*mathematics*

MDPI

*Article*

# Doubly Stochastic Hidden Semi-Markov Model Optimized by Swarm Intelligence for Railway Electronic Equipment Health Management

**Yanzhi Pang [1], Renwei Kang [2], Jianqiu Chen [1], Lei Liu [2,*] and Jidong Lv [3]**

[1] Guangxi Key Laboratory of International Join for China-ASEAN Comprehensive Transportation, Nanning University, Nanning 530000, China
[2] Signal & Communication Research Institute, China Academy of Railway Sciences Co., Ltd., Beijing 100081, China
[3] School of Automation and Intelligence, Beijing Jiaotong University, Beijing 100044, China
* Correspondence: leiliu@rails.cn

**Abstract**

Electronic equipment is essential to railway transport, as its performance directly impacts the safety of train operations. However, current operation and maintenance practices face several problems, including unrecognized degradation of equipment performance, difficulty in diagnosing faults after they occur, and coarse-grained maintenance strategies, These problems may lead to an excessively broad scope of fault impact, the continued operation of defective equipment, and other serious consequences. In this context, the paper presents an intermittent fault (IF)-based health management approach for railway electronic equipment. Building on fault mechanism analysis for electronic equipment, a mathematical parameterized characterization of intermittent faults is achieved by integrating the fault frequency and intensity vectors into the state transition probability matrix of a hidden Semi-Markov model (HSMM). To solve the non-convex parameter estimation problem, a global optimization algorithm based on swarm intelligence is constructed to maximize the log-likelihood objective function, so as to evaluate the current operational state of the equipment. The paper then uses the equipment observation sequences predicted by the long short-term memory network (LSTM) to obtain the remaining useful life (RUL) values at any timepoint. The estimation serves as the basis for proactively adjusting maintenance decisions. In experiments based on real-world power module faults, compared with the conventional Baum–Welch trained HSMM method, the proposed method in this paper improves state recognition accuracy by 2.15% and reduces the RMSE of RUL prediction by 16.67%, verifying its effectiveness.

**Keywords:** health management; doubly stochastic process; HSMM; swarm intelligence optimization; LSTM; state recognition; life prediction

**MSC:** 49Q22

Check for updates

Academic Editor: Alexander Zeifman

Received: 25 August 2026
Revised: 1 October 2026
Accepted: 4 October 2026
Published: 23 October 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

By the end of 2025, China's operational railways have reached 165,000 km in length, among which 50,400 km are high-speed railways. As the country that operates the largest railway network, China serves the safety and convenience of thousands of passengers with its technologies and services. Within this network, electronic equipment—installed

to perform train control, monitoring, protection, etc.—has seen an increase in both its application scale and quantity. From a life-cycle perspective, the repetition of in-service operation and out-of-service maintenance constitute the life cycle of electronic equipment until its obsolescence. During the in-service operation phase, the equipment—under normal circumstances—operates as planned and in an orderly fashion to ensure train control and protection [1]. However, when faults occur, the established order and logic are disrupted. Regardless of the strategies employed, out-of-service maintenance primarily focuses on fault prevention. From this perspective, the health state of electronic equipment serves as the precondition that decides any Operation and Maintenance (O&M) methods applied on-site. Unfortunately, the state is not effectively evaluated throughout the life cycle, as manifested by the following two aspects:

(1) Fault diagnosis and response of electronic equipment are typically conducted after the fault has occurred, resulting in a lack of technical means for proactive fault prevention. The approach leads to difficult diagnosis and localization, inefficiency, and low accuracy. A typical example that unveils the relatively large fault impact scope occurred on 26 February 2021, when faults of electronic equipment resulted in train delays lasting 189 min. In this incident, the train control system issued an alarm and the train executed an emergency brake to a stop. The cause was later identified as an internal short circuit in the safety digital input/output unit of the control system. However, at the moment the fault occurred, there were no technical means available to assist in locating or diagnosing the specific fault point.

(2) The current maintenance strategies are coarse-grained and undifferentiated. Different types of electronic equipment used on a massive scale across China are subject to one-size-fits-all maintenance strategies throughout their full life cycles. For example, there are five types of onboard equipment used at scale for train operation control. Although each type differs in technical architecture, development platform, and manufacturing process, they all undergo the same maintenance tasks within the same maintenance cycle. In addition, the development of the maintenance strategies is primarily based on experience rather than the health state. And there are no technical means to prompt warnings or alerts in advance when the equipment degrades in performance or is insufficient to perform the transport tasks. On 19 May 2023, an anomaly occurred in the track-circuit decoding board of a train during operation. It represents a typical example where the operation of key equipment beyond its service life and untimely maintenance led to health degradation.

Current fault response and maintenance strategies are insufficient to meet the transport requirements. Therefore, to ensure the safe and efficient operation of trains, it is crucial to establish methods to evaluate the health state and develop sound maintenance strategies that allow for timely warnings before major accidents occur. Achieving full-life-cycle health management of transport infrastructure and equipment has become a common consensus and development direction in relevant domains [2]. With the aid of intelligent reasoning algorithms, the health management of electronic equipment facilitates a shift from passive fault response to proactive prevention. This transition is achieved by acquiring predictive information on equipment faults and health state, which is then integrated with the maintenance resources concerned. The integration provides a basis for implementing measures before faults occur, aiming to improve equipment reliability and reduce O&M costs [3].

The premise of health management lies in fault prediction, which primarily involves recognizing the operational state of equipment for a forthcoming period and predicting its Remaining Useful Life (RUL) [4]. It is important to note that, in engineering practice, Intermittent Faults (IFs) constitute the most common type of fault encountered during the mid-to-late stage of service, accounting for approximately 40% to 70%. The non-

permanent faults that occur repeatedly and stochastically last for a limited amount of time. The equipment concerned can often return to normal operation within a short timeframe without maintenance. Such faults pose significant maintenance and diagnosis challenges, leading to issues like no fault found and retest OK. The consequences—increased costs, unnecessary and repeated testing, and excessive maintenance alike—severely disrupt the availability and integrity of railway equipment.

Relevant research indicates that the characteristics of IFs offer a glimpse of the health state of the electronic equipment concerned. In this sense, IF-based health management provided a valid solution [5], but today, the diagnosis and recognition of IFs remain in the initial stages. Additionally, the complexity and diversity of O&M scenarios for railway electronic equipment make IF study all the more difficult, presenting greater challenges to equipment health management. The complexity of these O&M scenarios [6] is reflected in the following aspects:

(1) Complex operational environments. China's crisscrossed network of railways covers a vast land with significantly diverse geographical conditions. As a result, both fixed electronic equipment along the railway lines and on-board equipment work in complex environments, facing thunderstorms, rain, snow, extreme cold, high temperatures, electromagnetic interference, frequent vibrations, etc. These harsh conditions complicate the correlations between the electronic equipment faults or reliability indices on one hand and the operational environments on the other.

(2) Alternation of multiple maintenance methods. The on-board electronic equipment travels around the country with the trains. With different destinations come different maintenance strategies characterized by the diversity of maintenance items, cycles, operating standards, procedures, and operator competencies. As a result, multiple maintenance methods are implemented alternately throughout the life cycle of the equipment, making it difficult to evaluate the impact of different strategies on equipment health.

(3) The spatial and temporal multi-scale nature of data and information. A vast amount of information—in-service monitoring data, out-of-service maintenance records, and event logs—is generated around electronic equipment in China using a variety of acquisition devices and means to reflect the service conditions of different components or equipment. The data collection—being either objective or subjective, qualitative or quantitative—may occur at different intervals and levels of accuracy, adhere to different measurement units, be presented in textual, visual, or auditory formats, and be stored on different physical carriers in a distributed manner. The variation contributes to the temporal and spatial multi-scale nature of the data and information available for health management, making analysis and data mining challenging.

In the context of Industry 5.0, it is an inevitable trend in equipment health management to explore preventive maintenance strategies that balance task performance and cost-efficiency. Consequently, the ability to predict faults based on the estimated degradation of equipment and to propose optimal time windows for maintenance interventions become crucial [7]. The paper studies the health management methodology for railway electronic equipment throughout its full life cycle and in complex O&M scenarios. An IF-based degradation model is developed to recognize the operational state of the equipment and predict its RUL. Building on this model, proactive preventive maintenance strategies can be developed in advance to enable timely warnings of any potential safety hazards and the mitigation of any unexpected circumstances—performance degradation, task defects, and unscheduled operations—caused by the sudden faults of railway transport orders. The main contributions of this paper are as follows:

An IF-based health management approach is proposed for railway electronic equipment. The approach utilizes a Hidden Semi-Markov Model (HSMM) [8] to characterize

the performance degradation process of the equipment, particularly the stochasticity in both operational state transitions and state durations. The effectiveness of the proposed methodology is validated through experiments, supporting the transformation from passive response to proactive prevention.

The paper proposes a recognition strategy where HSMM parameters are optimized with swarm intelligence to estimate the equipment state. This enables the accurate operational state recognition of railway electronic equipment.

A long short-term memory (LSTM) method to predict the observation sequence for equipment is proposed, which enables RUL prediction for electronic equipment in different states and at various timepoints throughout its full life cycle. This allows for proactive interventions in maintenance decision-making.

The remaining part of the paper is organized as follows. In Section 2, the related work is reviewed, while in Section 3, the IFs mechanism of railway electronic equipment is analyzed. Section 4 elaborates on the establishment of the health management model for electronic equipment, while Section 5 conducts experiment and analysis. Section 6 discusses application prospects, and Section 7 concludes the work of this paper.

## 2. Related Work

Health management represents an upgrade of the traditional built-in testing and state monitoring approach. The progress is to facilitate the transition from conventional sensor-based diagnosis to intelligent prediction. It encompasses pre-fault forecasting, comprehensive health state control, and future reliability prediction. The ultimate goal is to develop a predictive maintenance approach based on current equipment capacity, balancing task performance with cost-efficiency [9]. Aligned with the pressing demands of industrial sites, this transition signifies the trend toward equipment autonomy, integration, precision, and an all-in-one design, which can be essential for future progress. As researchers in China and abroad have conducted extensive studies on equipment health management, some progress has been achieved, which is reviewed below from three perspectives, namely research methods, research tasks, and engineering practice.

From a method point of view, three types of relatively mature and commonly used approaches for health management have emerged. They are based on failure physical models, statistical reliability, and data-driven techniques, respectively. The first, which leverages the failure physical models, focuses on the prediction and diagnosis at the electronic component level, utilizing the life-cycle loads and the failure mechanisms of the products. The approach enables safe and effective state monitoring for electronic equipment using highly reliable correlations between characteristic parameters on one hand and the internal junction temperature or aging state of the equipment on the other. However, for complex electronic equipment composed of thousands of components, the challenges posed by accurate mathematical modeling and substantial computational demands result in limited operability for practical implementation. Consequently, the approach has seen less study and application in the industrial sector. The second, based on statistical reliability, relies on historical data to fit the life distribution of objects from a reliability standpoint. While these methods provide a basis for estimating average RUL, they often generate significant errors. The third data-driven approach involves installing various sensors at critical locations of the research subjects or conducting diverse experiments. The data collected from the sensors or experiments helps enable an indirect understanding of equipment state and operational patterns and inform maintenance decisions [10].

The data-driven approach represents the mainstream technology developed thus far. In the face of increasingly complex and tedious health management tasks, the traditional approaches of offline testing, signal processing, and diagnostic expert-based systems for

vulnerable components have revealed their significant disadvantages of high costs and low efficiency. Additionally, the advancements in information and IoT technologies translate into a vast amount of data accumulated across various information platforms, which offer a robust data source for data-driven health management. Enhanced by deep learning and other AI technologies, health management has achieved remarkable results in railways, energy, aviation [3], and other domains, leading to the emergence of convolutional networks [11], LSTM networks [12], Transformer networks [13], digital twins [2], etc. It is essential for health management to continuously adapt to business requirements, driving the evolution of deep models from research paradigms characterized by single modality, single task, and limited data to those embracing multiple modalities, multiple tasks, massive data, and ultra-large models. Large-scale foundation models, exemplified by ChatGPT, shed light on new research directions for developing tasks applicable to equipment health management [3]. Researchers have devised dedicated solutions to address the domain-specific requirements for health management. For instance, Zhou et al. [14] introduced a deep neural network model emphasizing the interactions among similar features and the significance of different time steps, known as the similar feature interaction model with distance self-attention. To tackle errors and uncertainties in physical measurements, as well as the heterogeneous degradation of equipment under time-varying operating conditions, Xue et al. [15] proposed a prediction framework that combines the two-phase clustering with particle filtering and LSTM, applicable to the degradation research of lithium-ion batteries. In addressing the random uncertainty of internal variations during equipment degradation, Cao and Peng [16] proposed a maximum likelihood estimation method for random uncertainties based on probability functions and uncertainty distributions to predict RUL in gallium arsenide lasers and gyroscopes. Targeting the uncertainties in predictions that significantly affect industrial equipment, Li et al. [17] proposed an innovative Bayesian deep learning framework that captures the combined effects of epistemic and aleatoric uncertainties. The arrangement is applicable to the health management of hydraulic mechanisms used in circuit breakers.

From the perspective of research tasks, health management for electronic equipment primarily comprises two elements: current state recognition and future fault prediction [15,18]. The former mainly focuses on IFs [19], which have been identified as the most significant type of fault encountered during the mid-to-late periods of equipment service in engineering practice [20]. IFs exhibit a discrete and stochastic pattern, alternating between fault and non-fault behaviors [21]. Although they often disappear without any maintenance or intervention [22], they pose a serious threat to the operational safety of the systems [23]. Statistical data from industrial sites indicate that IFs primarily occur at the equipment hardware level. Therefore, the paper focuses exclusively on the hardware-related IFs rather than the software-related ones. Numerous methods have been developed for state recognition, using multi-agent system [24], improved deep forest [25], multi-scale enhanced convolution-transformer network [13], the combination of Walsh transform and a multilayer perceptron [26], and the integrated data detection via dynamic time warping and a variational autoencoder of feature engineering. Generally speaking, deep learning is the dominant approach [27]. Fault prediction for electronic equipment aims to accurately estimate the RUL based on state recognition to determine the need for targeted preventive measures, such as maintenance or replacement. The basic concept is to establish an equipment degradation model [16,28]. For example, Liu et al. [29] proposed a memory network-based time series multi-step fault forecasting model dedicated to the air braking system on-board. Built on the LSTM network and the attention mechanism, Cheng et al. [30] proposed the use of contact resistance characteristics and intermittent faults for the RUL prediction of electrical connectors. Rahman et al. [31] introduced a framework

for the central management of vehicle health. The proposal aimed to monitor the health state of individual vehicles using an IoT-driven multi-layer heterogeneous network and machine learning.

