scientific reports

Check for updates

OPEN

# Evaluation framework for domain-specific digital twin platforms

Zhao Tang[1]✉, Dayuan Zhuang[1] & Jianjun Zhang[2]

**As digital twin (DT) applications continue to proliferate across diverse industries, a noticeable gap exists in the availability of evaluation methods or frameworks to aid in the selection and development of DT platforms, particularly for Domain-Specific applications. To address this gap, this paper proposes a comprehensive evaluation framework for DT platforms, with a focus on Domain-Specific applications. The framework uses the Best-Worst Method and Fuzzy Comprehensive Evaluation method (BWM-FCE) to assess the performance, user experience, and economic effects of DT platforms. The proposed framework is applied to a case study of a high-speed train DT platform and compared with other evaluation methods AHP(Analytic Hierarchy Process) and BWM-SPA (Best-Worst Method-Set Pair Analysis). The results demonstrate the feasibility and effectiveness of the proposed framework and highlight its potential for guiding the development and selection of DT platforms for Domain-Specific applications.**

**Keywords** Digital twin platforms, Evaluation framework, Fuzzy comprehensive evaluation, Domain-specific digital twin, BWM-FCE

A DT is a virtual model of physical object or system that can replicate its behaviour, anticipate its performance, and detect potential problems. It is a key tool in the digital transformation of industries, helping organizations meet new competitive demands[1]. Over the last few years, there has been a significant focus on developing DT solutions to improve performance, reliability, and safety in various sectors such as transportation[2,3], manufacturing[4,5], healthcare[6,7], and energy[8]. As the interest in DT-based applications grows, the significance of DT platforms as the fundamental basis for creating, implementing, and managing DT solutions becomes increasingly apparent. Therefore, with the continued adoption of DT technology in the engineering domain, the need for robust DT platforms is set to increase.

However, at present, the focus of research efforts in the field of DTs is primarily on the creation and refinement of the DT model[9,10], rather than the underlying DT platform, especially the evaluation methods for selecting components to construct them. The DT platforms, as an important tool used in Industry 4.0, stemming from the previous modelling and simulation platforms, enhanced with new sensors and control technologies, and DT-based applications typically require Domain-Specific software platforms to effectively build and deploy them[11,12].

The rest of this article is organized as follows. Section "Related works" reports on the literature review of existing DT platforms and now available evaluation methods of DT. Section "Concept of domain-specific DT platform" focuses on the concept of Domain-Specific DT platform and distinguishes them from other similar concepts. Section "Evaluation framework" describes the comprehensive evaluation model based on the Best-Worst Method and Fuzzy Comprehensive Evaluation (BWM-FCE), followed by the Sect. "Case study", which verifies the rationality and practicality of the BWM-FCE evaluation model and the established framework in the former section by evaluating real cases of high-speed train DT platforms. In Sect. "Results analysis and discussion", not only the weight distribution results of the indicator system for the real case are compared between the BWM and AHP methods, but also the comprehensive evaluation results of FCE and SPA are compared. At last, Sect. "Conclusions and future works" draws conclusions and provides direction for future research.

## Related works

In this section, we explore thoroughly on the existing DT platforms and evaluation methods of DT platform, which help to select relevant functional components and assess the functionality suitability of a DT platform for a Domain-Specific application.

[1]State Key Laboratory of Rail Transit Vehicle System, Southwest Jiaotong University, Chengdu 610031, China. [2]National Centre for Computer Animation, Bournemouth University, Poole BH125BB, UK. ✉email: tangzhao@swjtu.edu.cn

### DT platforms

A DT platform serves as the foundational infrastructure for creating and managing DTs and DT-based applications. We gather literature related to existing DT platforms[13–15], and examine their distinct features, capabilities, and specific offerings in detail.

Ansys Twin Builder is a dedicated product for DT creation and control. It supports the full cycle of DT work, including planning, creation, verification, validation, and implementation. Mathworks' Matlab (Simulink) provides a comprehensive environment for creating, validating, and optimizing DT models. Microsoft's Azure Digital Twins is an IoT (Internet of Things) platform that enables the creation of complex digital environment models. Oracle Corporation's Oracle Cloud provides a robust cloud computing service that offers scalable and secure infrastructure for hosting and processing DT data. Its global data center network delivers servers, storage, networks, applications, and services.

Bosch IoT Suite, offered by Bosch, provides a collection of cloud services and software packages tailored to IoT projects[16]. It offers the necessary infrastructure and tools for developing and managing DT applications. Bosch IoT Suite focuses on delivering end-to-end solutions for IoT deployments, including the creation and management of DTs. Siemens AG, through Siemens PLM Software, provides a robust platform for product lifecycle management and process control. Solutions like HEEDS AI Simulation Predictor provides organizations with the capability to optimize products through state-of-the-art AI with built-in accuracy awareness, maximizing the benefits of the DT[17].General Electric's DT product, Predix, offers a specialized solution that integrates DT technology with industrial assets, making it particularly suitable for sectors such as manufacturing[18], energy, and transportation[19].

In addition to the aforementioned commercial DT Platforms, there are currently three well-known open-source DT Platforms, each with its unique features and advantages. ThingJS offers a low-code IoT 3D visualization development platform, enabling users to build DT visualization applications. It provides development tools such as 3D scene editors, 2D logic diagram editors, and map model systems. Eclipse Ditto is designed to assist users in creating DTs of devices connected to the internet, supporting the management of device states and implementing fine-grained access control. OpenTwins is an open-source framework for the design, development, and integration of DTs. It supports enhanced data analysis and visualization through 3D technologies and streaming machine learning. Naturally, there is also the IoT platform - FIWARE, which has been tested through real-world cases and is transitioning from a research to a commercial level[20].

Overall, the DT platforms provided by different suppliers for constructing a DT, coupled with the usage needs of different users, present a variety of possibilities. Nonetheless, this wide selection presents a significant challenge for organizations when it comes to selecting the most suitable platform that aligns with their Domain-Specific requirements.

### Evaluation methods

Another related work is research on the evaluation methods for DT platform, especially for DT platform components picking, which is essential to determine DT platform's effectiveness and applicability to specific applications. However, despite the increasing interest in DT technology, research on the evaluation of DT platform is limited. Furthermore, most existing evaluations focus on the DT models or DT-based applications[21,22], whereas only a few studies have explored on the comprehensive evaluation method of components for DT Platforms construction.

Two types of evaluation methods for DT platforms are found in the literature, including the assessment of DT platforms against requirements and the development of an evaluation framework that considers multiple factors[23]. For example, through empirical analysis of workflow process step functions, Prasad Talasila et al conducted a comparative evaluation between two DT platforms, HUBCAP and DIGITBrain[24].

From the literature review, we can find that the previous studies have mainly focused on developing and implementing DT platforms for various industries and applications, using different modelling and simulation tools, IoT platforms, and cloud computing technologies. However, they have paid little attention to the evaluation methods or frameworks for DT platforms, especially for Domain-Specific applications. Another way, the few existing evaluation methods are mostly based on subjective opinions, simple decision-making methods, and general-purpose criteria and indicators, which may not be suitable or applicable for complex and uncertain evaluation problems for Domain-Specific DT applications.

Therefore, there exists a research gap in the evaluation framework for DT platforms that specifically considers the quality of service, data quality, and system architecture in the context of Domain-Specific applications. To address these gap, two key challenges must be tackled. Firstly, there is a need to develop a comprehensive and standardized evaluation framework that can encompass all aspects of a DT platform. Secondly, clear and well-defined evaluation criteria and indicator system that are specifically tailored to Domain-Specific applications must be established.

## Concept of domain-specific DT platform

The term "Domain-Specific" here refers to a focus within a specific field or industry. In the context of DT applications, Domain-Specific DT is proposed in contrast to General-purpose DT and Multi-domain ubiquitous DT. General-purpose DT is defined as a real-time decision support sandbox[25,26], which assembles components such as data acquisition, system modelling, and intelligence applications into a DT based on modular and distributed thinking in the software domain. Multi-domain ubiquitous DT (UDT)[27] is a complex infrastructure system information management model based on Domain-Driven Design (DDD), while Domain-Specific DT is a digital solution designed for a specific industry or domain requirements. A fundamental technology for DT is the Digital Twin Graph (DTG)[28], an origin concept, that represents a versatile data structure that is integral to a processing framework designed to automate the creation of DTs across various domains without bias toward any specific field. For example, The primary high-speed railway of DTG lies in its facilitation of the automated and streamlined integration of domain knowledge, enabling the construction of virtual DTs for physical entities. In contrast, the capability of Domain-Specific digital twins is particularly evident in their ability to address novel issues within professional industries that cannot be resolved using existing procedural or process-driven knowledge, such as detecting the location and intensity of impacts in structural safety analysis[29]. Each industry or domain has its specific challenges, processes, and variables that need to be accurately represented and simulated within a DT. Therefore, a Domain-Specific DT is designed to closely mimic and replicate the characteristics of a specific industry or domain application, allowing for more accurate and relevant analysis, decision-making, and optimization within that particular context.

