Device Health Status Assessment Under the Influence of Multiple Exception Modes

Xue-feng Yuan[1], Fei-long Liu[1], Yong-jun Qie[1], Shuai Sun[1], Jie Ren[1]

Sany Heavy Industry Co. LTD, Digital Twin Research Institute, Beijing 102200, China

yuanxuefeng112@gmail.com
15618589645@163.com
qieyj2@sany.com.cn
suns10@sany.com.cn
renj28@sany.com.cn

ABSTRACT

Equipment reliability is the key feature to ensure the equipment operation for a long time. It is difficult to determine the overall reliability of industrial equipment due to the different reliability states of different subsystems. A device abnormality identification method based on JS (Jenson's Shannon) divergence and a health status assessment technology based on FMECA (failure mode, effect and criticality analysis) are proposed. This method enables an accurate assessment of the current health status of the device. First, the historical operation data is preprocessed according to the characteristics of the equipment to improve the data quality. The JS divergence method is reused to extract the similarity between the key feature data distribution and the benchmark data distribution. Then, the FMECA report is established using the real running data of the device combined with expert experience. Gray theory was used to determine the degree of association between one-way health state membership vector and different health state rank vector. Finally, the health status level was comprehensively evaluated by the fuzzy membership method. Taking the mechanical arm component of a 100-ton crane as an example, the results show that this method can effectively evaluate the current health state of the equipment, and provide power for the abnormal advance disposal and auxiliary management decisions.

1. INTRODUCTION

Preventive maintenance, predictive maintenance and other means have gradually become an important means to reduce maintenance costs under the premise of ensuring equipment reliability. Fault prediction and health management technology is to predict the possibility of equipment failure in a future period of time, and take reasonable maintenance measures to reduce the overall cost. This method has the functions of data collection and sorting, condition monitoring and early warning, fault diagnosis, fault prediction, health management, life estimation (Tan, 2021). Fault prediction and health management techniques first originated in the military field, such as helicopter Health and Usage Monitoring system, aircraft condition monitoring system, US Army Diagnostic Improvement Program (Wang, 2016);. In the 1990s, the condition-based maintenance requirements were put forward in the joint strike fighter projects of the US Army and the British Army, and the fault prediction and health management technology was formally proposed. In the military field, the application effect and technical level of fault prediction and health management technology are the highest. Among them, the most representative is the F-35 fighter failure prediction and health management system, which reduces the maintenance manpower by 20% to 40%, and reduces the total logistics support cost by 50% (Zhao, 2019; Qin, 2016; Hu et al., 2017; Yu et al., 2018).

Fault prediction is a theoretical and technical means to estimate the type, time, probability and location of equipment failure based on various internal and external factors such as equipment running characteristics, parts and components materials and stress characteristics, equipment working environment. The accuracy of fault prediction is based on reasonable and scientific physical, mathematical and other models. In order to improve the accuracy of prediction information and maintenance measures of engineering equipment, it is necessary to comprehensively and deeply analyze all kinds of influencing factors of engineering equipment faults, including operating

https://doi.org/10.36001/IJPHM.2023.v14i2.3533

environment, composition structure, working principle, fault characteristics, fault classification (Pan et al., 2021; Zhang et al., 2017).

In recent years, the fault diagnosis and Health assessment methods of machinery equipment have emerged one after another, which has a positive impact on the construction machinery. Scholars have used deep reinforcement learning (Zhong et al., 2023), machine learning algorithms (Bhat et al., 2023), principal component analysis (Bencheikh et al., 2020), wavelet packet decomposition (Hamadouche et al., 2018) and other methods to realize the fault diagnosis of cement rotary kiln. The diversity of PHM algorithms and the complexity of design factors make it challenging to choose an appropriate algorithm for a specific application (Zou et al., 2023). At the same time, the faults of transformer (Zhang et al., 2021), gear (Goswami et al., 2023), centrifugal pump (Goncalves et al., 2021) involved in the construction machinery have also been widely studied by scholars.

Health assessment is the key technology of equipment health management. By evaluating the current health state of the equipment, that is, the ability of the system to maintain a certain reliability and continue to complete the predetermined function, it provides a decision-making basis for the implementation of condition-based maintenance, reduces the probability of equipment failure, and then improves the safety, reliability and availability of the equipment. At present, there are mainly two types of research methods related to system health assessment: one is based on component failure analysis, and the other is based on system performance degradation data . For engineering equipment with long life and high reliability, limited by economic cost and maintenance technology development, it is difficult to obtain complete failure data samples of equipment, and it is relatively easy to obtain performance degradation data.

To achieve the equipment health assessment for complex systems, the system needs to be split, and the comprehensive assessment can only be carried out after the subsystem health assessment is completed. Even at the component level, there are different types of equipment failure modes. Therefore, the key technical difficulties in the realization of equipment health assessment lie in two aspects: one is the identification of failure mode and abnormal state under less data; the other is the comprehensive equipment health assessment that integrates multiple abnormal states.

In order to solve the above two key problems, the evaluation method of device health status assessment under the influence of multiple exception modes is proposed to comprehensively evaluate the equipment health status. The innovation point of this method is that it integrates the anomaly recognition technology driven by data and expert experience, fully discovers abnormal patterns, and conducts comprehensive health evaluation of equipment under multi-anomaly patterns, which solves the limitation problem of unilateral anomaly recognition method and the difficulty of comprehensive evaluation of multi-anomaly patterns.

2. METHODOLOGY

2.1. Anomaly recognition method

2.1.1. Overview

The anomaly of the equipment refers to the deviation of the equipment operation from the normal range, which leads to the degradation of the equipment performance or failure. The equipment anomaly recognition mainly includes two aspects, one is the determination of the exception mode and its severity measure, the other is the identification of the abnormal state and the calculation of the probability of occurrence. The abnormal mode reflects the degree of deviation from the normal state during the operation of the equipment, and the identification of the abnormal state of the equipment determines whether it is abnormal. There are many methods for anomaly identification, and this paper focuses on the analysis of methods based on real-time data collected at the device end. We determine the exception modes and severity according to the characteristics of the equipment, and mainly use the JS divergence of key features to distinguish the normal and abnormal state. The exception mode recognition method is added by combining the mechanism logic abnormal judgment and the confidence interval abnormal judgment method. We use the above method to identify the abnormal state and calculate the probability of occurrence in the period. Our equipment health assessment flow chart is shown in Figure 1.

