OPEN ACCESS
IOP Publishing | ISECAEP
Journal of Reliability Science and Engineering
J. Reliab. Sci. Eng. 1 (2025) 025006 (10pp)
https://doi.org/10.1088/3050-2454/ade0c2

Research on the evaluation of machine tool service performance based on precision retention and reliability

Guobin Yan[1], Qiang Cheng[1,*], Yuegang Hu[1,2] and Guodong Sun[3]

[1] Machinery Industry Key Laboratory of Heavy Machine Tool Digital Design and Testing Technology, Beijing University of Technology, Beijing 100124, People's Republic of China
[2] Ya'an Polytechnic College, Ya'an 625000, People's Republic of China
[3] Shandong Institute of Scientific and Technical Information, Jinan 250101, People's Republic of China

E-mail: chengqiang@bjut.edu.cn

Received 28 March 2025, revised 7 May 2025
Accepted for publication 3 June 2025
Published 24 June 2025

Abstract

The precision retention and reliability of CNC machine tools are critical indicators affecting machining quality and production efficiency. To address the limitations of existing evaluation methods that fail to comprehensively consider expert authority and indicator weights, this paper proposes a performance evaluation system for machine tools in service, integrating the entropy weight method, analytic hierarchy process, and fuzzy comprehensive evaluation. An evaluation model is constructed with two first-level indicators, six second-level indicators, and third-level indicators, incorporating expert authority to adjust weights, and the performance of five domestic machine tools is ranked. The results demonstrate that Machine Tool A exhibits the best overall performance, but domestic machine tools generally have shortcomings in accuracy retention and adaptability, primarily due to insufficient material processes for key components and a lack of dynamic compensation technology. This study provides a quantitative basis for the optimal design and procurement decisions of machine tools.

Keywords: machine tools, reliability, precision retention, evaluation method

1. Introduction

The service performance of machine tools refers to their ability to operate stably and meet expected functional requirements in real-world working environments over extended periods. It focuses on the machine tools' capacity to maintain stability, reliability, and durability during prolonged usage, as well as their adaptability to various practical operating conditions. With the rapid development of Industry 4.0 and intelligent manufacturing, the performance requirements for CNC machine tools are increasingly demanding. However, in actual production processes, the performance of CNC machine tools gradually degrades over time, leading to issues such as reduced machining accuracy, increased failure rates, and higher maintenance costs. Therefore, establishing a scientific and comprehensive evaluation system for the service reliability of CNC machine tools is of great significance for optimizing machine tool usage, extending equipment lifespan, and reducing production costs.

In the 1950s, reliability research began to incorporate statistical and mathematical modeling methods, applying mathematical models to analyze and predict system reliability. Tools such as failure mode and effects analysis (FMEA) and fault tree analysis were developed during this period. By

* Author to whom any correspondence should be addressed.

the 1970s, researchers at the 'Scientific Research Institute for Metal Cutting Machine Tools' had initiated specialized studies on the reliability of machine tools, publishing classic works such as Accuracy and Reliability of CNC Machine Tools. Reliability is one of the critical attributes of CNC machine tools and holds a pivotal position in their design. Reliability design encompasses various aspects, including reliability allocation [1], reliability assessment [2], remaining useful life prediction [3], and risk analysis [4].

In terms of reliability research methods, Qiang *et al* combined trapezoidal fuzzy numbers with the maximum entropy-ordered weighted operator to handle fuzzy information in reliability allocation [5]. Congbin Yang *et al* proposed a reliability improvement method based on the coupling of spatiotemporal factors for fault hierarchical propagation. Considering the coupled effects of component comprehensive importance, fault tolerance, and failure modes on the machine tool system, they established a spatiotemporal fault hierarchical propagation topology-directed graph model [6]. Chang used an ordered weighted averaging tree to achieve reliability allocation for CNC machine tools, reducing the difficulty of reliability allocation [7]. Li *et al* considered the weight relationships among risk factors and conducted FMEA for offshore wind turbines, concluding that the support structure and energy production system are more critical than other systems [8]. Additionally, many researchers have attempted to use fuzzy numbers to address the shortcomings of traditional FMEA methods in accurately expressing fuzzy information [9, 10]. Sriramdas *et al* proposed a linear programming-based approximation method for trapezoidal fuzzy numbers and applied it to reliability allocation [11]. Feng *et al* innovatively introduced factors such as carbon emissions and module types into reliability allocation and used the maximum entropy ordered weighted operator to calculate reliability allocation results [12]. Zhang and Liao applied the fuzzy analytic hierarchy process (AHP) method to reliability allocation for direct-drive gear hobbing machines [13]. Bai *et al* used AHP to calculate the weights of influencing factors and employed fuzzy fault trees to compute subsystem failure probabilities [14]. Du *et al* evaluated the importance of subsystems based on structural complexity, technical maturity, failure severity, maintenance difficulty, and service condition severity, and determined the weights for machine tool subsystem reliability reallocation using AHP [15]. The results obtained from the AHP method belong to proportional information, where data relationships are proportional. Using AHP results as weights enhances the rationality of data utilization. Peng *et al* proposed a multi-objective optimization tolerance allocation design method for machine tools based on NSGA-II algorithm and thermal characteristic analysis [16]. Wei *et al* developed a series of new acquisition functions with closed-form expressions to accelerate approximate Bayesian quadrature, thereby solving the BMU problem with the required accuracy [17].

