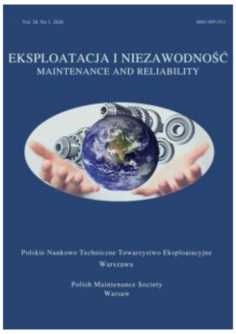


Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 28 (2026), Issue 1

journal homepage: http://www.ein.org.pl

Article citation info:
Li S, Yang Z, Fang C, Luo H, He J, Li J, Nie X, Reliability Estimation of Retraction Mechanism Kinematic Accuracy under Small Sample, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(1) http://doi.org/10.17531/ein/207148

Reliability Estimation of Retraction Mechanism Kinematic Accuracy under Small Sample

Indexed by: Web of Science Group

Sen Li[a], Zhaojun Yang[a], Chuanyun Fang[b], Huaiyong Luo[c], Jialong He[a,c,*], Jia Li[a,], Xinyu Nie[d]

[a] School of Mechanical and Aerospace Engineering, Jilin University, China
[b] CHONGQING QINGPING MACHINERY CO.,LTD, China
[c] Chongqing Research Institute, Jilin University, China
[d] Faw Jiefang Automotive Co.,Ltd, China

Highlights

- Maintenance strategy integrates operational availability under delay propagation.
- Introducing delay time as a state variable to handle operational availability impacts.
- A Markov Decision Process–Based Condition-Based Maintenance Model.
- A scenario-based value iteration method to solve discrete health states effectively.

Abstract

Manufacturing system degradation can damage its reliability, resulting in decreased product quality and delayed deliveries. These challenges are characteristic of imperfect manufacturing systems. Moreover, the propagation of delay time across task periods may reduce the operational availability in future periods. Condition-based maintenance is an effective method for mitigating system degradation and enhancing reliability. However, existing condition-based maintenance studies often overlook the impact of delay propagation on operational availability. To address this issue, this paper proposes a condition-based maintenance model based on a Markov decision process. By introducing delay time as a state variable to capture changes in operational availability and incorporating it into the reward model, the proposed strategy aims to maximize enterprise profit. A case study and comparative analysis using data from a manufacturing enterprise validate the effectiveness and superiority of the proposed model in improving economic performance.

Keywords

reliability, condition-based maintenance, imperfect manufacturing system, operational availability, delay propagation, Markov decision process

1. Introduction

The manufacturing system is central to enterprise operations, and its reliability refers to the ability to complete specified tasks under given conditions within a defined time frame[1]. Over time, system health degrades, reducing reliability and potentially leading to quality deterioration and production delays[2]. In severe cases, degradation causes unexpected shutdowns, which reduce the average proportion of time the system operates effectively during a task period—referred to as operational availability[1]. Such degradation and quality variability reflect the characteristics of imperfect manufacturing systems[3, 4], where issues arise from both machine unreliability and the probability of producing nonconforming products. As these imperfections threaten both operational availability and product quality, it is essential to develop maintenance strategies that ensure the reliability of imperfect manufacturing systems.

(*) Corresponding author.
E-mail addresses: S. Li (ORCID: 0009-0008-6186-6314) lisen92@foxmail.com, Z. Yang, yzj@jlu.edu.cn, C. Fang, scyy_2018@163.com, H. Luo, luohuaiyong@cqrijlu.com, J. He (ORCID: 0000-0001-6598-6091) hejl@jlu.edu.cn, J. Li, lijia199407@163.com, X. Nie, niexinyu@fawjiefang.com.cn

To preserve system reliability, effective maintenance strategies are essential for mitigating degradation and sustaining performance[5, 6]. Traditional strategies like Time-Based and Failure-Based Maintenance[7, 8] are simple but may cause unnecessary losses due to untimely or excessive interventions[9]. Condition-Based Maintenance (CBM) improves efficiency by using sensor data to model degradation and schedule actions based on system health[10]. Advances in sensing and artificial intelligence have further enhanced the practicality of CBM[11]. Central to CBM is accurately modeling degradation: recent CBM methods increasingly adopt continuous-state stochastic processes rather than discrete-state models, effectively capturing system dynamics[12, 13]. Among these methods, Markov Decision Processes (MDP) stand out due to their capability to model complex state transitions and optimize long-term maintenance decisions[14].

Recent CBM studies have begun to jointly consider product quality and production for imperfect manufacturing system. However, most neglect operational availability, and even those considering availability generally ignore the impact of delay propagation, i.e., delays in one task period inevitably spilling over into subsequent periods and further impairing availability[15]. Given the complexity introduced by system degradation, uncertain quality, and delay propagation, CBM is particularly well-suited to be modeled using MDP.

Therefore, targeting imperfect manufacturing systems, this paper proposes an MDP-based CBM strategy that introduces delay time as a state variable to reflect variations in operational availability. This approach enables the model to explicitly capture delay propagation effects and optimize maintenance decisions to maximize long-term enterprise profitability.

2. Literature review and contributions

This paper presents a review of the relevant literature in the context of the problem to be addressed. In the first subsection, current research on maintenance strategy for imperfect manufacturing systems is reviewed. In the second subsection, studies on maintenance strategies regarding availability are discussed. In the third subsection, current research on solving CBM problems for manufacturing systems using MDP is presented. Finally, the technical contributions of this paper are summarized.

2.1. Maintenance strategy for imperfect manufacturing system

In imperfect manufacturing systems, the coexistence of equipment degradation and quality uncertainty poses significant challenges for maintenance strategy design. Unlike idealized systems, these imperfections require maintenance policies to jointly consider system reliability, product quality, and production continuity. As a result, increasing attention has been paid to integrated strategies that combine production, maintenance, and quality control: Li et al.[16] proposed a CBM strategy that incorporates the working schedule to balance maintenance with production while maintaining product quality. Shi et al.[17] developed a model for imperfect systems that integrates decisions on production, maintenance, and quality under inventory constraints. Ait-El-Cadi et al.[18] introduced a joint control policy for production, maintenance, and dynamic quality inspection in failure-prone systems. Zhang et al.[19] used a digital twin to co-optimize predictive maintenance and production scheduling. Guendouli et al. [20]proposed an integrated policy considering production-related parameters affecting degradation and final product quality.

Although these studies reflect the complexity of imperfect manufacturing systems, they primarily focus on optimizing cost, quality, and scheduling outcomes, while few explicitly incorporate operational availability into the decision-making process. As discussed in the introduction, availability is essential for understanding how degradation and maintenance affect system performance over time. Therefore, there remains a gap in addressing availability within integrated CBM strategies for imperfect systems.

2.2. Maintenance strategy regarding availability

As discussed in the introduction, availability characterizes the proportion of time a system remains functionally operational within a task period, thereby directly influencing production output and timely order fulfillment. Some studies have explored maintenance strategies with a focus on maintaining high availability levels, highlighting its role as a critical indicator of system performance over time. The following works represent recent developments in this direction: Yin et al.[21] examined how to optimize the periodic preventive maintenance rate to maximize availability in Markov systems with multiple degraded states. Chalabi et al.[22] proposed a grouping strategy for multi-unit systems to improve availability while reducing preventive maintenance costs. Lotovskyi et al.[23] developed an age-based maintenance model that evaluates the effect of perfect and imperfect maintenance on the availability of offshore oil production systems. Yang et al.[24] proposed an availability-oriented maintenance strategy for automated production lines that accounts for performance degradation. An et al.[25] introduced a condition-based maintenance model that minimizes cost risk under operational availability constraints.

However, these studies are generally not framed within the context of imperfect manufacturing systems, nor do they consider the propagation of delays across task periods—an omission that hinders their ability to capture how current delays reduce future operational availability.

2.3. Condition-based maintenance studies utilizing MDP

The degradation of manufacturing systems often exhibits Markovian properties, making Markov Decision Processes (MDP) a suitable framework for modeling CBM strategies based on system state information. Numerous studies have applied MDP or reinforcement learning to optimize maintenance decisions[26-28]. For example, Zhang et al.[29] proposed a semi-Markov decision model for a two-component series system that adaptively determines optimal strategies without a predefined structure. Zhou et al.[30] developed an MDP-based method for series-parallel systems with intermediate buffers to identify optimal maintenance actions. Tang et al.[31] introduced a semi-MDP model that integrates dynamic maintenance policies with mean residual life estimation. Liu et al.[32] formulated a selective maintenance strategy for missions under imperfect maintenance using Q-learning within an MDP framework.

While these studies demonstrate the applicability of MDP-based frameworks in maintenance optimization, few of them explicitly incorporate operational availability or model the cumulative effect of delay propagation—key factors in imperfect manufacturing systems addressed in this study.

2.4. Contributions and outline

Targeting imperfect manufacturing systems characterized by both equipment degradation and product quality uncertainty, this paper addresses the overlooked issue of delay propagation in operational availability. The main contributions are as follows:

(1)A condition-based maintenance (CBM) strategy is proposed for imperfect manufacturing systems, which explicitly considers operational availability while ensuring product quality. To capture the effect of delay propagation, delay time is introduced as a state variable and combined with health state to form a two-dimensional state space.

