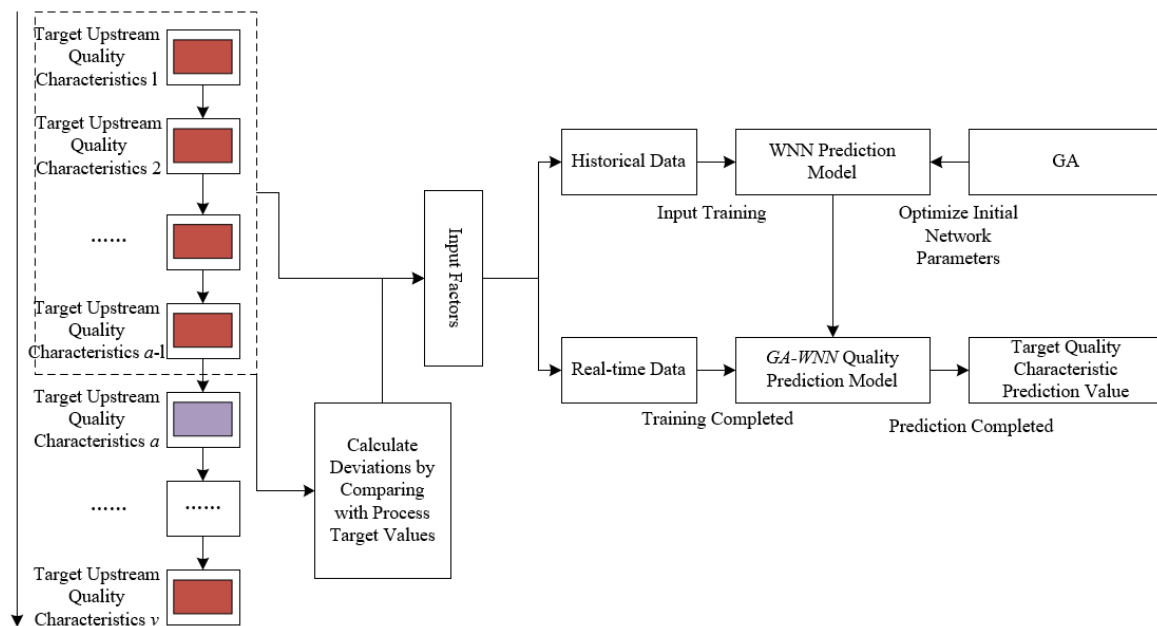


Fig. 1 Schematic diagram of the manufacturing process quality prediction model

This study introduces initial optimization parameters in the WNN, including connection weights μ_{uk} and μ_{kj} , translation factors y_k , and scale factors x_k . To adapt to this high-dimensional data structure and simplify the coding process, this paper employs a real-number encoding strategy, using floating-point numbers to represent the network's weights and factors. Additionally, considering the two-dimensional nature of the weight matrix and the one-dimensional nature of chromosomes, this study unfolds the two-dimensional weight matrix into a one-dimensional array and concatenates it with translation and scale factors, forming a complete chromosome representation for the implementation and calculation of GA. The total length of individual encoding is represented by $u \times k + k + k + k \times j$. Fig. 2 shows the schematic diagram of the encoding strategy.

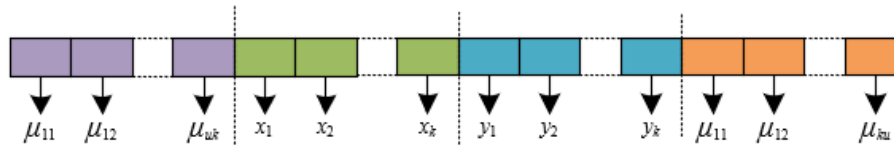


Fig. 2 Encoding strategy schematic diagram

In the implementation of the GA, the reciprocal of the error between the predicted values and actual values serves as the fitness function to measure the performance of the prediction model, where a higher fitness value indicates higher model prediction accuracy and superior solution quality. Assuming the prediction output error of the WNN is represented by r , the expression for the fitness function d is as follows:

$$d = \frac{1}{r} \quad (3)$$

To ensure that individuals with high fitness are prioritized for selection, this study employs the widely used roulette wheel selection method to determine the probability of an individual being selected, ensuring the inheritance and optimization process's effectiveness of high-quality solutions. Assuming the fitness value of the u -th individual is represented by D_u , and the coefficient is represented by j , with the total number of individuals in the population represented by V , the calculation formulas are as follows:

$$d_u = \frac{j}{D_u} \quad (4)$$

$$o_u = \frac{d_u}{\sum_{u=1}^V d_u} \quad (5)$$

For two chromosomes X and Y of length m , assuming a random number between $[0,1]$ is represented by e , the crossover operation process is as follows:

$$x_j = (1 - e)x_j + ey_j \quad (6)$$

$$y_j = ex_j + (1 - e)y_j \quad (7)$$

When a mutation operator mutates the j -th gene of chromosome X with a certain probability, assuming the upper and lower bounds of gene value are represented by x_{MAX} and x_{MIN} , the current iteration number by H , the maximum number of network iterations by H_{MAX} , and a random number between $[0,1]$ by e_1 and e_2 , the mutation operation expression is as follows:

$$x_j = \begin{cases} x_j + (x_{MAX} - x_j) \cdot e_1 \cdot \left(1 - \frac{H}{H_{MAX}}\right)^2 & e_2 > 0.5 \\ x_j + (x_{MIN} - x_j) \cdot e_1 \cdot \left(1 - \frac{H}{H_{MAX}}\right)^2 & e_2 \leq 0.5 \end{cases} \quad (8)$$

3. Construction of a quality stability assessment system for the manufacturing process

To construct a reasonable quality stability assessment system for the manufacturing process, it's first necessary to determine multiple influencing factors. This study is not limited to a specific purpose but extends the view to multiple factors that may affect quality stability throughout the manufacturing process. Therefore, the factor set N is defined as a collection of quality characteristics encompassing all stages from raw material selection, processing technology, assembly process, to final inspection, i.e., $N = \{n_1, \dots, n_u, \dots, n_l\}$, where l represents the number of quality characteristics being assessed in the manufacturing process. For the evaluation set I , this study uses the degree of deviation of quality characteristic values from process standards as a key indicator to measure the quality stability of the manufacturing process. The evaluation set I is defined as three levels: $I = \{i_1, i_2, i_3\} = \{\text{significant deviation, general deviation, minor deviation}\}$,

reflecting the degree of deviation of various quality characteristics from their specified standards in the manufacturing process.

In the construction of the quality stability assessment system for the manufacturing process, selecting appropriate membership functions to quantify the degree of deviation of various quality characteristics is crucial. This step aims to transform ultrasonic NDT data and other quality characteristic data into dimensionless values that can reflect quality stability. Unlike assessments in other assembly processes that may only focus on specific dimensions or characteristics, the assessment of quality stability in the manufacturing process needs to consider comprehensive quality characteristics from raw material inspection, intermediate process control, to final product inspection, which may include dimensions, strength, surface roughness, and other dimensions.

To quantify the reasonableness of these quality characteristic deviations, this study uses normalization to convert data into values within the $[0,1]$ interval. The closer the normalized value is to 1, the closer the quality characteristic of the process is to the set standard value, indicating a smaller actual deviation and higher quality stability. Conversely, the closer the value is to 0, the larger the actual deviation and the lower the quality stability. Since quality characteristics in manufacturing often include various types, and each type of characteristic may have different levels of acceptability and standards, traditional normalization methods may not meet the needs of all situations. Therefore, this paper proposes a flexible normalization method that considers the upper and lower deviations from the standard values of quality characteristics in the manufacturing process, more comprehensively and accurately reflecting the quality stability of the manufacturing process. Assuming the normalized quality characteristic value is represented by a^* , the actual quality characteristic value by a , the standard value of the quality characteristic by i , the maximum upper deviation of the quality characteristic value by i_G , and the maximum lower deviation of the quality characteristic value by i_M . The following is the normalization equation:

$$a^* = 1 - \frac{|a - i|}{i_G - i_M} \quad (9)$$

Assuming the membership degree corresponding to parameter a is denoted by $e(a)$, the expected value of the distribution is denoted by i , and the width of the curve is denoted by δ , then the membership function expression is as follows:

$$e(a) = \exp\left[\frac{-(a - i)^2}{2\delta^2}\right] \quad (10)$$

Further assuming the distribution expected values for the three intervals are denoted by i_x , ω_y , and i_z , with i_x and i_z being the maximum and minimum values of that parameter attribute, and ω_y being the given expected value under the *general deviation* condition, the specific function expressions are as follows:

$$e(a) = \begin{cases} 1 & a \leq i_x \\ \exp\left[\frac{-(a - i_z)^2}{2\delta_x^2}\right] & i_x < a \leq i_z, \delta_x = \frac{i_y - i_x}{3} \end{cases} \quad (11)$$

$$e(a) = \begin{cases} \exp\left[\frac{-(i_y - a)^2}{2\delta_{y1}^2}\right] & i_x < a \leq i_z, \delta_{y1} = \frac{i_y - i_x}{3} \\ \exp\left[\frac{-(a - i_y)^2}{2\delta_{y2}^2}\right] & i_y < a \leq i_z, \delta_{y2} = \frac{i_z - i_y}{3} \end{cases} \quad (12)$$

$$e(a) = \begin{cases} \exp\left[\frac{-(a - i_z)^2}{2\delta_z^2}\right] & i_x < a \leq i_z, \delta_z = \frac{i_z - i_y}{3} \\ 1 & i_z \leq a \end{cases} \quad (13)$$

Assuming the number of quality characteristics involved in the assessment is denoted by l , a quality stability fuzzy matrix for the manufacturing process can then be established as:

$$E = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ \vdots & e_{u2} & \vdots \\ e_{l1} & e_{l2} & e_{l3} \end{bmatrix} \quad (14)$$

4. Optimization modelling for manufacturing process quality control

In the entire production process of manufacturing, the complexity of quality control far exceeds that of a single manufacturing stage, as it encompasses multiple stages, from raw material processing and component manufacturing to final product assembly. Each stage can introduce quality deviations, which may accumulate along the production chain, affecting the final product's quality stability. Especially when using high-precision techniques such as ultrasonic NDT for quality assessment, even minor deviations can have a significant impact on the final product's quality. Therefore, traditional quality control thresholds based on manufacturing standards often fail to manage the entire manufacturing process's quality stability comprehensively and effectively. Addressing this issue, this paper aims to optimize quality control thresholds by real-time prediction of quality characteristics at various stages of the manufacturing process and comprehensive assessment of the entire manufacturing process's quality stability. Fig. 3 shows the schematic diagram of the optimization model for manufacturing process quality control.

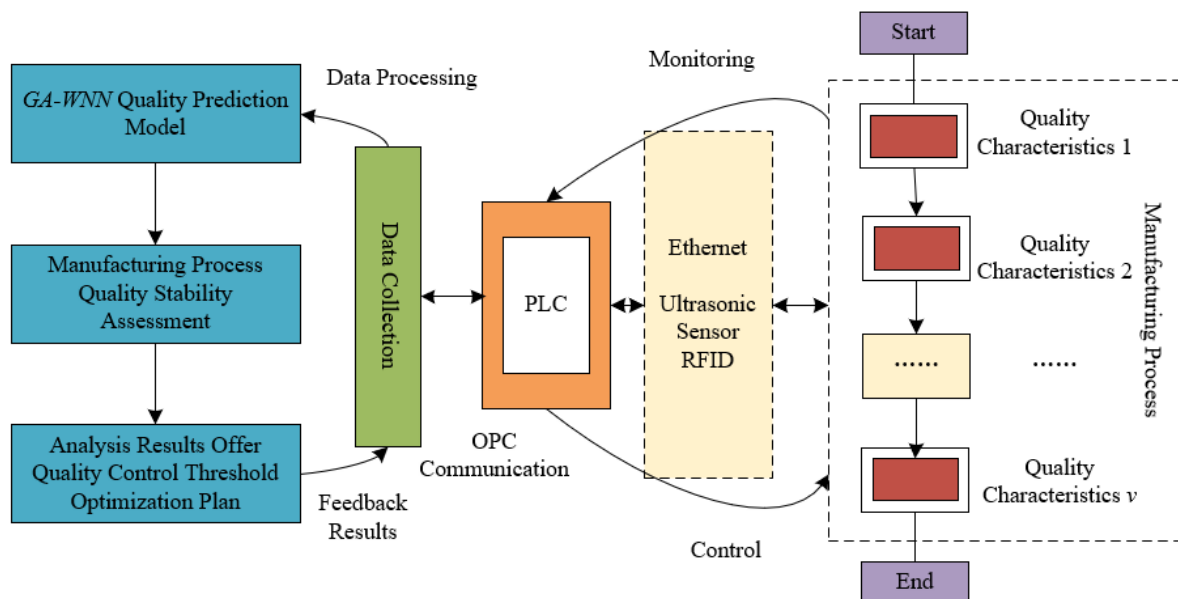


Fig. 3 Schematic diagram of the optimization model for manufacturing process quality control

In the manufacturing process, due to the diversity and complexity of the involved processes, the impact of each process's quality characteristics on the overall product quality stability varies. Especially when high-precision quality assessment technologies such as ultrasonic NDT are introduced, quality deviations in each process and their impact on the final product quality can be identified more accurately. In this context, controlling all quality characteristics of every process to the same extent would lead to a tremendous waste of resources and may not achieve effective real-time quality control. Therefore, this paper proposes a quality control optimization principle, which prioritizes controlling those quality characteristics of processes that have a significant impact on overall quality stability.

Assuming the initial process requirements for quality characteristics are represented by $H \pm e$, the process target value by H , and the permissible range within the process by $[H-e, H+e]$. Meeting the initial process requirements, the control threshold is divided into j levels, with the quali-

ty characteristic control threshold at the j -th level represented by H_j , and the threshold change degree at the j -th level by S_j , where $S_j \in [0,1]$, then there is:

$$H_j = [H - H \times S_j, H + H \times S_j] \quad (15)$$

In the optimization modelling for quality control in the manufacturing process, optimizing the quality characteristic thresholds is a key step aimed at precisely controlling the quality characteristics of each production stage to improve the overall manufacturing process's quality stability. Below are the five key steps for optimizing quality characteristic thresholds in the manufacturing process:

- **Data Collection and Model Training:** First, collect ultrasonic NDT data and the quality characteristics data of corresponding processes from the real-time database of the Manufacturing Execution System (MES). This data reflects the historical quality situation in the manufacturing process and is used to train the constructed GA-WNN model. Through this training, the model learns the relationship between quality characteristics in the manufacturing process and the final product quality, providing a basis for subsequent quality predictions.
- **Quality Characteristics Prediction and Assessment:** Next, use the real-time collected manufacturing process data as input to predict the quality characteristics of target control points using the trained GA-WNN model. By comprehensively assessing these prediction results using the established manufacturing process quality stability assessment method, a quantified score about the entire manufacturing process's stability can be obtained.
- **Optimization Adjustment of Quality Control Thresholds:** If the assessment results do not meet the preset quality stability standards, it indicates that optimization adjustments of quality control thresholds are needed. Priority is given to adjusting the quality characteristics with greater weight and currently showing larger deviations in their predicted values. By progressively narrowing the range of control thresholds, i.e., adjusting from looser to stricter standards, quality deviations are reduced, and quality stability is improved.
- **Cyclical Optimization Process:** If the quality stability still does not meet the standards even after adjusting to the strictest control threshold range, the same adjustment process is applied to the quality characteristics with the next highest weight. This cyclical process ensures that by sequentially optimizing the control thresholds of key quality characteristics, the entire manufacturing process's quality stability can gradually improve to meet the assessment standards.
- **Real-time Feedback and Process Adjustment:** Finally, the optimized quality control threshold results are promptly fed back to the Programmable Logic Controller (PLC) on the manufacturing floor to guide actual manufacturing operations. This step ensures that the optimizations of quality control thresholds can be quickly implemented, impacting the manufacturing process, and effectively improving the manufacturing process's quality stability and product quality.

To validate the practical application of the GA-WNN model, we collaborated with an automotive parts manufacturing company and implemented the GA-WNN model on their production line. The company uses ultrasonic non-destructive testing technology to monitor product quality during production, and we applied the GA-WNN model to analyse the data from this process. The experimental results demonstrated that GA-WNN significantly improved the accuracy of quality predictions and was able to identify potential quality issues in advance. In terms of quality stability assessment, GA-WNN accurately captured production fluctuations, enabling the management team to adjust production parameters in time, thus preventing large-scale production losses. Specifically, within two months of implementing GA-WNN, production efficiency increased by 10 %, and the product rework rate decreased by 15 %. This case study validates the effectiveness of GA-WNN in real manufacturing environments.

5. Experimental results, analysis, and discussion

Comparing the manufacturing process quality prediction results shown in Fig. 4, differences in predictive performance between the GA-WNN and the traditional WNN are observable. It's evident from the experimental data that GA-WNN provided closer to actual value predictions for most samples. For instance, on sample 3, the GA-WNN predicted a value of 24.3, while the actual value was 24.6, resulting in a small difference of only 0.3. In comparison, the WNN predicted 24, showing a discrepancy of 0.6 from the actual value. Similarly, for sample 6, the GA-WNN's prediction of 23.4 was very close to the actual value of 23.8, whereas the WNN predicted 23, resulting in a larger discrepancy. Overall, GA-WNN consistently showed lower prediction errors across all samples than WNN, indicating higher accuracy in manufacturing quality predictions. These experimental results highlight the effectiveness of the GA-WNN model in predicting quality in the manufacturing process. Through GA optimization, GA-WNN can more accurately adjust network parameters, better capturing the patterns of quality variation in the manufacturing process, and providing more precise prediction results. This improvement in accuracy is crucial for early identification of potential quality issues in the manufacturing process, offering strong data support for timely intervention measures.

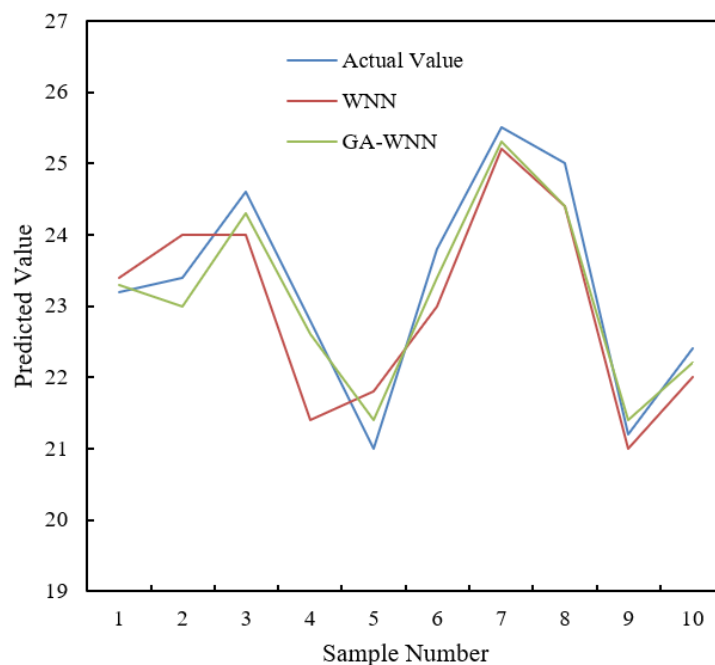


Fig. 4 Comparison of quality prediction results for manufacturing process test set

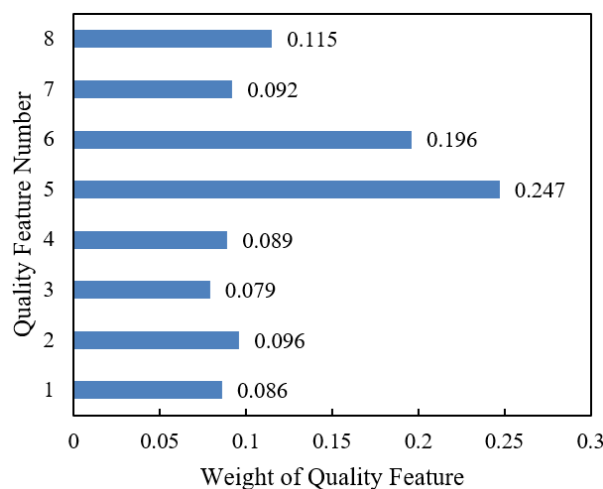


Fig. 5 Weight distribution for quality characteristics in manufacturing process

Fig. 5 shows the weight distribution data for various quality characteristics in the manufacturing process, indicating that among all considered quality characteristics, the fifth and sixth quality characteristics occupy significantly higher weights of 24.7 % and 19.6 % respectively, which are noticeably higher than other quality characteristics. This suggests that these two quality characteristics have a greater impact on the final product's quality within the manufacturing process; therefore, their stability and accuracy should be prioritized during quality control and optimization. Conversely, the third quality characteristic holds the lowest weight at 7.9 %, indicating its relatively minor impact on the final product's quality. Such weight distribution helps identify key quality characteristics that significantly affect the final product's quality in the manufacturing process, providing a basis for the rational allocation of resources and optimization adjustments. Through precise allocation of quality characteristic weights, the proposed optimization model for quality control in the manufacturing process can guide quality management activities more effectively during the production process. This optimization strategy, based on weight distribution, ensures that limited resources are focused on key stages that most significantly affect the final product's quality, thus improving the efficiency and effectiveness of quality control throughout the manufacturing process. Furthermore, real-time feedback of optimization measures to the production line's PLC further ensures real-time monitoring and adjustment of the manufacturing process, enhancing the production process's flexibility and responsiveness.

Table 1 presents the final evaluation vectors for 100 sets of quality characteristic data in the manufacturing process, reflecting the membership degrees of three different quality characteristics (Characteristic 1, Characteristic 2, Characteristic 3) across various sample groups. It can be observed that Characteristic 3 has higher membership degrees in most groups, such as 0.784 in Group 1, 0.784 in Group 3, and 0.725 in Group 100, showing that Characteristic 3 deviates from its target values more frequently compared to Characteristics 1 and 2 throughout the manufacturing process. In contrast, Characteristic 1 generally has lower membership degrees, indicating higher stability and lower likelihood of deviation during the manufacturing process. Characteristic 2's membership degrees are at a medium level, indicating its probability of deviating from target values falls between Characteristics 1 and 3. Analysis of this data demonstrates the effectiveness of the proposed optimization model for quality control in the manufacturing process. Firstly, the analysis of the final evaluation vectors for quality characteristic data accurately identifies quality characteristics prone to deviation during the manufacturing process, providing a clear basis for optimizing adjustment of quality control thresholds.

Table 1 Final evaluation vector for quality characteristic data in manufacturing process

Group	Characteristic 1	Characteristic 2	Characteristic 3
1	0.021	0.134	0.784
2	0	0.265	0.562
3	0.002	0.153	0.784
...
48	0.095	0.314	0.445
49	0.115	0.258	0.451
50	0.223	0.178	0.432
51	0	0.101	0.835
52	0	0.162	0.784
...
99	0.124	0.332	0.421
100	0	0.124	0.725

Table 2 Stability scores for the manufacturing process

Group	Characteristic 1	Characteristic 2	Characteristic 3
1	83.25	51	58.94
2	71.24	52	86.25
3	82.36	53	82.31
...
48	65.25	99	63.21
49	62.39	100	77.89

Table 2 provides stability scores for various groups of data concerning three different quality characteristics (Characteristic 1, Characteristic 2, Characteristic 3) in the manufacturing process. The data shows significant differences in stability scores for different quality characteristics across groups, reflecting the quality fluctuations during the production process. For instance, Characteristic 1 scored 83.25 in Group 1 and 62.39 in Group 49, illustrating variability in its stability across different production batches. Characteristic 2 scored relatively high in some groups, such as 100 in Group 49, indicating higher stability in these specific production batches. In contrast, Characteristic 3's scores vary greatly across groups, such as 86.25 in Group 2 and 63.21 in Group 48, showing notable variations in stability during different production processes. These findings suggest that the developed quality stability assessment system for the manufacturing process effectively evaluates and reflects the stability of various quality characteristics during production. Through this assessment system, potential quality fluctuation issues in the production process can be accurately identified, providing a scientific basis for optimizing adjustment of quality control thresholds and formulating corresponding quality improvement measures.

6. Conclusion

This paper has successfully developed a comprehensive system for quality prediction, stability assessment, and control optimization in the manufacturing industry by integrating machine learning technology, particularly the GA-WNN. The research focuses on three core aspects: firstly, developing a manufacturing quality prediction model based on GA-WNN, achieving precise prediction of potential quality issues in the manufacturing process; secondly, constructing a quality stability assessment system for the manufacturing process, evaluating production stability through the analysis of production data fluctuations; and lastly, studying an optimization model for manufacturing process quality control, enhancing product quality and production efficiency through algorithmic optimization of production parameters.

The experimental results demonstrate the comprehensiveness and depth of this research. Analysis of the training set error convergence graph for different manufacturing process quality prediction models, test set quality prediction results, weight distribution of quality characteristics, final evaluation vectors for quality characteristics, stability scores for the manufacturing process, predicted results and comprehensive scores for actual manufacturing process quality characteristic data, and levels of quality control thresholds not only proves the effectiveness of the GA-WNN model in quality prediction and control within the manufacturing industry but also shows the feasibility of improving manufacturing process stability and product quality consistency through optimized control thresholds and real-time feedback adjustment measures.

Although the GA-WNN model demonstrates excellent performance in quality prediction and stability assessment, there are still some limitations. Firstly, the performance of the GA-WNN model is highly dependent on the quality of the input data. In cases where there is significant noise or missing data, the model may experience a decrease in accuracy. Secondly, the training process of the GA-WNN model requires a considerable amount of time, especially when the dataset is large. The computational complexity of the genetic algorithm can be high, which may impact the real-time performance in practical applications. Additionally, GA-WNN is highly sensitive to high-frequency noise in ultrasonic signals, which means that more data preprocessing work may be needed when dealing with complex noise environments. Future research could focus on the following aspects for improvement: exploring more efficient algorithms to reduce training time, developing robust optimization methods for noisy data, and further enhancing the model's adaptability to ensure stable operation even when data quality is poor.

Acknowledgments

This research project is supported by the Natural Science Foundation of Hebei Province (Grant No.: E2019210309), and the technology development project of China Energy Investment Group Co., Ltd (Grant No.: 20230336, 20240010).