In the field of machine tool evaluation, Huang, XQ proposed an evaluation method based on gray clustering analysis and fuzzy comprehensive evaluation. They classified the health status levels of in-service CNC machine tools and constructed a performance indicator system for CNC machine tools [18]. Yang *et al* combined fuzzy theory with comprehensive evaluation theory to build a fuzzy comprehensive evaluation mathematical model for CNC machine tools [19]. Cui and Li established a service quality evaluation indicator system for machine tools. The evaluation indicator system includes five aspects: processing cost, processing time, processing accuracy, green attributes, and reliability. They used the least squares method to establish an optimization model for determining indicator weights based on subjective and objective weighting [20]. Tian *et al* proposed a fuzzy comprehensive evaluation method based on AHP to address the diversity and fuzziness of machine tool selection in CAPP systems [21]. Kumar *et al* conducted a comprehensive literature review on the application of entropy weight method (EWM) in machining operations, demonstrating its implementation process and methodological development through a representative case study [22]. Li *et al* proposed a safety assessment method for CNC machine tools based on EWM, and verified its application on the XKA28 heavy-duty CNC gantry milling machine [23]. Li established an intelligent scheduling model of Rail Guided Vehicle in intelligent machining system using EWM and expert ranking method [24] Chen *et al* proposed a generalized reliability evaluation method for CNC machine tools based on improved entropy weight extensible matter-element model. Introducing contrast intensity and conflict intensity to consider the interrelationships between indicators to improve the entropy weighting method [25]. Kumar *et al* implemented a multi-objective optimization of turning parameters for EN 353 alloy steel by integrating EWM with TOPSIS. Their study validated the objective effectiveness of EWM in trade-off analysis among energy consumption, surface roughness, and material removal rate, demonstrating its superior performance over both equal-weight method and AHP in machining optimization decision-making [26].

A comprehensive literature review and analysis revealed that although AHP, EWM and fuzzy comprehensive evaluation have been widely applied in machine tool reliability research, significant limitations were identified: (1) the evaluation coverage was found to be inadequate, with existing studies predominantly focusing on single dimensions while lacking synergistic analysis of comprehensive indicators including precision retention, reliability, environmental performance, and adaptability; (2) expert authority quantification was absent in traditional methods, where AHP's subjective weighting was directly applied without quantifying differences in expert qualifications, resulting in biased weight allocation. The unique contributions of this study are demonstrated through three key aspects: First, an innovative integration of EWM and AHP was developed, where expert authority was quantified using EWM to calibrate AHP's subjective weights, leading to significantly enhanced evaluation objectivity. Second, a lifecycle-oriented indicator expansion was achieved through the construction of a comprehensive evaluation system comprising 26 tertiary indicators, with Industry 4.0-critical dimensions such as 'environmental adaptability' and 'maintenance convenience' being incorporated to address deficiencies in previous systems. Third, dynamic coupling analysis was conducted through correlational modeling between precision lifespan and MTBF, through which the interaction mechanisms between precision and reliability during long-term service were revealed, providing novel insights for maintenance strategy optimization.

2. Manufacturing maturity evaluation methods

2.1. Establishing an indicator evaluation system

To establish a scientifically validated evaluation index system, eighteen domain experts were rigorously selected to participate in this study, consisting of both professors specializing in CNC machine tool research from leading academic institutions and senior engineers with practical industry experience from prominent machine tool manufacturers. The methodological workflow was implemented through the following standardized procedure: Initially, relevant industrial standards were comprehensively reviewed to develop a preliminary Delphi questionnaire, with specific reference to the following normative documents: (1) ISO230-1:1996 Test code for machine tools—Part 1: Geometric accuracy of machines operating under no-load or finishing conditions. (2) ISO 14955-3:2020 Machine tools—Environmental evaluation of machine tools—Part 3: Principles for testing metal-cutting machine tools with respect to energy efficiency. (3) Report on the Development of China's Intelligent Manufacturing Industry for the Year 2023–2024. (4) GBT23567.2-2018- Reliability assessment of CNC machine tools—Part 2- Machining centers. Subsequently, a pre-test was conducted with five industry experts, and their feedback was compiled to form the final expert consultation questionnaire. The questionnaire consisted of:

(1) Questionnaire instructions, which introduced the research objectives, relevant background information, and guidelines for completing the form.
(2) The main body of the questionnaire, included two first-level indicators, five second-level indicators, and thirty third-level indicators for the machine tool service performance index system. The questionnaire also featured a section for modification suggestions, allowing experts to supplement or revise the entries.
(3) A basic information survey form for experts, covering their educational background, years of experience in machine tool manufacturing, number of participations in the formulation or revision of technical standards in the machine tool industry, number of authorized patents related to machine tools, income from technology transfers, number of new technologies developed, number of high-quality papers published, number of participations in municipal or higher-level projects, number of municipal or higher-level awards and honors, and a self-assessment score of expert authority.

The research team distributed and collected two consultation questionnaires through on-site distribution, remote video, and WeChat. A performance evaluation system for machine tool service based on precision retention and reliability has been established. The first-level indicators are: A1-Precision Retention, B1-Reliability. Second-level and Third-level indicators are shown in table 1.

