International Journal of Mechanical and Electrical Engineering
ISSN: 3005-9615 (Print), ISSN: 3005-7132 (Online) | Volume 7, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijmee.v7n3.12
Journal homepage: https://wepub.org/index.php/IJMEE/index

Research on Multi-dimensional Health Evaluation Method for Complex Machinery

-- Taking Fracturing Pump as an Example

Yunlu Zhang, Chuan Wang and Huawei Wang

School of Southwest Petroleum University, Chengdu Sichuan, 610000, China

ABSTRACT

This paper focuses on the multi-dimensional health state assessment of complex mechanical equipment, using a fracturing pump as a case study. It proposes a comprehensive assessment framework that integrates the analysis of multiple data dimensions, including real-time operational data, vibration data, and historical maintenance data. A health baseline is constructed by obtaining residual data of the equipment in a healthy state using a Generalized Regression Neural Network (GRNN) and extracting time-frequency domain features. The health assessment model for the vibration data dimension is built by calculating and normalizing the Mahalanobis Distance between the vibration data to be assessed and the health baseline. An equipment operational state evaluation model based on multiple characteristic parameters is established, using these parameters as failure criteria, to quantitatively characterize the Health Index from the real-time operational data dimension. Based on reliability theory, a multi-state transition probability for the main components of the equipment is obtained using a Markov model. The probability is mapped to the Health Index using a weighted average method for defuzzification, thereby constructing a health assessment model for the historical maintenance data dimension. Finally, the health state of the equipment is graded and comprehensively evaluated. This method can provide theoretical support for the preventive maintenance and health management of equipment, while also offering a reference for the intelligent operation and maintenance of complex mechanical equipment.

KEYWORDS

Fracturing Pump; Multi-Dimensional Health State Assessment; Mahalanobis Distance; Generalized Regression Neural Network (GRNN); Markov Model Introduction.

1. INTRODUCTION

Health management technology involves using techniques such as measurement, control, and data analysis to perform fault diagnosis and prediction in the early stages of failure, preventing mechanical equipment from failing during operation. With the increasing demands for longer lifespan and higher reliability of complex machinery, in-depth research on health management technology for complex mechanical equipment has become more urgent. [1-2]

There is considerable research on the health state assessment of single equipment or systems, but most focus on a certain key characteristic parameter to evaluate the equipment's health state, lacking a comprehensive and holistic health assessment and future state prediction scheme. For example, Literature introduced the concepts of membership degree and grey correlation, applied fuzzy hierarchical relationships to describe the relationship between evaluation factors and equipment state, and established a comprehensive evaluation model for equipment operational state. Analysis through a high-voltage circuit breaker case demonstrated the model's feasibility, but further research is needed on more reasonable calculation of weight coefficients. Literature took a single equipment system as the research object, starting from the equipment's operational state, improved the training algorithm of the Hidden Markov Model and introduced an aging factor to construct the equipment degradation process and proposed a method for equipment state diagnosis research. However, the determination of the optimal form of the aging factor and the prediction of equipment state values were not explained. Literature, based on attribute data of distribution transformer operation status and information entropy theory, combined with indicator thresholds and weight values of multiple attributes, proposed a big data analysis and evaluation method for distribution transformer operation status, which can effectively reflect the overall operational state of distribution transformers and correspondingly improve fault prediction levels.

The health state assessment of complex mechanical equipment is a key link in ensuring reliable equipment operation, reducing maintenance costs, and improving production efficiency. As a core equipment in shale gas extraction, the health state of a fracturing pump directly affects operational efficiency and safety. However, traditional single-dimensional assessment methods struggle to comprehensively reflect the complex state of the equipment. Therefore, this paper proposes a multi-dimensional health state assessment framework that comprehensively considers real-time monitoring data, historical data, and maintenance data, establishing corresponding mathematical models to achieve a comprehensive assessment of the health state of complex mechanical systems.

2. METHOD CONSTRUCTION

The multi-dimensional health state assessment framework proposed in this paper covers the following three core dimensions:

1.Real-time Operational Data Dimension: Based on real-time data such as temperature, pressure, and flow rate collected during equipment operation, a health index model is constructed by establishing an operational state evaluation of the equipment's multiple characteristic parameters.

2.Historical Maintenance Data Dimension: A Markov model is used to model historical maintenance data to predict the future state of the equipment. The probability is mapped to the health index using a weighted average method for defuzzification.

3.Vibration Data Dimension: Residual data of the equipment in a healthy state is obtained via a Generalized Regression Neural Network (GRNN). Time-frequency domain features are extracted to construct a health baseline. The health assessment model for the vibration data dimension is built by calculating and normalizing the Mahalanobis Distance between the vibration data to be assessed and the health baseline.

4.Comprehensive Health Evaluation: The CRITIC weight method is used to assign weights to each dimension, obtaining the equipment's comprehensive health index. The equipment's health state is then graded for a comprehensive evaluation.

Each dimension employs different health assessment methods to ensure the comprehensiveness and accuracy of the evaluation results. [3]

2.1. High-Frequency Measured Data Assessment Method

The process for the health assessment method based on high-frequency measured data is shown in Figure 1. Data preprocessing is first performed. A health baseline construction method based on an observer model is adopted, establishing a fault observer based on GRNN to obtain the system's residual data in a healthy state. Time-domain features are then extracted from the residual data. The Mahalanobis space is constructed using various feature points from the healthy state, serving as the health baseline for the mechanical system. During assessment, the observer is used to obtain the residual of the data to be assessed, features are extracted to form a feature sequence, and then distance measurement is conducted. The Mahalanobis distance between the current state and the health baseline is calculated and normalized into a health degree, which is the performance assessment result for the vibration data dimension of the mechanical system.

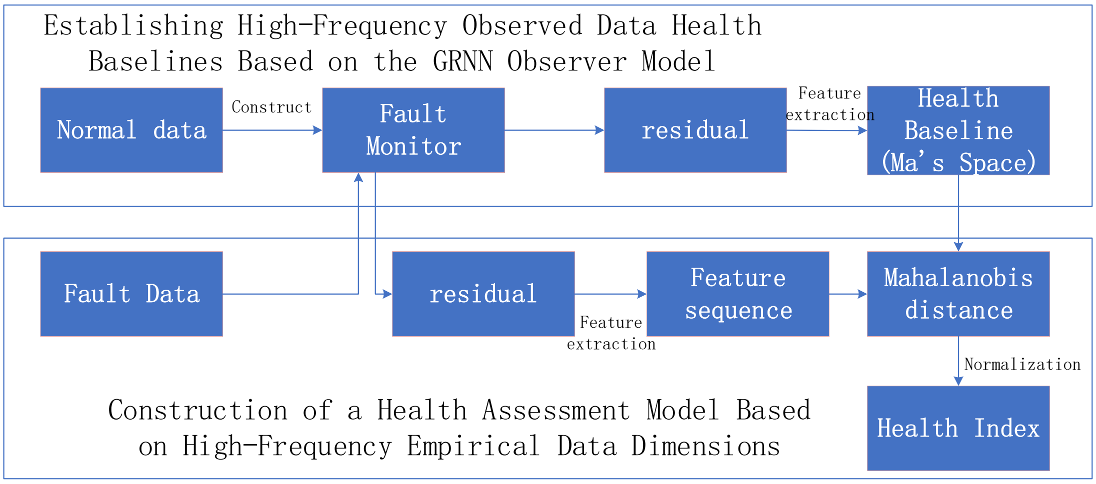

Fig 1. Flowchart for High-Frequency Actual Measurement Data Health Assessment Method

2.1.1. Data Preprocessing

1.Deletion of Abnormal Data Collected by Sensors: In daily data recorded by sensors of complex mechanical equipment, some data measured multiple times a day leads to duplicate records. This affects subsequent diagnostic research, so repeatedly measured data needs to be deleted.