Traditional RUL prediction methods emphasize the nonlinear mapping relationships of time series data. Due to the lack of adequate consideration of the structural information of the equipment, the prediction accuracy tends to be low [32]. To enhance the accuracy and effectiveness of RUL prediction, researchers in the field target their studies on two aspects, namely, the model itself and the processing of data involved in model computation. The first focuses on better computational power, including the spatio-temporal fusion attention model based on a graph neural network [32], the cloud-edge-based lightweight temporal convolutional network model [11], and the health model designed to maximize system utility by leveraging the dependencies between measuring and measured systems [33]. The second addresses sample data imbalance and data compression. Notable solutions include eXtreme gradient boosting enhanced by a conditional generative adversarial network [34], the multi-sensor graph transfer network approach [35], and the efficient data compression technique with the encryption capability [36].

The state recognition of faults and RUL prediction are two important yet distinct subjects. The deep learning-based methods applied today tend to resolve the two issues separately. In other words, one network targets one single task, leading to low training efficiency, the waste of computational resources, and other drawbacks [37]. However, Literature [10] introduces a multi-task deep learning model designed for the simultaneous implementation of equipment healthy state evaluation and RUL prediction. The model, validated using a public dataset of engines, has demonstrated its effectiveness and advancement. The concept of integration presents a novel pathway for improving equipment health management.

From the perspective of engineering practice, health management for most electronic equipment (with the exception of electronic units with clearly identified degradation patterns, such as batteries) remains in the exploratory stage, which translates into unsatisfactory results in engineering applications [5,38]. This is because the electronic equipment deteriorates in an unpredictable pattern, and faults tend to appear in a sudden manner. Given the absence of any noticeable signs of degradation for most electronic components, it is difficult to capture the characteristic parameters useful for health management in engineering practice. The aforementioned reasons lead to significant challenges to health management [39].

To sum up, previous research on the health management of railway infrastructure has primarily focused on equipment with clearly defined degradation patterns, such as the wear of rails and switch machines, while there has been limited attention given to electronic equipment. The complex structural composition, unique business characteristics, and time-varying O&M scenarios of this equipment continue to pose significant challenges to the study of the health degradation of railway electronic equipment. In this context, the paper explores an IF-based health management approach for electronic equipment, which is of substantial practical significance.

## 3. Mechanisms Analysis of Intermittent Faults in Railway Electronic Equipment

For any electronic equipment, regardless of its operating environments, its electronic components are sure to degrade to varying degrees over time, during which IFs—within a specific tolerance range—occur to the equipment and/or the system. In general, when the environmental stress disappears, so does the intermittent fault. The reverse is also true—when the environmental stress is reapplied, the fault emerges again. The pattern

indicates that the root cause of IFs stems from certain defects or damages within the electronic equipment, constituting the internal cause, while the excitation of environmental stress serves as the external cause.

In the case of railway electronic equipment, it suffers from internal damages amid its exposure to temperature changes, vibration, corrosion, and other environmental and operational stresses over its long-term service life. The inherent defects of some equipment further complicate the situation, constituting the internal causes. For example, IFs in electrical connectors arise when contact resistance exceeds the design threshold within a short timeframe. That aside, environmental stress—temperature fluctuations and train vibrations in particular—may excite the equipment with damages and/or defects in real-time. These excited damages and/or defects result in a change to transmission signals, manifested as observable IFs. In other words, the occurrence of environmental stress can be understood as a stochastic event and process.

During train operation, the faults of certain control equipment can often be restored by a simple restart of the equipment, which exemplifies this stochastic process.

It is evident that environmental stress and component degradation contribute to IFs as the major factors. Temperature and vibration are the most important manifestations among all environmental stress factors, accounting for 40% and 28% [5], respectively. This can be explained by the presence of wheel–rail vibrations of trains and the fact that railway equipment is installed in various geographical regions.

The incidence rate and duration of IFs are abstracted as a random degradation process, and the Inhomogeneous Poisson Process is used to characterize the random degradation process of the equipment. The intensity function of fault occurrence is as follows:

$$\lambda(t) > 0,\ t \geq 0 \tag{1}$$

The cumulative intensity function is as follows:

$$\Lambda(t) = \int_0^t \lambda(s)ds \tag{2}$$

(1) Define boundary 1: normal state -> intermittent fault

It represents the time distribution of the first fault occurring from a normal state, where the probability density function (PDF) of the first fault time is the first passage time density of the intensity function:

$$f_{N\to I}(t) = \lambda(t)e^{-\Lambda(t)} \tag{3}$$

The mode $t_1^*$ of this PDF is defined as the mathematical boundary between normal state and intermittent fault:

$$t_1^* = \text{argmax} f_{N\to I}(t) \tag{4}$$

That is, when $t < t_1^*$, the equipment is in a normal state; when $t \geq t_1^*$, the equipment is in an intermittent fault state.

(2) Define boundary 2: intermittent fault -> permanent fault

It represents that the cumulative number of faults occurring within the time interval $[0, t]$ exceeds a specified critical threshold $K$, and the equipment cannot recover after the last fault. The occurrence time of a permanent fault is defined as the first passage time of the $K$th fault. The PDF of the occurrence time of the $K$th fault is calculated by the following equation:

$$f_{I\to P}(t) = \frac{[\Lambda(t)]^{K-1}}{(K-1)!}\lambda(t)e^{-\Lambda(t)} \tag{5}$$

The mode $t_2^*$ of this PDF is defined as the mathematical boundary between intermittent fault and permanent fault:

$$t_2^* = \text{argmax} f_{I \to P}(t) \tag{6}$$

That is, when $t_1^* < t < t_2^*$, the equipment is in an intermittent fault state; when $t \geq t_2^*$, the equipment is in a permanent fault state.

By defining the mode of the PDF as the mathematical boundary between the normal state and the intermittent fault state, and between the intermittent fault state and the permanent fault state, we do not treat physical faults as deterministic events. Rather, this definition provides a deterministic counterpart for random degradation processes, based on the risk-minimization principle for engineering decisions. This boundary represents the state transition point with the lowest misjudgment cost in a statistical sense and offers an actionable time point for maintenance decisions. In engineering practice, this treatment is logically consistent with first-passage-time-based life prediction and existing vibration severity classification standards.

In summary, in the framework of the Inhomogeneous Poisson Process, the mathematical boundaries of normal state, intermittent fault and permanent fault are uniquely determined by the probability density function.

Under interactions among the initial design and manufacturing defects, external environmental stress, and aging of components, the occurrence of IFs is characterized by stochasticity, repeatability, and recoverability [13].

(1) Stochasticity: In the time dimension, IFs occur at varying timepoints and for different durations;

(2) Repeatability: IFs occur repeatedly at identical positions;

(3) Recoverability: Faults and errors can be eliminated.

Corresponding fault cases are common in the actual applications of railway equipment. An example of IFs based on Automatic Train Protection (ATP) equipment is illustrated in Table 1.

**Table 1.** Examples of IFs of ATP equipment.

| Fault Time | $t_d$ | $n_a$ | Cause of Fault | Inducing Factor | Fault Characteristic |
|---|---|---|---|---|---|
| 28 October 2023 16:00 | 34 | 6 | Bad contact of solder joint on chip | Manufacturing process | Stochasticity |
| 22 May 2024 7:25 | 33 | 7 | Abnormal hardware circuit | Design flaw | Repeatability |
| 20 August 2024 10:08 | 28 | 2 | Abnormal cable connector | Vehicle vibration | Recoverability |
| 22 August 2024 14:07 | 17 | 1 | Abnormal crystal oscillator in communication board | Temperature change | Repeatability |
| 30 August 2024 18:59 | 14 | 2 | Abnormal capacitance of a CPU board | Equipment aging | Recoverability |

In Table 1, $t_d$ represents delay minute due to fault, and $n_a$ represents the number of affected following trains. As shown in Table 1, IFs with various characteristics and inducing factors often occur in railway equipment. Their presence can lead to significant delays, which to varying degrees affect the normal operation of subsequent trains and severely disrupt the transport order. It should be pointed out that the repair of such IFs is minimal, i.e., the repair only restores the equipment to its operational state without altering the underlying degradation process. To address these issues, it is crucial to establish a comprehensive health management model, supported by effective methods for accurately extracting characteristic parameters that account for the various characteristics of IFs. The following section provides a detailed analysis.

## 4. Health Management Model

This section starts by outlining the overall framework and process of health management for railway electronic equipment based on IFs. It then elaborates on three key techniques, including parameter extraction for the characteristics of IFs, state recognition, and RUL prediction.

### *4.1. Overall Framework and Process of Health Management*

To implement state recognition for railway electronic equipment based on IFs, it is essential to develop a rational health model that incorporates the dynamic characteristics of these faults. It is well established that the performance of operating electronic equipment gradually deteriorates, increasing the likelihood of faults due to the effects of operational and environmental stresses. However, there are currently no effective means to observe these faults directly. The characteristics of IFs are capable of characterizing the deteriorating operational state of the equipment and can be made observable through reproduction experiments. In other words, the objective is to infer the performance state of the equipment by observing the characteristics of IFs. The HSMM is a modeling approach that infers potential and hidden state sequences from observation sequences [40], of which the core lies in estimating the conditional probability distribution of unknown variables based on observable variables. In this study, the HSMM is utilized to establish the health management model for railway electronic equipment.

HSMM facilitates a doubly stochastic process that comprises a hidden Markov chain with a specific number of states, along with a set of explicit stochastic functions [41,42]. In this model, both transitions between states and between states and observations are stochastic. During the operation of electronic equipment, its degraded state is not directly observable—referred to as the hidden state—and can only be inferred through the stochastic process of observation sequences.

The HSMM is expressed as follows:

$$\lambda = [\boldsymbol{\pi}, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}] \tag{7}$$

where

$\boldsymbol{\pi}$ represents the initial state probability, which indicates the probability of various states at the initial time, and is denoted as $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_N)$, where:

$$\pi_i = P(y_1 = s_i), 1 \leq i \leq N \tag{8}$$

This indicates the probability that the initial state is $s_i$.

$\boldsymbol{A}$ represents the state transition matrix, which indicates the probability of transition between states, and is denoted as $\boldsymbol{A} = [a_{ij}]_{N \times N}$, where:

$$a_{ij} = P(y_{t+1} = s_j \mid y_t = s_i), 1 \leq i, j \leq N \tag{9}$$

This indicates the probability that the state is $s_j$ at the timepoint following any timepoint $t$ when the state is $s_i$.

$\boldsymbol{B}$ represents the state observation probability matrix, which indicates the probability of obtaining each observation based on the current state, and is denoted as $\boldsymbol{B} = [b_{ij}]_{N \times M}$, where:

$$b_{ij} = b_i(x_j) = P(x_t = o_j \mid y_t = s_i), 1 \leq i \leq N, 1 \leq j \leq M \tag{10}$$

This indicates the probability of obtaining the observation $o_j$ (or $x_j$) at any time $t$ when the state is $s_i$. The set of all observations is denoted as $O$, i.e., $O = \{o_1, o_2, \ldots, o_M\}$.

Once a state $s_i$ starts, it lasts for $d$ time steps and then transitions to the next state $s_j$. The observation probability during this period is denoted as:

$$b_i(x_{t:t+d-1}) = \prod_{s=t}^{t+d-1} b_i(x_s) \tag{11}$$

$\boldsymbol{D}$ represents a state duration matrix, which indicates the duration of the system at a given state, and is denoted as $\boldsymbol{D} = [p_i(d)]$, where:

$$p_i(d) = P(D(s_i) = d \mid y_t = s_i), 1 \le i \le N \tag{12}$$

This indicates the probability that the occurrence of state $s_i$ lasts for $d$ periods of observation, which usually follows a probability distribution such as Gaussian, Gamma, and Weibull.

The duration matrix $\boldsymbol{D}$ is defined over discrete time steps (e.g., number of sampling intervals). When the dwell-time is modeled by a continuous distribution $f(\tau)$ (e.g., Gaussian, Gamma), we discretize it as $P(D = d) = \int_{d-1}^{d} f(\tau)d\tau$ for $d = 1, 2, \ldots, D_{\max}$, with truncation at $D_{\max}$ and subsequent renormalization. This ensures that the discrete duration probabilities used in the HSMM are consistent with the discretized version of the specified continuous dwell-time distribution.

HSMM allows for perceiving the state variation and transition process of equipment based on observation sequences. The entire model is illustrated in Figure 1.

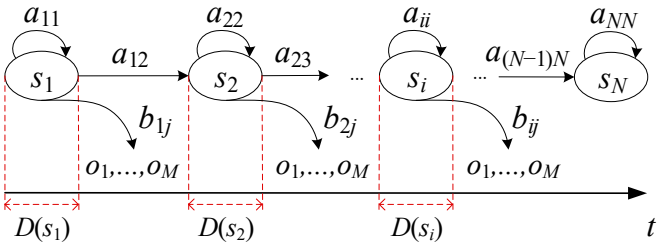



**Figure 1.** Schematic diagram of HSMM.

As shown in Figure 1, with the given HSMM $\lambda = [\boldsymbol{\pi}, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}]$, the observation sequences $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ are generated as follows.

(1) Set $t = 1$ and select the initial state $y_1$ according to its probability $\boldsymbol{\pi}$.

(2) Select the observation variable value $x_t$ according to the state $y_t$ and the state observation probability $\boldsymbol{B}$.

