ENGINEERING Management 2026, 13(1): 17–41
https://doi.org/10.1007/s42524-026-5184-5

RESEARCH ARTICLE

Jin LI, Qi LI, Wei ZHANG, Zhen HE, Shuguang HE

Condition-based maintenance for serial multistage manufacturing system considering reliability and quality over a finite horizon

Abstract Serial multistage manufacturing systems (SMMS), comprising multiple consecutive stages, are widely adopted in modern industry. At each stage, quality-related components (QRCs) refer to machine parts that directly impact product quality, while key product characteristics (KPCs) reflect both product performance and customer requirements. The quality of KPCs serves as an indicator of both product quality and machine condition, whereas the condition of QRCs reflects component health and provides early warnings of potential quality issues. Simultaneously monitoring both KPCs and QRCs across all stages is vital for ensuring system reliability and maintaining consistent product quality. To best of our knowledge, in SMMS, existing studies have primarily focused on the economic design of condition-based maintenance (CBM) strategies for either monitoring QRCs or KPCs individually, while their joint monitoring has received limited attention. To address this gap, this study proposes a cost-effective joint CBM strategy for SMMS over a finite horizon. A multivariate generalized likelihood ratio (MGLR) chart is employed to monitor the variations in KPCs, and the hazard rate of QRCs is evaluated using a proportional hazards model (PHM). Alarm scenarios and cost formulations are developed over a finite horizon, and a simulation framework is established to calculate the expected total cost (ETC). Subsequently, a simulation-based genetic algorithm is employed to minimize the ETC. Finally, the proposed strategy is validated on the four-stage scroll machining process in a compressor manufacturing system, demonstrating its effectiveness.

Keywords condition-based maintenance, statistical process control, serial multistage manufacturing system, stochastic processes, cost optimization

1 Introduction

1.1 Background and motivation

Serial multistage manufacturing system (SMMS) is a common type of discrete manufacturing system and plays a critical role in modern industry. The SMMS consists of sequentially connected stages that facilitate the flow of materials and information. The collective performance of these stages ultimately determines the quality of final product (Bera and Mukherjee, 2016; Feng et al., 2025). In SMMS, the condition of machine at each stage has a direct impact on product quality, while non-conforming items generated at one stage may adversely affect the operational condition of downstream machine. Moreover, the widespread application of sensors in manufacturing processes has greatly enhanced data acquisition, thereby providing robust support for the joint monitoring of machine condition and product quality. With the integration of machine condition data and product quality data, early indicators of quality deterioration or machine failure can be effectively identified. This enables manufacturers to implement both feedforward and feedback control strategies for timely and targeted maintenance actions (Shi, 2023). This joint monitoring approach reduces defects, rework, and scrap, leading to significant cost savings and higher economic efficiency, while improving machine stability and utilization to enhance overall

Received Jul. 9, 2025; revised Oct. 7, 2025; accepted Oct. 11, 2025

Jin LI, Zhen HE, Shuguang HE (✉)
College of Management and Economics, Tianjin University, Tianjin 300072, China
E-mail: shuguanghe@tju.edu.cn

Qi LI
School of Management, Tianjin University of Technology, Tianjin 300384, China

Wei ZHANG
School of Economics and Management, Inner Mongolia University of Technology, Hohhot 010051, China

This research was supported by the National Natural Science Foundation of China (Grant Nos. 72032005, 72231005, and 72401008).

production performance and system reliability (Boumallessa et al., 2023; Wang et al., 2025).

In each stage of SMMS, the quality-related components (QRCs) refer to components of machine that impact product quality, while the key product characteristics (KPCs) represent the critical indicators reflecting product performance and customer requirements (Lu and Zhou, 2017). In real production, relying solely on monitoring the condition of QRCs may overlook real-time variations in product quality, potentially resulting in delayed detection of non-conforming items. This weakens both the responsiveness and effectiveness of quality control. Conversely, monitoring only KPCs may fail to detect machine abnormalities in a timely manner, thereby increasing the risk of sudden failures and production interruptions. For example, in the scroll machining process of compressors, the quality control of scroll primarily relies on product sampling, while maintenance decisions are typically based on either sampled results or cumulative production counts of scroll. However, this approach overlooks the real-time condition of QRCs (such as cutting tools and fixtures). As a result, these components may deteriorate to a high-risk state even when sampled products still meet specifications, potentially leading to unexpected machine failures and unplanned downtime. The joint monitoring scheme can effectively overcome these limitations by integrating QRCs and KPCs across all stages. Real-time monitoring of KPCs ensures process stability and reduces the occurrence of non-conforming products, whereas real-time monitoring of QRCs enables the detection of degradation trends and mitigates risks in SMMS. The integration of these monitoring insights supports more precise process control and continuous optimization, ultimately enhancing system reliability and operational efficiency.

When the SMMS stops due to machine condition or product quality issues, effective maintenance of the machines is essential to restore their reliability and ensure consistent product quality. Since machine condition directly affects product quality at each stage, and the production of non-conforming items can further impair the operational condition of downstream machines, a scientific maintenance strategy is essential to prevent quality degradation and avoid cascading failures across stages. Condition-based maintenance (CBM), as a vital and widely implemented maintenance strategy, plays a crucial role in enhancing system reliability, reducing maintenance costs, and promoting intelligent manufacturing. CBM leverages real-time condition data such as vibration, acoustic emission and oil analysis to continuously assess machine health, facilitating more accurate maintenance decisions and improving cost efficiency (Alaswad and Xiang, 2017; Zheng and Zhou, 2021). Within a joint monitoring framework, simultaneous acquisition and analysis of machine condition data and product quality data allow for a more comprehensive understanding of deterioration process and quality variation across stages. Based on this framework, a joint CBM strategy is formulated to integrate machine health with product quality. By leveraging real-time information on machine condition and product quality, this approach enables proactive and targeted maintenance, effectively mitigating unexpected failures. Consequently, it reduces operational costs, enhances machine reliability and utilization, ensures consistent product quality, and significantly improves the overall economic performance of intelligent manufacturing systems.

1.2 Literature review

In recent years, numerous studies have investigated CBM strategies that incorporate product quality considerations. These studies are generally categorized into two main approaches: (1) Equipment condition assessment through product quality monitoring. This approach leverages real-time quality monitoring to infer machine condition, allowing maintenance decisions to be triggered by deterioration in product quality. (2) Quality inference from machine condition monitoring. This method establishes a quantitative relationship between machine condition and product quality. By directly monitoring the machine condition, the corresponding product quality level is inferred, and maintenance actions are subsequently determined.

In practical manufacturing settings, product quality is often prioritized as it directly influences customer satisfaction and compliance with industry standards (Liu et al., 2023). As a result, quality data, such as dimensional measurements and non-conforming rates, are routinely collected using cost-effective tools and are often required by regulations. Thus, many studies have developed CBM strategies based on quality monitoring. Statistical process control (SPC) is the most commonly used method in the literature for real-time monitoring of product quality (Montgomery, 2019; Qiu, 2013). For example, in the early years, Cassady et al. (2000) combining proposed an integrated control chart and maintenance strategy by combining an X-bar chart with an age-replacement policy, and demonstrated through simulation-based optimization that this joint approach achieved greater cost reduction than employing either method independently. Yeung et al. (2007) simultaneous proposed integrating preventive maintenance (PM) with SPC by formulating it as a partially observable Markov decision process to derive near-optimal joint policies that minimize maintenance, sampling, and quality-related costs. Further studies on early efforts to integrate CBM and SPC can be found in these works (Mehrafrooz and Noorossana, 2011; Panagiotidou and Tagaras, 2010; Xiang, 2013). In the past decade, Naderkhani and Makis (2016) proposed an economic design framework for multivariate Bayesian control charts using a semi-Markov decision process.

This method employs a dual-sampling interval strategy and jointly optimizes sample size, sampling intervals, control limits, and maintenance decisions to minimize the long-run expected cost per unit time. Bahria et al. (2019) utilized X-bar control charts to monitor the production process and developed an integrated strategy model for coordinating production-inventory management, maintenance scheduling, and process quality control in a single-unit manufacturing system. Kampitsis and Panagiotidou (2022) proposed a new double-sampling Bayesian control chart with state-dependent variable inspection frequency, which considered a single-unit process prone to operational deterioration and catastrophic failures. Tasias (2022) introduced a Bayesian model that could jointly optimize processes such as production, maintenance and quality, which consider the process has multiple correlated quality characteristics. Wan et al. (2023) proposed an integrated approach that simultaneously addresses economic production quantity, maintenance policy and quality control for an imperfect continuous flow process. Shojaee et al. (2024) proposed an integrated model combining Hotelling T^2 control chart for linear profiles, production cost evaluation, and maintenance policies to minimize total cost. For more information on assessing the machine condition through product quality monitoring, refer to the following works (Cao et al., 2025; He et al., 2019; Lv et al., 2024; Rasay et al., 2022; Shi et al., 2024; Zhang et al., 2024).

Over the past two decades, extensive research has been devoted to CBM of degrading systems through real-time machine condition monitoring (Alaswad and Xiang, 2017; Li et al., 2020; Mohamed-Larbi and Daoud, 2024; Olde Keizer et al., 2017; Zhao et al., 2025). Among these efforts, significant work has also been devoted to the joint evaluation of machine condition and product quality based on machine condition monitoring. Such evaluation typically relies on the prior establishment of a quantitative relationship between machine condition and product quality. For example, Chen and Jin (2005) proposed a general linear model to construct the effect of QRCs degradation on KPCs and the KPCs is described as a covariate to assess the failure rates of QRCs. Jin et al. (2019) proposed a dynamic quality model that characterizes the varying influence of machine degradation on product quality, with process effects represented as piecewise linear functions. Boumallessa et al. (2023) proposed a coupling function links the product quality indicator and machine degradation level for joint control of production, quality, and maintenance. Based on the general linear model proposed in Chen and Jin (2005), aiming to simultaneously improve process reliability and final product quality, Lu and Zhou (2017) introduced a novel opportunistic maintenance scheduling method for SMMS. Ye et al. (2019) considered the effect of imperfect inspection process and a new reliability model is developed based on the interaction between machine performance and product quality. Shi et al. (2023) considered a maintenance strategy that simultaneously considers machine availability, reliability and product quality. More recently, Lu and Luo (2024) modeled the influence of tool wear on product quality deterioration and proposed a dynamic CBM policy for heterogeneous-wearing tools. Zhang et al. (2025) joint proposed three joint optimization models integrating PM and product quality improvement for deteriorating manufacturing systems with quality–reliability dependency. For more on assessing machine condition and product quality by monitoring machine degradation, one can refer to these works (Cheng and Li, 2020; Han et al., 2021; Khatab et al., 2019; Li et al., 2023; Lu and Zhou, 2019; Majdouline et al., 2022; Wang et al., 2020; Wang et al., 2024).