2.Data Interpolation: For data of complex mechanical equipment, missing data sometimes occurs. Interpolation of the equipment dataset is an effective method to handle missing data. This paper uses linear interpolation to add missing data. Linear interpolation is a common numerical analysis method that estimates missing data values based on the straight line between two adjacent data points. It is simpler and more convenient compared to other interpolation methods. Given known points, to interpolate at x and find the value of y, the mathematical expression is:

$$\mathbf{y} = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 \tag{1}$$

3.Wavelet Threshold Denoising: Wavelet Domain Denoising (WDD) uses Wavelet Decomposition (WD) to decompose a noisy signal into several wavelet coefficients. Threshold processing is then used to convert the wavelet coefficients into estimated wavelet coefficients. Finally, reconstruction is performed to obtain the denoised signal.

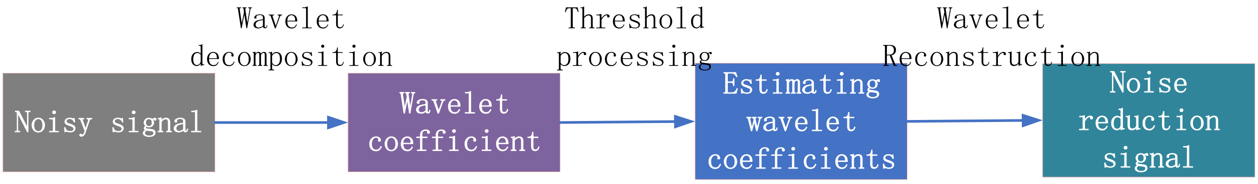

Fig 2. Wavelet Threshold Denoising Process

1) Wavelet Decomposition: WDD is essentially a band-pass filter. When processing a time-domain signal S(t), the low-pass filter in WDD filters out the low-frequency information of the signal, and then the high-pass filter filters out the high-frequency information. When using WD, each layer has corresponding approximation coefficients and detail coefficients. Figure 3 shows the structure diagram of a three-layer wavelet decomposition.

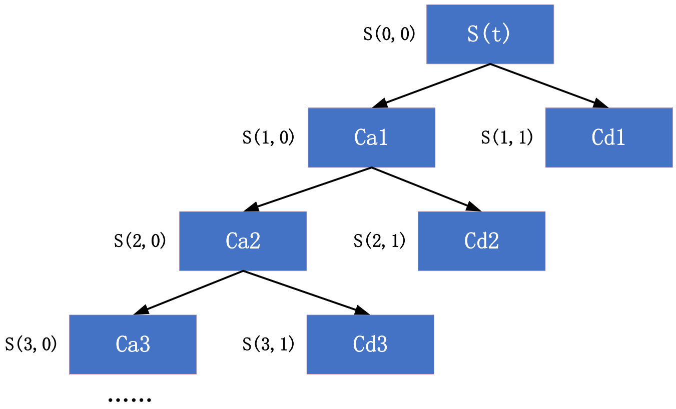

Fig 3. Three-Level Wavelet Decomposition Structure Diagram

n-layer wavelet decomposition:

$$S(t) = Ca_n + \sum_{i=1}^{n} Cd_i \tag{2}$$

2) Threshold Selection: The threshold is a judgment value for processing wavelet coefficients. Depending on the relationship between the wavelet coefficient and the threshold, different treatments can be applied to the wavelet coefficients. The selection of the threshold affects the denoising effect. The Fixed Threshold (Fthr), as the most commonly used global threshold for processing wavelet coefficients, is expressed as:

$$F_{thr} = \delta\sqrt{2\ln N} \tag{3}$$

$$\delta = \frac{median(|Cd_i|)}{0.6745} \tag{4}$$

Where: δis the noise standard deviation; N is the sample size; i is the wavelet decomposition layer number; meadian(|Cdi|) is the median of the amplitudes of all wavelet coefficients.

Although the fixed threshold is effective and convenient, the noise content in each layer of wavelet coefficients varies after multi-layer decomposition. Using a fixed threshold method may not only remove the noise part of the wavelet coefficients but also remove some non-noise information. Therefore, this paper improves it. The new Local Adaptive Threshold (LAthr) can compensate for the shortcomings of the fixed threshold as a judgment value. It posits that each layer of wavelet decomposition has a threshold most suitable for processing the wavelet coefficients. As the decomposition layer increases, the corresponding threshold decreases, thus avoiding the loss of effective information caused by using a fixed threshold. LAthr can be expressed as:

$$LA_{thr_i} = \frac{\delta\sqrt{2\ln N}}{\ln\sqrt[3]{i}+1} \tag{5}$$

3) Selecting the Threshold Function: Commonly used wavelet threshold functions include the hard threshold, soft threshold, and semi-soft threshold functions.

Hard threshold function:

$$Cd_i' = \begin{cases} |Cd_i| & |Cd_i| \geq thr \\ 0 & |Cd_i| < thr \end{cases} \tag{6}$$

Soft threshold function:

$$Cd_i' = \begin{cases} \operatorname{sgn}(Cd_i)(|Cd_i| - thr) & |Cd_i| \ge thr \\ 0 & |Cd_i| < thr \end{cases} \tag{7}$$

Semi-soft threshold function:

$$Cd_i' = \begin{cases} |Cd_i| & |Cd_i| \ge thr_2 \\ \operatorname{sgn}(Cd_i)\dfrac{thr_2(|Cd_i| - thr_1)}{thr_2 - thr_1} & thr_2 < |Cd_i| < thr_1 \\ 0 & |Cd_i| \le thr1 \end{cases} \tag{8}$$

Where: thr is the threshold value; thr1 is the upper threshold value; thr2 is the lower threshold value.

After processing by the hard threshold function, the wavelet coefficients retain the signal characteristics, but the continuity is poor, and the signal is prone to fluctuation during reconstruction, resulting in poor smoothness of the denoised signal. After soft threshold processing, the continuity is better, but when the wavelet coefficient is greater than the threshold, there is a certain error between and, causing deviation in the signal reconstruction process. The semi-soft threshold function combines the advantages of hard and soft thresholds, not only retaining the characteristics of the signal itself but also having good continuity. However, since it needs to determine upper and lower threshold values, it greatly increases the calculation process, making it less practical. To address this issue, the semi-soft threshold function is improved. The function expression is as follows:

$$Cd_i' = \begin{cases} \operatorname{sgn}(Cd_i)(|Cd_i| - thr + \dfrac{2thr}{1 + e^{2Cd_i/thr}}) & |Cd_i| \ge thr \\ 0 & |Cd_i| < thr \end{cases} \tag{9}$$

The improved threshold function uses as an adjustment coefficient to quantize the wavelet coefficients with the threshold, making the estimated wavelet coefficients after threshold processing approach the real wavelet coefficients. This method has been proven to have better denoising effects than hard, soft, and conventional semi-soft threshold functions.

4) Wavelet Reconstruction: The denoised signal is obtained by reconstructing the approximation coefficients obtained from wavelet decomposition and the detail coefficients after threshold processing. Signal-to-Noise Ratio (SNR) or Mean Square Error (MSE) can be used to evaluate the effect of the denoising method. Generally, a larger SNR and a smaller MSE indicate less white noise content in the denoised signal and a better denoising effect.