(3) Generate the state $y_{t+1}$ at the next timepoint of the model according to the state $y_t$ and the state transition matrix $\boldsymbol{A}$.

(4) If $t < n$, set $t = t + 1$ and proceed to Step (2); otherwise stop the process.

In this process, three fundamental points need to be addressed for HSMM:

(1) With the given model $\lambda$, how do we effectively calculate the probability $P(\boldsymbol{X} \mid \lambda)$ of generating the observation sequence $\boldsymbol{X}$? This involves calculating probabilities to evaluate the degree of match between the model and the observation sequence. This aspect is typically addressed using the Forward–Backward algorithm.

(2) With the given observation sequence $\boldsymbol{X}$, how do we adjust the parameters of the model $\lambda$, to maximize the conditional probability $P(\boldsymbol{X} \mid \lambda)$ of this observation sequence? This involves learning to predict the RUL of the equipment. In other words, how do we train the model to better describe the observation data? This is typically addressed by the Baum–Welch algorithm.

(3) With the given model $\lambda$ and the observation sequence $\boldsymbol{X}$, how do we determine the most matching state sequence $\boldsymbol{Y} = \{y_1, y_2, \ldots, y_n\}$, which corresponds to the maximum

conditional probability $P(\boldsymbol{Y} \mid \lambda, \boldsymbol{X})$? This involves predicting the hidden equipment state from the observation sequence. It is typically addressed using the Viterbi algorithm.

These 3 problems are progressive: the first is a calculation problem, which is to calculate the likelihood probability of the observation sequence $\boldsymbol{X}$ under model $\lambda$; the second is a learning problem, which is to select the parameter corresponding to the maximum likelihood probability as the model parameter based on the previously obtained likelihood probability; the third is a prediction problem, which is to estimate the state sequence $\boldsymbol{Y}$ that best matches the observation sequence $\boldsymbol{X}$ based on the previously obtained model parameters.

The overall process of equipment health management based on IFs is shown in Figure 2, and the specific steps are as follows:

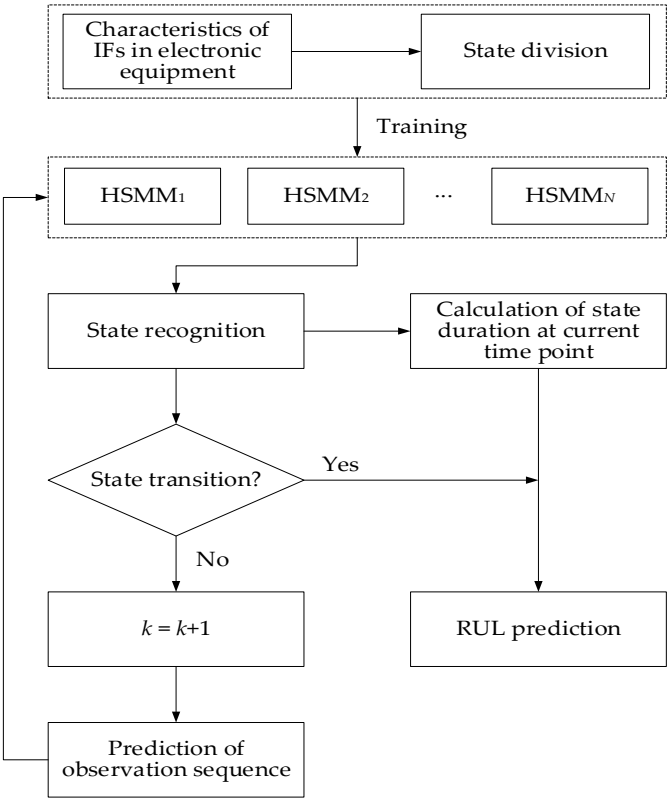



**Figure 2.** Overall process of health management for electronic equipment.

Step 1: Extract the characteristics of IFs and classify the operational state of the equipment.

Step 2: Estimate the parameters of the HSMM according to observation sequences, i.e., the characteristic parameters of IFs, using the forward–backward algorithm, and solve the model $\lambda$. Train the model using the Baum–Welch algorithm, and output the model parameters corresponding to the maximum probability of the observation sequences.

Step 3: Input the observation sequences into the HSMM of each operational state, calculate the probability $P(\boldsymbol{X} \mid \lambda)$ of the observation sequences in different HSMMs, and output the model state with the maximum probability, i.e., the current operational state of the equipment.

Step 4: Calculate the state duration of the equipment at the current timepoint and predict the RUL of the equipment.

Step 5: Evaluate whether the state outputted by the model indicates a transition. If so, update the RUL of the equipment according to the current state and the calculated state

duration. If not, increase the time step, transfer to the observation sequence prediction for the next timepoint, and proceed to Step 2.

To summarize, in order to address the health management requirements for railway electronic equipment, the key technical issues that need to be resolved include parameter extraction for the characteristics of IFs, operational state recognition, and RUL prediction. It is important to highlight that the solutions are tailored to fit the actual scenarios encountered in railway operations. Specifically, for characteristic parameter extraction, a fuzzy clustering algorithm is utilized to classify the observation data corresponding to each operational state. For state recognition, a swarm intelligence optimization algorithm is implemented to enhance the parameter calculation process for HSMM. Regarding RUL prediction, the focus is on addressing RUL equivalence for identical states across different timepoints. These three aspects are discussed in detail below.

### *4.2. Parameters for Characteristic of IFs*

Although faults in electronic equipment typically occur suddenly, they are fundamentally connected to a gradual process of state degradation influenced by factors such as temperature, humidity, vibration, and the operating stress experienced by the internal components. In other words, electronic equipment experiences a series of health states with progressive deterioration throughout the operational lifespan. This understanding offers a theoretical foundation for effective health management. The specific process of state degradation is detailed as follows:

(1) In the initial stage of degradation, the equipment performance parameters change slightly, without inducing faults.

(2) When degradation reaches a certain level, continuous changes in equipment performance parameters under the action of environmental stress result in IFs.

(3) As degradation progresses, the equipment performance parameters change significantly, consequently increasing the intensity, frequency, and duration of IFs. When the duration of faults increases to a critical threshold, the frequency of faults begins to decrease, ultimately causing IFs to evolve into permanent faults.

The degraded state experienced by devices throughout the fault evolution process is not directly observable in the real world. Instead, it can only be perceived from the symptoms manifested by state observations. From this perspective, the strong correlations between the IFs of electronic equipment and its degradation and failure provide significant advantages for evaluating and predicting its operational state based on the state of IFs and the characteristics of its lifespan. Given that aging is a primary cause of IFs in electronic equipment and that IFs serve as precursors to permanent faults throughout the full life cycle [20], the characteristics of IFs can be utilized to effectively characterize and evaluate the health state of the equipment.

The health state of equipment is the result of the combined effects of multiple physical parameters. These parameters are often multidimensionally distributed, making their measurement challenging. As a result, quantitatively evaluating the health state is both difficult and complex. Typically, the health state is reflected in the operational characteristics under specific stress conditions. It can be characterized by the frequency, duration, and amplitude of sudden changes in operational performance parameters, which correspond to the frequency, duration, and intensity of faults, particularly IFs. Consequently, these three characteristic parameters can serve as indices for the degradation process of a health state and can be used to quantitatively describe IFs.

The frequency of IFs $N_f$ represents the count of IFs occurrences within a fault burst.

The fault duration $T_A$ represents the length of time for a single IFs occurrence, i.e., the timeframe from the appearance to the disappearance of the fault symptom. The mean duration of all fault occurrences can be expressed as follows:

$$T_A^m = \frac{\sum_{i=1}^{N_f} T_A^i}{N_f} \tag{13}$$

where $T_A^i$ represents the duration of the $i$-th IFs occurrence.

Fault intensity represents the deviation intensity of a fault from the normal level, expressed as the ratio of the operational performance parameter to the fault threshold:

$$\psi_f = P_w / P_{th} \tag{14}$$

where $P_w$ represents an operational performance parameter of the equipment component, and $Pth$ represents the failure threshold. Given $\psi_f \in (0, \infty)$, $\psi_f = 1$ indicates fault occurrences, and a larger $\psi_f$ indicates more intense faults. For "intermittent faults" in electronic equipment, the fault threshold essentially defines the boundary between the "normal state" and "failure". Since gaps in electronic equipment (e.g., solder joint voids, loose connectors, and PCB delamination) are difficult to observe directly, the corresponding thresholds are rarely characterized by a single physical dimension. Instead, they are more commonly defined as statistical boundaries derived from electrical signal measurements, and abnormalities are typically identified using the 3σ rule.

From the point of view of faults, as a health state deteriorates, the frequency, duration, and intensity of IFs increase under identical stress conditions.

After determining the characteristic parameters of faults, the next step is to correlate these parameters with the operational state of the equipment. However, the complex structure of railway equipment results in relatively weak correlations. The Fuzzy C-Means (FCM) clustering algorithm is well-suited for handling uncertain and stochastic data. It utilizes characteristic data to identify fuzzy relationships among data points and to explore potential data structures without supervision. Consequently, this algorithm is employed to correlate observation data with the operational state of the equipment, thereby generating sample data for each operational state.

The essence of the FCM clustering algorithm lies in first categorizing the observation sequence $\boldsymbol{X}$ into c-type fuzzy subsets, and then computing their cluster centers. This process minimizes the value of the objective function, which can be expressed as follows:

$$J = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m \left\| x_j - c_i \right\|^2 \tag{15}$$

where $u_{ij}$ represents the membership, indicating the degree to which each sample $x_j$ is subordinate to a category $c_i$; $c_i$ represents the center of the $i$-th cluster; $m$ represents the fuzzy factor, which defines the extent of fuzziness in the membership of data points to the cluster center. The value of $c$ is taken as the number of operational states.

### *4.3. Operational State Recognition of Equipment*

The fundamental concept for recognizing the operational state of the equipment based on HSMM is as follows: The equipment experiences a total of $N$ operational states from the initiation of operations to complete failure. For each state, the HSMM parameters generated through training are denoted as $\lambda_1, \lambda_2, \ldots, \lambda_N$. The probability $P(\boldsymbol{X} \mid \lambda_i)$ of the observation sequence $\boldsymbol{X}$ in the model $\lambda_i$ is calculated using the Forward–Backward algorithm. The

state $s_i$ corresponding to the maximum $P(\boldsymbol{X} \mid \lambda_i)$ is then considered the current state of the equipment.

The Gaussian Mixture Model (GMM) is employed to fit the observation probability density function for various operational states, utilizing its capability to approximate a finitely continuous probability density function with infinite precision. Furthermore, GMM effectively handles parameter training for HSMM, enhancing the robustness of the model parameters.

To effectively reduce computational complexity during model training and state recognition, the observation sequences $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ are subjected to piecewise integration over every $s$ consecutive observations. The original observation sequence $\boldsymbol{X}$ can be represented as $\boldsymbol{X'}$, which can be expressed as follows:

$$X' = \{x_1, x_2, \ldots, x_t, \ldots, x_{T'}\} \tag{16}$$

where

$$T' = \begin{cases} T/s, & T\%s = 0 \\ [T/s] + 1, & T\%s \neq 0 \end{cases} \tag{17}$$

$$x_t = \sum_{k=(t-1)s+1}^{\min(ts,T)} o_k / s \tag{18}$$

where $x_t$ represents the observation after integration, with "%" indicating the remainder and "[ ]" indicating the rounded number.

The model training data consists solely of observation sequences. In this context, HSMM learning is classified as an unsupervised task, and the Baum–Welch algorithm is typically employed. The essence of this algorithm lies in updating the model parameter $\lambda$, to maximize the conditional probability $P(\boldsymbol{X} \mid \lambda)$ under the observation sequence $\boldsymbol{X}$.

Therefore, the log-likelihood function $L(\lambda)$ is constructed to realize the maximization.

$$L(\lambda) = \log P(\boldsymbol{X}|\lambda) = \log \sum_{\boldsymbol{Y}} P(\boldsymbol{X}, \boldsymbol{Y}|\lambda) \tag{19}$$

(1) Define the forward probability [8] $\alpha_t(i)$, which represents the probability that at time $t$, the hidden state $y_t = s_i$ and the observation sequence $\boldsymbol{X}$ is $x_1, x_2, \ldots, x_t$ expressed as follows:

$$\alpha_t(i) = P(x_1, x_2, \ldots, x_t, y_t = s_i|\lambda) \tag{20}$$

First, at the initialization time, the first state segment starts at time 1, lasts for $d$ time steps, and ends at time $d$. Thus:

$$\alpha_d(i) = \pi_i p_i(d) b_i(x_{1:d}), d = 1, \ldots, D_i \tag{21}$$

Assume that the previous state $s_i$ ends at time $t - d$ and then transitions to state $s_j$, where $s_j$ lasts for $d$ time steps and ends at time $t$. Recursion gives:

$$\alpha_t(i) = \sum_{j=1}^{N} \sum_{d=1}^{D_i} \alpha_{t-d}(j) a_{ji} p_i(d) \prod_{s=t-d+1}^{t} b_i(x_s) \tag{22}$$

where $t - d \geq 1$, $i \neq j$, $d \leq D_i$.

The entire observation sequence ends at time $T$, and the last state ends at time $T$. The conditional probability $P(\boldsymbol{X} \mid \lambda)$ is calculated as:

$$P(\boldsymbol{X}|\lambda) = \sum_{i=1}^{N} P(x_1, x_2, \ldots, x_T, y_T = s_i|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{23}$$

(2) Define the backward probability $\beta_t(i)$, which represents the probability that at time $t$, the hidden state $y_t = s_i$, and the observation sequence $\boldsymbol{X}$ from $t + 1$ to $T$ is $x_{t+1}$, $x_{t+2}$, …, $x_T$, expressed as follows:

$$\beta_t(i) = P(x_{t+1}, x_{t+2}, \ldots, x_T|y_t = s_i, \lambda) \tag{24}$$

First, initialize the backward probability $\beta_T(i)$ of each hidden state at time $T$ as follows:

$$\beta_T(i) = 1,\ 1 \le i \le N \tag{25}$$

Similarly, recursion gives:

$$\beta_t(i) = \sum_{j=1}^{N} \sum_{d=1}^{D_j} a_{ij} p_j(d) \prod_{s=t+1}^{t+d} b_j(x_s) \beta_{t+d}(j) \tag{26}$$

where $t + d \le T$, $i \ne j$, $d \le D_j$.