Moreover, when it comes to the monitoring of multistage manufacturing systems, Zou and Tsung (2008) integrated the multivariate change point detection scheme with specific directional information based on a multistage state-space model. Shang et al. (2013) proposed a multivariate generalized linear mixed-effects model to characterize multistage processes with categorical data, effectively integrating physical laws and engineering knowledge. Yeganeh et al. (2024) proposed a multistage control chart tailored for healthcare data, combining statistical control charts with machine learning techniques to enhance early detection, confirmation, and patient safety. For more on multistage monitoring, refer to the these works (Ebadi and Ahmadi-Javid, 2020; Liu et al., 2025; Odom et al., 2018; Park and Lee, 2022). However, to best of our knowledge, no existing studies have simultaneously integrated product quality and machine condition monitoring within the framework of CBM for SMMS. This study aims to bridge this research gap by developing an integrated monitoring and early warning framework over a finite horizon, with the goal of enhancing machine reliability and ensuring consistent product quality in SMMS.

Table 1 presents a detailed comparison between the proposed strategy and representative studies, highlighting a more refined classification of the key differences. Evidently, CBM strategies jointly incorporating machine condition and product quality data in SMMS remain unexplored in existing literature.

1.3 Overview

This study investigates the monitoring of product quality variation and the assessment of machine hazard rates in SMMS over a finite horizon. Based on the monitoring results, it identifies the stage requiring maintenance and accordingly optimizes the maintenance strategy over a finite horizon. The main contributions of this study are summarized as follows: (1) Existing models often oversimplify the interdependence between QRCs degradation

Table 1 Review of the relevant works to this research

	Considering product quality		Considering machine condition		Data source		Multistage
	SPC	Quality loss function; Non-conforming rate	Degradation process	Failure	Product	Machine	
Bahria et al. (2019); Hadian et al. (2021); Huang et al.(2020); Salmasnia et al. (2017); Shojaee et al. (2024)	√				√		
He et al. (2019); Kampitsis and Panagiotidou (2022); Panagiotidou and Tagaras (2010); Rasay et al. (2022); Sabri-Laghaie et al. (2022); Shi et al. (2025); Tasias (2024); Tasias (2022)	√			√	√		
Bahrami et al. (2021); Bera and Mukherjee (2016); Liu et al. (2024); Shang et al. (2013); Yeganeh et al. (2024); Zou and Tsung (2008)	√				√		√
Cheng and Li (2020); Khatab et al. (2019); Lu and Zhou (2019); Lu and Luo (2024); Lu and Zhou (2017); Shi et al. (2023); Wang et al. (2024); Ye et al. (2019); Zhang et al. (2025)		√	√	√		√	√
Boumallessa et al. (2023); Wang et al. (2025)	√		√		√	√	
Proposed joint CBM strategy	√		√	√	√	√	√

and KPCs quality in SMMS (Chen and Jin, 2005; Lu and Luo, 2024; Shi et al., 2023). To address this limitation, we develop a coupled degradation-quality framework that explicitly incorporates the effect of non-conforming items on QRCs degradation and links the impact of QRCs degradation on KPCs quality through a generalized linear model. (2) Conventional monitoring approaches in SMMS typically only consider KPCs quality and QRCs degradation separately, which limits their ability to accurately capture QRCs hazard rates and KPCs quality variations simultaneously (Khatab et al., 2019; Liu et al., 2023). To overcome this shortcoming, we propose an integrated joint monitoring and diagnosis scheme that enables simultaneous monitoring and stage diagnosis, thereby improving responsiveness to abnormal changes in both QRCs hazard rates and KPCs quality variations. (3) Existing CBM strategies often overlook the joint feedback effects of maintenance actions on both KPCs quality and QRCs condition (Bahria et al., 2019; Zhang et al., 2025). In contrast, we formulate a joint CBM strategy in which maintenance actions are triggered by joint monitoring results, and their effects on QRCs hazard rates and KPCs quality variations are quantitatively evaluated. (4) Previous studies have rarely addressed the optimization of joint CBM strategies for SMMS over a finite horizon (Lu and Zhou, 2019; Lu and Luo, 2024; Shi et al., 2023). This study fills that gap by simulating the total cost over a finite horizon and optimizing the joint CBM strategy, thereby ensuring cost-effective decision-making and demonstrating the practical feasibility of the proposed framework for SMMS. This research offers a new approach to quality control and machine management in manufacturing. It enables intelligent monitoring, accurate diagnosis, and optimized decision-making, thereby enhancing system reliability and cost efficiency.

The remainder of this paper is organized as follows. Section 2 presents the problem statement, including the presentation of SMMS and the assumptions used in this study. Section 3 models the degradation of QRCs and KPCs quality interactions of SMMS. Section 4 models the proposed joint CBM scheme over a finite horizon, including the modeling of system hazard rate, MGLR chart and maintenance effect. In Section 5, we give the different scenarios of alarm results and derive the expression of expected total cost (ETC). The simulation-based optimisation of ETC over a finite horizon is introduced. In Section 6, the applicability of proposed strategy is validated by a four-stage scroll machining process. Finally, some concluding remarks are given in Section 7.

2 Problem statement

We consider an SMMS operating over a finite time horizon T, consisting of N consecutive stages, where each stage is equipped with a single machine. In each stage, the machine comprises p QRCs, while the product produced exhibits q KPCs. The degradation of QRCs directly impacts the corresponding KPCs at the same stage, and this relationship is referred to as the degradation-quality (D-Q) effect. Additionally, non-conforming items from upstream stages can adversely affect the condition of downstream QRCs, and accelerating the degradation of QRCs. This relationship is referred to as the quality-degradation (Q-D) effect. Figure 1 schematically illustrates the interactions between degradation of QRCs and KPCs across the stages of SMMS. To monitor both QRCs condition and KPCs quality, a periodic inspection scheme is employed with a fixed interval h. At each inspection epoch $t_j = jh, j = 0, 1, \ldots, \left\lfloor \frac{T}{h} \right\rfloor$, the degradation of QRCs and the stability of KPCs at each stage are systematically

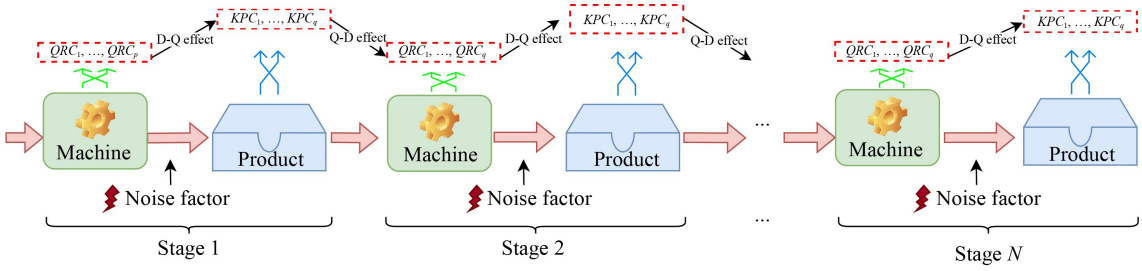


Fig. 1 Interaction of degradation and quality in SMMS.

assessed. A PM action is initiated whenever the QRCs at any stage enter an undesired condition or the stability of KPCs exceeds the defined tolerance levels.

In SMMS, checking and repairing machines at each stage when an alarm occurs is time-consuming and costly, and it often leads to unnecessary resource consumption. To address this issue, this study proposes a diagnostic strategy based on alarm signals. By leveraging these signals, the strategy identifies the machine that has the greatest impact on system stability and reliability, enabling targeted maintenance action. This diagnostic strategy facilitates the timely restoration of the health of QRCs and mitigates further quality deterioration, thereby preserving the performance of the SMMS. Moreover, some assumptions are presented to facilitate further elaborations:

Assumption 1: It is assumed that the number of non-conforming items propagated to the downstream stage follows a Poisson process. The effects of the dimensional deviations of these non-conforming items on QRCs degradation are characterized by mutually independent truncated Normal distributions. This assumption is consistent with that adopted in (Ye et al., 2019).

Assumption 2: The inherent degradation process driven by their operational conditions is assumed to be independent of the degradation process caused by the influence of non-conforming items, and all the degradation level is initialized to zero at time 0. This assumption is consistent with the literature (Cao et al., 2020; Kong et al., 2017; Ye et al., 2019).

Assumption 3: It is assumed that the hazards rate of QRCs at each stage is mutually independent, and that the effect of QRCs degradation caused by non-conforming items on the hazards rate of downstream QRCs is negligible. This assumption of hazards rate independence in SMMS is widely adopted in the literature (Jafari et al., 2018; Lu and Zhou, 2019; Shi et al., 2023).

Assumption 4: It is assumed that the state of KPCs at each stage can be categorized into two distinct conditions: in control (IC) and out of control (OC). This classification aligns with the piecewise characterization of process quality states proposed by (Jin et al., 2019), which effectively captures the dynamic shifts in quality state during manufacturing.

Assumption 5: When the control chart signals an alarm while the process is actually IC, the alarm is considered a false alarm, which incurs false cost but does not cause downtime. After the process state is verified, production continues without interruption. This assumption reflects practical operational procedures aimed at minimizing unnecessary stoppages and is widely adopted in the SPC (Bahria et al., 2019; He et al., 2019; Kampitsis and Panagiotidou, 2022; Shojaee et al., 2024).

The notations involved in this study are summarized in Table 2.

3 Modeling degradation-quality interactions

3.1 Influence of non-conforming items on QRCs degradation

In this section, the QRCs degradation process is decomposed into two contributing factors: the inherent degradation driven by their operational conditions and the degradation caused by the influence of non-conforming items propagated from upstream stages. At stage k, the inherent degradation of QRCs driven by their operational conditions, are represented by $\boldsymbol{D}_t^k = (D_{1,t}^k, \ldots, D_{p,t}^k)^\top$. The inherent degradation of QRCs is modeled by p independent Gamma processes. The Gamma process is appropriate for degradations such as tool wear, bearing fatigue, and thermal aging, which are typically irreversible and cumulative, as its strictly increasing paths can effectively capture such deterioration (Lu and Zhou, 2017; Wang et al., 2024; Ye et al., 2019; Zhang et al., 2025). Specifically, the degradation of lth QRC is described as follows

1） $D_{l,t_0}^k = 0$,

2） For $\forall t_1 < t_2 < t_3 < \cdots$, $D_{l,t_2}^k - D_{l,t_1}^k, D_{l,t_3}^k - D_{l,t_2}^k, \cdots$ are independent,

3） For $\forall t \geqslant s$, $D_{l,t}^k - D_{l,s}^k \sim Ga(\alpha_l^k(t-s), \beta_l^k)$,

where α_l^k and β_l^k are shape and scale parameters for the lth QRC of stage k, respectively.