SNR can be expressed as:

$$SNR = 10\log_{10}(P_{sig} / P_{noise}) \tag{10}$$

MSE is expressed as:

$$MSE = mean[(S(t) - S'(t))^2] \tag{11}$$

Where: Psig is the signal active power; Pnoise is the noise power; S(t)is the ideal signal; S'(t)is the denoised signal.

2.1.2. GRNN Baseline Construction

GRNN has strong nonlinear mapping ability and learning ability, can quickly and accurately describe nonlinear objects, and GRNN training is relatively fast. Therefore, GRNN is used to construct the fault observer for mechanical equipment vibration data. [4-5]

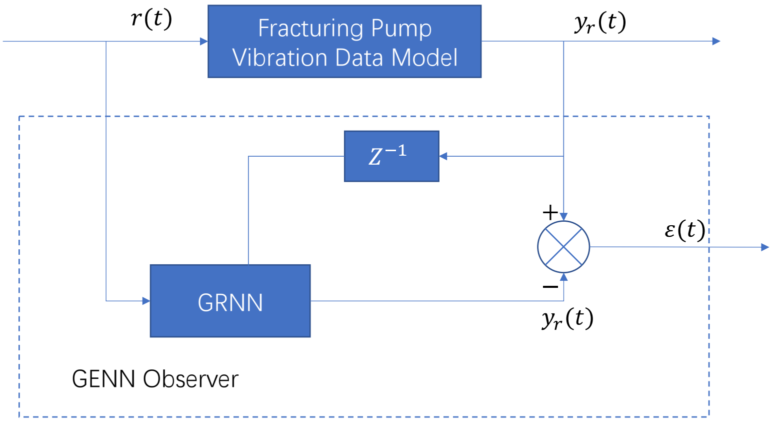

Fig 4. Based on the GRNN Observer Model

The structure of the GRNN-based fault observer is shown in Figure 4. The input signals of the neural network are the original input signal and the actual output signal of the mechanical system. The network's estimated output signal is. represents a delay, meaning the data input to the neural network at this moment is the model's output data from the previous moment. The difference between the mechanical system's actual output and the network's estimated output is defined as the residual, i.e., residual.

$$\varepsilon(t) = y_r(t) - \tilde{y}_r(t) \tag{12}$$

In the training phase, GRNN uses a large amount of vibration data from the healthy state as training data, i.e., using the input signal r(t) and the output signal as the input training samples for the network, and the network's output signal as the estimated output signal. After training, the structure and related parameters of the neural network are determined, i.e., an observer model that can characterize the normal state of the mechanical system is obtained. The output of this model in the normal state, i.e., the residual of the mechanical equipment in a healthy state, can be used to construct the system's health baseline.

To more effectively use the residual signal to establish the system's health baseline, time-domain and frequency-domain features need to be extracted from the residual. The extracted features include peak value, root mean square value, and average absolute value. The peak value is the maximum absolute value of the signal in one cycle, reflecting the performance lower limit of the mechanical equipment; the root mean square value (i.e., the effective value) reflects the energy of the signal and the corresponding power size; the average absolute value is the average value of the area enclosed by the signal relative to the coordinate axis in the time domain, which can be used to characterize the average performance of the mechanical system in one cycle. Assuming the peak value, root mean square value, and average absolute value of a set of signals form a feature point, under a typical working condition, the healthy state of the system is characterized by obtaining N feature points to form a feature sequence. Based on the feature sequence, features with significant degradation trends are screened to construct a Mahalanobis space as the normal reference space, and this Mahalanobis space is used as the system's health basel

2.1.3. Mahalanobis Distance Health Assessment

The health assessment of a complex mechanical system is a standardized measurement of the degree of deviation from the health baseline. Taking the health baseline of the complex mechanical system as the evaluation benchmark, the distance between the feature point of the condition to be assessed and the health baseline is calculated and normalized into a health degree to represent the performance of the mechanical system.

The calculation method for Mahalanobis Distance is as follows:

(1) Calculate the average value of each vector, as follows:

$$\bar{x}_i = \frac{\sum_{j=1}^{n} x_{ij}}{n} \tag{13}$$

(2) Calculate the standard deviation of each vector, as follows:

$$s_i = \sqrt{\frac{\sum_{j=1}^{n} \left(x_{ij} - \bar{x}_i\right)^2}{n-1}} \tag{14}$$

(3) Orthogonalize the feature vectors to obtain Z_{ij}, and find its transpose matrix Z^T, as follows:

$$Z_{ij} = \frac{\left(x_{ij} - \bar{x}_i\right)}{s_i} \tag{15}$$

(4) Calculate the correlation matrix A of the orthogonal matrix, where each element is:

$$a_{ij} = \frac{\sum_{m=1}^{n} Z_{im} Z_{mj}}{n-1} \tag{16}$$

(5) The Mahalanobis Distance is:

$$d_{MD,j} = Z^T A^{-1} Z \tag{17}$$

During health assessment, first use the constructed fault observer to obtain the residual value of the state to be assessed, then extract features to construct a feature point, and finally calculate the Mahalanobis distance between the feature point and the health baseline. A larger Mahalanobis distance indicates that the feature point to be assessed is farther from the health baseline, and the system performance has degraded more severely; a smaller Mahalanobis distance indicates that the feature point to be assessed is closer to the health baseline, and the system performance is relatively good. To describe the system's performance more intuitively, the Mahalanobis distance is normalized into a health degree, as follows:

$$H_1 = 1 - \frac{\arctan(d_{MD,j} + b) - \arctan b}{\pi/2 - \arctan b} \tag{18}$$

Where b is the normalization parameter. By adjusting the value of parameter b, the sensitivity of to different fault stages changes.H_1 is a value between 0 and 1. The larger is H_1, the better the system performance; conversely, the more severe the system performance degradation.

2.2. Low-Frequency Measured Data Assessment Method

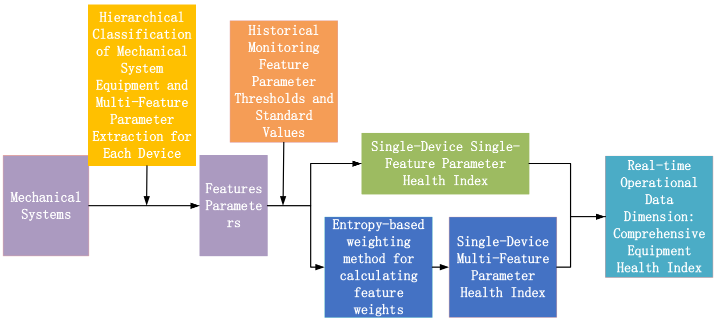

Fig 5. Basic Process for Evaluating Equipment Health Status Based on Low-Frequency Actual Measurement Data Dimensions

The Health Index is defined as the degree of satisfaction that the equipment maintains specific performance during production operation. It is a single value between 0 and 1 obtained by comprehensively analyzing various state information during equipment operation and the operational conditions of the field equipment. A higher value indicates a better operational state of the equipment. 1 represents the best state, and 0 represents complete failure state. The basic process for equipment health state evaluation based on the real-time operational data dimension is shown in Figure 5.

2.2.1. Single Equipment Single Characteristic Parameter Health Index

During the service process of equipment, the standard values of each indicator constantly change over time. Through the equipment's design parameters and certain historical statistical data, a standard time function $\tilde{f}_{ij}(t)$ can be obtained to describe the dynamic standard value of each characteristic parameter, along with the minimum limit $x_{\min}$ and maximum limit $x_{\min}$ for normal operation of this indicator. The specific process is as follows:

(1) Based on design parameters: Design parameters provide initial values, which are usually the standard values under ideal working conditions.

