scientific reports

OPEN

A digital twin modeling and application for gear rack drilling rigs lifting system

Wang Jiangang[1,2], Shi Lei[1,2], Feng Ding[1,2]✉, Liang Jinli[1,2], Hou Lingxia[1,2] & Miao Enming[3]

A comprehensive digital transformation has been undergone by the oil and gas industry, wherein digital twins are leveraged to enable real-time data analysis, providing predictive and diagnostic engineering insights. The potential for developing intelligent oil and gas fields is substantial with the implementation of digital twins. A digital twin framework for gear rack drilling rigs is proposed, built upon an understanding of the digital twin composition and characteristics of the gear rack drilling rig lifting system. The framework encompasses descriptions of digital twin characteristics specific to drilling rigs, the application environment, and behavioral rules. The modeling approach integrates mechanism modeling, real-time performance response, instantaneous data transmission, and data visualization. To illustrate this framework, exemplary case studies involving the transmission unit and support unit of the lifting system are presented. Mechanism models are constructed to analyze dynamic gear performance and support unit response. Real-time data transmission is facilitated through sensor-based monitoring, enhancing the prediction speed and accuracy of dynamic performance through a synergy of mechanism modeling, machine learning, and real-time data analysis. The digital twin of the lifting system is visualized utilizing the Unity3D platform. Furthermore, functionalities on data acquisition, processing, and visualization across diverse application scenarios are encapsulated into modular components, streamlining the creation of high-fidelity digital twins. The frameworks and modeling methodologies presented herein can serve as a foundational and methodological guide for the exploration and implementation of digital twin technology within the oil and gas industry, ultimately fostering its advancement in this sector.

Keywords Drilling and completion engineering, Gear rack drilling rig, Digital twin, Real-time prediction.

The gear rack drilling rig represents a novel drilling machinery, wherein it replaces traditional lifting structures, such as cranes and winches, with a gear rack transmission system[1,2]. This system can achieve the elevation of the drilling column through the meshing of gears and racks. The modular design of the drilling rig enables it to be compact, highly interchangeable, and easily disassembled, thereby contributing to its digital technical characteristics[3]. In the face of the emerging trend of deep integration between the oil and gas industry and information technology, the utilization of digital twin technology for the realization of intelligent and unmanned drilling rigs has garnered significant attention. The digital technology, facilitated by real-time data analysis, plays a pivotal role in both the design and operational phases of the drilling rig, offering extensive possibilities for the industry[4,5].

Digital twins, characterized as a pioneering technology for comprehending, describing, and evolving target objects, find apt application in model-based services. Notably, digital twin prove beneficial in diverse areas such as production management[6,7], design optimization[8], performance monitoring[9,10], and personnel training[11]. Since the inception of the digital twin concept in 2003, Grieves[12] delineated its three integral components: physical entity, virtual twin, and connectivity. Tao Fei[13] expanded upon this framework by introducing the components of data and services, emphasizing the integration of data from multiple sources to yield more precise and comprehensive information. As simulation technology, the Internet of Things, and artificial intelligence evolve, scholars propose diverse ideas, design solutions, and functionalities for digital twins within their domains, addressing both process and function, and adopting various approaches and concepts for digital twins in design, monitoring, and operation.

[1]School of Mechanical Engineering, Yangtze University, Jingzhou 434023, Hubei, China. [2]Hubei Engineering Research Center for Oil and Gas Drilling and Completion Tools, Jingzhou 434023, Hubei, China. [3]College of Mechanical Engineering, Chongqing University of Technology, Chongqing 401135, China. ✉email:

Wang[14] proposed an interactive digital twin in various scenarios, integrating the digital twins of components or subsystems through assembly. This enriches the application environment and functionalities of digital twins. Aivaliotis[15] combines the degradation curve of structural component data in the mechanism model to evaluate the remaining service life of DT in predicting and maintaining robots. Farsi[16] proposed a method based on an ontology model that analyzes product lifecycle data, updates the ontology model, and assists designers in estimating costs during the early decision-making process. Liang[17] developed a digital twin considering the impact of vibration on machine tool structures. This model associates position variables with cutting forces, visualizing the effects of variables on machine tool vibration during the manufacturing process, thereby optimizing manufacturing accuracy and efficiency. Martín[18] addresses the challenge of creating digital twin fusion for multiple subsystems by integrating the data platform Kafka with Unity 3D. This promotes knowledge sharing, data interconnection, and the creation of complex digital twins in the petrochemical industry. Hu[19] conducted research on digital twin-driven manufacturing equipment, utilizing finite element and other simulation methods to provide training data for model construction methods and evolution mechanisms. This research culminated in the establishment of digital twins for manufacturing equipment. However, further improvements are needed to enhance the prediction accuracy of surface quality. SONG[20,21] proposed a method that combines dynamic compensation of collected data to enhance prediction speed. This approach is based on the results of structural static simulation calculations for real-time performance response in telescopic boom cranes. P. Stavropoulos[22]has developed a digital twin model integrating Multi-Body Simulation to depict the dynamic behaviors of industrial robotic machining processes.

Digital twin technology plays a pivotal role in advancing the digitalization of equipment design, monitoring, and predictive maintenance in the oil and gas industry. According to the Society of Petroleum Engineers (SPE)[23], digital twin technology can reduce operational failures by 20–25% and decrease engineering time by up to 70%.

Various approaches and concepts have been adopted for object-oriented digital twins in design, monitoring, and operation. In the realm of design, Naik[24] focused on digital twins of wave steel catenary risers. Utilizing a multi-objective optimization method integrating genetic algorithms and RBF, the optimal design of watertight pipe structural parameters is achieved, with the objective of minimizing fatigue damage. Li[25] designed a subsea tree, considering the influence of materials, structure, and process on the overall device performance. Addressing monitoring aspects, Anwesha[26] employed artificial intelligence to parametrically learn numerous drilling parameters, enabling real-time optimization and calibration of borehole trajectories. Zhang[27] discussed the technical advantages of digital twin technology in monitoring and predicting the performance behavior of pipelines and other devices, providing theoretical and technical references for storage and transportation equipment.

Edward[28] collected field data from offset wells and created a statistical database based on well depth and time. A drill bit digital twin for offset wells is developed, establishing the relationship between bit wear, drilling pressure, and well depth. NOV's digital twin for downhole drilling during the drilling process[23] utilized a large dataset, combined with artificial intelligence algorithms, to establish a drilling knowledge base, enabling real-time response prediction of optimal drilling pressure and mechanical speed. Maycon[29] created a wellbore digital twin, allowing real-time monitoring through multidisciplinary numerical simulation data to observe pressure and safety factors. However, despite these advancements, the exploration of constructing multi-layered, multi-scale digital twins specifically for drilling rigs continues to represent an unresolved and paramount challenge. This challenge necessitates addressing complex subsystem integration, real-time data transmission, and the accuracy of predictive models tailored to the dynamic and diverse operational environment of drilling rigs.

In this paper, a comprehensive digital twin framework tailored to the lifting systems of gear rack drilling rigs is proposed. In Sect. 3, the definition of constituent elements within these lifting systems is meticulously addressed, drawing upon digital twin characteristics, application environments, and behavioral rules. In Sect. 4, the intricate construction methodology, encompassing mechanism modeling, real-time performance response, data transmission, and visualization, is described. Case studies are presented in the subsequent sections, systematically elucidating the research methodology of the digital twin framework proposed within a specific context.

Digital twin framework for gear rack drilling rig lifting systems

A gear rack drilling rig is a complex system of mechanical, electrical and hydraulic coupling, and a multi-layered (parts-agencies-systems-equipment) complex equipment containing mechanical mechanisms (e.g., derrick and rig substructure, etc.), people, environment, and auxiliary hardware (pipe wrench, automatic catwalks). Drilling rigs typically consist of three components: run in hole, pull out of hole and drilling is usually completed by a large number of equipment/systems. Typical application scenarios, in terms of macro-model scales, contain management, and, in terms of functional rules, typical application scenarios contain equipment status data collection and monitoring and monitoring. The three contents have multiple scales and scenario characteristics, and the elements appear cyclically and iteratively during the drilling process, which involves the coupling of the systems of the drilling rig; therefore, the construction of the digital twin of the gear rack drilling rig requires multidomain model coupling, multiscale construction methods, and multifunctional interactive modeling.