Let θ_j^k denote the dimensional deviation of the jth non-conforming item from the specification limit, which is propagated to stage k. The θ_j^k is given by

Table 2 Notations

Symbols	Description
Abbreviations	
SMMS	Serial multistage manufacturing system
CBM	Condition-based maintenance
PM	Preventive maintenance
QRCs; KPCs	Quality-related components; Key product characteristics
D-Q effect	The degradation of QRCs impacts the quality of KPCs
Q-D effect	Non-conforming items from upstream stage impact the degradation of downstream QRCs
MGLR	Multivariate generalized likelihood ratio
PHM	Proportional hazards model
IC; OC	In control; out of control
ETC; $\mathrm{ETC}_{\min}$	The expected total cost; The optimal expected total cost
Process parameters	
N	Number of stages
T	Finite time horizon of SMMS
h	Inspection interval
n	Sample size of KPCs
$p;q$	Number of QRCs at each stage; Number of KPCs at each stage
ζ_l	The contribution weight of upstream non-conforming items to the degradation of the lth QRC
$a_l^k;b_l^k;C_l^k$	The effect coefficient of the 1th QRC degradation process, the effect coefficient of noise variables (e.g., man, material), and their interaction effect coefficient at stage k
ω^k	The effects of degradation on hazard rate at stage k
e	Age reduction coefficient of QRCs
α	Probability of type I error
δ	Shift magnitude of the random error covariance matrix
$l_{\mathrm{pm}};l_{i,S_1}^k;l_{i,S_2}^k;l_{i,S_3}^k$	Index of PM actions; Indices of PM actions for S_1, S_2, and S_3 at stage k
Stochastic variables	
$D_{l,t}^k$	The inherent degradation process experienced by the lth QRC in stage k at time t, which is solely driven by its own operational conditions
$Q_{l,t}^k$	The degradation process caused by the influence of non-conforming items on the lth QRCs in stage k at time t
$X_{l,t}^k$	The total degradation process of the lth QRC in stage k at time t, subject to both the Q-D effect and the D-Q effect
$Y_{l,t}^k$	The lth KPC value in stage k at time t
$f_{l,t}^k;\epsilon_{l,t}^k$	Conditional expectation of $Y_{l,t}^k$ given $X_{l,t}^k$; Random error of the lth KPC in stage k at time t
z_t^k	Noise variables in stage k at time t
$\theta_j^k;s(\theta_{j,k})$	Dimensional deviation of the jth non-conforming item propagated to stage k; its impact on the degradation of QRCs at stage k
N_t^k	Number of non-conforming items propagated to stage k within the time interval $[0,t]$
ξ_t^k	Cumulative impact of upstream non-conforming items on the degradation of QRCs at stage k within the time interval $[0,t]$
$M_{S_1}^k;M_{S_2}^k$	Total number of PM actions for S_1 and S_2 at stage k
$M_{S_1};M_{S_2};M$	Total number of PM actions for S_1 and S_2; Total number of PM actions over a finite horizon
$T_{l_{\mathrm{pm}}};V_{l_{\mathrm{pm}}}$	Time of l_{pm}th PM action; Virtual age of l_{pm}th PM action
$T_{l_{i,S_1}^k};T_{l_{i,S_2}^k};T_{l_{i,S_3}^k}$	Time of l_{i,S_1}^k, l_{i,S_2}^k, l_{i,S_3}^k th PM action at stage k
Parameters of stochastic variables	
$\alpha_l^k;\beta_l^k$	Shape and scale parameters of Gamma process for $D_{l,t}^k$
$\mu_{nc}^k;\sigma_{nc}^k$	Mean and standard deviation of truncated Normal distribution for $s(\theta_{j,k})$

(*Continued*)

Symbols	Description
ρ^k	Rate of non-conforming items propagated to stage k
Σ_z^k	Covariance matrix of q-variate Normal distribution for z_t^k
$\lambda^k;\eta^k$	Shape and scale parameters of Weibull distribution for baseline hazard rate
$\Sigma_0^k;\Sigma_1^k$	Covariance matrix of random error vector when process is IC; Covariance matrix of random error vector when process is OC
Cost parameters	
c_{in}	Cost of each inspection
$c_{\text{pm}};c_{\text{mm}},c_{\text{r}}$	PM cost; Minimal repair cost; Replacement cost
$c_{\text{ic}};c_{\text{oc}}$	Unit time cost of IC state; Unit time cost of OC state
c_{f}	Cost of each false alarm

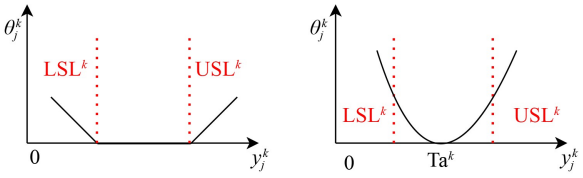
$$\theta_j^k=\begin{cases}\dfrac{y_j^k-USL^k}{USL^k}, & y_j^k>USL^k,\\ 0, & LSL^k\leqslant y_j^k\leqslant USL^k,\\ \dfrac{LSL^k-y_j^k}{LSL^k}, & y_j^k<LSL^k,\end{cases}\tag{1}$$

where USL^k and LSL^k are the upper and lower specification limits, respectively, and y_j^k is the jth observation of stage k within the inspection interval h. In SMMS, the upstream dimensional deviations of non-conforming items will accelerate the degradation of downstream machine. Let $s(\theta_j^k)$ represent the degradation increment at stage k caused by the jth non-conforming item from stage $(k-1)$. A truncated Normal distribution is employed to model the influence of the dimensional deviation of the jth non-conforming item on the degradation of QRCs, i.e., $s(\theta_j^k)\sim TN(\mu_{nc}^k,\sigma_{nc}^{k^2};0,\infty)$, where μ_{nc}^k and σ_{nc}^k are the mean and variance parameters, respectively.

Remark: In Eq. (1), the definition of θ_j^k assumes that the deviation is zero when the observed value falls within the specification limits. Once the observation exceeds these limits, the deviation is measured by the difference between the observation y_j^k and the specification limit. However, in a more general context, θ_j^k can be defined as the deviation of the observation from the target value, which is consistent with the formulation of the Taguchi quality loss function (Taguchi, 1986). Considering a target value Ta^k, the Taguchi loss function can be expressed as

$$\theta_j^k=K(y_j^k-\text{Ta}^k)^2,\tag{2}$$

where K is the loss coefficient. The definition of θ_j^k based on the Taguchi quality loss function is more general, as it implies that any deviation of KPCs from the target value contributes to QRCs degradation. Figure 2 illustrates the different definitions of the dimensional deviation θ_j^k. In this study, we adopt the definition in Eq. (1), which is reasonable in certain machining scenarios. For example, in turning operations, a workpiece is considered acceptable



(a) Defined by specification limits. (b) Defined by quality loss.

Fig. 2 Different definitions of dimensional deviation $\boldsymbol{\theta_j^k}$.

as long as its diameter remains within the specified tolerance range, even if it deviates from the nominal target value, since such deviations have a negligible impact on downstream assembly (Ye et al., 2019).

Let N_t^k denote the number of non-conforming items propagated to stage k during the time interval $[0,t]$, and $\xi_t^k=\sum_{j=1}^{N_t^k}s(\theta_j^k)$ represent the accumulated discrete degradation at stage k caused by these non-conforming items. In this study, the counting process $\{N_t^k;t\geqslant 0\}$ is modeled as a Poisson process with a rate parameter ρ^k, i.e., $N_t^k\sim\text{Poisson}(\rho^k t)$. For stage k, the degradation of QRCs caused by the influence of non-conforming items propagated from upstream stages, denoted as $\boldsymbol{Q}_t^k=(Q_{1,t}^k,\ldots,Q_{p,t}^k)^\top$. The lth QRC degradation which account for the influence of non-conforming items is defined as

$$Q_{l,t}^k=\zeta_l\times\xi_t^k,\ k=1,\ldots,N,\tag{3}$$

where the coefficient $\zeta_l\in[0,1]$ represents the weight capturing the contribution of upstream non-conforming items to the degradation of the lth QRC, with $\sum_{l=1}^p\zeta_l=1$.

Considering both the inherent degradation and the degradation induced by the influence of non-conforming items, and under the **Assumption 2**, the total degradation of QRCs $\boldsymbol{X}_t^k=(X_{1,t}^k,\ldots,X_{p,t}^k)^\top$ is given by

$$\boldsymbol{X}_t^k=\boldsymbol{D}_t^k+\boldsymbol{Q}_t^k,\ k=1,\ldots,N,\tag{4}$$

where $\boldsymbol{D}_t^k=(D_{1,t}^k,\ldots,D_{p,t}^k)^\top$ and $\boldsymbol{Q}_t^k=(Q_{1,t}^k,\ldots,Q_{p,t}^k)^\top$.

3.2 Influence of QRCs degradation on downstream KPCs

In modern manufacturing systems, the condition of machines has become an increasingly critical factor in determining product quality. As machine operate over time, its QRCs gradually degrade, which often affects the corresponding KPCs quality. Let $\boldsymbol{Y}_t^k = (Y_{1,t}^k, \ldots, Y_{q,t}^k)^\top$ denote the KPCs value vector in stage k at time t, with its observation abbreviated as $\boldsymbol{y}_t^k$. The observation of $\boldsymbol{X}_t^k$ is abbreviated as $\boldsymbol{x}_t^k$. The relationship between $\boldsymbol{X}_t^k$ and $\boldsymbol{Y}_t^k$ can be modeled by

$$\boldsymbol{Y}_t^k = \boldsymbol{f}(\boldsymbol{X}_t^k) + \boldsymbol{\varepsilon}_t^k,\ k = 1, 2, \ldots, N, \tag{5}$$

where $\boldsymbol{\varepsilon}_t^k = (\varepsilon_{1,t}^k, \varepsilon_{2,t}^k, \ldots \varepsilon_{q,t}^k)^\top$ is the vector of random errors for stage k and $\boldsymbol{\varepsilon}_t^k \sim \boldsymbol{N}_q(0, \boldsymbol{\Sigma}_0^k)$. The function $\boldsymbol{f}(\cdot)$ characterizes the relationship between $\boldsymbol{X}_t^k$ and $\boldsymbol{Y}_t^k$, and for notational simplicity, $\boldsymbol{f}(\boldsymbol{X}_t^k)$ is abbreviated as $\boldsymbol{f}_t^k$. Specifically, $\boldsymbol{f}_t^k = (f_{1,t}^k, \ldots, f_{q,t}^k)$ represents the conditional expectation of $\boldsymbol{Y}_t^k$ given $\boldsymbol{X}_t^k$, where each element is defined as $f_{i,t}^k = E(Y_{i,t}^k \mid \boldsymbol{X}_t^k = \boldsymbol{x}_t^k), i = 1, \ldots, q$.

Moreover, with the development of intelligent manufacturing technology, the role of these non-machine factors (e.g., man, material, method, measurement and environment) on the quality is gradually weakened, and they are referred to as the noise factors $\boldsymbol{z}_t^k$. Considering the interaction between the QRCs degradation $\boldsymbol{X}_t^k$ and the noise factors $\boldsymbol{z}_t^k$ (Chen and Jin, 2005), the general linear model is assumed as

$$\boldsymbol{Y}_t^k = \varphi^k + \boldsymbol{a}^k \boldsymbol{X}_t^k + \boldsymbol{b}^k \boldsymbol{z}_t^k + \boldsymbol{X}_t^{k\top} \boldsymbol{C}^k \boldsymbol{z}_t^k,\ k = 1, ..., N, \tag{6}$$

where $\varphi^k = (\varphi_1^k, \ldots, \varphi_q^k)^\top \in \mathbb{R}^{q\times 1}$, $\boldsymbol{a}^k = (\boldsymbol{a}_1^k, \ldots, \boldsymbol{a}_q^k)^\top \in \mathbb{R}^{q\times p}$, $\boldsymbol{b}^k = (\boldsymbol{b}_1^k, \ldots, \boldsymbol{b}_q^k)^\top \in \mathbb{R}^{q\times q}$ and $\boldsymbol{C}^k = (\boldsymbol{C}_1^k, \ldots, \boldsymbol{C}_q^k)^\top \in \mathbb{R}^{q\times p\times q}$. $\boldsymbol{a}^k$ and $\boldsymbol{b}^k$ represent the effect coefficient of $\boldsymbol{X}_t^k$ and $\boldsymbol{z}_t^k$, respectively; $\boldsymbol{C}^k$ represents the effect coefficient of interactions between $\boldsymbol{X}_t^k$ and $\boldsymbol{z}_t^k$.