(2) Combined with historical data, analyze the trend of each characteristic parameter changing over time, and fit the standard time function $\tilde{f}_{ij}(t)$ using relevant mathematical models. Possible mathematical models used include:

Linear model: $\tilde{f}_{ij}(t) = a_i t + b_i$, suitable for slowly changing indicators.

Exponential decay model: $\tilde{f}_{ij}(t) = a_i e^{-b_i t} + c_i$, suitable for performance degradation caused by aging.

Periodic model: $\tilde{f}_{ij}(t) = a_i \sin(\omega t + \phi) + b_i$, suitable for periodically changing working conditions.

Polynomial model: $\tilde{f}_{ij}(t) = a_n t^n + \cdots + a_1 t + a_0$, suitable for complex trends.

When the actual measured value exceeds the threshold range, the health index is 0, and the equipment should be shut down immediately for inspection and repair. When the actual measured value coincides with the standard value, the health index is 1, which is the best state of the equipment. When the actual measured value is within the threshold range, by comparing the distance to the standard value, a value between (0,1) can be derived to quantify the equipment state. The specific calculation is as follows:

$$h_{ij}(t) = \begin{cases} \dfrac{f_{ij}(t) - x_{\min}}{\tilde{f}_{ij}(t) - x_{\min}}, f_{ij}(t) \in \left[x_{\min}, \tilde{f}_{ij}(t)\right]; \\ \dfrac{x_{\max} - f_{ij}(t)}{x_{\max} - \tilde{f}_{ij}(t)}, f_{ij}(t) \in \left[\tilde{f}_{ij}(t), x_{\max}\right]; \\ 0, f_{ij}(t) \rangle x_{\min} \text{或} f_{ij}(t) \langle x_{\max} \end{cases} \tag{19}$$

Where, $h_{ij}(t)$ is the health index based on the single characteristic parameter x_j of equipment i; $f_{ij}(t)$ is the measured value of characteristic parameter x_j at time t, $\tilde{f}_{ij}(t)$ is the standard value of this parameter at time t.

2.2.2. Single Equipment Multiple Characteristic Parameters Health Index

During equipment operation, abundant state information is generated, such as noise, stress, strain, temperature, pressure, etc. Due to the complex operating environment of the equipment, it is often difficult to evaluate the entire equipment's operational state using the change of a single state characteristic parameter. This paper uses a comprehensive weighting method to evaluate the health state of equipment based on multiple state characteristic parameters. The health index expression for a single equipment based on multiple characteristic parameters is as follows:

$$h_i(t) = \sum_{j=1}^{m} \rho_{ij} h_{ij}(t) \tag{20}$$

Where, $h_i(t)$ is the health index of equipment i based on multiple characteristic parameters; m is the number of characteristic parameters; ρ_{ij} is the weight of characteristic parameter x_j, indicating the influence of this parameter on the equipment state, $0 \le \rho_{ij} \le 1$, and $\sum_{j=1}^{m} \rho_{ij} = 1$.

Equipment health state evaluation is a typical multi-attribute decision-making problem. Determining the calculation method for the weight of each characteristic parameter of equipment i is the core of this type of problem. In equipment health state evaluation, when certain characteristic parameters deviate severely from normal values, it often indicates that a certain part of the equipment's performance has deteriorated sharply, requiring enhanced monitoring or even shutdown for maintenance. However, in a constant-weight evaluation model, it might be overlooked due to its small weight, and the overall evaluation might still be in the normal range, failing to reflect the true state of the equipment. Therefore, when evaluating the comprehensive health state of system equipment, the balance between various characteristic parameters needs to be considered. According to the entropy weight method proposed in literature, the steps to obtain the weights of each monitoring characteristic parameter are as follows:

Assume that the evaluation indicator values at n time points are selected as samples, and there are m evaluation indicators. represents the evaluation indicator value of the j-th indicator at the i-th time, where i=1,2,3,..., n; j=1,2,3..., m.

I After dimensionless processing of the indicator values, calculate the proportion of the i-th time under the j-th indicator:

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}} \tag{21}$$

II Calculate the entropy value of the j-th indicator:

$$\mathrm{e_j} = -\frac{1}{\ln \mathrm{n}} \sum_{i=1}^{n} \left(p_{ij} \ln p_{ij} \right), 0 \le \mathrm{e_j} \le 1 \tag{22}$$

III Calculate the difference coefficient:

$$h_j = 1 - e_j \tag{23}$$

IV Calculate the entropy weight of the j-th indicator:

$$\rho_{ij} = \frac{h_j}{\sum_{j=1}^{m} h_j} \tag{24}$$

Where, $0 < \rho_{ij} < 1, \sum_{j=1}^{m} \rho_{ij} = 1$

2.3. High-Frequency Measured Data Assessment Method

2.3.1. Determining Prior Probability

Determining the prior probability generally involves two scenarios: Availability of fault data from the machine itself or existing equipment of the same batch. In this case, the failure probability of the component can be directly determined from the fault data. For some components without available fault data for reference. In this case, the prior probability needs to be determined through expert opinion.

The specific method is as follows: The fuzzy interval group decision-making method is used. The prior probability of the root node is comprehensively derived by having multiple experts provide fuzzy intervals. This method better reflects the experts' fuzzy understanding of the prior probability and can also reduce the subjectivity of expert opinions, making the results more consistent with reality. When using the fuzzy interval group decision-making method to find the root node's prior probability, the fuzzy intervals given by multiple experts are comprehensively processed using set-valued statistics to obtain the prior probability value that synthesizes the opinions of multiple experts.

2.3.2. Determining State Transition Probability and Future State Probability

As shown in the figure below, it is the state transition diagram of a multi-state degradation component. It is assumed that each component includes four states: No, DS1, DS2, and Yes. State No refers to a good operating state. State Yes refers to the failure state. States DS1 and DS2 represent the 1st and 2nd degradation states, respectively. The first degradation state indicates the initial degradation of the equipment, i.e., the degradation when the equipment is just put into use. This degradation has a minimal impact on the equipment's reliability and is the running-in stage of the equipment. The second degradation state indicates further degradation of the equipment. At this point, the degradation will cause the equipment's performance to decline but will not directly lead to equipment failure. Each parent node is initially in its perfect functional state (No). Over time, it can enter the first (DS1) or second (DS2) degradation state, or enter the failure state (Yes), as shown in Figure 6.