References

- [1] Yang, S.Y., Zhang, M.F. (2023). Blockchain-driven optimization in intelligent manufacturing, *International Journal of Simulation Modelling*, Vol. 22, No. 2, 350-361, [doi: 10.2507/IJSIMM22-2-C010](#).
- [2] Li, Z.P. (2022). Management decisions in multi-variety small-batch product manufacturing process, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 537-547, [doi: 10.2507/IJSIMM21-3-C015](#).
- [3] Del Prete, G., Leone, F., Dentico, V., Gallo, N., Maio, L., Memmolo, V., Vespini, V., Coppola, S., Stella, E., Ferraro, P. (2023). Laser ultrasonic non-destructive testing for aerospace applications, In: *Proceedings of 2023 IEEE 10th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, Milan, Italy, 525-528, [doi: 10.1109/MetroAeroSpace57412.2023.10189937](#).
- [4] Rahman, S.Z. Vatambeti, R. (2023). Enhanced fault diagnosis in motor bearings: Leveraging optimized wavelet transform and non-local attention, *International Journal of Knowledge and Innovation Studies*, Vol. 1, No. 2, 127-138, [doi: 10.56578/ijkis010205](#).
- [5] Shaloo, M., Schnall, M., Klein, T., Huber, N., Reitingner, B. (2022). A review of non-destructive testing (NDT) techniques for defect detection: Application to fusion welding and future wire arc additive manufacturing processes, *Materials*, Vol. 15, No. 10, Article No. 3697, [doi: 10.3390/ma15103697](#).
- [6] Quaegebeur, N., Masson, P., Berry, A., Ardin, C., D'Anglade, P.-M. (2018). Ultrasonic non-destructive testing of cardboard tubes using air-coupled transducers, *NDT & E International*, Vol. 93, 18-23, [doi: 10.1016/j.ndteint.2017.09.011](#).
- [7] Helsel, M.A., Popovics, J.S., Stynoski, P.B., Kreiger, E. (2021). Non-destructive testing to characterize interlayer bonds of idealized concrete additive manufacturing products, *NDT & E International*, Vol. 121, Article No. 102443, [doi: 10.1016/j.ndteint.2021.102443](#).
- [8] Situmorang, N., Alisjahbana, S.W., Riyanto, H., Najid. (2024). Effect of data division on classification performance model prediction of specified compressive strength core concrete using Ultrasonic Pulse Velocity in tandem with machine learning, *Mathematical Modelling of Engineering Problems*, Vol. 11, No. 2, 430-438, [doi: 10.18280/mmep.110214](#).
- [9] Zhang, L., Cong, Y., Meng, F., Wang, Z., Zhang, P., Gao, S. (2021). Energy evolution analysis and failure criteria for rock under different stress paths, *Acta Geotechnica*, Vol. 16, No. 2, 569-580, [doi: 10.1007/s11440-020-01028-1](#).
- [10] Carboni, M., Zamorano, M. (2024). On monitoring fretting fatigue damage in solid railway axles by acoustic emission with unsupervised machine learning and comparison to non-destructive testing techniques, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 238, No. 2, 256-267, [doi: 10.1177/09544097231193186](#).
- [11] Wu, J.-D., Huang, Y.-H. (2023). Enhanced identification of internal casting defects in vehicle wheels using YOLO object detection and X-ray inspection, *Traitement du Signal*, Vol. 40, No. 5, 1909-1920, [doi: 10.18280/ts.400511](#).
- [12] Günaydin, O., Akbaş, E., Özbeyaz, A., Güçlüer, K. (2023). Machine learning based evaluation of concrete strength from saturated to dry by non-destructive methods, *Journal of Building Engineering*, Vol. 76, Article No. 107174, [doi: 10.1016/j.jobe.2023.107174](#).
- [13] Nagawkar, J., Leifsson, L. (2021). Efficient global sensitivity analysis of model-based ultrasonic nondestructive testing systems using machine learning and sobol'indices, *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, Vol. 4, No. 4, Article No. 041008, [doi: 10.1115/1.4051100](#).
- [14] Silva, M.I., Malitckii, E., Santos, T.G., Vilaça, P. (2023). Review of conventional and advanced non-destructive testing techniques for detection and characterization of small-scale defects, *Progress in Materials Science*, Vol. 138, Article No. 101155, [doi: 10.1016/j.pmatsci.2023.101155](#).
- [15] Peng, L., Ji, P., Qian, Y., Yang, Y., Tian, H., Han, J. (2023). Quality control of small batch product in manufacturing process, In: Liu, C. (ed.), *Proceedings of the 23rd Pacific Basin Nuclear Conference*, Vol. 2, Springer Proceedings in Physics, Vol. 284, Springer, Singapore. 973-983, [doi: 10.1007/978-981-19-8780-9_92](#).
- [16] Zhang, L., Chao, W., Liu, Z., Cong, Y., Wang, Z. (2022). Crack propagation characteristics during progressive failure of circular tunnels and the early warning thereof based on multi-sensor data fusion, *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, Vol. 8, Article No. 172, [doi: 10.1007/s40948-022-00482-3](#).
- [17] Erdinç, P., Buduneli, Z., Erton, C., Gerşil, Ç., Paldrak, M., Staiou, E. (2023). Applications of statistical process control, quality improvement tools and techniques, and a simulation model in a garment manufacturing company, In: Durakbasa, N.M., Gençylmaz, M.G. (eds.), *Industrial engineering in the Industry 4.0 era. ISPR 2023. Lecture notes in mechanical engineering*, Springer, Cham, Switzerland, 409-422, [doi: 10.1007/978-3-031-53991-6_31](#).
- [18] Yu, J.-Y., Wang, L.-N. (2022). Quality intelligent control method of railway vehicle wheel axle manufacturing process based on part quality index, *International Journal of Manufacturing Technology and Management*, Vol. 36, No. 5-6, 326-338, [doi: 10.1504/IJMTM.2022.128726](#).
- [19] Gu, J., Zhao, L., Yue, X., Arshad, N.I., Mohamad, U.H. (2023). Multistage quality control in manufacturing process using blockchain with machine learning technique, *Information Processing & Management*, Vol. 60, No. 4, Article No. 103341, [doi: 10.1016/j.ipm.2023.103341](#).
- [20] Zhao, L.-P., Li, B.-H., Yao, Y.-Y. (2023). A novel predict-prevention quality control method of multi-stage manufacturing process towards zero defect manufacturing, *Advances in Manufacturing*, Vol. 11, No. 2, 280-294, [doi: 10.1007/s40436-022-00427-9](#).
- [21] Tiryaki, S., Aydn, A. (2022). Multivariate hotelling T2 control chart for monitoring some quality characteristics in medium density fiberboard manufacturing process, *Drvna Industrija*, Vol. 73, No. 1, 35-46, [doi: 10.5552/drvinid.2022.2046](#).

Optimizing obstacle avoidance path planning for intelligent mobile robots in multi-obstacle environments

Zhang, H.Q.^{a,*}

^aSchool of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, P.R. China

ABSTRACT

To enhance the obstacle avoidance capabilities of intelligent mobile robots in environments with multiple obstacles, a method for optimal path planning has been developed and analysed. Multi-obstacle scene maps were generated using the safety boundary condition method and envelope method, with obstacle distribution abstractly represented by an obstacle rate. The robot's pose and orientation were determined through the application of Adaptive Monte Carlo localization. Speed sampling was conducted to identify the optimal speed combination, establishing a set of candidate trajectories. A trajectory evaluation function was constructed, and the trajectory with the highest score was selected. The improved A* algorithm was integrated into the evaluation function, incorporating the obstacle rate to refine path planning. To eliminate redundant sections, a key point extraction algorithm was employed to adjust and smooth the path, ensuring optimal obstacle avoidance. The proposed approach demonstrated superior performance, achieving a shorter planned path, a rapid obstacle avoidance reaction time of 32 milliseconds, a 100 % success rate in obstacle avoidance, and a path smoothness of 97 %. These results indicate that the method is highly effective for obstacle avoidance, with significant improvements in both safety and efficiency.

ARTICLE INFO

Keywords:

Intelligent mobile robots;
Obstacle avoidance path planning;
Adaptive Monte Carlo;
Envelope method

*Corresponding author:

zhq@jsut.edu.cn
(Zhang, H.Q.)

Article history:

Received 5 September 2024

Revised 20 October 2024

Accepted 27 October 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

With the continuous progress and innovation of science and technology, mobile robot technology has become one of the hotspots in today's research field [1-8], which has a wide range of applications, including but not limited to industrial manufacturing, logistics distribution, medical care, rescue exploration, etc. [9, 10]. In these complex and dynamic environments, mobile robots need to be able to quickly and accurately plan the optimal or feasible path from the starting point to the target point while avoiding various obstacles. As the core issue of mobile robot research, path planning has far-reaching significance and wide application value. It is not only about how the robot reaches the designated destination from a starting point but also an important embodiment of the robot's autonomous navigation, environmental perception, and decision-making ability. In the process of solving this problem, it is necessary to fully consider the complex factors, such as the motion characteristics of the robot, the determination of the starting and target positions, and the possible obstacles in the surrounding environment, so as to plan a safe, effective, and optimized path. With continuous progress in the field of artificial intelligence, robots have become increasingly capable of environmental perception and decision-making [3]. However, how to achieve optimal obstacle avoidance path planning in complex and changeable environments is still a major challenge in the field of robotics. Carrying out relevant

research not only helps to promote the progress of related technologies but also provides more possibilities for the practical application of robot technology.

Therefore, relevant research scholars have carried out research on this. For example, Li *et al.* [10] proposed an improved nature-inspired swarm intelligence (NISI) meta-heuristic algorithm called the Slime Mold Optimization Algorithm (SMOA), which was inspired by the oscillation characteristics of slime moulds when they met prey. The ability to read workspace maps and plan collision-free paths with minimal processing time and cost under constraints. However, the SMOA algorithm may be optimized for specific types of environments or obstacle layouts at the beginning of design. In practical applications, robots may face a variety of complex and dynamic environments, including different types of obstacles, irregular shapes and sizes, and dynamic obstacles. If the SMOA algorithm cannot adapt to these changes well, its performance may be affected, which makes it difficult for the robot to effectively find the optimal obstacle avoidance path. Chen *et al.* [11] proposed a three-neighbourhood search A* algorithm combined with an artificial potential field to optimize the path planning of mobile robots. The algorithm integrates and improves part of the artificial potential field and the A* algorithm to solve the irregular obstacles in the forward direction. In the process of obstacle avoidance, the current pose vector of the mobile robot is constructed, which reduces the search scope to less than three neighbours and avoids repeated searches. However, in the practical application of this method, if there are narrow channels or dense areas of obstacles in the environment, the three-neighbourhood search strategy may make the robot fall into local optimization, and the effect is not good. Mohanraj *et al.* [12] developed a hybrid algorithm to find the approximate optimal solution of global path planning and enhanced the algorithm by considering the improved region and real-time constraints to avoid collision with obstacles and find the best path. However, this algorithm is computationally burdensome, especially in large-scale or complex environments, which leads to long computing times and cannot achieve the desired real-time performance. Chen *et al.* [13] proposed an artificial potential field method based on global path guidance to solve the problem of target inaccessibility and local minima of the traditional artificial potential field method in complex environments with dynamic obstacles. Firstly, to solve the problem of target inaccessibility, global path attraction is added to APF. Secondly, an obstacle detection optimization method is proposed, which selects the optimal virtual target point by setting the evaluation function to improve the local minimum problem. Finally, based on the optimization method of obstacle detection, the processes of gravity and repulsion are improved so that the path can pass through the narrow channel smoothly and keep collision-free. Because different environments and tasks may require different evaluation functions and parameter settings, if there are dynamic obstacles in the environment or the shape and size of obstacles are complex and changeable, the performance of obstacle detection and optimization methods may be affected, resulting in the robot being unable to accurately identify and avoid obstacles.

In order to further improve the effect of path planning of intelligent mobile robots, the optimal obstacle avoidance path planning of intelligent mobile robots in multi-obstacle scenes is proposed. The innovative technological roadmap of this method is as follows:

- Environmental modelling: Using the method of adding safety boundary conditions and the envelope method to construct a multi obstacle scene map, setting safety boundaries for each obstacle, and enveloping complex shaped obstacles with simple geometric shapes. Simultaneously calculate the obstacle rate, which is obtained by the ratio of obstacle area to the total map area, to abstractly represent the distribution of obstacles.
- Robot localization: The adaptive Monte Carlo localization (AMCL) algorithm is used to determine the pose and direction of the robot. Using scattered particles to represent the possible position and direction of the robot, updating particle weights based on sensor information, and adaptively adjusting the number and distribution of particles based on the robot's motion state and sensor reliability.
- Trajectory generation speed sampling and candidate trajectories: Based on the current pose and positioning information of the robot, speed sampling is performed, and speed values are selected according to rules within the feasible speed range. Based on the robot

motion model, predict the motion trajectory for each sampling speed within a certain period of time and construct a candidate trajectory set.

- Trajectory generation evaluation and selection: Construct a trajectory evaluation function that comprehensively considers factors such as the distance between the trajectory and obstacles (reference obstacle rate), trajectory length, and motion smoothness. Apply this function to the candidate trajectory set and select the trajectory with the highest score as the preliminary optimal trajectory.
- Path optimization: Improve the A* algorithm by introducing obstacle rates into the evaluation function, making the algorithm more focused on avoiding areas with dense obstacles. Integrate key point extraction algorithms to identify and eliminate redundant road segments. Finally, the optimized path is recombined with the preliminary optimal trajectory to obtain the final optimal obstacle avoidance path planning result.

2. Multi-obstacle scene map generation

In an environment of irregular obstacles, irregular obstacles are assumed to be two types of obstacles, namely convex polygon obstacles and concave polygon obstacles. There are two methods to deal with these two kinds of obstacles, which can be transformed into regular obstacles by adding the safety boundary condition method and the envelope method [14].

For complex scenes, in addition to environmental factors, the irregularity of obstacles is the main influencing factor. To optimize the path, especially to judge whether the robot can move in an irregular, complex obstacle environment, it is necessary to obtain the size of the safety boundary conditions and the robot body, and build the safety distance model d_{safe} as follows:

$$d_{safe} = r + l \quad (1)$$

where, d_{safe} is the safe distance, r is the working radius of the robot, l is an important constant. The value of l should neither be too small (which may cause collisions with obstacles) nor too large (which may hinder the ability to find an optimal path). After adding safety boundary conditions to irregular multi-obstacles, the robot can be imagined as a particle, which can effectively prevent the robot from touching obstacles, as shown in Fig. 1.

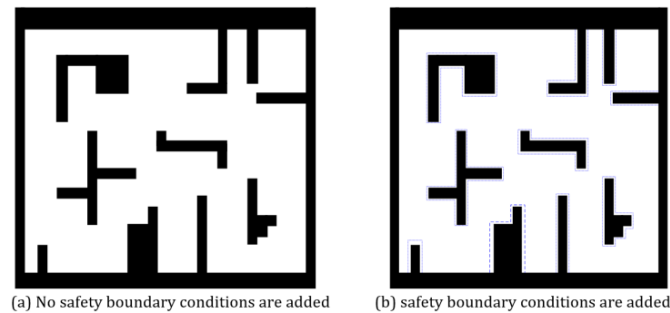


Fig. 1 Multi obstacle scene map

As shown in Fig. 1, in generating a multi-obstacle scene map, it is necessary to quantify its relevant information. The traditional A* algorithm for path planning tends to increase the search space and reduce efficiency, especially in environments with a large number of obstacles [15, 16]. If the obstacle information can be integrated into the evaluation function and the heuristic function can be dynamically adjusted according to the distribution of obstacles, the search space can be greatly reduced, and the search efficiency can be improved. Because the distribution of obstacles on different maps is also different, it is necessary to formulate a standard to represent the distribution of obstacles.

Set an obstacle rate P in the grid map to describe the distribution of obstacles. Set the starting node coordinate as (x_s, y_s) , the target node coordinate as (x_g, y_g) , and the number of obstacle grids as N , then the expression of obstacle rate is as follows:

$$P = \frac{N}{(|x_s - x_g| + 1) \times (|y_s - y_g| + 1)} (P \in (0,1)) \quad (2)$$

3. Positioning and speed sampling

When the robot is in motion, the speed is related to the problem of minimizing time consumption, and the angle is related to the obstacle. Generally speaking, both are also closely related to the path optimization problem. In this paper, the fuzzy logic strategy is introduced. By adjusting the speed and angle of the robot, an asymptotically optimal and smooth path is planned, which makes the path more in line with the dynamic stability of the robot.

3.1 Adaptive Monte Carlo localization

The basic idea of robot obstacle avoidance is to ensure that it is close to the target while avoiding obstacles. The fuzzy logic strategy is realized by the fuzzy logic controller. The fuzzy logic controller is activated when the robot's obstacle sensor detects any obstacle. The fuzzy logic controller will produce corresponding changes in the speed and angle of the robot according to different conditions of the environment [17, 18]. The robot abduction problem refers to the situation where, even though the robot's position and posture are initially known, they are altered due to external interference. As a result, the odometer may provide inaccurate data or fail to offer any control information altogether. The adaptive Monte Carlo localization method solves the problem that the robot cannot use the odometer model to update the particle swarm due to the change of position and posture caused by external forces during the movement of the robot, which eventually leads to the failure of localization [19, 20]. To solve this problem, the adaptive Monte Carlo positioning method randomly scatters particles with a certain probability in the grid map, and the calculation formula is as follows:

$$\begin{cases} w_{slow} = w_s + \partial_{slow}(w_{avg} - w_{slow}) \\ w_{fast} = w_f + \partial_{fast}(w_{avg} - w_{fast}) \end{cases} \quad (3)$$

where, w_{slow} is the long-term estimated weight, w_{fast} is the short-term estimated weight, w_{avg} is the average weight value of the particle swarm, and ∂_{fast} and ∂_{slow} generally take the default values of 0.1 and 0.001.

The formula for calculating the probability of randomly sowing seeds is as follows:

$$\max\left\{0, 1 - \frac{w_{fast}}{w_{slow}}\right\} \quad (4)$$

In the kidnapping problem of mobile robot, the weight value at this time is calculated with the observation value of lidar, the value of w_{avg} will be greatly reduced, and w_{fast} will also be far less than the value of w_{slow} . At this time, the particle swarm will appear randomly in the map, and then focus on global positioning. Finally, the particle swarm will match the real position.

Another advantage of adaptive Monte Carlo localization is that it uses the KLD sampling method, which can adjust the number of particles generated autonomously, so that the efficiency of particle filtering can be further improved. The calculation formula of the maximum number of particles M_x is as follows:

$$M_x = \frac{k-1}{2\varepsilon} \left(1 - \frac{2}{9(k-1)} + \sqrt{\frac{2}{9(k-1)}} Z_{1-\delta} \right)^3 \quad (5)$$

where, k is the number of nonempty particle state spaces, $Z_{1-\delta}$ is the $1 - \delta$ quantile on the normal distribution, and ε is the given value. During initialization, because particles are scattered due to random seeding, the value of k is relatively large, and the required value of M_x is also large. After a long time of continuous updating and convergence, the value of k will gradually decrease, and the value of M_x will also decrease. This summarizes the way of automatically adjusting the number of particles, reduces the amount of calculation, improves efficiency, and shortens the time of the algorithm.

Based on the above, the robot's position and pose positioning in the grid map is completed. Then, to plan an asymptotically optimal and smooth path to make the path more consistent with the stability of the robot's dynamics, the speed and angle of the robot's movement are obtained to determine its direction. Therefore, according to the kinematics formula, the kinematics model of the robot is constructed, which can be expressed as:

$$\dot{v} = \frac{1}{\tau_v} V + \frac{1}{\tau_v} v_c \quad (6)$$

$$\ddot{\theta} = \frac{1}{\tau_\theta} \dot{\theta} + \frac{1}{\tau_\theta} \theta_c \quad (7)$$

where, v_c is the motion speed of the robot, θ_c is the direction angle of the robot, τ_v is the time speed constant of the robot's autonomous motion, V is the initial speed of the robot, τ_θ is the time constant of the angular speed, and θ is the initial angular speed of the robot.

Set the distance ρ from the mobile robot to the starting position. At this time, the relative distance from the mobile robot to the starting position is:

$$x_0 = \rho \cos \theta \quad (8)$$

$$y_0 = \rho \sin \theta \quad (9)$$

where, x_0 is the horizontal distance from the mobile robot to the starting position, y_0 is the vertical distance from the mobile robot to the starting position, and the derivative of displacement $\rho' = v$ is obtained:

$$\dot{x}_0 = v \cos \theta - \rho \dot{\theta} \sin \theta \quad (10)$$

$$\dot{y}_0 = v \sin \theta - \rho \dot{\theta} \cos \theta \quad (11)$$

where, \dot{x}_0 is the horizontal speed of the mobile robot to the starting position, and \dot{y}_0 is the vertical speed of the mobile robot to the starting position.

To sum up, adaptive Monte Carlo is used to complete the positioning of the robot's pose and direction, so as to determine its position in the environment more accurately, and then find the optimal obstacle avoidance path from the starting point to the end point according to the environmental map and obstacle information.

3.2 Speed sampling

The precise control of speed allows the robot to effectively avoid obstacles, thereby enhancing the success rate of obstacle avoidance [21]. Thus, velocity sampling is performed to obtain the optimal velocity combination, and a candidate trajectory set is established based on the trajectory under this velocity constraint. Through subsequent evaluation, the target position can be achieved more efficiently.

In the two-dimensional velocity space, let the velocity set be (v_i, w_i) , where v_i is the linear velocity and w_i is the angular velocity. Because the mobile robot is constrained by its own indoor structure and the surrounding environment, there is a limit to the sampling speed, which mainly has the following three constraints:

- Maximum and minimum speed constraints. In practical applications, due to the limited hardware performance of mobile robots, there is a limit value to the speed. V_m is defined as the set of the maximum and minimum speed constraints of mobile robots, and V_m is expressed as:

$$V_m = \{(v, w) | v \in [v_{min}, v_{max}] \cap w \in [w_{MIN}, w_{max}]\} \quad (12)$$

where, v_{min} is the minimum linear velocity, v_{max} is the maximum linear velocity, w_{min} is the minimum angular velocity, and w_{max} is the maximum angular velocity.

- Motor acceleration and deceleration constraints. Due to the limited torque of the motor and other tools, there are restrictions on the acceleration and deceleration of the motor. V_d is defined as the set of the maximum and minimum speeds that the mobile robot can produce within the predicted time interval, and V_d is expressed as:

$$V_d = \{(v, w) | v \in [v_t - \dot{v}\Delta t, v_t + \dot{v}\Delta t] \cap w \in [w_t - \dot{w}\Delta t, w_t + \dot{w}\Delta t]\} \quad (13)$$

where, v_t is the linear velocity of the mobile robot at time t , w_t is the angular velocity at time t , \dot{v} is the maximum linear velocity of the mobile robot, \dot{w} is the maximum angular velocity, and Δt is the simulation prediction time interval.

- Braking distance constraint. This constraint requires the robot to effectively reduce its speed to zero before the actual collision according to its maximum deceleration when facing the potential collision risk, to avoid any form of collision accident. Define V_a as the speed set of the braking distance constraint, then V_a is represented as:

$$V_a = \{(v, w) \mid v \leq \sqrt{2 \text{dist}(v, w) \cdot \dot{v}} \cap w \leq \sqrt{2 \text{dist}(v, w) \cdot \dot{w}}\} \quad (14)$$

where, $\text{dist}(v, w)$ is the nearest distance from the mobile robot to the obstacle in state (v, w) . Therefore, the speed sampling interval V of the mobile robot in time period Δt is:

$$V = V_m \cap V_d \cap V_a \quad (15)$$

After obtaining the velocity sampling spatial model, according to the current state and environmental information of the mobile robot, the possible trajectories under constraint conditions are predicted to establish a candidate trajectory set. Calculate the cost of the trajectory to score, and select the trajectory with the highest score as the next movement scheme of the mobile robot.

4. Optimal obstacle avoidance path planning for mobile robot

4.1 Trajectory evaluation function

After determining the motion model of the mobile robot and the velocity sampling space V , several feasible trajectories are simulated through multiple groups of selectable velocity combinations in the velocity sampling space. At this time, it is necessary to establish a trajectory evaluation function for selecting the optimal trajectory from these feasible trajectories. In the dynamic serial port algorithm, the trajectory evaluation function consists of the target azimuth evaluation function, the obstacle clearance evaluation function, and the speed evaluation function, as follows:

$$G(v, \omega) = \varepsilon(\alpha \cdot \text{heading}(v, \omega) + \beta \cdot \text{dis}(v, \omega) + \gamma \cdot \text{vel}(v, \omega)) \quad (16)$$

where, $\text{heading}(v, \omega)$ is the target azimuth evaluation function, $\text{dis}(v, \omega)$ is the obstacle clearance evaluation function, $\text{vel}(v, \omega)$ is the velocity evaluation function, α , β and γ are the weights of the target azimuth evaluation function, obstacle clearance evaluation function, and velocity evaluation function respectively, and ε is the normalization coefficient.

- The target azimuth evaluation function $\text{heading}(v, \omega)$ is used to measure the degree to which the end point of the simulated trajectory of the mobile robot faces the target point, as shown in Eq. 17. The included angle between the direction of the end point of the mobile robot trajectory and the connecting line of the target point is θ . The smaller the θ value, the more the mobile robot tends to move towards the target point, and the larger $\text{heading}(v, \omega)$.

$$\text{heading}(v, \omega) = 180^\circ - \theta \quad (17)$$

- The obstacle clearance function $\text{dis}(v, \omega)$ represents the distance from the point on the simulated trajectory of the mobile robot to the nearest obstacle. The larger the value, the farther the mobile robot is from the obstacle on the simulated trajectory, and the higher the safety of mobile robot navigation [22]. The function of d_{\max} in Eq. 18 is to avoid the absence of obstacles on a simulated trajectory, which leads to an excessive gap function of obstacles and accounts for too much in the total trajectory evaluation function.

$$\text{dis}(v, \omega) = \begin{cases} d, & d < d_{\max} \\ d_{\max}, & d \geq d_{\max} \end{cases} \quad (18)$$

where, d is the distance from the point on the trajectory of the mobile robot to the nearest obstacle, and d_{\max} is the threshold value of the distance from the mobile robot to the obstacle.

- The speed evaluation function $vel(v, \omega)$ is used to evaluate the forward speed of the robot in the process of moving towards the target point, as shown in Eq. 19. The larger the value of $vel(v, \omega)$, the faster the mobile robot can approach the target point on the simulated trajectory.

$$vel(v, \omega) = |v_g| \quad (19)$$

where, v_g is the linear velocity corresponding to the simulated trajectory of the mobile robot.

4.2 Improved evaluation function

The heuristic function of A* algorithm can be dynamically adjusted according to the distribution of obstacles. When there are fewer obstacles, the heuristic function will be increased, the search space will be reduced, and the operation speed will be improved [23]. When there are many obstacles, the heuristic function cannot be increased all the time, otherwise, the algorithm will fall into local optimization, and there will be unnecessary turns, which will increase the path cost and make it difficult to search for the optimal path. Therefore, when there are many obstacles, the value of the heuristic function is reduced to make the path more accurate. The search space of the A* algorithm changes with the position of the starting node and the target node [24, 25]. The improved A* algorithm introduces the obstacle rate P into the evaluation function, as shown in Eq. 20, and changes the weights of $g(n)$ and $h(n)$ according to the obstacle rate P , so as to improve the adaptability of the evaluation function $G(v, \omega)$ and ensure that the planned path is globally optimal.

$$G'(v, \omega) = \frac{g(n) + (1 - \ln P)h(m)}{G(v, \omega)} \quad (20)$$

$$g(n) = \sum_{(x_n, y_n) - (x_s, y_s)}^{(x_g, y_g)} \sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2} \quad ((x_0, y_0) = (x_s, y_s)) \quad (21)$$

$$h(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2} \quad (22)$$

where, $g(n)$ is the movement cost from the starting node (x_s, y_s) to the current node (x_n, y_n) , and $h(n)$ is the estimated cost from the current node (x_n, y_n) to the target node (x_g, y_g) .

There are many redundant sections in the path planned by the improved A* algorithm. These redundant sections not only increase the total length of the path, but also may affect the driving efficiency and energy consumption of the robot. To solve this problem, a key point extraction algorithm is introduced. The core idea of this algorithm is to identify and extract the key turning points from the planned path. These turning points are usually the ones with large direction changes or dense obstacle distribution in the path. By extracting these key points, redundant sections in the path can be eliminated, and a more concise, efficient path with fewer inflection points can be obtained. The specific implementation is as follows:

First, all nodes $\{x_1, x_2, \dots, x_n\}$ in the path are obtained. x_1 is the starting point, x_n is the target node, and V is the key speed sampling set. Then, starting from x_1 , connect x_2, x_3, \dots in a straight line until there is an obstacle in the connecting line with x_m , then add x_{m-1} to V . The nodes from x_1 to x_{m-1} are redundant nodes. At this time, set x_{m-1} as the starting point and connect x_m, x_{m+1}, \dots in a straight line in turn, until there is a straight line connection between the starting point and the target point. Last, the target node is added to V , and all nodes in V are connected in a straight line in turn to complete the optimization.

4.3 Smooth path

In actual path planning should meet the following requirements: the path is continuous, the path curvature is continuous, and the path is easy to follow by the mobile robot. In the grid environment, although the path screened by the improved algorithm has achieved remarkable results in reducing the number of paths turns and turning angles, the planned path still shows the characteristics of an unsmooth, that is, the path curvature is discontinuous. Therefore, to ensure that the mobile robot can travel along the planned path more smoothly, it is particularly important to

smooth the filtered path. Among many smoothing methods, the Bezier curve is favoured for its excellent smoothing performance and flexibility.

The Bezier curve can flexibly control the trend and smoothness of the curve by adjusting the position and number of these control points. The filtered path key points are used as the control points of the Bezier curve and connected to form the control polygon, and then a smooth new path is generated by using the generation algorithm of the Bezier curve. This effectively solves the problem of discontinuous path curvature, makes the planned path more in line with the motion characteristics of the mobile robot, and improves the driving stability of the robot.