Among them, machine tool accuracy fluctuation refers to the degree of dynamic variation in accuracy metrics within the specified tolerance range. Accuracy retentivity degree measures the extent to which accuracy metrics remain at their initial state over time. Accuracy lifespan is the total operational duration during which all accuracy metrics remain within the required tolerance limits. Accuracy fluctuation reflects short-term dynamic stability during machining. Accuracy retentivity degree indicates the rate of long-term accuracy degradation. Accuracy lifespan represents the total operational time before accuracy failure occurs. Machine tool lifespan refers to the total service time or workload from initial operation until the machine completely loses its fundamental functionality. In most cases, the accuracy lifespan ends before the machine tool lifespan—meaning that even after the accuracy lifespan is exceeded, the machine may still be used for machining parts with lower precision requirements until final decommissioning. Machine tool compatibility describes the capability of a machine to accommodate varying machining tasks or technological requirements, including compatibility and flexibility with different workpieces, processes, and production modes (e.g. batch or single-piece production). A highly adaptable machine can respond more efficiently to diverse production demands, reducing setup time and costs.

B131-Processing energy consumption ratio, standby energy consumption: According to the ISO 14955-1 standard, calculate the electrical energy consumption (kWh) per unit material removal amount (cm^3), and the data is collected in real time through the built-in electric meter and cutting volume monitoring system of the machine tool. B132-Carbon emissions during the service life of machine tools: Based on the IPCC National Greenhouse Gas Inventory Guidelines, energy consumption is converted into CO_2 equivalent (kg), distinguishing between usage stages (processing/standby) and energy types (such as China's power grid carbon emission factor of 0.583 kg CO_2 kWh^{-1}). B133-Water resource consumption: Record the average annual replenishment amount (liters) of the coolant circulation system, converted to freshwater consumption equivalent according to ISO 14046 standard, including indirect water consumption (such as electricity production water consumption).

2.2. EWM

The EWM is a method that calculates objective weights through information entropy without subjective judgment and is suitable for data-driven decision-making. Entropy reflects the disorder of indicator data. If the entropy value of a certain indicator is high, it indicates that the indicator has significant differences in different samples and contains more effective information, so it should be given higher weight. On the contrary, if the entropy value is small, the weight decreases. AHP relies on subjective judgment from experts and calculates weights by constructing a judgment matrix, making it suitable for complex decision-making problems with multiple criteria

Table 1. Evaluation indicator system.

First-level	Second-level	Third-level indicators layer	SN
A1	A11-Machine tool accuracy	A111-Machine tool accuracy fluctuation	1
		A112-Accuracy retentivity degree	2
		A113-Accuracy lifespan of machine tools	3
	A12-Precision of processed parts	A121-Consistency of dimensional accuracy of processed parts	4
		A122-Consistency of shape accuracy of processed parts	5
		A123-Consistency of surface roughness of processed parts	6
B1	B11-Fault	B111-Mean time between failures (MTBF)	7
		B112-Mean time to repair (MTTR)	8
		B113-Fault rate	9
		B114-Availability	10
		B115-Time to first failure (TTFF)	11
		B116-Machine tool lifespan	12
		B117-Severity of malfunction	13
	B12-Dynamic and static characteristics and efficiency	B121-Stiffness	14
		B122-Vibration characteristics	15
		B123-Structural stability	16
		B124-Cutting speed	17
		B125-Feed speed	18
		B126-Material removal rate	19
		B127-Compatibility	20
	B13-Environment protection	B131-Processing energy consumption ratio, standby energy consumption	21
		B132-Carbon emissions during the service life of machine tools	22
		B133-Water resource consumption	23
	B14-Guarantee	B141-Maintenance convenience and maintenance cost	24
		B142-Safety	25
		B143-Environmental adaptability	26

and levels. Because the basic information of experts can be expressed through data, which is more objective than AHP, the EWM is used to determine the authority of experts. The application process of the EWM is as follows:

(1) Expert authoritative evaluation form constructed.
(2) Eliminate the influence of dimensionality and standardize the data based on the original data matrix. There are n samples and m indicators, and the original data matrix is:

$$X = (x_{ij})_{n\times m}. \quad (1)$$

(3) For positive indicators (the larger the better):

$$r_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}. \quad (2)$$

Obtain standardized matrix:

$$R = (r_{ij})_{n\times m}. \quad (3)$$

(4) After standardizing the data, a standardized data table can be obtained. Subsequently, the proportion P_{ij} of the j indicator in the evaluation of the i expert can be calculated,

$$P_{ij} = \frac{r_{ij}}{\sum_{i=1}^{n} r_{ij}}. \quad (4)$$

(5) Based on the calculated proportions, the information entropy of each indicator can be derived,

$$e_j = -\frac{1}{\ln(n)}\sum_{i=1}^{n} p_{ij}\ln(p_{ij}). \quad (5)$$

In this context, e_j represents the entropy value of the j indicator, and $\ln(n)$ is utilized for normalization purposes.
(6) The weight w_j for each indicator is calculated, which represents the entropy weight of the j indicator:

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{m}(1 - e_j)} \quad (6)$$

(7) Based on the aforementioned weights and the standardized data, the authority score for each expert can be derived,

$$S_i = \sum_{j=1}^{m} w_j \cdot r_{ij}. \tag{7}$$

2.3. AHP

Due to the fact that most of the evaluation indicators in the established evaluation system cannot be quantified, weights need to be calculated through the subjective assignment method. AHP is currently the most widely used subjective assignment method. This article aims to evaluate the service performance of machine tools and assign values based on the hierarchical structure model of the indicators in table 1.