(2)The CBM strategy is formulated within a Markov Decision Process (MDP) framework that models both degradation-induced failures and quality deterioration. A scenario-based value function is constructed to solve the Bellman optimal equation and derive the optimal maintenance policy.

The remainder of the paper is as follows: Section 3 presents the problems and assumptions to be addressed by the model; Section 4 models the maintenance strategy of manufacturing systems through MDP; Section 5 represents the solution method of the maintenance strategy model; Section 6 conducts a numerical study and comparative analysis based on the historical data of a commercial vehicle manufacturing company to obtain the optimal maintenance strategy, demonstrating the effectiveness and superiority of the proposed approach; Finally, Section 7 concludes this paper.

3. Problem description and assumptions

In this section, an imperfect single-machine manufacturing system that consistently produces a single product type is considered as the study object. The flowchart of the maintenance strategy is presented in Fig.1.

The duration of the task period for the single-machine manufacturing system is denoted as δ, starting with a maintenance period. During this period, the maintainer performs system inspections, which is assumed to require negligible time and cost. Based on the observed health state s_n in the nth task period and the delay time T_n^{ot} carried over from the previous period, the manager selects a maintenance action a_n: No Maintenance(Null, no action is taken), Imperfect Maintenance(IM, imperfectly restores the system), or Overhaul(restores the system to a "good as new" state). After the maintenance action is executed, the system transitions to a new health state $s_n^{'}$, and the corresponding maintenance time is denoted as T_{Null}, T_{IM} or T_{O}. The associated costs are C_{Null}, C_{IM} and C_{O}. The relationship between time and cost is described in Eq.(1):

$$\begin{cases} T_O \gg T_{\text{IM}} > T_{\text{Null}} = 0 \\ C_O \gg C_{\text{IM}} > C_{\text{Null}} = 0 \end{cases} \tag{1}$$

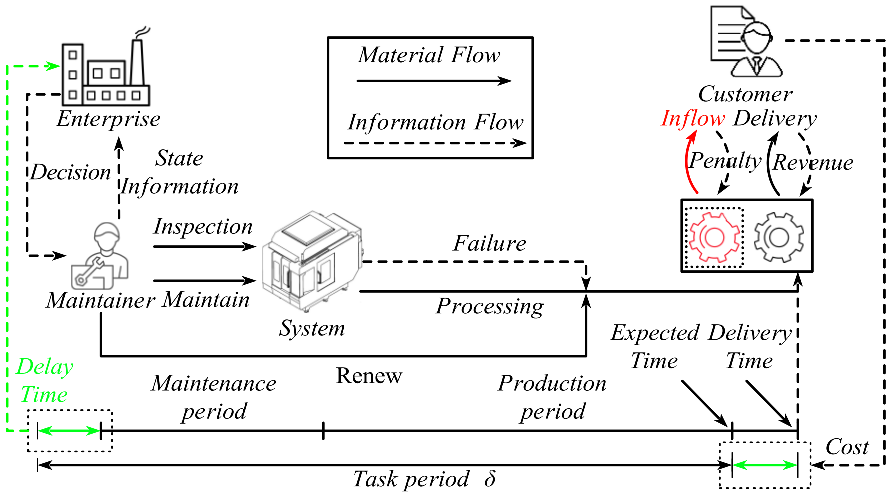


Fig.1. Maintenance strategy flowchart.

After the maintenance period concludes, the production period begins, during which the operator runs the system at a constant takt time t_{akt}. Customer demand is defined as a fixed upper-bounded constant D, with the range specified in Eq.(2):

$$D \le \left\lfloor \frac{\delta - T_O}{t_{akt}} \right\rfloor \tag{2}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. This ensures that customer demand D can be satisfied within the remaining time of the task period, even if the most time-consuming maintenance action is executed, without causing delays.

If a degradation failure occurs during the production period, i.e., the system degrades to the threshold L, the maintainer performs a Renew action. After Renew, the system returns to an optimal state, and no further degradation failures are assumed to occur within the current task period. Both Overhaul and Renew restore the system to a "good as new" condition, with their costs (C_{O} and C_R) being identical. However, the time required for Renew (T_{R}) is longer than that for Overhaul (T_{O}) because maintenance conditions during the maintenance period are more favorable. During the production period, the probability of producing nonconforming products increases as the system's health state deteriorates. These nonconforming products eventually reach the customer, incurring a unit penalty cost C_{pena}. The task period ends with the production period, and the enterprise is expected to deliver products by the customer's expected delivery time (before the end of the task period). Therefore, the ideal maximum operational availability of the system in the current period is given by the following equation:

$$A_n^{max \frac{\delta - T_n^{ot} - T_{a_n}}{\delta}} \tag{3}$$

where T_{a_n} is the required time for maintenance action a_n.

Each delivered product yields revenue r. If production is not completed on time, delayed delivery occurs, incurring a unit delay cost C_{del}. Within a finite time horizon, the delay time T_{n+1}^{ot} generated in the current task period carries over to the next, causing delay propagation and reduced the operational availability in the subsequent period. Consequently, the nth task period effectively begins with the delay time T_n^{ot} from the previous task period (indexed as n for consistency). The manager must account for both the health state s_n and the delay time when formulating the maintenance strategy to prevent further delay propagation to future periods.

4. Optimal condition-based maintenance strategy model

To construct the proposed CBM strategy, this section formulates the decision-making process within the MDP framework by aligning it with the CBM execution logic described in Section 3. In this process, the system state is first diagnosed, followed by maintenance decision-making, after which the system enters a production cycle where degradation and nonconforming products may occur due to system imperfections. Based on these outcomes, corresponding rewards are computed and state transitions occur.

Accordingly, Section 4.1 defines the state space and available maintenance actions, including delay time as a state variable to reflect variations in operational availability. Section 4.2 details the imperfect manufacturing system, focusing on degradation and quality-related uncertainties. Section 4.3 formulates the reward and state transition functions, capturing the feedback and dynamics for policy optimization. This structure ensures that all MDP elements—state, action, reward, and transition—are coherently embedded into the CBM process, highlighting the contribution of modeling delay propagation and imperfect manufacturing characteristics.

4.1. State space and Action

Let the manufacturing system health state s_n, and the delay duration T_n^{ot} constitute the state of the system in the nth task period, denoted by italicized S_n. The state S_n is an element of the state space $\boldsymbol{S_n}$, which is represented using italicized bold letters (the state space and the action space involved in the following are denoted by italicized bold letters).

The value of s_n ranges from $[0, L]$, where 0 represents the optimal health state and L represents the worst state, which is the threshold for system degradation failure. Based on the system state s_n, the corresponding maintenance action a_n is performed, leading to a transition in the system state to s_n'. Since T_n^{ot} remain unchanged, the state space of the system at the start of production period is $S_n' = \{S_n' | S_n' = (s_n', T_n^{ot})\}$, where the non-bold S_n' represents an element of the state space, i.e., the state of the system.

This section defines a generalized set $\boldsymbol{a}$ of maintenance actions and explores the corresponding health state transition functions:

(1) Null: No intervention is performed on the manufacturing system.

(2) Overhaul or Renew: The manufacturing system is restored to its optimal health state.

(3) Imperfect Maintenance (IM): Maintenance actions in real manufacturing scenarios are influenced by various factors, leading to inconsistent maintenance effects.

the probability mass function of the health state transition when a_n = Null is described by Eq.(4):

$$P(s_n' | s_n, a_n = \text{Null}) = \begin{cases} 1, s_n = s_n' \\ 0, \text{otherwise} \end{cases} \tag{4}$$

The health state transition probability mass function for a_n = Overhaul or Renew is given by Eq.(5):

$$P(s_n' | s_n, a_n = \text{Overhaul or Renew}) = \begin{cases} 1, s_n' = 0 \\ 0, \text{otherwise} \end{cases} \tag{5}$$

When a_n = IM, the health state s_1 after system maintenance is assumed to follow a Beta distribution constrained to the interval $(0, s_n)$, and the health state transition probability density function is described by Eq.(6)[33]:

$$y(s_n' | s_n, a_n = \text{IM}) = \frac{1}{s_n} \frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)} \left(\frac{s_n'}{s_n}\right)^{u-1} \left(1 - \frac{s_n'}{s_n}\right)^{v-1}, s_n' \in (0, s_n) \tag{6}$$

where u and v are positive parameters determined using maximum likelihood estimation from historical data.

The manager's decision regarding a in any system state forms the maintenance strategy presented in this paper, based on the available state information.