For $n + 1$ control point P_1, P_2, \dots, P_n , the Bezier curve is:

$$P(t) = \sum_{i=0}^n B_{i,n}(t) \cdot P_i, t \in [0,1] \quad (23)$$

where, P_i is the control point of Bezier curve, t is the independent parameter, and $B_{i,n}(t)$ is the Bernstein basis function of degree n , and it meets the following requirements:

$$B_{i,n}(t) = \sum_{i=0}^n C_n^i \cdot (1-t)^{n-1} \cdot t^i, i = 0,1,2,\dots,n \quad (24)$$

According to the definition of the Bezier curve, P_0 and P_n are the start and end points of the control polygon, respectively. When $n = 1$, the first-order Bezier curve is a straight line with two position points. When $n = 2$, the second-order Bezier curve is a parabola, and there are three position points. When $n \geq 3$, the curve is a higher-order Bezier curve with $n + 1$ position points. To make the path of the Bezier curve fit closer to the path filtered by the improved A* algorithm, the midpoint of each two key path points and the 8 equidistant points close to them are added between each two key path points, and this series of path points is set as the control points of the Bezier curve to generate a smooth path. To sum up, the final optimal obstacle avoidance path planning of mobile robot is achieved.

5. Simulation results and analysis

5.1 Simulation preparation and test indicators

To test the feasibility and effectiveness of local dynamic planning in multi obstacle scenes, that is, whether the mobile robot can avoid obstacles, the proposed algorithm is simulated and tested. Establish a grid map with a side length of 1 m and a size of 50×50 . Set the parameters of the evaluation function as $\alpha = 0.2, \beta = 0.3, \gamma = 0.5$, and the parameters of the improved evaluation function as $\alpha = 0.2, \beta = 0.3, \gamma = 0.15, \lambda = 0.35$. The settings of various speed parameters and time parameters are shown in Table 1. Under the above parameter settings, the global path dynamic obstacle avoidance parameter test software setting interface is shown in Fig. 2.

To provide a comprehensive evaluation of performance and validate the proposed algorithm, the following indicators are defined:

- Path length: it is the key index to evaluate the actual moving distance of the robot from the starting point to the end point. In the field of robot navigation and path planning, the shorter path length shows that the robot can reach the destination more efficiently.
- Obstacle avoidance ability: evaluate the reaction speed and the success rate of obstacle avoidance when the robot encounters obstacles. The reaction speed can be measured by time. The shorter the reaction time, the better the obstacle avoidance capability. The calculation formula for the success rate of obstacle avoidance is as follows.

$$Z_k = \frac{Z_1}{Z} \times 100 \% \quad (25)$$

- Path smoothness: evaluate whether the planned path is smooth and whether there are too many transitions and jitters. The smooth path can reduce the energy consumption and mechanical wear of the robot, and improve the stability and comfort of the movement. The calculation formula is as follows:

$$S = 1 / \sum |L_i - L_{i-1}| / \sum L_i \quad (26)$$

where, L_i is the length of the i -th segment in the path. The higher the path smoothness result, the smoother the path.

Table 1 simulation parameter settings

Parameter name	Numerical value	Parameter name	Numerical value
v_{max}	2 m/s	v_{min}	0
w_{max}	0.7 rad/s	w_{min}	-0.7 rad/s
\dot{v}_{max}	0.4 m/s ²	\dot{v}_{min}	0
\dot{w}_{max}	1.7 rad/s ²	\dot{w}_{min}	-1.7 rad/s ²
prediction time T	3 s	time interval Δt	0.1 s



Fig. 2 Global path dynamic obstacle avoidance parameter test software setting interface

5.2 Effectiveness result analysis

To verify the effectiveness of the proposed algorithm, the improved A* algorithm and the algorithm fused with key point extraction are respectively used for obstacle avoidance path planning, and the results are shown in Fig. 3.

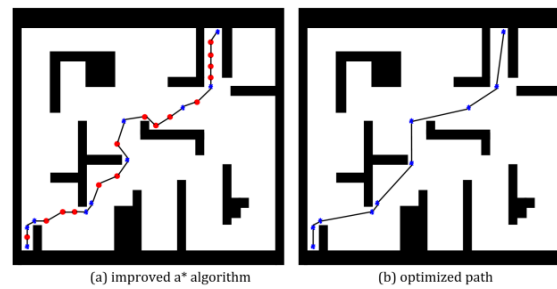


Fig. 3 Obstacle avoidance path planning results

According to the results in Fig. 3, the corresponding distance results can be obtained through the test software. After optimization, compared with the path planned by the improved A* algorithm, the results of the proposed fusion algorithm have been significantly improved. Specifically, the path length is reduced from 39.2536 to 36.9805, which is 5.79 %. This shows that the robot can reach the destination faster and improve the overall execution efficiency. Moreover, the optimized path also significantly reduces the number of nodes. The original 25 nodes have been reduced to only 10 nodes, and the optimization range is as high as 60 %. This not only reduces the complexity of the algorithm, but also helps to improve the stability and reliability of the robot in actual operation. In terms of path inflection points, the number of inflection points was effectively reduced from the original 17 to 8, with an optimization rate as high as 52.94 %. This

shows that the path becomes smoother with the proposed hybrid algorithm, and the robot can reduce unnecessary steering and pauses during driving, further improving the driving efficiency.

To sum up, the results show that the improved A* algorithm and the key point extraction algorithm are better than the improved A* algorithm, which can successfully shorten the path length, reduce the number of nodes and inflection points, thus achieving a significant improvement in robot path planning. This shows that the proposed method is effective and can realize the optimal obstacle avoidance path planning of an intelligent mobile robot.

5.3 Path length result analysis

Based on the above, 10 different 50×50 grid maps are randomly simulated, and the MGWO algorithm in literature [12] and the hybrid algorithm in literature [13] are used as comparison algorithms to test the path length index with the proposed algorithm. The results are shown in Table 2.

According to the results obtained in Table 2, the proposed algorithm is used for obstacle avoidance path planning of 10 scenes, and the path length of each scene after planning is lower than that of another comparison algorithm. The shortest path length after planning using the MGWO algorithm in the literature [12] is 39.658 m, the shortest path length after planning using the hybrid algorithm in the literature [13] is 38.976 m, and the path lengths after planning using the proposed algorithm are less than 37 m. Therefore, using the proposed algorithm for the obstacle avoidance path planning, the path lengths are the shortest, and it can arrive at the target location more quickly and with higher efficiency.

Table 2 Path length results

Scene map number	Path length of the proposed algorithm (m)	MGWO algorithm path length (m)	Hybrid algorithm path length (m)
1	35.765	40.657	39.864
2	35.981	40.879	39.958
3	34.867	39.658	38.976
4	36.724	42.784	40.659
5	35.241	40.823	39.463
6	35.547	40.985	39.769
7	36.837	42.859	40.858
8	35.531	40.991	39.713
9	36.794	41.759	40.957
10	35.785	41.348	39.897

5.4 Analysis of obstacle avoidance results

Based on the above test scenario, the proposed algorithm, the MGWO algorithm in literature [12], and the hybrid algorithm in literature [13] are used to test the obstacle avoidance ability index. The results are shown in Fig. 4.

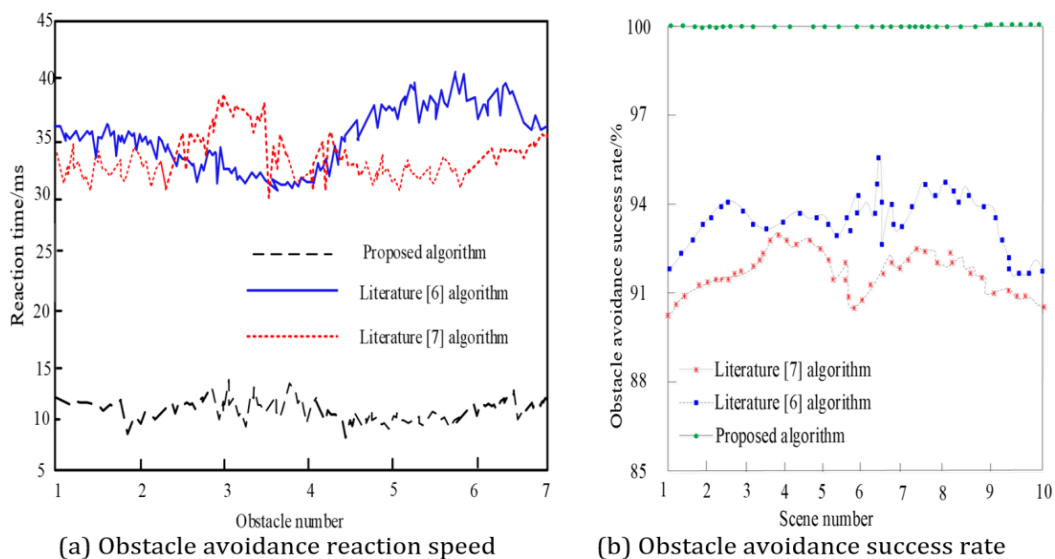


Fig. 4 obstacle avoidance results

According to the results obtained in Fig. 4(a), when the above three algorithms are used for obstacle avoidance path planning, the obstacle avoidance reaction speed is less than 1s, but the comparison of the three results shows that the obstacle avoidance reaction speed of the proposed algorithm is lower than that of the two literature comparison algorithms. When the proposed algorithm is used for obstacle avoidance path planning, its obstacle avoidance response speed can always be maintained within 15 ms. The MGWO algorithm in literature [12] has the lowest obstacle avoidance reaction speed of 32 ms, and the hybrid algorithm in literature [13] has the lowest obstacle avoidance reaction speed of 30 ms. After comparing the three algorithms, the proposed algorithm has a lower obstacle avoidance reaction speed when facing obstacles. According to the results in Fig. 4(b), using the proposed algorithm for obstacle avoidance path planning, the success rate of obstacle avoidance can reach 100 %, and the effect of obstacle avoidance is better. The highest success rate of obstacle avoidance is 93 % using MGWO algorithm in literature [12] and 95 % using hybrid algorithm in literature [13]. After comparing the success rate of obstacle avoidance, it can be concluded that the proposed algorithm can effectively complete the obstacle avoidance path planning. The above results show that the proposed algorithm has a strong obstacle avoidance capability for obstacle avoidance path planning.

5.5 Analysis of path smoothness results

Based on the above test scenario, the proposed algorithm, the MGWO algorithm in literature [12] and the hybrid algorithm in literature [13] are used to test the path smoothness index. The results are shown in Fig. 5.

According to the results obtained in Fig. 5, after the proposed algorithm is used for obstacle avoidance path planning for 10 scenes, the path smoothness results are higher than 97 %. While using the MGWO algorithm in literature [12], its path smoothness result is up to 88 %, and using the hybrid algorithm in literature [13] its path smoothness result is up to 90 %. Therefore, using the proposed algorithm for planning with high path smoothness can effectively reduce the energy consumption and mechanical wear and tear of the robot, and improve the stability and comfort of the movement.

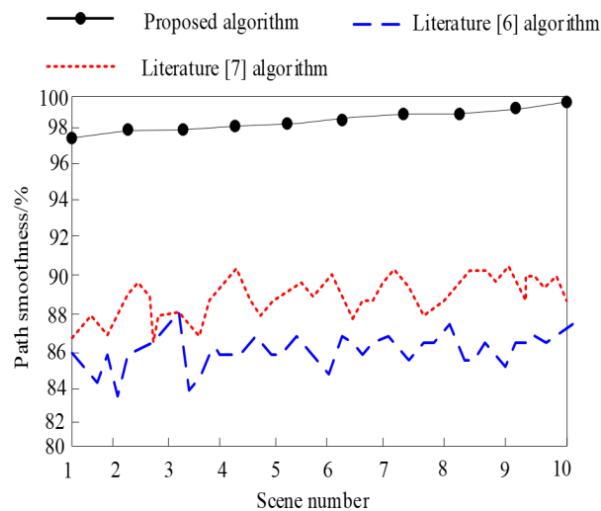


Fig. 5 Path smoothness results

6. Conclusion

With the development of technology, intelligent mobile robots are widely used, but safe and efficient task execution in multi obstacle environments is an urgent problem that needs to be solved in the field of robotics. This study conducted optimal obstacle avoidance path planning for robots in multi obstacle scenarios. The safety boundary condition method and envelope method were used to generate maps and obtain obstacle rates. Adaptive Monte Carlo localization was used, and after velocity sampling and constructing evaluation functions to select trajectories, the

improved A* algorithm was used to introduce obstacle rates and incorporate keypoint extraction algorithms to re plan the path and smooth it out. Although the path planned by this algorithm is short, with an obstacle avoidance response speed of 32 ms, a success rate of 100 %, and a path smoothness of 97 %, it has a driving effect on the development of the robotics field. However, there are potential weaknesses: the construction of maps relies on empirical values, the particle convergence speed of positioning algorithms in special environments may slow down, and the accuracy may be affected. The sampling range and method when constructing trajectories through speed sampling may also affect the results.

Acknowledgement

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (Grant No.: 62001197).

References

- [1] Gültekin, A., Diri, S., Becerikli, Y. (2023). Simplified and smoothed rapidly-exploring random tree algorithm for robot path planning, *Tehnički Vjesnik – Technical Gazette*, Vol. 30, No. 3, 891-898, doi: [10.17559/TV-20221015080721](https://doi.org/10.17559/TV-20221015080721).
- [2] Khaleel, H.Z., Oleiwi, B.K. (2024). Design and implementation low cost smart cleaner mobile robot in complex environment, *Mathematical Modelling of Engineering Problems*, Vol. 11, No. 10, 2869-2877, doi: [10.18280/mmep.111030](https://doi.org/10.18280/mmep.111030).
- [3] Jiang, T., Zhang, S., Wang, R., Wang, S. (2023). Development and verification of an autonomous and controllable mobile robot platform, *Mechatronics and Intelligent Transportation Systems*, Vol. 2, No. 1, 11-19, doi: [10.56578/imits020102](https://doi.org/10.56578/imits020102).
- [4] Muthukumaran, S., Sivaramakrishnan, R. (2019). Optimal path planning for an autonomous mobile robot using dragonfly algorithm, *International Journal of Simulation Modelling*, Vol. 18, No. 3, 397-407, doi: [10.2507/IJSIMM18\(3\)474](https://doi.org/10.2507/IJSIMM18(3)474).
- [5] Khaleel, R.Z., Khaleel, H.Z., Al-Hareeri, A.A.A., Al-Obaidi, A.S.M., Humaidi, A.J. (2024). Improved trajectory planning of mobile robot based on pelican optimization algorithm, *Journal Européen des Systèmes Automatisés*, Vol. 57, No. 4, 1005-1013, doi: [10.18280/jesa.570408](https://doi.org/10.18280/jesa.570408).
- [6] Dabic-Miletic, S. (2024). The challenges of integrating AI and robotics in sustainable WMS to improve supply chain economic resilience, *Journal of Industrial Intelligence*, Vol. 2, No. 2, 119-131, doi: [10.56578/jii020205](https://doi.org/10.56578/jii020205).
- [7] Korabayev, Y., Kosbolov, S., Kubesova, G., Shurenov, M., Duisebayeva, K. (2024). Investigating the design and application of mobile robotic devices with manipulation devices for space technology, *Journal Européen des Systèmes Automatisés*, Vol. 57, No. 2, 345-352, doi: [10.18280/jesa.570204](https://doi.org/10.18280/jesa.570204).
- [8] Jiang, T., Zhang, S., Wang, R., Wang, S. (2023). Development and verification of an autonomous and controllable mobile robot platform, *Mechatronics and Intelligent Transportation Systems*, Vol. 2, No. 1, 11-19, doi: [10.56578/imits020102](https://doi.org/10.56578/imits020102).
- [9] Shi, W., Tang, D.B., Zou, P. (2021). Multi-objective automated guided vehicle scheduling based on MapReduce framework, *Advances in Production Engineering & Management*, Vol. 16, No. 1, 37-46, doi: [10.14743/apem2021.1.383](https://doi.org/10.14743/apem2021.1.383).
- [10] Li, J., Li, J., Fang, H., Jiang, J. (2024). Dynamic energy-efficient path planning for electric vehicles using an enhanced ant colony algorithm, *Tehnički Vjesnik – Technical Gazette*, Vol. 31, No. 2, 434-441, doi: [10.17559/TV-20230510000618](https://doi.org/10.17559/TV-20230510000618).
- [11] Chen, J., Tan, C., Mo, R., Zhang, H., Cai, G., Li, H. (2021). Research on path planning of three-neighbor search A* algorithm combined with artificial potential field, *International Journal of Advanced Robotic Systems*, Vol. 18, No. 3, Article No. 17298814211026449, doi: [10.1177/17298814211026449](https://doi.org/10.1177/17298814211026449).
- [12] Mohanraj, T., Dinesh, T., Guruchandhramavli, B., Sanjai, S., Sheshadhri, B. (2023). Mobile robot path planning and obstacle avoidance using hybrid algorithm, *International Journal of Information Technology*, Vol. 15, No. 8, 4481-4490, doi: [10.1007/s41870-023-01475-5](https://doi.org/10.1007/s41870-023-01475-5).
- [13] Chen, Y., Wang, P., Lin, Z., Sun, C. (2023). Global path guided vehicle obstacle avoidance path planning with artificial potential field method, *IET Cyber-Systems and Robotics*, Vol. 5, No. 1, Article No. e12082, doi: [10.1049/csy2.12082](https://doi.org/10.1049/csy2.12082).
- [14] Xia, W., Wang, W., Gao, C. (2022). Trajectory optimization with obstacles avoidance via strong duality equivalent and hp-pseudospectral sequential convex programming, *Optimal Control Applications and Methods*, Vol. 43, No. 2, 566-587, doi: [10.1002/oca.2839](https://doi.org/10.1002/oca.2839).
- [15] Qiu, G., Li, J. (2024). Path planning for unified scheduling of multi-robot based on BSO algorithm, *Journal of Circuits, Systems and Computers*, Vol. 33, No. 7, Article No. 2450133, doi: [10.1142/S0218126624501330](https://doi.org/10.1142/S0218126624501330).
- [16] Jin, X., Wang, Z. (2022). Proximal policy optimization based dynamic path planning algorithm for mobile robots, *Electronics Letters*, Vol. 58, No. 1, 13-15, doi: [10.1049/ell2.12342](https://doi.org/10.1049/ell2.12342).

- [17] Satir, E., Baser, E. (2022). Optimization of interval type-2 fuzzy logic controller using real-coded quantum clonal selection algorithm, *Elektronika Ir Elektrotechnika*, Vol. 28, No. 3, 4-14, [doi: 10.5755/j02.eie.31148](#).
- [18] Chang, C.-K., Elmashtoly, A.M. (2022). Protection coordination index assessment using fuzzy logic controller, *Energies*, Vol. 15, No. 4, Article No. 1377, [doi: 10.3390/en15041377](#).
- [19] Akai, N. (2023). Reliable Monte Carlo localization for mobile robots, *Journal of Field Robotics*, Vol. 40, No. 3, 595-613, [doi: 10.1002/rob.22149](#).
- [20] Jaenal, A., Moreno, F.-A., Gonzalez-Jimenez, J. (2023). Sequential Monte Carlo localization in topometric appearance maps, *The International Journal of Robotics Research*, Vol. 42, No. 13, 1117-1132, [doi: 10.1177/02783649231197723](#).
- [21] Xue, D., Wu, D., Yamashita, A.S., Li, Z. (2023). Proximal policy optimization with reciprocal velocity obstacle based collision avoidance path planning for multi-unmanned surface vehicles, *Ocean Engineering*, Vol. 273, Article No. 114005, [doi: 10.1016/j.oceaneng.2023.114005](#).
- [22] Fang, Z., Liang, X. (2022). Intelligent obstacle avoidance path planning method for picking manipulator combined with artificial potential field method, *Industrial Robot*, Vol. 49, No. 5, 835-850, [doi: 10.1108/IR-09-2021-0194](#).
- [23] Sui, F., Tang, X., Dong, Z., Gan, X., Luo, P., Sun, J. (2023). ACO+PSO+A*: A bi-layer hybrid algorithm for multi-task path planning of an AUV, *Computers & Industrial Engineering*, Vol. 175, Article No. 108905, [doi: 10.1016/j.cie.2022.108905](#).
- [24] Li, C., Huang, X., Ding, J., Song, K., Lu, S. (2022). Global path planning based on a bidirectional alternating search A* algorithm for mobile robots, *Computers & Industrial Engineering*, Vol. 168, Article No. 108123, [doi: 10.1016/j.cie.2022.108123](#).
- [25] Wang, Y.J., Liu, X.Q., Leng, J.Y., Wang, J.J., Meng, Q.N., Zhou, M.J. (2022). Study on scheduling and path planning problems of multi-AGVs based on a heuristic algorithm in intelligent manufacturing workshop, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 505-513, [doi: 10.14743/apem2022.4.452](#).

Modelling and optimization of dimensional accuracy and surface roughness in dry turning of Inconel 625 alloy

Vukelic, D.^{a,*}, Milosevic, A.^a, Ivanov, V.^b, Kocovic, V.^c, Santosi, Z.^a, Sokac, M.^a, Simunovic, G.^d

^aUniversity of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia

^bSumy State University, Faculty of Technical Systems and Energy Efficient Technologies, Sumy, Ukraine

^cUniversity of Kragujevac, Faculty of Engineering, Kragujevac, Serbia

^dUniversity of Slavonski Brod, Mechanical Engineering Faculty in Slavonski Brod, Slavonski Brod, Croatia

ABSTRACT

The CNC dry turning of Inconel 625 alloy was investigated in this study. Turning was performed using different feeds, corner radii, and insert types. The dimensions and the arithmetic mean surface roughness were measured before and after turning. The influence of the input parameters and the process modelling were evaluated using a full factorial design of experiments. Prediction models were developed, and process optimization was carried out to simultaneously maximize dimensional accuracy and surface quality. The optimal values of the input parameters were identified for the wiper insert, a feed rate of 0.1 mm/rev, and a corner radius of 0.8 mm. Under these conditions, the optimal deviation from the specified dimension was 0.2 mm, while the surface roughness was 0.297 μm . During the confirmation phase, the mean percentage errors were 0.9 % for surface roughness and 3.45 % for dimensional deviation. The percentage errors observed in the confirmation experiments, all of which were below 5 %, demonstrate the feasibility of using the proposed approach for modelling and optimizing the turning of Inconel 625 alloy.

ARTICLE INFO

Keywords:

CNC dry turning;
Inconel 625 alloy;
Surface roughness;
Dimensional accuracy;
Optimization;
Modelling

*Corresponding author:

vukelic@uns.ac.rs
(Vukelic, D.)

Article history:

Received 2 August 2024
Revised 13 September 2024
Accepted 18 September 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Turning is one of the machining processes most frequently used for machining cylindrical workpieces. It can often be used as a finish operation without the need for further machining. In this case, it is crucial to perform the machining within the tolerance limits and the required surface roughness [1,2]. In other words, it is crucial to solve the problem of dimensional accuracy and the surface quality [3]. Considering all variables is very difficult due to their large number. Therefore, it is necessary to identify and evaluate the most influential input parameters. Different materials can be machined by finish turning using different machine tools, cutting tools and fixtures [4,5]. The problem of achieving the required dimension and surface roughness becomes much more complicated when machining Inconel alloys, especially under dry conditions. Inconel alloys are difficult-to-cut materials because turning is associated with high temperatures and many accompanying effects such as poor surface finish, intense tool wear, microstructural changes in the material, etc. [6].

In previous research, the turning of Inconel 625 alloy was studied under different aspects and conditions.

Kushwaha and Singh [7] investigated the influence of cutting speed, feed and depth of cut on surface roughness during dry and wet turning using the Taguchi method. The lowest surface roughness was achieved when all input parameters were at the lowest level during wet turning. Kosaraju *et al.* [8] investigated the influence of cutting speed, feed and depth of cut on cutting force and surface roughness using the Taguchi method. Jeykrishnan *et al.* [9] evaluated the influence of cutting speed, feed, and depth of cut on surface roughness using the Taguchi method. The results showed a dominant influence of feed on surface roughness, with increasing feed the surface roughness deteriorated. Dhananchezian [10] investigated the influence of cutting speed and type of cooling on surface roughness and tool wear. The surface roughness decreased with the increased speed and the used of cryogenic cooling. Liquid nitrogen had a significant effect on reduced tool wear. Magri *et al.* [11] investigated the influence of coolant direction on surface roughness and tool life. The coolant directed to the flank face contributed to the improvement of surface roughness, while directing the coolant to the flank and rake faces simultaneously did not contributed to the increase in tool life. Singh and Padhy [12] investigated the effects of cooling and lubrication conditions, cutting speed, feed and depth of cut on surface roughness, tool wear, cutting forces and temperature. Dry turning showed the worst results and wet or minimum quantity lubrication (MQL) turning showed the best results. Yildirim [13] investigated the influence of lubrication and cooling conditions on tool wear, temperature and surface roughness. The hexagonal boron nitride (hBN) liquid nitrogen cooling method showed the best performance. Yildirim *et al.* [14] investigated the effect of hBN nanoparticles at MQL on surface roughness, tool wear and temperature. The use of hBN nanofluid effectively reduced both tool wear and surface roughness in machining processes. Singh and Padhy [15] investigated the effect of lubrication method, cutting speed, feed and depth of cut on temperature. The regression model showed that the temperature increased with the increased machining parameters and that the feed had the greatest influence. Nano MQL lubrication conditions also showed better results in terms of temperature, tool wear and chip morphology. Yagmur [16] presented the dependence of tool life, temperature and surface roughness in relation to cutting conditions, cutting speed and feed. MQL provided the best results compared to dry and vortex cooling methods. Padhy and Singh [17] presented a finite element analysis (FEA) of the influence of the type of lubrication (dry, MQL, NMQL) on cutting force, temperature and tool wear. The NMQL conditions showed the best results in all aspects. Rakesh and Chakradhar [18] investigated the influence of cutting speed, feed, depth of cut and cooling conditions on surface roughness, tool wear and cutting force. Cryogenic cooling (LN₂ air) produced the best parameters. With increased feed and depth of cut, surface roughness, cutting forces and tool wear increased. With increased speed, surface roughness and cutting forces decreased, while tool wear increased. Surface roughness improved with increasing cutting speed and decreasing feed and depth of cut. Makhesana *et al.* [19] investigated the influence of different lubrication and cooling conditions on surface roughness, tool wear, chip morphology, hardness and power consumption. The most favourable results were obtained with nMQL with graphite and MoS₂.

As can be seen, the turning of Inconel 625 alloy has been intensively studied using near-dry turning methods. The results obtained indicate the efficiency of these methods, as the results are very close to wet turning. However, a lower amount of coolant and lubricant results in less contamination, which is still present to a lesser extent. In addition, equipment, fluid, fluid treatment requirements and the like incur additional costs. Therefore, dry turning should not be eliminated and replaced by near-dry turning.

Several studies were carried out exclusively under the conditions of dry turning of the Inconel 625 alloy. Marimuthu and Baskaran [20] evaluated the effect of cutting speed, feed and depth of cut on surface roughness and material removal rate (MRR) using the Taguchi method. Feed and cutting speed had a significant effect on surface roughness, and feed and depth of cut had a significant effect on MRR. Venkatesan *et al.* [21] investigated the influence of cutting speed, feed and depth of cut on cutting force and surface roughness using the Taguchi method. Feed and cutting speed had the greatest influence on surface roughness, and feed had the greatest influence on cutting force components. Ramanujam *et al.* [22] investigated the influence of cutting speed, feed

and depth of cut on surface roughness, power consumption and MRR using the Taguchi method and fuzzy logic. The greatest influence on the output parameters had feed. Jain *et al.* [23] presented the influence of spindle speed, feed and depth of cut on MRR. Spindle speed and feed had the greatest influence on MRR. Prokes *et al.* [24] presented the dependence of surface roughness on cutting speed, feed and depth of cut. Only the feed showed a significant influence on the surface roughness. Lotfi *et al.* [25] presented a FEA to predict tool wear as a function of cutting speed, feed and depth of cut for two types of cutting tools (coated carbide and ceramic). The depth of cut had a significant effect, and the feed had no effect on tool wear. Cutting speed had an effect on tool flank wear, but only for coated carbide tools. Hemakumar and Kuppan [26] presented the effects of cutting speed and feed on cutting force, surface roughness and flank wear using a full factorial experimental design. They found that feed had the greatest effect on cutting force and surface roughness, and that cutting speed had the greatest effect on flank wear. Vasudevan *et al.* [27] presented the influence of cutting speed, feed and depth of cut on surface roughness and MRR. Feed was the most influential factor on both output parameters. Waghmode and Dabade [28] investigated the dependence of cutting speed, feed and depth of cut on cutting forces and surface roughness. The cutting forces increased with increased depth of cut and feed. The increased in feed also had an effect on the increased in surface roughness. Padhy and Singh [29] investigated the effects of cutting speed, feed and depth of cut on surface roughness, cutting force and MRR using Taguchi method. Depth of cut and feed had the greatest effect on forces, cutting speed and feed had the greatest effect on surface roughness, and cutting speed had the greatest effect on MRR. Narkhede *et al.* [30] investigated the influence of cutting speed, feed and depth of cut on surface roughness during dry turning and under cryogenic coolant conditions. The lowest surface roughness was found at the lowest feed and the smallest depth of cut as well as the highest cutting speed. Cutting forces and temperature were measured under the most favourable turning conditions. It was proved that the performance parameters are improved by using cryogenic coolant. Padhy and Singh [31] applied the Taguchi method to analyse the influence of cutting speed, feed and depth of cut on cutting force, surface roughness and MRR. The most influential parameter for surface roughness and MRR was cutting speed, and the most important parameter for cutting force was feed. Sim *et al.* [32] investigated the effects of cutting speed, feed and depth of cut on cutting force and temperature, using fuzzy logic to optimize the process. The depth of cut had the greatest effect on cutting forces and temperature.