Accordingly, platforms that facilitate the development of both Domain-Specific Digital DTs and Domain-General DTs can be classified into two categories: Domain-Specific DT platforms and Domain-General platforms. To gain a deeper understanding of the concept of a Domain-Specific DT platform, it is essential to distinguish between Domain-Specific DT platforms and Domain-General platforms. Domain-Specific DT platforms are designed to support the construction of DTs within specific fields, incorporating established theoretical methods as well as emerging theories from interdisciplinary domains. In contrast, Domain-General platforms primarily serve to simulate and analyze various engineering challenges, thereby assisting engineers and scientists in addressing particular engineering issues across areas such as structural analysis, fluid dynamics, heat conduction, among others. These types of platforms are typically offered by traditional software providers, notable examples include Ansys Twin Builder by Ansys and Azure Digital Twins by Microsoft Inc. Furthermore, unlike their Domain-General counterparts, Domain-Specific DT platforms demonstrate a greater degree of scalability and flexibility while being more easily integrated with enterprise systems such as Enterprise Resource Planning (ERP) and Manufacturing Execution Systems (MES). This degree of customization facilitates a smooth adaptation to the changing demands of the industry.

Researchers emphasize the significance of DT platforms tailored to specific domains, along with their supporting technologies, models, and functions, as this approach allows for the construction of DTs that meet the unique demands of particular industries. In contrast, Domain-General platforms only facilitate the simulation of laws related to real-world objects, highlighting the critical role of effective evaluation tools and methodologies in adapting DT platforms to industry needs, and leading us to propose a comprehensive assessment framework that serves as a reference guide for researchers, practitioners, and developers to create effective digital twin platforms for specific applications.

## Evaluation framework

The process of using the constructed framework involves three steps: 1) classifying the requirements and functional components, 2) determining the evaluation indicators, and 3) applying the BWM-FCE evaluation model to assess the platform.

Implementation of the evaluation framework for different applications of DT platforms have various specific requirements and fundmental component that must be considered. For instance, in the case of the high-speed railway, DT Platform is designed to perform a range of functions, including data integration, model synchronization, visualization, simulation, and performance evaluation. The platform offers numerous benefits, including increased efficiency, enhanced safety, reduced costs, improved decision-making, and optimized asset management. Subsequently, classify the requirements and components of the DT platform. The requirements analysis, evaluation procedure, and evaluation results analysis involved in the framework are shown in Fig. 2a.

### Requirements and components of a DT platform

Based on the methodologies outlined in relevant literature[30], this paper examines the requirements for constructing a DT platform. Table 1 presents an overview of the requirements for a DT platform. These requirements are categorized into four groups based on the Analytic Hierarchy Process (AHP)[31,32] and Best-Worst Method (BWM)[33,34] and other research in the intersection of Management Science and DT technology[35]: basic function requirements, functional performance requirements, user experience, and profitability and efficiency. To distinguish whether it pertains to Domain-General or Domain-Specific functionalities, we label it as either 'general' or 'specific', highlighting the diverse set of features necessary for an effective DT platform based on the literature review and experts' advice[36]. Domain-Specific requirements refer to the unique needs and characteristics of a particular industry or application domain that the DT platform should address. Domain-Specific functional components refer to the components of the DT platform that provide specialized capabilities and functionalities to meet the Domain-Specific requirements.

Existing research has analyzed factors affecting usage intention of DT Technology in the manufacturing industry, the study examines the relationship between the factors and elements required to build a DT platform. To meet the requirements mentioned in the Table 1, based on the core functional components described by Tao et al, in their DT software platform reference architecture called makeTwin[37], we summarize all the functional components of a DT platform, as shown in Fig. 1, each of them plays a crucial role in enabling the platform. These functional components are divided into five high-level categories: data collection, communication, processing, visualization and interaction management component.

The typical functional components of a DT platform support the bi-directional working cycle (Bi-directional Input-Process-Output Cycle) , generally including the input unit, the central processing unit, and the output unit.

The workflow and concept of the DT are depicted on the respective sides of Fig. 1, illustrating their interrelation and functionality. The DT initiates with input from the physical model, which involves data collection from various sources such as physical sensors ,IoT devices and virtual sensors. The subsequent step involves processing data through simulation models and algorithms to create or update the DT model. Additionally, the visualization

| Category | Requirements | Description | Domain |
|---|---|---|---|
| Basic function requirements | Data capture | The ability to collect multi-source heterogeneous data, historical data. | General |
| | Data storage | Robust data storage capabilities to support large amounts of data. | General |
| | Data retrieval | Efficient data retrieval mechanisms to allow to access and analyze data quickly. | General/specific |
| | Data analysis | Include: real-time analytics, advanced analytics and predictive analytics | General/specific |
| | Visualization | The platform provide tools to visualize DT models, data, and algorithms. | General |
| | Interaction | Not only digital models can represent physical models, but also in the fact that digital models can predict the performance of physical models in reverse. | General |
| | Integration | The ability to integrate with other systems and platforms, such as IoT devices, cloud systems, and enterprise applications. | General/specific |
| | Interoperability | Support for industry standards for data exchange and interoperability, to allow for seamless integration with other systems. | General |
| | Generalization | It can make platform components more usable and reusable in different scenarios to better meet diverse needs. | General/specific |
| Functional performance requirements | High fidelity | Provide an accurate and scalable representation of the physical object. | General/specific |
| | Real-time simulation | The platform should support real-time simulations to enable users to analyse and understand the behaviour of systems and processes in real-time. | Specific |
| | Real-time processing | The ability to process and analyse large amounts of real-time data. | Specific |
| | Scalability | Handle increasing amounts of data and DT models. | General/specific |
| | Security | Robust security features to protect sensitive data and ensure that the DT models are protected from unauthorized access. | Specific |
| | Synchronization | The ability to ensure that the DT accurately represents the real-world system in real-time by continuously updating its data and state. | Specific |
| | Specialization | The ability to offer professional models, data and algorithms in specific domain. | Specific |
| User experience | User interface | A user-friendly interface that makes it easy for users to access, analyse, and interact with DT models. | Specific |
| | Interaction | Designed to be intuitive to minimize the learning curve for users. | General |
| | Ease of use | Designed to be user-friendly, to minimize the learning curve for users. | General/specific |
| | Customization | Ability to be customized and configured to meet the specific needs. | General |
| Profitability and Efficiency | Cost-effectiveness | The platform provides a positive return on investment for the business. | General |
| | Market demand fit | Understand of the market demand and the specific requirements. | Specific |
| | Domain flexibility | Flexibility to accommodate different types of models and simulations, as well as support for multiple domains and use cases. | General |
| | Market share | The percentage of total sales in a particular market that a DT platform has | General/specific |

**Table 1**. The requirements of a DT platform.

component enhances the accessibility and user-friendliness of the DT platform. The interaction management component is crucial for a DT platform, as it facilitates user engagement and ensures bidirectional influence between the digital and physical models. Given the numerous components and diverse functionalities of the DT platform, it is essential to establish an indicator system that considers these aforementioned sub-factors.

## Evaluation indicator system

A suitable and comprehensive indicators system is vital to ensure a thorough assessment of DT platforms' build tools. The evaluation indicator system is derived from the requirements analysis of a DT platform in Sect. "Requirements and Components of a DT Platform", which identifies the basic function requirements, functional performance requirements, user experience, and profitability and efficiency as the main categories of requirements. These categories are further divided into sub-categories based on the literature[38] and expert opinions. The draft provided by the expert opinion employs the K-means method and the Random Forest method to categorize the evaluation indicators under each first-level indicator. The selection of experts is based on three criteria: years of professional experience, academic qualifications, field of expertise and job title. Take case studies as an example, the information of the participating experts was organized along three dimensions as presented in Table 2.

The evaluation indicators are also classified into Domain-General and Domain-Specific indicators, depending on whether they are applicable to all domains or only to a specific domain. Four main categories of evaluation indicators have been extracted from these requirements and functional components, including Domain-Specific functional component performance, Domain-General functional component performance, economic effects, and user experience as the first-level indicators. And incorporate a user-centric evaluation that involves feedback from potential users of DT platforms to gauge user-friendliness, practicality, and real-world applicability. The first two categories consist of four second-level indicators each, while the latter two comprise three second-level indicators each, as shown in Fig. 2b.