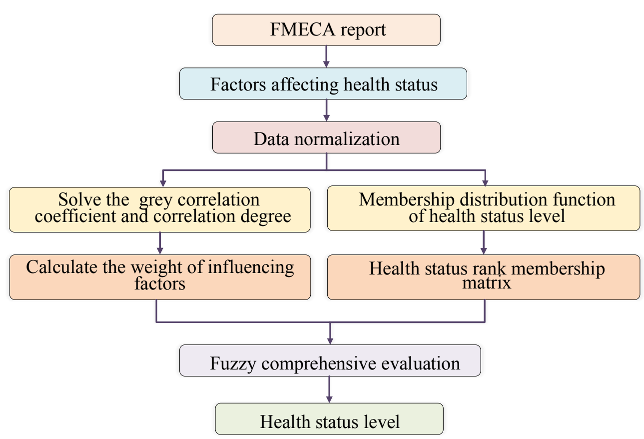


Figure 1. Equipment health assessment flow chart

2.1.2. JS divergence

The basic principle of the abnormal identification method based on JS divergence: JS divergence is mainly used to measure the similarity of the probability distribution of two sets of data. This method sets the distribution of a set of normal data as a reference benchmark, and calculates the JS divergence between the data distribution of other periods and the benchmark distribution. The higher the JS divergence, the more normal it is.

JS divergence measures the similarity of two probability distributions, which was developed from information entropy and KL divergence. This method is based on a variant of KL divergence, which solves the problem of asymmetric KL divergence. The entropy of a random variable is Eq.(1). KL divergence can calculate the similarity of two probability distributions, which is defined as Eq.(2).KL divergence has two properties:

1. $D_{\mathrm{KL}}(P,Q) \geq 0$
2. It doesn't satisfy the symmetry, with $D_{\mathrm{KL}}(P,Q) \neq D_{\mathrm{KL}}(Q,P)$. So choose carefully as a measure of the gap between the two distributions.

The asymmetry of the KL divergence can sometimes cause problems, so to achieve symmetry, the JS divergence is calculated as Eq.(3).

$$H[x] = -\sum_{x} p(x)\log p(x) \tag{1}$$

where $p(x)$ is the probability distribution of the variable, and x is a discrete random variable.

$$\begin{aligned} D_{\mathrm{KL}}(P,Q) &= -\int p(x)\log q(x)dx - (-\int p(x)\log p(x)dx) \\ &= -\int p(x)\log\frac{q(x)}{p(x)}dx \end{aligned} \tag{2}$$

where $p(x)$ and $q(x)$ are probability distributions of two variables.

$$D_{\mathrm{JS}}(P,Q) = \frac{1}{2}D_{\mathrm{KL}}(P,M) + \frac{1}{2}D_{\mathrm{KL}}(Q,M) \tag{3}$$

where $M = 1/2(P+Q)$. According to Eq. (3), three properties of JS divergence can be obtained:

1. JS divergence can be applied to sets of more than two probability distributions;
2. JS divergence values are nonnegative and bounded;
3. The JS divergence is symmetric with respect to the parameter order, that is, $D_{\mathrm{JS}}(P,Q) = D_{\mathrm{JS}}(Q,P)$.

2.1.3. Anomaly detection by exponential smoothing

Exponential smoothing method is used for short-term trend prediction, which is a time series analysis and prediction method developed on the basis of moving average method. Exponential smoothing usually has a third-order model, and the first-order model can predict the next data in real time according to the recent historical data, and the further away from the current time, the less influence of historical data. Compared with the first exponential smoothing method, the second exponential smoothing method makes another smoothing, and this method can realize the seasonal trend prediction. The cubic exponential smoothing method solves the problem of periodic superimposed seasonal data fluctuation prediction for the trend prediction of the quadratic curve, and the calculation formula is Eq. (4) .

$$\begin{aligned} &Y_{t+T} = a_t + b_t T + c_t T^2 \\ &S_t^{(3)} = aS_t^{(2)} + (1-a)S_{t-1}^{(3)} \\ &a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)} \\ &b_t = \frac{a}{2(1-a)^2}[(6-5a)S_t^{(1)} - 2(5-4a)S_t^{(2)} + (4-3a)S_t^{(3)}] \\ &c_t = \frac{a}{2(1-a)^2}[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}] \end{aligned} \tag{4}$$

where a is the smoothing factor, y_{t+1}' is the predicted value in $t+1$ period, which is the smoothed value S_t in this period (t period). y_t is the actual value in period t ; y_t' is the predicted value in period t , which is the smoothing value $S_{t\text{-}1}$ in the previous period. Y_{t+T} is the predicted value at period $t+T$; T is the number of periods that the t period moves backward; $S_t^{(2)}$ is the second exponential smoothing value of the t period; $S_{t\text{-}1}^{(2)}$ is the second exponential smoothing value of the t - 1 period; $S_t^{(3)}$ is the cubic exponential smoothing value of the t period.

The predicted value of the next cycle is calculated according to the exponential smoothing third-order model, and then the residual between the predicted value and the actual value is calculated. If the residual exceeds the threshold, the anomaly is marked. The threshold can be determined by the n-times standard deviation method and the confidence interval method.

2.1.4. Mechanism approach

The most important difference between big data in industrial environment and Internet big data is the extraction of data features. The characteristics of industrial big data have more physical meaning and the mechanism logic of the correlation between the characteristics. Therefore, according to the constraints of the physical mechanism, we make full use of the coupling relationship between sensor measurement data to construct logical rules to form the criterion.

2.1.5. Data pre-processing

Data quality problems are caused by external disturbance, sensor problems and network anomalies of the data collected on the actual site. There will be some phenomena such as the constant value of the numerical duration and the data obviously exceeding the operating range of the equip. Therefore, it is necessary to filter and preprocess the data. The criterion that the maximum value is equal to the minimum value in the period is adopted to eliminate the outliers with constant value, as shown in Eq. (5).

Significant outliers were eliminated by reliability calculation based on 3sigma criterion, as shown in Eq. (6). The physical boundary method is used to filter the parameters, and the conditions are satisfied as shown in Eq. (7).

$$\begin{aligned} & if\ Max[V_n(t)] = Min[V_n(t)]\ t \in [t_l, t_h]\ n = 1,2,3,\cdots n \\ & \quad delete\ V_n(t) \end{aligned} \tag{5}$$

Where $Max[V_n(t)]$ and $Min[V_n(t)]$ are the maximum and minimum value in the current period, and $[t_l, t_h]$ is the time cycle range, and n is to select parameters of different dimensions.

$$\begin{aligned} & if\ \left|V_n(t) - AVG(t_l, t_h)\right| > 3 \times STD(t_l, t_h) \\ & \quad delete\ V(t) \end{aligned} \tag{6}$$

$$V_l < V_n(t) < V_h \tag{7}$$

where $AVG(t_l, t_h)$ and $STD(t_l, t_h)$ are the average value and standard deviation within the current period, and V_l and V_h are the boundaries of parameters respectively.