Compared to alternative techniques such as minimum quantity lubrication, minimum quantity cooling lubrication or cryogenic cooling, dry turning of the alloy Inconel 625 is the most environmentally friendly and the chips are the easiest to recycle. With dry turning, there are no problems associated with the cutting fluid (disposal, recycling, reuse), the costs of which can be considerable. The disadvantages of dry turning are the occurrence of higher temperatures in the cutting zone, which leads to higher tool wear and shorter tool life. Dry turning also causes higher energy consumption due to the higher cutting forces. Finally, surface roughness and dimensional accuracy can deteriorate due to the above-mentioned phenomena. These negative effects of dry turning do not occur if the machining parameters (cutting speed, feed and depth of cut) have lower values, and the machining time is not long. For sustainable production, however, it is necessary to find a balance between the required characteristics of accuracy and quality, productivity and also the environmental requirements of modern production systems. To eliminate the previous problems in dry turning, high-precision machine tools and rigid fixtures must be used to ensure reliable locating and secure clamping. In addition, coated cutting tools with suitable geometry should be used to improve tribological effects, i.e. to keep cutting temperatures and energy consumption as low as possible. Finally, it is necessary to optimize the turning parameters to minimize temperature and reduce wear, especially during prolonged dry turning.

It can be noted that in previous research in the field of dry turning, various methods (Taguchi, RSM, fuzzy logic, etc.) have been used to analyse and optimise the output parameters of turning (surface roughness, cutting force, temperature) based on the input parameters. All the research carried out so far in the field of dry turning of Inconel 625 alloy has its advantages and disadvantages. It should be noted that there has been no research to date in which dimensional accuracy and surface roughness have been integrally investigated. The interactions between the input

parameters and their influence on the output parameters have also not been investigated, nor has the fact that the input parameters are usually turning modes (cutting speed, feed and depth of cut).

In contrast to the previous work, the main contribution of this study is the evaluation of the influence of the dominant factors and their interactions on dimensional accuracy and surface roughness. Feed, corner radius and insert type were selected as input parameters. The output parameters, which were evaluated, estimated, modelled and then optimised, were the dimensional deviation and the arithmetic mean of the surface roughness. First, the influence of feed, corner radius and insert type on dimensional accuracy and surface roughness was evaluated. Then the effects were evaluated, and the most important factors and their interactions were selected. The turning process was then modelled and optimised. The validity of the modelling and optimization was evaluated by confirmatory tests. In view of the increasing environmental protection requirements and legal obligations in many countries, research was carried out under dry conditions.

2. Materials and methods

The flowchart according to which the investigation was carried out is shown in Fig. 1.

The investigations were carried out on workpieces made of Inconel 625 alloy, whose chemical composition is as follows: 58-71 % Ni, 21-23 % Cr, 8-10 % Mo, 5 % Fe, 3.2-3.8 % Nb + Ta, ≤ 1 % Co, ≤ 1 % Mn and ≤ 0.40 % Al. The properties of the Inconel 625 alloy are: modulus of elasticity $2.1 \cdot 10^5$ MPa, density 7.8 g/cm^3 , tensile strength 990 MPa, yield strength 516 MPa, hardness 160 HB, thermal expansion $15.8 \text{ } \mu\text{m/m}^\circ\text{C}$, and thermal conductivity 8.5 W/mK . The dimensions of the workpieces are $\varnothing 44 \times 420 \text{ mm}$.

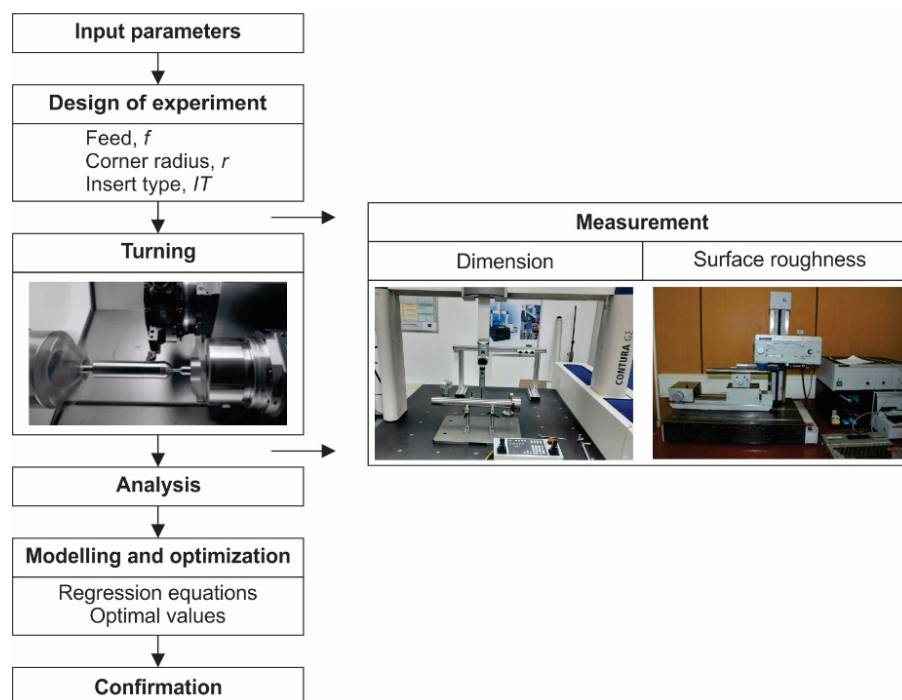


Fig. 1 Flowchart of the methodology

The finish longitudinal turning is carried out on a CNC lathe. The workpiece is locating and clamped using a chuck and a turning centre. The turning parameters and the inserts were selected in accordance with the recommendations of the insert manufacturer and based on the properties of the workpiece material, the workpiece geometry, the type of process, the properties of the technological equipment and the process stability. The cutting speed and feed were selected at lower but recommended levels in order to minimise the effects of tool wear on the results obtained. The cutting speed is 60 m/min . The insert size was selected to achieve the required depth of cut and

the mutual relationship between cutting edge length and effective cutting edge length was considered. The assumed depth of cut $a_p = 1.5$ mm is greater than the corner radius of any insert. A new PVD TiAlN+TiAlN insert was used for each experiment. The common characteristics of the inserts are: square shape, thickness 4.8 mm, cutting edge length 12.3 mm, fixing hole diameter 5.2 mm and inscribed circle diameter 12.7 mm.

Three input parameters were varied during the experimental research. The feed (f) was treated as a continuous factor, the corner radius (r) as a categorical numerical factor and the insert type (IT) as a categorical attributive factor. The input parameters used in this study are as follows: $f = 0.1, 0.15, 0.2, 0.25, 0.3$ mm/rev, $r = 0.2, 0.4, 0.8, 1.2$ mm and $IT = \text{standard, wiper}$.

The dimensional accuracy is quantified by the deviation of the theoretical dimension in relation to the realized dimension. Since the diameters of the workpieces differ by small values defined by the width of the tolerance field, the measurement deviation for each experiment is calculated as follows:

$$\Delta D_i = D_{2ti} - D_{2mi} \quad i=1, 2, \dots, 40 \quad (1)$$

where D_{2t} is the theoretical value of the diameter after turning (required nominal dimension) and D_{2m} is the measured (real) value of the diameter after turning. The theoretical value of the diameter before turning is calculated as follows:

$$\Delta D_{2ti} = D_{1i} - 2 \cdot a_p \quad i=1, 2, \dots, 40 \quad (2)$$

where D_{1i} is the measured diameter value before turning and a_p is the depth of cut.

The measurements of dimensions and surface roughness were carried out under controlled microclimatic conditions at a temperature of $20^\circ\text{C} \pm 1^\circ\text{C}$.

The surface roughness was measured with a Talysurf measuring device. The measurements were carried out in the feed direction with a sampling length of 0.8 mm and an evaluation length of 4 mm using a Gaussian filter. Five measurements were taken to reduce the relative measurement error. The surface roughness is quantified by the arithmetic mean surface roughness (Ra).

The dimensions were measured on a Carl Zeiss Contura G2 CMM. The accuracy of the CMM is $1.9 + L/300$ μm (L – measuring length). A measuring probe with a length of 75 mm was used. The diameter of the ball was 5 mm. The measurement strategy used is point by point. By measuring five reference diameters, the relative error of the diameter change is reduced.

A randomized experimental design is used to modelling the turning process. During modelling, an evaluation of the main effects and interactions is first carried out. This makes it possible to exclude the main factors and interactions from the regression due to their lack of significance and to create the conditions for conducting the experiment with minimal prediction error. An analysis of variance and an analysis of model adequacy are then carried out, as well as the creation of regression equations for dimensional accuracy (dimensional deviation, ΔD) and the quality of the treated surface (surface roughness, Ra).

The turning process is optimized using D-optimal design of experiments, which enables a systematic investigation of the input parameters and their interactions to determine the optimal setting of the input parameters to achieve the specified objective function – simultaneous maximization of accuracy and quality (minimization of dimensional deviation and surface roughness).

3. Results

3.1 Experiments

The experimental investigations were carried out according to the full factorial experimental design, which allows the investigation of all combinations of levels of the input parameters. Considering the fact that 5 levels were assumed for the feed, 4 levels for the corner radius and 2 levels for the insert type, a total of $5 \times 4 \times 2 = 40$ tests were carried out. The measurement results for different combinations of input parameters performed according to the randomized experimental design are shown in Table 1.

Table 1 Experimental results

No.	f (mm/rev)	r (mm)	Insert	Ra (μ m)	ΔD (mm)
1	0.1	0.2	Wiper	1.316	0.101
2	0.1	0.4	Standard	1.251	0.048
3	0.25	0.4	Wiper	3.908	0.312
4	0.15	0.2	Wiper	2.961	0.225
5	0.25	0.2	Wiper	8.224	0.617
6	0.25	0.2	Standard	15.625	0.586
7	0.15	0.2	Standard	5.625	0.213
8	0.15	0.4	Wiper	1.406	0.112
9	0.3	1.2	Wiper	2.814	0.367
10	0.15	0.8	Wiper	0.669	0.056
11	0.1	1.2	Standard	0.665	0.035
12	0.15	0.8	Standard	1.406	0.052
13	0.1	0.2	Standard	2.503	0.095
14	0.2	0.8	Standard	2.502	0.091
15	0.3	0.2	Standard	22.503	0.859
16	0.2	1.2	Standard	2.664	0.148
17	0.25	0.8	Standard	3.906	0.141
18	0.15	0.4	Standard	2.812	0.104
19	0.25	1.2	Wiper	1.953	0.252
20	0.1	0.4	Wiper	0.625	0.052
21	0.3	0.4	Wiper	5.622	0.452
22	0.2	0.2	Wiper	5.263	0.401
23	0.3	0.2	Wiper	11.844	0.905
24	0.2	0.8	Wiper	1.192	0.101
25	0.2	0.4	Standard	5.004	0.187
26	0.3	0.8	Wiper	2.678	0.227
27	0.1	1.2	Wiper	0.314	0.039
28	0.1	0.8	Standard	0.627	0.022
29	0.15	1.2	Wiper	0.705	0.091
30	0.25	0.8	Wiper	1.861	0.157
31	0.2	0.2	Standard	10.008	0.382
32	0.25	0.4	Standard	7.813	0.292
33	0.2	1.2	Wiper	1.254	0.161
34	0.2	0.4	Wiper	2.503	0.202
35	0.15	1.2	Standard	1.502	0.083
36	0.3	1.2	Standard	6.004	0.336
37	0.25	1.2	Standard	4.163	0.231
38	0.1	0.8	Wiper	0.297	0.025
39	0.3	0.8	Standard	5.626	0.204
40	0.3	0.4	Standard	11.252	0.421

3.2 Process analysis and modelling

The statistical analysis of the measurement results was carried out using the JMP software, whereby a model with main effects and two-factor interactions was selected. In this way, a model is obtained that allows the modelling of complex relationships between input and output parameters. Given the irregular distribution of the data and the excessive variation between values, a logarithmic transformation of Ra was performed to allow interpretation and statistical analysis. The estimated input parameters of the model, sorted by statistical significance, are shown in Table 2. As can be seen, all input parameters and the interaction of feed and corner radius are statistically significant. Other two-factor interactions are not significant.

Table 2 Effect summary

Source	Log Worth	P Value
$f(0.1,0.3)$	25.304	0.00000
r	20.530	0.00000
Insert	15.528	0.00000
$f \times r$	14.713	0.00000

Table 3 shows the summary of fit of the selected factor models of surface roughness and dimensional deviation. The predictions of the coefficient of determination (RSquare, RSquare Adj) are close to 1, which means that the model fits the data very well. In both cases, RSquare values (0.985099, 0.984298) are close to the RSquare Adj values (0.981253, 0.980246). The coefficients of determination show that more than 98 % of the variability in surface roughness and dimensional deviation is determined by the feed, corner radius, insert type and the interaction of feed and corner radius. The root mean square error is small compared to the mean of the response, indicating a good fit and accuracy of the prediction model, i.e. that the model fits the measured data well.

Table 4 shows the analysis of variance for the selected regression models. A high model F-value of 256.1723 for surface roughness and 242.9135 for dimensional deviation with a low p-value (< 0.0001) indicates that the model effects are statistically significant.

Table 5 shows the estimated regression coefficients of the parameters. The parameters provide an estimate of the effects of the model input parameters on the dimensional deviation and the logarithmically transformed value of the surface roughness. P-values (Prob > |t|) of less than 0.05 indicate that the model factors are significant (marked with * in Table 4).

Table 3 Summary of fit

Parameter	log <i>Ra</i>	ΔD
RSquare	0.985099	0.984298
RSquare Adj	0.981253	0.980246
Root Mean Square Error	0.143263	0.029879
Mean of Response	0.957479	0.234625

Table 4 Analysis of variance

Source	DF	log <i>Ra</i>			ΔD		
		Sum of Squares	Mean Square	F Ratio	Sum of Squares	Mean Square	F Ratio
Model	8	42.061990	5.25775	256.1723	1.7348960	0.216862	242.9135
Error	31	0.636252	0.02052	Prob > F	0.0276754	0.000893	Prob > F
C. Total	39	42.698242		<.0001*	1.7625714		<.0001*

Table 5 Parameter estimates

Term	log <i>Ra</i>				ΔD			
	Parameter Estimate	Std. Error	t Ratio	Prob> t	Parameter Estimate	Std. Error	t Ratio	Prob> t
Intercept	0.9574789	0.022652	42.27	<.0001*	0.234625	0.004724	49.66	<.0001*
<i>f</i> (0.1,0.3)	1.0829312	0.032035	33.81	<.0001*	0.209	0.006681	31.28	<.0001*
<i>r</i> [0.2]	0.883553	0.039234	22.52	<.0001*	0.203775	0.008183	24.90	<.0001*
<i>r</i> [0.4]	0.1647571	0.039234	4.20	0.0002*	-0.016425	0.008183	-2.01	0.0535
<i>r</i> [0.8]	-0.552802	0.039234	-14.09	<.0001*	-0.127025	0.008183	-15.52	<.0001*
<i>r</i> [1.2]	-0.495508	0.039234	-12.63	<.0001*	-0.060325	0.008183	-7.37	<.0001*
Insert [Standard]	0.3541856	0.022652	15.64	<.0001*	-0.008125	0.004724	-1.72	0.0954
Insert [Wiper]	-0.354186	0.022652	-15.64	<.0001*	0.008125	0.004724	1.72	0.0954
<i>f</i> × <i>r</i> [0.2]	6.3579e-5	0.055485	0.00	0.9991	0.1811	0.011572	15.65	<.0001*
<i>f</i> × <i>r</i> [0.4]	0.0001446	0.055485	0.00	0.9979	-0.0156	0.011572	-1.35	0.1874
<i>f</i> × <i>r</i> [0.8]	0.000214	0.055485	0.00	0.9969	-0.1132	0.011572	-9.78	<.0001*
<i>f</i> × <i>r</i> [1.2]	-0.000422	0.055485	-0.01	0.9940	-0.0523	0.011572	-4.52	<.0001*

Feed, corner radius and insert type have statistical significance for surface roughness. Feed has the highest statistical significance, followed by corner radius and then insert type. The interactions between the factors have no statistical significance for surface roughness.

Feed, three levels of corner radius (0.2, 0.4 and 1.2) and their interactions have statistical significance for the dimensional deviation. A corner radius of 0.4 mm is neither statistically significant independently nor in interaction with any other parameter. The insert type has no statistical significance either independently or in interaction with other parameters.

Fig. 2 illustrates the relationship between the actual and predicted values. The resulting plots demonstrate that the data points closely align with the line, indicating a strong fit of the model. This suggests that the predicted values are accurate and closely match the actual values for both surface roughness and dimensional deviation. The narrow range of the confidence interval further supports the precision of the predicted values.

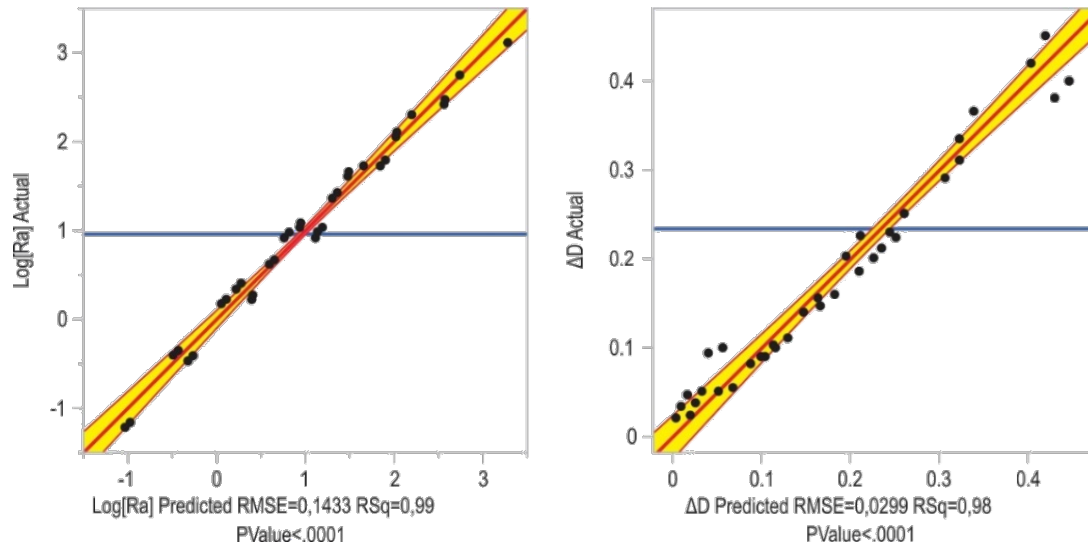


Fig. 2 Actual by predicted plot

Fig. 3 shows the dependence of the residuals on the predicted value. It can be seen that the residuals are independent in both cases, i.e. they do not correlate with a parameter, which indicates that there was no systematic increase or decrease in the residuals during the experiment.

Fig. 4 shows that the studentized residuals are not outside the interval and are randomly distributed around the value zero, which not only confirms the normal distribution, but also means that there are no significantly deviating values, i.e. no potential outliers.

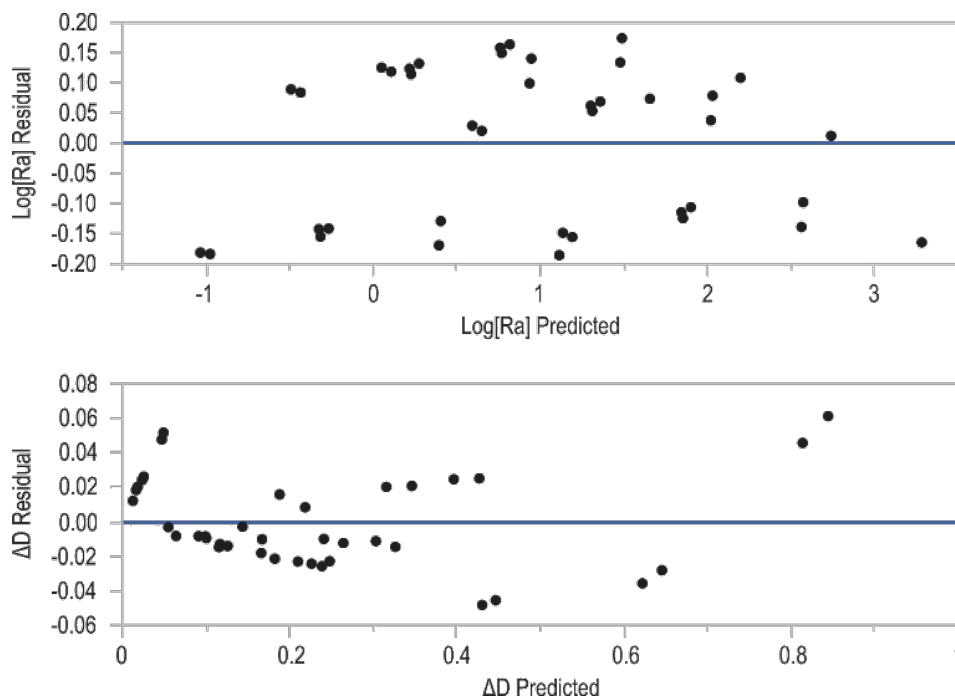


Fig. 3 Residual by predicted plot

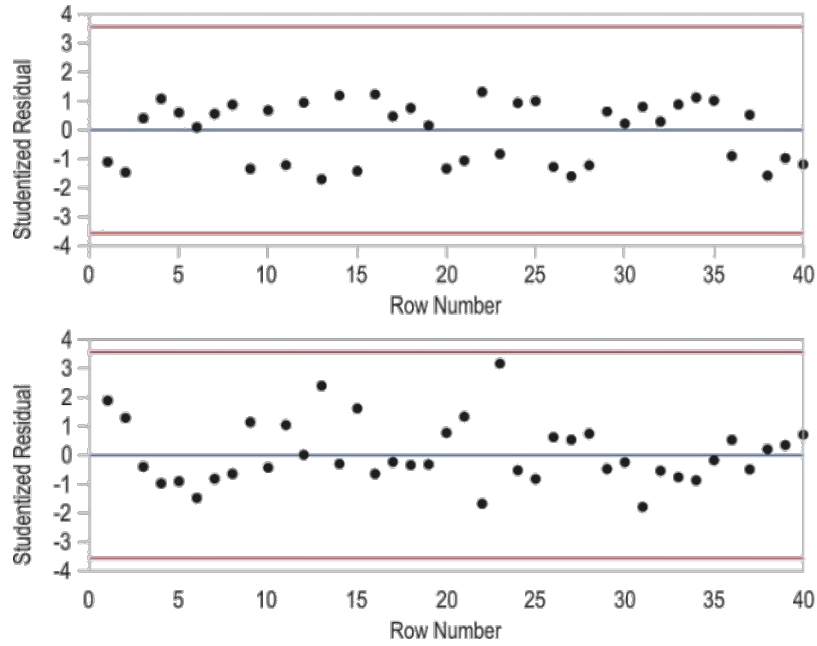


Fig. 4 Studentized residuals

For the selected models, the statistical analysis of the measured data yields the following regression equations for roughness and dimensional deviation:

$$\log Ra = 0.9574789383 + 1.0829311516 \cdot \left(\frac{f-0.2}{0.1}\right) + Match(r) \begin{pmatrix} "0.2" & \Rightarrow & 0.8835530387 \\ "0.4" & \Rightarrow & 0.1647570702 \\ "0.8" & \Rightarrow & -0.552802394 \\ "1.2" & \Rightarrow & -0.495507715 \\ else & \Rightarrow & . \end{pmatrix} +$$

$$\left(\frac{f-0.2}{0.1}\right) \cdot Match(r) \begin{pmatrix} "0.2" & \Rightarrow & 0.0000635788 \\ "0.4" & \Rightarrow & 0.0001445986 \\ "0.8" & \Rightarrow & 0.0002140391 \\ "1.2" & \Rightarrow & -0.000422217 \\ else & \Rightarrow & . \end{pmatrix} +$$

$$Match(Insert) \begin{pmatrix} "Standard" & \Rightarrow & 0.3541856229 \\ "Wiper" & \Rightarrow & -0.3541856229 \\ else & \Rightarrow & . \end{pmatrix} \quad (3)$$

$$\Delta D = 0.234625 + 0.209 \cdot \left(\frac{f-0.2}{0.1}\right) + Match(r) \begin{pmatrix} "0.2" & \Rightarrow & 0.203775 \\ "0.4" & \Rightarrow & -0.015425 \\ "0.8" & \Rightarrow & -0.127025 \\ "1.2" & \Rightarrow & -0.060325 \\ else & \Rightarrow & . \end{pmatrix} + \left(\frac{f-0.2}{0.1}\right) \cdot$$

$$Match(r) \begin{pmatrix} "0.2" & \Rightarrow & 0.1811 \\ "0.4" & \Rightarrow & -0.0156 \\ "0.8" & \Rightarrow & -0.1132 \\ "1.2" & \Rightarrow & -0.0523 \\ else & \Rightarrow & . \end{pmatrix} + Match(Insert) \begin{pmatrix} "Standard" & \Rightarrow & -0.008125 \\ "Wiper" & \Rightarrow & 0.008125 \\ else & \Rightarrow & . \end{pmatrix} \quad (4)$$

The graphical interpretation of the influence of the input parameters on the output parameters is illustrated by surface plots in Fig. 5.