The performance of Domain-General functional components is assessed in four areas: scalability, security, synchronisation, and visualisation, while Domain-Specific functional component performance is assessed in four metrics: modelling, simulation, algorithm and interaction. Beneath these metrics, third-level indicators are further delineated to maintainability, generality, stability, immersion, accuracy, templating, simulation consistency, real-time response time, algorithmic configurability and interaction robustness. In the assessment

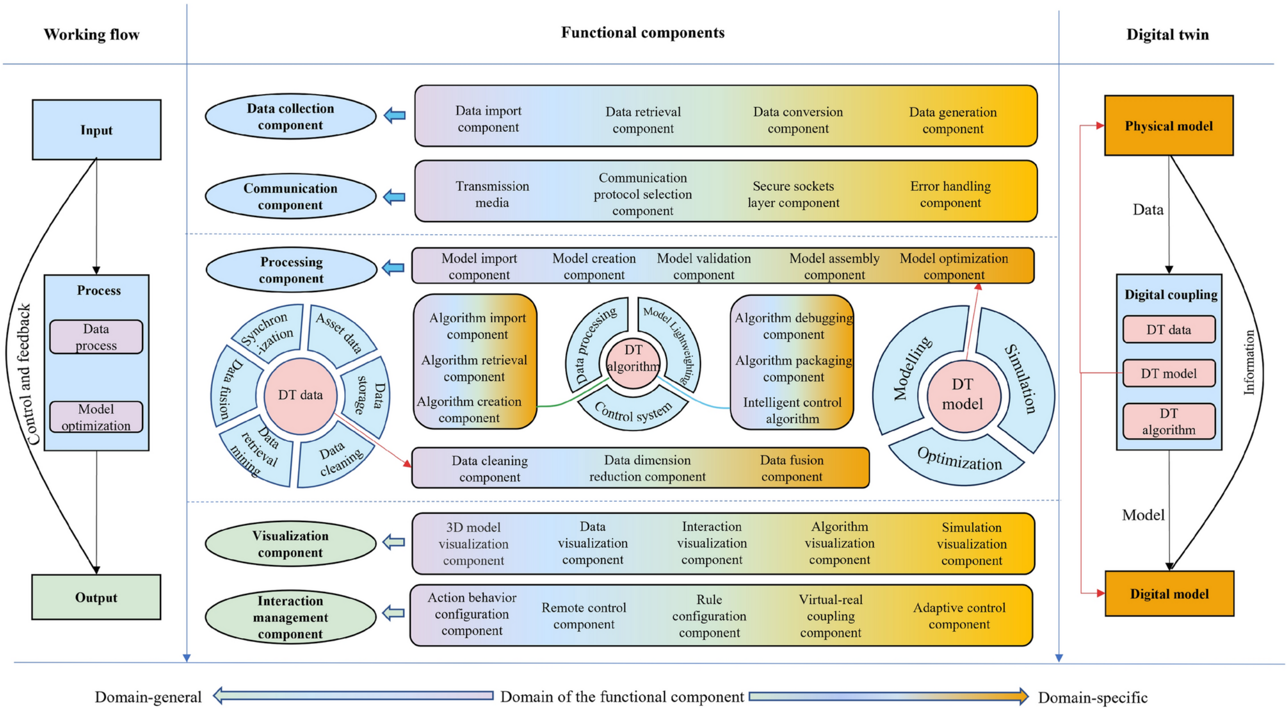


**Figure 1**. High-level functional components of a DT platform illustrating the Bi-directional Input-Process-Output Cycle, distinguishing between Domain-General and Domain-Specific functional components.

| No. | Years of service | Academic qualifications | Field of expertise |
|---|---|---|---|
| 1 | 8 | Bachelor's Degree | Computer Science |
| 2 | 3 | Master's Degree | Mechanical Engineering |
| 3 | 5 | Doctorate | Physics |
| 4 | 4 | Bachelor's Degree | Management Science |
| 5 | 2 | Master's Degree | Statics |
| 6 | 6 | Bachelor's Degree | Economics |
| 7 | 8 | Master's Degree | Mechanical Engineering |

**Table 2**. Experts' information.

process utilizing these metrics, integrating the third-level metrics into the second-level indicators, thereby enhancing the clarity and comprehensibility of the evaluation results while minimizing duplication and redundancy among the metrics. It is worth emphasizing that the core of the framework is the following BWM-FCE Evaluation Model.

## BWM-FCE evaluation model

Utilizing the BWM-FCE evaluation model to assess a DT platform based on the evaluation indicator system outlined in the preceding section. BWM, is employed to determine the weights of a set of criteria or factors, providing clarity on their impact in decision-making problems. This method offers the advantage of yielding consistent results with minimal comparative information and has found widespread application in project selection[39], supplier evaluation[40,41], project subcontracting[42], etc. By integrating the BWM with FCE, the novel evaluation model effectively incorporates quantitative and qualitative factors, thereby rendering it suitable for the intricate and multidisciplinary nature of DT platforms. This algorithm adeptly considers quantitative elements in its assessment, further affirming its appropriateness for the complex landscape of DT platforms. A significant advantage of employing the BWM-FCE evaluation model in assessing DT platforms is its capacity to integrate subjective insights from experts and stakeholders throughout the evaluation process. Methods for gathering both subjective opinions and objective variables include expert surveys, performance metrics, and economic statistics. This approach facilitates a more holistic evaluation that acknowledges the diverse perspectives and requirements of all involved stakeholders. To implement the BWM-FCE model for evaluating DT platforms, several systematic steps must be adhered to.

First, the evaluation indicator and sub-indicator are defined based on the characteristics of the platform and the requirements of the DT platforms' construction. And the best/worst criteria should be determined, the

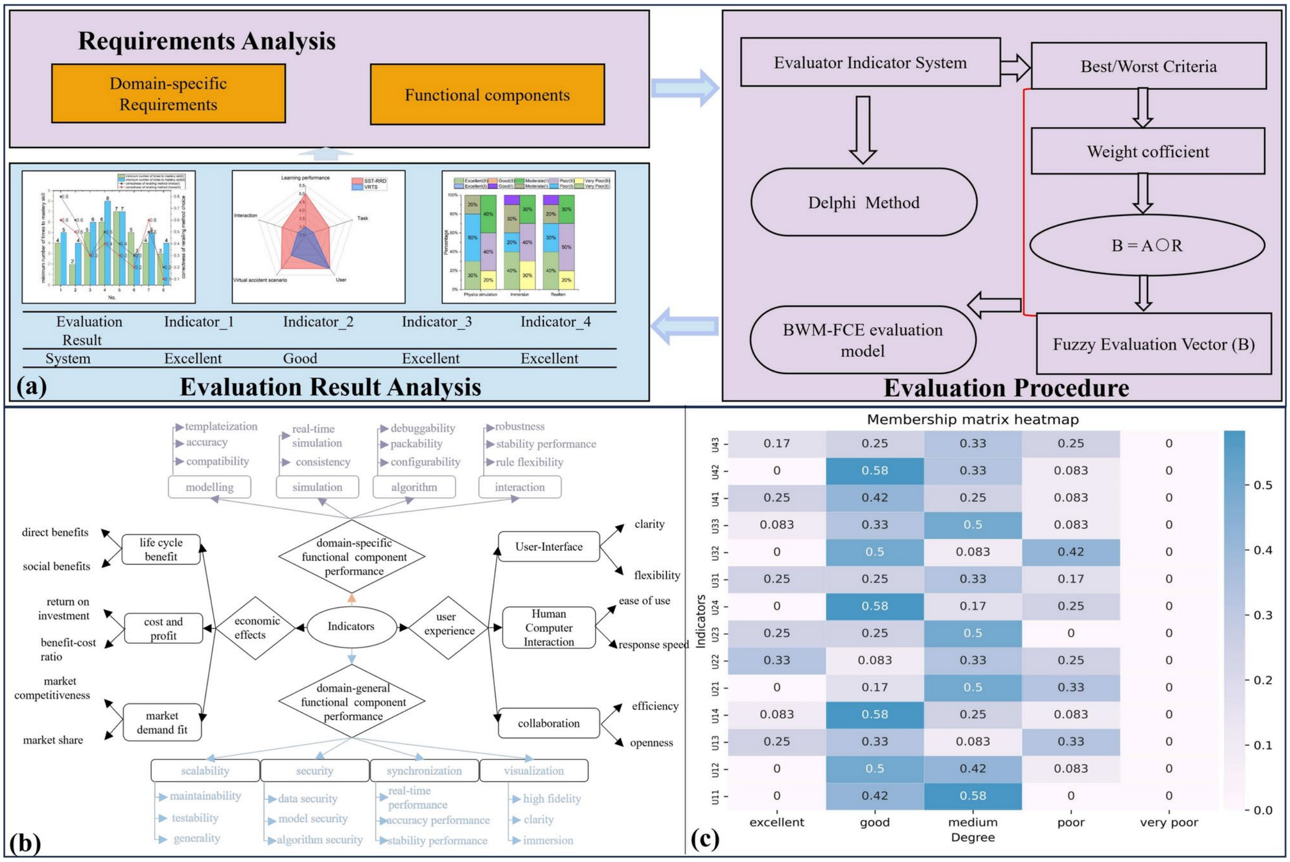


**Figure 2**. (**a**) A comprehensive evaluation framework for evaluating Domain-Specific DT platforms, (**b**) Established indicator system of framework, (**c**) Framework's membership matrix.

best criterion serves as the determinant that exerts the most postive effect on decision-making , while the worst criterion demonstrates the opposite effect. The decision makers (DMs) then give preferences of the best criterion over all the other criteria and also preferences of all the criteria over the worst criterion using a number from a predefined scale (e.g. 1 to 9), the indicator with a score of 1 in the optimal scoring table is identified as the best (most important, most ideal) indicator, and the indicator with a score of 1 in the worst scoring table is identified as the worst (least important, least ideal) indicator.