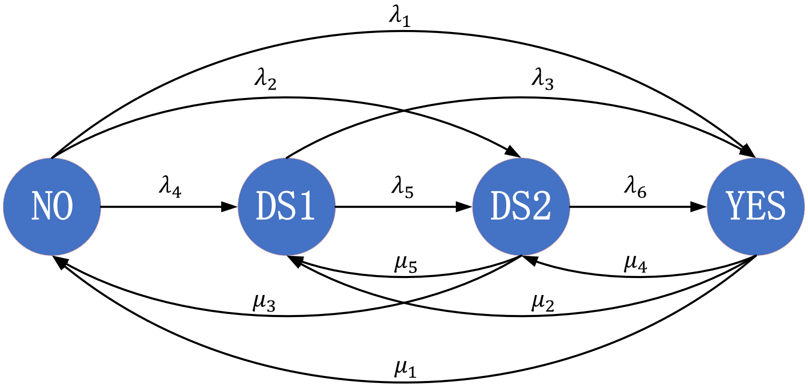

Fig 6. State Transition Diagram for Multi-State Degenerate Components

When a node fails, it needs to be repaired. If the component's performance is restored to the perfect functional state after repair, it is called a complete repair; if after repair, the component's performance can reach both the perfect functional state and the two degradation states, it is considered an incomplete repair. This paper only considers the case of incomplete repair. The failure rate and repair rate are given on the state transition curve. [6]

For simplicity, considering the failure rateλ and repair rate μ for each parent node, the failure rates, λ_1, λ_2, λ_3, λ_4, λ_5, λ_6, and the repair ratesμ_1, μ_2, μ_3, μ_4, μ_5, the following assumed equations can be obtained:

$$\lambda_2 = \lambda_3 \tag{25}$$

$$\lambda_4 = \lambda_5 = \lambda_6 \tag{26}$$

$$\lambda_1 + \lambda_2 + \lambda_4 = \lambda \tag{27}$$

$$\lambda_1 : \lambda_2 : \lambda_4 = 1:6:3 \tag{28}$$

$$\mu_2 = \mu_3 \tag{29}$$

$$\mu_4 = \mu_5 \tag{30}$$

$$\mu_1 + \mu_2 + \mu_4 = \lambda \tag{31}$$

$$\mu_1 : \mu_2 : \mu_4 = 7:2:1 \tag{32}$$

Assume the current time is t, and the time interval between two consecutive trials is Δt. According to equations (25) to (32), the failure rate and repair rate for each state of each node can be calculated. The failure rate and repair rate are calculated as follows:

$$\lambda = \frac{1}{MTBF} \tag{33}$$

$$\mu = \frac{1}{MTTR} \tag{34}$$

Where MTBF is the ratio of the component's cumulative operating time to the number of failures; MTTR is the Mean Time to Repair.

Then, Table 1 gives the transition relationships between consecutive nodes under incomplete repair.

Table 1. Transition Matrix Between Consecutive Nodes

t	t+Δt			
	NO	DS1	DS2	YES
NO	$e^{-\lambda \Delta t}$	$0.3(1-e^{-\lambda \Delta t})$	$0.6(1-e^{-\lambda \Delta t})$	$0.1(1-e^{-\lambda \Delta t})$
DS1	0	$e^{-0.9\lambda \Delta t}$	$(1-e^{-0.9\lambda \Delta t})/3$	$2(1-e^{-0.9\lambda \Delta t})/3$
DS2	0	0	$e^{-0.3\lambda \Delta t}$	$1-e^{-0.3\lambda \Delta t}$
YES	$0.7(1-e^{-\mu \Delta t})$	$0.2(1-e^{-\mu \Delta t})$	$0.1(1-e^{-\mu \Delta t})$	$e^{-\mu \Delta t}$

2.3.3. Health Index Based on Weighted Average Method for Defuzzification

Because the Markov model represents the node state in the form of probability, for the node representing the health state of the evaluation object, the probability values of different health states can be analogous to the membership degrees of the evaluated object in different health states. Methods of fuzzy mathematics can be used to defuzzify this fuzzy probability information.

For example, the target node has m states$(0,1,\ldots.,m-1)$. The set of score values corresponding to each state is $V=(v_1,v_2,\ldots,v_m)$, where $v_i[0,1]$. The set of probability values for the different states of the target node obtained from the Markov model is $W=(w_1,w_2,\ldots,w_m)$. Then, according to the weighted average method, the module health index hi can be calculated by the following formula:

$$h_3=\frac{\sum_{i=1}^{m}v_iw_i}{\sum_{i=1}^{m}w_i}=\sum_{i=1}^{m}v_iw_i \tag{35}$$

2.4. Comprehensive Equipment Health State Evaluation Based on CRITIC Weight Method

2.4.1. CRITIC Weight Method

The CRITIC (CRiteria Importance Through Intercriteria Correlation) weight method is an objective Multi-Criteria Decision Making (MCDM) method used to determine the weights of various criteria (or features). It was proposed by Diakoulaki et al. in 1995 and is particularly suitable for scenarios where the importance of criteria needs to be quantified based on the data itself, especially when there are correlations between features.

Basic Principles of the CRITIC Weight Method:

The CRITIC method determines weights by analyzing the contrast intensity and conflicting character of the criteria, aiming to objectively reflect the importance of each criterion to the decision. Its core ideas are:

(1) Contrast Intensity: Quantified by the standard deviation of the criterion. A larger standard deviation indicates greater variation in the values of that criterion across different alternatives, providing more information, and thus should be assigned a higher weight.

(2) Conflict: Measured by the Pearson correlation coefficient between criteria. If a criterion has a low correlation with other criteria (i.e., high conflict), it indicates that it provides unique information and should be assigned a higher weight.

The CRITIC weight method synthesizes these two factors to generate weights through mathematical calculation, ensuring that the weights reflect the intrinsic characteristics of the data. The specific calculation steps of the CRITIC weight method are as follows:

(1) Construct the Decision Matrix: Assume there are m samples (time points) and n=3dimensions (features F_1, F_2, F_3), forming a decision matrix:

$$X=\left[x_{ij}\right]_{m\times n},\qquad i=1,2,\ldots,m,\qquad j=1,2,3 \tag{36}$$

Where x_{ij} represents the health index value of the i-th sample on the j-th dimension.

(2) Data Normalization: To eliminate the influence of dimensions, normalize the decision matrix to obtain the normalized matrix. Since all dimensions are benefit-type criteria (the larger the value, the better), the normalization formula is:

$$r_{ij}=\frac{x_{ij}-\overset{\min}{\underset{j}{}}\left\{x_{ij}\right\}}{\overset{\max}{\underset{j}{}}\left\{x_{ij}\right\}-\overset{\min}{\underset{j}{}}\left\{x_{ij}\right\}} \tag{37}$$

Where $\min_{j}\left\{x_{ij}\right\}$ and $\max_{j}\left\{x_{ij}\right\}$ are the minimum and maximum values of the j-th dimension, respectively.

(3) Calculate Standard Deviation (Contrast Intensity): For each dimension j, calculate the standard deviation σ_j of the normalized data:

$$\sigma_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(r_{ij}-\overline{r}_j)^2}, \quad \overline{r}_j = \frac{1}{m}\sum_{i=1}^{m} r_{ij} \tag{38}$$

Where $\overline{r}_j$ is the mean of the j-th dimension, and σ_j reflects the degree of variation (contrast) of this dimension.

(4) Calculate Correlation Coefficient (Basis for Conflict): Calculate the Pearson correlation coefficient ρ_{jk} between dimensions:

$$\rho_{jk} = \frac{\sum_{i=1}^{m}(r_{ij}-\overline{r}_j)(r_{ik}-\overline{r}_k)}{\sqrt{\sum_{i=1}^{m}(r_{ij}-\overline{r}_j)^2 \cdot \sum_{i=1}^{m}(r_{ik}-\overline{r}_k)^2}} \tag{39}$$

Where $j,k = 1,2,3, j \neq k$. ρ_{jk} reflects the correlation between dimension j and k.

(5) Calculate Conflict: The conflict for the j-th dimension is:

$$\text{Conflict}_j = \sum_{k=1,k\neq j}^{n}(1-\rho_{jk}) \tag{40}$$

Higher conflict (i.e., lower correlation) indicates that the dimension provides more unique information.