Then, calculate the conditional probability $P(\boldsymbol{X} \mid \lambda)$ as follows:

$$P(\boldsymbol{X}|\lambda) = \sum_{j=1}^{N} \pi_j \sum_{d=1}^{D_j} p_j(d) \prod_{s=1}^{d} b_j(x_s) \beta_d(j) \tag{27}$$

By combining Equations (23) and (27), the probability $P(\boldsymbol{X} \mid \lambda)$ of the observation sequence $\boldsymbol{X}$ is expressed as a combination of the forward probability and the backward probability:

$$P(\boldsymbol{X}|\lambda) = \sum_{i=1}^{N} \alpha_t(i) \beta_t(i) \tag{28}$$

At this point, solve for the model $\lambda$ at the maximum $P(\boldsymbol{X} \mid \lambda)$, with the following boundary conditions:

$$\alpha_t(i) = \begin{cases} \pi_i, & t = 1 \\ 0, & t < 1 \end{cases} \tag{29}$$

$$\beta_t(i) = \begin{cases} 1, & t = T \\ 0, & t > T \end{cases} \tag{30}$$

During the solution process, define intermediate variables:

$$\gamma_t(i) = P(y_t = s_i|\boldsymbol{X}, \lambda) = \frac{P(y_t = s_i, \boldsymbol{X}|\lambda)}{P(\boldsymbol{X}|\lambda)} = \frac{1}{P(\boldsymbol{X}|\lambda)} \alpha_t(i) \beta_t(i) \tag{31}$$

$$\begin{aligned} \xi_t(i, j, d) &= P\left(y_{t-d} = s_i, y_t = s_j \middle| \boldsymbol{X}, \lambda\right) = \frac{P\left(y_{t-d} = s_i, y_t = s_j, \boldsymbol{X}|\lambda\right)}{P(\boldsymbol{X}|\lambda)} \\ &= \frac{\alpha_{t-d}(i) a_{ij} p_j(d) b_j(x_{t-d+1:t}) \beta_t(j)}{P(\boldsymbol{X}|\lambda)} \end{aligned} \tag{32}$$

where $\gamma_t(i)$ represents the probability of being in the state $s_i$ at time $t$ for a given $\lambda$. $\xi_t(i,j,d)$ denotes the probability that the state is $s_i$ at time $t - d$, $s_j$ at time $t$, and $s_j$ lasts for $d$ time steps.

Maximizing the $L(\lambda)$ function yields the expressions for $\boldsymbol{\pi}$, $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{D}$:

The initial state probability $\pi_i$ actually represents the posterior probability that the first segment belongs to state $s_i$, and its value is the sum of $\gamma_d(i)$ over $d$, i.e.,

$$\overline{\pi}_i = \sum_{d=1}^{D_i} \gamma_d^{\text{init}}(i) \tag{33}$$

where $\overline{\pi}_i$ is the estimated value of $\pi_i$, and $\gamma_d^{\text{init}}(i)$ is calculated as follows:

$$\gamma_d^{\text{init}}(i) = \frac{\pi_i p_i(d) b_i(x_{1:d}) \beta_d(i)}{P(\boldsymbol{X}|\lambda)} \tag{34}$$

The estimated value of the state transition probability $a_{ij}$ is calculated by the following equation:

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T} \sum_{d=1}^{D_j} \xi_t(i,j,d)}{\sum_{j'=1}^{N} \sum_{t=1}^{T} \sum_{d=1}^{D_{j'}} \xi_t(i,j',d)} \tag{35}$$

The key to calculating the observation probability $b_i(x_j)$ is to obtain the posterior probability that the state $s_i$ is occupied at time $u$. The probability is denoted as $\eta_u(i)$:

$$\eta_u(i) = \sum_{t=u}^{T} \sum_{d=t-u+1}^{D_i} \gamma_t(i,d) \tag{36}$$

where $\gamma_t(i,d)$ is defined as the posterior probability that, given the model parameters $\lambda$, the system is in state $s_i$ at time $t-d$, remains in this state for $d$ consecutive time steps, and the segment terminates at time $t$. From this definition, $\gamma_t(i,d)$ is a conditional probability, whose Bayesian form is given by:

$$\gamma_t(i,d) = \frac{P_j}{P(\boldsymbol{X}|\lambda)} \tag{37}$$

where $P_j$ denotes the joint probability, which consists of three components:

(1) The probability from time 1 to time $t-d$, i.e., the forward probability $\alpha_{t-d}(i)$, computed recursively via Equation (22).

(2) The probability from time $t-d+1$ to time $t$, i.e., the product of the observation probability $b_i(x_{t-d+1:t})$ of the current state and the probability $p_i(d)$ that the state lasts for $d$ time steps.

(3) The probability from time $t+1$ to time $T$, i.e., the backward probability $\beta_t(i)$.

Under the Markov property, these three observation segments are conditionally independent. Therefore, $P_j$ can be expressed as the product of the three components:

$$p_j = \alpha_{t-d}(i) p_i(d) b_i(x_{t-d+1:t}) \beta_t(i) \tag{38}$$

Combining Equations (37) and (38), $\gamma_t(i,d)$ can be expressed as follows:

$$\gamma_t(i,d) = \frac{\alpha_{t-d}(i) p_i(d) b_i(x_{t-d+1:t}) \beta_t(i)}{P(\boldsymbol{X}|\lambda)} \tag{39}$$

Now, the estimated value of the observation probability $b_i(x_j)$ is calculated by the following equation:

$$\overline{b}_i(x_j) = \frac{\sum_{u=1}^{T} \delta(x_t, v_j) \eta_u(i)}{\sum_{u=1}^{T} \eta_u(i)} \tag{40}$$

where

$$\delta(x_t, v_j) = \begin{cases} 1, & x_t = v_j \\ 0, & x_t \neq v_j \end{cases} \tag{41}$$

The probability distribution of state duration $\boldsymbol{D} = [p_i(d)]$ is a ratio, where the numerator of the ratio is the expected number of times the equipment stays in hidden state $s_i$ for duration $d$, while the denominator is the expected total number of times the equipment is in hidden state $s_i$. Define $\overline{p}_i(d)$ as the re-estimated value of probability $p_i(d)$:

$$\overline{p}_i(d) = \frac{\sum_{t=d}^{T} \gamma_t(i,d)}{\sum_{d'=1}^{D_i} \sum_{t=d'}^{T} \gamma_t(i,d')} \tag{42}$$

In this way, after calculating $\pi_i$, $a_{ij}$, $b_i(x_j)$, $p_i(d)$, we obtain the model $\lambda = [\boldsymbol{\pi}, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}]$.

After segment-wise integration of the observation sequence, the resulting segments are fed into the HSMMs corresponding to each state. The maximum likelihood probability is then computed for each state, and the state associated with the model yielding the highest likelihood is taken as the current state of the equipment. In practical applications, the accuracy of HSMM relies on its parameters. Hence, estimating the optimal model parameters is considered a critical issue that must be addressed. The Baum–Welch algorithm, employed in the process of maximizing the $L$ function, is essentially an Expectation Maximal (EM) algorithm; however, it has the drawback of converging to local optima.

By simulating the survival behaviors of colonizing species in nature, swarm intelligent optimization algorithms (SIOA) [43] accomplish optimization tasks through swarm evolution, which is grounded in communication and cooperation among individuals. Initially, a set number of feasible solutions are randomly generated based on the established constraints. Subsequently, each stage of the algorithm is employed to iteratively search for the global optimal solution and identify the optimal hyperparameter within a specified range. With enhanced robustness and global search capabilities, these algorithms are well-suited for tackling high-dimensional, nonlinear, and complex optimization problems.

Several classical SIOAs are incorporated to develop the best HSMM by estimating HSMM parameters, including the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Whale Optimization Algorithm (WOA).

### *4.4. RUL Prediction*

The equipment is generally still in service when its RUL is predicted, making it impossible to obtain observations throughout the full life cycle of the equipment, from its initial production to the end of its life. Therefore, historical data from the same type of equipment are typically utilized to establish a model, with an observation sequence prior to a given time inputted to predict the hidden state. Subsequently, the RUL is estimated based on the time distribution of service age. The essence of prediction involves dividing the full life cycle of electronic equipment into several operational states, identifying these states, and predicting the RUL based on the current state. This approach results in RUL equivalence for the equipment in identical states across different timepoints. As illustrated in Figure 3, the equipment state is $s_i$ at timepoints $t_a$ and $t_b$. However, it is evident that these two timepoints should correspond to different RULs, revealing a discrepancy between the prediction concept and the actual situation.

The key of addressing this problem lies in identifying the transition point into degradation states, such as the timepoint $t_c$ shown in Figure 3, which marks the transition from the current state $s_i$ to the next degraded state $s_{i+1}$. First, in the temporal dimension, several time intervals $\{k_1, k_2, \ldots, k_i, \ldots, k_n\}$ are defined, letting $k_i = k_{i+1,}$ which indicates these identical intervals represent the minimum unit time, such as 1 month. In this way, the duration $D(s_i)$

of any state $s_i$ must be the sum of several intervals $k_i$, and similarly, the state transition point $t_c$ is also the sum of several intervals $k_i$. Then, based on the LSTM method [44], the observation sequence for the next timepoint is predicted with a step size taken as the unit time $k_i$. Finally, state recognition and RUL prediction are conducted using the HSMM. The specific steps are as follows:

Step 1: Use LSTM to make a single-step prediction of observation sequences.

Step 2: Input the prediction result into HSMM to identify the degraded state at the next timepoint.

Step 3: Determine the timepoint of state transition between different adjacent states.

Step 4: Predict the RUL of the equipment.

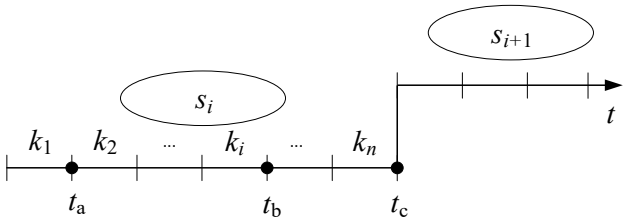



**Figure 3.** Diagram of transition between adjacent states.

(1) Prediction of observation sequence

Use the LSTM method, and take the observed values $\{x_1, x_2, \ldots, x_{t-1}\}$ before time $t - 1$ as input to predict the observed value $x_t$ at the next moment. LSTM is a special RNN network structure that introduces a new internal state and a gating mechanism within the hidden layer to regulate the information transfer path, effectively addressing gradient explosion and vanishing that occur in traditional RNNs. The internal state $\boldsymbol{c}_t$ and output $\boldsymbol{h}_t$ are expressed as follows:

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \widetilde{\boldsymbol{c}}_t \tag{43}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh(\boldsymbol{c}_t) \tag{44}$$

where $\boldsymbol{f}_t$ represents the forget gate, $\boldsymbol{i}_t$ represents the input gate, $\boldsymbol{o}_t$ represents the output gate, $\odot$ represents the element-wise product of vectors, and $\boldsymbol{c}_{t-1}$ is the memory unit at the previous moment, $\widetilde{\boldsymbol{c}}_t$ is the candidate state obtained via the tanh activation function:

$$\widetilde{\boldsymbol{c}}_t = \tanh(\boldsymbol{W}_c \boldsymbol{x}_t + \boldsymbol{U}_c \boldsymbol{h}_{t-1} + \boldsymbol{b}_c) \tag{45}$$

In the gating mechanism, $\boldsymbol{f}_t$, $\boldsymbol{i}_t$, $\boldsymbol{o}_t$ are calculated as follows:

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f \boldsymbol{x}_t + \boldsymbol{U}_f \boldsymbol{h}_{t-1} + \boldsymbol{b}_f\right) \tag{46}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \boldsymbol{x}_t + \boldsymbol{U}_i \boldsymbol{h}_{t-1} + \boldsymbol{b}_i) \tag{47}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \boldsymbol{x}_t + \boldsymbol{U}_o \boldsymbol{h}_{t-1} + \boldsymbol{b}_o) \tag{48}$$

where $\sigma$ represents the Logistic function, with an output interval of (0, 1). The symbols $W$, $U$, and $b$ denote the learnable network parameters.

(2) RUL prediction

Among them, the RUL prediction method is detailed as follows:

The durations $D(s_i)$ of all operational states throughout the full life cycle constitute the useful life of the equipment. The probability distribution function $P(d_n/s_i)$ for durations among all operational states follows a single Gaussian distribution. Under the constraint

$T = \sum_{i=1}^{N} D(s_i)$, the duration $D(s_i)$ of each state can be expressed as follows by maximizing $\sum_{i=1}^{N} \log P(d_n/s_i)$:

$$D(s_i) = \mu(s_i) + \rho\sigma^2(s_i) \tag{49}$$

where $\mu$ and $\sigma^2$ represent the mean and variance of $D(s_i)$, respectively, and $\rho$ can be expressed as follows:

$$\rho = \left(T - \sum_{i=1}^{N} \mu(s_i)\right) / \sum_{i=1}^{N} \sigma^2(s_i) \tag{50}$$

Use a recursive method to calculate the $RUL_i$ of the equipment in the state $s_i$.

(1) In the state $s_{i-1}$:

$$RUL_{i-1} = a_{(i-1,i-1)}[D(s_{i-1}) + D(s_i)] + a_{(i-1,i)}D(s_i) \tag{51}$$

(2) In the state $s_{i-k}$ ($k > 1$):

$$RUL_{i-k} = a_{(i-k,i-k)}[D(s_{i-k}) + RUL_{i-k+1}] + a_{(i-k,i-k+1)}RUL_{i-k+1} \tag{52}$$

(3) In the state $s_i$:

$$RUL_i = a_{ii}[D(s_i) + RUL_{i+1}] + a_{(i,i+1)}RUL_{i+1} \tag{53}$$

Combining Equations (49) and (53), the relationship between the equipment state observation sequence and the RUL is determined.