2.2. Health status assessment methods

There are many factors that affect the health of equipment, including the reliability of the equipment itself, including external environmental factors and environmental factors, so it is difficult to make a quantitative description directly. Fortunately, the FMECA method can obtain the influencing factors of the equipment from normal to failure by agreeing on the failure mode and impact analysis of the equipment. Therefore, we can use the FMECA method for equipment state health assessment.

FMECA is an inductive analysis method that analyzes all possible failure modes and their possible effects, and classifies them according to the severity and occurrence probability of each failure mode. It is a single factor analysis method. FMECA consists of two parts: Failure Mode and Effect Analysis (FMEA) and criticality Analysis (CA). CA can only be performed on the basis of FMEA. FMECA is an important work item of product reliability analysis, and it is also the basis of carrying out maintenance analysis, safety analysis, test analysis and support analysis.

Grey relational analysis method can measure the degree of association between factors according to the degree of trend similarity between factors. This paper applies this method combined with fuzzy comprehensive evaluation to quantify the influence of different exception mode dimensions, abnormal severity and frequency dimensions on health status. Figure 2 is a schematic diagram of device health assessment based on FMECA.

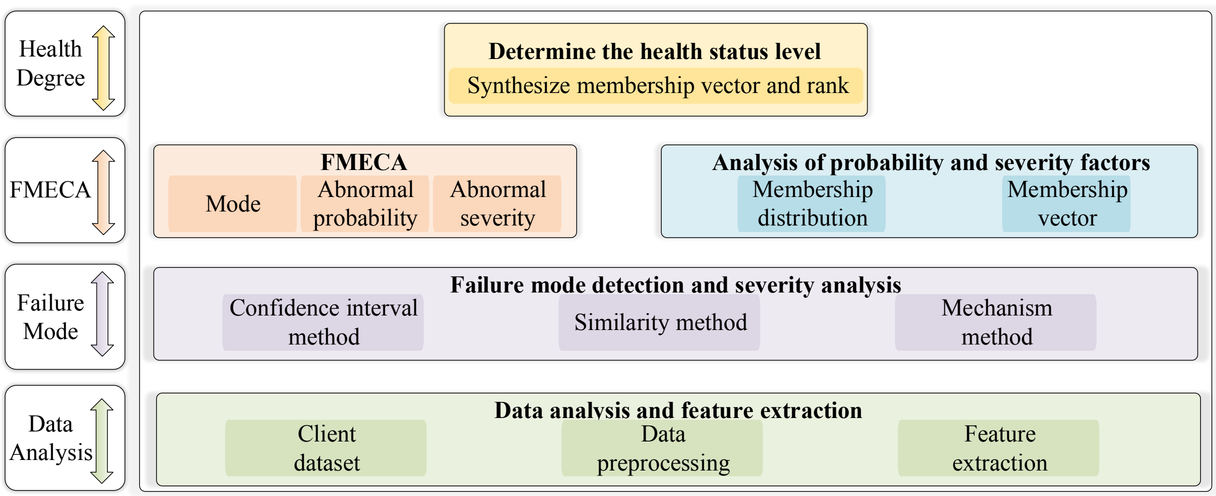


Figure 2. Device health assessment based on FMECA

2.2.1. Health status classification

The purpose of health status classification is to realize health status assessment and serve for fault prediction and maintenance decision. Health status classification should match the purpose of health status assessment. If it is only to judge whether the equipment is good or bad, it can be divided into two levels: normal and fault. If the state warning is included, the health state can be divided into three levels: normal, attention and deterioration. More levels of division can also be achieved for other purposes.

There are several steps and methods to classify and confirm the health status level.

1. Select the state feature parameters: The selection of characteristic parameters should screen multiple performance parameters of the equipment, and select the parameters that can be used to characterize the health state of the equipment.
2. Analyze the law of abnormal development: For different exception modes, the evolution process should be analyzed for different equipment and component

mechanism characteristics, and the evolution law of data and data characteristics should be analyzed.

3. Determine the anomaly threshold: Different characteristic parameters of the equipment have different normal ranges in different operating states. The normal value, attention value and fault value should be determined according to the mechanism, statistical distribution of data and other methods.
4. Dividing health states: According to the evaluation purpose and classification principle of equipment health status, we reasonably determine the number of health levels and the definition of each level, and give the judgment rules. Combined with the test data, we analyze the rationality of the grade division and adjust it appropriately to adapt to the target.

2.2.2. Evaluation parameter

We use the device sensors and related measurement parameters to monitor the device state and thus determine different device exception modes. At the same time, the probability of abnormal occurrence and its impact will affect the equipment health status assessment. Therefore, in this paper, we use the exception mode probability and exception mode harshness as the input of equipment health state assessment.

The exception mode probability refers to the occurrence probability of different exception modes monitored in a unit cycle, which is mainly calculated by the sample data in the collection cycle. The smaller the abnormal probability is, the better the health state of the equipment is.

Exception modes harshness refers to the severity of the consequences caused by the occurrence of this exception mode on the device. For different exception modes, the impact on the device is not the same. Since the device is composed of different subsystems, and different subsystems are composed of different components, the exception modes also belong to different predetermined levels. The lower the harshness, the smaller the impact and the healthier the device. The severity of abnormal modes is mainly assessed by experts.

2.2.3. Evaluation parameter

Normalization of influencing factors: The factors affecting the health status of equipment are diverse, which will lead to the problem of dimensional inconsistency. In order to carry out quantitative analysis, the influencing factors need to be normalized. Eq. (8) is used for index normalization, so that the dimensional data becomes the dimensionless data. For non-digital indicators, expert scoring can be used to determine.

$$x_i^{'} = \frac{x_i - x_i^{\min}}{x_i^{\max} - x_i^{\min}} \tag{8}$$

where: x_i is the actual value of the i influencing factor and $x_i^{'}$ is the normalized value of x_i . $x_i^{\max}, x_i^{\min}$ are the maximum and minimum values of x_i .

Vectorization of health status level: If the health status of equipment is classified by three levels: health, attention, and deterioration (fault), the vectorization of health status level is expressed as $v_{01} = (1,0,0)$, $v_{02} = (0,1,0)$, $v_{03} = (0,0,1)$.