Second, the BWM is applied to formulate the weights of each criterion and sub-criterion, these weights reflect the priority of each criterion and sub-criterion in the evaluation process is as follow:

$$\begin{cases} \min\limits_{\epsilon} \quad \epsilon \\ \text{s.t.} \quad \left| \dfrac{V_{best}}{w_j} - C_{bj} \right| \leq \epsilon \\ \qquad \left| \dfrac{w_j}{V_{worst}} - C_{wj} \right| \leq \epsilon \quad \forall j \in \{1, 2, \dots, n\} \\ \qquad \sum\limits_{j=1}^{n} w_j = 1 \\ \qquad w_j \geq 0 \quad \forall j \in \{1, 2, \dots, n\} \end{cases} \tag{1}$$

In the equation, $\epsilon$ is the target value of planning problems, $V_{best}$ / $V_{worst}$ is the best/worst criteria, $w_j$, $C_{bj}$ and $C_{wj}$ are the weights of the remaining indicators, the weight of the best indicator, and the weight of the worst indicator. **Third**, a membership matrix must be constructed for each indicator and sub-indicator by identifying the concepts and causal relationships that affect them. The membership matrix represents the causal relationships among the concepts in a qualitative way. Finally, the BWM weights and membership matrix are combined to obtain a comprehensive evaluation score for the DT platform using fuzzy arithmetic operations (such as addition, multiplication, inverse operation) or other aggregation methods. The score indicates how well the platform meets the evaluation criteria and sub-criteria. The procedure of BWM-FCE algorithm used in the BWM-FCE evaluation model proposed framework is represented bellow.

**Input:** DT platform requirements list ($R_l$), components list ($C_l$), best/worst criteria vector ($V_{best}/V_{worst}$), experts' information ($E_i$), and questionnaire results ($Q_r$);
**Output:** DT platform's evaluation result ($R_{ce}$) and weight (A);
**for** *each episode* **do**
    Initialize $R_l$, $C_l$, and $V_{best}/V_{worst}$, obtain experts' information ($E_i$), and questionnaire results ($Q_r$).
    **repeat**
        **if** *!exists weight A or* $CR > 1$ **then**
            ($R_l$, $C_l$) reasonable;
            construct the Indicator System;
            do the BWM with Eq. (1);
        **else**
            do the Consistence Check with Eq. (2);
    **until** $CR < 1$;
    Store the intermediate result ( $V_{best}/V_{worst}$, A);
    **repeat**
        do the FCE;
        Processing $Q_r$ to obtain a membership matrix (R);
        **if** $\sum_{i=1}^{n} R_{ij} \neq 1$ *or* $R_{ij} \notin [0, 1]$ **then**
            reconstruct R;
        **else**
            do the Fuzzy Operation with Eq. (3);
    **until** *Conclusion aligns with expert validation*;
    Store the evaluation result ($R_{ce}$);
**return** *DT platform's evaluation result ($R_{ce}$) and weight (A).*

**Algorithm 1**. BWM-FCE Evaluation Algorithm

In the BWM-FCE evaluation algorithm, $V_{best}/V_{worst}$ is the best/worst criteria, $A$ introduces the weight of each indicator being formulated, *CI* and *CR* in equation are the parameters of consistency check about weight; The preliminary $R$ represents the membership matrix of specific level indicators, the element r of it is very important, so it needs expert validation.

$$\text{consitency ratio (CR)} = \frac{\epsilon}{CI} \tag{2}$$

CI is the consistency index. The calculation of the Consistency Index (CI) can be derived from Table 3 as outlined by Rezaei[33].

In addition, the *CI* in AHP is as follows:

$$CI = \left(\frac{\lambda_{\max} - n}{n - 1}\right) \tag{3}$$

In the equation, n is the order of the matrix and $\lambda_{\max}$ is the maximum eigenvalue of the matrix.

## Case study

Assessing DT platforms can be prohibitively expensive and time-consuming, often exceeding the budgetary and temporal constraints of research projects. Additionally, many platforms in other industries are not fully accessible for comprehensive evaluation due to confidentiality concerns, which further complicates the assessment process. In this section, we apply the proposed framework to assess a DT platform designed for high-speed railway, called DT Train. The platform is required to meet the operation and maintenance practice standards and provide fundamental components for simulating the train operations.

### Background

Creating a DT platform for high-speed railways is vital for the industry's digital transformation. The platform acts as a comprehensive tool to boost operational efficiency through virtual modeling, real-time monitoring, optimized scheduling, and predictive maintenance. It enhances safety by preemptively identifying and mitigating hazards, thus optimizing the rail network. DT Train Platform developed by SWJTU represent a revolutionary step in train design, optimization and operation management[43]. By developing virtual replicas of physical trains, this platform facilitates comprehensive simulation and analysis of train performance. The process commences with a physics-based generic model, which is subsequently customized using specific train data. Continuous updates

| $a_{BW}$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| *CI* | 0 | 0.44 | 1 | 1.63 | 2.3 | 3 | 3.73 | 4.47 | 5.23 |

**Table 3.** Consistency index.

with real-time data ensure the DT's accuracy. The platform's vehicle system dynamics, which are integrated with machine learning and sensor data, facilitate precise replication and predictive capabilities. This advanced technology paves the way for applications such as predictive maintenance and performance optimization, thereby promising enhanced efficiency and safety in train operations. The platform serves as a foundational infrastructure that facilitates comprehensive testing in intricate scenarios and fosters industry collaboration, thereby ensuring scalability and practical application within the rolling stock sector.

The DT Train Platform is used as a case study for a Domain-Specific DT platform, as it reflects the core ability and construction process of DT technology. Compared with DT platform of cities, aviation, and factories, although there are differences in application areas and specific technological implementations, they have commonalities in terms of requirements and construction processes:

1. Faced with the challenge of modeling and simulating complex systems, high-speed railways, like cities, aviation, and factories, are composed of multiple complex systems that involve equipment, facilities, and personnel from different fields.
2. DT constructed by DT platform require real-time collection and monitoring of system operation data, including sensor data, real-time monitoring data, etc. For example, high-speed railway DT platforms need to integrate data collected by various sensors and monitoring equipment to achieve real-time monitoring. Similarly, DT cities, aviation, and factories also need this.
3. DT platforms can utilize real-time data and simulation results to achieve predictive maintenance and intelligent decision support for the system. Through data analysis and algorithmic models, potential faults are predicted and corresponding maintenance recommendations are provided, thereby reducing downtime and maintenance costs. The high-speed railway DT platform provides decision support, helping decision-makers understand complex situations and make more scientific and reasonable decisions. DT cities, aviation, and factories also require.

Although the general requirements of the DT platform are universal to all applications, the DT Train Platform also have specific requirements and components that meet these requirements, which are different from those of Domain-General platforms. The selection of the DT Train Platform for evaluating the relevant characteristics of Domain-Specific DT using the proposed BWM-FCE framework has universal significance.

## Specific requirements and components of the DT Train Platform

The objective of the DT Train Platform is to supply a set of cmmponents to construct a digital representation of the high-speed railway system that accurately reflects its physical counterpart. And the DT Train Platform is designed explicitly to address the high-speed railway domain. The specific requirements for DT Train Platform's model processing component are more specialized. For example, it must be able to simulate the dynamics of high-speed trains and provide accurate simulations of the vehicle's behaviour, therefore, professional simulation software packages such as Simpack, Ansys and other self-developed software packages need to be integrated into this platform. Besides, the simulation capabilities, the DT TrainPlatform's data collection component must also be able to handle real-time data from onboard physical sensors, as well as provide real-time control and feedback to the train.

Another distinction between the general requirements and the specific requirements for the DT Train Platform lies in the necessity for advanced algorithms and models tailored for data analysis. The algorithms and models implemented within this platform must possess the capability to analyze substantial volumes of data and generate predictions regarding the behavior of high-speed trains in real-world scenarios.

To meet the other special requirements of DT Train, a component library offered by the DT Train Platform is constructed based on the component categories and requirements in Sect. "Requirements and components of a DT platform", the detailed components are listed in the following Fig. 3.

The primary components of the DT Train Platform are categorized into five distinct categories: data collection component, data processing component, model processing component, visualization component, and interaction management component. The data collection component includes track-beside equipment, smart devices, on-board sensors, bench devices, and virtual sensors. It is capable of not only acquiring data from laboratory benches and comprehensive testing vehicles but also utilizing intelligent equipment and virtual sensors to gather information on operating trains and simulation models. The data processing component involves Big Data technology and dynamic simulation in the field of rail transit, with TPLDna (a train dynamics simulation system) being more representative. The model processing component ranges from general CAE software to commonly used Simpack and UM software in the rail industry. The visualization component not only uses the Unity3D engine but also utilizes special cloud rendering technology. Finally, the interaction management component extends from the interaction management component PyQt/Flutter on the two-dimensional plane to the Extended Reality (XR) interaction in the three-dimensional space, and the dedicated component is based on VRCTS (a railway virtual rescue training system).