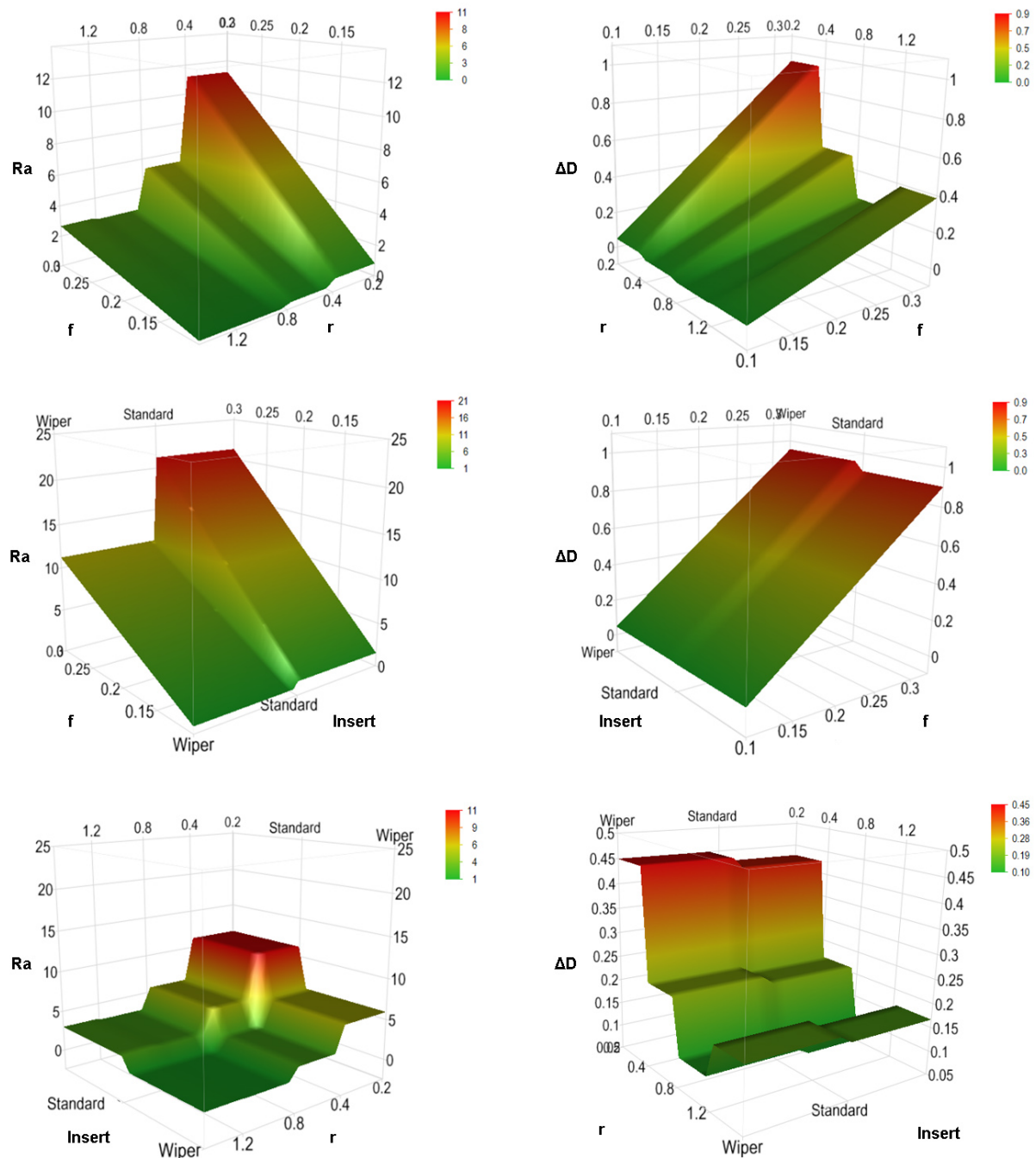


Fig. 5 Surface plot

3.3. Process optimization

After modelling, an optimization is carried out. The optimization problem was solved using the Wolfe reduced-gradient approach in the JMP software. The constraints are set in accordance with the conditions of the experimental research. The objective function is the simultaneous minimization of the surface roughness and the dimensional deviation. Since these two requirements must be met together in practice, both output parameters are given the same importance (same weighting coefficients).

The quantification of the optimization can be illustrated most easily with the profile plot of the prediction profiles (Fig. 6). The profiler plot shows the prediction of the optimal input variables. The optimal predicted value of the surface roughness is $\log Ra = -1.03265 \mu\text{m}$, i.e. $Ra \approx 0.356 \mu\text{m}$. The optimal predicted value of the measurement deviation is $\Delta D \approx 0.02 \text{ mm}$. The optimum predicted values can be determined with a confidence interval of 95 % for the following optimum

input parameters $f = 0.1$ mm/rev, $r = 0.8$ and wiper insert. The desirability function shows how the desirability value changes when the input parameters change, allowing quick (visual) identification of the combination of input parameters that best meet the set optimization goal. The desirability index provides a comprehensive assessment of the effectiveness of the solution obtained, taking into account the optimization goal and its importance. A high value of the overall desirability measure (close to 1) indicates that the optimal values of the parameters were achieved in accordance with the set optimization goals. This means that the values achieved are satisfactory in terms of product or process quality.

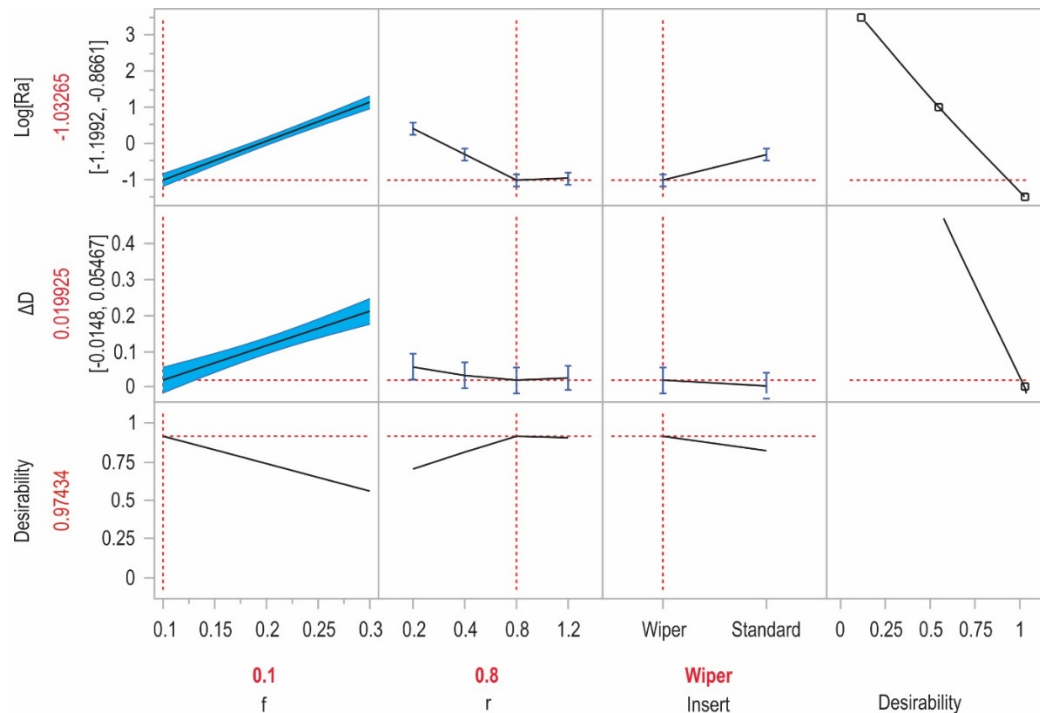


Fig. 6 Prediction profiler

3.4. Confirmation experiments

The regression models obtained and the optimum values for surface roughness and dimensional deviation were confirmed by a further 9 tests.

The confirmation of the regression models obtained was carried out with unknown combinations of input parameters in 8 trials. Since corner radius and insert type are categorical values, they were not varied. The feed was varied in 4 levels (0.125, 0.175, 0.225, 0.275 mm/rev).

The optimum values for surface roughness and dimensional deviation were confirmed at the theoretical optimum for the parameters $f = 0.1$ mm/rev, $r = 0.8$ and wiper insert. The results obtained together with the calculated percentage errors are shown in Table 6.

Table 6 The results of the confirmation experiments

No.	f (mm/rev)	r (mm)	Insert	Measurement		Prediction		Percentage error	
				Ra (μm)	ΔD (mm)	Ra (μm)	ΔD (mm)	PE_{Ra} (%)	$PE_{\Delta D}$ (%)
1	0.125	0.2	Standard	3.994	0.143	3.971	0.138	0.58	3.50
2	0.125	0.4	Wiper	0.969	0.086	0.956	0.082	1.34	4.65
3	0.175	0.2	Wiper	3.389	0.354	3.374	0.349	0.44	1.41
4	0.175	0.4	Standard	3.358	0.170	3.339	0.162	0.57	4.71
5	0.225	0.8	Wiper	1.395	0.148	1.379	0.141	1.15	4.73
6	0.225	1.2	Standard	2.976	0.211	2.965	0.205	0.37	2.84
7	0.275	0.8	Standard	4.824	0.177	4.811	0.172	0.27	2.82
8	0.275	1.2	Wiper	2.533	0.306	2.509	0.301	0.95	1.63
9	0.1	0.8	Wiper	0.365	0.021	0.356	0.020	2.47	4.76

4. Discussion

The results of experimental research show that the obtained output parameters achieved within very wide range. For surface roughness, the results vary in a range of 0.297-22.503 μm and with a ratio between the maximum and minimum values of 75.77 (7577 %). For dimensional deviation, the results vary in a range of 0.022-0.905 mm and a ratio between the maximum and minimum values of 41.14 (4134 %). High ratio values indicate the possibility of controlling the output parameters within very wide limits depending on the combination of input parameters. In addition, the minimum values of surface roughness and dimensional deviation show that it is possible to achieve high surface quality and high dimensional accuracy in dry turning if the machining parameters are optimized.

The results of the statistical analysis of the experimental results show that the regression models are appropriate and significant, which means that they describe the process of dry turning of Inconel 625 alloy well. The statistical analysis made it possible to identify significant factors and their interactions on the output parameters of the process.

Based on the estimated factors, the regression equations and the surface plots, it can be observed that the value of surface roughness increases with increasing feed and decreasing corner radius. The reverse is also true. At higher feeds and smaller corner radii, the peaks and valleys formed on the workpiece are deeper and wider. The theoretical geometric surface roughness is a function of the feed for a specific corner radius. Smaller corner radii lead to a poorer quality of the machined surface at the same feed. This is due to the different values of the approach angle along the cutting edge. Increasing the corner radius at the same feed improves the quality of the machined surface (up to the limit value of 0.8). A further increase in the corner radius (after the value of 0.8 mm) of the cutting insert affected the likely occurrence of chatter and vibration [33], leading to a deterioration in surface roughness and dimensional accuracy (at a radius of 1.2 mm). This conclusion is also confirmed by the flank wear, which is close to the smallest and largest value of the dimensional deviation, which was to be expected considering the short cutting time and the identically smaller value of the cutting speed, which in previous studies proved to be the most influential parameter for intensifying the wear mechanisms of the cutting insert.

Larger values of the feed and smaller values of the corner radius, individually or in interaction, form an uneven surface, resulting in a significant deviation from the circularity and cylindricity, and the waviness of the surface increases, worsening the dimensional accuracy. Similar to the surface roughness, after reaching the critical value of the corner radius of 0.8 mm, there is a slight deterioration in dimensional accuracy due to the likely occurrence of vibration. In addition to the deterioration in surface roughness, the induced vibrations also influenced the deterioration in dimensional accuracy.

Furthermore, the results showed that there is a significant interaction between feed and corner radius on dimensional deviation (Fig. 7). When the feed increases, the dimensional accuracy deteriorates. With an increase in the corner radius, the dimensional accuracy initially improves significantly and then deteriorates slightly. The smallest dimensional deviations are achieved with a corner radius of 0.8 mm. For each combination of feed and corner radius, the standard insert achieves a slightly better dimensional accuracy. The interaction of feed and corner radius follows the trends for feed and corner radius. The influence of the interaction is more pronounced for larger values of the feed and smaller values of the corner radius. This means that a smaller corner radius (0.2 mm) in combination with an increase in the feed has a linear (worse) effect on the dimensional accuracy. The smallest effect of this interaction on the dimensional accuracy is given at the smallest feed (0.1 mm/rev) and at a corner radius of 0.8 mm.

The insert type showed a significant effect on the surface roughness. Wiper inserts are designed to improve the quality of the machined surface. Therefore, for the same feed and corner radius, wiper inserts always produce a better quality turned surface, i.e. a lower surface roughness value. The insert type had no significant influence on the realisation of the dimension, i.e. on the dimensional deviation, therefore in machining operations where a certain surface roughness is not required, it is more rational to use standard inserts as they are cheaper.

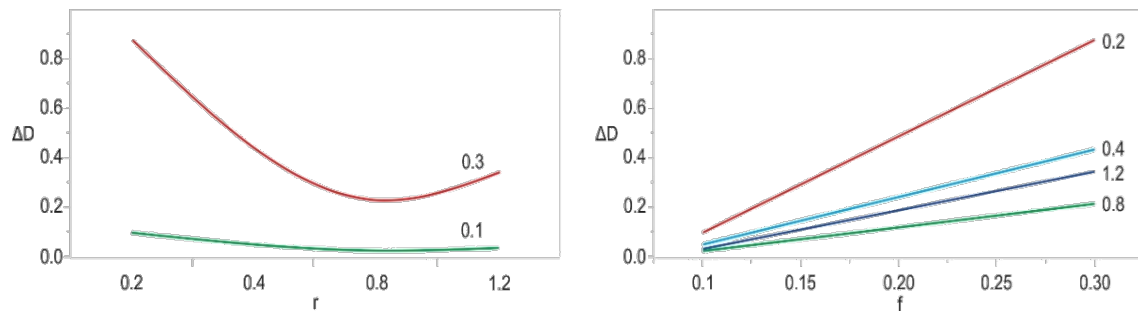


Fig. 7 Interaction profiles

The results of the confirmation experiments have shown that dry turning of Inconel 625 alloy is modelled sufficiently accurately for the given input and output parameters, as the percentage errors are in the range of 0.27-2.47 % for the surface roughness, i.e. in the range of 1.41-4.76 % for the measurement deviation.

5. Conclusion

The results of the experimental investigations and the statistical analysis of the measurement results have shown that there is a mutual functional dependency between the feed, the corner radius and the insert type on the one hand and the dimensional deviation and surface roughness on the other.

The prediction and validation of the created regression models demonstrated that it is possible to model both surface roughness and dimensional deviation.

The optimization results further confirmed that it is possible to optimize the turning process by considering the combined influence of feed, corner radius, and insert type on both surface roughness and dimensional deviation.

As the corner radius increases up to a certain value, both dimensional deviation and surface roughness improve, but after reaching a threshold, they begin to deteriorate slightly. As the feed increases, both dimensional deviation and surface roughness increase. The type of cutting insert affects surface roughness but does not influence dimensional deviation.

The percentage errors achieved, which are less than 5 % in the worst case, show that the process is correctly modelled and that the regression models created can be practically applied in a real production environment.

For future research, it is planned to include further input and output parameters in the prediction and optimization model.

Acknowledgement

This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-65/2024-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project "Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad" (No. 01-3394/1).

References

- [1] Tomov, M., Gecevska, V., Vasileska, E. (2022). Modelling of multiple surface roughness parameters during hard turning: A comparative study between the kinematical-geometrical copying approach and the design of experiments method (DOE), *Advances in Production Engineering & Management*, Vol. 17, No. 1, 75-88, [doi: 10.14743/apem2022.1.422](https://doi.org/10.14743/apem2022.1.422).
- [2] Vukelic, D., Prica, M., Ivanov, V., Jovicic, G., Budak, I., Luzanin, O. (2022). Optimization of surface roughness based on turning parameters and insert geometry, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 417-428, [doi: 10.2507/IJSIMM21-3-607](https://doi.org/10.2507/IJSIMM21-3-607).
- [3] Kang, W.T., Derani, M.N., Ratnam, M.M. (2020). Effect of vibration on surface roughness in finish turning: Simulation study, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 595-606, [doi: 10.2507/IJSIMM19-4-531](https://doi.org/10.2507/IJSIMM19-4-531).

- [4] Šarić, T., Vukelić, D., Šimunović, K., Svalina, I., Tadić, B., Prica, M., Šimunović, G. (2020). Modelling and prediction of surface roughness in CNC turning process using neural networks, *Tehnički Vjesnik – Technical Gazette*, Vol. 27, No. 6, 1923-1930, doi: [10.17559/TV-20200818114207](https://doi.org/10.17559/TV-20200818114207).
- [5] Vukelic, D., Simunovic, K., Kanovic, Z., Saric, T., Doroslovacki, K., Prica, M., Simunovic, G. (2022). Modelling surface roughness in finish turning as a function of cutting tool geometry using the response surface method, Gaussian process regression and decision tree regression, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 367-380, doi: [10.14743/apem2022.3.442](https://doi.org/10.14743/apem2022.3.442).
- [6] Qadri, S.I.A., Harmain, G.A., Wani, M.F. (2020). Influence of tool tip temperature on crater wear of ceramic inserts during turning process of Inconel-718 at varying hardness, *Tribology in Industry*, Vol. 42, No. 2, 310-326, doi: [10.24874/ti.776.10.19.05](https://doi.org/10.24874/ti.776.10.19.05).
- [7] Kushwaha, A., Singh, R. (2017). Optimization of CNC process parameters on Inconel 625 using response surface methodology, *International Journal of Mechanical Engineering and Technology*, Vol. 8, No. 7, 1830-1836.
- [8] Kosaraju, S., Vijay Kumar, M., Sateesh, N. (2018). Optimization of machining parameter in turning Inconel 625, *Materials Today: Proceedings*, Vol. 5, No. 2, 5343-5348, doi: [10.1016/j.matpr.2017.12.119](https://doi.org/10.1016/j.matpr.2017.12.119).
- [9] Jeykrishnan, J., Sai, M.M., Bareeth, M.S., Babu, B.S.M., Prikeshwar, S., Shrinivas, V.A. (2018). Parametric analysis of turning process characteristics in Inconel 625 alloy using traditional techniques, *International Journal of Mechanical Engineering and Technology*, Vol. 9, No. 8, 386-390.
- [10] Dhananchezian, M. (2019). Effectiveness of cryogenic cooling in turning of Inconel 625 alloy, In: Vijay Sekar, K., Gupta, M., Arockiarajan, A. (eds.), *Advances in Manufacturing Processes, Lecture Notes in Mechanical Engineering*, Springer, Singapore, 591-597, doi: [10.1007/978-981-13-1724-8_54](https://doi.org/10.1007/978-981-13-1724-8_54).
- [11] Magri, A., Diniz, A.E., Suyama, D.I. (2018). Evaluating the use of high-pressure coolant in turning process of Inconel 625 nickel-based alloy, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, Vol. 232, No. 7, 1182-1192, doi: [10.1177/0954405416664373](https://doi.org/10.1177/0954405416664373).
- [12] Singh, P., Padhy, C.P. (2019). Influence of nano (h-BN) cutting fluid on machinability of Inconel 625, *Journal of Physics: Conference Series*, Vol. 1355, Article No. 012033, doi: [10.1088/1742-6596/1355/1/012033](https://doi.org/10.1088/1742-6596/1355/1/012033).
- [13] Yildirim, C.V. (2019). Experimental comparison of the performance of nanofluids, cryogenic and hybrid cooling in turning of Inconel 625, *Tribology International*, Vol. 137, 366-378, doi: [10.1016/j.triboint.2019.05.014](https://doi.org/10.1016/j.triboint.2019.05.014).
- [14] Yildirim, C.V., Sarıkaya, M., Kivak, T., Sirin, S. (2019). The effect of addition of hBN nanoparticles to nanofluid-MQL on tool wear patterns, tool life, roughness and temperature in turning of Ni-based Inconel 625, *Tribology International*, Vol. 134, 443-456, doi: [10.1016/j.triboint.2019.02.027](https://doi.org/10.1016/j.triboint.2019.02.027).
- [15] Singh, P., Padhy, C. (2021). An improvised method of machinability evaluation with predictive temperature model for Inconel 625: A holistic machinist's perspective, *International Journal of Automotive and Mechanical Engineering*, Vol. 18, No. 2, 8802-8813.
- [16] Yagmur, S. (2021). The effects of cooling applications on tool life, surface quality, cutting forces, and cutting zone temperature in turning of Ni-based Inconel 625, *The International Journal of Advanced Manufacturing Technology*, Vol. 116, No. 3-4, 821-833, doi: [10.1007/s00170-021-07489-2](https://doi.org/10.1007/s00170-021-07489-2).
- [17] Padhy, C., Singh, P. (2022). Estimating machinability factors in turning of Inconel 625 under different lubricating conditions using 3D DEFORM FE analysis, *INCAS Bulletin*, Vol. 14, No. 1, 97-107, doi: [10.13111/2066-8201.2022.14.1.8](https://doi.org/10.13111/2066-8201.2022.14.1.8).
- [18] Rakesh, P.R., Chakradhar, D. (2023). Machining performance comparison of Inconel 625 superalloy under sustainable machining environments, *Journal of Manufacturing Processes*, Vol. 85, 742-755, doi: [10.1016/j.jmapro.2022.11.080](https://doi.org/10.1016/j.jmapro.2022.11.080).
- [19] Makhesana, M.A., Patel, K.M., Krolczyk, G.M., Danish, M., Singla, A.K., Khanna, N. (2023). Influence of MoS2 and graphite-reinforced nanofluid-MQL on surface roughness, tool wear, cutting temperature and microhardness in machining of Inconel 625, *CIRP Journal of Manufacturing Science and Technology*, Vol. 41, 225-238, doi: [10.1016/j.cirpj.2022.12.015](https://doi.org/10.1016/j.cirpj.2022.12.015).
- [20] Marimuthu, P., Baskaran, R. (2014). Optimal setting of machining parameters for turning Inconel 625 using coated tool, *Applied Mechanics and Materials*, Vol. 573, 632-637, doi: [10.4028/www.scientific.net/AMM.573.632](https://doi.org/10.4028/www.scientific.net/AMM.573.632).
- [21] Venkatesan, K., Ramanujam, R., Saxena, V., Chawdhury, N., Choudhary, V. (2014). Influence of cutting parameters on dry machining of Inconel 625 alloy with coated carbide insert - A statistical approach, *ARPJ Journal of Engineering and Applied Sciences*, Vol. 9, No. 3, 250-258.
- [22] Ramanujam, R., Venkatesan, K., Saxena, V., Pandey, R., Harsha, T., Kumar, G. (2014). Optimization of machining parameters using fuzzy based principal component analysis during dry turning operation of Inconel 625 - A hybrid approach, *Procedia Engineering*, Vol. 97, 668-676, doi: [10.1016/j.proeng.2014.12.296](https://doi.org/10.1016/j.proeng.2014.12.296).
- [23] Jain, H., Tripathi, J., Bharilya, R., Jain, S., Kumar, A. (2015). Optimisation and evaluation of machining parameters for turning operation of Inconel-625, *Materials Today: Proceedings*, Vol. 2, No. 4-5, 2306-2313, doi: [10.1016/j.matpr.2015.07.273](https://doi.org/10.1016/j.matpr.2015.07.273).
- [24] Prokes, T., Muralova, K., Kovar, J. (2016). Experimental evaluation on the quality of machined surface after turning of material Inconel 625, *MM Science Journal*, 1617-1620, doi: [10.17973/MMSJ.2016_12_2016197](https://doi.org/10.17973/MMSJ.2016_12_2016197).
- [25] Lotfi, M., Jahanbakhsh, M., Akhavan Farid, A. (2016). Wear estimation of ceramic and coated carbide tools in turning of Inconel 625: 3D FE analysis, *Tribology International*, Vol. 99, 107-116, doi: [10.1016/j.triboint.2016.03.008](https://doi.org/10.1016/j.triboint.2016.03.008).
- [26] Hemakumar, S., Kuppan, P. (2019). Experimental investigations and optimisation of process parameters in dry finish turning of Inconel 625 super alloy, *International Journal of Materials and Product Technology*, Vol. 59, No. 4, 303-320, doi: [10.1504/IJMPT.2019.104583](https://doi.org/10.1504/IJMPT.2019.104583).
- [27] Vasudevan, H., Rajguru, R., Shaikh, M., Shaikh, A. (2019). Optimization of process parameters in the turning operation of Inconel 625, *Materials Science Forum*, Vol. 969, 756-761, doi: [10.4028/www.scientific.net/MSF.969.756](https://doi.org/10.4028/www.scientific.net/MSF.969.756).

- [28] Waghmode, S.P., Dabade, U.A. (2019). Optimization of process parameters during turning of Inconel 625, *Materials Today: Proceedings*, Vol. 19, Part 2, 823-826, [doi: 10.1016/j.matpr.2019.08.138](https://doi.org/10.1016/j.matpr.2019.08.138).
- [29] Padhy, C., Singh, P. (2020). Use of multi-objective optimization technique (Taguchi-GRA approach) in dry hard turning of Inconel 625, *INCAS Bulletin*, Vol. 12, No. 2, 133-142, [doi: 10.13111/2066-8201.2020.12.2.11](https://doi.org/10.13111/2066-8201.2020.12.2.11).
- [30] Narkhede, G.B., Chinchani, S., Vadgeri, S.S., Rathod, B.S. (2020). Comparative evaluation of machining performance of Inconel 625 under dry and cryogenic cutting conditions, *IOP Conference Series: Materials Science and Engineering*, Vol. 810, No. 1, Article No. 012036, [doi: 10.1088/1757-899X/810/1/012036](https://doi.org/10.1088/1757-899X/810/1/012036).
- [31] Padhy, C., Singh, P. (2021). Optimization of machining parameters using Taguchi coupled grey relational approach while turning Inconel 625, *Journal of Mechanical Engineering*, Vol. 18 No. 2, 161-176, [doi: 10.24191/jmeche.v18i2.15151](https://doi.org/10.24191/jmeche.v18i2.15151).
- [32] Sim, C.Y., Reddy, M.M., Sethuramalingam, P. (2022). Fuzzy logic expert system based machinability simulation analysis on nickel based alloy steel, *International Journal on Interactive Design and Manufacturing*, Vol. 16, No. 3, 1063-1077 [doi: 10.1007/s12008-021-00804-x](https://doi.org/10.1007/s12008-021-00804-x).
- [33] Kang, W.T., Derani, M.N., Ratnam, M.M. (2020). Effect of vibration on surface roughness in finish turning: Simulation study, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 595-606, [doi: 10.2507/IJSIMM19-4-531](https://doi.org/10.2507/IJSIMM19-4-531).

Enhancing calibration accuracy with laser interferometry for high-resolution measuring systems

Lipus, L.C.^a, Acko, B.^{a,*}, Klobucar, R.^a

^aUniversity of Maribor, Faculty of Mechanical Engineering, Slovenia

ABSTRACT

This paper presents an analysis of the measurement capability of laser interferometry for calibrating high-resolution measuring systems, focusing on potential errors that need to be carefully controlled to ensure adequate metrological traceability. The primary scientific research focus of our National Dimensional Metrology Laboratory is the development of metrological applications for industry, with ongoing improvements in calibration procedures for measuring machines and tools. Through a review of possible error sources, an innovative approach to reducing the most significant factors is proposed. This is specifically applied to the following calibration cases: field calibration of coordinate measuring machines, and laboratory calibration of precision probes and line scales. Instrumental and environmental errors can be effectively mitigated by using periodically calibrated laser interferometers in well-controlled air conditions, ensuring an uncertainty of 0.2 $\mu\text{m}/\text{m}$. Most geometrical errors can be minimized by precise adjustments to the interferometry and positioning systems, achieving an uncertainty of 0.3 $\mu\text{m}/\text{m}$. However, errors caused by temperature differences in the material along the measuring path remain the most influential. These arise due to the high expansion coefficient of the material and some uncertainty in its properties. After several hours of temperature stabilization, using three temperature sensors along the displacement range for software compensation, temperature differences still contribute significantly to measurement uncertainty. For example, the error is 0.5 $\mu\text{m}/\text{min}$ in the case of line scale calibration and 1.1 $\mu\text{m}/\text{m}$ for coordinate measuring machine calibration.

ARTICLE INFO

Keywords:

Laser interferometry;
High resolution measurements;
Measurement uncertainty;
Dimensional metrology;
Coordinate measuring machine;
Precise probe;
Line scale

*Corresponding author:

bojan.acko@um.si
(Acko, B.)

Article history:

Received 24 July 2024
Revised 26 October 2024
Accepted 28 October 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

A wide range of verification methods, based on various measurement standards, is used for calibrating and determining geometric errors in high-resolution measuring systems. Among these, laser-based methods are the most accurate. Precision measuring machines used in industry typically have an uncertainty of around μm per meter in the measuring range. Sub-micrometer accurate measures, such as gauge blocks, setting rings, cylinders, step gauges, and line scales, are commonly employed to ensure their measurement traceability. Displacement lasers, known for their exceptional accuracy, often exceed manufacturing tolerance requirements. They are used in situations where more precise and detailed information about the properties of a metrological system is needed, such as in advanced monitoring of workpiece geometry in the automotive industry [1-3].

Our National Dimensional Metrology Laboratory is constantly improving the capability of our calibration and measurement procedures. The core of our research is characterization of current laboratory equipment to minimize errors, and development of new measurement procedures for

calibrating highly precise instruments and machines. Primary task in the field of scientific and Industry metrology of the laboratory is maintaining the national standard for length and assuring measurements traceability in Slovenian industry. For this purpose, the laboratory is accredited for length-calibrations of standards and measuring equipment by Slovenian Accreditation (SA). Its metrology capabilities are published in the database of national metrology institutes at the International Bureau of Weights and Measures (KCDB BIPM). The main focus of scientific research work of the laboratory is development of metrological applications for practical use in industry, science and other branches. We are also performing activities for customers in the field of industry and science metrology, as well as education for calibrations.