Digital twin composition for gear rack drilling rig lifting system

The research focuses on monitoring the operational behavior of the gear rack drilling rig lifting system during drilling operations, with a digital twin development framework illustrated in Fig. 1. The process begins with Entity Collection and Data Inputs, where sensors and data acquisition systems are installed on the drilling rig to gather operational data such as structural loads and motion states. Mechanism Calculation and Data Processing follow, involving initial data analysis to understand the rig's operational principles, including subsystem equipment coupling and performance responses to process flows. Subsequently, dynamic calculations

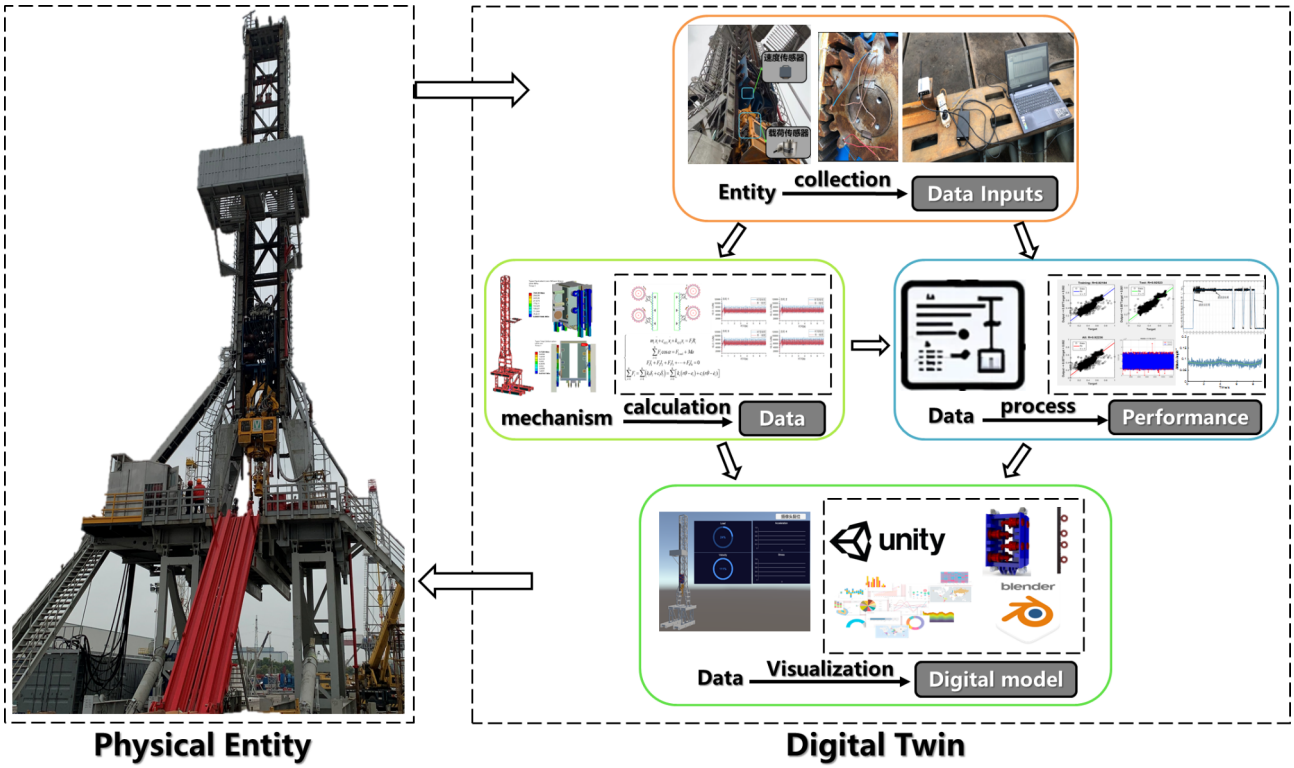

Fig. 1. Digital twin framework of rack and pinion drilling rig.

are performed, incorporating data analytics and machine learning algorithms to model the rig's response under motion excitation, resulting in a performance agent model. The processed data is then visualized and integrated into a Digital Model environment using tools like Unity and Blender. This step facilitates the creation of high-fidelity digital representations of the rig, combining geometric monitoring data with computational insights for real-time monitoring and interactive performance analysis. By integrating these stages, the framework ensures a comprehensive approach to monitoring, analyzing, and visualizing the gear rack drilling rig's performance, meeting the functional requirements of its digital twin.

Digital twin characteristics of gear rack drilling rig lifting systems

Digital twin technology for oil and gas drilling and completion entails creating a digital representation of physical objects and systems, such as drilling rigs, drilling strings, wellbores, and process flows. The digital twin aim to closely mirror their physical counterparts in real time, ensuring dynamic consistency during long-term operations. Notably, digital twin technology for oil and gas equipment systems exhibits several technical characteristics, including closed-loop control (real-virtual interaction control), multi-scaling (parts, equipment, units, systems), multi-physical fields (flow, solid, electricity, etc.), and multi-sourcing (data from tests/simulations), maintaining high fidelity to the physical entities. Additionally, the technology considers the effects of randomness (uncertainty in physical properties, loads, environment, etc.), complexity (functions of different units/bodies, operating characteristics, etc.), dynamics (location, real-time data updates, etc.), and interference (noise, transmission packet loss, etc.). For instance, in the context of a complex rack rig lifting system, Eq. (1) expresses the key elements of the digital twin.

$$X_{DT}=(X_E, X_M, X_D, X_C, X_F, X_V)$$
$$\begin{cases} X_E=(E_{closed-loop}, E_{trend}, E_{state}, E_{random}, \ldots) \\ X_M=(M_{transmission}, M_{support}, M_{couple}, \ldots) \\ X_D=(D_{speed}, D_{load}, D_{noise}, D_{splice}, \ldots) \\ X_C=(C_{wireless}, C_{wired}, C_{multiple}, C_{delay}, \ldots) \\ X_F=(F_{meshingforce}, F_{stressprediction}, F_{Filter}, \ldots) \\ X_V=(E_{component}, E_{behavioralrule}, E_{chart}, \ldots) \end{cases} \quad (1)$$

Here, X_E, X_M, X_D, X_C, X_F, and X_V represent the physical entity of the lifting system, the mechanism model, data detection, data transmission, data fusion, and the digital twin, respectively. Each element exhibits technical characteristics such as multi-scaling, multi-disciplinarity, and multi-physical field capabilities.

The data-driven digital twin technology can make real-time response information fusion in the process monitoring of the drilling rig lifting system. In terms of design, the design scheme can be evaluated and calibrated through the coupling and constraint relationships between the internal components of the system, replacing the mechanism model with a long design cycle, identifying possible defective problems in the design process, and then optimizing the design; in terms of process monitoring, real-time diagnosis is carried out in terms of the performance and life span of the physical entity based on the response of the mechanism model of the digital twin under different working conditions of the object-oriented, real-time prediction, discover and solve possible problems in advance, adjust and optimize the drilling strategy to reduce the problem of abnormal downtime. Digital twin technology has the following characteristics:

1. Fidelity: The digital twin is a digital model created with physical entities as objects, which is used to simulate and map the behavior of physical entities of research objects in drilling and completion engineering. The accuracy of the function, working conditions and performance of the lifting system is a necessary factor to ensure the fidelity of the digital model, the twin describes the physical entity not only contains geometric properties but also needs to consider the coupling of equipment and process technology in multi-disciplinary, multi-physical field, to ensure that the digital twin can truly reflect the rules of behavior of the lifting system and its intrinsic response.
2. Real-time: the digital twin and the physical entity through the sensor and the system simulation data acquisition, data fusion and other stages of data interaction, and then realize the real-time response and mapping of the digital twin to the physical entity. The data transmission capability and real-time resource scheduling capability are necessary factors to guarantee the timeliness of the digital twin, such as the collected data (device output data, detection data) combined with system simulation technology to achieve a dynamic response to the state of the physical entity.
3. Fusion: The digital twin gathers multi-dimensional and massive data such as information acquisition, system simulation, expert knowledge, etc., which involves the correlation data between various components of oil and gas equipment. The digital twin can clean, integrate, classify and reorganize the historical data of wellsite, real-time data, design standards and operation specifications to form a new dataset, which is reconstructed to be composed of the correlation data of the consistent rules of the research object, so that more information can be obtained from the new data set based on the reduced data scale. The data set reconstruction consists of associated data with consistent rules for the study object, and more information is obtained from the new data set on the basis of reducing the data scale.
4. Visualization: the digital twin visualization should be able to visually reflect the progress of oil and gas well drilling and production, the state of the equipment, especially the work that is difficult to observe directly downhole. Based on real-time data to establish dynamic data visualization graphics, take multi-scale modeling of equipment and processes, establish intuitive links between parts within the equipment, with the help of data graphics, visualization of data will be clear and effective display.
5. Expansion: The digital twin of the gear rack drilling rig can be enriched by combining the digital twins of each internal system. It can be integrated into the digital twin of a single well site by integrating the geological, equipment, process and other twins of a single well site, and it can also be integrated into the digital twins of multiple well sites to achieve remote central monitoring through remote communication.
6. Decision-making: the digital twin can achieve two-way communication with physical entities, combine sensor detection data, mechanism model data to take system simulation, agent model and other algorithms to describe the intrinsic mechanism of the physical entity, identify the key faults of the physical entity and the knowledge base, implement iterative updating based on the measurement data to achieve the closed-loop function of intelligent decision-making optimization of the physical entity, predict the abnormal downtime events and guide which parameters need to be adjusted.