The platform is designed to perform a variety of functions, including data integration, model synchronization, visualization, simulation, and performance evaluation. It offers numerous advantages such as increased efficiency, enhanced safety, reduced costs, improved decision-making capabilities, and optimized asset management. Furthermore, it distinguishes itself from Domain-General platforms through its ability to address the complexities associated with high-speed railways. This capability is achieved via the integration of specialized components, a hybrid modeling approach, and self-developed enabling tools that effectively capture and process large volumes of data while simulating complex operational behaviors in real-time. Due to the complex operating environment of high-speed railways and the varying requirements for using DT Train Platforms, as well as the numerous

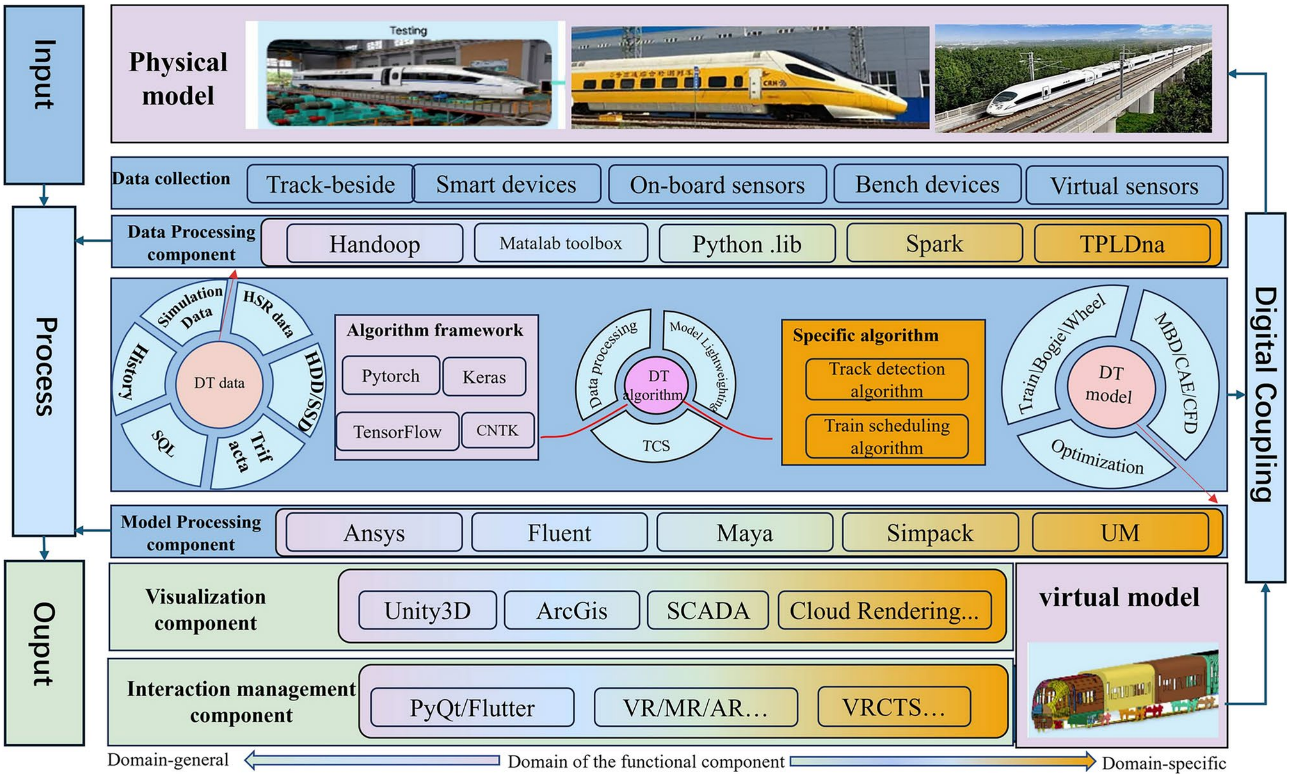


**Figure 3**. Main components of the DT Train Platform.

| Decision makers | Best criterion | Worst criterion | U1 | U2 | U3 | U4 |
|---|---|---|---|---|---|---|
| $D_1$ | U2 | – | 3 | 1 | 7 | 5 |
| | – | U3 | 5 | 9 | 1 | 3 |
| $D_2$ | U2 | – | 3 | 1 | 5 | 7 |
| | – | U3 | 5 | 7 | 1 | 3 |
| $D_3$ | U2 | – | 3 | 1 | 9 | 7 |
| | – | U3 | 5 | 9 | 1 | 3 |

**Table 4**. Pairwise comparison vectors of first indicators.

platform components, it is necessary to use the aforementioned framework to evaluate existing DT Trains and select conventional methods to validate the effectiveness of the framework.

## Evaluation process

Following the reference framework and steps proposed in the previous sections, which primarily involve identifying evaluation indicators, constructing the best/worst criteria vector, and utilizing the BWM-FCE evaluation model. This section implements the comprehensive evaluation of the DT Train Platform. And publicly available information about the experts as shown in Table 4.

*Implementing BWM*

The weights of the 14 second-level indicators and 4 first-level indicators are determined using the Delphi method and BWM. The best/worst criteria of first indicators are given by three experts are shown in Table 4.

Overall, consensus among experts is that the importance of factors decreases in the following order: Domain-Specific functional component performance, Domain-General functional component performance, economic effects, and user experience, Table 5 provides weights for all indicators in the first and second layers.

*Implementing FCE*

Following the application of the BWM, the weights of the indices are determined. Subsequently, the FCE procedure is employed to obtain the membership matrix and the resultant comprehensive evaluation results.

To objectively assess user experience, we gathered quantitative feedback regarding user satisfaction and usability through the utilization of standardized questionnaires, encompassing the System Usability Scale (SUS) and Net Promoter Score (NPS). Collecting and analyzing objective indicators of developer interaction with the

| First level indexes | 1st level weight | 2nd level indexes | 2nd level weight |
|---|---|---|---|
| U1 Domain-General functional component performance | 0.3178 | U11 scalability | 0.1208 |
| | | U12 security | 0.1208 |
| | | U13 synchronization | 0.0502 |
| | | U14 visualization | 0.0261 |
| U2 Domain-Specific functional component performance | 0.4686 | U21 modelling | 0.0239 |
| | | U22 simulation | 0.0731 |
| | | U23 algorithm | 0.2160 |
| | | U24 interaction | 0.1556 |
| U3 user-experience | 0.0978 | U31 user-interface | 0.0335 |
| | | U32 HCI | 0.0309 |
| | | U33 collaboration | 0.0333 |
| U4 economic effects | 0.114 | U41 life cycle benefit | 0.0131 |
| | | U42 cost and profit | 0.0228 |
| | | U43 market demand | 0.0799 |

**Table 5**. Final weight distribution of evaluation model. [1] Ui corresponds to the four first level indicators, while Uij represents the j-th secondary indicator under the indicator Ui

system, such as the completion time of individual development tasks, bug-related downtime, and the frequency of utilization of individual components or modules. And evaluating subsystem performance metrics, including response time, fault tolerance, and reliability, which directly influence the user experience. Additionally, we employ qualitative metrics such as user satisfaction, engagement, and retention to appraise the overall performance of the platform. Collectively, these metrics offer insights into user experience and platform usage, enabling us to enhance system design and augment user satisfaction. After analyzing the data collected from the questionnaire, The platform's performance on 14 second-level indicators and their degree of membership to 5 levels (excellent, good, medium, poor, very poor) are displayed in the Fig. 2c. The membership matrix has been expertly validated, using the weights of each indicator (see Table 5) and the membership matrix, the evaluation reissults vector are formulated. The first level membership matrix R is as follow:

$$R = A_i \circ R_i = \begin{bmatrix} 0.0349 & 0.1303 & 0.1061 & 0.0466 & 0.0000 \\ 0.0572 & 0.1951 & 0.1402 & 0.0761 & 0.0000 \\ 0.0139 & 0.0274 & 0.0349 & 0.0215 & 0.0000 \\ 0.0033 & 0.0491 & 0.0572 & 0.0063 & 0.0000 \end{bmatrix} (i = 1, 2, 3, 4) \tag{4}$$

Based on the maximum membership principle, a radar chart is drawn as shown in Fig. 4, the final fuzzy comprehensive evaluation result is formulated as follow:

$$B = \begin{bmatrix} 0.0396 & 0.1412 & 0.1095 & 0.0533 & 0.0000 \end{bmatrix} \tag{5}$$

The comprehensive evaluation of the high-speed railway DT platform is moderate. As shown in Fig. 4b, the platform's 14 second level indicators are mainly distributed between good and moderate levels. Among these indicators, $U_{22}$ (simulation) demonstrates exceptional performance, indicating that the platform possesses robust simulation capabilities. This feature is particularly critical for this Domain-Specific DT platform, which has been meticulously designed to address the requirements associated with the development of high-speed trains.