Single-factor health status level: For the abnormal probability distribution characteristics, the smaller the probability value is, the more the health status level tends to be "healthy". Therefore, the membership distribution function is used to quantify the influence of single factor on health status. We calculate the membership vector v_i (for example $v_1 = (\mu_{11}, \mu_{12}, \mu_{13})$) of single influencing factor by selecting the appropriate health membership distribution function of abnormal probability and abnormal severity.We determine the health level under single factor according to the principle of maximum membership degree, where the triangular distribution function used to divide the three categories is referred to as Eq. (9).

$$\mu_1 = \begin{cases} 1 & (p = 0) \\ \dfrac{a_1 - p}{a_1} & (0 < p < 0.41) \\ 0 & (0.41 \le p) \end{cases}$$

$$\mu_2 = \begin{cases} 0 & (0 \le p < a_1) \\ \dfrac{p - a_1}{a_2 - a_1} & (a_1 \le p < a_2) \\ \dfrac{a_3 - p}{a_3 - a_2} & (a_2 \le p < a_3) \\ 0 & (a_3 \le p) \end{cases} \tag{9}$$

$$\mu_3 = \begin{cases} 0 & (0 \le p < a_4) \\ \dfrac{p - a_4}{1 - a_4} & (a_4 \le p < 1) \\ 1 (p = 1) \end{cases}$$

2.2.4. Calculate the health correlation

We use the grey relational analysis method to calculate the correlation coefficient of the above health status level vector and the single factor membership vector, as shown in Eq. (10).

$$\theta_{kij}(l) = \frac{\min\limits_i \min\limits_j |v_{0k}(j) - v_i(j)| + \rho \max\limits_i \max\limits_j |v_{0k}(j) - v_i(j)|}{|v_{0k}(l) - v_i(l)| + \rho \max\limits_i \max\limits_j |v_{0k}(j) - v_i(j)|} \tag{10}$$

where: $\theta_{kij}(l)$ is the correlation coefficient between comparison sequence v_i and reference sequence v_{0k} on the $j-th$ index; $v_i(l)$ is the $i-th$ index of membership vector under the influence of the $j-th$ factor. v_{0k} is the $k-\text{th}$ health status level vector; ρ is the resolution coefficient in the range [0,1].

2.2.5. Evaluate the health status level

Fuzzy comprehensive evaluation is a comprehensive evaluation method based on fuzzy mathematics. Its advantage is that it can quantitatively evaluate some factors that have unclear boundaries and are not easy to quantify. The main steps are as follows:

1. Determine the set of factors for the evaluation object;
2. Determine the comment set of the evaluation object;
3. Determine the weight vector of the evaluation factors A ;
4. Determine the fuzzy comprehensive judgment matrix R ;
5. Fuzzy comprehensive evaluation, as shown in Eq. (11).

$$B = A \cdot R \tag{11}$$

3. RESULTS AND DISCUSSION

In the maintenance and support work of a single cylinder bolt boom system of a large tonnage trunk crane, there are problems such as various exception modes, difficult to determine the degree of failure impact, high failure rate, and the system state is affected by the operation intensity, so it is difficult to determine the health state of the equipment. In this paper, we select 8-week data of a single vehicle to perform anomaly identification, probability calculation, and then FMECA analysis on fault samples for this system. We combined the maintenance records and client data to determine the FMECA analysis results and establish the health evaluation model of the large arm. Finally, the model was extended to 63 truck cranes for health evaluation. We deployed a real-time system to evaluate the real equipment health status level, and verified the false positive rate and false negative rate of the model. Figure 3 is the schematic diagram of the crane single cylinder latch telesopic arm.

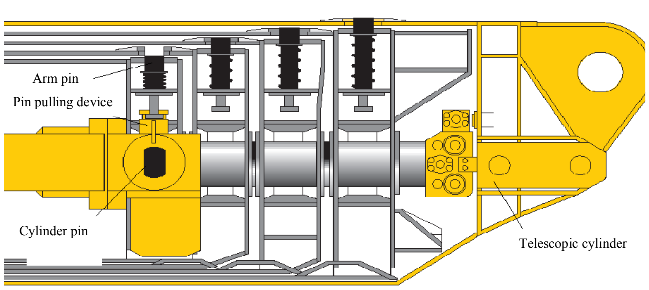


Figure 3. The schematic diagram of the crane single cylinder latch telescopic arm

3.1. Data Collection

At present, the mainstream large tonnage truck crane has installed T-Box for data acquisition and application of CAN protocol communication. Of course, the crane fleet analyzed in this paper has transferred the data collection to the cloud Apache Hive database storage through 4G. We applied the data collected by the existing sensors of the crane equipment for calculation, avoiding the addition of new sensors. The main selected fields are shown in Table 1, and the signal acquisition period is 1s.

The first arm extension process and the first arm contraction process are shown in Figure 4. It can be seen from the figure that each arm is extended and retracted in sequence. At the same time, the cylinder pin and the arm pin are required to cooperate in the process.

Serial number	Signal
1	Second arm percentage(%)
2	Third arm percentage(%)
3	Forth arm percentage(%)
4	Fifth arm percentage(%)
5	Sixth arm percentage(%)
6	Seventh arm percentage(%)
7	Telescopic cylinder Length(mm)
8	Right cylinder pin unlock(0,1)
9	Right cylinder pin lock(0,1)

Table 1. Description of vehicle crane parameters

3.2. Abnormal Recognition

Before anomaly identification, we define the exception modes as Table 2. Mode A uses JS divergence to calculate the distance between each speed feature and the benchmark feature (normal speed can be screened as the benchmark) to determine the criterion. Mode B determines the criterion using temporal anomaly detection. C, D, and E use logical rules to form criteria.

Exception Modes	Definition
Mode-A	The similarity between the velocity feature and the baseline feature is low
Mode-B	Abnormal speed data of arm extension and arm contraction
Mode-C	The unlocking signal of cylinder pin and arm pin is equal to the locking signal
Mode-D	Long duration of cylinder pin lock and arm pin unlock
Mode-E	Signal is lost in the arm

Table 2. Defining exception modes

The principle of applying JS divergence to identify large arm anomalies is: For the abnormal large-arm system, it is often reflected in the arm extension speed. Therefore, we extract the velocity features and construct the velocity distribution of the normal sample telescopic arm process as the benchmark model.

Figure 5 is the arm extension velocity and arm contraction velocity over time. The process of arm extension and arm contraction can be divided into three processes: acceleration phase, constant speed phase and deceleration phase.