The RUL of a state $s_i$ at a given time, denoted as $RUL_{(i,k)}$, consists of two parts: the remaining duration of the current state and the sum of the durations of all subsequent states, as shown in Figure 3. Accordingly, a more concise formulation has been provided.

$$RUL_{(i,k)} = \sum_{l=i+1}^{n} k_l + \sum_{l=i+1}^{N} D(s_l) \tag{54}$$

where $D(s_l)$ is calculated by Equation (49).

As indicated by Equation (54), the equipment has a unique predicted RUL value at each timepoint in any state before permanent failure. This approach aligns with the actual railway situation by solving the abovementioned problem where the same RUL is yielded for the equipment under identical states at different timepoints.

## 5. Experiment and Discussion

This section begins by presenting a case study involving real-world IFs in electronic equipment used for on-site railway operations. This example provides an experimental scenario to validate the state recognition and RUL prediction processes, so as to evaluate the effectiveness of the proposed health management approach for electronic equipment.

### *5.1. Case Study of IFs*

IFs in railway electronic equipment often occur, disrupting the normal operation of trains. Only when failures are identified as IFs can the HSMM method be applied to evaluate the equipment state and predict the RUL, thereby providing support for maintenance decisions. The following section presents an IF case of electronic equipment used in railway applications, establishing a benchmark for subsequent experiments.

(1) Fault introduction

At 13:00:51 on 15 July 2023, the driver machine interface (DMI) monitor on a running train displayed a blank screen, which was accompanied by an interruption of communication between the DMI and the ATP machine. Consequently, the ATP initiated emergency braking to stop the train. The driver attempted to restart the ATP multiple times before normal operation was restored.

(2) Cause analysis

Based on the ATP data and DMI log records, the fault is caused by abnormal operation of the DMI equipment, resulting in a blank display and a communication interruption of the DMI.

According to the internal processing logic of the DMI monitor, when communication between the DMI and the ATP machine is interrupted, the DMI should turn off the backlight, activate the fault indicator, and log the corresponding event. However, despite a communication interruption record during this incident, the monitor fails to turn off the backlight or activate the fault indicator, and there are no abnormal records in the DMI log. The blank display at the time of the fault indicates that normal 12 V voltage was present within the DMI, along with the subsequent normal 5 V voltage converted by the U18 chip. The failure to turn off the backlight and activate the fault indicator suggests an abnormality in the 24 V-to-5 V power module on the DMI's bottom board, leading to a failure of the CPU board and ultimately resulting in the inability to generate a log entry.

To sum up, the root cause of this fault has been identified as an abnormal 24 V-to-5 V power module within the DMI.

(3) Fault testing

(1) Upon return of the faulty DMI to the laboratory, an inspection confirms that it is normal in appearance, showing no abnormal symptoms such as dents or bumps. Additionally, the internal connectors remain securely connected, with no signs of looseness.

(2) During the power-on verification of the faulty DMI, the monitor boots up normally, showing no abnormal symptoms, such as a blank display.

(3) The faulty DMI is subjected to burn-in testing at an atmospheric temperature (25 °C) for 12 h, during which the monitor functions normally with no abnormal symptoms, such as a blank display.

(4) The internal 5 V power supply of the monitor is subjected to voltage testing at an atmospheric temperature and is found to be normal during the power-on process.

(5) During a high-temperature test (60 °C) conducted on the faulty DMI for 6 h, the 5 V power supply is subjected to voltage testing. Noticeable voltage fluctuations are observed between 355.4 ms and 355.8 ms during the rising stage of the voltage waveform, as shown in Figure 4.

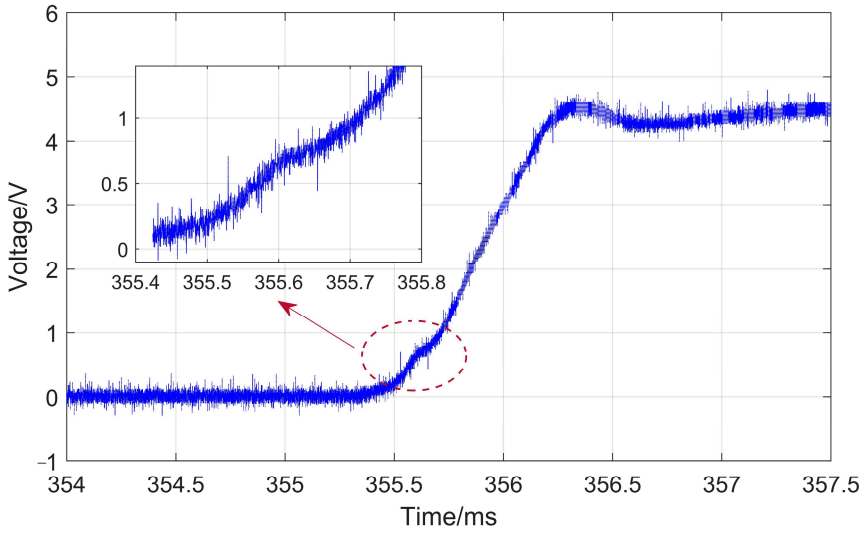



**Figure 4.** 5 V voltage waveform during high-temperature testing.

Based on the laboratory fault testing, the cause of the blank display in the DMI is identified as the accidental failure of the 24 V-to-5 V power module on the internal bottom board of the DMI under high-temperature conditions.

Based on further analysis of the cause and mechanism, it is concluded that under temperature stress, thermal stress due to high and low temperatures affects the connections within the electronic equipment, impacting the localized mechanical, physical, chemical, and electrical properties. Specifically, different conductor materials have different thermal expansion coefficients. During temperature changes, there is a thermal strain difference within the solder joint, leading to mechanical stress. This alternating thermal stress negatively affects the connections between different conductor materials, resulting in cracks within the solder joint and subsequent IFs in terms of electrical performance during the temperature transition.

Based on fault symptoms and characteristics observed during cause analysis, such as an occasional failure, reproduction during high-temperature testing, and normal operation after restart by the driver, it is sufficient to identify the randomness, repeatability, and recoverability exhibited by the electronic equipment under external excitations. This is a typical IF case.

### *5.2. Data Pre-Processing and Model Initialization*

Taking this fault case of the power supply module in a real-world application as an example, its lifecycle operational states are defined as follows: healthy states, minor faults, faulty states, and permanent failure, denoted as $s_1$, $s_2$, $s_3$, and $s_4$ respectively. In railway transport scenarios, any IFs in electronic equipment, regardless of its duration and intensity, will affect the transport order. Therefore, throughout the full life cycle of the equipment, greater emphasis is placed on the frequency of IFs.

In this experiment, 10 power modules from the same batch are selected for a 48-h accelerated degradation test under high-temperature conditions. The sampling interval is 30 min, and the observation object is the number of intermittent faults. Since the entire experiment targets the full life cycle of the power modules and already includes the permanent failure state, all 10 power modules complete the full 48-h experiment, with 96 observation samples obtained for each power module, for a total of 960 observation samples. As shown in Table 2, all datasets are divided by physical module: modules 1–6 (576 observations, 516 windows) are used for training, modules 7–8 (192 observations, 172 windows) for validation, and modules 9–10 (192 observations, 172 windows) for testing. The sliding window has a length of $L = 10$ and slides only within each respective module, with no overlap of any original observations between different sets.

**Table 2.** Dataset division scheme.

| Set | Module | Number of Observations | Number of Windows | Proportion |
|---|---|---|---|---|
| Training set | M1–M6 | 576 | 516 | 60% |
| Validation set | M7–M8 | 192 | 172 | 20% |
| Test set | M9–M10 | 192 | 172 | 20% |
| Total | 10 | 960 | 860 | 100% |

In the experiment, the frequencies of IFs experienced by the equipment over time are shown in Figure 5, where both the horizontal and vertical axes are normalized using the Min–Max method, the frequency and intensity of IFs in the power module gradually increase over time, accompanied by fluctuations to certain extents. The intensity of faults is indicated by frequency. When the intensity reaches a certain level, it is primarily manifested

through extended durations and increased amplitudes, which contribute to the gradual evolution into equipment faults and eventual unserviceability, followed by a decline in IF frequency. Therefore, the number of IFs is used as the primary parameter for assessing the performance state of the power module.

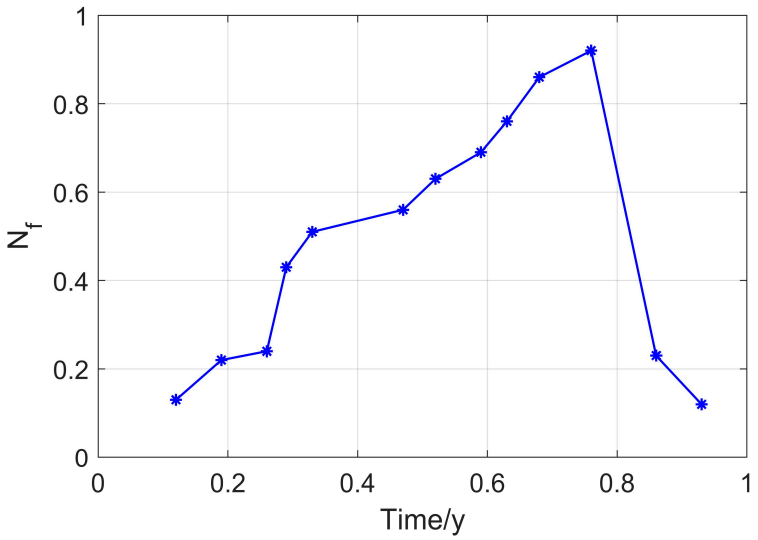



**Figure 5.** Time-varying IF frequencies.

The next work is to define the correspondence between IF frequencies and 4 states of the power modules. The cluster center of subsets is calculated using Equation (15), with *c*-type set to 4. When the clustering frequency *c* is fixed, the value of the fuzzy factor *m* becomes a key parameter influencing the clustering effect. A larger value of *m* results in a fuzzier clustering outcome and a more uniform membership of data points across cluster centers.

The *m*-value is set within the range of [1.5, 2.5] to calculate the silhouette coefficient, a common metric for evaluating clustering effect that combines cohesion and separation within the clusters. The *m*-value is adjusted based on the magnitude of this index to maximize the silhouette coefficient in each experiment. The FCM method is applied repeatedly for 3 to 10 iterations, and the silhouette coefficients from each experiment are presented in Table 3. When the number of experimental runs is 6, the silhouette coefficient reaches its maximum, indicating that this clustering effect is optimal. Consequently, this clustering outcome is selected as the training sample for the observations of each state, defining the correlations between the characteristic parameters of the power module and the four states, as shown in Table 4. In the table, the IF frequencies are normalized using the Min–Max method.

**Table 3.** Silhouette coefficients of FCM clustering.

| Number of Experimental Runs | Silhouette Coefficient | Number of Experimental Runs | Silhouette Coefficient |
|---|---|---|---|
| 3 | 0.3671 | 7 | 0.4265 |
| 4 | 0.5279 | 8 | 0.3698 |
| 5 | 0.6147 | 9 | 0.3244 |
| 6 | 0.7263 | 10 | 0.2986 |

**Table 4.** Correspondence between IF frequencies and individual states.

| State | $s_1$ | $s_2$ | $s_3$ | $s_4$ |
|---|---|---|---|---|
| The number of IFs | [0,0.24] | (0.24,0.51] | (0.51,0.92] | (0.92,1] |

In this experiment, the main technical measures adopted to prevent time data leakage are as follows:

(1) Dataset division level: The training set, validation set, and test set are divided according to physical modules rather than windows or time points, so as to ensure that observations from the same degradation path do not appear in different sets at the same time.

(2) Window construction level: Windows slide only within the modules, and all 10 observations in each window come from the same physical module; there is no overlap of any original observations between different sets.

(3) Clustering level: The FCM cluster centers are fitted only with the training set. The thresholds of the characteristic parameters corresponding to each state are determined by the training set, and then the states of the validation/test observations are classified.

(4) Feature engineering level: The normalization parameters (mean, standard deviation) are fitted only with the training set, and only the transform is applied to the validation set and the test set.

(5) Model training and evaluation level: The validation set is used for parameter tuning, and the test set is used only for final evaluation.

Under these principles and constraints, the distribution of data samples for each state is shown in Table 5.

**Table 5.** Distribution of data samples in each state.

| Set | $s_1$ | $s_2$ | $s_3$ | $s_4$ | Total |
|---|---|---|---|---|---|
| Training set | 219 | 132 | 118 | 107 | 576 |
| Validation set | 73 | 44 | 39 | 36 | 192 |
| Test set | 73 | 44 | 39 | 36 | 192 |
| Total | 365 | 220 | 196 | 179 | 960 |

The power module is in a healthy state as its initial state, which is expressed as follows:

$$\boldsymbol{\pi} = [1\ 0\ 0\ 0] \tag{55}$$

The state transition matrix $\boldsymbol{A}$ is initialized, allowing transitions only from the current state to either the current or the next state, with their transition probabilities assumed to be equal, expressed as follows:

$$\boldsymbol{A} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{56}$$

GMM parameters used to represent the state observation probability matrix $\boldsymbol{B}$ include the number of Gaussian components $G$, mean value $\mu_{jg}$, and standard deviation $\sigma_{jg}$; and the Gaussian distribution parameters used to represent the probability distribution of state durations $p_i(d)$ include mean value $\mu_{id}$ and standard deviation $\sigma_{id}$. The K-means cluster-

ing algorithm is employed to initialize these parameters. The number for the piecewise integration of observation sequences is set to 3.

### *5.3. Experimental Results*

(1) Model Training

The degradation data of the equipment throughout its full life cycle is used to train the HSMM. Following the piecewise integration of observation sequences, the HSMM is trained for each state. The full-life-cycle model is intended to calculate the state durations of the equipment and the transition probabilities between states. The models for various states aim to recognize operational states at various timepoints.