Digital twin context of gear rack drilling rig lifting systems

The complex digital twin has the integrated characteristics of multi-physics, multi-module and multi-system. The modeling of the digital twin of the gear rack drilling rig integrates the idea of modularity, decomposing the rig into individual simple objects and also merging the relationships between these objects into a whole. The digital twin divides the individual functional components, but the complete digital twin generation is composed of model components, application environment, behavioral rules and also is the result of the modular integration of the individual modular-simple digital twin modularity. For the complete digital twin, the spatial scale consists of multiple layers from systems to parts, and each layer usually contains multiple application scenarios. The application scenarios are constructed based on the rules of behavior of different equipment and components.

Multi-level model

The multi-level modeling approach of the gear rack drilling rig lifting system is shown in Fig. 2, which constructs the physical model of each multi-level based on the spatial structural relationship and functional coordination relationship of the equipment, which can be divided into: system, unit, equipment and parts. equipment features and part dimensions are adopted, and the parts are divided and clustered by the way of the smallest subset. Digital twin multi-level modeling of the gear rack drilling rig is achieved using the parts and functional content as the basis. Specifically, the system is the gear rack drilling rig, which reflects the overall equipment state, management and other macro states. The unit layer includes transmission unit, support unit, etc. These units are composed of multiple entities, which reflect the collaboration between the equipment. The equipment layer includes derrick, lifting box, etc., which realizes the functions of pipe string lifting and fixing. The part layer focuses on the micro-scale, including derrick vibration, gear meshing state, etc.

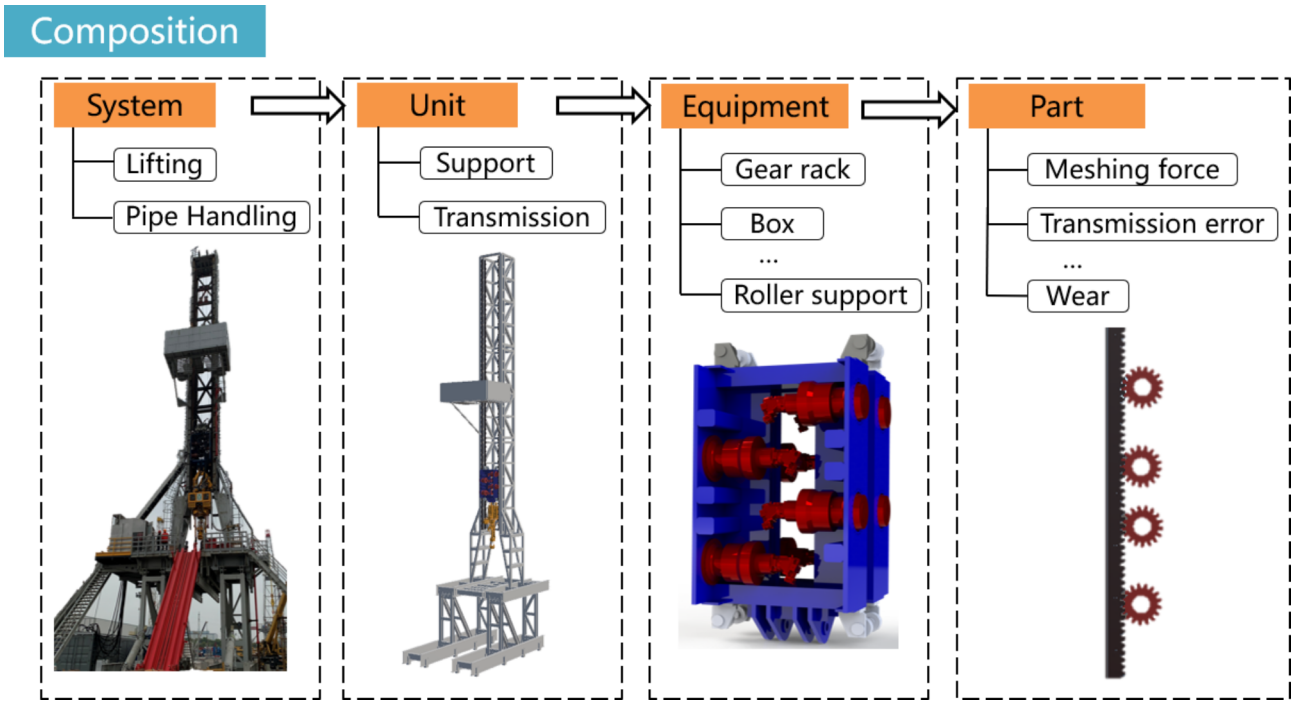

Fig. 2. Multi-level modeling of lifting system.

Application environment

The application environment of the gear rack drilling rig is related to the processes of raising and lowering drills and drilling wells in drilling engineering. The digital twin model modeling content is determined in the process monitoring of the rig lifting system. For example, during the drilling process, the lifting system making a connection based on the depth of the drill column to ensure the continuity of the pipe column lowering. At the transmission unit level, the input real-time position is used to determine the depth of the column and output control commands. At the equipment level, in the mesh state monitoring in the rack and pinion transmission mechanism, the load, position, acceleration and other signals are collected to predict the gear mesh state, and real-time prediction of the mesh force is carried out by combining artificial intelligence and deep learning algorithms. application environment contexts are included in each layer.

Behavioral rule

Behavioral rule is the process of reflecting functions such as data characteristics, motion gestures and physical rules of physical entities to the digital twin. Based on the structure, function and assembly relationship of the physical entity of the equipment. Spatial geometric position and motion fit relationships are set up. The operational state of the real physical entity is reflected to the model behavior in digital space. Combine the physical parameters such as mass, material and friction coefficient of physical entities with the laws of kinematics (e.g., inertia law, etc.), laws of mechanics (e.g., Hooke's law, etc.) and laws of motion, etc., to define the rules of synchronous mapping of the model in digital space. Based on the fact that the functional implementation of digital twins for gear rack drilling rigs relies on the combination of different types of rules, encapsulating the commonly used functions of digital twins into components helps in the rapid modeling and implementation of digital twins.

Components are usually built according to the way of "data input - data processing - data output", which correspond to the basic model, function and visualization respectively. Programming and module integration is one of the fastest ways to develop components, and the code languages include c#, c++, Python and so on. Components are regarded as services to accomplish specific functions in the process of modeling behavioral rules, so the development of components is not limited to a specific programming language, but to achieve the required functions. The programming languages and software that should be commonly used for the development of data input components, data processing components and data output components. SolidWorks and Blender are used to develop geometric models of physical entities that are imported into Unity3D to build virtual scenes. The data input component is used to collect and transfer data. c# and python are used to develop the data collection program related to the device, and Redis and MySQL are used to transfer and store the data after acquisition. The data processing component is used for simulation, prediction, optimization and other functions. Ansys and Abaqus are both commonly used finite element simulation platforms. Deep learning algorithms and optimization algorithms are developed using Pytorch, Python and Matlab. The Data Output component is used to visualize data and models.