Proposition 1 When $\boldsymbol{z}_t^k \sim \boldsymbol{N}_q(0, \boldsymbol{\Sigma}_z^k)$, we have the following results:

(1) For ith KPC of stage k, the conditional expectation $f_{i,t}^k = \varphi_i^k + \boldsymbol{a}_i^k \boldsymbol{x}_t^k$.

(2) For stage k, the random errors $\boldsymbol{\varepsilon}_t^k$ follow the multivariate normal distribution, i.e., $\boldsymbol{\varepsilon}_t^k \sim N_q(0, \mathrm{diag}((\boldsymbol{b}_1^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_1^k)\Sigma_z^k(\boldsymbol{b}_1^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_1^k)^\top, ..., (\boldsymbol{b}_q^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_q^k)\Sigma_z^k(\boldsymbol{b}_q^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_q^k)^\top))$, where $\boldsymbol{b}_i^k \in \mathbb{R}^{1\times q}$ and $\boldsymbol{C}_i^k \in \mathbb{R}^{p\times q}, i = 1, \ldots, q$.

Proof: For (1), according to Eq. (6), we get $f_{i,t}^k = E(Y_{i,t}^k | \boldsymbol{X}_t^k = \boldsymbol{x}_t^k) = \varphi_i^k + \boldsymbol{a}_i^k \boldsymbol{x}_t^k + \boldsymbol{b}_i^k \mathrm{E}(\boldsymbol{z}_t^k) + \boldsymbol{x}_t^k \boldsymbol{C}^k \mathrm{E}(\boldsymbol{z}_t^k)$. Due to $\mathrm{E}(\boldsymbol{z}_t^k) = 0$, we get $f_{i,t}^k = \varphi_i^k + \boldsymbol{a}_i^k \boldsymbol{x}_t^k$. This proves (1). According to (1), Eqs. (5) and (6), we have $\varepsilon_{i,t}^k = (\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\boldsymbol{z}_t^k$. Thus, $\varepsilon_{i,t}^k \sim N(0, (\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top)$. Since the $\varepsilon_{i,t}^k, i = 1, \ldots, p$ are independent of each other, we get $\boldsymbol{\varepsilon}_t^k \sim N_q(0, diag((\boldsymbol{b}_1^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_1^k)\Sigma_z^k(\boldsymbol{b}_1^k + x_t^{k\top}\boldsymbol{C}_1^k)^\top, ..., (\boldsymbol{b}_q^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_q^k)\Sigma_z^k(\boldsymbol{b}_q^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_q^k)^\top))$. This proves (2).

Remark: In SMMS, the quality of KPCs at each stage is continuously influenced by the degradation of QRCs. First, according to (1) in **Proposition 1**, as the degradation of QRCs accumulates across stages, the conditional expectation tends to increase. This may lead to greater quality uncertainty and a higher likelihood of exceeding tolerance limits, thereby increasing the number of non-conforming items. Second, according to (2) in **Proposition 1**, the progressive degradation of QRCs also leads to increasing variation of KPCs. As a result, the KPCs gradually deviate from their nominal targets, ultimately undermining the stability and consistency of the final product.

As established in (2), the term $\boldsymbol{x}_t^k$ within the quadratic form is a key role in increasing the variance $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top$. Suppose that l_i^k is the acceptable variance threshold of $\varepsilon_{i,t}^k$. When $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top \leqslant l_i^k$, the shift magnitude of variance is small and does not differ much each. When $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top > l_i^k$, the shift magnitude of variance gradually becomes larger. When the variance of all the random errors in each stage does not exceed the corresponding acceptable threshold, i.e., $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top \leqslant l_i^k, \forall i = 1, \ldots, q$, the process is considered to be IC and the IC covariance matrix of the random errors $\boldsymbol{\varepsilon}_t^k$ is denoted by $\boldsymbol{\Sigma}_0^k$. Conversely, if there exists at least one random error term whose variance exceeds the acceptable threshold in its corresponding stage, i.e., $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^\top > l_i^k, \exists i = 1, \ldots, q$, the process is considered as the OC state and the OC covariance matrix of the random errors $\boldsymbol{\varepsilon}_t^k$ is denoted by $\boldsymbol{\Sigma}_1^k$.

4 Modeling of the joint CBM strategy

In this section, a novel quality and reliability oriented joint CBM strategy is proposed to facilitate the PM decision-making. The PHM is developed to monitor the hazard rate of QRCs, and the MGLR control chart is employed to capture the variation of KPCs in the SMMS. Based on the monitoring results, a diagnostic strategy is developed to identify the stages requiring PM over a finite horizon, and the effect of PM actions on QRCs is quantitatively assessed.

4.1 Modeling of system hazard rate

The PHM is widely adopted in reliability engineering for its capability to capture the effects of time-varying and condition-dependent factors on machine reliability (Zheng and Zhou, 2021). This flexibility makes the PHM particularly suitable for analyzing how different degradation indicators contribute to the increasing hazard rate over time. As a result, it allows for more accurate failure predictions and the development of effective maintenance

strategies (Hu and Chen, 2020; Pedersen et al., 2023; Zheng et al., 2023). In each stage of the SMMS, the degradation states are incorporated as covariates to enable a quantitative evaluation of their impact on failure risk. Accordingly, the hazard rate function of stage k is given by

$$\begin{aligned} h_t^k &= \frac{\lambda^k}{\eta^k}\left(\frac{t}{\eta^k}\right)^{\lambda^k-1}\mathrm{E}[\exp(\omega^k \boldsymbol{X}_t^k)] \\ &= \frac{\lambda^k t^{\lambda^k-1}}{\eta^{k\lambda^k}}\prod_{l=1}^{p}(1-\omega_l^k\beta_l^k)^{-a_l^k t}\exp(\lambda(M_l^k-1)t), \end{aligned} \tag{7}$$

where $M_l^k = \exp\left(\zeta_l\omega_l^k\mu_{nc}^k + \frac{1}{2}(\zeta_l\omega_l^k\sigma_{nc}^k)^2\right)\Phi\left(\frac{\mu_{nc}^k}{\sigma_{nc}^k} + \zeta_l\omega_l^k\sigma_{nc}^k\right)\Big/\Phi\left(\frac{\mu_{nc}^k}{\sigma_{nc}^k}\right)$. The $\frac{\lambda^k}{\eta^k}\left(\frac{t}{\eta^k}\right)^{\lambda^k-1}$ represents the Weibull baseline-hazard function, where λ^k and η^k are the shape and scaleparameters for stage k, respectively. The $\mathrm{E}[\exp(\omega^k \boldsymbol{X}_t^k)]$ represents the expected impact of QRCs degradation on hazard, where $\omega^k \in \mathbb{R}^{1\times p}$ is regression coefficient vector. The parameters of hazard function in Eq. (7) can be estimated using the maximum likelihood estimation method with lifetime histories (Vlok et al., 2002). From Eq. (7), it can be seen that h_t^k is the function of t, and its derivation is given in Appendix A.

According to **Assumption 3**, each stage operates independently, meaning the hazard rate of one machine does not influence another. This independence simplifies the modeling process by enabling the hazard rates of individual machines to be evaluated separately (Lu and Zhou, 2019). Consequently, the system hazard rate function H_t of the SMMS is given by

$$H_t = \sum_{k=1}^{N} h_t^k. \tag{8}$$

The system PHM signals an alarm when $H_t > L_D$, where L_D is the threshold for system hazard rate.

4.2 MGLR chart for monitoring the covariance

In SMMS of intelligent manufacturing, the KPCs data $\boldsymbol{y}_t^k$ at each stage can be efficiently collected through advanced sensing and data acquisition technologies. The random errors $\boldsymbol{\varepsilon}_t^k = (\varepsilon_{1,t}^k, \varepsilon_{2,t}^k, \cdots, \varepsilon_{q,t}^k)^\top$ of stage k is computed according to **Proposition 1**, i.e., $\boldsymbol{\varepsilon}_t^k = \boldsymbol{y}_t^k - \boldsymbol{f}_t^k$. We regard $\boldsymbol{\varepsilon}_t^k = (\varepsilon_{1,t}^k, \varepsilon_{2,t}^k, \cdots, \varepsilon_{q,t}^k)^\top$ as the q-dimensional multivariate normal vectors and $\boldsymbol{\varepsilon}_t^k \sim \boldsymbol{N}_q(0, \boldsymbol{\Sigma}_0^k)$. If a shift of stage k occurs at time τ^k, then the covariance matrix will change from $\boldsymbol{\Sigma}_0^k$ to $\boldsymbol{\Sigma}_1^k$. The hypothesis test of stage k is designed by

$$\mathcal{H}_0 : \boldsymbol{\Sigma}^k = \boldsymbol{\Sigma}_0^k \quad vs \quad \mathcal{H}_1 : \boldsymbol{\Sigma}^k = \boldsymbol{\Sigma}_1^k. \tag{9}$$

Let a sample of size n be collected at time t. The corresponding log-likelihood function of $\boldsymbol{\Sigma}^k$, based on the sample $\boldsymbol{\varepsilon}_{t,1}^k, \ldots, \boldsymbol{\varepsilon}_{t,n}^k$, is given by (Jing et al., 2022)

$$l(\boldsymbol{\varepsilon}_t^k|\boldsymbol{\Sigma}^k) = -\frac{n}{2}(\log|\boldsymbol{\Sigma}^k| + tr(\boldsymbol{\Sigma}^{k-1}\boldsymbol{S}_t^k) + p\log(2\pi)), \tag{10}$$

where $|A|$ is the determinant of matrix A, $\mathrm{tr}(A)$ represents the trace of A, and $\boldsymbol{S}_t^k = \sum_{j=1}^{n}\boldsymbol{\varepsilon}_{t,j}^{k\ T}\boldsymbol{\varepsilon}_{t,j}^k/n$. If $t < \tau^k$, all the samples up to time t follow the IC distribution and the log-likelihood is

$$l_{\infty,t}(\boldsymbol{\Sigma}_0^k) = \sum_{i=1}^{t} l(\boldsymbol{\varepsilon}_i^k|\boldsymbol{\Sigma}_0^k). \tag{11}$$

If $t \geqslant \tau^k + 1$, then the log-likelihood is

$$l_{\tau^k,t}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k) = \sum_{i=1}^{\tau^k} l(\boldsymbol{\varepsilon}_i^k|\boldsymbol{\Sigma}_0^k) + \sum_{i=\tau^k+1}^{t} l(\boldsymbol{\varepsilon}_i^k|\boldsymbol{\Sigma}_1^k). \tag{12}$$

The shift time τ^k can be estimated by maximizing the log-likelihood $l_{\tau^k,t}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k)$ and is given by

$$\hat{\tau}^k = \max_{1\leqslant\tau^k\leqslant t} l_{\tau^k,t}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k). \tag{13}$$

Considering the log-likelihood ratio for testing whether there has been a change in the process prior to the current time point, the MGLR test statistic at time t is

$$\begin{aligned} R_t^k &= \max_{1\leqslant\tau^k\leqslant t}[l_{\tau^k,t}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k) - l_{\infty,t}(\boldsymbol{\Sigma}_0^k)] \\ &= \max_{1\leqslant\tau^k\leqslant t}\frac{n(t-\tau^k)}{4}tr(\boldsymbol{\Sigma}_0^{k-1}\boldsymbol{\Sigma}_1^k - I)^2, \end{aligned} \tag{14}$$

where $\boldsymbol{\Sigma}_1^k = \sum_{i=\tau^k+1}^{t}\sum_{j=1}^{n}\boldsymbol{\varepsilon}_{i,j}^{k\ \top}\boldsymbol{\varepsilon}_{i,j}^k/n(t-\tau^k)$ can be computed under the hypothesis $\mathcal{H}_1$. The proof of Eq. (14) is given by Appendix B. Furthermore, the monitoring statistic for the covariance in SMMS at time t is given by

$$R_t = \max_{1\leqslant k\leqslant N} R_t^k. \tag{15}$$

The MGLR chart signals an alarm if $R_t > L_Q$, where L_Q is a pre-specified control limit.