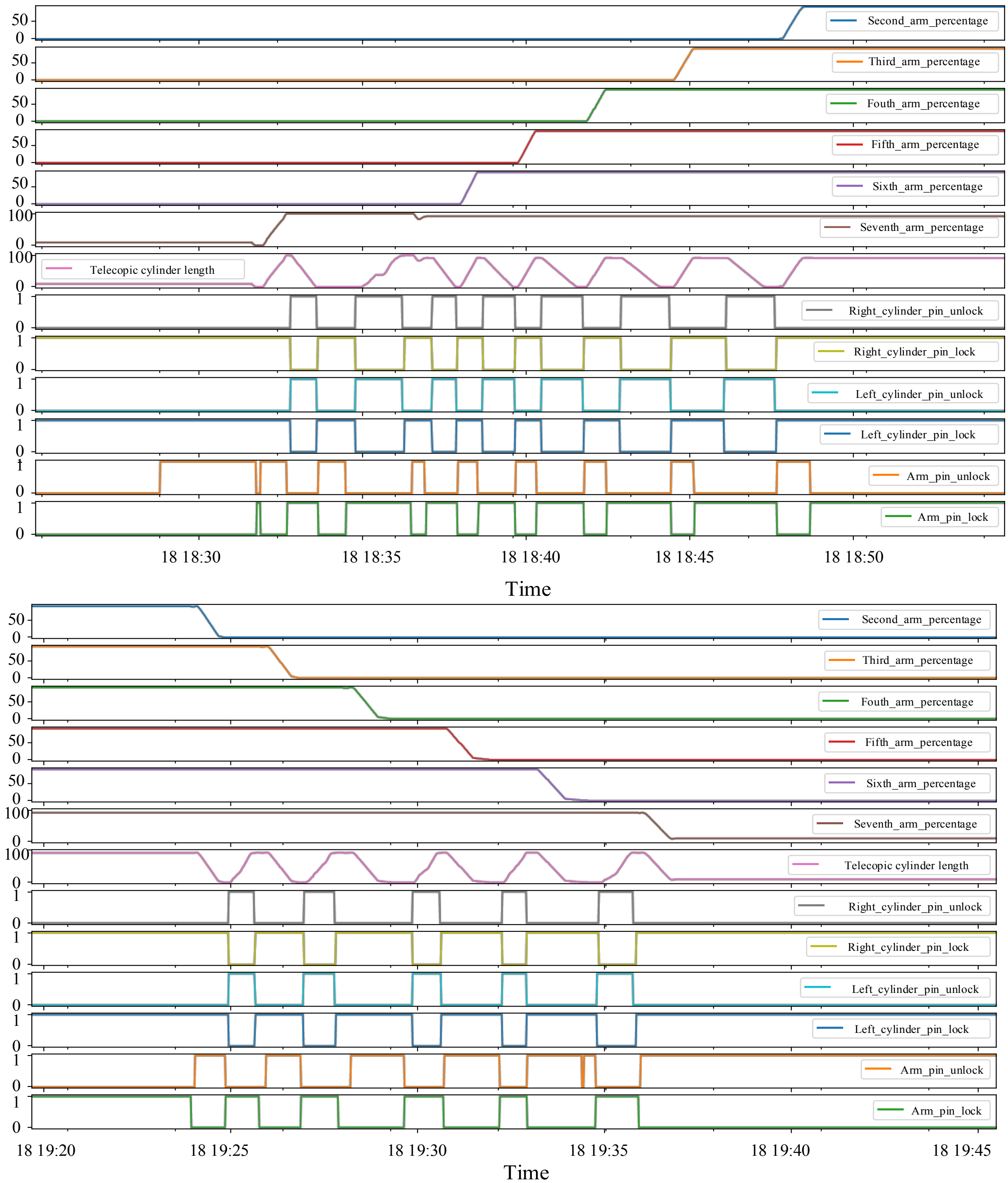


Figure 4. The arm extension process and the arm contraction process

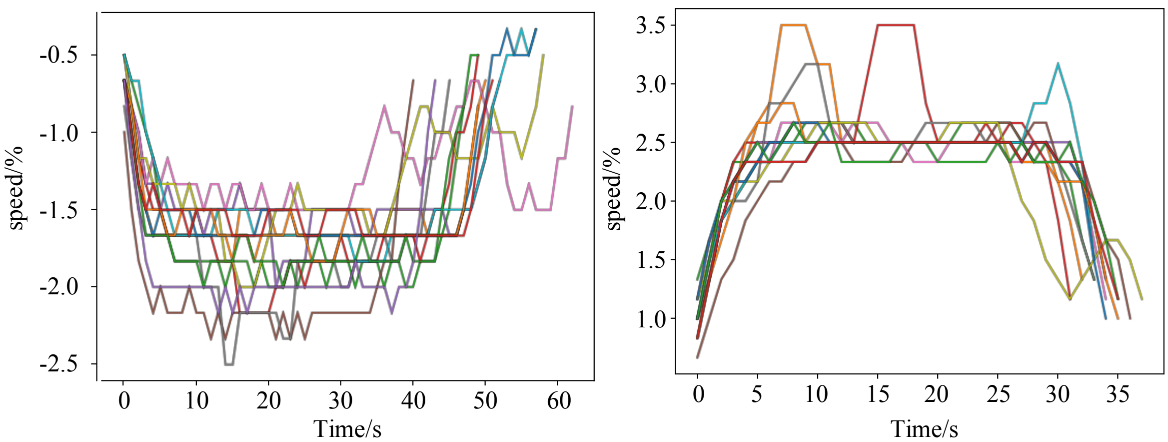


Figure 5. The arm extension velocity and arm contraction velocity over time

Taking the fourth arm as an example, the median velocity of the three stages during each action is extracted as the sample. Among the known eight weekly samples, two weeks week_3 and week_4 are fault samples. Figure 6 is the box plot, it can be initially seen that the speed distribution range of the fault samples is wide.

According to the Gaussian kernel density estimation method, we calculated the velocity distribution function of each sample at three stages, and its distribution is shown in Figure 7. Comparing the benchmark with normal samples and fault samples respectively, it can be seen that there are clear differences between the distribution.