The components and connections used in the digital twin of the lifting system are shown in Fig. 3. In the offline deep learning model training phase, the historical data collected by the sensors are stored in a data pool,

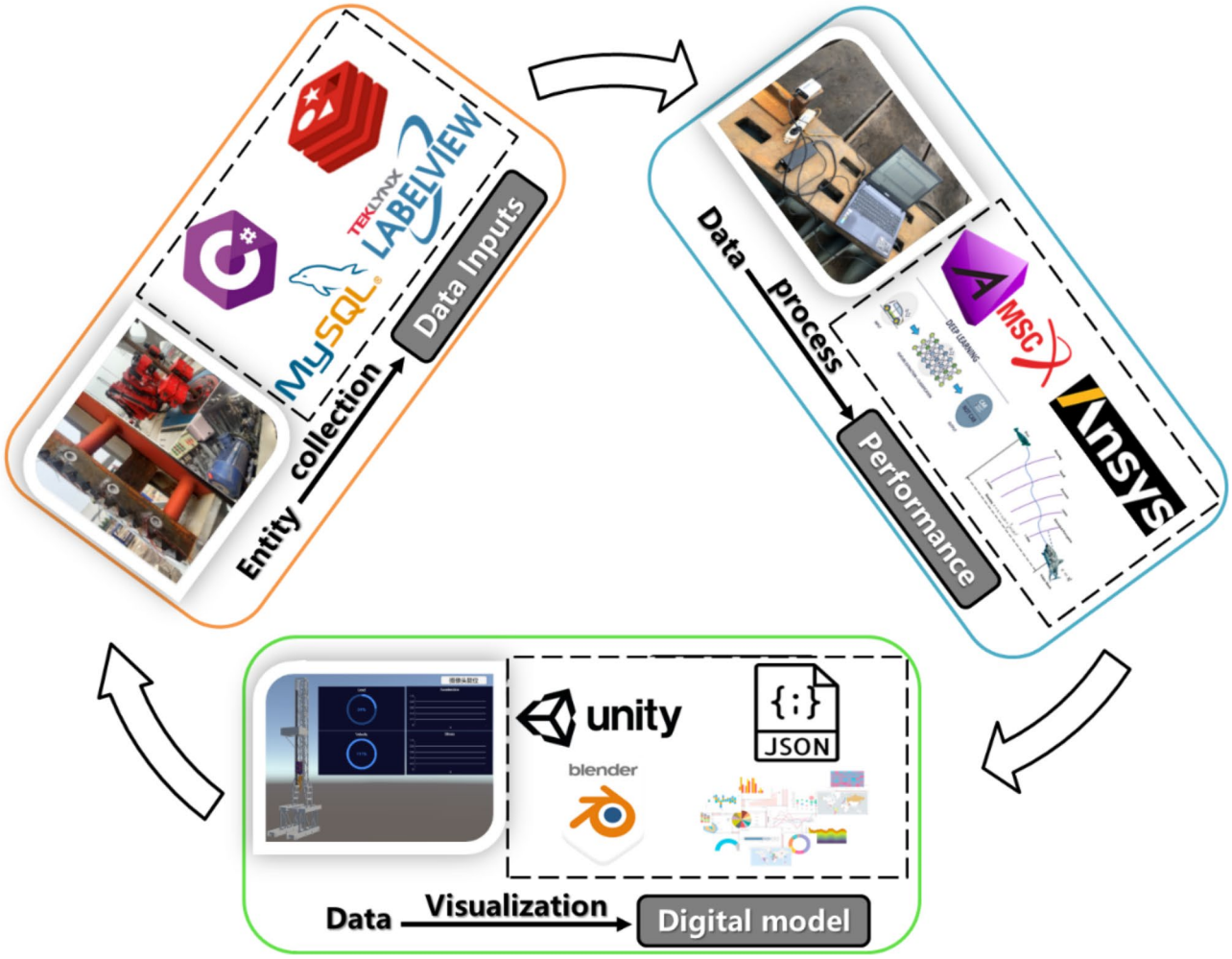

Fig. 3. Models behavioral rule components.

and after data preprocessing, they are used to train a variety of deep learning models[30], including BP Neural Network (BPNN), Long Short-Term Memory (LSTM), and Residual Network (ResNet). Then the deep learning network with the highest accuracy rate is selected as the prediction model. In the model performance prediction phase, the real-time data is also preprocessed and fed into the selected deep learning model to output the structural performance response.

Model iteration

Real-time mapping of physical entity states via digital twin models is essential for ensuring high accuracy. The iterative updates, based on real-time models, constitute the core technology of digital twin symbiosis evolution. Building upon the definition of modeling methods, application environments, and behavioral rules outlined previously, a database containing structural parameters, operational performance, and rule features is established for parts, mechanisms, units, and systems. The model undergoes iterative real-time updates facilitated by standardization, parameterization, and classification methods. For instance, the finite element modeling method is utilized to analyze stress and displacement in the drilling rig structure. Consequently, real-time mapping of the digital frame model to physical entities is achieved, wherein the calculation results of the digital model and its geometric structure, material properties, and performance responses are corrected through test data integration.

Construction method for gear rack drilling rig lifting system

In this paper, with reference to the 5-dimensional DT model and the division of digital twin body models and functions, we propose a digital twin model construction method based on the expandability of the association interaction process for the single-function digital twin composition or application involving only one scene. These simple digital twins are expected to be assembled into multi-functional, multi-scene digital twins through multi-scale association, data processing and multi-context interaction. Driven by sensor updates and historical data, all the different parts of the physical system are modeled and mirror synchronized. In addition, correlated but single-function digital twins should be able to interact with each other to achieve more complex functionality, which in turn enables the digital twin to be continuously extended and enriched.

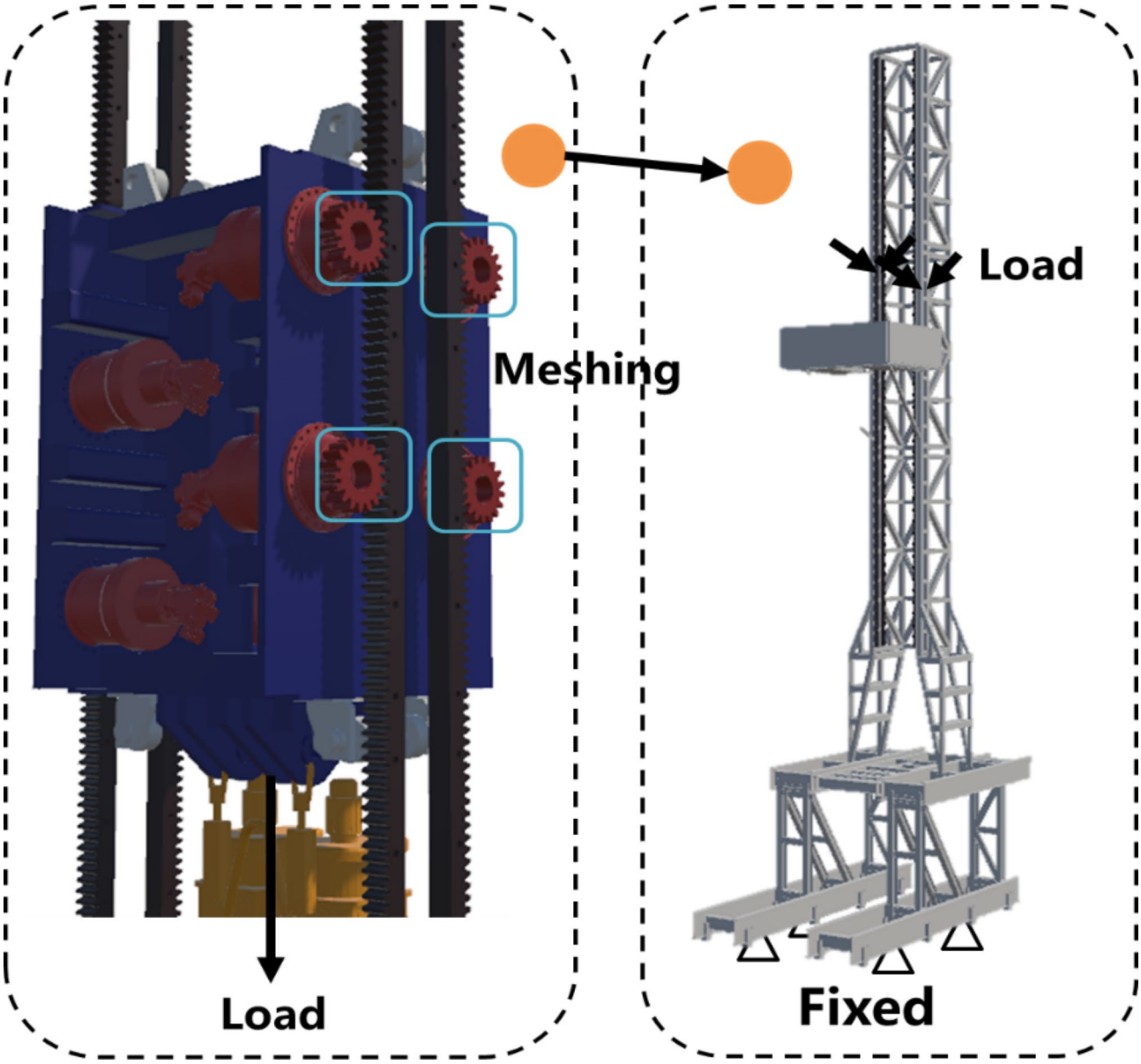

Fig. 4. Lifting system mechanism model.