4.3 The joint CBM strategy for finite horizon

Over the finite horizon T, an alarm may be triggered by the MGLR chart, by the system PHM, or by both simultaneously. If the first alarm is triggered by the MGLR chart, it reflects a significant variation in the KPCs of the SMMS. If the first alarm is triggered by the system PHM, it indicates that the hazard rate of the QRCs in the SMMS has reached an unacceptable level. If both alarms are triggered simultaneously, it implies that both a significant variation and an unacceptable hazard rate have occurred. In all cases, according to Eq. (7) and **Proposition 1**, these alarms indicate that the QRCs have degraded to an undesired level, thereby requiring maintenance action.

This study considers two types of maintenance action: minimal repair and PM. When machine fails, the minimal repair action is performed to restore it to a normal operational condition. The minimal repair action neither

changes the condition of the QRCs nor affects their deterioration states, and its duration is assumed to be negligible. When an alarm is triggered, the PM action is conducted proactively to mitigate QRCs deterioration. The PM action restores the QRCs to an as-good-as-new condition in terms of degradation levels, which are reset to zero, while leaving the age of machine unchanged. It is assumed that the degradation levels of QRCs are zero at the beginning of their service life.

Let l_{pm} denote the index of l_{pm}th PM action over a finite horizon, and let $T_{l_{\text{pm}}}$ represent the time at which this PM action is performed. Accordingly, the degradation of the QRCs after the l_{pm}th PM action is given by

$$\boldsymbol{X}_t^k = \begin{cases} 0, & \text{if } t = T_{l_{\text{pm}}}, \\ \boldsymbol{D}_{t-T_{l_{\text{pm}}}}^k + \boldsymbol{Q}_{t-T_{l_{\text{pm}}}}^k, & \text{if } t \in (T_{l_{\text{pm}}}, T_{l_{\text{pm}}+1}). \end{cases} \tag{16}$$

According to **Proposition 1**, after maintenance, both the conditional expectation $f_{i,t}^k = \varphi_i^k + \boldsymbol{a}_i^k \boldsymbol{x}_t^k$ and the variance of random errors $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top} \boldsymbol{C}_i^k)\boldsymbol{\Sigma}_z^k (\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top} \boldsymbol{C}_i^k)^\top$ at stage k are updated according to Eq. (16). Let $V_{l_{\text{pm}}+1}$ denote as the virtual age after the $(l_{\text{pm}}+1)$th PM action, which is given by

$$V_{l_{\text{pm}}+1} = V_{l_{\text{pm}}} + e(T_{l_{\text{pm}}+1} - T_{l_{\text{pm}}}), \tag{17}$$

where e is the age reduction coefficient and $e \in [0,1]$. $e = 0$ and $e = 1$ indicate that the PM action restores the QRCs condition to as good as new and does not change the QRCs condition, respectively. The evolution of hazard rate at stage k under PM action can be given by

$$h_t^k = \begin{cases} \dfrac{\lambda_k}{\eta_k^{\lambda_k}} V_{l_{\text{pm}}}^{\lambda_k - 1}, & \text{if } t = T_{l_{\text{pm}}}, \\ \dfrac{\lambda_k}{\eta_k^{\lambda_k}} (t - T_{l_{\text{pm}}} + V_{l_{\text{pm}}})^{\lambda_k - 1} \mathrm{E}\left[\exp(\omega^k \boldsymbol{X}_t^k)\right], & \text{if } t \in (T_{l_{\text{pm}}}, T_{l_{\text{pm}}+1}), \end{cases} \tag{18}$$

where $\mathrm{E}[\exp(\omega^k \boldsymbol{X}_t^k)] = \prod_{l=1}^{p} (1 - \omega_l^k \beta_l^k)^{-\alpha_l^k (t - T_{l_{\text{pm}}})} \exp(\lambda (M_l^k - 1)(t - T_{l_{\text{pm}}}))$.

However, in practical production processes, simultaneously inspecting and repairing all stages of the machine is both time-consuming and costly. To address this issue, we propose a diagnostic strategy that identifies the specific stage requiring maintenance, thereby ensuring the overall reliability and stability of the SMMS while reducing unnecessary maintenance efforts.

Proposition 2 When the joint monitoring alarms at time $T_{l_{\text{pm}}}$ in SMMS, a diagnostic strategy to determine which stage needs PM action is given by

$$k = \begin{cases} \arg\max\limits_{1 \leqslant k \leqslant N} R_{T_{l_{\text{pm}}}}^k, & \text{if } H_{T_{l_{\text{pm}}}} < L_{\text{D}} \,\&\, R_{T_{l_{\text{pm}}}} \geqslant L_{\text{Q}}, \\ \arg\max\limits_{k \in Des} |h_{T_{l_{\text{pm}}}}^k - \hat{h}_{T_{l_{\text{pm}}}}^k|, & \text{if } H_{T_{l_{\text{pm}}}} \geqslant L_{\text{D}} \,\&\, R_{T_{l_{\text{pm}}}} < L_{\text{Q}}, \\ \{\arg\max\limits_{1 \leqslant k \leqslant N} R_{T_{l_{\text{pm}}}}^k, \arg\max\limits_{k \in Des} |h_{T_{l_{\text{pm}}}}^k - \hat{h}_{T_{l_{\text{pm}}}}^k|\}, & \text{if } H_{T_{l_{\text{pm}}}} \geqslant L_{\text{D}} \,\&\, R_{T_{l_{\text{pm}}}} \geqslant L_{\text{Q}}, \end{cases} \tag{19}$$

where $\hat{h}_{T_{l_{\text{pm}}}}^k = \dfrac{\lambda_k}{\eta_k^{\lambda_k}} V_{l_{\text{pm}}-1}^{\lambda_k - 1}$ and $Des = \left\{k \mid \sum_{i=1, i \neq k}^{N} h_{T_{l_{\text{pm}}}}^i + \hat{h}_{T_{l_{\text{pm}}}}^k < L_{\text{D}}, 1 \leqslant k \leqslant N\right\}$. $\hat{h}_{T_{l_{\text{pm}}}}^k$ is the hazard rate when the PM action is taken at stage k and Des is the stage set of system hazard restoring to the desired value when PM action is taken at stage k.

In **Proposition 2**, the diagnostic scheme of the control chart identifies the stage with the highest statistic $R_{T_{l_{\text{pm}}}}^k$, which reflects the most pronounced KPCs variation and thus requires PM action. By contrast, the diagnostic scheme of the system PHM aims to keep the system hazard rate within acceptable thresholds, leading to the maintenance of the stage that yields the largest hazard rate improvement, i.e., $h_{T_{l_{\text{pm}}}}^k - \hat{h}_{T_{l_{\text{pm}}}}^k$. When alarms are simultaneously raised by both the MGLR chart and the system PHM at time $T_{l_{\text{pm}}}$, PM action is concurrently taken on the stage with the maximum $R_{T_{l_{\text{pm}}}}^k$ and on the stage providing the greatest reduction in hazard rate.

Over the finite horizon, three possible scenarios of joint monitoring alarm results may occur and are illustrated in Fig. 3(a). Specifically:

- Scenario S_1 corresponds to the MGLR chart alarms and S_1 arises under the condition $\{R_t \geqslant L_{\text{Q}} \,\&\, H_t < L_{\text{D}} \,\&\, t < T\}$. When no shift occurs over a finite horizon, i.e., the process remains IC, no action is taken and production continues as usual. Conversely, if a shift occurs prior to an alarm from the MGLR chart, a PM action is carried out.
- Scenario S_2 corresponds to the system PHM alarms and S_2 arises under the condition $\{H_t \geqslant L_{\text{D}} \,\&\, R_t < L_{\text{Q}} \,\&\, t < T\}$. PM action is performed regardless of whether a shift occurs.
- Scenario S_3 corresponds to the case where no alarm is triggered before the horizon ends and S_3 arises under the condition $\{R_t < L_{\text{Q}} \,\&\, H_t < L_{\text{D}} \,\&\, t = T\}$, and a final replacement action is taken.

Within the finite horizon, scenarios S_1 and S_2 may occur multiple times, whereas scenario S_3 occurs at most once, typically at the end of finite horizon. Let l_{i,S_1}^k, $i = 1, 2, \ldots, M_{S_1}^k$, denote the index of the ith PM action at

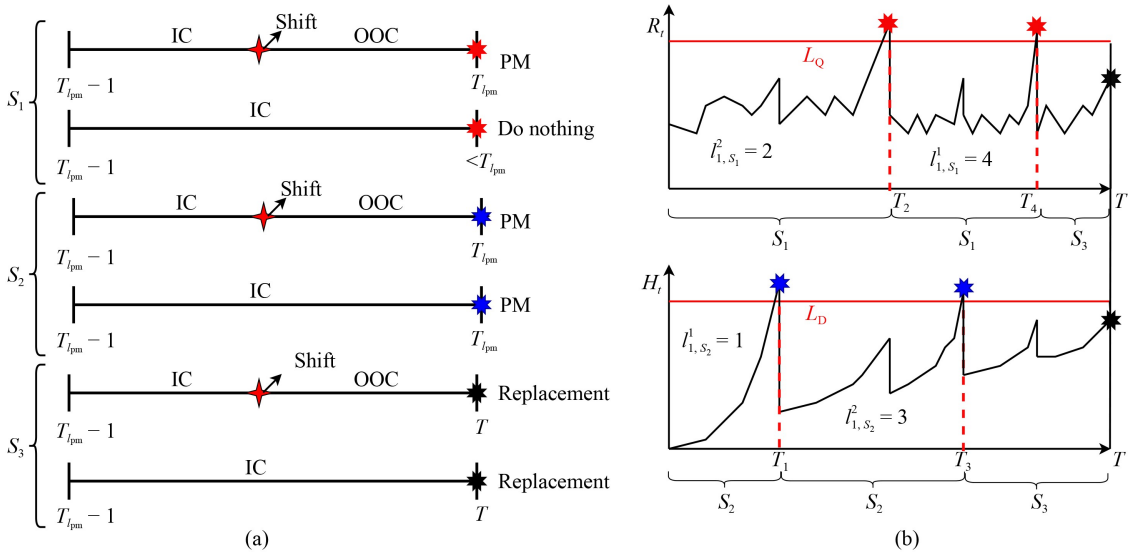


Fig. 3 Results of joint monitoring for finite horizon: (a) Scenarios of monitoring result, and (b) Example of online monitoring.

stage k associated with scenario S_1, where $M_{S_1}^k$ is the total number of such actions at stage k over a finite horizon. Let l_{i,S_2}^k, $i=1,2,\ldots,M_{S_2}^k$, denote the index of the ith PM action at stage k associated with scenario S_2, where $M_{S_2}^k$ is the total number of such actions at stage k over a finite horizon. The total number of PM actions for S_1 and S_2 across all stages are $M_{S_1}=\sum_{k=1}^{N} M_{S_1}^k$, $M_{S_2}=\sum_{k=1}^{N} M_{S_2}^k$. The total number of PM actions over a finite horizon is $M=M_{S_1}+M_{S_2}$.