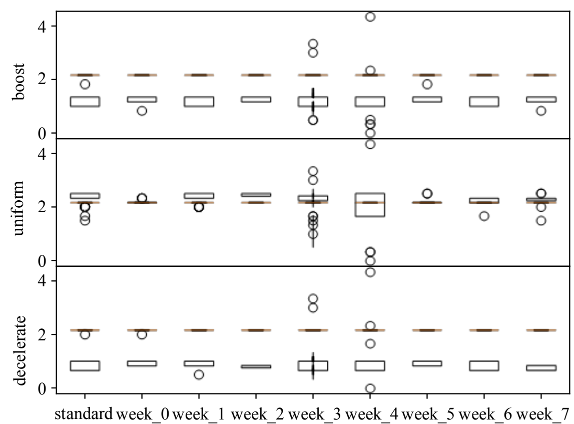
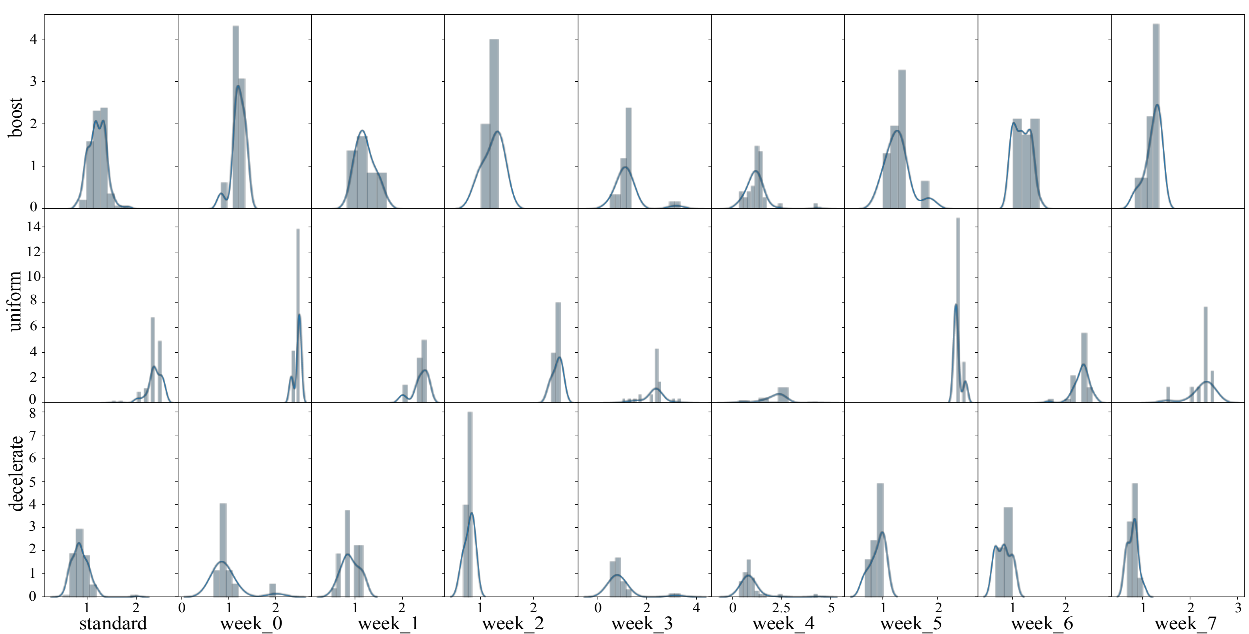


Figure 6. Box patterns of different samples



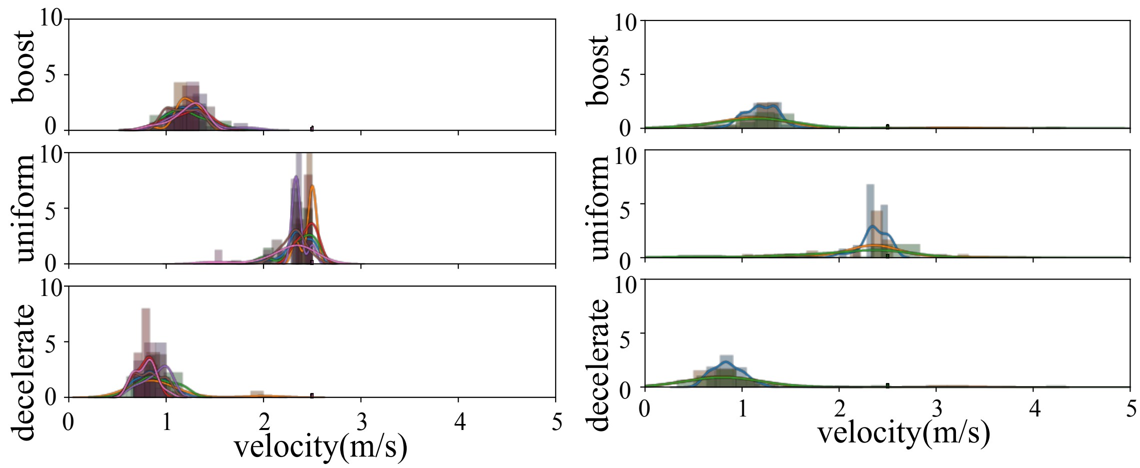


Figure 7. Velocity distribution of each sample in three phases

After obtaining the distribution of each sample data, we calculate the JS divergence between the sample data and the benchmark distribution according to Eq. (5). The JS divergence of the three stages of each sample is averaged to index this metric so that the numerical range is [0,1], as shown in Figure 8. We take the average of the three indexation indexes as the anomaly evaluation index, and determine whether it is abnormal according to the set threshold. It can be seen in Figure 9 that there is a significant decrease at week_3 and week_4.

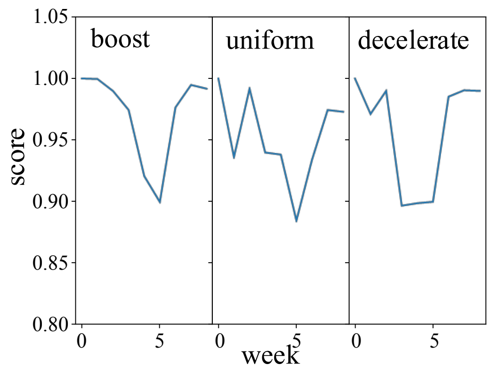


Figure 8. The JS divergence of the three stages

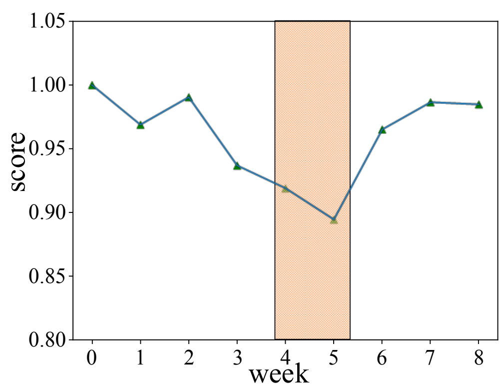


Figure 9. the average of the three indexation indexes

We calculate the abnormal probability of different joint arm, extension arm and contraction arm processes and obtain the probability of exception mode A. According to the exponential smoothing method for time series data anomaly detection and the mechanism rule to determine the anomaly method, we calculate the anomaly probability of mode B, mode C, mode D, mode E respectively. The abnormal severity is determined according to the degree of influence on the equipment after the occurrence of exception modes. Finally, the FMECA analysis result table is formed as shown in Table 3.

Serial number	Exception nodes	Abnormal probability	Abnormal severity
1	Mode-A	0.0296	0.7
2	Mode-B	0	0.8
3	Mode-C	2.8521	0.8
4	Mode-D	1.6154	0.4
5	Mode-E	0	0.7

Table 3. FMECA analysis result

3.3. Single-factor health status rating

In this paper, the membership distribution function is used to quantify the influence of a single factor on health status. We used a triangular distribution to calculate the health state membership values for the parameter abnormal probability factor and the abnormal severity factor.

In Eq.(9), the coefficient a_i needs to be adjusted according to the effect of the distribution function on the accuracy of the final result. In this paper, genetic algorithm(GA) is used to find the optimization coefficient, and its objective function should make the model output consistent with the sample label. GA is an adaptive and global optimizing probability search method. Detailed steps are performed as in Table 4.