With the maximum number of iterations set at 20 and the convergence error at 0.001, these models are provided with observation sequences correlated with various states along with the initialized model parameters λ as inputs. The Baum–Welch algorithm is used to estimate the HSMM parameters using Equations (33), (35), (40) and (42), yielding the $i$-th parameter update $\overline{\lambda}_i$. The Forward–Backward algorithm is used to calculate the likelihood probability $L(\overline{\lambda}_i)$ of the model with $\overline{\lambda}_i$ using Equation (19). The iteration continues if it is indicated that the Baum–Welch algorithm has not yet converged or reached the maximum number of iterations; otherwise, the final HSMM parameters $\overline{\lambda}$ are generated by suspending the algorithm.

Model training is conducted with observation sequences correlated with various states, and the HSMM training curves for the four states are illustrated in Figure 6. As the number of iterations increases, the log-likelihood values produced for each state also gradually rise, indicating an improvement in the accuracy of state recognition. The tendency to converge after 10 iterations demonstrates the rapid learning capability of the HSMM, enhanced by the incorporation of a swarm intelligence optimization algorithm.

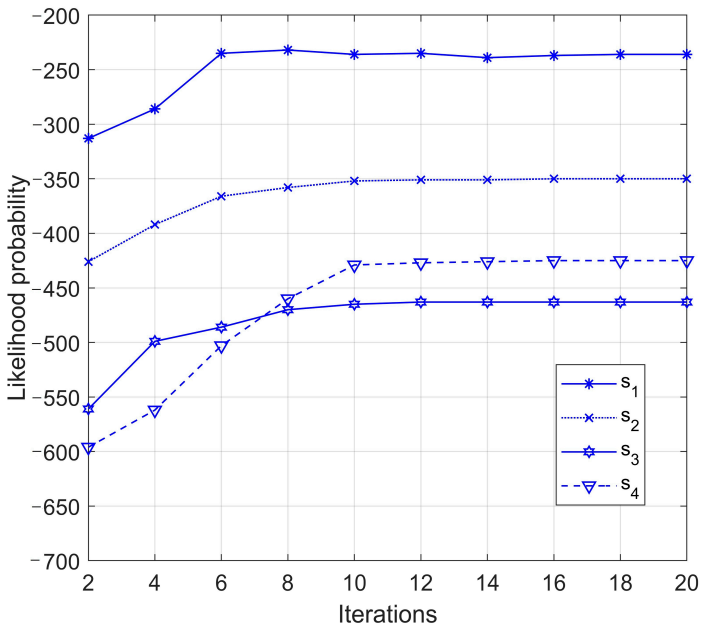



**Figure 6.** HSMM training curves in four states.

(2) Analysis of State Recognition Results

The trained HSMM is utilized to recognize the operational states of the power module. In Table 5, the minimum size of test samples representing various states is 36 groups. Therefore, after segmented integration, 36 groups of test samples representing various states are input into the HSMM corresponding to respective states. The Forward–Backward algorithm is then employed to calculate the maximum likelihood probability. The state

associated with the highest probability value is regarded as the current state of the power supply module. The state recognition results in the test set are shown in Table 6.

**Table 6.** Operational state recognition results based on HSMM.

| State | $s_1$ | $s_2$ | $s_3$ | $s_4$ | Accuracy (%) |
|---|---|---|---|---|---|
| $s_1$ | 35 | 1 | 0 | 0 | 97.22 |
| $s_2$ | 0 | 34 | 1 | 1 | 94.44 |
| $s_3$ | 0 | 0 | 35 | 1 | 97.22 |
| $s_4$ | 0 | 1 | 0 | 35 | 97.22 |

In Table 3, the traditional Baum–Welch algorithm is used to estimate HSMM parameters. For state $s_1$, 1 out of 36 samples is incorrectly identified as $s_2$, and the rest are correctly identified, resulting in a recognition accuracy of 97.22%. Similarly, the recognition accuracies for states $s_2$, $s_3$, and $s_4$ are 94.44%, 97.22%, and 97.22%, respectively. The average recognition accuracy across the four states is 96.53%. Subsequently, GA, PSO, and WOA algorithms are employed to maximize the likelihood function and re-estimate the HSMM parameters, thereby creating the optimal HSMM. The experimental results on the same test set are displayed in Table 7.

**Table 7.** State recognition accuracy by different optimization algorithms (%).

| Methods | $s_1$ | $s_2$ | $s_3$ | $s_4$ | Average |
|---|---|---|---|---|---|
| Conventional HSMM | 97.22 | 94.44 | 97.22 | 97.22 | 96.53 |
| GA-based | 97.22 | 97.22 | 100.00 | 94.44 | 97.22 |
| PSO-based | 100.00 | 97.22 | 100.00 | 97.22 | 98.61 |
| WOA-based | 100.00 | 97.22 | 100.00 | 94.44 | 97.92 |

As shown in Table 7, the average state recognition rates calculated by swarm intelligence optimization algorithms are 97.22%, 98.61%, and 97.92%, respectively, all of which exceed the traditional HSMM method's rate of 96.53%, up 0.71%, 2.15% and 1.44% respectively. Notably, the recognition rate for state $s_3$ reaches 100%. These differences indicate that incorporating swarm intelligence optimization algorithms is effective in estimating HSMM parameters and improves the accuracy of state recognition compared to the traditional Baum–Welch algorithm. Furthermore, the PSO algorithm shows the most significant improvement in recognition accuracy and is more suitable for the state recognition of power modules.

(3) Analysis of RUL Prediction Results

Taking into account the differences in HSMM classifiers trained under various states of the equipment, a full-life-cycle HSMM classifier is developed based on the HSMMs corresponding to different states, trained using the complete data from all states. Additionally, the optimal PSO algorithm is employed to estimate the HSMM parameters, and the state durations at different timepoints are calculated.

After recognizing the current state of the power module, the observation sequence predicted by the LSTM is utilized to identify the degraded state of the equipment at the next timepoint. If the state transition point has not yet been reached, observation prediction and state recognition continue. Meanwhile, the RUL is calculated using Equation (54).

To obtain RUL values in identical states across different timepoints, it is necessary to predict the observation sequence of the equipment at the next timepoint. This essentially involves a time series single-step prediction method. The tf.keras sequential model used for sequence modeling in TensorFlow is utilized to construct the LSTM prediction model.

First, an LSTM layer is added by defining the number of units and the input shape. Then, a fully connected layer is included to output the prediction results.

During LSTM training, the initial model parameters are set as shown in Table 8. The mini-batch gradient descent method is employed to optimize the learning process by identifying parameters that minimize structural risk. A mean squared error loss function is defined to measure the differences between the predicted values and the actual values. Additionally, the training loss curve is plotted, as illustrated in Figure 7, to reflect the convergence of the model during training.

**Table 8.** Parameters of observation sequence prediction model.

| Parameter Meaning | Parameter Expression | Parameter Value |
|---|---|---|
| Sliding window size | window_size | 6 |
| Epoch | epoch | 200 |
| Batch size | batch_size | 6 |
| Basic learning rate | base_lr | 0.002 |
| Optimizer | optimizer | Adam |

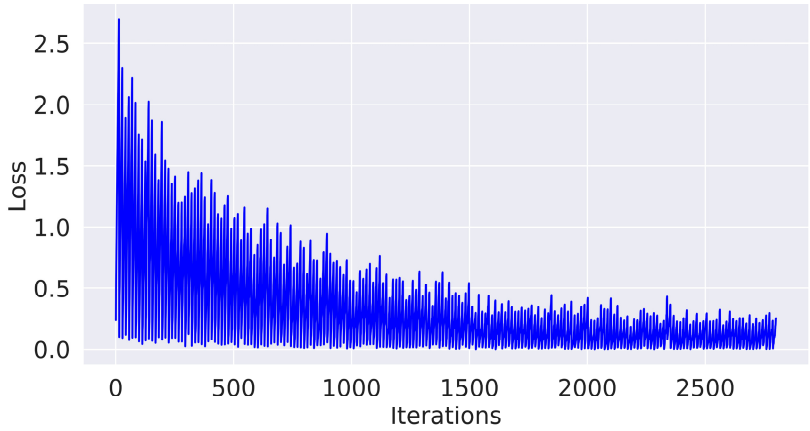



**Figure 7.** Training loss curve of observation sequence prediction model.

As shown in Figure 7, the loss value decreases with an increasing number of iterations, ultimately reaching stable convergence. This trend indicates that the pattern information contained in the observation sequences is accurately perceived. The mean squared error and coefficient of determination of the LSTM prediction model on the test set are recorded as 0.014 and 0.946, respectively. These results suggest that future observation sequences are well predicted, demonstrating the effectiveness and feasibility of the proposed LSTM-based method for predicting observation sequences in power modules.

Based on the generated state durations and the predicted future observation sequences, the RUL values for the power module are calculated using Equation (54).

### *5.4. Discussion and Analysis*

This section mainly discusses two issues. Firstly, from the perspective of computational complexity, it discusses the advantages of swarm intelligence optimization algorithms over traditional EM methods in HSMM parameter estimation. Secondly, from the perspective of statistical testing, it discusses the robustness of the proposed health management method.

#### 5.4.1. Computational Complexity Analysis

In order to compare the computational complexity of the EM, GA, PSO, and WOA methods for estimating HSMM parameters, general complexity variables are first defined as shown in Table 9.

**Table 9.** Common variables in HSMM parameter estimation.

| Symbol | Meaning | Typical Value |
|---|---|---|
| $L$ | Length of observation sequence | $10^3$~$10^5$ |
| $S$ | Number of hidden states | 3~20 |
| $D$ | Maximum state duration | 20~100 |
| $V$ | Observation dimension | 10~$10^3$ |
| $K$ | Number of EM algorithm iterations | 50~200 |
| $R$ | Number of EM random initializations | 20~50 |
| $N_{iter}$ | Number of nested numerical optimization iterations in EM | 30~200 |
| $M_s$ | Population size of swarm intelligence algorithms | 30~100 |
| $G$ | Number of evolution generations of swarm intelligence algorithm | 200~500 |

EM is an analytical iteration method, while the latter three are swarm intelligence (meta-heuristic) random search methods. This mechanism leads to a distinct difference in computational complexity between the two categories of methods. Regardless of the algorithm used, the log-likelihood function of HSMM must be calculated. The complexity of the forward algorithm (including duration distribution) for HSMM is calculated by the following equation:

$$O_L = O(L \cdot S \cdot D \cdot (S + D)) \tag{57}$$

Each iteration of the EM algorithm consists of two core steps: the expectation step (E-step) and the maximization step (M-step). For the E-step, the main mathematical operations include calculating forward–backward probability, state occupancy posterior probability and transition count expectation, with a computational complexity of $O_L$; for the M-step, the main mathematical operations include updating transition probability $a_{ij}$, observation probability $b_i(x_j)$ and duration distribution $p_i(d)$ respectively, with a computational complexity of $O_M = O(N_{\text{iter}} \cdot L \cdot S \cdot D)$. Considering multiple iterations and multiple random initializations, the computational complexity of the EM algorithm in HSMM parameter estimation can be accurately expressed as follows:

$$O_{EM} = O(K \cdot R \cdot N_{iter} \cdot L \cdot S \cdot D \cdot (S + D)) \tag{58}$$

Taking the PSO algorithm as an example, in each iteration:

(1) For $M_s$ particles, calculate the likelihood value once for each particle, with a computational complexity of $M_s \cdot O_L$.

(2) Update the velocity and position of each particle, with a computational complexity of $O(M_s \cdot P)$, where $P = S^2 + S \cdot V + S \cdot D$.

(3) Update the global optimum and individual optimum, with a computational complexity of $O(M_s)$.

Since $P \ll L \cdot S \cdot D \cdot (S + D)$, the computational complexity of item (2) is negligible. In this case, the computational complexity of HSMM parameter estimation based on the PSO algorithm is calculated as follows:

$$O_{PSO} = O(G \cdot M_s \cdot L \cdot S \cdot D \cdot (S + D)) \tag{59}$$

Similarly, the computational complexity of GA and WOA is comparable to that of the PSO algorithm, namely:

$$O_{GA} = O_{WOA} = O_{PSO} \tag{60}$$

In swarm intelligence algorithms, in each generation, the calculation of the likelihood value for each particle/whale/individual depends only on its own position vector and

has nothing to do with other individuals. Therefore, there is an inherent high degree of parallelism. However, the EM algorithm has strict temporal data dependencies and cannot be effectively parallelized. This is the fundamental reason why the EM algorithm faces computational bottlenecks in scenarios involving long sequences (large $L$) and high-dimensional states (large $S$, $D$).

Under parallel conditions, the number of evolution generations of swarm intelligence algorithm ($G$) usually satisfies the following condition:

$$G < \frac{K \cdot R \cdot N_{iter}}{M_s} \tag{61}$$

Therefore, by combining Equations (58)–(61), the calculation result of Equation (59) is always smaller than that of Equation (58), i.e., $O_{GA} = O_{PSO} = O_{WOA} < O_{EM}$, which indicates that the effective time of GA/PSO/WOA is significantly shorter than that of the EM method. This fundamentally demonstrates the necessity of selecting swarm intelligence algorithms to optimize HSMM parameters.

#### 5.4.2. Method Robustness Analysis

To verify the stability or robustness of the proposed method—that is, to ensure that the RUL prediction results are independent of specific data or degradation stages—a resampling-based experimental approach is designed using the rolling prediction technique. The core ideas are "rolling-window resampling" and "distribution test of prediction errors". The specific steps are as follows:

(1) Data resampling. Over the entire life cycle of the equipment [1, $T$], select $m = 30$ different prediction start points at a fixed step size of $\Delta = 2$ time units, i.e., $t_s \in \{t_1, t_2, \ldots, t_m\}$.

(2) Generation of error sequences. Perform RUL prediction at each time point $t_s$, to obtain $RUL_i(t_s)$, calculate the prediction error $e_i = RUL_i(t_s) - RUL_t(t_s)$, and obtain an error sequence $\{e_1, e_2, \ldots, e_m\}$. Here, $RUL_t(t_s)$ denotes the actual RUL value at time $t_s$.

(3) Statistical inference.

First, evaluate any systematic bias induced by the prediction method using a one-sample $t$-test.