Figure 3(b) illustrates an example of a two-stage SMMS ($N=2$) with a total of four PM actions ($M=4$) executed over a finite horizon. Among them, two PM actions are triggered by alarms from the MGLR chart ($M_{S_1}=2$), and the other two by the system PHM ($M_{S_2}=2$). In Stage 1, two PM actions are performed: one ($M_{S_2}^1=1$) occurs at the first PM action ($l_{1,S_2}^1=1$), and one ($M_{S_1}^1=1$) occurs at the fourth PM action ($l_{1,S_1}^1=4$). In Stage 2, two PM actions are also executed, including one ($M_{S_1}^2=1$) at the second PM action ($l_{1,S_1}^2=2$) and one ($M_{S_2}^2=1$) at the third PM action ($l_{1,S_2}^2=3$).

5 Simulation-based optimization for finite horizon

In this section, we systematically characterize the cost associated with the joint monitoring strategy over a finite horizon. Specifically, the total cost comprises four major components: inspection cost, quality-related cost, maintenance cost, and false alarm cost. We provide detailed mathematical formulations for each of these cost elements incurred over a finite horizon.

5.1 Inspection cost

When employing the proposed monitoring scheme, each inspection incurs a cost associated with acquiring the degradation data of QRCs and KPCs data. Let c_{in} denote the cost per inspection. In scenario S_1, for the l_{i,S_1}^k th PM interval of stage k, $[T_{l_{i,S_1}^k-1}, T_{l_{i,S_1}^k}]$, the number of inspection within the l_{i,S_1}^k th PM interval of stage k is $\dfrac{T_{l_{i,S_1}^k}-T_{l_{i,S_1}^k-1}}{h}$. Hence, the total inspection cost $C_I^{S_1}$ over a finite horizon of S_1 is $C_I^{S_1}=c_{\text{in}}\sum_{k=1}^{N}\sum_{i=1}^{M_{S_1}^k}\dfrac{T_{l_{i,S_1}^k}-T_{l_{i,S_1}^k-1}}{h}$.

Similarly, the total inspection cost $C_I^{S_2}$ over a finite horizon of scenario S_2 is $C_I^{S_2}=c_{\text{in}}\sum_{k=1}^{N}\sum_{i=1}^{M_{S_2}^k}\dfrac{T_{l_{i,S_2}^k}-T_{l_{i,S_2}^k-1}}{h}$. Scenario S_3 represents the period from last maintenance time T_M until the end of the finite horizon T, with the number of inspection for S_3 is $\dfrac{T-T_M}{h}$. Since S_3 occurs only once over a finite horizon, the total inspection cost of S_3 is $C_I^{S_3}=c_{\text{in}}\dfrac{T-T_M}{h}$.

Thus, the total inspection cost C_{I} over a finite horizon is given by

$$C_I=p_1\times C_I^{S_1}+p_2\times C_I^{S_2}+C_I^{S_3}, \tag{20}$$

where $p_1=\dfrac{M_{S_1}}{M}$ and $p_2=\dfrac{M_{S_2}}{M}$ denote the occurrence probabilities of scenarios S_1 and S_2, respectively.

5.2 Maintenance cost

As the maintenance strategy proposed in Section 4.3, between two consecutive PM actions, two types of maintenance activities may be performed: the first is minimal repair, which is initiated upon failure and incurs a minimal repair cost c_{mm}; the second is PM, which is triggered by alarms from either the MGLR chart or the system PHM,

with an associated PM cost c_{pm}. At the end of the finite horizon, we will replace all the QRCs at each stage, with a replacement cost c_{r}.

For scenario S_1, the expected number of failures during the interval $[T_{l^k_{i,S_1}-1}, T_{l^k_{i,S_1}}]$ is $\sum_{j_1 \leqslant j \leqslant j_2} \int_{t_j}^{t_{j+1}} h_t^k \mathrm{d}t$, where $j_1 = T_{l^k_{i,S_1}-1}/h$ and $j_2 = T_{l^k_{i,S_1}}/h$ denote the inspection time points corresponding to the $(l^k_{i,S_1}-1)$th and l^k_{i,S_1}th PM actions, respectively. Therefore, the total maintenance cost $C_M^{S_1}$ over a finite horizon of S_1 is given by

$$C_M^{S_1} = c_{\text{mm}} \underbrace{\sum_{k=1}^{N} \sum_{i=1}^{M_{S_1}^k} \left(\sum_{j_1 \leqslant j \leqslant j_2} \int_{t_j}^{t_{j+1}} h_t^k \mathrm{d}t \right)}_{\text{Expected number of failures under } S_1} + \underbrace{c_{\text{pm}} M_{S_1}}_{\text{PM cost under } S_1}. \tag{21}$$

Similarly, the total maintenance cost $C_M^{S_2}$ over a finite horizon of scenario S_2 is given by

$$C_M^{S_2} = c_{mm} \underbrace{\sum_{k=1}^{N} \sum_{i=1}^{M_{S_2}^k} \left(\sum_{j_3 \leqslant j \leqslant j_4} \int_{t_j}^{t_{j+1}} h_t^k dt \right)}_{\text{Expected number of failures under } S_2} + \underbrace{c_{pm} M_{S_2}}_{\text{PM cost under } S_2}, \tag{22}$$

where $j_3 = T_{l^k_{i,S_2}-1}/h$ and $j_4 = T_{l^k_{i,S_2}}/h$ denote the inspection time points corresponding to the $(l^k_{i,S_2}-1)$th and l^k_{i,S_2}th PM actions, respectively.

For scenario S_3, the maintenance cost consists of the minimal repair cost c_{mm}, as well as the replacement cost c_r. Therefore, the total maintenance cost $C_M^{S_3}$ over a finite horizon of S_3 is given by

$$C_M^{S_3} = c_{\text{mm}} \underbrace{\sum_{k=1}^{N} \sum_{j_5 \leqslant j \leqslant \lfloor \frac{T}{h} \rfloor} \int_{t_j}^{t_{j+1}} h_t^k \mathrm{d}t}_{\text{Expected number of failures under } S_3} + \underbrace{c_r}_{\text{Replacement cost}}, \tag{23}$$

where $j_5 = T_M/h$ denotes the inspection time points corresponding to the Mth PM action.

Thus, the total maintenance cost C_M over a finite horizon is given by

$$C_M = p_1 \times C_M^{S_1} + p_2 \times C_M^{S_2} + C_M^{S_3}, \tag{24}$$

where $p_1 = \dfrac{M_{S_1}}{M}$ and $p_2 = \dfrac{M_{S_2}}{M}$ denote the occurrence probabilities of scenarios S_1 and S_2, respectively.

5.3 Quality cost

In the production process, the system typically transitions from IC to OC state, with the associated costs for each state being distinctly different. We assume the unit time cost of IC state is c_{ic}, while the unit time cost of OC state is c_{oc}.

For scenario S_1, the shift time of stage k corresponding to its l^k_{i,S_1}th PM action is denoted by τ^k_{i,S_1}. The total quality cost $C_Q^{S_1}$ over a finite horizon of S_1 is given by

$$C_Q^{S_1} = \sum_{k=1}^{N} \sum_{i=1}^{M_{S_1}^k} \left[\underbrace{(\tau^k_{i,S_1} - T_{l^k_{i,S_1}-1}) c_{ic}}_{\text{IC cost of } l^k_{i,S_1}\text{th PM}} + \underbrace{(T_{l^k_{i,S_1}} - \tau^k_{i,S_1}) c_{oc}}_{\text{OC cost of } l^k_{i,S_1}\text{th PM}} \right], \tag{25}$$

and according to Eq. (13), the shift time τ^k_{i,S_1} is estimated as

$$\tau^k_{i,S_1} = \underset{T_{l^k_{i,S_1}-1} \leqslant \tau_i^k \leqslant T_{l^k_{i,S_1}}}{\arg\max} \left\{ l_{\tau_i^k, T_{l^k_{i,S_1}}}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k) \right\}. \tag{26}$$

Similarly, for scenarios S_2 and S_3, the corresponding quality costs over a finite horizon, denoted by $C_Q^{S_2}$ and $C_Q^{S_3}$, are given by

$$\begin{aligned} C_Q^{S_2} &= \sum_{k=1}^{N} \sum_{i=1}^{M_{S_2}^k} \left[\underbrace{(\tau^k_{i,S_2} - T_{l^k_{i,S_2}-1}) c_{ic}}_{\text{IC cost of } l^k_{i,S_2}\text{th PM}} + \underbrace{(T_{l^k_{i,S_2}} - \tau^k_{i,S_2}) c_{oc}}_{\text{OC cost of } l^k_{i,S_2}\text{th PM}} \right], \\ C_Q^{S_3} &= \sum_{k=1}^{N} \left[\underbrace{(\tau^k_{S_3} - T_M) c_{ic}}_{\text{Terminal IC cost}} + \underbrace{(T - \tau^k_{S_3}) c_{oc}}_{\text{Terminal OC cost}} \right], \end{aligned} \tag{27}$$

and the shift time τ^k_{i,S_2} and $\tau^k_{S_3}$ are estimated as

$$\begin{aligned} \tau^k_{i,S_2} &= \underset{T_{l^k_{i,S_2}-1} \leqslant \tau_i^k \leqslant T_{l^k_{i,S_2}}}{\arg\max} \left\{ l_{\tau_i^k, T_{l^k_{i,S_2}}}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k) \right\}, \\ \tau^k_{S_3} &= \underset{T_M \leqslant \tau^k_{S_3} \leqslant T}{\arg\max} \left\{ l_{\tau^k_{S_3}, T}(\boldsymbol{\Sigma}_0^k, \boldsymbol{\Sigma}_1^k) \right\}. \end{aligned} \tag{28}$$

Thus, the total quality cost C_Q over a finite horizon is given by

$$C_Q = p_1 \times C_Q^{S_1} + p_2 \times C_Q^{S_2} + C_Q^{S_3}, \tag{29}$$

where $p_1 = \dfrac{M_{S_1}}{M}$ and $p_2 = \dfrac{M_{S_2}}{M}$ denote the occurrence probabilities of scenarios S_1 and S_2, respectively.

5.4 False alarm cost

When the control chart signals an alarm while the process is IC state, such an event is identified as a type I error, characterized by the false alarm rate α(Montgomery, 2019; Qiu, 2013). Although production can continue after confirming that the process is indeed IC, the associated inspections may still incur costs. Let c_f denote the cost incurred by each false alarm.