Mechanism model

The prerequisite for reliability analysis is the creation of a reasonable theoretical model, which divides the composition of the gear rack drilling rig lifting system according to its function, distinguishing between the transmission part of the movement and the support part that ensures its correct movement, which mainly consists of the transmission unit and the support unit. Support unit (derrick and base): It is the key mechanism of the gear rack drilling rig lifting system, which is used to support the lifting of the transmission system and all the weight of the drilling string in the well and on the first floor, and to carry out the drilling operations of lifting and lowering the drilling tools and lowering the casing, etc. The transmission unit (lifting box) is the key mechanism of gear rack drilling rig transmission and power output, in which the gears are installed on both sides of the hoisting box, and the internal drive consists of hydraulic motors, couplings reducers and other components, as shown in Fig. 4. The up and down drive of the pipe column is achieved by meshing with the gear mounted on the derrick. This paper ignores the factor of fasteners and seals, structural fits that are considered reliable.

Based on the construction content, the specific construction method of the dynamic mechanism model is proposed. The dynamic characteristics of the working process of the lifting system are affected by many factors, such as the position of the lifting box, the load, and the transmission speed. It is necessary to find a clue to associate these multidimensional variables for the high-fidelity description of the working process. Firstly, a structural decomposition of the lifting system is performed to describe the evolution of the system performance parameters. Various factors in a complex system can evolve on different time scales, for example, when the

position and excitation of the transmission unit change during operation, the stiffness matrix, damping matrix and mass matrix of the corresponding system structural dynamics undergo relevant changes.

The transmission unit (gear, rack and pinion), as a key part of the transmission unit, can reflect the change of dynamic excitation in the transmission process, and this is used as one of the boundary conditions of the support unit. The dynamic equations of gear meshing excitation force can be described as in Eq. (2):

The support unit, as a form of energy concentration in the working process, can reflect the changes in the dynamic characteristics of the working process. The dynamic equations of vibration can be described as in Eq. (3):

$$\begin{cases} m_i\ddot{x}_i + c_{i(t)}\dot{x}_i + k_{i(t)}x_i = F_iR_i \\ \sum_{i=1}^{8} F_i\cos\alpha = F_{load} + Ma \\ F_1l_1 + F_2l_2 + F_3l_3 + \cdots + F_8l_8 = 0 \\ \sum_{i=1}^{8} F_i = \sum_{i=1}^{8}\left(k_l\delta_l + c_t\dot{\delta}_l\right) = \sum_{i=1}^{8}\left[k_i\left(r\theta - e_i\right) + c_i(r\dot{\theta} - \dot{e}_i)\right] \end{cases} \tag{2}$$

$$m\ddot{x} + c\dot{x} + kx = F_{(t)} \tag{3}$$

M is the total mass of the transmission unit, F_{load} is the load-bearing capacity of the lifting device, F_i is the meshing force of each gear (i = 1–8), a is the acceleration of the lifting device, l_i is the distance from each gear to the center of gravity of the lifting device, m_i is the mass of each gear, R_i is the indexing circle of each gear, and $k_{i(t)}$ is the stiffness of each gear.

Real time performance response

In order to accurately describe the dynamic characteristics of the lifting system process, many sensors need to be arranged for data acquisition, and it is too costly to obtain the global performance of the lifting system through the sensors due to the harsh working environment, the complexity of the system structure, and the cost constraints. In the research of performance prediction of oil and gas major equipment, the data-based modeling approach is a new way to solve the performance prediction problem of complex equipment. Mechanical models and collected data are organically combined instantaneously to describe the mechanical properties of equipment real-time response[31,32]. Physical entity-based mechanistic models for data processing and prediction of dynamic structural properties. A machine learning based data model for tracking the multi-timescale evolution of system parameters in the mechanism model and capturing the actual change characteristics. The performance parameter library of the lifting system is built by using multi-technological means such as sensors, data transmission, and data processing techniques, and the real-time data-based performance evaluation system is established by combining the highly adaptive algorithmic model in machine learning. The performance evaluation indexes such as response time and maximum load are selected to achieve the performance prediction, structural parameter optimization and reliability analysis of the lifting system.

The performance response model is constructed with a recommended number of samples at least ten times the number of variable types. Based on the drive unit contents of the lifting system, input variables such as transmission unit position, load, and speed are defined. Structural performance is calculated using a finite element model and dynamics[3]. For instance, Adams is employed to compute meshing forces in gear transmissions, with these forces serving as boundary conditions input into ANSYS for stress analysis of the supporting elements. The results of structural analyses from these samples serve as training data for the machine learning model[33]. In finite element calculations, increasing the number of nodes benefits the structural performance model. However, higher node counts can challenge the accuracy of performance predictions. Equation (4) represents the performance response model, achieved by predicting the performance of each node and integrating to derive overall model performance.

$$M = \left\{P^1_{(1,2,\ldots,i)}, P^2_{(1,2,\ldots,i)}, ., P^n_{(1,2,\ldots,i)}\right\} \tag{4}$$

where M is the performance prediction data, P is the performance calculation model, P^n represent different models, $P_{(1,2,\ldots,i)}$ represent different boundary input conditions.

Real-time data transmission

Data acquisition and transfer is the technical method used to exchange data between the digital twin physical entities and the digital model. The types of data can be classified into ontological model, boundary attributes, and state attributes. The ontology model provides a general description of the physical entity, such as name, geometry, and weight. Boundary attributes describe the real-time data of the physical entity, including the overall operating status, position, angle, speed, etc. State attributes are mainly related to the physical structure and technical performance of the equipment, as well as the dynamic data of the model, such as structural mechanical performance, dynamic response, fatigue, etc. One of the key points of real-time data mining is to obtain complete, stable and effective data through reasonable sensor placement. This paper uses data that includes state data such as load and speed. Several types of sensors were already included in the initial design process of the drilling platform. The data is transmitted to the driller's room, and only the data interface needs to be connected

for reading the state data, that is, PLC connection. The second key point of real-time data mining is to achieve the fusion between the mechanism model of the gear rack drilling rig, the performance response model and the collected data. Based on filtering, modal decomposition and other means to achieve the segmentation and noise reduction of the collected data. In the process of information fusion, some data are redundant and need to be cleaned. For example, even in each single-function digital twin with the same equipment name, these same data need to be recorded only once. The effect of time-varying factors during the operation of the gear rack drilling rig is taken into account to update the information about the status of the gear rack drilling rig during operation. Combined with the performance real-time response component for dynamic calculation and evaluation, the realization process can be expressed with Eqs. (5–6):

$$M_P = \{D_{historical},\ D_{mechanism},\ D_{state},\ P\} \tag{5}$$

$$M_C = \{D_{online},\ D_{mechanism},\ M_P,\ P_C\} \tag{6}$$

where, $M_P$denotes the model composed of historical data, mechanism and state information of the gear rack drilling rig together with equipment characteristics, Historical, D_{online}, D_{state}, P, P_C denotes historical data, current data, mechanism model, state information, historical equipment characteristics (multi-scale, multi-disciplinary, multi-physical field and other parameters) and current equipment characteristics, respectively, and denotes the dynamic updating model under the drive of current data. Real-time data transmission is achieved under the condition of ensuring that the data information is not lost or damaged, and the authenticity of the data is maintained to the maximum extent. Through TCP, socket and other communication protocols for data exchange, to achieve a unified dynamic interaction and management of data. Extract the physical information of the lifting system through sensors and other data collection technologies, and store the collected data in JSON, XML and other development tools.

Visualization

Visualization of performance data can be observed intuitively is a key part of the digital twin model building process[34]. The use of three-dimensional cloud diagrams, charts and graphs to characterize the structural performance information of its components is able to ensure that the state variables collected in the physical entity are accurately transferred to the virtual digital space. Under the premise of ensuring model fidelity, local optimization of the rig lifting system model improves the image quality. Multi-detail stereoscopic display is achieved to obtain a realistic 3D model. Dynamic visualization in the digital twin implements the function of real-time mapping of mechanism model, performance response and data to dynamically present the data to the personal computer in real time.