To ensure measurement traceability, all potential error sources must be analyzed to identify and reduce the most influential factors in real-world environmental conditions, thereby achieving proper measurement accuracy. In a laser interferometric system, laser interferometers and corresponding air sensors should be periodically calibrated, while the entire instrumental system can be directly checked through comparison measurements with a reference laser interferometer [4]. This paper presents several metrological laser interferometry (LI) systems with innovative adjustments for error mitigation, developed in our laboratory, and summarizes the uncertainty budget, which is generally applicable to other similar high-precision metrological systems.

2. Measurement system and adjustments

Measurement systems with laser interferometer (LI) provide very precise position or distance information. It consists of a laser head, which produces the beam of highly stabilized, low-power light, and a variety of optical components and accessories such as air and material sensors. Laser interferometry is based on the principle of Michelson interferometer: a beam splitter divides the light beam into two beams, the reference beam is returned by a fixed reflector and another one returns from the retroreflector which is shifted along the measured path. Combined by the splitter, they interfere into pulses, which are counted by a photo-electro detector inside the laser head (Fig. 1).

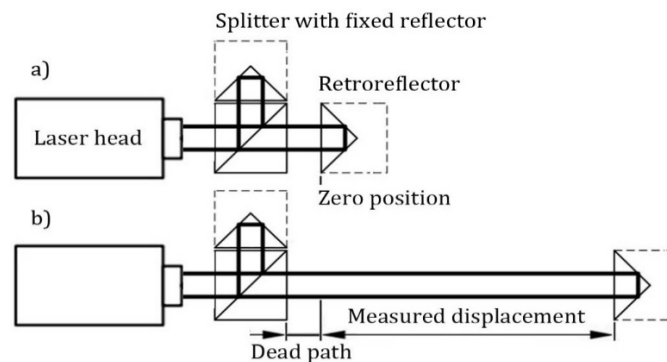


Fig. 1 The principle of Michelson interferometer

A calibration procedure generally comprehends the following basic steps:

- Installation and precise alignment of laser head and optics. Special mounting elements are constructed for the optics used for specific calibrations.
- The air conditions shall be stabilized for at least 5 hours. Three material temperature sensors are usual used for material and one air sensor to observe temperature, pressure and humidity along the measuring path.
- Deviations are then measured in eleven points (zero point and ten equal increments along the whole measuring length, or more by a customer's request), repeated three times in each point.

Generally, there are three categories of potential measurement errors: geometrical, environmental, and instrumental. The methods upgraded in our laboratory to mitigate these errors are presented in the following subsections.

2.1 Geometrical errors

The most of geometrical errors can be eliminated by applying a precise alignment procedure. The cosine error occurs when the beam from laser and the axis of the stage motion are not completely parallel (Eq. 1), while Abbe error occurs due to sloping of the retroreflector. Firstly, the beam is aligned by parallel shifts of the laser head into the centre of the target at close position, then by angle adjustments at maximum displacement.

$$e_{\cos} = L(\cos \varphi - 1) \approx -L\varphi^2/2 \quad (1)$$

where φ is angle between the beam and the axis of the moving stage, and L is the displacement.

When calibrating one-coordinate measuring machines (CCMs) of various types, the retroreflector is attached to the moving probe (Fig. 2), and the cosine error is eliminated by aligning the system at the maximum possible displacement. The LI system is used for calibration of CCMs with measuring length minimal 1 m, while CCMs with shorter measuring length are calibrated with gauge blocks. Assuming an adjustment at a minimum length of 1 m and visual centering of the beam within ± 1 mm, the maximum cosine error, as calculated by Eq. 1, is $0.5 \times 10^{-6} \times L$. The uncertainty is calculated by reduction with $\sqrt{3}$ at assumption of rectangular distribution, $u_{\cos} \approx 0.3 \times 10^{-6} \times L$.

Abbe error can be minimized within the resolution of CMM, which is minimally ± 5 nm for the best machine with digital display. The position of the optical elements and the moving parts are set in such way, that the measured object axis is set in the line with the center of the laser linear retroreflector. It allows the Abbe errors to be reduced to negligible levels even for the most demanding dimensional metrology tasks [5, 6]. Generally, for a maximal error 15 nm, the Abbe uncertainty is $u_a = 0.009$ nm. The best mitigation of the Abbe error can reduce it to just a few nanometers.

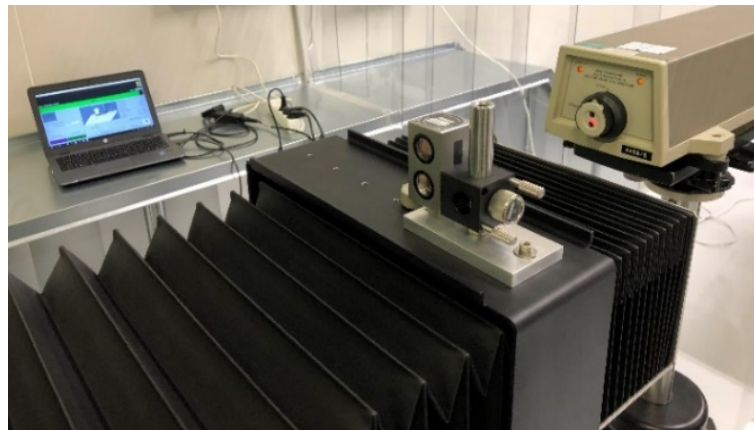


Fig. 2 Set-up for calibration of a one-coordinate measuring machine

When calibrating a precise probe, the retroreflector is installed below the probe (Fig. 3). The ceramic gauge is used as a base, and its surface imperfections are eliminated by setting the zero position. Precise probe is fixed, while the shift along its measurement range is performed by horizontal shifting the base with moving table, which has angle variation up to 15 micro-radians. When visual centering the probe onto retroreflector with offset ± 1 mm, the Abbe error is maximally ± 15 nm. The cosine error of the probe is eliminated by using dial indicator (Fig. 4). Within deviations $\pm 2 \mu\text{m}$ along the optic's frame 40 mm, the cosine error is negligible, $1.25 \times 10^{-9} \times L$, where L is probe's measurement range, commonly with resolution ± 5 nm. When calibrating the one-coordinate measuring machine, the cosine error of the laser head is adjusted horizontally along the moving table. By visual centering ± 1 mm at minimal distance 1 m, the maximal cosine error is achieved $0.5 \times 10^{-6} L$, and the uncertainty $u_{\cos} \approx 0.3 \times 10^{-6} \times L$.

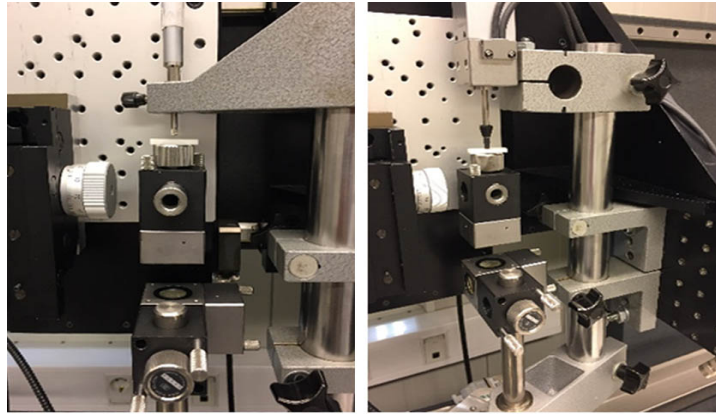


Fig. 3 Set-up for calibration of precise probe

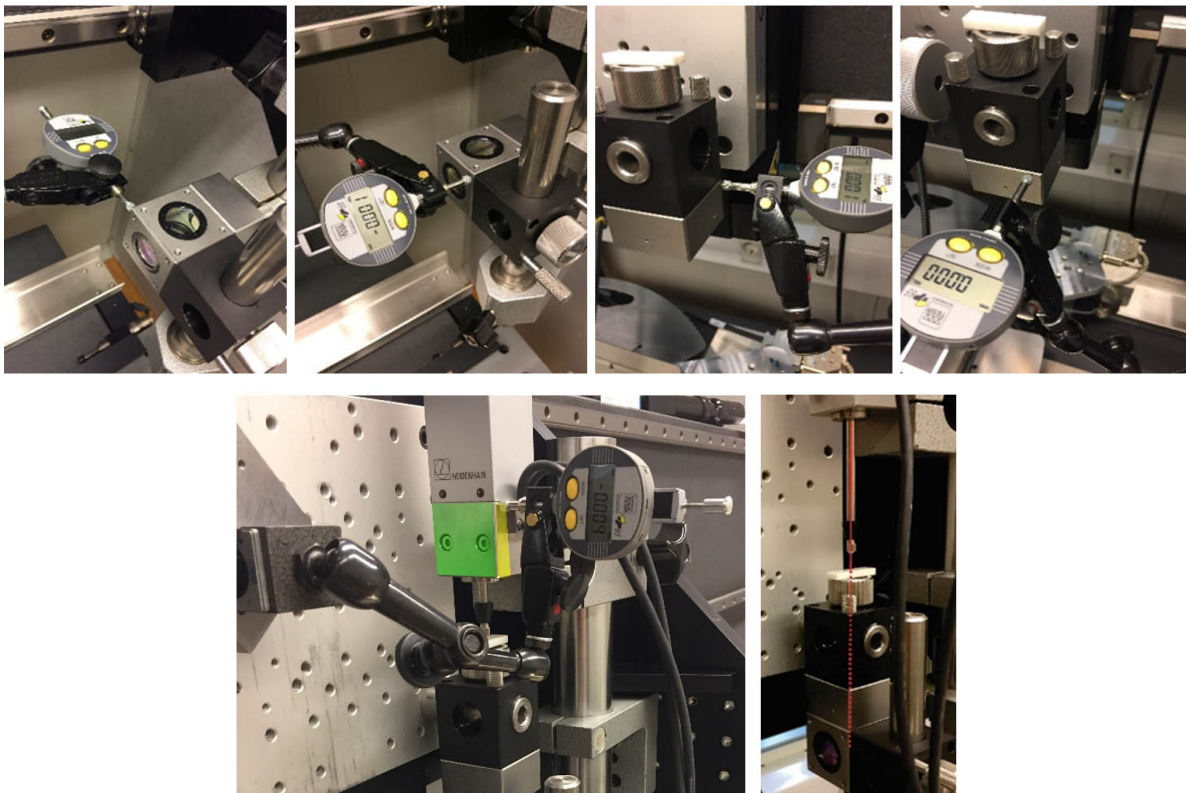


Fig. 4 Pitch and yaw measurements of the stage in y-direction

When calibrating a line scale, it is shifted using a movable table, and the centers of the line marks are precisely located by a video system with a high-resolution camera and software designed to detect the center of dark regions. A laser interferometer is then used to measure displacement from the zero position of the line scale, with a retroreflector fixed onto the moving table (Fig. 5). Similarly to the previous case, the beam is aligned along the moving table at maximal displacement and the camera's axis is adjusted by using dial indicator. The geometrical error of the moving table in the x- and z-directions (within the coordinate system that includes the optics, camera, and line scale, as shown in Fig. 6) was determined empirically [2]. For pitch, the largest angle difference along the 500 mm measurement path was $7 \mu\text{m}/\text{m}$. The maximum offset of the laser retroreflector in the z-direction was estimated to be 3 mm, resulting in an expected error interval for pitch of $7 \text{ nm}/\text{mm} \times 3 \text{ mm} = 21 \text{ nm}$. For yaw, the largest angle difference along the 500 mm measurement path was $10 \mu\text{m}/\text{m}$. The maximum offset of the laser retroreflector in the x-direction was estimated to be 1 mm, resulting in an expected error interval for yaw of $10 \text{ nm}/\text{mm} \times 1 \text{ mm} = 10 \text{ nm}$. Total error of table inclination, calculated by Pythagorean Theorem, is 23 nm, so the Abbe uncertainty is maximally $u_a = 0.013 \text{ nm}$.

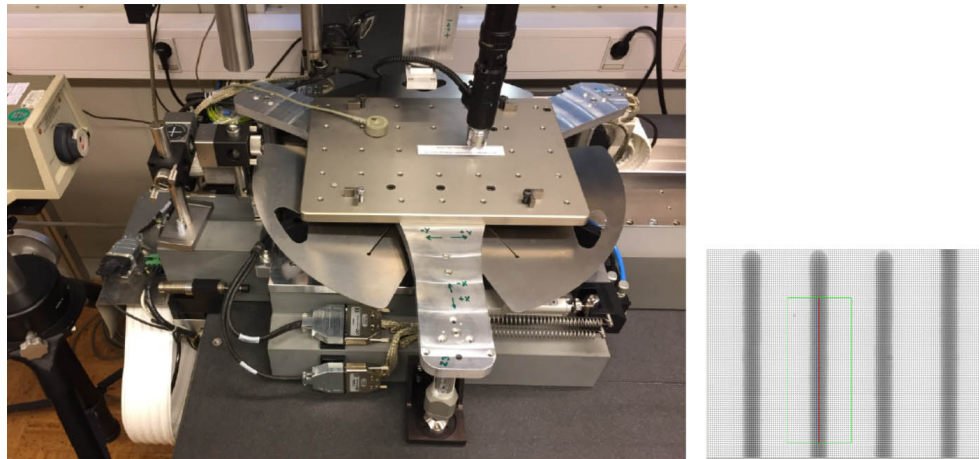


Fig. 5 Set-up for calibration of line scale with video detecting of line-mark position



Fig. 6 Alignment of optics, camera and the moving table with fixed line scale

2.2 Environmental errors

Environmental errors occur due to thermal properties of the machine under test and due to dependence of the light wavelength on the atmospheric conditions. Usually, three material temperature sensors are used along the measuring path, and one air sensor for measuring temperature, pressure and humidity. High-quality laser interferometers maintain a stable and repeatable light frequency, which is negligible compared to the uncertainty of the sensors. The LI software automatically compensates the material path length with the material sensors and the vacuum wavelength with air refractive index, which depends on the pressure p , temperature T and humidity H , and was empirically determined by Edlen's equation (with coefficients $-0.955 \times 10^{-6} / \text{K}$, $0.268 \times 10^{-6} / \text{hPa}$ and $-0.0085 \times 10^{-6} / \%$ of air humidity).

The compensation uncertainty depends on variation of air parameters along the measuring path and sensors' sensitivity. The sensitivity of common high-quality air sensors is around $0.15 \times 10^{-6} \times L$, including possible drift through years of application [3]. At common variations of air parameters in well-controlled conditions during field calibrations (0.5 K, 0.5 hPa, and 10 % humidity fluctuation) along the measuring path, and applying Edlen's coefficients along with a reduction by the square root of three for an assumed rectangular distribution [7], the total uncertainty for air compensation is $0.3 \times 10^{-6} \times L$. In well controlled conditions in laboratory chamber, where air temperature 20 °C varies just ± 0.1 K, the total uncertainty for air compensation is $0.2 \times 10^{-6} \times L$.

Dead path error is caused by an uncompensated length of the laser beam between the interferometer and the retroreflector, with the machine stage at zero position (Fig. 1). It gets negligible by placing the linear interferometer optics close to the zero point of the moveable retroreflector. At $L = 10$ mm in well controlled conditions, we get maximally 3 nm, which is practically negligible in comparison to the best resolution of the calibrated devices, described here. Also, the counting system of modern laser interferometers have negligible influence. Both contributes a variation of the displayed result L_{LI} on level of nm, usually eliminated by choosing LI resolution of $0.01 \mu\text{m}$. So,

the maximal permissible uncertainty of the indication of the laser interferometer, including air compensation, stability of light stability, counting system and dead path, is taken:

$$u_{LI} = 10 \text{ nm} + 0.3 \times 10^{-6} \times L \text{ at field conditions} \quad (2a)$$

$$u_{LI} = 10 \text{ nm} + 0.2 \times 10^{-6} \times L \text{ at conditions in chamber} \quad (2b)$$

where nominal length L is expressed in metres.

The software correction uncertainty strongly depends on the uncertainty of thermal expansion coefficient, α . For coordinate measuring machines, we take $\alpha = (11 \pm 1) \times 10^{-6}/\text{K}$, and for precise probes we take $\alpha = (8 \pm 1) \times 10^{-6}/\text{K}$. While in case of line scales made of glass, steel or other material, having quite different temperature expansion coefficients, in many cases not exactly known, we take $\alpha = (10 \pm 2) \times 10^{-6}/\text{K}$. Assuming rectangular distribution, their uncertainty u_α is calculated from deviations by reducing with $\sqrt{3}$. To provide minimal temperature deviations along the measuring path, calibrations are performed after temperature stabilization of the measuring machines, e.g. at least of 5 hours stabilization at field calibrations, and at least of 24 hour stabilization in the chamber. Temperature of the calibrated device is measured with three material temperature sensors, installed along the measuring path.

For our case of material temperature sensors, the uncertainty is maximally 0.015 K, evaluated from their calibration certificates (including expanded calibration uncertainty 0.015 K, reduced by 2, and maximal deviation 0.016 K and long term maximal drift 0.016 K, both reduced by $\sqrt{3}$). The temperature deviation of the device (from standard temperature 20 °C) is calculated as a mean value θ of the three measured deviations. Assuming maximal variation ± 0.3 K in field well controlled conditions, the uncertainty of the mean temperature is $0.3 \text{ K}/\sqrt{3 \times 3} = 0.1 \text{ K}$. The total standard uncertainty is less than $u_\theta = \sqrt{0.015^2 + 0.1^2} \text{ K} = 0.1 \text{ K}$, while for cases in the chamber with maximal temperature variation ± 0.1 K, we get $u_\theta = \sqrt{0.015^2 + 0.033^2} \text{ K} = 0.04 \text{ K}$ and take 0.05 K into calculation of total calibration uncertainty.

3. Results and discussion

The measurement accuracy of the specific calibration process is calculated from following mathematical model of the measured deviation:

$$e = L_i \cdot (1 + \alpha \cdot \theta) - (L_{LI} - e_g) \quad (3)$$

where parameters are:

L_i indicated value on the calibrated device

α thermal expansion coefficient of the measurement system of the device at 20 °C

θ temperature deviation of the measurement system of the device from 20 °C

L_{LI} length indicated by the laser interferometer

e_g geometrical error of the measurement system

Combined standard uncertainty is calculated according to GUM [7] by the equation:

$$u_e^2 = (c_i \cdot u_i)^2 + (c_\alpha \cdot u_\alpha)^2 + (c_\theta \cdot u_\theta)^2 + (c_{LI} \cdot u_{LI})^2 + (c_g \cdot u_g)^2 \quad (4)$$

where coefficients are partial derivatives of the expression (3):

$$c_i = \partial e / \partial L_i = 1 + \alpha_m \cdot \theta_m \approx 1 \text{ at well controlled temperature conditions}$$

$$c_\alpha = \partial e / \partial \alpha_m = \theta \cdot L_i \approx \theta \cdot L, \text{ where } L \text{ is measured length}$$

$$c_\theta = \partial e / \partial \theta_m = \alpha \cdot L_i \approx \alpha \cdot L$$

$$c_{LI} = \partial e / \partial L_{LI} = -1$$

$$c_g = \partial e / \partial e_g = 1$$

and partial uncertainties are evaluated from errors, discussed before. Resolution, repeatability, and actual temperature deviation of the calibrated instrument are considered.

For the best-known one-dimensional measuring device with digital reading, 10 nm resolution and 0 nm repeatability, the reading error is within $\pm 0.005 \text{ } \mu\text{m}$. At assumed rectangular distribution, the indication uncertainty is $u_i = 0.005 \text{ } \mu\text{m} / \sqrt{3} = 0.003 \text{ } \mu\text{m}$. At calibration of line scales using high-resolution video system (Table 3), the indication uncertainty is $u_i = 0.025 \text{ } \mu\text{m}$.

The combined standard uncertainty calculated by Eq. 3 is given in tables 1-3 for three different calibrations, for best presumed devices and conditions. The expanded uncertainty by Eqs. 5 is double value of the standard uncertainty, covering 95 % possibility of the measurement results. The first term in Eq. 5a is negligible yielding linear graph (Fig. 7a), while the sub-micrometre graphs (Figs. 7b and 7c) have comparable terms, with equal value at 30 mm in Eq. 5b and at 50 mm in Eq. 5c.

Tables 1-3 present the basic contributions of measurement uncertainty in length calibration procedures using the laser interferometry. With the presented measurement uncertainty, we can ensure traceability to calibration laboratories and industry. In the following, the methods and possible improvements to reduce the measurement uncertainty will be described.

For reducing the measurement uncertainty, we must always analyse and find the maximal contributions in the uncertainty budget. In Tables 1-3, where the measurement uncertainty budget is covered, the laser interferometry, the environmental temperature and the cosine error have the greatest influence on the measurement uncertainty. The impact of laser interferometry could be improved with better environmental sensors. Researchers [8] propose to measure the distance with laser interferometry in a vacuum, but such an approach is very expensive and impractical to implement. The uncertainty of the linear temperature coefficient could be reduced by an accurate knowledge of the material, but this determination is very critical in the length measurements. In the procedures described above, environmental temperature and cosine error for positioning the laser beam stand out the most. Minimizing the temperature effect could be eliminated with better environmental conditions, but it is very difficult to achieve environmental temperature deviation below 0.1 K. Temperatures could be measured in several points and perform real-time temperature compensation, as suggested by certain authors [1]. Even such a temperature measurement system is expensive and requires good handling and detailed analysis. Minimizing the cosine error could be eliminated with a better positioning system [3]. The authors presented a digital system for laser beam positioning. Such a system is sensitive to outside light and needs to be controlled. Proposed methods can contribute to more accurate measurement capabilities. However, they require additional investment and knowledge.

Table 1 Uncertainty budget for calibration of one coordinate measuring devices

Variable	Standard uncertainty	Distribution	Sensitivity coefficient	Uncertainty contribution
L_i	0.003 μm	Rectangular	1	0.003 μm
α	$0.58 \times 10^{-6} \text{ K}^{-1}$	Rectangular	$0.2 \text{ K} \times L$	$0.12 \times 10^{-6} \times L$
θ	0.1 K	Normal	$11 \times 10^{-6} \text{ K}^{-1} \times L$	$1.1 \times 10^{-6} \times L$
L_{Li}	$0.01 \mu\text{m} + 0.3 \times 10^{-6} \times L$	Normal	-1	$0.01 \mu\text{m} + 0.3 \times 10^{-6} \times L$
e_{\cos}	$0.3 \times 10^{-6} \times L$	Rectangular	1	$0.3 \times 10^{-6} \times L$
e_a	0.009 nm	Rectangular	1	0.009 nm
Total:				$\sqrt{(0.014 \mu\text{m})^2 + (1.19 \times 10^{-6} \times L)^2}$

The expanded uncertainty for calibration of one coordinate measuring devices is:

$$U = \sqrt{(0.03 \mu\text{m})^2 + (2.4 \times 10^{-6} \times L)^2} \quad (5a)$$

Table 2 Uncertainty budget for calibration of precise probes

Variable	Standard uncertainty	Distribution	Sensitivity coefficient	Uncertainty contribution
L_i	0.003 μm	Rectangular	1	0.003 μm
α	$0.58 \times 10^{-6} \text{ K}^{-1}$	Rectangular	$0.1 \text{ K} \times L$	$0.06 \times 10^{-6} \times L$
θ	0.05 K	Normal	$8 \times 10^{-6} \text{ K}^{-1} \times L$	$0.4 \times 10^{-6} \times L$
L_{Li}	$0.01 \mu\text{m} + 0.2 \times 10^{-6} \times L$	Normal	-1	$0.01 \mu\text{m} + 0.2 \times 10^{-6} \times L$
e_{\cos}	$0.3 \times 10^{-6} \times L$	Rectangular	1	$0.3 \times 10^{-6} \times L$
e_a	0.009 nm	Rectangular	1	0.009 nm
Total:				$\sqrt{(0.014 \mu\text{m})^2 + (0.54 \times 10^{-6} \times L)^2}$

The expanded uncertainty for calibration of precise probes is:

$$U = \sqrt{(0.03 \mu\text{m})^2 + (1.1 \times 10^{-6} \times L)^2} \quad (5b)$$

Table 3 Uncertainty budget for calibration of line scales

Variable	Standard uncertainty	Distribution	Sensitivity coefficient	Uncertainty contribution
L_i	0.025 μm	Normal	1	0.025 μm
α	$1.155 \times 10^{-6} \text{ K}^{-1}$	Rectangular	$0.1 \text{ K} \times L$	$0.1 \times 10^{-6} \times L$
θ	0.05 K	Normal	$10 \times 10^{-6} \text{ K}^{-1} \times L$	$0.5 \times 10^{-6} \times L$
L_{LI}	$0.01 \mu\text{m} + 0.2 \times 10^{-6} \times L$	Normal	-1	$0.01 \mu\text{m} + 0.2 \times 10^{-6} \times L$
e_{\cos}	$0.3 \times 10^{-6} \times L$	Rectangular	1	$0.3 \times 10^{-6} \times L$
e_a	0.013 nm	Rectangular	1	0.013 nm
Total:				$\sqrt{(0.03 \mu\text{m})^2 + (0.63 \times 10^{-6} \times L)^2}$

The expanded uncertainty for calibration of line scales is:

$$U = \sqrt{(0.06 \mu\text{m})^2 + (1.25 \times 10^{-6} \times L)^2} \quad (5c)$$

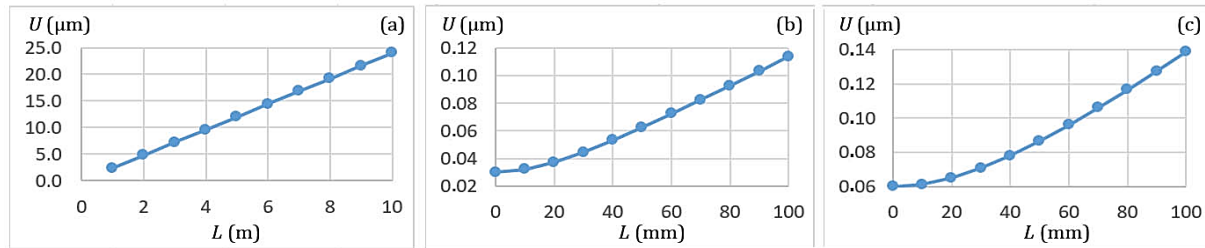


Fig. 7 Expanded calibration uncertainty: (a) one-dimensional coordinate measuring devices, (b) precise probes, (c) line scales

When measuring samples, the first term rises due to surface error, depending on sample's quality. While roughness of fine polished gauge blocks varies 15-30 nm [9, 10], geometrical imperfections have a sub-micrometre contribution [11]. Incrementally formed samples had roughness 0.4-0.75 μm [12], cylinders after turning had waviness 25-100 nm [13], and roughness after polishing 25-50 nm [14]. The waviness of a setting ring (Fig. 8) in contact with probes contributes 25-50 nm.

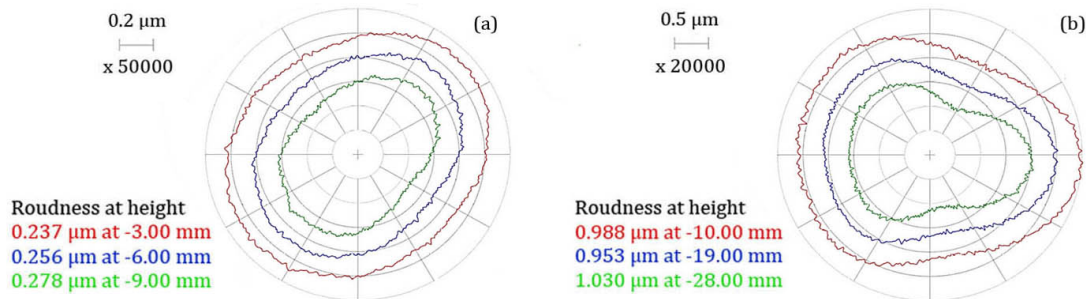


Fig. 8 Roundness profile of setting rings with diameter: (a) 50 mm, (b) 150 mm

4. Conclusion

Proper procedures with calibrated standards, deployable on machine tools, and effective error mitigation in real-world environmental conditions are essential to ensure the required measurement accuracy. Calibration procedures typically offer seven to ten times better accuracy than the calibrated device, ensuring proper metrological traceability. The uncertainty budget was presented for three specific cases of improved setups, chosen among the procedures used in our laboratory and verified by the national accreditation body. For highly precise calibration, our laboratory developed specific positioning systems applying a precise moving table, upgraded with clamping system and video probe system, adjusted with dial gauges, as detailed in the paper. By reviewing potential error sources in the measuring system, we found that using high-quality, periodically calibrated laser interferometers and sensors in well-controlled air conditions effectively mitigates instrumental and environmental errors. Geometrical errors can also be significantly reduced through precise adjustments of optics and probes. The majority of errors in the uncertainty budget are proportional to the displacement length, with the absolute term only becoming noticeable at shorter displacements (up to 100 mm). When calibrating gauge samples, such as cylinders

and rings, the surface error contributes essentially, while Abbe and dead-path errors in the absolute term become negligible. The most significant errors arise from material temperature differences along the measuring path. After several hours of stabilization and using three material temperature sensors along the path, this error remains the primary contributor. When further mitigation is necessary, such as for the calibration of step gauges, a more advanced monitoring system for material temperature was developed in our laboratory. Based on the content and measurement uncertainty calculations presented in the article, we provide the current needs of industry and calibration laboratories. Improvements and future work are also proposed. These enhancements could reduce measurement uncertainties but would require additional investments and expertise for implementation.