For scenario S_1, the number of inspections conducted while the process is IC state during the interval $[T_{l^k_{i,S_1}-1}, T_{l^k_{i,S_1}}]$ is given by $\dfrac{\tau^k_{i,S_1} - T_{l^k_{i,S_1}-1}}{h}$. Accordingly, the number of false alarms during this interval is $\dfrac{\tau^k_{i,S_1} - T_{l^k_{i,S_1}-1}}{h}\alpha$. Thus, the total false alarm cost $C_F^{S_1}$ over a finite horizon is given by

$$C_F^{S_1}=\sum_{k=1}^{N}\sum_{i=1}^{M_{S_1}^k}\underbrace{\frac{\tau_{i,S_1}^k-T_{l_{i,S_1}^k-1}}{h}\alpha c_f}_{\text{False alarm cost of } l_{i,S_1}^k \text{ th PM}}. \tag{30}$$

Similarly, for scenarios S_2 and S_3, the false alarm costs $C_F^{S_2}$ and $C_F^{S_3}$ over a finite horizon are given by

$$\begin{aligned} C_F^{S_2}&=\sum_{k=1}^{N}\sum_{i=1}^{M_{S_2}^k}\underbrace{\frac{\tau_{i,S_2}^k-T_{l_{i,S_2}^k-1}}{h}\alpha c_f}_{\text{False alarm cost of } l_{i,S_2}^k \text{ th PM}},\\ C_F^{S_3}&=\sum_{k=1}^{N}\underbrace{\frac{\tau_{S_3}^k-T_M}{h}\alpha c_f}_{\text{Terminal false alarm cost}}. \end{aligned} \tag{31}$$

Thus, the total false alarm cost C_F over a finite horizon is given by

$$C_F=p_1\times C_F^{S_1}+p_2\times C_F^{S_2}+C_F^{S_3}, \tag{32}$$

where $p_1=\dfrac{M_{S_1}}{M}$ and $p_2=\dfrac{M_{S_2}}{M}$ denote the occurrence probabilities of scenarios S_1 and S_2, respectively.

5.5 Optimization of ETC over a finite horizon

According to the cost formulations presented in Sections 5.1–5.4, the total cost over a finite horizon is expressed by the following equation

$$TC=C_I+C_Q+C_M+C_F, \tag{33}$$

where C_I, C_Q, C_M, and C_F denote the inspection cost, quality cost, maintenance cost, and false alarm cost, respectively. Since T_{l_{i,S_1}^k}, T_{l_{i,S_2}^k}, $M_{S_1}^k$, and $M_{S_2}^k$ are stochastic variables over a finite horizon, the total cost TC is also stochastic. Due to the randomness and complexity inherent in these variables, it is difficult to compute the total cost analytically. Therefore, a simulation-based optimization approach is employed to approximate the total cost and determine the optimal decision (Cheng et al., 2018).

For each candidate set of decision variables $L=(L_D,L_Q)$, multiple simulation runs are conducted to estimate the ETC. Regarding the simulation method for calculating the ETC[L], we give it in Algorithm 1. In Algorithm 1, the TC_i is the total cost for ith simulation. The optimal thresholds $L^*=(L_D^*,L_Q^*)$ are obtained by minimizing the ETC, denoted as $ETC_{\min}$, as follows

$$\begin{aligned} ETC_{\min}&=\min_{L_D,L_Q}\frac{1}{s}\sum_{i=1}^{s}TC_i\\ &\text{s.t. } L_D^{\min}<L_D\leqslant L_D^{\max},\\ &L_Q\in\{L_{Q_1},L_{Q_2},\dots\},\\ &\delta>0,\ n\in\mathbb{Z}^+. \end{aligned} \tag{34}$$

Here, $L_{\mathrm{D}}^{\min}$ and $L_{\mathrm{D}}^{\max}$ denote the lower and upper bounds of the continuous threshold L_D, and L_Q is selected from a discrete control limit set $\{L_{Q_1},L_{Q_2},\dots\}$. The set $\{L_{Q_1},L_{Q_2},\dots\}$ is related to the type I error α, and this relationship can be established through Monte Carlo simulation. This approach is widely used in control chart research to assess the statistical performance of monitoring schemes, especially in balancing detection sensitivity against the false alarm frequency (Qiu, 2013).

To enhance the efficiency of identifying the optimal policy, a simulation-based optimization approach using a genetic algorithm is employed. Genetic algorithms are heuristic methods designed to explore large solution spaces efficiently. The algorithm begins by generating an initial population of candidate solutions, followed by iterative evaluation of their fitness to guide the construction of subsequent generations. Genetic operations, such as crossover and mutation, are employed to evolve the population toward better solutions. The specific implementation and modifications of the genetic algorithm in this study are detailed as follows:

(1) **Coding:** Solutions are represented using real-number encoding.

(2) **Generation of initial population:** The initial population, consisting of n_{pop} candidate solutions, is generated randomly.

(3) **Fitness evaluation:** The fitness of a solution $L=(L_D,L_Q)$ is computed as ETC[L]$^{-1}$, where ETC[L] is obtained by Algorithm 1. To improve computational efficiency, once ETC[L] is calculated for a solution, its value is stored and directly retrieved in later generations if the same solution appears.

(4) **Crossover:** Parent solutions are selected using the roulette wheel selection method. Arithmetical crossover is applied to generate new candidate solutions as follows:

$$\begin{aligned} L_1'&=\nu_1L_1+\nu_2L_2,\\ L_2'&=\nu_1L_2+\nu_2L_1, \end{aligned} \tag{35}$$

where L_1 and L_2 are two parent solutions, $\nu_1\in(0,1)$ is a randomly chosen coefficient, and $\nu_2=1-\nu_1$.

(5) **Mutation:** To preserve valuable genetic information while maintaining diversity, a non-uniform mutation method is adopted. The mutation is applied to a randomly selected variable L_i of a solution $L=(L_1,L_2,\dots,L_{n_{\text{pop}}})$, generating a new solution $L'=(L_1',\dots,L_i',\dots,L_{n_{\text{pop}}}')$. The mutated value L_i' is determined as follows

$$L_i'=L_i+\Delta(L_i^U-L_i,t)\ \text{or}\ L_i'=L_i-\Delta(L_i-L_i^L,t), \tag{36}$$

where L_i^U and L_i^L are the upper and lower bounds of L_i, respectively. The $\Delta(L,t)$ is the non-uniform perturbation function.

Algorithm 1 Simulation and optimization of joint CBM for finite horizon.

Initialize parameters:
Set initial value of $j \leftarrow 1$, $M_{S_1}^k \leftarrow 0$, $M_{S_2}^k \leftarrow 0$, $l_{pm} \leftarrow 0$, $s \leftarrow 5000$.

Simulation procedure:
for k in $1 : N$ **do**
  Generate random values $\boldsymbol{D}_{t_j}^k$ and $\boldsymbol{Q}_{t_j}^k$.
  Compute $\boldsymbol{f}_{t_j}^k$ and $h_{t_j}^k$ by **Proposition** 1 and Equation (7).
  Generate n random samples $\boldsymbol{Y}_{t_j}^k$.
  Compute n random samples $\boldsymbol{\varepsilon}_{t_j}^k$ by $\boldsymbol{\varepsilon}_{t_j}^k = \boldsymbol{Y}_{t_j}^k - \boldsymbol{f}_{t_j}^k$.
end for
Compute R_{t_j} and H_{t_j} by Equations (15) and (8).
if $R_{t_j} \geq L_Q$ & $H_{t_j} < L_D$ & $t_j < T$ **then** ▷ Control chart signal.
  $k \leftarrow \arg\max_{1 \leq k \leq N} R_{t_j}^k$. ▷ Diagnosis of stage.
  $T_{l_{pm}} \leftarrow t_j$, $l_{pm} \leftarrow l_{pm} + 1$.
  $T_{l_{i,S_1}^k} \leftarrow T_{l_{pm}}$, $M_{S_1}^k \leftarrow M_{S_1}^k + 1$.
  $\boldsymbol{X}_{t_j}^k \leftarrow 0$, $V_{l_{pm}} \leftarrow V_{l_{pm}-1} + e(T_{l_{pm}} - T_{l_{pm}-1})$, to step 4.
else if $H_{t_j} \geq L_D$ & $R_{t_j} < L_Q$ & $t_j < T$ **then** ▷ System PHM signal.
  $k \leftarrow \arg\max_{k \in Des} |h_{t_j}^k - \hat{h}_{t_j}^k|$. ▷ Diagnosis of stage.
  $T_{l_{pm}} \leftarrow t_j$, $l_{pm} \leftarrow l_{pm} + 1$.
  $T_{l_{i,S_2}^k} \leftarrow T_{l_{pm}}$, $M_{S_2}^k \leftarrow M_{S_2}^k + 1$.
  $\boldsymbol{X}_{t_j}^k \leftarrow 0$, $V_{l_{pm}} \leftarrow V_{l_{pm}-1} + e(T_{l_{pm}} - T_{l_{pm}-1})$, to step 4.
else if $H_{t_j} \geq L_D$ & $R_{t_j} \geq L_Q$ & $t_j < T$ **then** ▷ Signal simultaneously.
  To steps $12 - 15$ and $17 - 20$.
else if $H_{t_j} < L_D$ & $R_{t_j} < L_Q$ & $t_j < T$ **then** ▷ No signal.
  No action, $j \leftarrow j + 1$.
  To step 4.
else if $t_j \geq T$ **then** ▷ End of inspection.
  $M_{S_1} \leftarrow \sum_{k=1}^N M_{S_1}^k$, $M_{S_2} \leftarrow \sum_{k=1}^N M_{S_2}^k$, $M \leftarrow M_{S_1} + M_{S_2}$.
  Compute $p_1 = \frac{M_{S_1}}{M}$ and $p_2 = \frac{M_{S_2}}{M}$.
  End of inspection, to step 31.
end if
Compute $TC_{S_1} = p_1 \times \left[C_I^{S_1} + C_Q^{S_1} + C_M^{S_1} + C_F^{S_1}\right]$.
Compute $TC_{S_2} = p_2 \times \left[C_I^{S_2} + C_Q^{S_2} + C_M^{S_2} + C_F^{S_2}\right]$.
Compute $TC_{S_3} = C_I^{S_3} + C_Q^{S_3} + C_M^{S_3} + C_F^{S_3}$.
$TC \leftarrow TC_{S_1} + TC_{S_2} + TC_{S_3}$. ▷ Total cost.
Repeat above steps s times and compute ETC $\leftarrow \sum_{i=1}^s \frac{TC_i}{s}$.

Optimization procedure:
Compute the $\text{ETC}_{\min}$ using the genetic algorithm.
Output the optimal values $L^* \leftarrow (L_D^*, L_Q^*)$.