(6) Calculate Information Content: The information content C_j for the j-th dimension combines contrast intensity and conflict:

$$C_j = \sigma_j \cdot \sum_{k=1,k\neq j}^{n}(1-\rho_{jk}) \tag{41}$$

(7) Calculate Weights: The weight w_j for the j-th dimension is:

$$w_j = \frac{C_j}{\sum_{k=1}^{n} C_k} \tag{42}$$

The weights satisfy $\sum_{j=1}^{n} w_j = 1$。

2.4.2. Equipment Multi-Dimensional Health Comprehensive Evaluation

Experts or decision-makers subjectively assign weights to factors based on experience and knowledge. That is, assign the proportion of health evaluation based on measured data and health evaluation based on reliability. Let $h_1(t)$ be the health index based on the vibration data dimension, and γ be the weight of the vibration data dimension; $h_2(t)$ be the health index based on the measured data dimension, and δ be the weight of the real-time data dimension; $h_3(t)$ be the health index based on the historical-maintenance data dimension, and μ be the weight of the historical-maintenance data dimension; $HI(t)$ is the equipment comprehensive health index.

$$HI(t) = \gamma h_1(\mathrm{t}) + \delta h_2(\mathrm{t}) + \mu h_3(\mathrm{t}) \tag{43}$$

Where $h_1(t)\in[0,1], h_2(t)\in[0,1], h_3(t)\in[0,1], HI(t)\in[0,1]$.

The system equipment comprehensive health index HI(t) takes a value between 0-1. However, in engineering practice, qualitative methods are often used to divide the system's health status into several levels, such as Healthy, Sub-healthy, Unhealthy (slight fault state), Sick (fault state), and Critically Sick (severe fault state).

Based on actual production experience data from the enterprise, a mapping relationship between the system health index and the qualitative evaluation level of the operational status is established. The comprehensive health status of the system equipment is divided into 5 levels, as listed in Table 2.

Table 2. System Equipment Comprehensive Health Index Grade Division

Health Status Level	HI(t) Value Range	Health Status Description
Healthy	(0.9,1]	Health status is very good.
Sub-healthy	(0.8,0.9]	Abnormal signs appear, should not run continuously for long periods.
Unhealthy	(0.6,0.8]	Relatively serious abnormal signs appear, adjustment measures should be taken.
Sick	(0.4,0.6]	Serious abnormal signs appear, should be shut down for maintenance in a short time.
Critically Sick	[0,0.4]	Cannot operate, should be shut down immediately for maintenance.

3. MULTI-DIMENSIONAL HEALTH STATE ASSESSMENT PROCESS FOR COMPLEX MECHANICAL EQUIPMENT - FRACTURING PUMP CASE ANALYSIS

To verify the effectiveness and practicality of the complex mechanical multi-dimensional health state assessment model, a 5000-type fracturing pump from a certain oilfield's fracturing equipment was selected for analysis.

3.1. Comprehensive Equipment Health State Evaluation Based on CRITIC Weight Method

The on-site measurement point arrangement for the fracturing pump is shown in the following figures:

Fig 7. Vibration measurement point installation location

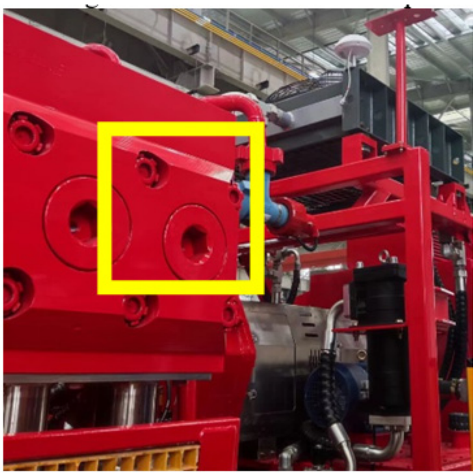

Fig 8. The left image shows the installation location of the pressure measurement point at the rear end of the hydraulic end.

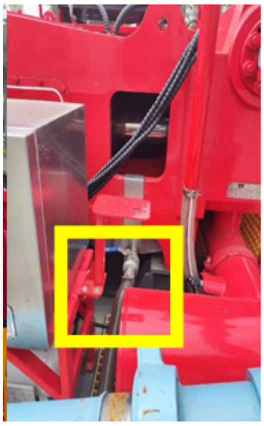

Fig 9. The right figure shows the installation location of the pressure measurement point at the front end of the suction manifold.

The fracturing pump is the core component of the entire fracturing truck. Based on manual experience, the plunger position of the fracturing pump is prone to failure, but the severity is generally low, requiring only pump inspection and replacement of parts. The place with the highest failure severity is the transmission box, but the probability of failure is very low. Therefore, after comprehensive consideration, it was decided to focus the vibration measurement points near the plunger. One measurement point was arranged at the bottom of cylinders 1, 3, and 5. These 3 points can cover abnormal vibration detection for 5 cylinders. Additionally, one measurement point was arranged vertically on the protective cover of the gear reducer or planetary reducer to assist in judgment. One pressure measurement point was arranged at the end of the power end and one at the front end of the suction manifold. [7-8]

3.2. Fracturing Pump Dataset Construction and Division

This paper mainly studies the fracturing pump of fracturing equipment. The data samples come from field detection data, historical data, and maintenance data, specifically including: Vibration signals from cylinder 1, 3, 5; Vibration signal at the connection point of the left and right ends of the power end; Discharge pressure; Suction pressure; Lubricating oil temperature; Lubricating oil pressure; Failure rate, repair rate, and prior probability of vulnerable parts.

The division of each dimension dataset is as follows:

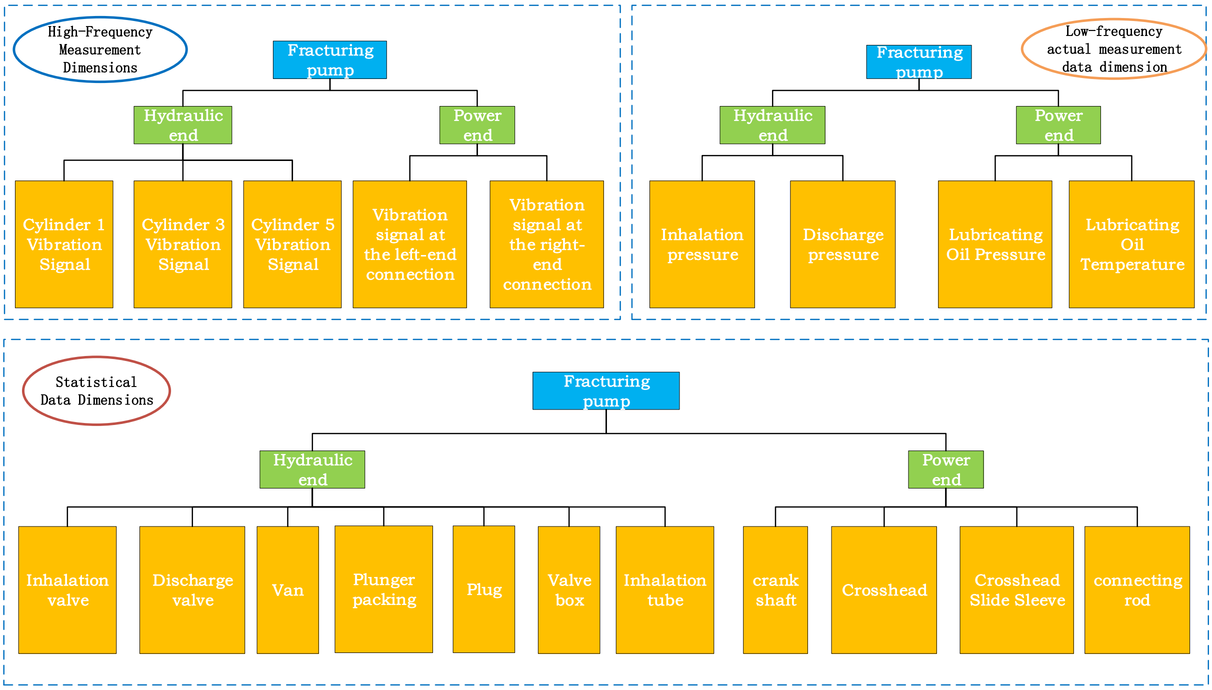

Fig 10. Multi-Dimensional Health Evaluation Dataset Partitioning for Fracturing Pumps

3.3. Fracturing Pump Multi-Dimensional Health Evaluation

3.3.1. Vibration Data Dimension Health Degree

1000 samples each were taken from the collected vibration signals of cylinder 1, 2, 3, 4, 5, and the vibration signal at the connection point of the left and right ends of the power end for denoising. Time-domain analysis was performed on the denoised vibration data, and time-domain features with obvious degradation trends were selected.