Input: FMECA analysis result;Health and failure samples
Output: Coefficients of the probability and severity distribution function

$$Sample = \mathrm{n} \times (data, label)$$
$$a_{11} \sim a_{14} \in [0,1]; a_{21} \sim a_{24} \in [0,1] \ \#\text{parameter space}$$
$$v_{A1} = (\mu_{A11}, \mu_{A12}, \mu_{A13}) \cdots v_{E1} = (\mu_{E11}, \mu_{E12}, \mu_{E13})$$
$$v_{A2} = (\mu_{A21}, \mu_{A22}, \mu_{A23}) \cdots v_{E2} = (\mu_{E21}, \mu_{E22}, \mu_{E23})$$
$$v_{A1} \sim v_{E1}; v_{A2} \sim v_{E2} \xrightarrow{substituting} \xi_{kij} \to r_{kij} \to r_{kij}^{'}$$
$$A_A = [r_{kA1}^{'}, r_{kA2}^{'}] \cdots A_E = [r_{kE1}^{'}, r_{kE2}^{'}]$$
$$R_A = \begin{bmatrix} v_{A1} \\ v_{A2} \end{bmatrix} \cdots R_E = \begin{bmatrix} v_{E1} \\ v_{E2} \end{bmatrix}$$
$$B_A = A_A \bullet R_A \cdots B_E = A_E \bullet R_E$$
$$\textit{if } index(\max(B_i)) = 0 : \mathrm{mode}(i) \text{ is 'health'\#model\_output}$$
$$\textit{if } \text{Sample\_label} = \text{model\_output} : num(correct) + 1$$
$$G = \max(num(correct)) \#\text{GA objective function}$$

Table 4. Pseudo code for coefficient optimization of distribution function

First, n data samples with fault or health labels are prepared. Then, a_i coefficient and parameter space are initialized, and B is constructed according to formulas (8) ~ (11). The health state is obtained by calculating the index of the largest element in the vector B . Determine whether the prediction is accurate according to the comparison of the sample label and the model prediction result label. If the judgment is correct, the number of correct predictions is added to one. Finally,taking the exceptional severity as an example, we applied the GA to optimize in the parameter space of a_i to obtain $a_1 = 0.41, a_2 = 0.45, a_3 = 0.65$, and the distribution function is:

$$\mu_1 = \begin{cases} 1 & (p = 0) \\ \dfrac{0.41 - p}{0.41} & (0 < p < 0.41) \\ 0 & (0.41 \le p) \end{cases}$$

$$\mu_2 = \begin{cases} 0 & (0 \le p < 0.41) \\ \dfrac{p - 0.41}{0.04} & (0.41 \le p < 0.45) \\ \dfrac{0.65 - p}{0.2} & (0.45 \le p < 0.65) \\ 0 & (0.65 \le p) \end{cases}$$

$$\mu_3 = \begin{cases} 0 & (0 \le p < 0.6) \\ \dfrac{p - 0.6}{0.4} & (0.6 \le p < 1) \\ 1 & (p = 1) \end{cases}$$

The distribution function calculated according to the method in Table 4. Table 1 was brought into the distribution function to calculate the membership vector of abnormal harshness single influence factor v_i .

$$v_{A2} = (0, 0, 0.25)$$
$$v_{B2} = (0, 0, 0.5)$$
$$v_{C2} = (0, 0.75, 0)$$
$$v_{D2} = (0, 0.25, 0)$$
$$v_{E2} = (0, 0, 0.25)$$

Similarly, for the abnormal probability of different exception modes, we choose the appropriate membership distribution function, in which the parameters of the distribution function are optimized by the labeling of labeled samples. In Figure 10, the membership vector is calculated as v_i :

$$v_{A1} = (0.15, 0.85, 0.03)$$
$$v_{B1} = (1, 0, 0)$$
$$v_{C1} = (0, 0.94, 0)$$
$$v_{D1} = (0.35, 0.30, 0)$$
$$v_{E1} = (1, 0, 0)$$

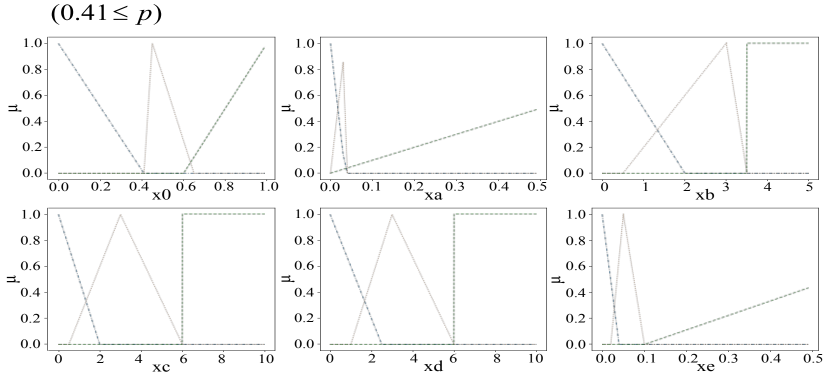


Figure 10. Membership distribution map

3.4. Assess the degree of health

We use the grey relational analysis method to calculate the health correlation degree, and calculate the correlation coefficient of the membership vectors v_{A1} - v_{E1} and v_{A2} - v_{E2} in turn. Finally, the probability weight and severity weight are obtained as follows:

	A	B	C	D	E
Probability weight	0.6587	0.6928	0.7037	0.6884	0.7037
Severity weight	0.7111	0.6852	0.6815	0.7111	0.7111

Table 5. Single factor weight

We use the grey relational analysis method to calculate the health correlation degree, and calculate the correlation coefficient of the membership vectors v_{A1} - v_{E1} and v_{A2} - v_{E2} in turn. Finally, the probability weight and severity weight are obtained as follows:

$$B_A = (0.1016, 0.5571, 0.1972)$$
$$B_B = (0.7037, 0.0, 0.3426)$$
$$B_C = (0.0, 0.6539, 0.3426)$$
$$B_D = (0.2388, 0.1914, 0.0)$$
$$B_E = (0.7037, 0.0, 0.1778)$$

In this way, it can be judged that the health status grades of the five states of the system under the influence of the comprehensive factors of abnormal probability and abnormal severity are 'attention', 'health', 'attention', 'health' and 'health', and the overall evaluation result is 'attention'.

3.5. Model results

We apply the method to batch equipment, covering 63 car cranes. We deploy this method in the big data environment, using Hive as the database, Spark as the batch calculation engine, BI as the presentation tool. We take the weekly data of a single vehicle as a sample, select 63 equipment for 7 months comprehensive evaluation, and form 2016 samples. Combined with the verification of fault maintenance records and the confirmation of service engineers, we calculate the false positive rate and false negative rate of the samples. The results show that the false negative rate is 4.76%, the false positive rate is 5.45%, so the algorithm has a good effect.