Then, compare the homogeneity of error variances across the early, middle, and late periods of the equipment life cycle using Levene's test.

Finally, examine the correlation between the error $e_i$ and the start point $t_s$ using Spearman's rank correlation test.

Through resampling, the calculated RUL values for the power module are shown in Table 10. The error sequence listed in this table shows: a mean of 0.024, a sample standard deviation of 0.340, a MAPE of 0.826%, and an RMSE of 0.335. From the perspective of business requirements, these errors are acceptable, indicating the effectiveness and feasibility of using HSMM to predict the RUL of power modules. The prediction error sequence and autocorrelation function are represented in Figure 8.

**Table 10.** RUL prediction results based on HSMM.

| Sampling Point | $t_1$ | $t_2$ | $t_3$ | $t_4$ | ... | $t_{30}$ |
|---|---|---|---|---|---|---|
| $RUL_i$ | 72.45 | 70.38 | 68.25 | 66.43 | ... | 14.27 |
| $RUL_t$ | 72.82 | 70.19 | 68.65 | 65.97 | ... | 14.36 |
| $e_i$ | −0.37 | 0.19 | −0.40 | 0.46 | ... | −0.09 |

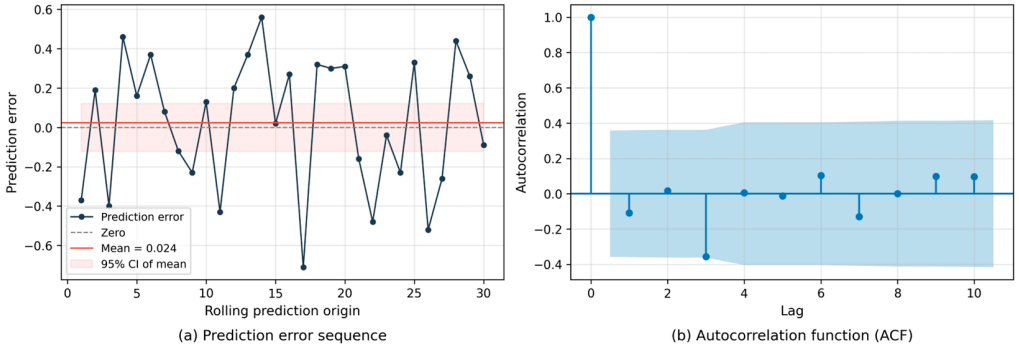



**Figure 8.** Prediction error sequence and autocorrelation function.

Figure 8a shows the error sequence of 30 rolling prediction points. The dotted line denotes the zero baseline, the solid line indicates a mean of 0.024, and the shaded area represents the 95% confidence interval for the mean. The error fluctuates randomly around zero, showing no evidence of systematic bias. Figure 8b presents the autocorrelation function of the error sequence, with the blue shaded area denoting the 95% confidence band. All autocorrelation coefficients for lags up to 10 fall within the confidence band, exhibiting neither slow decay nor periodic peaks, which indicates that no significant autocorrelation is present.

Table 10 shows the results obtained after optimizing HSMM parameter estimation using the PSO algorithm, with the confidence interval set at 95%. Since PSO has already been identified as the most suitable optimization algorithm for state recognition in Section 5.3, this comparison is intended to show that the PSO-optimized HSMM achieves better RUL prediction performance than the conventional HSMM. The results are shown in Table 11. Compared with conventional methods, the swarm intelligence optimization method has better prediction results; specifically, MAPE is reduced by 34.75% and RMSE is reduced by 16.67%.

**Table 11.** Comparison of RUL prediction results before and after adopting swarm intelligence optimization method.

| Method | MAPE | RMSE |
|---|---|---|
| Conventional HSMM | 0.613% | 0.402 |
| PSO-based | 0.826% | 0.335 |

To verify whether the HSMM method improved by the swarm intelligence optimization algorithm exhibits a statistically significant advantage, firstly, the normality of the error sequence is evaluated using a Shapiro–Wilk test, yielding a test statistic of W = 0.951 and a significance level $p = 0.181 > 0.05$, indicating that the data follow a normal distribution. Therefore, a one-sample $t$-test is used to evaluate whether the mean of the errors deviates significantly from 0, yielding $t(29) = 0.392$, $p = 0.698$, and a 95% confidence interval for the mean difference of $[-0.103, 0.151]$, which includes 0.

The 30 prediction errors are obtained from adjacent rolling points over the life cycle of the same equipment and are therefore temporally dependent. If a one-sample $t$-test is performed directly, the standard error will be underestimated, leading to an artificially small $p$ value. To address this issue, we adopt the effective-sample-size-corrected $t$-test [45]. The key idea is not to change the calculation of the t-statistic, but to replace the nominal sample size of 30 with the effective sample size $n_{eff}$ and then refer to the $t$-distribution.

First, the lag-1 autocorrelation coefficient $\rho_1 = -0.109$ is calculated for the 30 prediction error series. The calculation method for $\rho_1$ is as follows:

$$\rho_1 = \frac{\sum_{i=1}^{n-1}(e_i - \bar{e})(e_{i+1} - \bar{e})}{\sum_{i=1}^{n}(e_i - \bar{e})^2} \tag{62}$$

where $\bar{e}$ represents the mean of the error series.

Then, the effective sample size $n_{eff} = 37.34$ is obtained. The calculation method for $n_{eff}$ is as follows:

$$n_{eff} = n \cdot \frac{1 - \rho_1}{1 + \rho_1} \tag{63}$$

The lag-1 autocorrelation coefficient of the error series is $\rho_1 = -0.109$, indicating weak negative autocorrelation. Negative autocorrelation implies that adjacent prediction errors tend to deviate from the mean in opposite directions. In other words, the error series exhibits slight mean reversion or oscillation rather than persistent deviations, and its information content is no lower than that in the independent case. From a statistical inference perspective, negative autocorrelation can lead to overestimation of the standard error, making the *t*-test conservative rather than producing spurious significance. The calculated effective sample size is $n_{eff} = 37.34$. Since negative autocorrelation does not increase the amount of sample information, we conservatively set $n_{eff} = n = 30$. Accordingly, the effective-sample-size-corrected *t*-test yields the same results as the one-sample *t*-test; in both cases, the null hypothesis that the error mean is zero cannot be rejected.

A Wilcoxon signed-rank test is conducted for robustness verification, and the result ($p = 0.614$) also indicates non-significance.

In summary, both tests indicate the absence of significant deviation from 0 in the prediction errors, demonstrating that there is insufficient evidence that the method exhibits systematic overestimation or underestimation bias.

To evaluate the performance stability of the prediction method across different degradation stages, the 30 predictions are divided into three stages based on the RUL percentage corresponding to each prediction start point: early period (RUL > 2/3 of the total life cycle), middle period (RUL between 1/3 and 2/3 of the total life cycle), and late period (RUL < 1/3 of the total life cycle). Here, the total life cycle of the equipment is defined as the sum of the first prediction start point and its corresponding actual RUL. Levene's test (median-based) compares the homoscedasticity of prediction errors across the three stages.

Based on the results of Levene's test (test statistic W = 0.368, $p = 0.696 > 0.05$), the null hypothesis of homoscedasticity cannot be rejected. This indicates that, regardless of whether the equipment is in its early healthy state or approaches failure, the method produces prediction errors without significant differences in fluctuation range, i.e., it maintains prediction stability independent of specific degradation stages.

Similarly, the results of Spearman's rank correlation test (correlation coefficient $\rho = -0.037$, $p = 0.847 > 0.05$) indicate that there is no statistically significant monotonic correlation between the error sequence and the prediction time sequence. In other words, the method does not produce systematic differences in predictions between the early and late periods of the equipment life cycle, and the prediction accuracy is independent of specific degradation stages, further verifying the robustness of the method.

The statistical test results provide a preliminary validation of the feasibility of the proposed method on a single equipment life cycle. The method explicitly models state dwell time and offers an interpretable probabilistic framework for RUL prediction. Given the manufacturing differences, operating condition variations, and failure mode differences across equipment, a single-case result is insufficient to establish the robustness of the

method. Future work will systematically assess its generalization capability on multi-equipment and multi-operating-condition datasets.

The predicted RUL values provide a foundation for dynamically adjusting equipment maintenance strategies in actual railway transport, including maintenance schedules, maintenance items, risk warnings, and the storage of spare parts. Such adjustments facilitate advanced intervention in the equipment degradation process and help proactively prevent equipment faults, thereby ensuring the safe and stable operation of the equipment.

To sum up, taking the IFs in the power module used in a railway application as an example, this paper demonstrates through testing that the HSMM-based approach for recognizing equipment operational states and predicting RUL is effective. The prediction results are considered valuable for guiding equipment maintenance decision-making. The proposed method offers a feasible solution for equipment health management throughout the entire life cycle. It should be noted that, to provide a proof of concept, the experiments detailed in this section focus on power modules—a typical equipment type—and are conducted using characteristic parameters that are common to all electronic equipment, rather than those unique to power modules. Therefore, the experimental method is generally applicable and provides a foundation for subsequent experiments involving a wider variety of electronic equipment.

## 6. Prospects for Practical Application

To achieve a closed-loop implementation of the proposed method, spanning from equipment performance evaluation and RUL prediction to on-site maintenance, this paper develops a health management platform for railway electronic equipment and proposes a closed-loop implementation path based on the "prediction–decision–execution–feedback" process. The core task of this platform is to transform the technical metric of "how much life remains for the equipment" into an economic and dispatching decision regarding "when to repair, what to repair, and what spare parts to stock". The specific implementation path is shown in Figure 9.

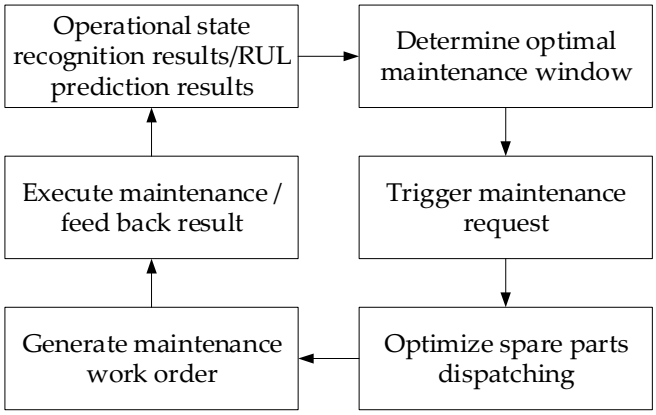



**Figure 9.** Implementation path of health management for electronic equipment.

The implementation steps in Figure 9 are detailed as follows:

(1) Receive equipment state recognition results and RUL prediction results.

(2) Determine the optimal maintenance window: Most railway electronic equipment is part of safety-critical systems, and its maintenance activities are premised on safe and reliable equipment operation. Therefore, safety and risk tolerance are the primary considerations in determining maintenance windows. Under the constraint that risk remains manageable, a time window of $\pm 10\%$ around the predicted RUL value is set as the optimal maintenance window. A maintenance request is triggered when the equipment's

operating time reaches the start of this window. Furthermore, when the lower confidence limit of the predicted RUL approaches the threshold for safe operation, the time window is unconditionally advanced to ensure operational safety.

(3) Trigger a maintenance request: The proposed trigger mechanism does not rely on a single signal source; rather, it consists of two components: a warning trigger and a decision trigger. The warning trigger is intended to allow sufficient time for tasks such as spare part procurement and staff scheduling. When the predicted RUL value is less than twice the spare part lead time, the platform issues a yellow alert and initiates the spare part provisioning process. The decision trigger works primarily when the equipment's operating time reaches the left endpoint of the optimal maintenance window, at which point the platform issues a red decision command. This dual-trigger mechanism works synergistically to avoid maintenance delays caused by resource shortages and to prevent premature maintenance that would unnecessarily shorten the useful life.

(4) Optimize spare part dispatching: The essence of spare part dispatching optimization is to establish a risk-oriented dynamic balance between maintenance availability and inventory turnover. Its core objective is to minimize the total life-cycle inventory cost of spare parts while meeting the target spare part availability rate (e.g., $\geq$98%). Specifically, instead of relying on fixed-cycle dispatching, the probability of future demand for spare parts is derived from the RUL probability distribution. This optimization strategy shifts spare part inventory from "passive response to failures" to "proactive alignment to forecasted demand".

(5) Generate a maintenance work order: The platform converts decision instructions into structured maintenance work orders that specify maintenance strategies, timelines, and resource lists, and pushes them to field workers' handheld devices in a timely manner. It is important to note that the push timing is selected based on transport schedules. For trackside electronic equipment, maintenance must be performed during maintenance windows—track closure periods designated for construction and maintenance activities, during which no regular trains operate in the given sections. For onboard electronic equipment, maintenance can only be performed when the train is taken out of service and is parked in a maintenance workshop, where maintenance activities are carried out in a centralized manner.

(6) Execute maintenance/feed back result: This step links predictions and decisions with actual field operations. At the execution level, employees follow guidance received on handheld devices to track the progress and process of maintenance activities in real time; at the feedback level, the platform records data such as actual repair durations, spare part consumption, critical items, and types of anomalies to create comprehensive execution logs. Furthermore, this paper proposes using deviations between the actual useful life and the predicted RUL as a benchmark for model evaluation and applying the Bayesian updating method to incorporate new observation data into the degradation model, thereby enabling online correction of RUL prediction parameters and adaptive adjustment of decision thresholds. This feedback mechanism ensures that the health management model can learn and evolve from every maintenance operation, forming a closed-loop proactive preventive maintenance cycle of "prediction–decision–execution–feedback".

In summary, the health management platform for railway electronic equipment based on the model presented in Section 4 continuously monitors equipment operational states, automatically generates tailored maintenance recommendations, and undergoes continuous feedback optimization. It establishes a complete decision chain for precision maintenance aimed at maximizing life-cycle operational efficiency and supplies the technical support required for field application.