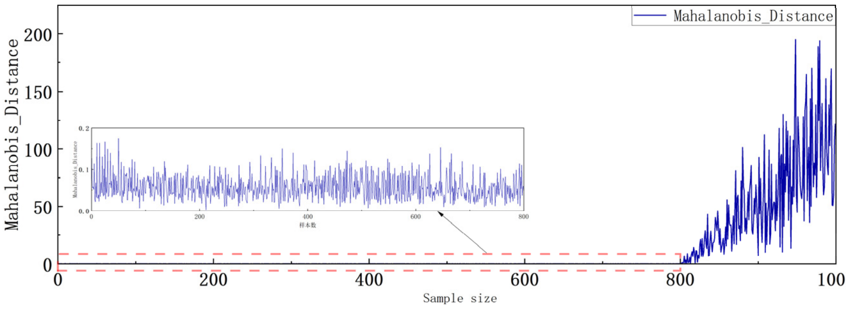

Fig 11. Variation in Martens Distance of Fracturing Pump Across Vibration Data Dimensions

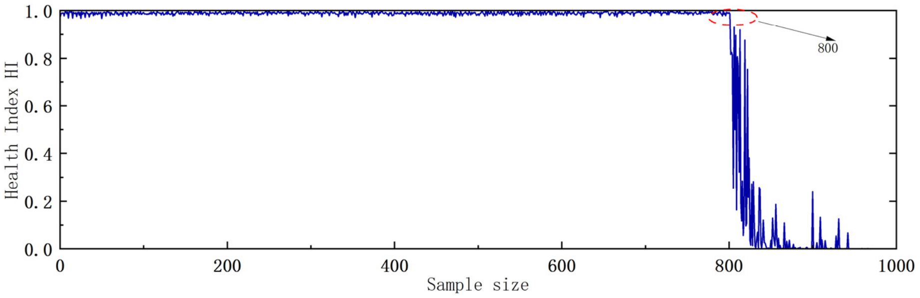

Fig 12. Changes in Fracturing Pump Health Index Based on High-Frequency Actual Measurement Data Dimensions

3.3.2. Measured Data Dimension Health Degree

Based on the real-time data measured on-site, the time function fij(t)of the characteristic parameter x of equipment i over a period of stable operation was fitted. The characteristic parameters of each equipment at time t were extracted, as listed in Table 3. For the characteristic parameters of the equipment in Table 3, the health index hij(t)of each equipment at time t was calculated according to formula (2-1), and the results are also listed in Table 3.

Table 3. Characteristic Parameter Values of Fracturing Pump Equipment

Equipment	Equipment Weight	Characteristic Parameter	Standard Value	Threshold	Char. Param. Weight	Measured Value	Health Index
Power End	0.55	Suction Pressure / kPa	250	230-280	0.6	240.1	0.977985
		Discharge Pressure / kPa	110	95-120	0.15	109.1	0.924806
		Instantaneous Flow Rate / (m3 · min−1)	1.0	0.9-1.1	0.1	0.998	0.645961
		Water Horsepower / HP	2400	2200-2550	0.15	2433.3	0.983327
Fluid End	0.45	Lubricating Oil Temp. / ∘C	50	35-80	0.65	51.7	0.999988
		Lubricating Oil Pressure / kPa	665	650-680	0.35	672.56	0.464976

The comprehensive health index $h_i(t)$for this dimension's equipment is calculated by the formula:

$$h_i(t)=\sum_{j=1}^{m}\rho_{ij}h_{ij}(t)=0.89$$

500 sets of raw data collected during normal operation of the system were used as samples, with a sampling interval of 24 hours. Using the health index calculation steps described above and applying Python, the corresponding equipment health index sample data was fitted, as shown in Figure 13:

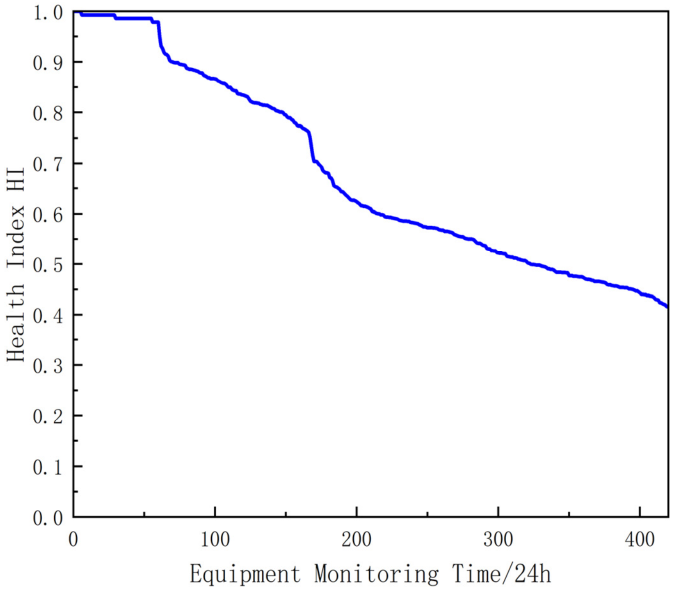

Fig 13. Changes in the Comprehensive Health Index of Low-Frequency Measured Data Dimension Equipment

3.3.3. Historical Maintenance Data Dimension Health Degree

Based on typical values from "OREDA" and "IEEE493" and expert experience, considering component characteristics and industrial application scenarios, reasonable failure rates (λ) and repair rates (μ) are provided for key components. The units are both h^{-1}. Failure rates are based on statistical data, and repair rates are estimated based on maintenance complexity. [9]

Table 4. Basic Component Failure Probability and Repair Rate

Component Name	**Failure Rate** (λ,h^{-1})	**Repair Rate** (μ,h^{-1})	**Mean Time To Repair**(MTTR,h)	**Availability** (A)
Suction Valve	1.5E-6	0.25	4.0	0.999994
Discharge Valve	2.0E-6	0.20	5.0	0.999990
Plunger Packing	5.0E-6	0.10	10.0	0.999950
Valve Chamber	3.0E-5	0.05	20.0	0.999400
Crankshaft	5.0E-6	0.04	25.0	0.999875
Crosshead Sleeve	2.0E-5	0.033	30.3	0.999395
Crosshead	1.5E-5	0.04	25.0	0.999625
Connecting Rod	5.0E-5	0.02	50.0	0.998001