(1) Pairwise discriminant matrices are established based on expert X's scoring:

$$A = \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{m1} & \cdots & B_{mn} \end{bmatrix}, \tag{8}$$

where B_{mn} represents the degree of importance of the m indicator compared to the n indicator at the same level. When the evaluation matrix's consistency is poor, experts must renegotiate and evaluate it.

(2) Hierarchical single sorting and consistency verification. Column vectors are normalized to obtain

$$\overline{B_{mn}} = \frac{B_{mn}}{\sum_{e=1}^{f} B_{en}} \; m,n = 1,2,\cdots,f. \tag{9}$$

All columns in the same row are added up as follows:

$$\overline{w_m} = \sum_{n=1}^{f} \overline{B_{mn}} \; n = 1,2,\cdots,f. \tag{10}$$

The weight is obtained as follows:

$$w_m = \frac{\overline{w_m}}{f}. \tag{11}$$

The maximum eigenvalue is calculated as follows:

$$\lambda_{\max} = \frac{1}{f} \sum_{m=1}^{f} \frac{(Aw)_m}{w_m} \frac{1}{2}, \tag{12}$$

where $(Aw)_m$ represents the m component of vector Aw.

When the evaluation matrix provided by experts exhibits poor consistency, the experts must renegotiate and reassess the evaluations. The procedure for this process is outlined below. Initially, the maximum eigenvalue of each pairwise comparison matrix is computed. Subsequently, consistency testing is conducted using the random consistency index (RI), the consistency index (CI), and the consistency ratio (CR). If the calculated CR value is less than 0.1, the consistency test is considered satisfactory. However, if the CR value exceeds this threshold, the data must be adjusted, and the evaluation matrix reconstructed. The formulas for calculating the consistency index (CI) and the random consistency ratio (RI) are as follows:

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1}, \tag{13}$$

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}}. \tag{14}$$

The value of RI was obtained using the Monte Carlo method, and the average random consistency index was constructed according to table 2.

Table 2. Consistency index RI.

Multi-order evaluation matrix RI value									
n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45

2.4. Fuzzy comprehensive evaluation method

The fuzzy comprehensive evaluation method is an evaluation approach grounded in fuzzy mathematical theory, which is particularly suitable for addressing multi-factor, fuzzy, and uncertain problems. In the context of machine tool service performance evaluation, this method proves effective in quantifying evaluation content due to the involvement of both explicit and implicit factors, as well as the inherent fuzziness of certain activities or phenomena. The key components of this method are as follows:

(1) Factor set: this refers to the set of factor indicators used in the fuzzy comprehensive evaluation, specifically the second-level indicators of the 'Machine Tool Service Performance Evaluation System.'
(2) Evaluation set: this represents the degree of quality or performance for each evaluation factor, forming a collection of various assessment levels or results.
(3) Weight set: the weight values indicate the relative importance of each indicator within the overall evaluation system. The weight set is a collection of these weight values, specifically the combined weights of the second-level indicators.
(4) Fuzzy relation matrix: the fuzzy relation matrix is used to quantify qualitative indicators. Through the fuzzy statistical method, evaluators assign scores to each factor based on the evaluation set, determining the degree of membership of each factor to different evaluation levels. By combining these membership degrees, the fuzzy rela-

Table 3. Comparison of fuzzy operators.

Characteristic	Reflect the weight effect	Comprehensive degree	Utilize R's information
M (•,+)	Obvious	Strong	Sufficient
M (•,∨)	Obvious	Weak	Insufficiency
M (∧,+)	Obvious	Strong	Relatively sufficient
M (∧,∨)	Not obvious	Weak	Insufficiency

tion matrix is constructed. Its mathematical model can be expressed as:

$$R_{h\times j}=\begin{bmatrix} R_1 \\ R_2 \\ \cdots \\ R_N \end{bmatrix}=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{h1} & r_{h2} & \cdots & r_{hj} \end{bmatrix}=(r_{ik})_{h\times j}, \tag{15}$$

where $(r_{i1} \quad r_{i2} \quad \cdots \quad r_{ij})$ represents the evaluation set for a single factor indicator u_i $(i=1, \quad 2, \quad \cdots, \quad n)$, and R_{hj} denotes the collection of all evaluation sets for individual factor indicators.

(5) The fuzzy relation matrix is calculated as follows:

$$B=A\times R_{h\times j}=A\times\begin{bmatrix} R_1 \\ R_2 \\ \cdots \\ R_N \end{bmatrix}=A\times\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{h1} & r_{h2} & \cdots & r_{hj} \end{bmatrix} \tag{16}$$

where 'A' is the combined weight set of the three-level evaluation indicators and 'B' is the assessment set. This set is calculated by allocating the combined weights of each evaluation indicator and the assessment matrix.

In the application of fuzzy comprehensive evaluation for machine tool assessment, selecting an appropriate fuzzy operator is crucial to maintain the rationality of the evaluation process and strengthen the logical coherence of computations. Through comparative analysis of various common fuzzy operators—examining their information processing mechanisms, weight utilization, and comprehensiveness—the most suitable operator type is identified to maximize the incorporation of evaluation information. The distinct features of these operators are summarized in table 3.