4.2. Imperfect Manufacturing System

An imperfect manufacturing system is characterized by the simultaneous presence of machine degradation and quality variability, where both reliability loss and the production of nonconforming products can occur during operation. In this paper, we use the Gamma process to model the degradation of the manufacturing system[34]. The system will enter the production period with the health state s_n', and the degradation of the system occurs only during the production period. For clarity, this subsection will use s_n' as the initial reference state for discussion. Given that the state of the system entering the production period of the nth work period is s_n', the probability density function of the state at the next work period denoted as s_{n+1}, is provided by Eq.(7):

$$f(s_{n+1}, t | s_n') = \frac{\beta^{\alpha t}(s_{n+1} - s_n')^{\alpha t - 1} e^{-\beta(s_{n+1} - s_n')}}{\Gamma(\alpha t)} \tag{7}$$

where α is the shape parameter, β is the scale parameter, and $\Gamma(\cdot)$ represents the Gamma function, as defined in Eq.(8):

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du, x > 0 \tag{8}$$

The Bernoulli process is a discrete-time stochastic process[35] consisting of an infinite number or a finite number of mutually independent random variables, each taking on one of two values: 0 or 1, where 0 represents "failure" and 1 represents "success". If the probability of a manufacturing system producing nonconforming products is constant, the system's production process can be regarded as a Bernoulli process with a stable parameter, where 0 stands for conforming products and 1 for nonconformities. According to the literature[36], the probability of a manufacturing system producing nonconformities is positively correlated with its state s_n', increasing as s_n' becomes larger. The probability at state s_n' is described by Eq.(9):

$$p(s_n') = p_0 + \eta\left(1 - exp\left(-\lambda_q (s_n')^{\gamma_q}\right)\right) \quad (9)$$

where p_0 is the probability of nonconforming product when the system is in its optimal state, η denotes the quality degradation boundary, and λ_q and γ_q are constants determined from historical data.

The Gamma process is a Lévy process, which, by definition, continuous-time Markov processes[37, 38]. Consequently, the production process of the single-machine manufacturing system can be regarded as a special Bernoulli process, where the probability of success varies according to the Markov chain. This stochastic process is referred to as a Markov-modulated Bernoulli process (MMBP)[39]. However, MMBP is a Bernoulli process with non-smooth parameters, which is difficult to analyze and compute. The literature[40] proposes a smooth binomial approximation method for the discrete Markov Modulated Bernoulli Process, allowing it to be approximated as a smooth-parameter Bernoulli process. The approximated probability of success is expressed by Eq.(10):

$$\hat{p}_{(s_n')} = D^{-1} \sum_{i=1}^{D} \sum_{s=s_n'}^{L} p_d(s_{id}) p_{s_n', s_{id}}^{(i-1)} \quad (10)$$

where $p_d(s_{id})$ represents the probability that the system will produce a nonconforming product (probability of success) at state s_{id}, and s_{id} denotes the health state of the system after producing the ith product.

In this paper, the health state of the manufacturing system is a continuous random variable, and therefore the above equation is transformed into a continuous form, as shown in Eq.(11):

$$\hat{p}_{(s_n')} = D^{-1} \sum_{i=0}^{D-1} \int_{s_n'}^{L} p_d(s_{id}) \frac{f(s_{id}, i \cdot d | s_n')}{\int_{s_n'}^{L} f(s_{id}, i \cdot d | s_n') ds_{id}} ds_{id} \quad (11)$$

4.3. Reward and state transition function

In the nth task period, when the state of the system is S_n, an immediate reward $r(S_n, a_n)$ is obtained upon executing the action a_n, as described in Eq.(12):

$$r(S_n, a_n) = -C_{a_n} \quad (12)$$

where C_{a_n} is the cost of the action a_n.

The system begins production with a health state s_n', and as described in Section 3, two scenarios may occur during the production period. These scenarios are discussed in detail below:

Scenario 1 (ω_1): The system degrades to failure during the production period, leading to downtime.

Assuming the system shuts down at time T, the probability density function of the system entering Scenario 1 at health state s_n' is given by Eq.(13) :

$$P(\omega_1 | s_n') = f(L, T | s_n') \quad (13)$$

where $f(\cdot)$ is the Eq.(7).

Assuming that $w_1^{(1)}$ denotes the number of nonconforming products produced by the system before entering Scenario 1, with its probability given by Eq.(14):

$$p(w_1^{(1)} | s_n') = \binom{\left\lfloor \frac{T}{t_{akt}} \right\rfloor}{w_1^{(1)}} \hat{p}_{(s_n')}^{\ w_1^{(1)}} (1 - \hat{p}_{(s_n')})^{\left\lfloor \frac{T}{t_{akt}} \right\rfloor - w_1^{(1)}} \quad (14)$$

where $\hat{p}_{(s_n')}$ represents the approximate smooth parameter probability of producing a nonconforming product when the health state of the system is s_n', obtained from Eq.(11).

Assuming that s_n'' denotes the system state upon entering each scenario. Upon entering Scenario 1, the system is restored to an "as good as new" condition, transitioning the health state from the threshold L to s_n'' (for 0), as shown in Eq.(15):

$$P(s_n'' | \omega_1) = \begin{cases} 1, s_n'' = 0 \\ 0, \text{otherwise} \end{cases} \quad (15)$$

After Renew is executed, the system resumes production until customer demand is fulfilled. Assuming that $w_1^{(2)}$ nonconforming products are produced in the remaining demand, with its probability given by Eq.(16):

$$p(w_1^{(2)} | 0) = \binom{D - \left\lfloor \frac{T}{t_{akt}} \right\rfloor}{w_1^{(2)}} \hat{p}_{(0)}^{\ w_1^{(2)}} (1 - \hat{p}_{(0)})^{D - \left\lfloor \frac{T}{t_{akt}} \right\rfloor - w_1^{(2)}} \quad (16)$$

Once the remaining production is completed, the health state s_n'' (which is 0) transitions to s_{n+1}, as shown in Eq.(17) :

$$P(s_{n+1} | s_n'') = f(s_{n+1}, D \cdot t_{akt} - T | s_n'' = 0) \quad (17)$$

Based on the maintenance action a_n, the delivery time upon entering Scenario 1 can be determined using Eq.(18):

$$T_n(\omega_1) = T_{a_n} + D \cdot t_{akt} + T_R + T_n^{ot} \quad (18)$$

Therefore, the operational availability of the system in Scenario 1 can be shown as Eq.(19):

$$\frac{\delta - T_n^{ot} - T_{a_n} - min\{\delta - T_n^{ot} - T_{a_n} - T, T_R\}}{\delta} \quad (19)$$

In summary, the total reward for the system in Scenario 1 is expressed in Eq. (20):

$$R(\omega_1) = r \cdot D - C_R - C_{del} \cdot \max(0, T_n(\omega_1) - \delta) - C_{pena} \cdot (w_1^{(1)} + w_1^{(2)}) \quad (20)$$

Scenario 2 (ω_2): Smooth production until the end of the task period.

In this scenario, the system operates smoothly until the end of the task period. The probability of the system entering Scenario 2 at health state $s_n^{'}$ is given by Eq.(21):

$$P(\omega_2|s_n') = 1 - \int_0^{D\cdot t_{akt}} f(L,T|s_n')dT \quad (21)$$

Assuming that w_2 denotes the number of nonconforming products produced by the system under Scenario 2, with its probability given by Eq.(22) :

$$p(w_2|s_n') = \binom{D}{w_2} \hat{p}_{(s_n')}{}^{w_2}(1-\hat{p}_{(s_n')})^{D-w_2} \quad (22)$$

Since no maintenance is performed during the production period, the transition function of the system's health state from s_n' to $s_n^{''}$ is expressed as Eq.(23):

$$P(s_n''|\omega_2) = \begin{cases} 1, s_n'' = s_n' \\ 0, \text{otherwise} \end{cases} \quad (23)$$

According to Eq.(7), the health state transition function can be further expressed as Eq.(24):

$$P(s_{n+1}|s_n'') = f(s_{n+1}, D\cdot t_{akt}|s_n'' = s_n') \quad (24)$$

Based on the maintenance action a_n, the delivery time after entering Scenario 2 is calculated by Eq.(25) :

$$T_n(\omega_2) = T_{a_n} + D\cdot t_{akt} + T_n^{ot} \quad (25)$$

Therefore, the operational availability of the system can obtain the maximum, which is shown as Eq.(3).