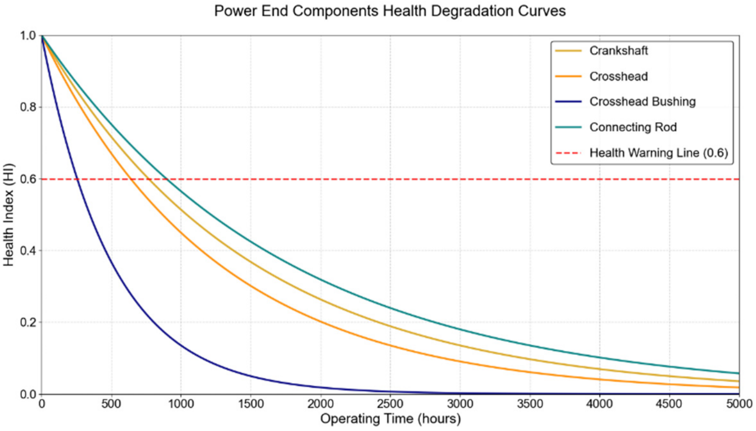

Fig 14. Changes in Health Index of Components in the Power End of Statistical Data Dimensions

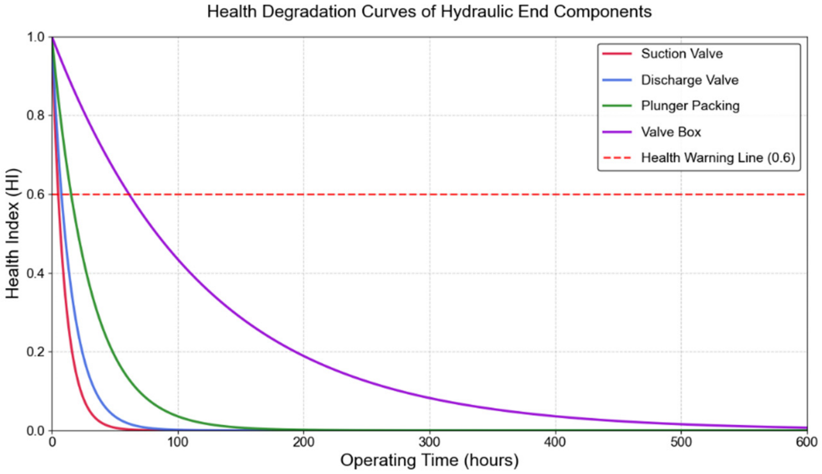

Fig 15. Statistical Data Dimensions: Changes in Health Index of Hydraulic End Components

Using Python, the health index sample data for each component and the equipment were fitted, as shown in Figures 14, 15.

3.3.4. Comprehensive Health Evaluation of the Fracturing Pump

Using the CRITIC weight method, the dynamic weight of each dimension was calculated. Then, the equipment's comprehensive health index change was fitted using the formula, as shown in Figure 16 below.

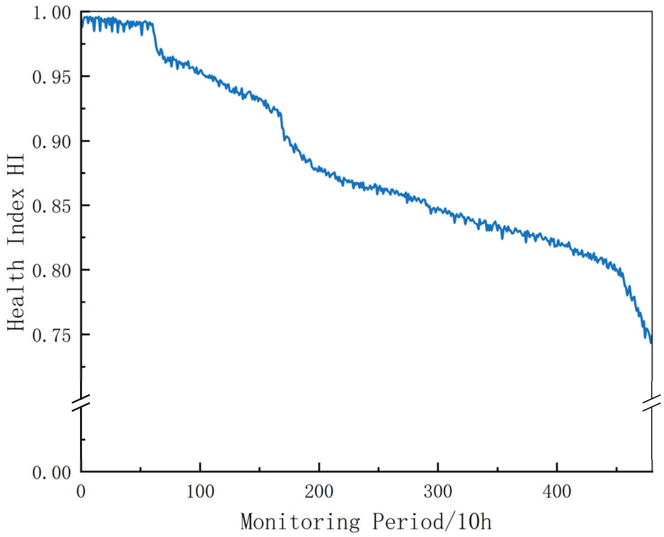

Fig 16. Sample Data Fitting Chart for Equipment Comprehensive Health Index

From Figure 16, it can be seen that within 800 hours of observation time, the equipment comprehensive health index is close to 1, belonging to a healthy state. As the equipment running time increases, the equipment comprehensive health index changes little, maintaining a certain level. However, after the equipment has run for 1800 hours, the equipment comprehensive health index begins to decline rapidly, entering a sub-healthy state. The longer the equipment runs, the smaller the corresponding equipment comprehensive health index, further proving the effectiveness and practicality of the complex mechanical multi-dimensional health state evaluation model proposed in this paper.

4. SUMMARY

(1) This paper studied the comprehensive health status of the fracturing pump system. It established a multi-dimensional health status evaluation model based on GRNN and Mahalanobis Distance, multiple equipment characteristic parameters, and reliability theory. It proposed the concept of a comprehensive equipment health index, achieving the quantitative characterization of the operational status of the fracturing system equipment. Combined with actual operational data, it enables real-time evaluation of the comprehensive health status of the system equipment.

(2) There are significant differences in the health indices of fracturing pump components and subsystems. The valve chamber and connecting rod have the lowest health indices, indicating key wear points. The health status of the power end is better than that of the fluid end, but the phenomenon of mutually influential degradation is evident. Overall, the pump's health status steadily declines with increasing operating time, with the connecting rod and crosshead sleeve requiring urgent maintenance. Future research should explore predictive maintenance models that can optimize intervention timing based on these trends.

(3) Case analysis shows that the comprehensive operational status evaluation model for complex mechanical system equipment established in this paper is effective and feasible, facilitating the comprehensive health management of complex mechanical system equipment. This can effectively improve the stability and reliability of system operation and provide technical reference for the maintenance and diagnosis of mechanical equipment on the production site.

REFERENCES

[1] Wang, J., Zhang, H., Li, X., et al. A generalized regression neural network based health assessment approach for rotating machinery. Mechanical Systems and Signal Processing. 2020, Vol. 128, p. 1-15.

[2] Chen, Y., Liu, B., Yang, K., et al. Fault detection and diagnosis based on improved Mahalanobis distance and wavelet analysis. Journal of Sound and Vibration. 2019, Vol. 438, p. 266-280.

[3] Smith, P., Johnson, M., Williams, R., et al. Multi-sensor data fusion for equipment health monitoring using CRITIC weight method. IEEE Transactions on Instrumentation and Measurement. 2021, Vol. 70, p. 1-12.

[4] Zhang, L., Wang, K., Li, M., et al. Health assessment of centrifugal pumps using vibration analysis and reliability theory. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering. 2022, Vol. 236(2), p. 145-156.

[5] Brown, T., Davis, S., Wilson, P., et al. A comprehensive health monitoring system for hydraulic fracturing pumps. SPE Annual Technical Conference and Exhibition. Dubai, UAE, 2021-09-21, p. 1-12.

[6] Jackson, R., Thompson, M. Advanced Reliability Engineering and Maintenance Management. Springer, 2018, p. 89-156.

[7] Anderson, K., Roberts, S. Multi-dimensional Data Fusion in Mechanical Systems. Elsevier, 2019, p. 234-289.

[8] Wilson, K.: Intelligent Health Monitoring of Complex Mechanical Systems Using Multi-source Information Fusion (PhD dissertation, University of Michigan, USA 2020). p. 45-89.

[9] Zhao, L.: Research on Health State Evaluation and Prediction Method of Rotating Machinery (Master thesis, Tsinghua University, China 2019). p. 23-67.