Here, the symbol '+' denotes addition, '•' signifies multiplication, while '∧' and '∨' represent the minimum and maximum operations, respectively. The mathematical model of the fuzzy operator M (·, +) is expressed by equation (17).

Unlike other operators, this model employs conventional real-number multiplication and addition for matrix synthesis. It achieves a balanced integration of all weighted indicators, ensuring that the influence of each factor is thoroughly considered in the evaluation. This approach enhances the robustness and inclusiveness of the assessment by systematically accounting for every contributing element,

$$b_{d,k}=\sum_{i=1}^{j}a_{d,i}r_{d,ik}, k=1,2,\cdots,n, \tag{17}$$

where $a_{d,i}$ represents the measure of the influence of factor $u_{d,i}$ within the evaluation factors, $b_{d,k}$ denotes the membership degree of the evaluation grade $v_{d,k}$ to the fuzzy operator set B_d, and $r_{d,ik}$ indicates the membership degree of the evaluation object to the evaluation grade $u_{d,i}$ when $v_{d,k}$ is considered independently.

3. Results of performance evaluation of machine tool service

3.1. Expert authority evaluation results based on the EWM

This study selected five experts from an 18-member expert database to participate in weighted scoring of the evaluation indicators. The selection process employed a rigorous two-stage screening approach (18 → 10 preliminary candidates → 5 final experts) based on the following criteria to ensure authority and representativeness: Industry Experience (ensuring extensive practical expertise): Experience in machine tool manufacturing/utilization must exceed 8 years, or exceed 5 years with outstanding achievements, and leadership tenure in a machine tool enterprise must be greater than 3 years.

Technical Achievements (reflecting innovation capability): Ownership of more than 5 machine tool-related patents is required, with technology transfer revenue reaching at least 300 000 USD. Academic Contributions (demonstrating theoretical proficiency): Publication of more than 3 high-quality research papers and participation in at least 5 municipal-level or higher research projects are mandatory.

Industry Influence (indicating industry recognition): Prior involvement in technical standard formulation and receipt of municipal-level or higher awards related to the industry are essential. Basic information of five experts is shown in table 4.

MATLAB software is used to perform calculations using the above methods. Then, the results are normalized to obtain expert authority weights: (0.122, 0.163, 0.214, 0.241, 0.260).

3.2. Weight calculation results based on AHP

A nine-level numerical scale layer is used in this article to illustrate the importance level between indicators, as shown in table 5.

The scoring of indicators at all levels was conducted by five selected experts. The evaluation forms were distributed and collected through on-site delivery, WeChat, and other methods to aggregate the scores provided by each expert. Taking the scoring method of Expert A (one of the five experts) for the four indicators B11, B12, B13, and B14 as an example, the scores are shown in table 6.

Table 4. Expert information table.

Expert number	1	2	3	4	5
Years of machine tool manufacturing and usage	9	11	12	13	21
Number of technical standards participated in formulating	4	1	3	3	2
Revenue from technology transfer (in USD 10 000)	30	45	92	101	97
Years of serving as a leader in machine tool enterprises	4	9	9	11	13
Number of machine tool-related patents (top 3 ranking)	8	15	17	16	30

Table 5. Scale of relative importance ratio.

Comparison of indicators	Extremely important	Very important	Important	Slightly important	Equally important	Slightly unequal	Not important
Value	9	7	5	3	1	1/3	1/5
Comparison of indicators	Very unimportant	Extremely unimportant	Take 8, 6, 4, 2, 1/2, 1/4, 1/6, and 1/8 as the middle values of the evaluation values				
Value	1/7	1/9					

Table 6. Scoring situation.

Expert A's rating	B11	B12	B13	B14
B11	1	2	9	7
B12	1/2	1	5	3
B13	1/9	1/5	1	1/2
B14	1/7	1/3	2	1

Based on the AHP method, the calculations were performed using MATLAB software. As shown in table 2, when $n = 4$, the random consistency index (RI) was 0.9, resulting in a consistency ratio (CR) of 0.073, which is less than 0.1. The consistency test was thus passed, confirming the reliability of the weight calculation results and obtain the weight values as follows: a_{ij} (0.570, 0.280, 0.055, 0.095). Each expert can apply the above methods to determine the weight assignment of each indicator.

Using the aforementioned method, the weight assignments for each indicator provided by the five experts can be calculated. These weights are subsequently multiplied by the respective authority coefficients of the experts. Following normalization, the final comprehensive weights for each indicator are derived as follows:

$$\omega_i = \sum_{k=1}^{8} c_k \omega_i^k, \tag{18}$$

where c_k represents the expert authority coefficient (satisfying $\sum_{k=1}^{8} c_k = 1$) and i denotes the number of indicators at each level. The weights for each indicator after normalizing ω are shown in figure 1.

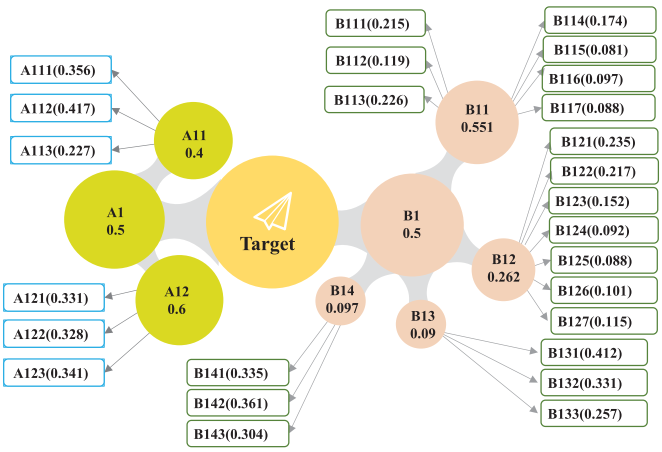

Figure 1. Related weights of each indicator.