## Results analysis and discussion

The effectiveness and advantages of the proposed framework is evaluated and discussed in this section. Main components of the proposed framework, including the weight determination method and the result analysis, are determined based on detailed comparison.

### Results analysis

*Results of BWM process*

Apart from BWM used in the proposed framework, AHP is another widely-used multi-criteria decision-making method[44,45]. Therefore, we use the AHP method to serve as the benchmark. Figure 4a shows the weight distribution of the first level and the second level indicators formulated using the two methods (AHP and BWM). Difference between the four indicators of the first level in the two methods is extremely small, with a range of 0.0553, although the range of the second level weight distribution is slightly larger (0.1015), after the joint calculation of the first level weight, the range of the weight distribution is 0.0323, the disparity is marginal. The Spidelman coefficients, correlation coefficients and covariance of AHP and BWM weight distribution results are generally formulated. Spidelman coefficients for the two-layer weight distribution of the two methods are 1 and 0.5297, respectively, which means the positive correlation between the two is very strong. But the covariance between the first level, second level weight distributions of the AHP and BWM methods are 0.021 and 0.003,

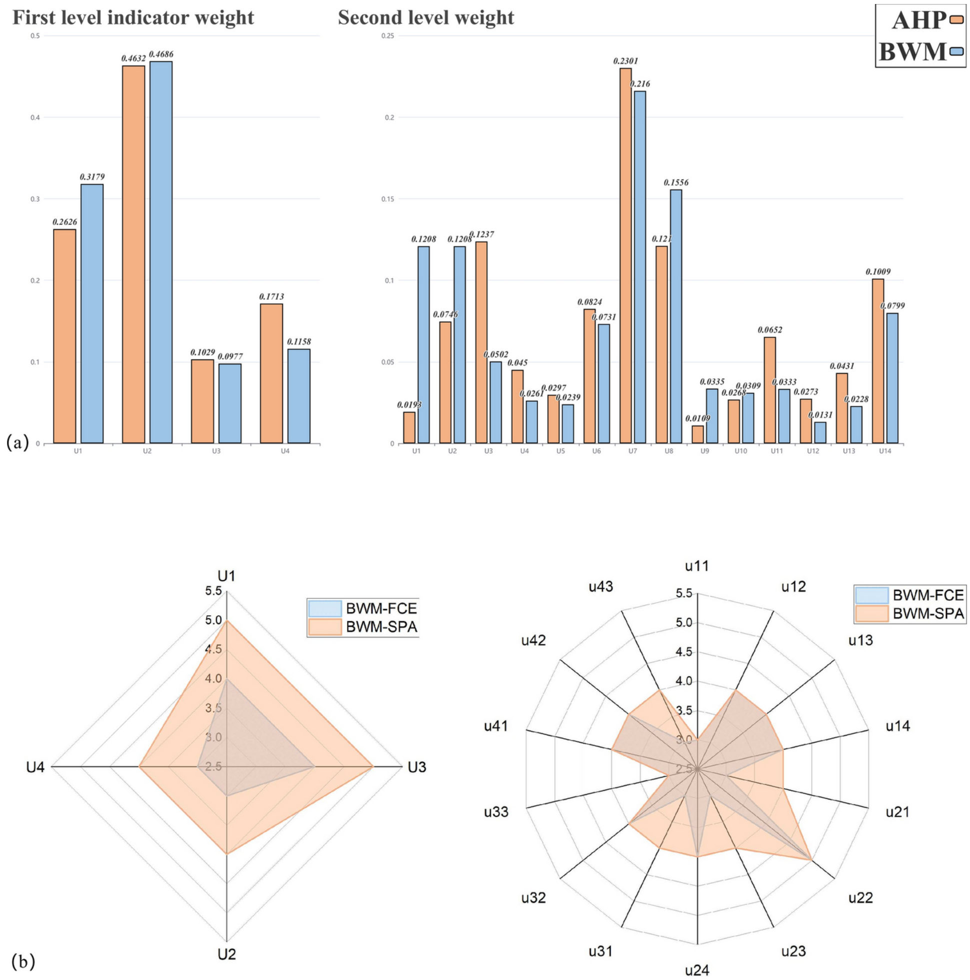


**Figure 4**. Comparison of the weights and evaluation results, (**a**) the first and second level weights of indicators (**b**) the evaluation results of DT Train.

indicating a certain linear relationship between the two sets of weights, however, this linear relationship is very weak. And the correlation coefficients for each level weight distribution calculated by the two methods are 0.97 and 0.76, respectively, this further illustrates that the data obtained by the two methods are positively correlated.

Furthermore, to analyze the trends and magnitudes of changes in the indicator layer in response to variations in specific criterion layer indicators, and to facilitate higher-level decision-making, 1,000 sampling points were selected for each criterion layer indicator within the aforementioned evaluation framework to assess its sensitivity. A critical ratio (CR) closer to zero signifies superior performance[34]. Perform weight sensitivity analysis on the indicators, and the results are shown in Fig. 5. As shown in Fig. 5a, as the weights assigned to Domain-General and Domain-Specific increase, CR gradually decreases, and the rate of decrease in Domain-Specific is significantly higher than the former, indicating that changes in Domain-Specific weight have a greater impact on the final result.

*Results of FCE process*

Fuzzy Comprehensive Evaluation (FCE) and Set Pair Analysis (SPA) are both methodologies employed for assessment and decision-making within complex systems characterized by uncertainty. To determine the more suitable method for evaluating a Domain-Specific DT platform, FCE and SPA are both applied to assess the DT train Platform. The weight distribution of the FCE is same as SPA, calculated by BWM, as shown in Table 5. The variables involved in the SPA are: the scores of 14 second level indicators, which are the weighted average of user evaluation scores, and the indicator level segmentation points selected based on expert advice. And the values of $i_1$, $i_2$, $i_3$, and j parameters in the five element connection number involved in SPA are 0.5, 0, -0.5, -1, respectively.

## Discussion

From Fig. 5b, it can be seen that when the weight of "Domain-General functional component performance" increases, only the weights of scalability, security, synchronization, and visualization will increase, while the weights of other indicator layers show a downward trend to varying degrees. Moreover, the slope of the straight