## 7. Conclusions

Considering the O&M characteristics of railway electronic equipment, such as complex operational environments and varying maintenance methods, a health management model is developed based on HSMM, introducing IFs to characterize the performance degraded states of electronic equipment. Focusing on the two key techniques of equipment state evaluation and RUL prediction, swarm intelligence optimization algorithms are incorporated to enhance the parameter estimation of HSMM. This integration can solve the problem of model parameters converging to local optima. Additionally, the issue of RUL equivalence for identical states across different timepoints is resolved by identifying state transition points based on observation sequences predicted using LSTM. Experiments on a power module IFs in a railway scenario show that the PSO-based HSMM parameter estimation has the highest accuracy of equipment state recognition, up 2.15% compared to the conventional method. With an RMSE of 0.335 for RUL prediction across various timepoints throughout the equipment's full life cycle, 16.67% lower than that of the conventional methods, the statistical tests provide insufficient evidence that the proposed method exhibits systematic overestimation or underestimation bias. It offers an effective and feasible solution for full-life-cycle health management of railway electronic equipment.

It is suggested that cross-equipment, cross-scenario and cross-failure-type experimental verification be performed in the future to optimize the model's generalization ability and improve its robustness, allowing for better adaptation to the randomness and diversity of IFs.

**Author Contributions:** Conceptualization, Y.P. and R.K.; methodology, Y.P. and R.K.; software, L.L., J.C. and J.L.; validation, R.K. and J.L.; formal analysis, L.L.; investigation, Y.P., R.K. and J.C.; resources, L.L. and J.L.; data curation, L.L., Y.P. and R.K.; writing—original draft preparation, R.K. and L.L.; writing—review and editing, R.K. and Y.P.; visualization, J.C.; supervision, J.L. and L.L.; project administration, J.C.; funding acquisition, R.K. and J.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Innovation Driven Development Special Project for Guangxi Science and Technology Plan Project: Construction of China-ASEAN International Joint Laboratory for Comprehensive Transportation (Phase I), grant number: Guike AA21077011; China Railway, grant number: J2024G005; China Academy of Railway Sciences Co., Ltd., grant number: 2025YJ212; Construction of Scientific and Technological Innovation Bases for Science Research and Technology Development Plan Project of Nanning City, grant number: 20223230; Beijing HUA-TIE Information Technology Co., Ltd., grant number: 2024HT12.

**Data Availability Statement:** The data presented in this study are not publicly available due to encryption and confidentiality restrictions. Specifically, the data are collected from actual railway operation and maintenance systems and are subject to confidentiality agreements with the operating company. The raw data contain sensitive equipment-failure records, and commercially relevant maintenance information. Subject to appropriate authorization and compliance with relevant security regulations, the data can be made available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors would like to express their gratitude to Ying Hu for translating and polishing the manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Song, H.; Gao, S.; Li, Y.; Liu, L.; Dong, H. Train-Centric Communication Based Autonomous Train Control System. *IEEE Trans. Intell. Veh.* **2023**, *8*, 721–731. [CrossRef]
2. Chang, X.; Zhang, R.; Mao, J.; Fu, Y. Digital Twins in Transportation Infrastructure: An Investigation of the Key Enabling Technologies, Applications, and Challenges. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 6449–6471. [CrossRef]

3. Li, Y.-F.; Wang, H.; Sun, M. ChatGPT-like large-scale foundation models for prognostics and health management: A survey and roadmaps. *Reliab. Eng. Syst. Saf.* **2024**, *243*, 109850. [CrossRef]
4. Hsu, C.-T.; Yao, K.-C.; Chang, T.-Y.; Hsu, B.-K.; Shyr, W.-J.; Chou, D.-F.; Lai, C.-C. A Mathematics-Oriented AI Iterative Prediction Framework Combining XGBoost and NARX: Application to the Remaining Useful Life and Availability of UAV BLDC Motors. *Mathematics* **2025**, *13*, 3460. [CrossRef]
5. Lyu, K.; Cheng, X.; Li, H.; Zhang, Y.; Qiu, J.; Liu, G. New developments of prognostic and health management technology for electronic equipment. *Astronaut. Sin.* **2019**, *40*, 23285. [CrossRef]
6. Song, H.; Zhang, H.; Wu, X.; Li, W.; Dong, H. Vision and Inertial Sensors Fusion for Train Positioning in GNSS-Denied Environments. *IEEE Sens. J.* **2025**, *25*, 35323–35334. [CrossRef]
7. Vrignat, P.; Kratz, F.; Avila, M. Sustainable manufacturing, maintenance policies, prognostics and health management: A literature review. *Reliab. Eng. Syst. Saf.* **2022**, *218*, 108140. [CrossRef]
8. Gao, T.; Neusypin, K.A.; Dmitriev, D.D.; Yang, B.; Rao, S. Unsupervised Segmentation and Alignment of Multi-Demonstration Trajectories via Multi-Feature Saliency and Duration-Explicit HSMMs. *Mathematics* **2025**, *13*, 3057. [CrossRef]
9. Forouzanfar, M.; Gagliardi, G.; Tedesco, F.; Casavola, A. Integrated Model-Based Control Allocation Strategies Oriented to Predictive Maintenance of Saturated Actuators. *IEEE Trans. Autom. Sci. Eng.* **2024**, *22*, 1045–1056. [CrossRef]
10. Yan, J.; He, Z.; He, S. Multitask learning of health state assessment and remaining useful life prediction for sensor-equipped machines. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109141. [CrossRef]
11. Ren, L.; Liu, Y.; Wang, X.; Lu, J.; Deen, M.J. Cloud-Edge-Based Lightweight Temporal Convolutional Networks for Remaining Useful Life Prediction in IIoT. *IEEE Internet Things* **2021**, *8*, 12578–12587. [CrossRef]
12. Bae, J.; Xi, Z. Learning of physical health timestep using the LSTM network for remaining useful life estimation. *Reliab. Eng. Syst. Saf.* **2022**, *226*, 108717. [CrossRef]
13. Wang, S.; Liu, Z.; Jia, Z.; Zhao, W.; Li, Z. Intermittent fault diagnosis for electronics-rich analog circuit systems based on multi-scale enhanced convolution transformer network with novel token fusion strategy. *Expert Syst. Appl.* **2024**, *238*, 121964. [CrossRef]
14. Zhou, H.; Li, Z.; Wu, T.; Wang, C.; Li, K. Prognostic and Health Management of CT Equipment via a Distance Self-Attention Network Using Internet of Things. *IEEE Internet Things* **2024**, *11*, 31338–31354. [CrossRef]
15. Xue, K.; Yang, J.; Yang, M.; Wang, D. An Improved Generic Hybrid Prognostic Method for RUL Prediction Based on PF-LSTM Learning. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–21. [CrossRef]
16. Cao, X.; Peng, K. Stochastic Uncertain Degradation Modeling and Remaining Useful Life Prediction Considering Aleatory and Epistemic Uncertainty. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–12. [CrossRef]
17. Li, G.; Yang, L.; Lee, C.-G.; Wang, X.; Rong, M. A Bayesian Deep Learning RUL Framework Integrating Epistemic and Aleatoric Uncertainties. *IEEE Trans. Ind. Electron.* **2021**, *68*, 8829–8841. [CrossRef]
18. Handayania, L.; Vrignat, P.; Kratz, F. HSMM multi-observations for prognostics and health management. *Proc. Inst. Mech. Eng. O J. Risk Reliab.* **2024**, *239*, 253–275. [CrossRef]
19. Ferreira, R.S.; Favoreto, H.H. Transient overvoltages due to intermittent-ground faults in an industrial power system grounded by a resistance connected to the secondary of a grounding transformer. *Electr. Pow. Syst. Res.* **2023**, *220*, 109327. [CrossRef]
20. Mulahuwaish, A.; Leduc, R.J. Fault-tolerant supervisory control with permanent faults. *Int. J. Control* **2023**, *96*, 822–838. [CrossRef]
21. Asaadi, M.; Izadi, I.; Hassanzadeh, A.; Yang, F. Assessment of alarm systems for mixture processes and intermittent faults. *J. Process Control* **2022**, *114*, 120–130. [CrossRef]
22. Ibrahim, M.; Rozas, H.; Gebraeel, N. An Integrated Detection-Prognostics Methodology for Components With Intermittent Faults. *J. Comput. Inf. Sci. Eng.* **2024**, *24*, 061003. [CrossRef]
23. Gao, S.G.; Zhao, K.B. Intermittent Fault Diagnosis of Dynamic Systems with Model Uncertainty and Disturbance: An Adaptive Nondeterministic Observer Approach. *IEEE Trans. Reliab.* **2024**, *74*, 4131–4142. [CrossRef]
24. Natan, A.; Kalech, M.; Bartak, R. Diagnosis of intermittent faults in Multi-Agent Systems: An SFL approach. *Appl. Intell.* **2023**, *324*, 103994. [CrossRef]
25. Huang, C.; Shen, Z.; Zhang, J.; Hou, G. BIT-Based Intermittent Fault Diagnosis of Analog Circuits by Improved Deep Forest Classifier. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–13. [CrossRef]
26. Amaral, T.G.; Pires, V.F.; Cordeiro, A.; Foito, D.; Martins, J.F.; Yamnenko, J.; Tereschenko, T.; Laikova, L.; Fedin, I. Incipient Fault Diagnosis of a Grid-Connected T-Type Multilevel Inverter Using Multilayer Perceptron and Walsh Transform. *Energies* **2023**, *16*, 2668. [CrossRef]
27. Hojabri, M.; Nowak, S.; Papaemmanouil, A. ML-Based Intermittent Fault Detection, Classification, and Branch Identification in a Distribution Network. *Energies* **2023**, *16*, 6023. [CrossRef]
28. Ramezani, S.; Moini, A.; Riahi, M.; Marquez, A.C. A model to determining the remaining useful life of rotating equipment, based on a new approach to determining state of degradation. *J. Cent. South Univ.* **2020**, *27*, 2291–2310. [CrossRef]

29. Liu, Z.; Wang, C.; Yang, X.; Zhang, N.; Liu, F.; Zhang, B. Time Series Multi-Step Forecasting Based on Memory Network for the Prognostics and Health Management in Freight Train Braking System. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 8149–8162. [CrossRef]
30. Cheng, X.; Lv, K.; Zhang, Y.; Wang, L.; Zhao, W.; Liu, G.; Qiu, J. RUL Prediction Method for Electrical Connectors With Intermittent Faults Based on an Attention-LSTM Model. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2023**, *13*, 628–637. [CrossRef]
31. Rahman, M.A.; Rahim, M.A.; Rahman, M.M.; Moustafa, N.; Razzak, I.; Ahmad, T.; Patwary, M.N. A Secure and Intelligent Framework for Vehicle Health Monitoring Exploiting Big-Data Analytics. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19727–19742. [CrossRef]
32. Kong, Z.; Jin, X.; Xu, Z.; Zhang, B. Spatio-Temporal Fusion Attention: A Novel Approach for Remaining Useful Life Prediction Based on Graph Neural Network. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–12. [CrossRef]
33. Mancuso, A.; Compare, M.; Salo, A.; Zio, E. Optimal Prognostics and Health Management-driven inspection and maintenance strategies for industrial systems. *Reliab. Eng. Syst. Saf.* **2021**, *210*, 107536. [CrossRef]
34. Liu, J.; Xu, K.Z.; Cai, B.G.; Guo, Z.B. Fault Prediction of On-Board Train Control Equipment Using a CGAN-Enhanced XGBoost Method with Unbalanced Samples. *Machines* **2023**, *11*, 114. [CrossRef]
35. Zhang, D.; Xie, M.; Yang, J.; Wen, T. Multi-Sensor Graph Transfer Network for Health Assessment of High-Speed Rail Suspension Systems. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 9425–9434. [CrossRef]
36. Yin, Y.; Liu, Z.; Zhang, Q.; Qin, Y.; Zuo, M. A Data Compression Method with an Encryption Feature for Safe and Lightweight Vibration Condition Monitoring. *IEEE Internet Things* **2024**, *11*, 30524–30535. [CrossRef]
37. Ruan, D.; Wu, Y.; Yan, J.; Guhmann, C. Fuzzy-Membership-Based Framework for Task Transfer Learning between Fault Diagnosis and RUL Prediction. *IEEE Trans. Reliab.* **2023**, *72*, 989–1002. [CrossRef]
38. Lewis, A.D.; Groth, K.M. Metrics for evaluating the performance of complex engineering system health monitoring models. *Reliab. Eng. Syst. Saf.* **2022**, *223*, 108473. [CrossRef]
39. Panmala, N.; Suwanasri, T.; Suwanasri, C. Condition Assessment of Gas Insulated Switchgear Using Health Index and Conditional Factor Method. *Energies* **2022**, *15*, 9393. [CrossRef]
40. Heimisch, L.; Preuss, K.; Russwinkel, N. Cognitive processing stages in mental rotation—How can cognitive modelling inform HsMM-EEG models? *Neuropsychologia* **2023**, *188*, 108615. [CrossRef] [PubMed]
41. Hadj-Amar, B.; Jewson, J.; Vannucci, M. Bayesian Sparse Vector Autoregressive Switching Models with Application to Human Gesture Phase Segmentation. *Ann. Appl. Stat.* **2024**, *18*, 2511–2531. [CrossRef]
42. Nishikimi, R.; Nakamura, E.; Goto, M.; Yoshii, K. Audio-to-score singing transcription based on a CRNN-HSMM hybrid model. *APSIPA Trans. Signal Inf. Process.* **2021**, *10*, e7. [CrossRef]
43. del Rio, A.; Barambones, O.; Uralde, J. Design and Experimental Validation of a Novel Particle Swarm Optimization Algorithm Designed to Optimize Solar Power Extraction. *Mathematics* **2026**, *14*, 2079. [CrossRef]
44. Behgouy, P.; Ugurenver, A. Hybrid Microgrid Power Management via a CNN–LSTM Centralized Controller Tuned with Imperialist Competitive Algorithm. *Mathematics* **2025**, *13*, 4030. [CrossRef]
45. O'Shaughnessy, P.; Cavanaugh, J.E. Performing T-tests to Compare Autocorrelated Time Series Data Collected from Direct-Reading Instruments. *J. Occup. Environ. Hyg.* **2015**, *12*, 743–752. [CrossRef] [PubMed]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.