In summary, the reward for the system upon entering Scenario 2 is obtained as Eq.(26) :

$$R(\omega_2) = r\cdot D - C_{del}\cdot max(0, T_n(\omega_2) - \delta) - C_{pena}\cdot w_2 \quad (26)$$

Using Eqs.(18) and (25), the delay time T_{n+1}^{ot} for the current task period is determined by Eq.(27) :

$$T_{n+1}^{ot} = max(0, T_n(\omega_i) - \delta) \quad (27)$$

The transition function for the delay time is described in Eq.(28) :

$$P(T_{n+1}^{ot}|T_n^{ot}) = \begin{cases} 1, T_{n+1}^{ot} = max(0, T_n(\omega_i) - \delta) \\ 0, \text{otherwise} \end{cases} \quad (28)$$

Summarizing the above, the state transition function of the system during the task period is provided by Eq.(29):

$$P(S_{n+1}|S_n, a_n) = P(S_{n+1}|S_n'') \times P(S_n''|S_n', \omega_i) \times P(\omega_i|S_n') \times P(S_n'|S_n, a_n) \quad (29)$$

Since T_n^{ot} remain unchanged, the state space of the system after entering Scenarios is $S_n'' = \{S_n''|S_n'' = (s_n'', T_n^{ot})\}$, where the non-bold S_n'' represents an element of the state space, i.e., the state of the system. Thus, $P(\omega_i|S_n')$ is equivalent to Eq.(13) and Eq.(21), while $P(S_n''|S_n', \omega_i)$ corresponds to Eq.(15) and Eq.(23). Using Eqs.(17) , (24) and (28), the state transition function $P(S_{n+1}|S_n^{''})$ for the interval between the system's entry into the scenario and the next task period is given by Eq.(30):

$$P(S_{n+1}|S_n'') = P(s_{n+1}|s_n'')\cdot P(T_{n+1}^{ot}|T_n^{ot}) \quad (30)$$

5. Solving the model by value iteration

The Markov Decision Process (MDP) is typically solved by formulating the Bellman equation, which links the state value function with the action value function. Once the optimal action for each state is identified, the optimal maintenance policy is derived by evaluating all possible system states.

5.1. Bellman optimal equation

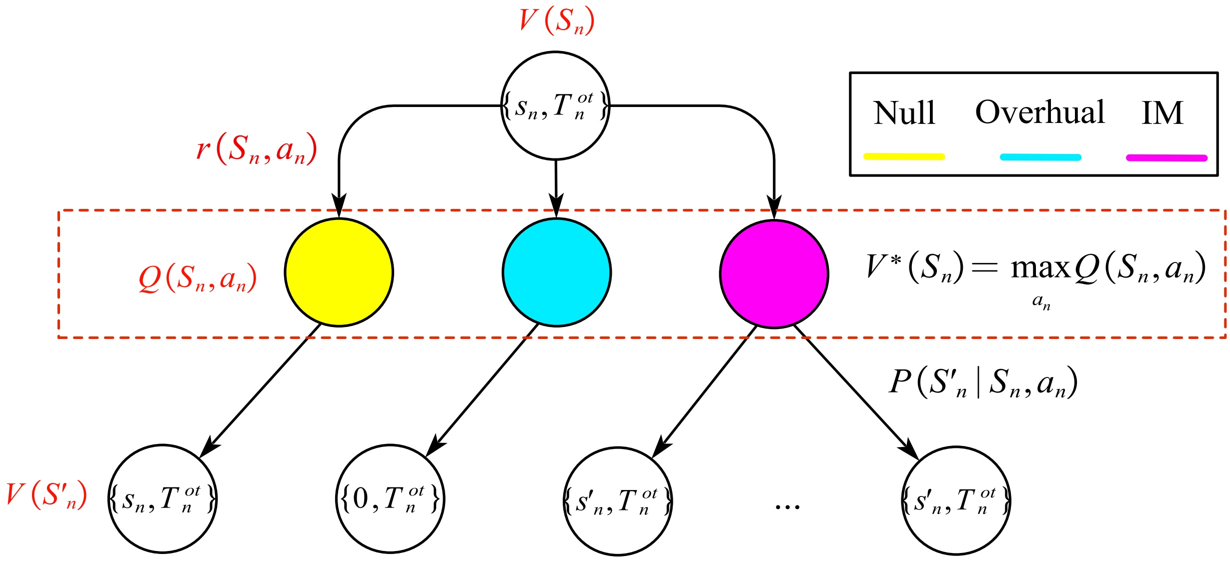


Fig.2. Bellman backup diagram of the maintenance period.

Based on the core components of the MDP, the Bellman equation is illustrated in Fig. 2 as a backup diagram for the maintenance period value function $V(S_n)$, leading to the expression given in Eq.(31):

$$V(S_n) = \sum_{a_n} \pi(a_n|S_n)\underbrace{\left(r(S_n,a_n) + \gamma \sum_{S_n'} P(S_n'|S_n,a_n)V(S_n')\right)}_{Q(S_n,a_n)} \quad (31)$$

where $V(S_n)$ represents the long-term expected reward during the nth task period when the system is in state S_n, while $Q(S_n, a_n)$ denotes the expected reward when action a_n is performed in state S_n.

Accordingly, the Bellman optimality equation is derived, as shown in Eq. (32):

$$V^*(S_n) = \max_{a_n} Q(S_n,a_n) \quad (32)$$

From this, the optimal maintenance strategy $\pi^*(S_n)$ can be determined, as shown in Eq.(33):

$$\pi * (S_n) = \underset{a_n}{argmax} Q * (S_n,a_n) \quad (33)$$

Fig.3 presents a Bellman Backup Diagram of the value function $V(S_n')$ during the production period. To further describe the system's expected performance in different operational scenarios, a scenario-based value function $Q(S_n', \omega_i)$ is introduced, as defined in Eq.(34):

$$V(S_n') = \sum_{i=1}^{2} P(\omega_i|S_n')Q(S_n',\omega_i) \quad (34)$$

where $Q(S_n', \omega_i)$ represents the long-term expected reward of the system when it enters scenario ω_i in state S_n' as shown in Eq.(35):

$$Q(S_n',\omega_i) = \overline{R(\omega_i)} + \gamma \sum_{S_{n+1}} \sum_{S_n''} P(S_n''|S_n',\omega_i)P(S_{n+1}|S_n'')V(S_{n+1}) \quad (35)$$

where $\overline{R(\omega_i)}$ denotes the expected reward in scenario ω_i, with its specific forms under Scenarios 1 and 2 derived from the expressions in Section 4.3, and given in Eqs.(36) and (37), respectively:

$$\overline{R(\omega_1)} = \sum_{w_1^{(2)}=0}^{D-\lfloor T/t_{akt}\rfloor} \sum_{w_1^{(1)}=0}^{\lfloor T/t_{akt}\rfloor} p(w_1^{(2)}|0)p(w_1^{(1)}|s_n')R(\omega_1) \quad (36)$$

$$\overline{R(\omega_2)} = \sum_{w_2=0}^{D} p(w_2|s_n')R(\omega_2) \quad (37)$$

Finally, by substituting Eq.(34) into Eq.(31), the specific form of the action value function $Q(S_n, a_n)$ is obtained, as expressed in Eq. (38):

$$Q(S_n,a_n) = r(S_n,a_n) + \sum_{S_n'} P(S_n'|S_n,a_n)\left(\sum_{i=1}^{2} P(\omega_i|S_n')Q(S_n',\omega_i)\right) \quad (38)$$

where $P(\omega_i|S_n')Q(S_n',\omega_i)$ as shown in Eq.(39)

$$\square(\square_\square|\square_\square') \cdot \square(\square_\square',\square_\square) = \begin{cases} \int_0^{\square\cdot\square_{\square\square\square}} \square(\square,\square|\square_\square')\left(\overline{\square(\square_\square)} + \square\sum_{\square_{\square+1}} \square(\square_{\square+1}|\square_\square'')\cdot\square(\square_{\square+1})\right)\square\square, \square = 1 \\ \square(\square_\square|\square_\square')\left(\overline{\square(\square_\square)} + \square\sum_{\square_{\square+1}} \square(\square_{\square+1}|\square_\square'')\cdot\square(\square_{\square+1})\right), \square = 2 \end{cases} \quad (39)$$

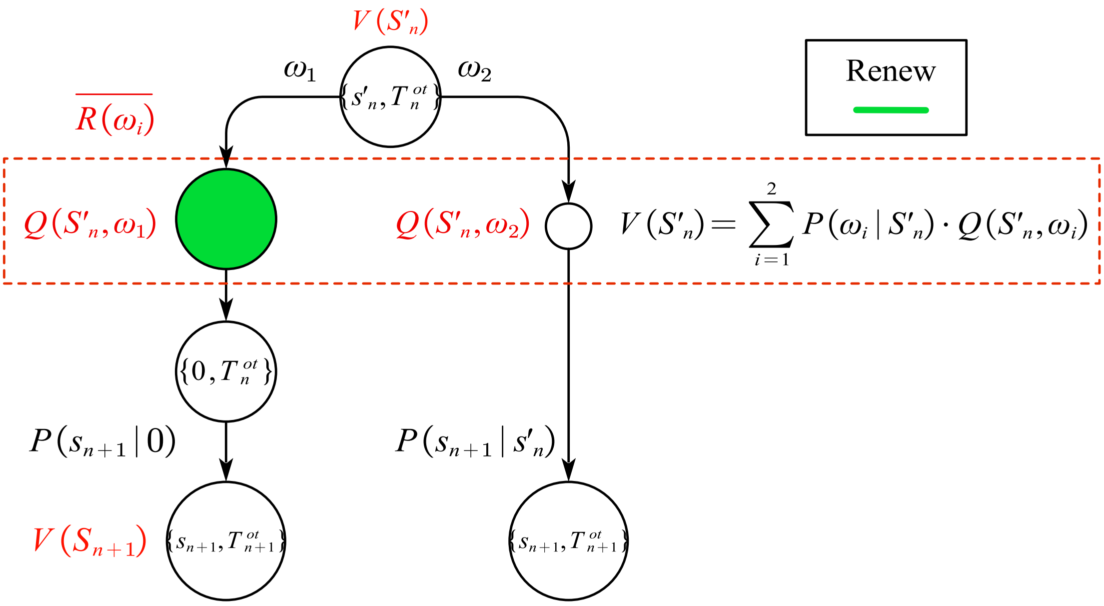


Fig.3. Bellman backup diagram of the production period.