The digital model of the gear rack drilling rig lifting system has the computational characteristic of a large number of nodes, which to some extent leads to the rate delay of the overall visualization. In this paper, for the large model visualization rate acceleration, with the key components of the lifting system (derrick and base), a multi-level 3D visualization method is adopted. In macroscopic modeling by surface stress/deformation information, internal node information is ignored to reduce the number of node modeling and accelerate the visualization time. In local model visualization, since the number of nodes is small at this point, the internal node information is supplemented to achieve full information visualization of the performance 3D model. The mechanical properties of the structure are dynamically presented to the PC in real time.

Case study

Scenario 1: structural performance prediction of transmission units

Application environment Model

The gear rack drilling rig transmission unit is characterized by low speeds and heavy loads, and the meshing performance of the gear rack is critical to the reliability of the transmission unit. Generally, the rack and pinion meshing force can be obtained by measurement and simulation. Wired and wireless strain gauges can directly measure the gear or rack contact, but they are affected by the moving position, noise and so on. In addition, the gear transmission meshing force can be simulated by dynamics simulation calculation, but for the unit of 8 gear-4 rack transmission, the dynamics cannot output the meshing performance in real time. For this reason, a method for predicting the performance of gear meshing force based on dynamics simulation calculation combined with machine learning is proposed. The results of dynamics simulation calculations are used as sample data for machine learning. The gear mesh performance prediction model trained with machine learning is used to output the gear mesh performance in real time.

According to the modeling approach proposed in this paper, the transmission unit is composed of a rack and pinion mechanism, lift box body parts, and the function is the prediction of rack and pinion meshing performance. In this case, the inputs include hoist box load, operating position, transmission speed and operating time, and the output is the predicted gear meshing force. The implementation of the transmission unit mesh performance prediction function can still be divided into the development of four components: mechanism model, real-time performance response, real-time data mapping and visualization. In addition to multiple data acquisition components and visualization components, data processing components, including a dynamics-based simulation calculation module and a BPNN performance prediction component are required to predict rack and pinion transmission mesh force performance.

Transmission unit component

The gear transmission performance prediction model is illustrated in Fig. 5. Initially, a physical model of the lifting box is constructed. Subsequently, dynamic simulation calculations are performed on the geometric model of the lifting box, with boundary conditions such as load and transmission speed set to obtain meshing force. These boundary conditions and simulation results are converted into sample data for training BPNN. The trained transmission unit can predict gear meshing performance under arbitrary working conditions. Data collection serves as the initial step in gear mesh performance prediction. In order to monitor the status of the lifting box, data from multi axis acceleration sensors and force sensors are read for digital volume mapping entities. A data acquisition component is developed for this purpose, utilizing C# and Python languages. In this setup, C# receives wireless and PLC data signals from the sensors, which are then forwarded to Python for data denoising. This component calculates the position and velocity of the lifting box based on acceleration data, facilitating the synchronization of physical transmission unit data with the digital model.

The trained performance prediction model is subsequently utilized to develop a data processing component, implemented in Python. The following steps are undertaken: Firstly, real-time parameters such as speed, position, and running time of the transmission unit are converted into real-time data and inputted into the trained transmission unit. Mean Absolute Error (MAE) and R-Squared (R^2) are employed to quantitatively evaluate the gear meshing force prediction algorithm. The formulas for and R^2 are presented in Eqs. (7–8).

Finally, the data output component is developed in Unity3D. The visualization component also includes charts and geometric models. Charts are drawn by using plug-ins (UGUI). Specifically, drilling rig parameters and environment parameters are displayed as line charts, bar charts and dashboards. And the data information is detailed in the text. The geometric model is transmitted to Unity via socket communication in the TCP protocol, based on data acquired by the data acquisition module. Within the software, the geometric model employs C# script mode to control the display of position and speed, thereby achieving real-time synchronization between the digital model and the physical entity's movement.

$$MAE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i) \tag{7}$$

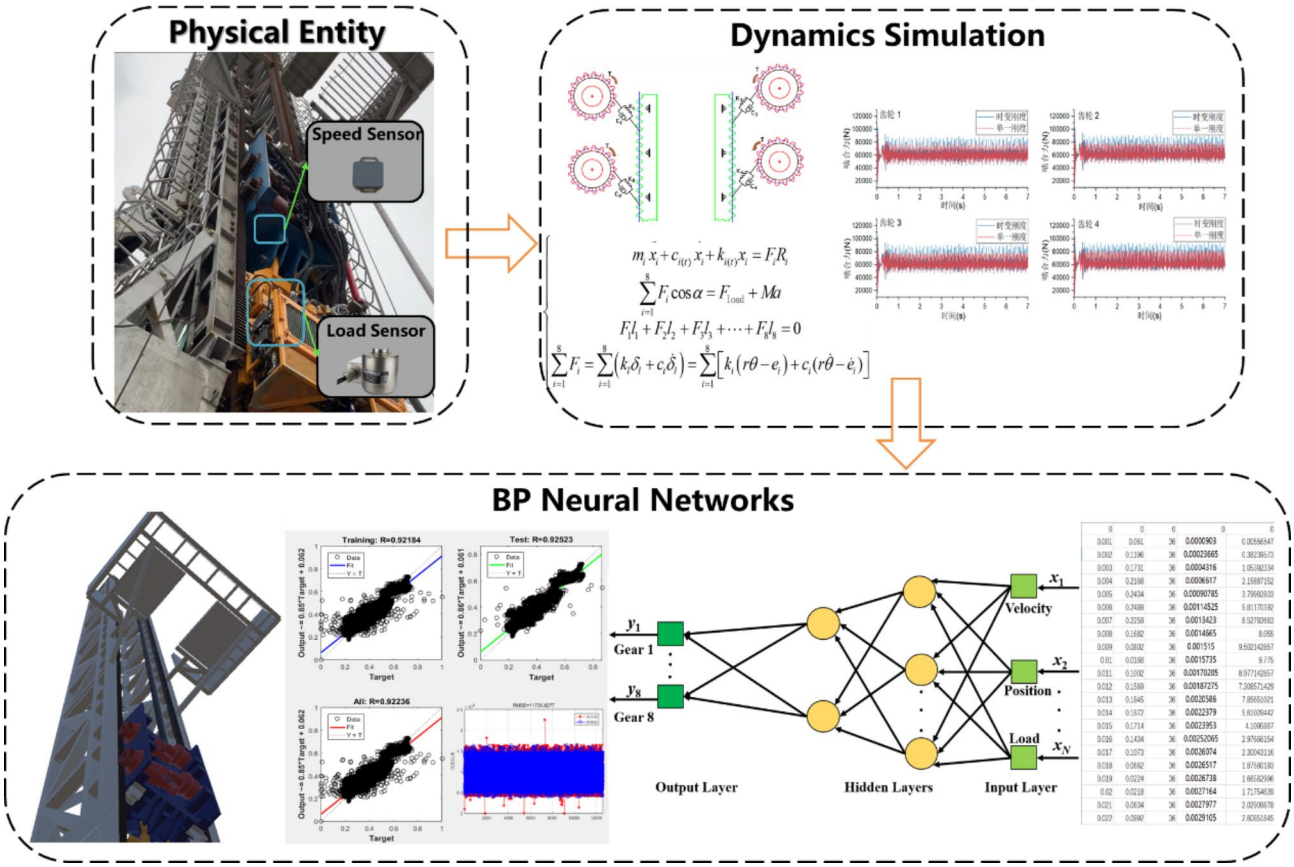

Fig. 5. Process of gear meshing force prediction.

$$R^2 = 1 - \frac{\sum_{1}^{m} (y_i - \hat{y}_i)^2}{\sum_{1}^{m} (y_i - \bar{y})^2} \tag{8}$$

Model output

The visualization interface of the gear rack transmission unit is illustrated in Fig. 6. Utilize speed, position, and load as inputs, and output the meshing force of each gear as the target. The BPNN is trained using the Bayesian regression algorithm, employing a learning rate of 0.01 and a maximum of 500 iterations. The gear meshing force prediction model reached convergence conditions after 80 training sessions. Through comparison and analysis