The comprehensive weight values of each three-level indicator are shown in table 7.

3.3. Evaluation results of machine tool service performance based on the fuzzy comprehensive evaluation method

A scoring panel consisting of 18 experts was established to evaluate five domestic machine tools (A, B, C, D, and E) of the same specification but from different manufacturers. The evaluation was conducted using a 100-point scale, and the average score from the 18 experts was taken as the final score for each machine tool under a specific indicator. The scoring results are presented in figure 2.

Based on the fuzzy comprehensive evaluation method, the scoring data were calculated using MATLAB software. The scores obtained for each machine tool are (75.903, 71.906,

Table 7. Comprehensive weight values of each three-level indicator.

Indicators	Comprehensive weight	Indicators	Comprehensive weight
A111	0.071	B121	0.031
A112	0.083	B122	0.028
A113	0.045	B123	0.02
A121	0.099	B124	0.012
A122	0.098	B125	0.012
A123	0.102	B126	0.013
B111	0.059	B127	0.015
B112	0.033	B131	0.019
B113	0.062	B132	0.015
B114	0.048	B133	0.012
B115	0.022	B141	0.016
B116	0.027	B142	0.018
B117	0.024	B143	0.015

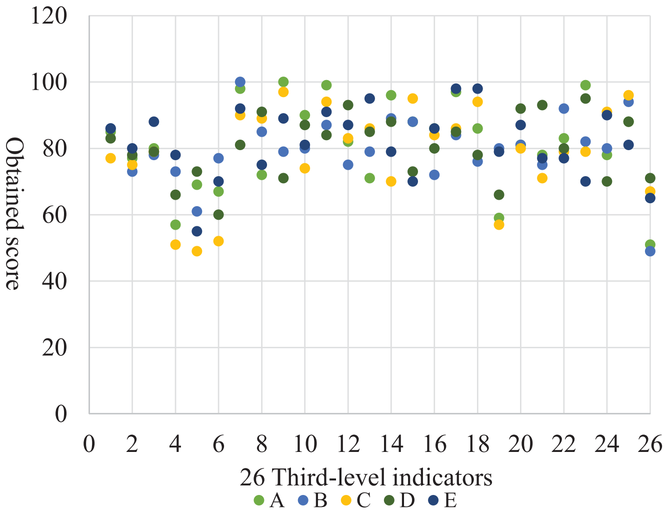

Figure 2. Machine tool service performance scores.

75.148, 74.558, 75.806). From these results, the ranking of the service performance of the five machine tools is as follows: Machine Tool A outperforms Machine Tool E, Machine Tool E outperforms Machine Tool C, Machine Tool C outperforms Machine Tool D, and Machine Tool B has the poorest service performance.

3.4. Weight sensitivity analysis and uncertainty quantification

The robustness of the evaluation results was verified using two methods: the single-index perturbation method and the Monte Carlo method.

Single-index perturbation method: the weights of first-level indicators (accuracy retention A1, reliability B1) were sequentially perturbed by ±10%, and the resulting changes in machine tool rankings were observed. It was found that when the weight of reliability was increased from 0.523 to 0.623, the ranking of Machine Tool E dropped from 2nd to 4th (due to its lower MTBF but higher maintenance convenience score), while Machine Tool A remained in 1st place (owing to its balanced performance in both failure rate and accuracy stability).

Monte Carlo method: Assuming a ±15% random error (uniform distribution) in expert scoring, 1000 random weight combinations were generated, and the probability distribution of each machine tool's ranking was statistically analyzed. The analysis revealed that Machine Tool A maintained its 1st-place ranking in 89.7% of simulations, while Machine Tool B ranked last in 72.3% of cases. When the weight of B11 (failure-related indicators) exceeded 35%, the ranking of Machine Tool E significantly declined, indicating that failure-related metrics were sensitive factors affecting the ranking. The 95% confidence interval for the final scores was ±3.5 points (e.g. Machine A: 75.9 ± 3.5), though no reversal in ranking order was observed.

4. Result analysis and improvement suggestions

Although the weights of each indicator have been determined, the application of this evaluation model must still be adapted to specific contexts. For instance, the environmental performance indicator (B13) is assigned a relatively low weight in the current model. However, if national policies impose mandatory carbon emission reduction requirements (e.g. a carbon tax exceeding 50 USD/ton), the weight of the environmental performance (B13) indicator should be increased accordingly. Similarly, in regions with water scarcity (e.g. Middle Eastern countries), the weight of the water resource efficiency sub-indicator (B133) would need to be adjusted upward to reflect local operational priorities.