5.2. Health state discretization

Reinforcement learning algorithms are commonly categorized into model-based and model-free types, depending on whether they rely on a model of the environment's dynamics[41] Model-based methods, such as value iteration[42] and policy iteration[43], use known or learned transition models for planning. In contrast, model-free methods like Q-learning[44] and REINFORCE[45] directly learn from interactions without modeling the environment. However, all the aforementioned algorithms are tabular and require the system state space to be enumerable[46]. in this paper, the system S_n comprises two components: the delay time T_n^{ot} from the previous task period and the health state s_n. While T_n^{ot} is a discrete variable measured in units of time, while s_n is a continuous variable governed by a Gamma process, representing the system's health evolution over time. Therefore, to apply standard reinforcement learning algorithms, the continuous health state s_n must be discretized.

Let the health state space of the system during the nth task period be denoted as $\mathbb{S}_n = \{s_n | 0 \leq s_n \leq L\}$. This interval is discretized into N equal segments, where N is a positive integer. The discretization step is defined as $\Delta = L/N$, yielding the discrete state space: $\mathbb{S}_n^{\Delta} = \{s_n^{\Delta} | 0, \Delta, 2\Delta, \ldots, L\}$, where the superscript Δ denotes the discretized form of the state.

It is worth noting that Eq.(11), which defines the smooth probabilistic relationship between the health state and the production of nonconforming products, is derived based on the continuous form of s_n. Since the approximation is more accurate in the continuous case, this equation is not discretized.

Based on the health state transition functions derived for each maintenance action in Section 4.2, this subsection will present their discretized forms. For the maintenance actions Null and Overhaul (or Renew), the discretized transition functions are expressed in Eqs.(40) and (41), respectively:

$$P(s'^{\Delta}_n | s^{\Delta}_n, a_n = \text{Null}) = \begin{cases} 1, s'^{\Delta}_n = s^{\Delta}_n \in \mathbb{S}^{\Delta}_n \\ 0, \text{otherwise} \end{cases} \tag{40}$$

$$P\left(s'^{\Delta}_n | s^{\Delta}_n, a_n = \text{Overhaul or Renew}\right) = \begin{cases} 1, s'^{\Delta}_n = 0, s^{\Delta}_n \in \mathbb{S}^{\Delta}_n \\ 0, \text{otherwise} \end{cases} \tag{41}$$

Since the health state transition function for the maintenance action IM is represented by a probability density function, it is converted into a discretized probability form by assigning the probability values within the interval $[s'^{\Delta}_n - \frac{\Delta}{2}, s'^{\Delta}_n + \frac{\Delta}{2}]$ to the discrete point s'^{Δ}_n, as shown in Eq. (42):

$$P(s'^{\Delta}_n | s^{\Delta}_n, a_n = \text{IM}) = \int_{s'^{\Delta}_n - \Delta/2}^{s'^{\Delta}_n + \Delta/2} y(s | s^{\Delta}_n, a_n = \text{IM})\, ds, s'^{\Delta}_n, s^{\Delta}_n \in \mathbb{S}^{\Delta}_n \tag{42}$$

Similarly, the health state transition functions, represented as probability density functions in Eqs.(17) and (24) for Scenarios 1 and 2, are discretized according to the different scenarios that the system may enter during the production period. These discretized forms are shown in Eqs.(43) and (44), respectively:

$$P(s^{\Delta}_{n+1} | s''^{\Delta}_n = 0) = \int_{s^{\Delta}_{n+1} - \frac{\Delta}{2}}^{s^{\Delta}_{n+1} + \frac{\Delta}{2}} f(s, D \cdot t_{akt} - T | 0)\, ds, s^{\Delta}_{n+1} \in \mathbb{S}^{\Delta}_n \tag{43}$$

$$P(s^{\Delta}_{n+1} | s'^{\Delta}_n) = \int_{s^{\Delta}_{n+1} - \frac{\Delta}{2}}^{s^{\Delta}_{n+1} + \frac{\Delta}{2}} f(s, D \cdot t_{akt} | s'^{\Delta}_n) ds,\ s'^{\Delta}_n, s^{\Delta}_{n+1} \in \mathbb{S}^{\Delta}_n \tag{44}$$

After discretizing the health state, the system's state is expressed as $S^{\Delta}_n = \{s^{\Delta}_n, T^{ot}_n\}$. Accordingly, the discretized form of the value function defined in Section 5.1 is given by Eq.(45)

$$Q(S^{\Delta}_n, a_n) = r(S^{\Delta}_n, a_n) + \sum_{S'^{\Delta}_n} P(S'^{\Delta}_n | S^{\Delta}_n, a_n) \left(\sum_{i=1}^{2} P(\omega_i | S'^{\Delta}_n) Q(S'^{\Delta}_n, \omega_i)\right) \tag{45}$$

In turn, the discretized Bellman optimal equation is derived and expressed in Eq.(46):

$$V^*(S^{\Delta}_n) = \max_{a_n} Q(S^{\Delta}_n, a_n) \tag{46}$$

According to Eq. (39), the discrete form of $P(\omega_i | S'^{\Delta}_n) Q(S'^{\Delta}_n, \omega_i)$ is obtained and presented in Eq.(47):

$$\begin{cases} \int_0^{D \cdot t_{akt}} f(L, T | s'^{\Delta}_n) \cdot \overline{R(\omega_i)}\, dT + \gamma \sum_{S^{\Delta}_{n+1}} \int_0^{D \cdot t_{akt}} P(S^{\Delta}_{n+1} | S'^{\Delta}_n) \cdot f(L, T | s'^{\Delta}_n) \cdot V(S^{\Delta}_{n+1}) dT, i = 1 \\ \left(1 - \int_0^{D \cdot t_{akt}} f(L, T | s'^{\Delta}_n)\, dT\right) \cdot \overline{(R(\omega_i)} + \gamma \sum_{S^{\Delta}_{n+1}} P(S^{\Delta}_{n+1} | S'^{\Delta}_n) \cdot V(S^{\Delta}_{n+1})), i = 2 \end{cases} \tag{47}$$

5.3. Value iteration

The value iteration algorithm is a classical tabular reinforcement learning algorithm that employs dynamic programming principles to efficiently and rapidly converge to the global optimal strategy. After discretizing the continuous and non-enumerable health states, the algorithm iteratively evaluates all possible system states to determine the optimal maintenance action for each one. This process constructs the system's optimal condition-based maintenance strategy $\pi * (S^{\Delta}_n)$, as detailed in Algorithm 1.

Algorithm 1 Value Iteration Algorithm.

Inputs: state of the system S_n^{Δ}; model parameters

Output: optimal condition-based maintenance strategy $\pi*(S_n^{\Delta})$ for the system in state S_n^{Δ}

Begin

Initialization: N-equalize $\mathbb{S}_n$; for anyS_{n+1}^{Δ}, make $V(S_{n+1}^{\Delta}) = 0$;

Calculation: Calculate$P(S'^{\Delta}_n|S_n^{\Delta}, a_n)$,$P(S_{n+1}^{\Delta}|S'^{\Delta}_n)$ and $P(\omega_i|S'^{\Delta}_n)Q(S'^{\Delta}_n, \omega_i)$ for anyS_n^{Δ};

for anyS_n^{Δ}

while $|V^M(S_n^{\Delta}) - V^{M-1}(S_n^{\Delta})| \geq \varepsilon$ **do**

Calculate $Q(S_n^{\Delta}, a_n)$ so that $V^*(S_n^{\Delta}) = \max_{a_n} Q(S_n^{\Delta}, a_n)$

end

end

$\pi*(S_n^{\Delta}) = argmax\, Q^M(S_n^{\Delta}, a_n)$

End

6. Numerical study

A commercial vehicle manufacturer utilizes a Mazak horizontal machining center HCN6800 to machine reducer housings, as illustrated in Fig. 4. To validate the proposed condition-based maintenance model, a numerical case study and comparative analysis are conducted using real-world data from this machining center.

(a) HCN 6800

(b) Reducer housings

Fig.4. Horizontal Machining Center and Reducer Housing.