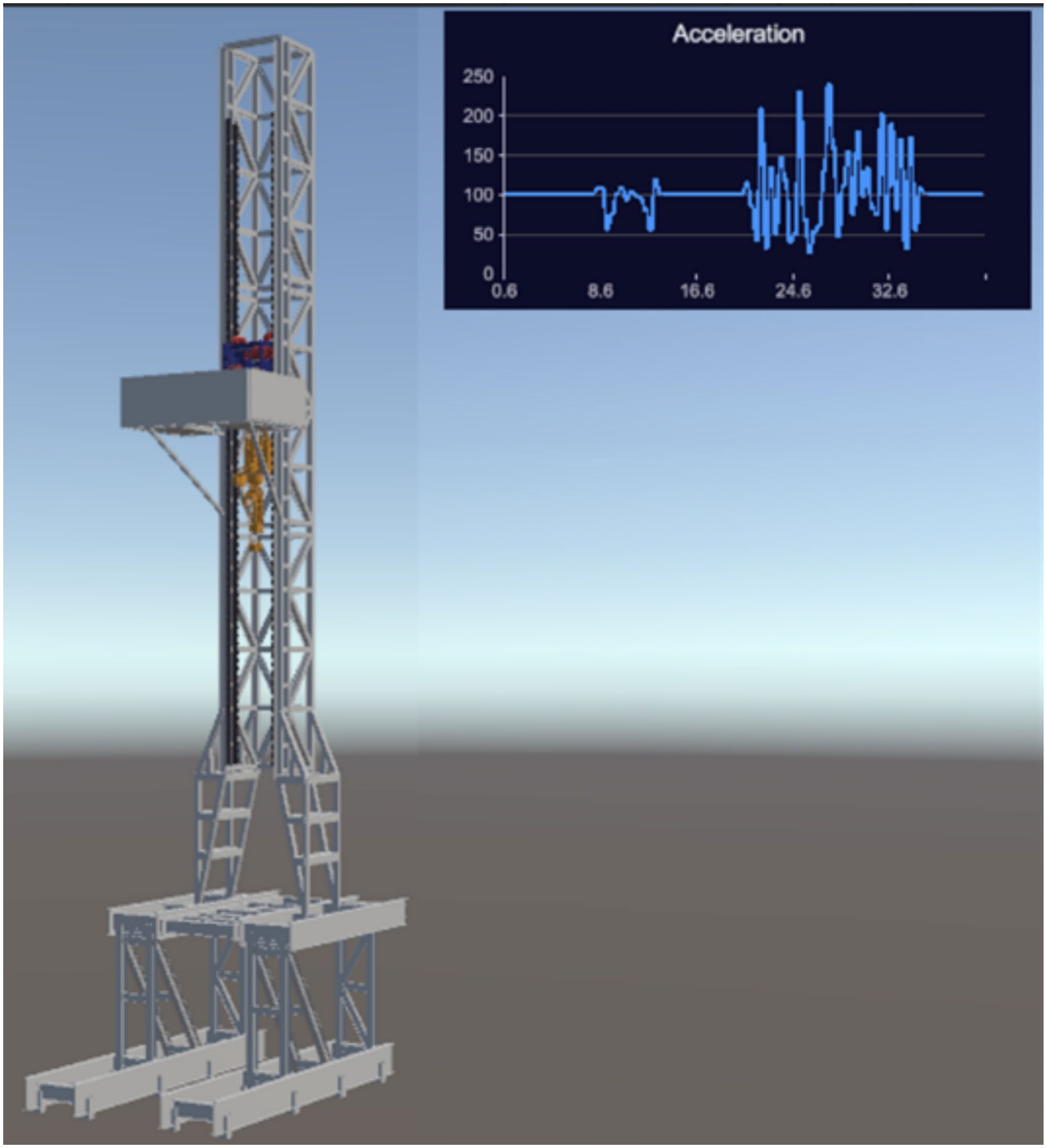

Fig. 6. Visualization of transmission unit.

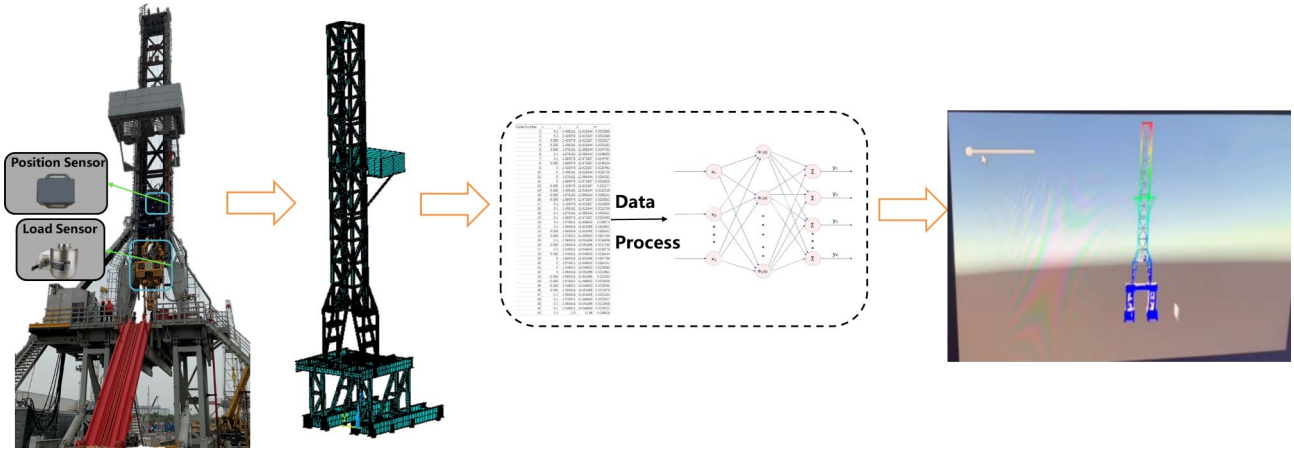

Fig. 7. Process of support system stress distribution prediction.

of the pre-divided test set data, the MAE of the prediction model is 0.87, and the R^2 value is 0.8457, surpassing the engineering requirement of 0.8.

Scenario 2: support unit structure performance prediction

Application environment Model

The support unit is a key component that affects the stability and reliability of drilling rig processing, and the load of the support unit is affected by the working environment, operating time and other factors. The traditional inspection usually adopts the method of periodic inspection, which is unable to predict the abnormal working conditions during the working process. There are two main modes of existing inspection methods: the direct method and the indirect method. In the direct method, stress/strain or displacement transducers are used to measure the support unit loads. Although the direct method has high measurement accuracy, the support unit is more than 10 m high and the harsh working conditions increase the maintenance cost of the measurement. In the indirect method, optical sensors are used to measure the support unit deformation. The indirect method is also through the data processing to feedback the load state of the support unit, and by the environmental light influence degree is strong. Therefore, this paper chooses the direct measurement method with data processing to reflect the overall load state of the support unit.

Support unit component

According to the modeling approach proposed in this paper, the component in this case is the part layer, whose function is the prediction of the load distribution of the support unit. In this case, the inputs include lifting load, lift box position, and the output is the predicted support unit load distribution. Due to the limited availability of experimental data, a finite element analysis model of the supporting element is established using finite element analysis method, and its effectiveness is verified by experimental data. Based on sample data, and utilizing the KNN + RBF machine learning algorithm, the stress distribution of support units is predicted. Establish different dynamic response models of drilling rigs through finite element analysis to obtain sample data. By combining data with KNN algorithm for position node data classification and RBF neural network for data regression, real-time prediction of stress distribution in training support units is completed.

Model output

The visualization interface of the support unit is shown in Fig. 7. The parameter settings for the KNN model specify selecting the nearest data as 3 and returning the nearest data as 15. Based on the load and position, which represent the input parameters, numerical model prediction results are obtained respectively. After training, out of sample data was compared within the interval, with a MAE of 0.1866 and a model R^2 of 0.9081, which is greater than the required 0.8 in engineering.

Conclusions and discussion

This paper introduces a novel approach to developing a digital twin for the lifting system in gear rack drilling rigs, focusing on both the support unit and transmission unit. Our methodology stands out for its integration of a five-dimensional digital twin framework and a modeling method incorporating multi-scale, multi-layer modeling, and environmental behavior rules. Unlike existing works, our construction framework emphasizes versatility, reusability, and scalability. The study concludes with the following key findings:

1. Digital Twin Construction Framework: The framework utilized for gear rack drilling rigs integrates a five-dimensional digital twin framework and a modeling method. It includes content definitions encompassing multi-scale, multi-layer modeling, and the application of environmental behavior rules. Construction methods involve the classification of mechanism models, performance responses, real-time data transmission, and visualization.