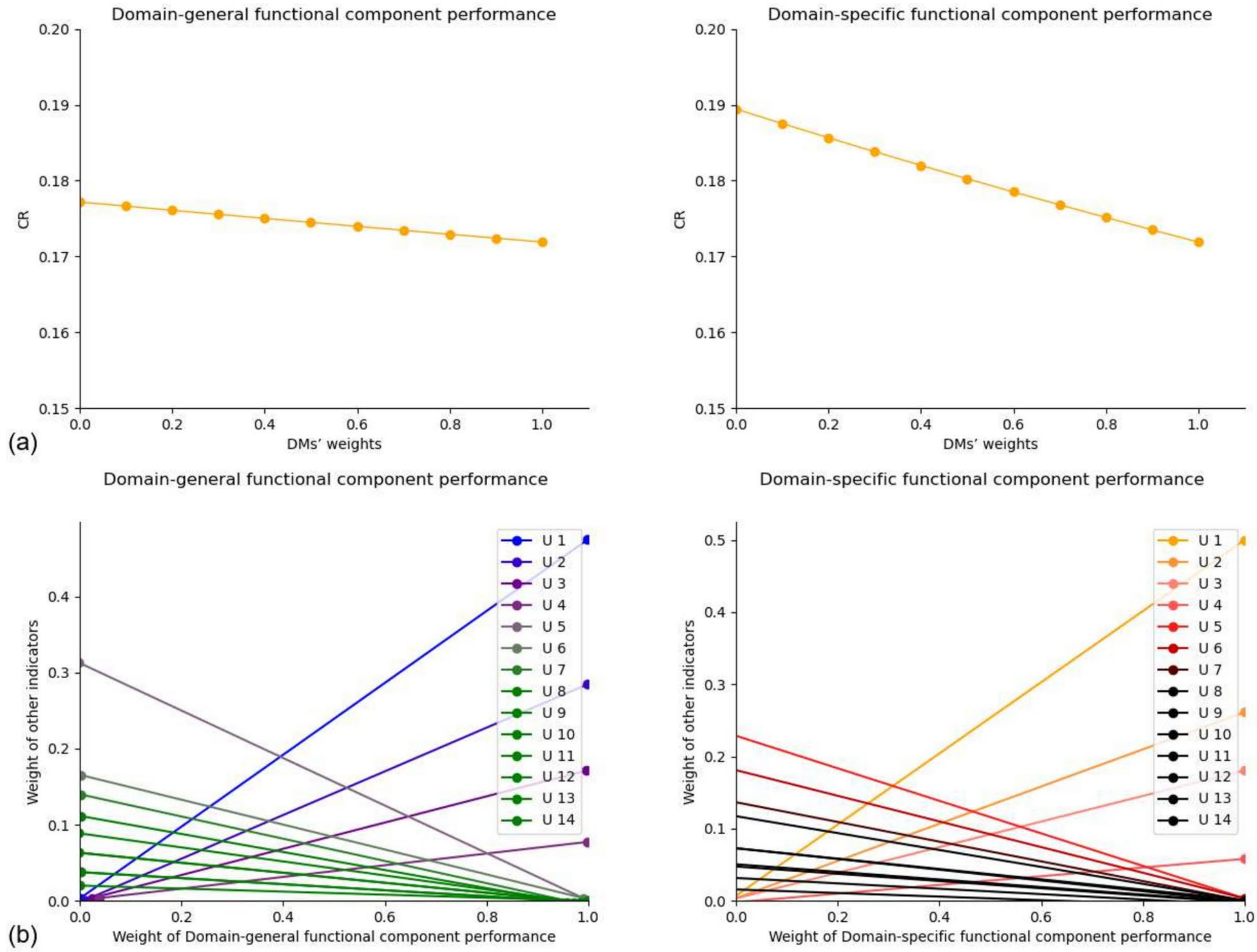


**Figure 5**. Sensitivity analysis of the Domain-Specific and Domain-General components' weights (**a**) CR accompanies changes in DMs' weights (**b**) Changes in weight of other indicators with the current indicator weight.

line in represents the speed at which the weights of each indicator change with the weight of "Domain-General functional component performance" or "Domain-Specific functional component performance".

The intersection of any two straight lines in Fig. 5 represents the inflection point at which the relative weight of the indicator changes. For example, the intersection of synchronization and algorithm (0.400, 0.1873) indicates that before this point, the importance of algorithm is greater than that of synchronization, and after this point, the importance of synchronization surpasses that of algorithm, highlighting the significance of indicator selection and weighting.

Although the final results using BWM-FCE and BWM-SPA differ, the overall trend is similar (Fig. 4b). Specifically, $U_1$ and $U_2$ score one level higher than $U_3$ and $U_4$, indicating that the high-speed train DT platform places more emphasis on the performance of Domain-General and Domain-Specific functional components performance compared to the user experience and economic benefits, which is consistent with the expected results. The variation in the assessment outcomes of the second level indicators is relatively minor, the four second-level indicators $U_{21}$, $U_{23}$, $U_{31}$, and $U_{43}$, is a one-degree difference, while the evaluation results of the remaining ten second-level indicators are consistent. It is worth mentioning that the assessment result of comprehensive evaluation framework on the simulation capabilities of DT Train Platform is consistent with the benchmark method, both are excellent.

As shown in Table 6, the comprehensive evaluation results of DT Trains obtained using the BWM-SPA method are: the comprehensive evaluation results of U1 and U2 are excellent, and the results of U3 and U4 are good. In summary, the evaluation results of BWM-FCE for the high-speed railway DT platform are more in line with the actual situation than those of BWM-SPA. Compared to the AHP method, BWM has a time complexity of $2 * n$, while the former has a time complexity of $n^2$, making it easier to implement in practical decision-making. Besides, SPA has a single data processing method, unlike FCE which can use fuzzy operators for calculations. Additionally, FCE has more ways to obtain fuzzy matrices and is more suitable for evaluating objects with multiple factors.

| 2nd indicators | Score value | Segmentation point | Five element contact number | Connectivity | CN |
|---|---|---|---|---|---|
| U11 | 73.334 | [90 80 70 60 50] | [0.33 0.67 0.00 0.00 0.00] | [0.4822 0.5178 0 0 0] | 0.741085 |
| U12 | 74.993 | | [0.50 0.50 0.00 0.00 0.00] | | |
| U13 | 78.334 | | [0.83 0.17 0.00 0.00 0.00] | | |
| U14 | 74.159 | | [0.42 0.58 0.00 0.00 0.00] | | |
| U21 | 70.833 | [90 80 70 60 50] | [0.08 0.92 0.00 0.00 0.00] | [0.4979 0.5021 0 0 0] | 0.748957 |
| U22 | 75.833 | | [0.58 0.42 0.00 0.00 0.00] | | |
| U23 | 73.333 | | [0.33 0.67 0.00 0.00 0.00] | | |
| U24 | 77.5 | | [0.75 0.25 0.00 0.00 0.00] | | |
| U31 | 7.4992 | [10 8 6 4 2] | [0.00 0.75 0.25 0.00 0.00] | [0 0.6727 0.3273 0 0] | 0.336361 |
| U32 | 6.8334 | | [0.00 0.42 0.58 0.00 0.00] | | |
| U33 | 7.6659 | | [0.00 0.83 0.17 0.00 0.00] | | |
| U41 | 3.74965 | [5 4 3 2 1] | [0.00 0.75 0.25 0.00 0.00] | [0 0.7130 0.2870 0 0] | 0.356511 |
| U42 | 3.70835 | | [0.00 0.71 0.29 0.00 0.00] | | |
| U43 | 3.70835 | | [0.00 0.71 0.29 0.00 0.00] | | |

**Table 6**. BWM-SPA evaluation results. [1] CN: Connection number

## Conclusions and future works

### Conclusions

In order to improve the efficiency of DT model development and optimize the process of Domain-driven DT design, this article proposes the term Domain-Specific DT platform. Researchers can improve the DT-driven research efficiency and process of DT development by utilizing the evaluation framework proposed in this article to identify components that meet specific requirements. They can achieve this by establishing an evaluation indicator system and a BWM-FCE evaluation model to evaluate and further develop the Domain-Specific DT platform. The construction of BWM-FCE evaluation model mainly includes the determination and verification of indicator weights, the determination of the membership matrix, and the acquisition of comprehensive evaluation results. In order to provide a detailed description of the research process of Domain-Specific DT platforms, taking the construction and evaluation of DT Trains platform as an example, the specific processes of requirement determination, component collection, and application evaluation were described. The main contributions of this research are as follows:

1) The concept of Domain-Specific DT platform and a comprehensive evaluation framework for it have been proposed, which can be used as a reference for researchers, practitioners, and developers to build and choose DT platforms that meet the needs of a Domain-Specific application.
2) BWM and FCE methods are integrated to form a BWM-FCE evaluation model in the framework, which can effectively determine the weights and scores of the evaluation indicators and sub-indicators, and incorporate the subjective opinions of experts and stakeholders in the evaluation process.
3) The proposed framework is applied to a case study of the DT Train Platform and its rationality and practicality are verified. We have also compared the proposed framework with other evaluation methods and discussed the advantages and limitations of each method.

The current scarcity of DT platform resources across various industries has hindered the validation of the framework's multi-functional extension through case studies. Future research will involve comprehensive evaluations of DT platforms within diverse sectors, including manufacturing, healthcare, and automotive. These studies aim to showcase the framework's adaptability and its capacity to meet the evolving technological demands of DT, as well as the emerging needs of different industries. Presently, the study remains confined to the laboratory setting, thus bypassing ethical and privacy concerns. However, as the deployment and utilization of DT platforms expand into broader fields, ethical and privacy considerations will become paramount, necessitating dedicated research to address these issues.

### Future works

The primary limitations of the proposed framework lie in the evaluation indicators and sub-indicators, which can vary significantly across domains due to their unique characteristics. This creates challenges in quickly adapting the evaluation indicator system to evolving DT platforms. Future work will focus on extending and refining the framework to better accommodate different types of DT platforms and applications. Additionally, we will aim to explore new evaluation methods and criteria that can effectively capture the multidimensional and dynamic nature of Domain-Specific DT platforms. To verify the generalization of the framework in other industries, future work will focus on extending and refining the evaluation framework to accommodate different types of DT platforms and applications, as well as exploring more evaluation methods and criteria that can capture the multidimensional and dynamic nature of Domain-Specific DT platforms. Additionally, we intend to develop a more agile evaluation indicator system that can rapidly adapt to changes in DT Platforms, ensuring that the framework remains both relevant and precise.

Furthermore, we will explore the integration of modular design concepts, facilitating the incorporation of advanced analytical techniques such as time series analysis, artificial neural networks (ANN), and cluster analysis. This integration is expected to enhance the framework's analytical capabilities while promoting its adaptability and robustness, ultimately improving our evaluation processes in diverse contexts.

## Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 9 May 2024; Accepted: 3 December 2024
Published online: 27 March 2025

## References

1. Liu, M., Fang, S., Dong, H. & Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **58**, 346–361 (2021).
2. Kušić, K., Schumann, R. & Ivanjko, E. A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics. *Adv. Eng. Inform.* **55**, 101858 (2023).
3. Ferrari, A. & Willcox, K. Digital twins in mechanical and aerospace engineering. *Nat. Comput. Sci.* **4**, 178–183 (2024).
4. Roy, R. B. et al. Digital twin: Current scenario and a case study on a manufacturing process. *Int. J. Adv. Manuf. Technol.* **107**, 3691–3714 (2020).
5. Huang, J. *et al.* Deep reinforcement learning-based dynamic reconfiguration planning for digital twin-driven smart manufacturing systems with reconfigurable machine tools. *IEEE Trans. Ind. Inf.* (2024).
6. Angulo, C., Ortega, J. A. & Gonzalez-Abril, L. Towards a healthcare digital twin 1. In *Artificial Intelligence Research and Development*, 312–315 (IOS Press, 2019).
7. Katsoulakis, E. et al. Digital twins for health: A scoping review. *NPJ Digit. Med.* **7**, 77 (2024).
8. Sleiti, A. K., Kapat, J. S. & Vesely, L. Digital twin in energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems. *Energy Rep.* **8**, 3704–3726. https://doi.org/10.1016/j.egyr.2022.02.305 (2022).
9. Azfar, T., Weidner, J., Raheem, A., Ke, R. & Cheu, R. L. Efficient procedure of building university campus models for digital twin simulation. *IEEE J. Radio Freq. Identif.* **6**, 769–773 (2022).
10. Zhang, B., Zhang, M., Dong, T., Lu, M. & Li, H. Design of digital twin system for DC contactor condition monitoring. *IEEE Trans. Ind. Appl.* (2023).
11. Redeker, M., Weskamp, J.N., Rössl, B. & Pethig, F. Towards a digital twin platform for industrie 4.0. In *2021 4th IEEE international conference on industrial cyber-physical systems (ICPS)*, 39–46 (IEEE, 2021).
12. Lehner, D. et al. Digital twin platforms: Requirements, capabilities, and future prospects. *IEEE Softw.* **39**, 53–61 (2021).
13. ANSYS. Ansys twin builder. [EB/OL]. https://cae-expert.ru/product/ansys-twin-builder.
14. Mathworks. Matlab. [EB/OL]. https://ch.mathworks.com/discovery/digital-twin.
15. Microsoft. Azure digital twins. [EB/OL]. https://azure.microsoft.com/ruru/services/digital-twins.
16. Yildiz, E., Møller, C. & Bilberg, A. Demonstration and evaluation of a digital twin-based virtual factory. *Int. J. Adv. Manuf. Technol.* **114**, 185–203 (2021).
17. Siemens. Siemens revolutionizes engineering simulation with heeds ai simulation predictor and simcenter reduced order modeling. [EB/OL]. https://newsroom.sw.siemens.com/en-US/heeds-ai-simulation-predictor-simcenter-rom/.
18. Aivaliotis, P., Georgoulias, K. & Chryssolouris, G. The use of digital twin for predictive maintenance in manufacturing. *Int. J. Comput. Integr. Manuf.* **32**, 1067–1080 (2019).
19. Electric, G. Digital twin software. [EB/OL]. https://www.ge.com/digital/iiot-platform.
20. Cirillo, F. et al. A standard-based open source iot platform: Fiware. *IEEE Internet Things Mag.* **2**, 12–18 (2019).
21. Seok, M. G., Tan, W. J., Cai, W. & Park, D. Digital-twin consistency checking based on observed timed events with unobservable transitions in smart manufacturing. *IEEE Trans. Industr. Inf.* **19**, 6208–6219 (2022).
22. Liu, J., Zhou, H., Tian, G., Liu, X. & Jing, X. Digital twin-based process reuse and evaluation approach for smart process planning. *Int. J. Adv. Manuf. Technol.* **100**, 1619–1634 (2019).
23. Wei, Y., Hu, T., Wei, S., Ma, S. & Wang, Y. Digital twin technology applicability evaluation method for cnc machine tool. *Int. J. Adv. Manuf. Technol.* 1–17 (2022).
24. Talasila, P. *et al.* Comparison between the hubcap and digitbrain platforms for model-based design and evaluation of digital twins. In Cerone, A. *et al.* (eds.) *Software Engineering and Formal Methods. SEFM 2021 Collocated Workshops*, 238–244 (Springer International Publishing, Cham, 2022).
25. AboElHassan, A., Sakr, A. H. & Yacout, S. General purpose digital twin framework using digital shadow and distributed system concepts. *Comput. Ind. Eng.* **183**, 109534 (2023).
26. AboElHassan, A., Sakr, A. & Yacout, S. A framework for digital twin deployment in production systems. In *Advances in Automotive Production Technology–Theory and Application: Stuttgart Conference on Automotive Production (SCAP2020)*, 145–152 (Springer, 2021).
27. Jiang, Y. et al. Multi-domain ubiquitous digital twin model for information management of complex infrastructure systems. *Adv. Eng. Inform.* **56**, 101951 (2023).
28. Du, J. & Luo, T. Digital twin graph: Automated domain-agnostic construction, fusion, and simulation of iot-enabled world. arXiv preprint arXiv:2304.10018 (2023).
29. Han, C., Wang, Z., Fu, Y., Dyke, S. & Shahriar, A. Transfer-ae: A novel autoencoder-based impact detection model for structural digital twin. *Appl. Soft Comput.* 112174 (2024).
30. Robertson, S. & Robertson, J. *Mastering the requirements process: Getting requirements right* (Addison-Wesley, 2012).
31. Golden, B. L., Wasil, E. A. & Harker, P. T. The analytic hierarchy process. *Appl. Stud.* **2**, 1–273 (1989).
32. Forman, E. H. & Gass, S. I. The analytic hierarchy process-an exposition. *Oper. Res.* **49**, 469–486 (2001).
33. Rezaei, J. Best-worst multi-criteria decision-making method. *Omega* **53**, 49–57 (2015).
34. Dehshiri, S. J. H., Emamat, M. S. M. M. & Amiri, M. A novel group bwm approach to evaluate the implementation criteria of blockchain technology in the automotive industry supply chain. *Expert Syst. Appl.* **198**, 116826 (2022).
35. Uhlenkamp, J.-F. *et al.* Digital twins: A maturity model for their classification and evaluation. *IEEE Access***10**, 69605–69635. https://doi.org/10.1109/ACCESS.2022.3186353 (2022).
36. Tao, F., Zhang, H. & Zhang, C. Advancements and challenges of digital twins in industry. *Nat. Comput. Sci.* **4**, 169–177 (2024).
37. Tao, F. et al. maketwin: A reference architecture for digital twin software platform. *Chin. J. Aeronaut.*[SPACE]https://doi.org/10.1016/j.cja.2023.05.002 (2023).
38. Psarommatis, F. & May, G. A standardized approach for measuring the performance and flexibility of digital twins. *Int. J. Prod. Res.* **61**, 6923–6938 (2023).

39. Wang, Z., Zhou, Y., Jin, X., Zhao, N. & Sun, J. Risk allocation and benefit distribution of ppp projects for construction waste recycling: A case study of China. *Eng. Constr. Archit. Manag.* **30**, 3927–3956 (2022).
40. Shaw, K., Lahri, V., Shankar, R. & Ishizaka, A. Joint multi-item multi-supplier sustainable lot-sizing model applying combined bwm, topsis, possibilistic programming, and $\epsilon$-constraint method. *IEEE Trans. Eng. Manag.* (2023).
41. Youssef, A. E. An integrated mcdm approach for cloud service selection based on topsis and bwm. *IEEE Access* **8**, 71851–71865 (2020).
42. Yadav, G., Mangla, S. K., Luthra, S. & Jakhar, S. Hybrid bwm-electre-based decision framework for effective offshore outsourcing adoption: A case study. *Int. J. Prod. Res.* **56**, 6259–6278 (2018).
43. Tang, Z. *et al.* Towards digital twin trains: implementing a cloud-based framework for railway vehicle dynamics simulation. *International Journal of Rail Transportation* 1–24 (2024).
44. Tan, Y. et al. Comparison of ahp and bwm methods based on arcgis for ecological suitability assessment of panax notoginseng in yunnan province, china. *Ind. Crops Prod.* **199**, 116737 (2023).
45. Ajrina, A. S., Sarno, R. & Ginardi, R. H. Comparison of ahp and bwm methods based on geographic information system for determining potential zone of pasir batu mining. In *2018 International Seminar on Application for Technology of Information and Communication*, 453–457 (IEEE, 2018).

## Acknowledgements
Thanks to the experts and engineers who contributed to the determination of the framework indicators and the development of the DT Train Platform, and thanks to Mr. Hao Hu for his editing and proofreading work.

## Author contributions
Z.T.: Conceptualization, Investigation, Writing, Methodology and Resources. D.Z.: , Writing, and Software Development,Experiment(s). J.Z.: Review and validation. All authors reviewed the manuscript.

## Funding
This work of the paper was supported by the Fundamental Research Funds for the State Key Laboratory of Rail Transit Vehicle System of Southwest Jiaotong University (No. 2024RVL-T14), and Science and Technology Research Plan Project of China Academy of Railway Science Corporation Limited (2023YJ036) and technology research and development plan project of China State Railway Group Co., Ltd (No.K2023J008).

## Declarations

## Competing interests
The authors declare no competing interests.

## Additional information
**Correspondence** and requests for materials should be addressed to Z.T.