6.1. Optimal strategies under the base model

Based on the company's product quality report, daily operational revenue data, and sensor-collected information, the parameters for the condition-based maintenance strategy model are determined, as shown in Table 1.

Table 1. model parameters.

parameters	value	parameters	value	parameters	value
L	10	λ_q	0.003	C_{pena}	8
D	30	γ_q	2.5	C_{del}	4
α	0.1	u	2	δ	60
β	0.9	v	2	T_{IM}	10
t_{akt}	1	r	10	T_R	$2T_O$
p_0	0.001	C_{IM}	10	T_O	20
η	0.9	C_R/C_O	30	N	50

Since the delay time T_n^{ot} is theoretically unbounded, which poses computational challenges and lacks practical feasibility, it is limited to the interval [0, δ/2] based on Table 1. The discount factor γ is set to 0.5, and the the iteration stopping threshold $\varepsilon = 10^{-6}$. Substituting the parameters into the model, all combinations of S_n^{Δ} are traversed, and the Bellman optimal equation in Section 5.2 is solved using Algorithm 1. After 6 iterations, the action value function $Q^M(S_n^{\Delta}, a_n)$ for all state combinations converges. Fig.5 shows the resulting optimal CBM strategy $\pi*(S_n^{\Delta})$.

If the system is not properly maintained during the current task period, the risk of system degradation failures during production increases. Furthermore, a large T_n^{ot} raises the likelihood of delay propagation, reducing the system's operational availability in subsequent periods and lowering the future state's expected reward $V(S_{n+1}^{\Delta})$. While maintenance actions such as IM and Overhaul improve $V(S_{n+1}^{\Delta})$ by restoring system health. However, according to Eq.(12) ,Null yields a higher immediate reward $r(S_n^{\Delta}$,Null) compared to IM, and the immediate reward $r(S_n^{\Delta}$,IM) of IM is higher than that of Overhaul. Since maintenance durations are deterministic, a large T_n^{ot} may cause the total maintenance and processing time to exceed the task period, increasing delay costs and further propagating delays. Consequently, a trade-off arises between the immediate rewards and the future state values. This trade-off is particularly evident in the following two specific state combinations, where balancing short-term and long-term benefits is critical.

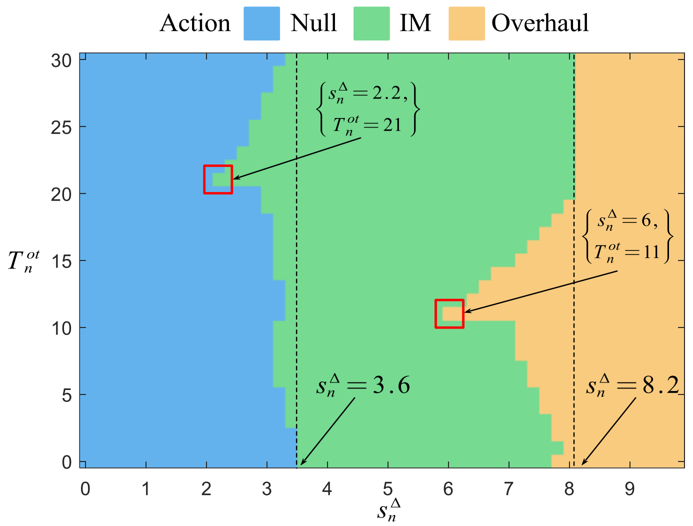


Fig.5. Optimal Condition-Based Maintenance Strategy $\pi^*(S_n^{\Delta})$.

IM's first execution point:$\{s_n^{\Delta}=2.2, T_n^{ot}=21\}$

As shown in Eq.(48), based on the model parameters, when $T_n^{ot}=21$ and the Null action is executed, if a degradation failure occurs in the system during processing, Renew will be performed. This results in the delay being propagated precisely to the (n+2)-th period, causing a further reduction in the expected future reward associated with Scenario 1.

$$T_n^{ot}+T_R+2D\cdot t_{akt}=21+40+60=121>2\cdot\delta=120 \quad (48)$$

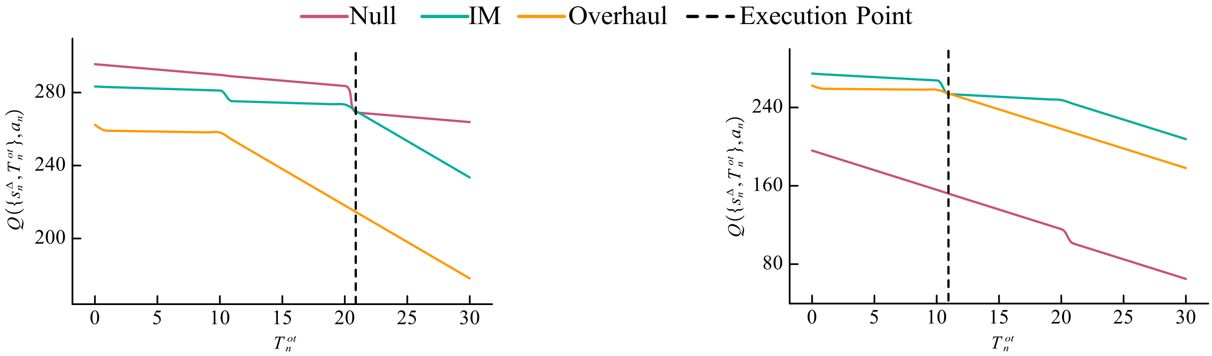
The probability of a healthy system experiencing a degradation failure (Scenario 1) during the processing period is low. Consequently, the future expected reward of Scenario 1 represents only a small percentage of the total expected reward, and changes in T_n^{ot} have minimal impact on the overall expected reward. Therefore, the optimal maintenance strategy $\pi*(S_n^{\Delta})$ is to execute Null when $s_n^{\Delta}<2.2$, regardless of T_n^{ot}. However, when the system state reaches $s_n^{\Delta}=2.2$ and Null is executed, the probability of the system entering Scenario 1 during the processing period increases. Based on this, at $T_n^{ot}=21$, the future expected reward outweighs the immediate reward, making this combination the first execution point of IM. When $T_n^{ot}>21$, the additional time consumption caused by executing IM further delays the current period, resulting in an immediate delay cost. In this case, Null regains its advantage in the trade-off between the immediate reward and the future expected reward, becoming the optimal action again. When $s_n^{\Delta}>2.2$, the advantage of IM gradually spreads around $T_n^{ot}=21$. As the health state degrades further to $s_n^{\Delta}=3.6$, the execution range of IM expands to cover all T_n^{ot}, which is referred to as the full execution point of IM. At this point, the system no longer selects Null and consistently executes IM to prevent further degradation, regardless of the delay time. This process continues until the health state reaches $s_n^{\Delta}=6$. Fig.6(a) illustrates the action value function curve for each maintenance action with respect to delay time T_n^{ot} at a system health state $s_n^{\Delta}=2.2$.

Overhaul's first execution point:$\{s_n^{\Delta}=6, T_n^{ot}=11\}$

As shown in Eq.(49), based on the base model parameters, when $T_n^{ot}=11$ and IM is executed, if the system still experiences a degradation failure during machining, the delay propagates precisely to the (n+2)-th period, resulting in a further reduction in the future expected reward associated with Scenario 1.

$$T_{\mathrm{IM}}+T_n^{ot}+T_R+2D\cdot t_{akt}=10+11+40+60=121>2\cdot\delta=120 \quad (49)$$

Generally, the probability of a system that has executed IM experiencing a degradation failure (Scenario 1) during the production period is low. Consequently, the future expected reward of Scenario 1 accounts for only a small portion of the total expected reward, and changes in T_n^{ot} have a limited impact on the overall expected reward. Therefore, when $s_n^{\Delta}\in[3.6,6)$, the optimal maintenance strategy $\pi*(S_n^{\Delta})$ is IM, regardless of T_n^{ot}. However, when the system health state reaches $s_n^{\Delta}=6$,the probability of entering Scenario 1 during the processing period increases, even if IM is performed. Based on this, at $T_n^{ot}=11$, the future expected reward outweighs the immediate reward, making this combination the first execution point of Overhaul. When $T_n^{ot}>11$, the time required for Overhaul further increases the delay in the current period, leading to an immediate delay cost. In such cases, IM regains its advantage in the trade-off between immediate and future rewards, becoming the optimal action again. As $s_n^{\Delta} > 6$, the execution region of Overhaul gradually expands around $T_n^{ot} = 11$. When the system health state degrades to $s_n^{\Delta} = 8.2$, Overhaul reaches its full execution point. Specifically, for $s_n^{\Delta} \in [8.2, 10)$, the system chooses Overhaul to restore its optimal health state, irrespective of T_n^{ot}. Fig.6(b) illustrates the action value function curve for each maintenance action with T_n^{ot} at $s_n^{\Delta} = 6$.