2. Application Scenarios: Two distinct application scenarios for the lifting system are formulated, focusing on predicting the structural performance of the transmission unit and the support unit. For each scenario, data acquisition, data processing, and model output identification are conducted. A simple digital twin is encapsulated into reusable components, categorized into three groups: data input for acquisition, core functionality for data processing, and visualization for data output. Importantly, data input and output components are designed for reusability across multiple scenarios.
3. Detailed Modeling Process: The construction of a digital twin for the gear rack drilling rig lifting system is presented, offering a comprehensive and implementable modeling process. The digital twin emphasizes different spatial scales and encompasses application scenarios with one or more kinetic energies, providing a realistic representation of the digital twin rig. Incorporating reusable components and interaction across diverse scenarios facilitate the extension of complex digital twins.

This study presents a detailed digital twin modeling process for the gear rack drilling rig lifting system, emphasizing versatility, reusability, and scalability. The constructed digital twin framework has significant potential for various applications, thereby advancing digital twin technology within drilling systems. Unlike gear and rack drilling machines, where loads are fixed at the Crown Block atop the rig, the digital twin model simplifies by ignoring the influence of transmission unit positioning on support units during its creation. Furthermore, the model focuses on the elastic stage, adjusting structural and material stiffness coefficients to simplify dynamic responses while overlooking the effects of wear and cracks, which intuitively affect stiffness. Despite impacting model accuracy, future research will explore integrating elastic-plastic behaviors to enhance fidelity.

Data availability

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Received: 22 March 2024; Accepted: 23 September 2024
Published online: 10 October 2024

References

1. Xu, C. Dynamic performance analysis of Jack-up platform's jacking system with staggering tooth. *J. Mech. Eng.***50**(19), 66–72 (2014).
2. Na, L. et al. *Meshing Characteristic Analysis of the Gear Rack Drilling Rig Hoisting System* (China Journal of Mechanical Transmission, 2015).
3. Wang, J. et al. Study the muti-bolt fastening under different load positions in gear rack drilling rig. *PLOS ONE*. **18**, e0290427 (2023).
4. Sircar, A. et al. Digital twin in hydrocarbon industry. *Petroleum Res.* (2022).
5. Papacharalampopoulos, A., Foteinopoulos, P. & Stavropoulos, P. Integration of Industry 5.0 requirements in digital twin-supported manufacturing process selection: a framework. *Procedia CIRP*. **119**, 545–551 (2023).
6. Gao, Z. & Hyder, S. Z. *Using Autonomous Control to Stabilize Well Performance without the Downhole Pressure Gauge* (ADIPEC, 2022).
7. Stavropoulos, P. et al. *Metamodelling of Manufacturing Processes and Automation Workflows towards Designing and Operating Digital Twins* (Applied Sciences, 2023).
8. Feder, J. BHA-design approach improves drilling performance and wellbore quality. *J. Petrol. Technol.***72**(12), 57–58 (2020).
9. Karpov, R. B. et al. Drilling Performance and Data Quality Control with Live Digital Twin. SPE Russian Petroleum Technology Conference, (2021).
10. Stavropoulos, P. et al. Multi-body dynamic simulation of a machining robot driven by CAM. *Procedia CIRP*. **107**, 764–769 (2022).
11. Burrafato, S. et al. Virtual Reality in D & C: New Approaches Towards Well Digital Twins. Offshore Mediterranean Conference and Exhibition, (2019).
12. Grieves, M. & Vickers, J. Digital Twin: Mitigating unpredictable, undesirable Emergent Behavior in Complex systems, (eds Kahlen, F-J., Flumerfelt, S. & Alves, A.) Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, Cham: Springer International Publishing, : 85–113. (2017).
13. Qi, Q. et al. Enabling technologies and tools for digital twin. *J. Manuf. Syst.***58**, 3–21 (2019).
14. Jia, W., Wang, W. & Zhang, Z. From simple digital twin to complex digital twin part II: Multi-scenario applications of digital twin shop floor. *Adv. Eng. Inform.***56**, 101915 (2023).
15. Aivaliotis, P. et al. Degradation curves integration in physics-based models: towards the predictive maintenance of industrial robots. *Robot. Comput. Integr. Manuf.***71**, 102177 (2021).
16. Farsi, M. et al. A digital twin architecture for effective product lifecycle cost estimation. *Procedia CIRP*. **100**, 506–511 (2021).
17. Liang, Z. et al. The process correlation interaction construction of Digital Twin for dynamic characteristics of machine tool structures with multi-dimensional variables. *J. Manuf. Syst.***63**, 78–94 (2022).
18. Robles, J., Martín, C., & Díaz, M. OpenTwins: An open-source framework for the development of next-gen compositional digital twins. *Comput. Ind.***152**, 104007 (2023).
19. Wei, Y. et al. Digital twin-driven manufacturing equipment development. *Robot. Comput. Integr. Manuf.***83**, 102557 (2023).
20. He, X. et al. *M-LFM: A multi-level fusion modeling method for shape-performance Integrated digital twin of complex structure* (Frontiers of mechanical engineering, 2022).
21. Lai, X. et al. Designing a shape–performance Integrated Digital Twin based on multiple models and dynamic data: a Boom Crane Example. *J. Mech. Des.***143**(7), 1–15 (2021).
22. Stavropoulos, P. et al. Integration of Machining Process Digital Twin in Early Design Stages of a Portable Robotic Machining Cell. 31st International Conference on Flexible Automation and Intelligent Manufacturing, : 301–315. (2023).
23. Jeffery, C. & Creegan, A. Adaptive drilling application uses AI to enhance On-Bottom drilling performance. *J. Petrol. Technol.***72**(08), 45–47 (2020).
24. Bhowmik, S., Naik, H. & Noiray, G. Lazy Wave Steel Catenary Riser Design Optimisation in Digital Field Twin. Offshore Technology Conference Asia, (2020).
25. Wei, L. et al. Applications of Digital Twins to Offshore Oil/Gas Exploitation: from visualization to Evaluation. *IFAC-PapersOnLine*, 738–743. (2020).
26. Mal, A. et al. Drilling parameter optimization in real-time. SPE/IADC International Drilling Conference and Exhibition (2023).

27. Zhang, L. & Wang, J. Intelligent operation and maintenance technology of oil & gas storage and transportation equipment based on industrial internet. *Oil Gas Storage Transp.***41**, 625–631 (2022).
28. Gharib Shirangi, M., Furlong, E. & Sims, K. S. Digital Twins for Well Planning and Bit Dull Grade Prediction. SPE Norway Subsurface Conference, (2020).
29. Vieira Dos Santos, M. et al. Development and Deployment of Digital Twin for Production and Well Integrity. SPE Annual Technical Conference and Exhibition, (2022).
30. Lu, C. et al. Shale oil production prediction and fracturing optimization based on machine learning. *J. Petrol. Sci. Eng.***217**, 110900 (2022).
31. Sun, Q. et al. Operation strength analysis of HXJ180 Offshore Workover Rig based on API Standard. *China Petroleum Mach.***50** (11), 58–65 (2022).
32. Wang, K. et al. Influence of rack side bending on rack transmission. *China Petroleum Mach.***47**, 55–66 (2019).
33. Chen, X. et al. A real-time drilling parameters optimization method for offshore large-scale cluster extended reach drilling based on intelligent optimization algorithm and machine learning. *Ocean Eng.***291**, 116375 (2024).
34. Liu, S., Bao, J. & Zheng, P. A review of digital twin-driven machining: from digitization to intellectualization. *J. Manuf. Syst.***67**, 361–378 (2023).

Author contributions

F. D. conceptualized this study and performed project administration. S.L.wrote the proposal and secured funding for this project. W. J. developed the methodology for this study. L. J. and H. L. collected data for test data. W. J. performed the simulations, data analysis, interpretation and visualization of results. W. J. wrote the original draft of the manuscript with key inputs from L. J. , F. D. , S.L. and M. E. critically reviewed and edited the manuscript.

Funding

This work of the paper was supported by the Major National Science and Technology Projects [2016ZX05038-002-LH001] and study on Nonlinear Coupling Vibration Mechanism and Intelligent Control Method of Steering Tool Combined Bearing-Rotor System [52204002].

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-73954-z.

Correspondence and requests for materials should be addressed to F.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024