Although Machine Tool A and E received similar overall scores, notable differences exist in key performance indicators. Machine Tool A demonstrated superior reliability with a lower failure rate and longer mean time between failures. However, it showed poorer performance in fault severity. During a five-year service period, Machine Tool A experienced only 5 failures, including one critical incident caused by design flaws in the cooling system and inadequate sealing of key components. This resulted in weeks of downtime, requiring replacement of both the cooling system and a servo motor. In contrast, Machine Tool E encountered dozens of easily repairable failures during the same period, contributing to its higher score in 'Maintenance convenience and maintenance cost'. Our investigation revealed that improper operation and negligence by Machine Tool E's users were significant contributing factors to its higher failure frequency. These findings were formally communicated to the enterprise operating Machine Tool E for corrective action.

As shown in figure 2, the five machine tools generally exhibit low scores in accuracy retention, adaptability, and environmental adaptability. The main reasons are as follows:

(1) The materials and manufacturing processes of key components in domestic machine tools are relatively inferior, leading to issues such as wear and deformation after prolonged use. For example, the performance of domestically produced guide rails and lead screws is inferior to those produced by THK.
(2) The control systems and software of domestic machine tools lack sufficient compatibility, making it difficult to

Table 8. Comparison of some indicators.

Indicator	Machine Tool A	DMG	Machine Tool A
MTBF (h)	1200	1800	−33%
Accuracy retentivity degree (μm year^{-1})	1.5	1.2	+25%
Maintenance cost (USD year^{-1})	8500	6000	+42%

seamlessly integrate with diverse processing requirements or third-party equipment.

(3) Compared to foreign products, the bed, column, and other basic components of domestic machine tools have shortcomings in material properties. First, they exhibit more significant thermal expansion and contraction under ambient temperature variations, leading to increased machining errors. Second, insufficient aging treatment results in greater creep deformation due to residual stress release.
(4) Foreign machine tools employ real-time sensor feedback and algorithmic compensation to correct errors caused by thermal deformation and vibration, whereas domestic models mostly rely on static calibration, resulting in poorer accuracy retention under complex working conditions.

A comparative analysis of selected performance indicators between domestic Machine Tool A and its counterpart model from DMG is presented in table 8.

Suggested improvement measures:

(1) Intensify the research and development efforts of domestic rolling components, adopt alloy materials and nano ceramic coating technology, optimize the heat treatment technology of rolling components, and adopt cryogenic treatment process.
(2) Deploy an adaptive machining system driven by digital twins, achieve self-optimization of process parameters, configure HSM high-speed machining interface modules, and be compatible with mainstream CAM software.
(3) Before assembling the machine tool, a composite process of vibration aging and thermal aging is adopted to fully release residual stress, integrate PID temperature control oil circuit, and build the working area into a constant temperature workshop.
(4) Develop a multi physics coupling compensation algorithm based on deep learning and deploy sensors with higher sensitivity.

5. Conclusion and prospect

This study focuses on machine tools as the research subject and establishes a scientifically sound and reasonable evaluation index system for assessing the service performance of machine tools based on accuracy retention and reliability. By combining the EWM and the AHP, the weights of each evaluation indicator were determined. The fuzzy comprehensive evaluation method was then applied to identify the machine tool with the best service performance. To some extent, this study addresses the shortcomings in the evaluation of machine tool service performance. The research presented in this paper can help machine tool enterprises identify their technical bottlenecks, promote innovation and upgrading of machine tool technology, and provide scientific guidance for machine tool users in equipment procurement and renewal decisions.

Future research will focus on the following aspects:

(1) Optimization and expansion of the evaluation index system: although the current evaluation index system covers multiple dimensions of accuracy retention and reliability, with the rapid development of intelligent manufacturing and Industry 4.0 technologies, future studies should incorporate emerging indicators such as intelligent diagnostics and data interaction capabilities to provide a more comprehensive assessment of machine tool performance.
(2) Development of a dynamic evaluation model: the performance of machine tools dynamically changes over their service life. Future research should explore real-time data-driven evaluation models, integrating sensor technology and the Internet of Things to enable online monitoring and dynamic assessment of machine tool performance.
(3) Enhanced application and promotion of the evaluation method: the proposed evaluation method can be extended to other types of machine tools or industrial equipment to verify its generalizability. Additionally, supporting software tools or platforms should be developed to facilitate practical implementation in industrial production, aiding enterprises in equipment evaluation and decision-making.

These efforts will contribute to refining the evaluation framework, improving its adaptability to technological advancements, and promoting its practical application in the industry.

References

[1] Jin T and Abedin A 2025 Co-allocation of reliability-redundancy, spare parts and spacecraft trajectory for lunar habitation *J. Reliab. Sci. Eng.* **1** 015005
[2] Xu J, Yan Q, Pei Y, Liu Z, Cheng Q, Chu H and Zhang T 2024 A statistical evaluation method based on fuzzy failure data for multi-state equipment reliability *Mathematics* **12** 1414
[3] Zhu Y, Cheng J, Liu Z, Zou X, Wang Z, Cheng Q, Xu H, Wang Y and Tao F 2024 Remaining useful life prediction approach based on data model fusion: an application in rolling bearings *IEEE Sens. J.* **24** 42230–44
[4] Compare M and Zio E 2025 Opportunities and risks of artificial intelligence for industry 5.0 in the context of reliability and maintenance engineering *J. Reliab. Sci. Eng.* **1** 015001
[5] Qiang C *et al* 2019 Reliability allocation method based on maximum entropy ordered weighted average and hesitant fuzzy Linguistic term set *Int. J. Fuzzy Syst.* **37** 7991–8004