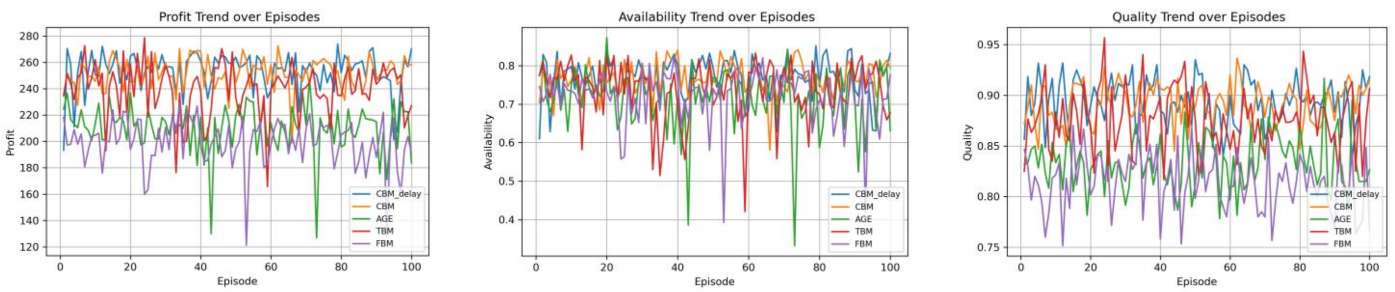


(a) $s_n^{\Delta} = 2.2$ (b) $s_n^{\Delta} = 6$

Fig.6. Action value function curves at the point of first execution.

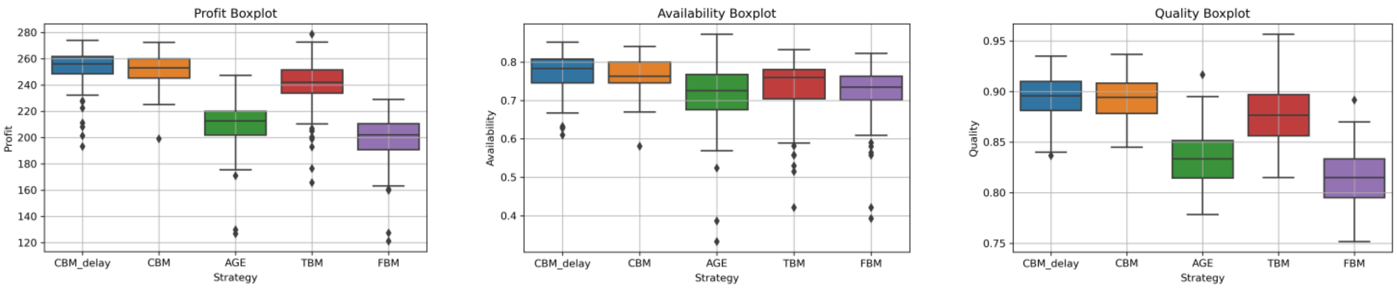
6.2. Comparative analysis of maintenance strategies

This subsection compares the optimal condition-based maintenance strategy proposed in this paper, denoted as CBM_delay, with several benchmark strategies, including failure-based maintenance (FBM), time-based maintenance (TBM), a baseline CBM strategy without considering delay propagation (denoted as CBM), and an age-based strategy (AGE). Under FBM, no preventive maintenance is performed. The system operates continuously until a shutdown occurs, after which a Renew action is executed. This approach is equivalent to consistently selecting the Null action for all state combinations. For TBM, preventive maintenance is performed on the manufacturing system at fixed intervals. If a downtime occurs during the production process, a Renew operation is executed. For CBM, the strategy is designed following existing studies that determine actions solely based on the system's health state[13, 16, 47]. Specifically, when the system health state is within [0, 3.2], Null is performed; when it is within [3.4, 7.8], IM is applied; and when it exceeds 7.8, an Overhaul is executed. For AGE, the strategy performs IM at fixed intervals based on the system's age, following the widely adopted industrial practice of applying preventive maintenance once a machine's age exceeds a predetermined threshold[48]. Based on the cost structure and degradation parameters defined in this study, the optimal age threshold is estimated to be 50, which is calculated using the formula proposed in [48]. Using the same environment and parameters as in Section 6.1, a comparative simulation test of 100 episodes was conducted for the manufacturing system using the five aforementioned strategies, with each episode consisting of 20 consecutive task periods. The trends of profit, availability, and quality for each strategy during operation are shown in Figure 7.To visually compare the performance of the strategies in terms of Profit, Availability, and Quality, Figure 8 presents the boxplots for the five strategies.



(a) Profit (b) Availability (c) Quality

Fig.7 Comparative Performance Trends Under Strategies



(a) Profit (b) Availability (c) Quality

Fig.8. Boxplots of Strategies on Profit, Availability, and Quality.

In the boxplots, the boxes represent the interquartile range (Q1 to Q3), the center line indicates the median, and the whiskers and outliers reflect the full range and rare cases. The distributions across Profit, Availability, and Quality reveal distinct differences among the five maintenance strategies. CBM_delay achieves the best performance overall, with the highest medians across all three metrics. Its boxplots show tight interquartile ranges and minimal low-end outliers, indicating strong and stable outcomes. CBM ranks second, with slightly lower medians and more spread in availability and quality, though it still maintains high overall effectiveness. TBM follows in third place. While its medians are not as high as CBM or CBM_delay, it maintains moderate performance across all metrics, with relatively narrow ranges and few extreme values, indicating consistent but less optimal outcomes. AGE, though a structured policy, underperforms TBM in all three indicators. Its distributions also show greater variability, especially in quality. FBM performs the weakest, with the lowest medians in every category and greater dispersion, particularly in profit and availability, pointing to high risk and inconsistent returns. In summary, CBM_delay achieves higher profit, more stable availability, and superior quality, confirming the superiority of the proposed strategy.

7. Conclusion

In this paper, we propose a condition-based maintenance strategy for imperfect manufacturing systems that aims to consider product quality and examine the impact of operational availability under delay propagation on the optimal maintenance decision. The strategy is modeled using a Markov decision process that accounts for system degradation failures and quality deterioration during production, with the production process simulated via a Markov-modulated Bernoulli process. To derive the strategy, we establish a scenario value function to support the Bellman optimal equation and discretize continuous health state variables, applying a value iteration algorithm to derive the optimal maintenance policy. A numerical case study and comparative analysis are conducted using data from the HCN6800 horizontal machining center of a commercial vehicle manufacturer. The results demonstrate that the proposed strategy outperforms traditional methods and condition-based maintenance strategies that ignore delay propagation, achieving better performance in terms of long-term profit, operational availability, and product quality.

Future research could expand the model to multi-machine manufacturing systems or refine it for multi-component systems, enhancing the model's applicability and realism.

References

1. Rausand M, Høyland A. System Reliability Theory: Models, Statistical Methods, and Applications2003.
2. Rivera-Gomez H, Gharbi A, Kenne JP, Montano-Arango O, Hernandez-Gress ES. Subcontracting strategies with production and maintenance policies for a manufacturing system subject to progressive deterioration. Int J Prod Econ. 2018;200:103-18. https://doi.org/10.1016/j.ijpe.2018.03.004
3. Taleizadeh AA, Sari-Khanbaglo MP, Cárdenas-Barrón LE. Outsourcing Rework of Imperfect Items in the Economic Production Quantity (EPQ) Inventory Model With Backordered Demand. IEEE Trans Syst Man Cybern -Syst. 2019;49:2688-99. https://doi.org/10.1109/TSMC.2017.2778943
4. Lai XF, Chen ZX, Bidanda B. Optimal decision of an economic production quantity model for imperfect manufacturing under hybrid maintenance policy with shortages and partial backlogging. Int J Prod Res. 2019;57:6061-85. https://doi.org/10.1080/00207543.2018.1562249