We compared this model with the model established by the "GMM + NSET" method (Yuan, 2021) for the false negative rate and the false positive rate. "GMM+NSET" is an equipment parameter early warning method based on Gaussian mixture model (GMM) and nonlinear state estimation (NSET) . In this method, it is necessary to set the threshold value of the output value of each model. If the deviation of the output value of the model and the input value exceeds the threshold value, the alarm occurs. We used the genetic algorithm to find the optimal threshold of each Signal, with the false positive rate plus the sum of the missing positive rate as the target value, and the results are shown in Table 6.

Signal number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Value(%)	6.2	5.2	6.3	5.4	4.9	5.5	7.6	9.5	9.6

Table 6. Average error of prediction value of each signal based on GMM + NSET method

The model constructed by the GMM + NSET method method predicts the samples and compares the results with the results of this method, as shown in Table 7.

	negative rate	positive rate
this method	**4.76%**	**5.45%**
GMM+NSET	15.62%	19.83%

Table 7. Comparison of the results between the two methods

4. CONCLUSION

The advantage of this method is that it evaluates the comprehensive health degree of equipment with multiple abnormal modes, which depends on the monitoring of various abnormal modes. Therefore, this method is difficult to adapt to the states of structural parts such as fatigue, wear and deformation that are difficult to be monitored by sensors.

The research of this method solves the problem that the health status of crane equipment can not be evaluated, the problem of "whether the equipment needs maintenance", and the problem of "what is the fault", which can help to deal with the deterioration of equipment in advance, avoid the cost of accidents, and guide the transition from post-maintenance to predictive maintenance.

More accurate equipment health status evaluation depends on the method of combining mechanism with data. We propose an equipment health status evaluation technology under the influence of multiple exception modes, which extracts multi-dimensional features to represent equipment exception modes and harshness, and uses grey relational analysis and fuzzy comprehensive evaluation to achieve objective and accurate equipment health status evaluation. This method solves the limitation problem of single anomaly and the difficulty of evaluating multiple exception modes. The actual data of the large tonnage crane equipment system is analyzed by this method. We contrasted this model with the "GMM + NSET" model. The results show that this technology has good universality and low false positive rate and false negative rate.

REFERENCES

Bencheikh, F.Harkat, M. F.Kouadri, A.Bensmail, A. (2020). New reduced kernel pca for fault detection and diagnosis in cement rotary kiln. *Chemometrics and Intelligent Laboratory Systems*, 204(1).DOI:10.1016/j.chemolab.2020.104091

Bhat D, Muench S, Roellig M(2023).Application of machine learning algorithms in prognostics and health monitoring of electronic systems: A review. *e-Prime - Advances in Electrical Engineering, Electronics and Energy*. Vol 4,2023,100166,https://doi.org/10.1016/j.prime.2023.100166.

Goncalves, J. P. S. , Fruett, F. , Dalfre Filho, J. G. , & Giesbrecht, M. . (2021). Faults detection and classification in a centrifugal pump from vibration data using markov parameters. *Mechanical Systems & SignalProcessing*(158),DOI:158.10.1016/j.ymssp.2021.107694

Guo S. (2014). Diagnosis of structural damage based on JS divergence. *Harbin Institute of Technology*. DOI：10.7666/d.D752859

Guo S. Diagnosis of structural damage based on JS divergence. *Harbin Institute of Technology*, 2014.

Hamadouche, A. , Kouadri, A. , & Bensmail, A. . (2018). Kernelized relative entropy for direct fault detection in industrial rotary kilns. *International Journal of Adaptive Control and Signal Processing*, 32(7).DOI:10.1002/acs.2879

Hu C H, Shi Q, & Si X S, et al. (2017). Progress in data-driven life span prediction and health management technologies. *Progress in data-driven life span prediction and health management technologies*, 46(1), 11.doi: 10.13976/j.cnki.xk.2017.0072

Pan F, Shen J X, Gao Y M, & Yang X Q. Research on engineering equipment fault prediction and health management technology.*Mechanical Management Development*,2021,36(05):98-100.DOI:10.16525/j.cnki.cn14-1134/th.2021.05.042.

Priyom G, Rajiv N R(2023).A systematic review on failure modes and proposed methodology to artificially seed faults for promoting PHM studies in laboratory environment for an industrial gearbox. *Engineering Failure Analysis*, Vol 146, 2023,107076,https://doi.org/10.1016/j.engfailanal.2023.107076.

Qin P. Functional reliability and health assessment of an explosive logic network system based on failure analysis. (Doctoral dissertation, University of Electronic Science and Technology of China). DOI：CNKI:CDMD:2.1016.172252

Tan H Z.(2021). Application of fault prediction and health management technology in the civilian field. Mechanical manufacturing, 59(6), 6.

Wang, W. , Lai, Y. C. , & Grebogi, C. . (2016). Data based identification and prediction of nonlinear and complex dynamical systems. *Physics Reports*, 644, 1-76. DOI：10.1016/j.physrep.2016.06.004

Yu Y, Si X S, Hu C H, Cui M Z, & Li H P. (2018). Progress in data-driven reliability assessment and life span prediction: Covariate-based methods. *Journal of Automation*, 44(2), 12. DOI：10.16383/j.aas.2018.c170005

Yuan Xuefeng, Chen Mubin, Ma Chenglong, & Zhou Yanwei.(2021).Study on Multi-Sensor Parameter Warning Based on Gaussian Mixed Model and Nonlinear State Estimation Technique. *2021 IEEE Sustainable Power and Energy Conference.* DOI:10.1109/iSPEC53008.2021.9735467

Zhang Xujin, Zhang Yun, Research on fault extraction of rotary kiln barrel based on wavelet packet decomposition, Mech. Eng. 09 (2021) 64－67. DOI：10.17583/ as0045

Zhang Y, Han G W, Lu N Y, Jiang B, & Zhi Y R. (2017). Health status assessment method of rail vehicle door system based on JS divergence. *Mechanical Design and Manufacturing Engineering*, 46(11), 6.

Zhao K. Study on failure mode and risk assessment of boiler bearing parts. (Doctoral dissertation, Shandong University).DOI：CNKI:CDMD:2.1015.531457

Zhong Xianping, Zhang Lin, Ban Heng. (2023).Deep reinforcement learning for class imbalance fault diagnosis of equipment in nuclear power plants, *Annals of Nuclear Energy*,184.DOI:10.1016/j.anucene.2023.109685

Zou X Y , Tao L F, Sun L L, et al.(2023).A case-learning-based paradigm for quantitative recommendation of fault diagnosis algorithms: A case study of gearbox.*Reliability Engineering & System Safety*.Vol 237,2023,109372,https://doi.org/10.1016/j.ress.2023.109372.