Acknowledgement

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core Funding No. P2-0190, Grant numbers 1000-20-0552, 6316-3/2018-255, 603-1/2018-16), as well as from the Metrology Institute of the Republic of Slovenia (funding of national standard of length; Contract No. C3212-10-000072). Furthermore, the authors would like to acknowledge the use of research equipment for processing and monitoring the machining and forming processes: Laser measuring systems for machine tools and coordinate measuring machines, procured within the project "Upgrading national research infrastructures – RIUM", which was co-financed by the Republic of Slovenia, the Ministry of Education, Science and Sport and the European Union from the European Regional Development Fund.

References

- [1] Šafarič, J., Dolšak, B., Klobučar, R., Ačko, B. (2020). Analysis of thermal contribution to the measurement uncertainty in step gauge calibration, *Precision Engineering*, Vol. 66, No. 52-61, doi: [10.1016/j.precisioneng.2020.06.012](https://doi.org/10.1016/j.precisioneng.2020.06.012).
- [2] Klobucar, R., Acko, B. (2017). Automatic high resolution measurement set-up for calibrating precise line scales, *Advances in Production Engineering & Management*, Vol. 12, No. 1, 88-96, doi: [10.14743/apem2017.1.242](https://doi.org/10.14743/apem2017.1.242).
- [3] Safaric, J., Klobucar, R., Acko, B. (2021). Measurement setup and procedure for precise step gauge calibration, *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, Article No. 1009610, doi: [10.1109/TIM.2021.3115579](https://doi.org/10.1109/TIM.2021.3115579).
- [4] Lipus, L.C., Budzyn, G., Acko, B. (2021). Analysis of laser interferometer measurement uncertainty by simulating error sources, *International Journal of Simulation Modelling*, Vol. 20, No. 2, 339-350, doi: [10.2507/IJSIMM20-2-563](https://doi.org/10.2507/IJSIMM20-2-563).
- [5] Köning, R., Weichert, C., Köchert, P., Guan, J., Flügge, J. (2013). Redetermination of the abbe errors' uncertainty contributions at the nanometer comparator, In: *Proceedings of the 9th International Conference on Measurement*, Smolenice, Slovakia, 171-174.
- [6] Leach, R. (2014). Abbe error/offset, In: Laperrière, L., Reinhart, G. (eds.), *CIRP Encyclopaedia of production engineering*, Springer, Berlin, Germany, 1-4, doi: [10.1007/978-3-642-35950-7_16793-1](https://doi.org/10.1007/978-3-642-35950-7_16793-1).
- [7] Guide to the expression of uncertainty in measurement GUM 1995, from https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf, accessed September 14, 2024.
- [8] Flügge, J., Weichert, C., Hu, H., Köning, R., Bosse, H., Wiegmann, A., Schulz, M., Elster, C., Geckeler, R.D. (2009). Interferometry at the PTB nanometer comparator: Design, status and development, In: *Proceedings of SPIE 7133, Fifth International Symposium on Instrumentation Science and Technology*, Shenyang, China, Article No. 713346, doi: [10.1117/12.821252](https://doi.org/10.1117/12.821252).
- [9] Buchta, Z., Šarbort, M., Čížek, M., Hucl, V., Řeřucha, Š., Pikálek, Z., Dvořáčková, Š., Dvořáček, F., Kůr, J., Konečný, P., Weigl, M., Lazar, J., Číp, O. (2017). System for automatic gauge block length measurement optimized for secondary length metrology, *Precision Engineering*, Vol. 49, 322-331, doi: [10.1016/j.precisioneng.2017.03.002](https://doi.org/10.1016/j.precisioneng.2017.03.002).
- [10] Ali, S.H.R., Naeim, I.H. (2014). Surface imperfection and wringing thickness in uncertainty estimation of end standards calibration, *Optics and Lasers in Engineering* Vol. 60, 25-31, doi: [10.1016/j.optlaseng.2014.03.012](https://doi.org/10.1016/j.optlaseng.2014.03.012).
- [11] Lipus, L.C., Acko, B., Tompa, J. (2022). Experimental determination of influences on a gauge block's stack length, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 339-349, doi: [10.14743/apem2022.3.440](https://doi.org/10.14743/apem2022.3.440).
- [12] Visagan, A., Ganesh, P. (2022). Parametric optimization of two point incremental forming using GRA and TOPSIS, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 615-626, doi: [10.2507/IJSIMM21-4-622](https://doi.org/10.2507/IJSIMM21-4-622).
- [13] Tomov, M., Gecevska, V., Vasileska, E. (2022). Modelling of multiple surface roughness parameters during hard turning: A comparative study between the kinematical-geometrical copying approach and the design of experiments method (DOE), *Advances in Production Engineering & Management*, Vol. 17, No. 1, 75-88, doi: [10.14743/apem2022.1.422](https://doi.org/10.14743/apem2022.1.422).
- [14] Yao, W., Lyu, B., Zhang, T., Guo, L., Song, Y. (2023). Effect of elastohydrodynamic characteristics on surface roughness in cylindrical shear thickening polishing process, *Wear*, Vol. 530-531, Article No. 205026, doi: [10.1016/j.wear.2023.205026](https://doi.org/10.1016/j.wear.2023.205026).

Improved Whale Optimization Algorithm for supply chain financial risk assessment of cloud warehouse platform

Zhang, H.^{a,b,*}, Guo, Y.W.^a, Hou, Y.^a, Tang, L.^a, Deveci, M.^c

^aBeijing Technology and Business University, P.R. China

^bBeijing Food Safety Research Base, P.R. China

^cRoyal School of Mines, Imperial College London, United Kingdom

ABSTRACT

This study provides an in-depth analysis of a new financial model for cloud warehouses and evaluates the associated credit risk within the context of supply chain financing, focusing on the intelligent transformation in this field. Concurrently, an optimization problem was derived from the evaluation issue, with the whale optimization algorithm (WOA) used to identify a reasonable default point and distance. To simplify the identification of these points, we enhanced the traditional WOA, resulting in an improved version, the IWOA, which demonstrated very good optimization performance. The IWOA's optimization capabilities were applied to determine the optimal ratio of short- and long-term debt coefficients, identifying the default point in the Kealhofer, McQuown, and Vasicek (KMV) credit monitoring model, replacing fixed values and yielding more precise results. Furthermore, this study introduces a novel analytical approach to credit risk measurement, advancing the development of related theories and methods. Accurate analysis of financial stability and risk is crucial in industrial sectors, including engineering and manufacturing. The simulation using specific data revealed that the IWOA-KMV model exhibited better and faster optimization capabilities, with greater discrimination ability compared to the KMV model. Overall, this study examines the risk factors in the cloud warehouse financing model, offers an improved version of the WOA, introduces a modified IWOA-KMV model to create a scientific, practical credit risk assessment framework, and provides guidance for risk control in cloud warehouse financing, a novel financing service.

ARTICLE INFO

Keywords:

Cloud warehouse platform;
Supply chain finance;
Risk assessment;
KMV model;
Swarm intelligence;
Improved Whale Optimization Algorithm

*Corresponding author:

btbuzhhaozhao@126.com
(Zhang, H.)

Article history:

Received 19 September 2024

Revised 9 October 2024

Accepted 15 October 2024



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Smart logistics reforms have promoted the development of supply chains. Adapting to the trend of intelligent warehousing has becoming increasingly difficult with advances in science and technology. In recent years, new warehousing services such as *pre-warehouse*, *lightning warehouse*, *unmanned warehouse*, and *satellite warehouse* have continuously emerged [1]. A cloud warehouse is a common storage solution that leverages cloud computing, the Internet, and other information technology and presents a management approach that extends beyond traditional warehousing services and idle storage facilities. The cloud warehouse platform for supply chain finance services utilizes information technologies such as cloud computing and the Internet of Things (IoT) to simultaneously collect and analyze supply chain information and provide financing channels for supply chain enterprises [2]. Through the standardization of warehouse receipt pledge financing, supply chain enterprises can offer credit guarantees to financial institutions, enhancing regulatory support for financial institutions and effectively reducing barriers associated with factual

information. However, the cloud warehouse platform that provides financial services for business enterprises also presents a new challenge in supply chain financing. The liquidity, uncertainty, and market volatility characteristics associated with using receipts as pledges in the supply chain financing of products from a warehouse make it difficult to assess pledge value and control credit risk [3]. Li *et al.* studied an algorithm that combines adaptive nonlinear convergence factor with variable gain compensation mechanism, adaptive weighting and advanced spiral convergence strategy, which improves the global search ability, convergence speed and accuracy of the WOA algorithm [4]. Therefore, considering the new characteristics and needs of the existing market, this study combines the improved whale optimization algorithm (IWOA) and the Kealhofer, McQuown, and Vasicek (KMV) model to carry out financial credit evaluation, which has more comprehensive search capabilities and obvious distinctions, creating a new model that evaluates the financial credit risks involved in cloud warehouse platforms more effectively.

The study focuses on the unique credit risk challenges in supply chain finance faced by small and medium enterprises (SMEs). It elaborates and analyzes existing credit risk assessment theories and strategies considering the risk factors in the supply chain analysis of a cloud warehouse. Based on this, the WOA was modified in three ways to address model calculation issues, and the KMV model was built to provide decision support for the supply chain financial risk assessment system. The usefulness of the proposed model for supply chain financial risk assessment was clearly confirmed through an example.

The remainder of the paper is structured as follows. In Section 2, a literature review is introduced, and in Section 3, the proposed approach is outlined. In Section 4, the IWOA-KMV credit risk assessment model is presented, and in Section 5, a simulation study of this model is presented. Finally, in Section 6, the conclusions are presented.

2. Literature review

2.1 Research status of the cloud warehouse platform

In recent years, the cloud warehouse platform has attracted widespread attention. Dobrescu *et al.* argued the need to combine cloud computing with supply chain networks to make them more intelligent, highlighting the importance of considering scientific and technological advancements in informed decision-making [5]. Song *et al.* proposed building a comprehensive information platform to solve information asymmetry problems in traditional supply chains [6]. In the era of information networking technology, virtual supply chains conduct transactions through information-sharing platforms, which can reduce information barriers and improve supply chain flexibility. Chai designed a conceptual model of an information platform based on blockchain technology by analyzing the factors affecting the construction of a supply chain information platform, emphasizing both the business process and economic model [7]. Noting the widespread implementation of blockchain technology in various territories, Lee highlighted that information system reviews needed to be more comprehensive to clearly assess the relative level of use of this technology [8]. Marza *et al.* established an evaluation framework based on the analytical hierarchy process (AHP) for coordinating the levels of finance and digitalization. Taking the data of five Eastern and Central European countries as an example, it was found that finance and digital water fortresses required government intervention for improved coordination results [9]. Deng introduced the concept of *cloud* technology to drive innovation in supply chain financial services and proposed the *cloud warehouse* model, emphasizing sharing as its core principle. He explored the integrated services and continuous spot trading mechanism enabled by this model, aiming to foster a multi-party win-win scenario. Furthermore, in terms of risk management, Pan *et al.* used the risk-adjusted return on capital (RAROC) method to establish a loan pricing model under the *cloud warehouse* financing mode, which exploited the powerful information collection function of a cloud warehouse and assisted financial institutions in avoiding risks and improving yield [11]. In the development direction of the cloud warehouse platform, Chinese enterprises have given more consideration to the diversity of services, combined with supply chain financial services, to solve the capital problem for small and medium-sized enterprises. Cloud warehouses, a new model of

warehouse management and supply chain financing services, along with the associated management risks and problems, are gradually becoming a popular research direction.

2.2 Research status of supply chain financial risk

In supply chain financing, warehouse receipt pledges help reduce the financial risks of SMEs. Wang *et al.* discussed the core index of warehouse receipt pledge risk control, specifically focusing on the pledge rate [3]. Wang *et al.* elaborated on the dangers of warehouse receipt pledge financing and the accompanying risk-management techniques, particularly those used by fourth-party logistics firms [12]. Zhang *et al.* developed a credit risk index system, a SVM, and a back propagation neural network model to assess the credit of top supply chain companies [2]. Li *et al.* use random forest and logistic regression classifications to evaluate the materiality and credit risk of P2P firms, identifying loan targets with a high probability of default [13]. Malhotra conducted an in-depth analysis of the identification of bad credit applications. Discovering that the neuro-fuzzy system outperformed other methods in identifying bad credit applications, he improved this system to assess the risk level of financial enterprises [14]. To address the capacity combination problem of contract manufacturers, Zheng and Bao proposed a uniform variation artificial fish swimming algorithm (AFSA_UM) with a good convergence performance and strong robustness [15]. Esmaili-Najafabadi *et al.* used the particle swarm optimization (PSO) algorithm to provide the solution to a mixed-integer nonlinear programming (MINLP) model for supplier selection and order allocation in a centralized supply chain that considered risk avoidance and interruption risk decision makers [16]. Liu and Huang introduced the integrated SVM model for risk assessment in supply chain finance [17]. Based on data from small and medium-sized enterprises in China, Du *et al.* compared six commonly used machine learning classification models and found that neural networks, support vector machines, and GBDT models performed better [18]. Warehouse receipt pledge financing is an important method in supply chain financing, and targeted evaluation models have been developed by many scholars to deal with the unique risk factors associated with this method. Traditional qualitative analysis models have stood the test of time, demonstrating a good robustness and wide applicability. However, these models require manual intervention and provide a low evaluation accuracy [19-20]. Modern risk assessment models have significantly improved risk quantification, assessment, and prediction. The advent of artificial intelligence algorithms has led to a notable improvement in their data analysis and mining capabilities [21], resulting in the application of intelligent optimization algorithms combined with evaluation models becoming a research hotspot.

The cloud warehouse financing mode theory represents the theoretical sublimation of the supply chain financing mode as a new financing mode. The practical development of cloud warehouse platforms is progressing rapidly, but theoretical research is lagging behind. In practice, enterprises primarily conduct research on cloud warehouse systems, and the unified standardized model has yet to be fully developed. The potential market environment is still actively being explored. Among the services provided by cloud warehouses, supply chain financial services, which form an important component of the enterprise capital chain, are more important than other warehousing services. Both opportunities and challenges coexist in these services, with existing management experience and theoretical guidance being limited. On the other hand, the WOA has gained significant attention from academics due to its small number of parameters, simplicity, and ease of use. An IWOA was applied to solve the supply chain financial risk assessment model of a cloud warehouse platform, providing a new solution and method for credit risk assessment.

3. Methodology

3.1 KMV model

Credit risk is a primary risk faced by commercial organizations. Effectively evaluating and judging credit risks and providing good advice to managers is important. While significant advancements have been made in China's financial management technology, the technical capabilities for risk quantification and the completeness of historical data remain relatively underdeveloped. The

KMV model is a well-established dynamic model, making it less influenced by subjective factors. Its parameters are straightforward to estimate, and it is capable of accurately obtaining the necessary data, offering both flexibility and timely results. Furthermore, it can comprehensively examine the financial information of listed companies and forecast the likelihood of enterprise default based on the closing price of a stock, reflecting the level of enterprise risk to a certain extent. After comprehensively comparing traditional and modern risk measurement models, we selected the KMV model for further study and modification in the evaluation of supply chain financial risk.

Modern asset-pricing and financial theories form the foundation of the KMV model. The influence of corporate debt structure on corporate default probability shows that the accurate assessment and analysis of business credit risk require the calculation of the default point and default distance [23-24]. The core principle of the KMV model is to estimate the expected default rate, followed by the calculation of corporate assets. This is done by adding the weighted sum of long-term debt, short-term debt, and equity value. The model can not only be adjusted dynamically according to historical data but also has a certain predictability and foresight, being able to reflect the comprehensive expectation of the future credit risk of a company in real time. The precise steps in the working of this model are as follows.

First, the net worth of the liabilities, stock value in the market, and volatility of stock market value are used to compute the value of assets and property worth fluctuations. The Black-Scholes-Merton option pricing model is used to estimate the market value of the equity of an enterprise. Then, the relationship between equity value volatility S_E and asset value volatility S_A is derived by solving a system of simultaneous equations. The unknown variables asset value V_A and asset value volatility S_A , which are required in the KMV model, can be obtained as shown in Eq. 1.

$$\begin{aligned} f(V_A, S_A, D, r, T) &= V_E = V_A \times N(d_1) - e^{-r \times T} \times D \times N(d_2) \\ d_1 &= \frac{\ln(\frac{V_A}{D}) + (r + \frac{S_A^2}{2}) \times T}{S_A \sqrt{T}} \\ d_2 &= d_1 - S_A \sqrt{T} \\ S_E &= \frac{V_A}{V_E} \times N(d_1) S_A \end{aligned} \quad (1)$$

Where V_A is the value of the assets, S_A is the value of the asset volatility, V_E is the market value of equity, S_E is the value of volatility in the stock market, $N(d)$ is the function of the cumulative probability distribution of the normal distribution, r is the risk-free interest rate, T is the debt maturity and D is the book value of debt.

The function $N(d)$ is calculated as shown in Eq. 2.

$$N(d) = \delta^d \frac{1}{\sqrt{2\pi}} e^{-0.5x^2} dx \quad (2)$$

Then, the future value is predicted according to the existing value of the company, and the default point (DPT) is set according to the debt structure of the company. In the KMV model, a default may occur if the corporate debt is close to the market value. Therefore, a comparison between the default point and corporate assets is performed to assess the likelihood of default. If a company's assets surpass the default point, the likelihood of a default event is reduced; if they fall below the default point, the likelihood increases. Through the analysis of massive data related to default events, the KMV model suggests that both short- and long-term liabilities are typically subject to default situations. Consider a case in which the default coefficient is set to (1,0.5). In this case, the default point can be obtained as shown in Eq. 3.

$$DPT = 1'STD + 0.5'LTD \quad (3)$$

The gap between the predetermined point and the anticipated value of the company and the volatility of its asset value, that is, the default distance DD , are calculated together. The default distance is the gap between a company's asset value and its default point, calculated as shown in Eq. 4.

$$DD = \frac{\text{Expected asset value} - \text{default point}}{\text{Expected asset value} \times \text{Asset volatility}} \quad (4)$$

According to Eq. 4, the final default distance is a dimensionless variable that represents the probability of default without considering factors such as the market environment, enterprise size, and future growth. The risk of overdue default is inversely proportional to the default distance. The specific expression for the default distance is given in Eq. 5.

$$DD = \frac{E(V_A) - DPT}{E(V_A)S_A} \quad (5)$$

Finally, the overdue default rate and expected default frequency EDF of the enterprise are calculated according to the default distance, and the default risk degree of the enterprise is obtained. The computing company for the expected default frequency is given by Eq. 6.

$$EDF = N(-DD) = 1 - N(DD) \quad (6)$$

The KMV model uses all its massive global default database resources to develop a precise relationship between default probability and default distance. This model provides results that are accurate and consistent with respect to real-world scenarios.

3.2 Improved whale optimization algorithm

In 2016, Mirjalili and Lewis proposed a new swarm intelligence optimization method called the WOA [25]. Whales use a distinctive hunting technique, known as the bubble-net hunting method, which is based on their observed social behavior. The working of the WOA is inspired by the behavior of whales. This technique may be roughly broken down into three stages: a bubble-net attack, surround hunting, and hunting for prey.

The whale surrounds the prey based on its location. Mimicking this, the WOA tries to identify the best search agent based on the location. The selected search agent is employed in the next step. Eq. 7 and Eq. 8 present the mathematical models of the process.

$$\vec{D} = |\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)| \quad (7)$$

where D is the individual whale distance vector, X^* is the localized ideal response, X is the position vector, t is the current number of iterations and C is the random number in the range $[0,2]$.

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (8)$$

where A is the random number in the range $[-2,2]$. The mathematical expression for A is shown in Eq. 9.

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (9)$$

where a is the linear decrease from 2 to 0, r is the random vector between $[0,1]$. The mathematical expression of C is shown in Eq. 10.

$$\vec{C} = 2 \cdot \vec{r} \quad (10)$$

Whales also use bubble-net strategies to attack their prey. The corresponding stage in the algorithm include contraction encircling and two spiral updates.

According to Eq. 8, when $-1 < A < 1$, mirroring the change in the position of a group of whales, the search agent moves to the best agent in the new location and contracts.

Eq. 11 describes the way a whale (agent) swims (moves) in a decreasing circle while adopting a spiral path.

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (11)$$

where D' is the distance in vector terms between the agent and the optimal agent in the current location, b is the parameter that establishes the logarithmic helix form, and l is the random number in $[-1,1]$. An agent is closest to the resource when $l = -1$, and farthest from the resource when $l = 1$.

In the WOA, spiral renewal behavior and contraction encircling behavior occur simultaneously. Mirjalili and Lewis assumed that when updating the positions of the search agents, there would be a 50 % chance of using either the spiral model or the contraction wrapping process during each iteration of the WOA [24]. The mathematical model of this behavior is described in Eq. 12.

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X^*(t)} - \overrightarrow{A} \cdot \overrightarrow{D'} & p < 0.5 \\ \overrightarrow{D'} \cdot e^{bl} \cdot \cos(2\pi l) + \overrightarrow{X^*(t)} & p \geq 0.5 \end{cases} \quad (12)$$

where p is the random number between $[0,1]$.

Mirjalili and Lewis (2016) used parameter A to describe the way whales hunt prey. When $|A| > 1$, the search agent is compelled to move toward another randomly chosen agent position. This process is called exploration. When $|A| < 1$, the search agent moves toward the optimal agent in its current location. This process is called development.

Theoretically, the WOA can be regarded as a metaphorical *machine* that optimizes an entire environment. The mathematical model is shown in Eqs. 13 and 14.

$$\overrightarrow{D} = \overrightarrow{C} \cdot \overrightarrow{X_{rand}} - \overrightarrow{X} \quad (13)$$

$$\overrightarrow{X(t+1)} = \overrightarrow{X_{rand}} - \overrightarrow{A} \cdot \overrightarrow{D} \quad (14)$$

where X_{rand} is the position vector for a randomly selected whale.

Being a representative of the swarm intelligence algorithm, the WOA includes a few parameters, has a straightforward algorithmic structure, and is easy to use. It falls under the category of bionic group optimization algorithms, exhibiting an outstanding performance in solving optimization problems. To address the shortcomings of the WOA and its frequent low convergence [26-27], this study proposes improvements to the algorithm from the following three aspects.

The WOA can suffer from getting trapped in local optima, and the algorithm can converge prematurely. It is necessary to balance the gradual convergence of the algorithm and its capacity for global search and overcome its premature convergence. The Gaussian mutation operation was introduced to tackle the slowing of personal evolution and the loss of population diversity. Eq. 15 provides the probability density function of the Gaussian distribution.

$$f(x) = \frac{1}{\sqrt{2ps}} e^{-\frac{(x-m)^2}{2s^2}} \quad (15)$$

where m is the expectation of the Gaussian distribution and s^2 is the variance of the Gaussian distribution. When $m = 0$ and $s = 1$, the distribution follows the conventional Gaussian distribution. The mathematical representations of the WOA improved by Gaussian mutation based on Eqs. 8, 11, and 14 are as follows.

$$\overrightarrow{X(t+1)} = (\overrightarrow{X^*(t)} - \overrightarrow{A} \cdot \overrightarrow{D}) \cdot (1 + k \cdot \text{Gauss}(0,1)) \quad (16)$$

$$\overrightarrow{X(t+1)} = (\overrightarrow{D'} \cdot e^{bl} \cdot \cos(2\pi l) + \overrightarrow{X^*(t)}) \cdot (1 + k \cdot \text{Gauss}(0,1)) \quad (17)$$

$$\overrightarrow{X(t+1)} = (\overrightarrow{X_{rand}} - \overrightarrow{A} \cdot \overrightarrow{D}) \cdot (1 + k \cdot \text{Gauss}(0,1)) \quad (18)$$

where k (0,1) is a decreasing random number and Gauss (0,1) is the standard Gaussian distribution.

Regarding DE, it is a powerful technique for addressing limitations in optimization, and it frequently been used in conjunction with other techniques in a variety of disciplines. The availability of a difference operator can lessen the blindness of the algorithm and speed up the optimization process. In the DE method, individual mutations are achieved by randomly selecting two individuals from the population, scaling their vector differences, and combining them with the vector of the individual to be modified. If the new solution has a higher fitness value than the current optimal solution, it replaces the current one; otherwise, the same solution is kept for the next iteration. Through this mechanism, individuals are guided to search for the optimal solution, accelerating the progress of the algorithm and lowering the likelihood of settling on a local optimal solution. The mechanism of DE correction is expressed in Eq. 19.

$$\overrightarrow{X(t+1)} = \overrightarrow{X_i(t)} + F(\overrightarrow{X_m(t)} - \overrightarrow{X_n(t)}) \quad (19)$$

where $\overrightarrow{X_i}$, $\overrightarrow{X_m}$, $\overrightarrow{X_n}$ are arbitrarily different individuals in the t -th generation population and F is the mutation factor.

The AFSA is an adaptive optimization technique whose working is based on the behavior of fish shoals. An artificial fish looks for the best value among the neighboring fish in its field of vision, avoids overcrowding, and moves in its preferred direction. In the basic WOA, an agent simply encircles its intended target and looks for the necessary resources. Therefore, it can easily collide

or even converge to the local optimal value prematurely. Exploiting the congestion factor and the rear-end collision behavior of the AFSA, premature phenomena can be avoided early in the developmental stages of the algorithm, and the capacity of the algorithm for optimization can be increased. The crowding degree discrimination method is expressed in Eq. 20.

$$Y_{max} / n_f > \delta \times Y_i \quad (20)$$

where Y_{max} are amount of the resources near the ideal place at present, Y_i is the self-state value of the artificial fish (agent), n_f is the number of individual artificial fish (agent) and δ is the congestion factor.

With an increase in the δ size, the capacity for congestion decreases while the ability of the algorithm to escape local optima increases. To avoid overcrowding around extreme values or strong points, the artificial fish (agent) will randomly depart or be rejected by other artificial fish (agent) when approaching such values, resulting in them being unable to precisely locate these values, slowing down the convergence speed of the algorithm. When the inequality holds, the artificial agent will consider the Y_{max} position to be too crowded.

4. IWOA-KMV credit risk assessment model

4.1 Parameter variables

Macroscopically, the enterprise stock market value V_E includes two parts: the market value of the enterprise's tradable shares V_{E1} , and the market value of the enterprise's non-tradable share V_{E2} , that is, the market value of equity V_{E2} , which is given in Eq. 21.

$$V_E = V_{E1} + V_{E2} \quad (21)$$

According to the Merton model, the enterprise stock market value can also be calculated using the company asset value V_A , as given in Eq. 22.

$$V_E = V_A \cdot N(d_1) - e^{-r \cdot T} \cdot D \cdot N(d_2) \quad (22)$$

where V_E is the equity market value, V_A is the cumulative probability distribution function of the asset value, $N(d)$ is the normal distribution variable, r is the risk-free interest rate T , is the debt maturity and D is the debt book value.

According to Ito's lemma, the stock price volatility S_E is given by Eq. 23.

$$S_E = \frac{V_A}{V_E} \cdot N(d_1) \cdot S_A \quad (23)$$

The volatility S_A of the asset values V_A and return on assets can be calculated according to Eq. 23, which represents the stock price volatility S_E .