5. Zhang C, Zhang YD, Dui H, Wang SP, Tomovic MM. Component Maintenance Strategies and Risk Analysis for Random Shock Effects Considering Maintenance Costs. Eksploat Niezawodn. 2023;25:12. https://doi.org/10.17531/ein/162011
6. Yang XZ, He YH, Liao RY, Cai YQ, Dai W. Mission reliability-centered opportunistic maintenance approach for multistate manufacturing systems. Reliab Eng Syst Saf. 2024;241:15. https://doi.org/10.1016/j.ress.2023.109693
7. Li H, Huang C-G, Guedes Soares C. A real-time inspection and opportunistic maintenance strategies for floating offshore wind turbines. Ocean Engineering. 2022;256. https://doi.org/10.1016/j.oceaneng.2022.111433
8. Li H, Teixeira AP, Guedes Soares C. A two-stage Failure Mode and Effect Analysis of offshore wind turbines. Renewable Energy. 2020;162:1438-61. https://doi.org/10.1016/j.renene.2020.08.001
9. Zhao HS, Xu FH, Liang BT, Zhang JP, Song P. A condition-based opportunistic maintenance strategy for multi-component system. Struct Health Monit. 2019;18:270-83. https://doi.org/10.1177/1475921717751871
10. Tsao YC, Pantisoontorn A, Vu TL, Chen TH. Optimal production and predictive maintenance decisions for deteriorated products under advance-cash-credit payments. Int J Prod Econ. 2024;269:12. https://doi.org/10.1016/j.ijpe.2023.109132
11. Song MQ, Zhang YZ, Yang F, Wang XF, Guo GM. Maintenance policy of degradation components based on the two-phase Wiener process. Eksploat Niezawodn. 2023;25:12. https://doi.org/10.17531/ein/172537
12. Xu WG, Cao L. Optimal maintenance control of machine tools for energy efficient manufacturing. Int J Adv Manuf Technol. 2019;104:3303-11. https://doi.org/10.1007/s00170-018-2233-1
13. Xu J, Liang ZL, Li YF, Wang KB. Generalized condition-based maintenance optimization for multi-component systems considering stochastic dependency and imperfect maintenance. Reliab Eng Syst Saf. 2021;211:20. https://doi.org/10.1016/j.ress.2021.107592
14. Zhang Q, Liu Y, Xiang Y, Xiahou T. Reinforcement learning in reliability and maintenance optimization: A tutorial. Reliab Eng Syst Saf. 2024;251. https://doi.org/10.1016/j.ress.2024.110401
15. Zhang JL, Lam WHK, Chen BY. On-time delivery probabilistic models for the vehicle routing problem with stochastic demands and time windows. Eur J Oper Res. 2016;249:144-54. https://doi.org/10.1016/j.ejor.2015.08.050
16. Li S, Yang Z, He J, Li G, Yang H, Liu T, et al. A novel maintenance strategy for manufacturing system considering working schedule and imperfect maintenance. Comput Ind Eng. 2023;185. https://doi.org/10.1016/j.cie.2023.109656
17. Shi LX, Lv XL, He YD, He Z. Optimising production, maintenance, and quality control for imperfect manufacturing systems considering timely replenishment. Int J Prod Res. 2024;62:3504-25. https://doi.org/10.1080/00207543.2023.2241563
18. Ait-El-Cadi A, Gharbi A, Dhouib K, Artiba A. Integrated production, maintenance and quality control policy for unreliable manufacturing systems under dynamic inspection. Int J Prod Econ. 2021;236:20. https://doi.org/10.1016/j.ijpe.2021.108140
19. Zhang QL, Yang L, Duan JG, Qin JY, Zhou Y. Research on integrated scheduling of equipment predictive maintenance and production decision based on physical modeling approach. Eksploat Niezawodn. 2024;26:15. https://doi.org/10.17531/ein/175409
20. Guendouli E, Mifdal L, Dellagi S, Kibbou E, Moufki A. Integrated production-maintenance strategy considering energy consumption and recycling constraints in dry machining. Int J Adv Manuf Technol. 2024:19. https://doi.org/10.21203/rs.3.rs-3982933/v1
21. Yin ML, Angus JE, Trivedi KS. Optimal Preventive Maintenance Rate for Best Availability With Hypo-Exponential Failure Distribution. Ieee Transactions on Reliability. 2013;62:351-61. https://doi.org/10.1109/TR.2013.2256672
22. Chalabi N, Dahane M, Beldjilali B, Neki A. Optimisation of preventive maintenance grouping strategy for multi-component series systems: Particle swarm based approach. Comput Ind Eng. 2016;102:440-51. https://doi.org/10.1016/j.cie.2016.04.018
23. Lotovskyi E, Teixeira AP, Soares CG. Availability analysis of an offshore oil and gas production system subjected to age-based preventive maintenance by Petri Nets. Eksploat Niezawodn. 2020;22:627-37. https://doi.org/10.17531/ein.2020.4.6
24. Yang M, Li CB, Tang Y, Xiong MK. Availability-Oriented Maintenance Strategy of Key Equipment in Automated Production Line Considering Performance Degradation. IEEE Robot Autom Lett. 2023;8:3182-9. https://doi.org/10.1109/LRA.2023.3266174
25. An D, Lee DJ. Optimal condition-based maintenance policy considering nested conditional value-at-risk and operational availability: A case study on semiconductor manufacturing equipment. IISE Trans. 2024:12. https://doi.org/10.1080/24725854.2024.2410334
26. Chen Y, Liu Y, Xiahou T. A Deep Reinforcement Learning Approach to Dynamic Loading Strategy of Repairable Multistate Systems. IEEE Transactions on Reliability. 2022;71:484-99. https://doi.org/10.1109/TR.2020.3044596
27. Wang J, Zhu XY. Joint optimization of condition-based maintenance and inventory control for a k-out-of-n:F system of multi-state degrading components. Eur J Oper Res. 2021;290:514-29. https://doi.org/10.1016/j.ejor.2020.08.016
28. Hu JW, Sun QZ, Ye ZS. Replacement and Repair Optimization for Production Systems Under Random Production Waits. Ieee Transactions on Reliability. 2022;71:1488-500. https://doi.org/10.1109/TR.2021.3111651
29. Zhang ZS, Zhou YF, Sun Y, Ma L. Condition-Based Maintenance Optimisation Without A Predetermined Strategy Structure For A Two-Component Series System. Eksploat Niezawodn. 2012;14:120-9.
30. Zhou YF, Zhang ZS. Optimal Maintenance Of A Series Production System With Two Multi-Component Subsystems And An Intermediate Buffer. Eksploat Niezawodn. 2015;17:314-25. https://doi.org/10.17531/ein.2015.2.20
31. Tang DY, Sheng WB, Yu JS. Dynamic Condition-Based Maintenance Policy For Degrading Systems Described By A Random-Coefficient Autoregressive Model: A Comparative Study. Eksploat Niezawodn. 2018;20:590-601. https://doi.org/10.17531/ein.2018.4.10
32. Liu YL, Qian XB. Selective maintenance optimization with stochastic break duration based on reinforcement learning. Eksploat Niezawodn. 2022;24:771-84. https://doi.org/10.17531/ein.2022.4.17
33. Liao HT, Elsayed EA, Chan LY. Maintenance of continuously monitored degrading systems. Eur J Oper Res. 2006;175:821-35. https://doi.org/10.1016/j.ejor.2005.05.017
34. Zhao X, Sun JL, Qiu QG, Chen K. Optimal inspection and mission abort policies for systems subject to degradation. Eur J Oper Res. 2021;292:610-21. https://doi.org/10.1016/j.ejor.2020.11.015
35. Bertsekas D, Tsitsiklis JN. Introduction to probability: Athena Scientific; 2008.

36. Bouslah B, Gharbi A, Pellerin R. Integrated production, sampling quality control and maintenance of deteriorating production systems with AOQL constraint. Omega-Int J Manage Sci. 2016;61:110-26. https://doi.org/10.1016/j.omega.2015.07.012
37. van Noortwijk JM. A survey of the application of gamma processes in maintenance. Reliab Eng Syst Saf. 2009;94:2-21. https://doi.org/10.1016/j.ress.2007.03.019
38. Applebaum D. Lévy Processes and Stochastic Calculus, Second Edition: Springer; 2009. https://doi.org/10.1017/CBO9780511809781
39. Özekici S. Markov modulated Bernoulli process. Mathematical Methods of Operations Research. 1997;45:311-24. https://doi.org/10.1007/BF01194782
40. Maillart LM, Cassady CR, Honeycutt J. A binomial approximation of lot yield under Markov modulated Bernoulli item yield. IIE Trans. 2008;40:459-67. https://doi.org/10.1080/07408170701592507
41. Sutton RS, Barto AG. Reinforcement learning: An introduction: MIT press; 2018.
42. Bellman RE. Dynamic Programming: Princeton Landmarks in Mathematics and Physics; 2021.08. https://doi.org/10.2307/j.ctv1nxcw0f
43. Howard RA. Dynamic programming and Markov processes: Technology Press of Massachusetts Institute of Technology; 1960.
44. Watkins CJCH, Dayan P. Q-learning. Machine Learning. 1992;8:279-92. https://doi.org/10.1007/BF00992698
45. Williams RJ. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine Learning. 1992;8:229-56. https://doi.org/10.1007/BF00992696
46. Liu Y, Chen Y, Jiang T. Dynamic selective maintenance optimization for multi-state systems over a finite horizon: A deep reinforcement learning approach. Eur J Oper Res. 2020;283:166-81. https://doi.org/10.1016/j.ejor.2019.10.049
47. Sun QZ, Ye ZS, Chen N. Optimal Inspection and Replacement Policies for Multi-Unit Systems Subject to Degradation. Ieee Transactions on Reliability. 2018;67:401-13. https://doi.org/10.1109/TR.2017.2778283
48. Huang J, Chang Q, Arinez J. Deep reinforcement learning based preventive maintenance policy for serial production lines. Expert Syst Appl. 2020;160:14. https://doi.org/10.1016/j.eswa.2020.113701