[6] Yang C, Wang Y, Yan J, Liu Z and Zhang T 2025 A fault hierarchical propagation reliability improvement method for CNC machine tools based on spatiotemporal factors coupling *Reliab. Eng. Syst. Saf.* **255** 110672
[7] Chang K-H 2016 A novel reliability allocation approach using the OWA tree and soft set *Ann. Oper. Res.* **244** 3–22
[8] Li H, Teixeira A P and Guedes Soares C 2022 An improved failure mode and effect analysis of floating offshore wind turbines *J. Mar. Sci. Eng.* **10** 1616
[9] Yuan Y and Tang Y 2022 Fusion of expert uncertain assessment in FMEA based on the negation of basic probability assignment and evidence distance *Sci. Rep.* **12** 8424
[10] Reda H and Dvivedi A 2022 Decision-making on the selection of lean tools using fuzzy QFD and FMEA approach in the manufacturing industry *Expert Syst. Appl.* **192** 116416
[11] Sriramdas V, Chaturvedi S and Gargama H 2014 Fuzzy arithmetic based reliability allocation approach during early design and development *Expert Syst. Appl.* **41** 3444–9
[12] Feng Y, Hong Z, Cheng J, Tian G, Zhang H, Tan J 2017 Environmental-friendly reliability allocation for product platform based on expert measurement and ICN *Comput. Electr. Eng.* **64** 132–144
[13] Zhang G B and Liao X B 2009 Research on reliability allocation of direct-drive hobbing machine *Mater. Sci. Forum* **858** 245–50
[14] Bai B, Xie C, Liu X, Li W and Zhong W 2022 Application of integrated factor evaluation–analytic hierarchy process–T-S fuzzy fault tree analysis in reliability allocation of industrial robot systems *Appl. Soft Comput. J.* **115** 108248
[15] Du Y, Wu G, Tang Y, Cao H and Liu S 2021 Reliability allocation method for remanufactured machine tools based on fuzzy evaluation importance and failure influence *Int. J. Precis. Eng. Manuf.—Green Technol.* **8** 1617–28
[16] Peng N *et al* 2025 Multi-objective optimal tolerance allocation design of machine tool based on NSGA-II algorithm and thermal characteristic analysis *Proc. Inst. Mech. Eng. B* (https://doi.org/10.1177/09544054241310330)
[17] Wei P, Kitahara M, Faes M G R and Beer M 2025 Probabilistic calibration of model parameters with approximate Bayesian quadrature and active machine learning *J. Reliab. Sci. Eng.* **1** 015003
[18] Xiaoqing H, Zhilong W and Shihao L 2022 Health status evaluation method of CNC machine tools based on grey clustering analysis and fuzzy comprehensive evaluation *Int. J. Fuzzy Syst.* **42** 4065–82
[19] Yang J Z *et al* 2011 Fuzzy-comprehensive evaluation of use reliability of CNC machine tools *Key Eng. Mater.* **1076** 374–8
[20] Cui P J and Li H 2011 The service quality evaluation of machine tools in manufacturing shop *Adv. Mater. Res.* **1261** 981–7
[21] Tian J P, Yang H L, Feng X Q, Deng M T and Tang P 2014 Study of machine tool selection based on fuzzy comprehensive evaluation *Appl. Mech. Mater.* **3558** 1421–5
[22] Kumar R *et al* 2021 Revealing the benefits of entropy weights method for multi-objective optimization in machining operations: a critical review *J. Mater. Res. Technol.* **10** 1471–92
[23] Li G *et al* 2016 A safety evaluation method for heavy-duty CNC machine tools for the total life cycle based on an entropy weight method*2016 11th Int. Conf. on Reliability, Maintainability and Safety (ICRMS)* (IEEE) pp 1–7
[24] Li Y 2019 Scheduling analysis of intelligent machining system based on combined weights *IOP Conf. Ser.: Mater. Sci. Eng.* **493** 012146
[25] Chen C *et al* 2024 A comprehensive evaluation method for generalized reliability of CNC machine tools based on improved entropy-weighted extensible matter-element method *Int. J. Precis. Eng. Manuf.* **26** 1–9
[26] Kumar R, Bilga S P and Singh S 2017 Multi objective optimization using different methods of assigning weights to energy consumption responses, surface roughness and material removal rate during rough turning operation *J. Clean. Prod.* **164** 45–57

Qiang Cheng, corresponding author of the manuscript, received the PhD degree from Huazhong University of Science and Technology in 2009. He currently serves as a professor and PhD supervisor at Beijing University of Technology. He has published more than 100 SCI papers and been authorized with over 80 invention patents in the UK, Japan, and China. His main research interests include high-end manufacturing equipment, intelligent manufacturing, and medical robots.

Guobin Yan The first author of the manuscript. PhD candidate at Beijing University of Technology.

Yuegang Hu The third author of the manuscript, is a recipient of the State Council Special Government Allowance. Bachelor degree. He has been honored with titles such as National Technical Expert, National Young Post Expert, Sichuan May Day Labor Medal Awardee, and Sichuan Outstanding Teacher. Currently, he serves as a Senior CNC Technician at Ya'an Polytechnic College and is an Exchange Scholar at Beijing University of Technology.

Guodong Sun Master's degree holder. Assistant Research Fellow at the Shandong Institute of Scientific and Technical Information. His main research direction is technology consulting and technology management services.