The book value of the debt of an enterprise can be obtained from financial reports. According to the option pricing theory, when the corporate asset value V_A is close to or less than the corporate debt book value D , the enterprise has a high probability of default risk, otherwise, the probability of default risk is small.

The results of the KMV model after many sample analyses and calculations revealed that when the enterprise asset value V_A was at a certain point between the short-term and total liabilities, the enterprise would default at a point called the default point. In the KMV model, the default point is defined as given in Eq. 24.

$$DPT = 1'STD + 0.5'LTD \quad (24)$$

where the short-term debt STD and long-term debt LTD have been added according to the coefficient combination (1,0.5), and the default trigger point DPT is the standard coefficient configuration of (1,0.5).

Since the thick-tailed distribution of the asset value was different from the actual default of the enterprise, the default distance DD in the KMV model was used to express the distance between the probability distribution of the asset value and the starting point of default and further calculate the expected default rate DEF of the enterprise.

The overdue default rate DEF of an enterprise indicates the possibility of default.

In this study, to facilitate a simple comparison, we chose the one-year time deposit benchmark interest rate of Renminbi (RMB) from the financial institutions of the People's Bank of China in 2019 as 1.50 %. Considering the 90-day treasury bond yield as a risk-free interest rate is a common approach in many financial markets around the world. However, when analyzing the Chinese financial market, it is important to consider the national conditions and the characteristics specific to it.

In the KMV model, the debt maturity of credit risk assessment is usually one year and accurately reflects the current state of short-term liabilities, where $T = 1$.

4.2 Model setting

In the traditional KMV model, the default point is usually calculated as the sum of the short-term debt multiplied by 1 and the long-term debt multiplied by 0.5. To explore the best choice of the default point for the KMV credit risk assessment model, we modified the KMV model and set the default point as the linear sum of the long-term debt multiplied by b and the short-term debt multiplied by a . The IWOA iteratively calculates the optimal combination of the coefficients (a and b) for short-term and long-term debt with the aim of determining the closest approximation to DPT , as shown in Eq. 25.

$$DPT = a'STD + b'LTD \quad (25)$$

An enterprise faces a higher default risk if its estimated asset value is below the default point. In the KMV model, the degree of default risk can be clearly measured through the default distance and default point. Under the general assumption of normal distribution, the default rate can be determined using a probability distribution function. Since there is currently no credible rating agency in China, it is essential to have a single measurement standard for the chosen samples. This study derives conclusions from the assessment of the China Securities Regulatory Commission regarding special treated (ST) shares. The CSRC marks identified listed companies with abnormal conditions and designates them as ST shares to alert investors of potential investment risks. The determination of the optimal default point of the KMV model can be transformed into an optimization problem by considering whether the evaluated enterprise has been categorized as an ST share or non-ST share.

A company with a good credit is considered a good credit company belonging to category G and one with a bad credit is considered a bad credit company belonging to category B. We chose equal amounts of data from companies with and without ST shares. The proportion of samples from firms with and without ST shares was $N/2$.

Since the number of samples of ST and non-ST shares was the same, according to the KMV model, the default distance DD of each company adopted the mean \overrightarrow{DD} as the discriminant criterion; that is, when the default distance DD of the evaluated company was less than or not less than \overrightarrow{DD} , the company was judged to be class B or class G, respectively, as shown in Eq. 26.

$$\text{if } \begin{cases} DD \geq \overrightarrow{DD} & \text{the company is judged as a class G company} \\ DD \leq \overrightarrow{DD} & \text{the company is judged as a class B company} \end{cases} \quad (26)$$

Correct classification:

- a. ST share companies are classified as class B
- b. Non-ST share companies are classified as class G

Classification errors:

- c. ST share companies are mistakenly classified as class G
- d. Non-ST share companies are mistakenly classified as class B

The probability of occurrence of the misclassification error m for the mistaken classification of companies with c ST shares as class G is given by Eq. 27.

$$P_1 = 1 - \frac{2m}{N} \quad (27)$$

The probability of occurrence of the misclassification error n for the mistaken classification of companies with d non-ST shares as class B in is given in Eq. 28.

$$P_2 = 1 - \frac{2n}{N} \quad (28)$$

The average misjudgment rate refers to the average probability of misjudgment caused by the two deviation coefficients a and b when the IWOA model is used to calculate the optimal combination of deviation coefficients. The average misjudgment rate is expressed by Eq. 29.

$$f(a, b) = \min \left(1 - \frac{m+n}{N} \right) \quad (29)$$

Fig. 1 illustrates the steps involved in the determination of the optimal default coefficient combination based on the IWOA-KMV model.

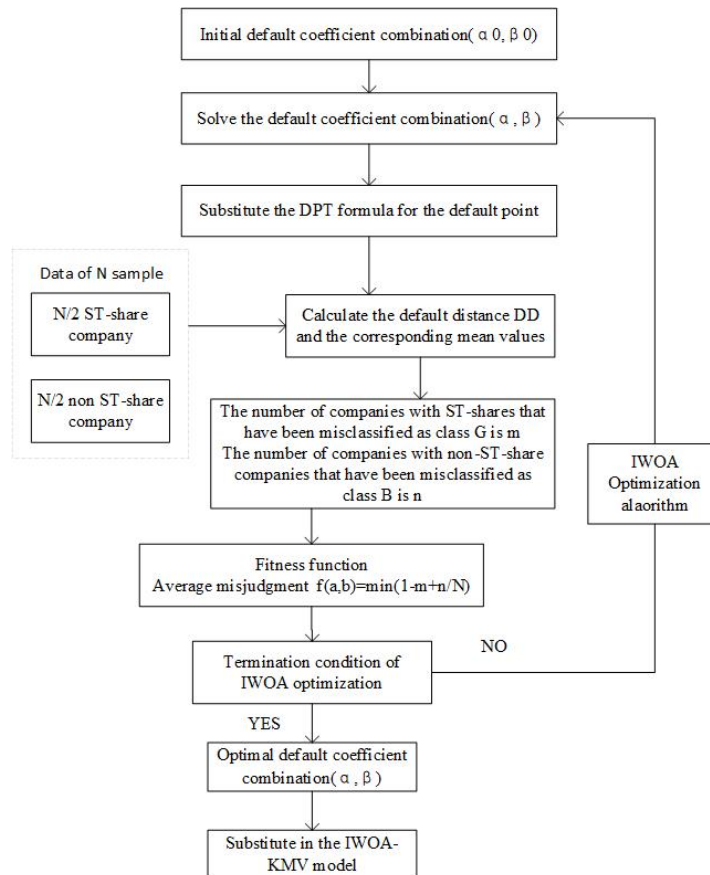


Fig. 1 Steps in the IWOA-KMV model

5. Simulation study on IWOA-KMV credit risk assessment model

5.1 Sample data acquisition and processing

The KMV model was selected to optimize and improve the IWOA-KMV model and assess the credit risk of businesses using real-time data from the capital market, which requires public information. Additionally, considering the challenges involved in gathering data samples and the accessibility of actual business financial data, we selected the data of listed companies as samples to simulate the IWOA-KMV model and verify its effectiveness. The data of risk warning (ST or * ST) companies were selected to conduct model simulation research. The ST or * ST mark indicates that a particular company has sustained a continuous operating loss for 2–3 years, potentially posing a risk of delisting. There are some measures, such as special treatment or a delisting warning, which are used to deal with a listed firm with an * ST designation; such a firm has experienced operating losses for two years in a row or has unusual circumstances, including financial data fraud, which may lead to the risk of delisting. * ST shares indicate a *delisting warning*. Stock exchanges take

special measures to address companies with investment risks, including poor operating conditions and financial anomalies, to provide warnings to investors.

In this study, we selected companies marked as ST and non-ST in the same industry from 2014 to 2019 for data analysis. Then, we determined the default distance and default probability of the two types of businesses using the IWOA-KMV model to assess whether they were consistent with the actual condition of the companies (in terms of the ST designation). The sample was divided into training and testing sets. For the training samples, 150 listed companies classified as ST and 150 companies classified as non-ST from China, covering the years 2014, 2015, and 2016, were selected. These samples were used to calculate the optimal coefficient of the default point. In 2017, 2018, and 2019, 150 Chinese listed companies in China were designated as ST, and the corresponding 150 listed companies in the north (not by ST) were selected as test samples, and the effectiveness of the model was statistically analyzed.

5.2 Model simulation

To explore the best combination of the optimal short-term debt coefficient a and long-term debt coefficient b in the KMV credit risk assessment model and modify the default point to improve the results of the model, the IWOA was utilized in this study to carry out the optimal iteration on the basis of the KMV model and construct the IWOA-KMV calculation model for the optimal default coefficient. Using the idea of optimization to determine the best combination of the optimal short- and long-term debt coefficients (a, b), we modified the default point for the model to obtain better results. The parameter settings of the IWOA and IWOA-KMV are listed in Table 1.

Table 1 IWOA and IWOA-KMV parameter settings

Parameters	Value	Meaning
SearchAgents_no	50	Population size
G	0	Initial number of iterations
Max_iteration	500	Maximum number of iterations
r	1.50 %	Risk-free rate of return
T	1	Debt maturity
(a_0, b_0)	(1,0.5)	Initial default coefficient combination

One hundred and fifty companies in China were listed with ST in 2014, 2015, and 2016. The corresponding unidentified companies (listing status is Norm) were selected as the development samples. The IWOA was applied to 50 iterations of the training samples to calculate the optimal coefficient combination of the default points. Finally, the optimal default coefficient combination of the current liability and long-term liability was obtained in the optimal default point calculation equation, and the ideal short-term debt to long-term liability default coefficient combinations were approximately 3.06 and 1.13, respectively.

The optimal default point of the IWOA-KMV model is calculated as shown in Eq. 30.

$$DPT = 3.06'STD + 1.13'LTD \quad (30)$$

In 2017, 2018, and 2019, 210 Chinese listed firms were chosen as test samples, and the corresponding unidentified companies (listing status is normal) were selected as test samples. The KMV and IWOA-KMV models were used to simulate the test samples. The calculation rules for the default point in the KMV model are given by Eq. 31.

$$DPT = 1'STD + 0.5'LTD \quad (31)$$

The calculation rules for the default points in the IWOA-KMV model are given by Eq. 32.

$$DPT = 3.06'STD + 1.13'LTD \quad (32)$$

To demonstrate the effect of the model correction more intuitively, a double comparison was used to simulate the model.

The default distance DD_{KMV} and $DD_{IWOA-KMV}$ of each enterprise were calculated by using the KMV model and IWOA-KMV model, respectively, to determine the extent of any potential improvements in the measurement accuracy of the default coefficient and whether the IWOA-KMV model had a better discrimination effect than that of the KMV model.

The data of the ST share and non-ST share companies were used for the simulation to verify whether the IWOA-KMV model had a better accuracy than that of the KMV model. The extent of the improvement in the performance of the KMV model through the IWOA could reveal a good distinction between the companies designated as ST (listing status is ST or * ST) and unmarked companies (listing status is Norm), which primarily depicted the amount of disparity between the default distances of the two types of firms, thus proving the goodness-of-fit and measurement performance of the IWOA-KMV model.

5.3 Analysis of simulation results

First, the calculation results of the test sample data were analyzed. The data with no obvious anomalies were marked as correct, and the correctness rate was calculated. A comparison of the correct rate of the default distance under different combinations of the default coefficients used by the KMV and IWOA-KMV models is given in Table 2.

According to the calculation results, the accuracy of the IWOA-KMV model was higher than that of the KMV model for both the ST share and non-ST share companies.

Using the mean difference test of the results of the test samples calculated by the IWOA-KMV model, the default distances under different combinations of the default coefficients used by the KMV and IWOA-KMV models were compared. The results are presented in Table 3.

It is widely known that the probability of the default increases as the default distance decreases. The default probability of the ST companies was higher than that of the norm companies in the calculation findings of the two models, adhering to practical scenarios. Furthermore, the average default distance of the IWOA-KMV model was 1.67, which was much greater than the 1.09 of the KMV model, from the perspective of mean difference. Therefore, based on the optimal default coefficient combination calculated by the IWOA algorithm, it was clear that the IWOA-KMV model and the KMV model had different results, and the default point correction had a better differentiation for the two types of companies with different degrees of risk.

Table 2 Comparison of the correct rates of the default distance under the KMV model and IWOA-KMV model

Name	Model Default coefficient combination	Correct rate of the default distance	
		ST share company (%)	Norm. corp. (%)
KMV	(1,0.5)	95.71	91.90
IWOA-KMV	(3.06, 1.13)	98.10	97.62

Table 3 Comparison of the mean difference of the default distance between the KMV model and IWOA-KMV model

Name	Model Default coefficient combination	Average default distance		Mean difference
		ST company	Norm. corp.	
KMV	(1,0.5)	1.79	2.88	1.09
IWOA-KMV	(3.06, 1.13)	1.54	3.21	1.67

6. Conclusion

Currently, China's cloud warehouse platform model is still in its early stages. The existing literature on cloud warehouse platforms, shared warehousing, and related supply chain financial services is limited, and publicly available data is scarce. Additionally, because the cloud warehouse model does not have a unified standard, there are many types of financial service products, and the links involved in financing risk control are complex. There are certain limitations in our analysis of financial risk in a supply chain using a cloud warehouse platform.

This study presents a thorough analysis of the novel financing model for the cloud warehouse platform from the perspective of the intelligent supply chain finance revolution and provides a preliminary discussion on the current situation and management methods of credit risk assessment using the cloud warehouse financing model. Considering the cloud warehouse platform, this study examines and assesses the financial credit risk associated with the supply chain and conducts an in-depth analysis of the unique individual risk of the operation mode of the cloud warehouse. The evaluation problem was simultaneously transformed into an optimization problem, and the optimization mechanism of the swarm intelligence optimization algorithm was used as a

reference to determine a reasonable default point and default distance. Additionally, the WOA was selected instead of the swarm intelligence optimization algorithm in enhancing the optimization performance of the conventional WOA and handling this problem more effectively. Based on the optimization capabilities of the enhanced algorithm and the analysis of China's credit risk situation, this study proposes the IWOA-KMV model, which is an improvement on the traditional KMV measurement model. This model is a new credit risk measurement method that can support advancements in existing credit risk measurement practices. In the early stage of listing, small and medium-sized enterprises may have a large number of non-tradable shares and imperfect data, and the value of non-tradable shares cannot be directly observed by the market.

Acknowledgement

This work was sponsored by the Beijing Natural Science Foundation (9222007), Beijing Municipal Social Science Fund Decision-making Consulting Key Project (Grant No. 21JCB059) and the Foundation Science and Technology Innovation Service Capacity Province (Grant No. 19002020217).

References

- [1] Yang, C., Lan, S., Lin, T., Wang, L., Zhuang, Z., Huang, G.Q. (2021). Transforming Hong Kong's warehousing industry with a novel business model: A game-theory analysis, *Robotics and Computer-Integrated Manufacturing*, Vol. 68, No. 2, Article No. 102073, doi: [10.1016/j.rcim.2020.102073](https://doi.org/10.1016/j.rcim.2020.102073).
- [2] Zhang, L., Hu, H., Zhang, D. (2015). A credit risk assessment model based on SVM for small and medium enterprises in supply chain finance, *Financial Innovation*, No. 1, Article No. 14, doi: [10.1186/s40854-015-0014-5](https://doi.org/10.1186/s40854-015-0014-5).
- [3] Huang, X., Sun, J., Zhao, X. (2021). Credit risk assessment of supply chain financing with a grey correlation model: An empirical study on China's home appliance industry, *Complexity*, Vol. 2021, No. 1, Article No. 9981019, doi: [10.1155/2021/9981019](https://doi.org/10.1155/2021/9981019).
- [4] Li, C., Yao, Y., Jiang, M., Zhang, X., Song, L., Zhang, Y., Zhao, B., Liu, J., Yu, Z., Du, X., Ruan, S. (2024). Evolving the whale optimization algorithm: The development and analysis of MISWOA, *Biomimetics*, Vol. 9, No. 10, Article No. 639, doi: [10.3390/biomimetics9100639](https://doi.org/10.3390/biomimetics9100639).
- [5] Dobrescu, R., Mocanu, S., Chenaru, O., Nicolae, M., Florea, G. (2021). Versatile edge gateway for improving manufacturing supply chain management via collaborative networks, *International Journal of Computer Integrated Manufacturing*, Vol. 34, No. 4, 407-421, doi: [10.1080/0951192X.2021.1879401](https://doi.org/10.1080/0951192X.2021.1879401).
- [6] Song, H., Lu, Q., Yu, K., Qian, C. (2019). How do knowledge spillover and access in supply chain network enhance SMEs' credit quality?, *Industrial Management & Data Systems*, Vol. 119 No. 2, 274-291, doi: [10.1108/IMDS-01-2018-0049](https://doi.org/10.1108/IMDS-01-2018-0049).
- [7] Chai, Y., Li, Q. (2021). Research on influencing factors of knowledge sharing in supply chain enterprises under blockchain environment, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 5, 1553-1559, doi: [10.17559/TV-20201116085005](https://doi.org/10.17559/TV-20201116085005).
- [8] Lee, S.Y. (2021). Blockchain-based medical information sharing service architecture, *International Journal of IT-based Public Health Management*, Vol. 8, No. 1, 27-32, doi: [10.21742/IJIPHM.2021.8.1.04](https://doi.org/10.21742/IJIPHM.2021.8.1.04).
- [9] Marza, B., Bratu, R., Șerbu, R., Stan, S., Oprean-Stan, C. (2021). Applying AHP and fuzzy AHP management methods to assess the level of financial and digital inclusion, *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 55, No. 4, 165-182, doi: [10.24818/18423264/55.4.21.11](https://doi.org/10.24818/18423264/55.4.21.11).
- [10] Deng, K. (2022). Research on evaluation of intelligent manufacturing capability and layout superiority of supply chains by big data analysis, *Journal of Global Information Management*, Vol. 30, No. 7, 1-20, doi: [10.4018/JGIM.294903](https://doi.org/10.4018/JGIM.294903).
- [11] Pan, Y.M., Tong, Y.L., Mi, G.X. (2013). The loan pricing study of supply chain finance based on cloud muster warehouse (CMW), *Applied Mechanics and Materials*, Vol. 462-463, 940-943, doi: [10.4028/www.scientific.net/AMM.462-463.940](https://doi.org/10.4028/www.scientific.net/AMM.462-463.940).
- [12] Wang, J., Zhang, H., Luo, H. (2022). Research on the construction of stock portfolios based on multiobjective water cycle algorithm and KMV algorithm, *Applied Soft Computing*, Vol. 115, Article No. 108186, doi: [10.1016/j.asoc.2021.108186](https://doi.org/10.1016/j.asoc.2021.108186).
- [13] Li, D., Na, S., Ding, T., Liu, C. (2021). Credit risk management of P2P network lending, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 4, 1145-1151, doi: [10.17559/TV-20200210110508](https://doi.org/10.17559/TV-20200210110508).
- [14] Malhotra, R., Malhotra, D.K. (2002). Differentiating between good credits and bad credits using neuro-fuzzy systems, *European Journal of Operational Research*, Vol. 136, No. 1, 190-211, doi: [10.1016/S0377-2217\(01\)00052-2](https://doi.org/10.1016/S0377-2217(01)00052-2).
- [15] Zheng, Z.L., Bao, X. (2019). The investment strategy and capacity portfolio optimization in the supply chain with spillover effect based on artificial fish swarm algorithm, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 239-250, doi: [10.14743/apem2019.2.325](https://doi.org/10.14743/apem2019.2.325).
- [16] Esmaeili-Najafabadi, E., Azad, N., Nezhad, M.S.F. (2021). Risk-averse supplier selection and order allocation in the centralized supply chains under disruption risks, *Expert Systems with Applications*, Vol. 175, Article No. 114691, doi: [10.1016/j.eswa.2021.114691](https://doi.org/10.1016/j.eswa.2021.114691).

- [17] Liu, Y., Huang, L. (2020). Supply chain finance credit risk assessment using support vector machine-based ensemble improved with noise elimination, *International Journal of Distributed Sensor Networks*, Vol. 16, No. 1, doi: [10.1177/1550147720903631](https://doi.org/10.1177/1550147720903631).
- [18] Du, M., Zhang, Z. (2023). Modeling selection for credit risk measurement: Based on meta path features, *Tehnički Vjesnik – Technical Gazette*, Vol. 30, No. 2, 545-554, doi: [10.17559/TV-20221107051221](https://doi.org/10.17559/TV-20221107051221).
- [19] Isoaho, K., Gritsenko, D., Mäkelä, E. (2021). Topic modeling and text analysis for qualitative policy research, *Policy Studies Journal*, Vol. 49, No. 1, 300-324, doi: [10.1111/psj.12343](https://doi.org/10.1111/psj.12343).
- [20] Szymkowiak, M., Rhodes-Reese, M. (2021). Examining the expansion of qualitative network models towards integrating multifaceted human dimensions, *ICES Journal of Marine Science*, Vol. 78, No. 7, 2395-2408, doi: [10.1093/icesjms/fsab105](https://doi.org/10.1093/icesjms/fsab105).
- [21] Shengkai, J. (2022). Research on bank credit risk assessment model based on artificial intelligence algorithm, In: *Proceedings of 2022 2nd International Symposium on Artificial Intelligence and its Application on Media (ISAIAM)*, Xi'an, China, 128-134, doi: [10.1109/ISAIAM55748.2022.00032](https://doi.org/10.1109/ISAIAM55748.2022.00032).
- [22] Çallı, B.A., Coşkun, E. (2021). A longitudinal systematic review of credit risk assessment and credit default predictors, *Sage Open*, Vol. 11, No. 4, doi: [10.1177/21582440211061333](https://doi.org/10.1177/21582440211061333).
- [23] Yeh, C.-C., Lin, F., Hsu, C.-Y. (2012). A hybrid KMV model, random forests and rough set theory approach for credit rating, *Knowledge-Based Systems*, Vol. 33, 166-172, doi: [10.1016/j.knosys.2012.04.004](https://doi.org/10.1016/j.knosys.2012.04.004).
- [24] Zhang, Y., Shi, B. (2016). Non-tradable shares pricing and optimal default point based on hybrid KMV models: Evidence from China, *Knowledge-Based Systems*, Vol. 110, 202-209, doi: [10.1016/j.knosys.2016.07.028](https://doi.org/10.1016/j.knosys.2016.07.028).
- [25] Mirjalili, S., Lewis, A. (2016). The whale optimization algorithm, *Advances in Engineering Software*, Vol. 95, 51-67, doi: [10.1016/j.advengsoft.2016.01.008](https://doi.org/10.1016/j.advengsoft.2016.01.008).
- [26] Rana, N., Latiff, M.S.A., Abdulhamid, S.M., Chiroma, H. (2020). Whale optimization algorithm: A systematic review of contemporary applications, modifications and developments, *Neural Computing and Applications*, Vol. 32, 16245-16277, doi: [10.1007/s00521-020-04849-z](https://doi.org/10.1007/s00521-020-04849-z).
- [27] Kaya, E., Kılıç, A., Babaoğlu, I., Babalık, A. (2021). Fuzzy adaptive whale optimization algorithm for numeric optimization, *Malaysian Journal of Computer Science*, Vol. 34, No. 2, 184-198, doi: [10.22452/mjcs.vol34no2.4](https://doi.org/10.22452/mjcs.vol34no2.4).

Calendar of events

- IMTS 2024, International Manufacturing Technology Show, September 9-14, 2024, Chicago, USA.
- 27TH European Conference on Artificial Intelligence, October 19-24, 2024, Santiago de Compostela, Spain.
- 16th EAI International Conference on Simulation Tools and Techniques, December 9-10, 2024, Bratislava, Slovakia.
- International Conference on Smart Materials & Structures, March 27-28, 2025, Berlin, Germany.
- 14th Spring World Congress on Engineering and Technology (SCET 2025), April 19-21, 2025, Guilin, China.
- International Connect & Expo on Material Science and Engineering, April 28-20, 2025, Rome, Italy.
- 2025 Annual Modeling and Simulation Conference (ANNSIM'25), May 26-29, 2025, Madrid, Spain.
- Fifth International Conference on Simulation for Additive Manufacturing (Sim-AM 2025), September 9-11, 2025, Pavia, Italy.

Notes for contributors

General

Articles submitted to the *APEM journal* should be original and unpublished contributions and should not be under consideration for any other publication at the same time. Manuscript should be written in English. Responsibility for the contents of the paper rests upon the authors and not upon the editors or the publisher. The content from published paper in the *APEM journal* may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). For most up-to-date information please see the APEM journal homepage apem-journal.org.

Submission of papers

A submission must include the corresponding author's complete name, affiliation, address, phone and fax numbers, and e-mail address. All papers for consideration by *Advances in Production Engineering & Management* should be submitted by e-mail to the journal Editor-in-Chief:

Miran Brezocnik, Editor-in-Chief
UNIVERSITY OF MARIBOR
Faculty of Mechanical Engineering
Chair of Production Engineering
Smetanova ulica 17, SI – 2000 Maribor
Slovenia, European Union
E-mail: apem@apem-journal.org

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2010* (or higher version) word processor. *Word .docx* format is required. Papers on A4 format, single-spaced, typed in one column, using body text font size of 11 pt, should not exceed 12 pages, including abstract, keywords, body text, figures, tables, acknowledgements (if any), references, and appendices (if any). The title of the paper, authors' names, affiliations and headings of the body text should be in *Calibri* font. Body text, figures and tables captions have to be written in *Cambria* font. Mathematical equations and expressions must be set in *Microsoft Word Equation Editor* and written in *Cambria Math* font. For detail instructions on manuscript preparation please see instruction for authors in the *APEM journal* homepage apem-journal.org.

The review process

Every manuscript submitted for possible publication in the *APEM journal* is first briefly reviewed by the editor for general suitability for the journal. Notification of successful submission is sent. After initial screening, and checking by a special plagiarism detection tool, the manuscript is passed on to at least two referees. A double-blind peer review process ensures the content's validity and relevance. Optionally, authors are invited to suggest up to three well-respected experts in the field discussed in the article who might act as reviewers. The review process can take up to eight weeks on average. Based on the comments of the referees, the editor will take a decision about the paper. The following decisions can be made: accepting the paper, reconsidering the paper after changes, or rejecting the paper. Accepted papers may not be offered elsewhere for publication. The editor may, in some circumstances, vary this process at his discretion.

Proofs

Proofs will be sent to the corresponding author and should be returned within 3 days of receipt. Corrections should be restricted to typesetting errors and minor changes.

Offprints

An e-offprint, i.e., a PDF version of the published article, will be sent by e-mail to the corresponding author. Additionally, one complete copy of the journal will be sent free of charge to the corresponding author of the published article.

APEM

journal

Advances in Production Engineering & Management

Chair of Production Engineering (CPE)
University of Maribor
APEM homepage: apem-journal.org

Volume 19 | Number 3 | September 2024 | pp 297-410

Contents

Scope and topics	300
Dynamic scheduling for manufacturing workshops using Digital Twins, Competitive Particle Swarm Optimization, and Siamese Neural Networks Weng, L.L.	301
Optimization of cold chain multimodal transportation routes considering carbon emissions under hybrid uncertainties Hou, D.N.; Liu, S.C.	315
Enhancing workplace safety and ergonomics with motion capture systems: Present state and a case study Vujica Herzog, N.; Buchmeister, B.; Jaehyun, P.; Kaya, Ö.	333
Machine learning for enhancing manufacturing quality control in ultrasonic nondestructive testing: A Wavelet Neural Network and Genetic Algorithm approach Song, W.T.; Huo, L.	347
Optimizing obstacle avoidance path planning for intelligent mobile robots in multi-obstacle environments Zhang, H.Q.	358
Modelling and optimization of dimensional accuracy and surface roughness in dry turning of Inconel 625 alloy Vukelic, D.; Milosevic, A.; Ivanov, V.; Kocovic, V.; Santosi, Z.; Sokac, M.; Simunovic, G.	371
Enhancing calibration accuracy with laser interferometry for high-resolution measuring systems Lipus, L.C.; Acko, B.; Klobucar, R.	386
Improved Whale Optimization Algorithm for supply chain financial risk assessment of cloud warehouse platform Zhang, H.; Guo, Y.W.; Hou, Y.; Tang, L.; Deveci, M.	395
Calendar of events	408
Notes for contributors	409

Published by CPE, University of Maribor



apem-journal.org