6 Numerical analysis

To validate the applicability of the proposed strategy, this paper conducts a numerical analysis based on the scroll machining process in compressor manufacturing. The analysis evaluates the proposed strategy's impact on both manufacturing efficiency and product quality, thereby demonstrating its effectiveness in optimizing overall production performance. The scroll machining process is a four-stage serial operation conducted over a finite time horizon of $T = 100$ hours. It involves several critical procedures, including milling, turning, drilling, and slotting, all of which are essential to ensuring the accuracy and performance of scroll components. At each stage, the QRCs refer to key machine components that influence product quality, such as cutting tools, fixtures, the spindle, and guideways. Meanwhile, the KPCs focus on the dimensional accuracy of critical features on the workpiece, which directly determine the functionality and overall quality of the scroll components. The schematic of compressor and scroll, along with the scroll machining process, are illustrated in Fig. 4. The relationship between QRCs degradation and the corresponding KPCs at each stage is described by the following expression, where the coefficients are obtained from field production records and refined through designed experiments (Taguchi, 1986; Wu and Hamada, 2011).

$$
\begin{aligned}
Y_t^1 &= \begin{bmatrix} 0.00036 \\ 0.00034 \\ 0.00032 \end{bmatrix} + \begin{bmatrix} 0.127 & 0.125 & 0.101 & 0.124 \\ 0.112 & 0.122 & 0.121 & 0.132 \\ 0.103 & 0.125 & 0.113 & 0.121 \end{bmatrix} \mathrm{X}_t^1 + \begin{bmatrix} 0.435 & 0.422 & 0.428 \\ -0.526 & -0.531 & -0.527 \\ 0.343 & 0.472 & 0.461 \end{bmatrix} \mathrm{z}_t^1 + \mathrm{X}_t^1 \boldsymbol{C}^1 \mathrm{z}_t^1, \\
Y_t^2 &= \begin{bmatrix} 0.00036 \\ 0.00034 \\ 0.00032 \end{bmatrix} + \begin{bmatrix} 0.127 & 0.125 & 0.101 & 0.124 \\ 0.112 & 0.122 & 0.121 & 0.132 \\ 0.103 & 0.125 & 0.113 & 0.121 \end{bmatrix} \mathrm{X}_t^2 + \begin{bmatrix} 0.435 & 0.422 & 0.428 \\ -0.526 & -0.531 & -0.527 \\ 0.343 & 0.472 & 0.461 \end{bmatrix} \mathrm{z}_t^2 + \mathrm{X}_t^2 \boldsymbol{C}^2 \mathrm{z}_t^2, \\
Y_t^3 &= \begin{bmatrix} 0.00036 \\ 0.00034 \\ 0.00032 \end{bmatrix} + \begin{bmatrix} 0.127 & 0.125 & 0.101 & 0.124 \\ 0.112 & 0.122 & 0.121 & 0.132 \\ 0.103 & 0.125 & 0.113 & 0.121 \end{bmatrix} \mathrm{X}_t^3 + \begin{bmatrix} 0.435 & 0.422 & 0.428 \\ -0.526 & -0.531 & -0.527 \\ 0.343 & 0.472 & 0.461 \end{bmatrix} \mathrm{z}_t^3 + \mathrm{X}_t^3 \boldsymbol{C}^3 \mathrm{z}_t^3, \\
Y_t^4 &= \begin{bmatrix} 0.00036 \\ 0.00034 \\ 0.00032 \end{bmatrix} + \begin{bmatrix} 0.127 & 0.125 & 0.101 & 0.124 \\ 0.112 & 0.122 & 0.121 & 0.132 \\ 0.103 & 0.125 & 0.113 & 0.121 \end{bmatrix} \mathrm{X}_t^4 + \begin{bmatrix} 0.435 & 0.422 & 0.428 \\ -0.526 & -0.531 & -0.527 \\ 0.343 & 0.472 & 0.461 \end{bmatrix} \mathrm{z}_t^4 + \mathrm{X}_t^4 \boldsymbol{C}^4 \mathrm{z}_t^4,
\end{aligned} \tag{37}
$$

where $\boldsymbol{C}^k \in \mathbb{R}^{3\times4\times3}, k = 1,2,3,4$ and

$$
\begin{aligned}
\boldsymbol{C}^1 &= \left[\begin{bmatrix} -0.081 & -0.078 & -0.082 \\ -0.068 & -0.069 & -0.069 \\ -0.082 & -0.085 & -0.093 \\ -0.078 & -0.061 & -0.087 \end{bmatrix}, \begin{bmatrix} 0.066 & 0.081 & 0.065 \\ 0.064 & 0.067 & 0.084 \\ 0.069 & 0.066 & 0.077 \\ 0.081 & 0.053 & 0.052 \end{bmatrix}, \begin{bmatrix} -0.062 & -0.078 & -0.046 \\ -0.061 & -0.061 & -0.074 \\ -0.051 & -0.065 & -0.073 \\ -0.081 & -0.073 & -0.051 \end{bmatrix} \right]^{\top}, \\
\boldsymbol{C}^2 &= \left[\begin{bmatrix} -0.062 & -0.072 & -0.052 \\ -0.071 & -0.064 & -0.064 \\ -0.072 & -0.081 & -0.057 \\ -0.071 & -0.071 & -0.061 \end{bmatrix}, \begin{bmatrix} 0.062 & 0.068 & 0.062 \\ 0.071 & 0.071 & 0.062 \\ 0.072 & 0.067 & 0.062 \\ 0.061 & 0.081 & 0.071 \end{bmatrix}, \begin{bmatrix} -0.062 & -0.078 & -0.057 \\ -0.055 & -0.064 & -0.044 \\ -0.072 & -0.067 & -0.067 \\ -0.062 & -0.072 & -0.061 \end{bmatrix} \right]^{\top}, \\
\boldsymbol{C}^3 &= \left[\begin{bmatrix} -0.062 & -0.068 & -0.086 \\ -0.064 & -0.064 & -0.064 \\ -0.072 & -0.077 & -0.087 \\ -0.061 & -0.065 & -0.065 \end{bmatrix}, \begin{bmatrix} 0.065 & 0.063 & 0.061 \\ 0.061 & 0.046 & 0.051 \\ 0.041 & 0.051 & 0.041 \\ 0.061 & 0.056 & 0.052 \end{bmatrix}, \begin{bmatrix} -0.062 & -0.068 & -0.067 \\ -0.065 & -0.064 & -0.044 \\ -0.072 & -0.067 & -0.067 \\ -0.061 & -0.061 & -0.071 \end{bmatrix} \right]^{\top}, \\
\boldsymbol{C}^4 &= \left[\begin{bmatrix} -0.072 & -0.067 & -0.056 \\ -0.074 & -0.054 & -0.064 \\ -0.062 & -0.076 & -0.087 \\ -0.063 & -0.071 & -0.051 \end{bmatrix}, \begin{bmatrix} 0.053 & 0.051 & 0.062 \\ 0.064 & 0.067 & 0.042 \\ 0.062 & 0.045 & 0.063 \\ 0.063 & 0.063 & 0.062 \end{bmatrix}, \begin{bmatrix} -0.042 & -0.068 & -0.022 \\ -0.043 & -0.041 & -0.084 \\ -0.042 & -0.046 & -0.067 \\ -0.061 & -0.032 & -0.071 \end{bmatrix} \right]^{\top}.
\end{aligned} \tag{38}
$$

In SMMS, the occurrence rate of non-conforming items and their effects on QRCs often vary across stages. This study considers three trends of the occurrence rate $\boldsymbol{\rho}$ with stage progression: increasing, constant, and decreasing. The corresponding results for different levels of the QRCs impact parameter $\boldsymbol{\mu}_{nc}$ are illustrated in Fig. 5. As shown in Fig. 5(a), the first stage is not considered, as it is assumed that the initial incoming product quality is stable, with zero occurrence rate of non-conforming items. When the occurrence rate of non-conforming items increases across the subsequent stages, QRCs degradation also rises, and higher impact levels of non-conforming items further amplify this effect. In Fig. 5(b), under a constant occurrence rate, QRCs degradation remains relatively stable across stages, although higher impact levels still lead to a gradual increase. In Fig. 5(c), when the occurrence rate decreases across stages, QRCs degradation correspondingly declines; however, higher impact levels

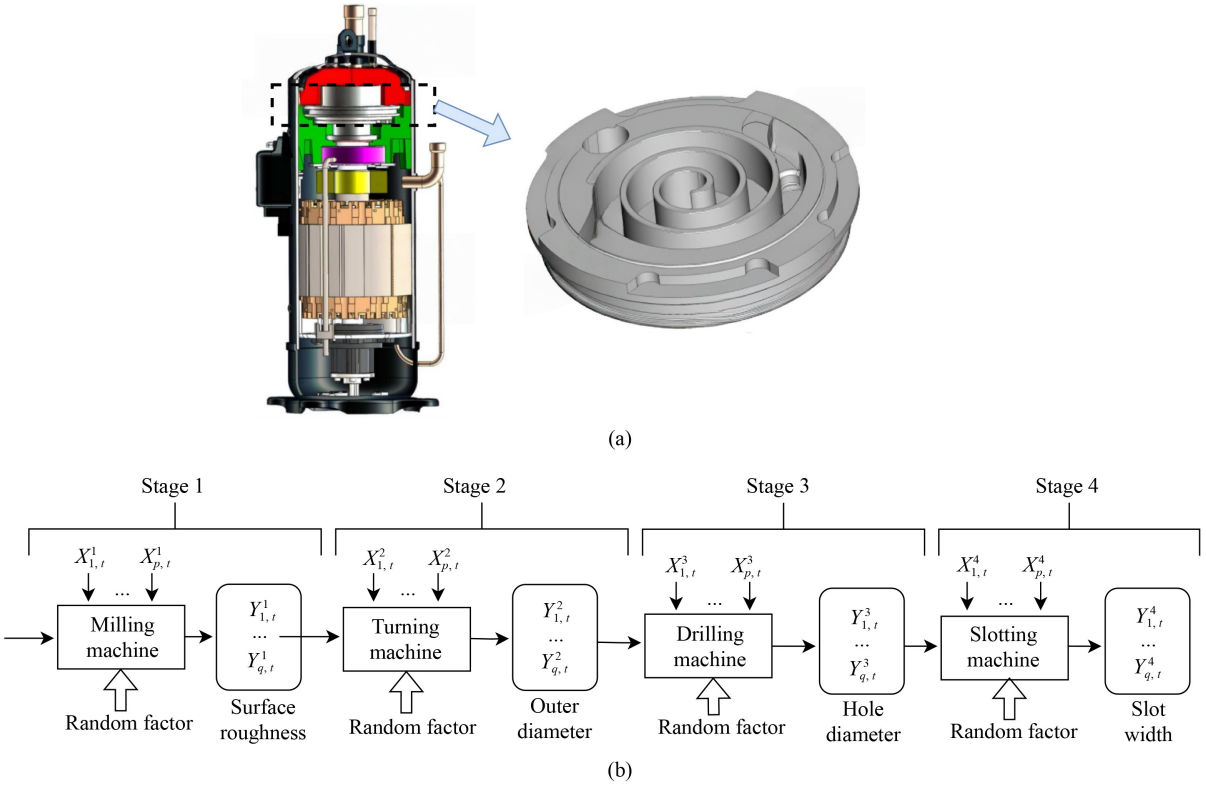


Fig. 4 (a) Schematic of compressor and scroll; (b) Machining process of scroll.

continue to intensify the degradation.

In this study, we assume the non-conforming rate at each stage is $\boldsymbol{\rho} = (0, 0.01, 0.02, 0.04)^{\mathrm{T}}$. The influence of each non-conforming item on the downstream QRCs follows a truncated Normal distribution with $\boldsymbol{\mu_{nc}} = (0.05, 0.05, 0.05, 0.05)^{\top}$ and $\boldsymbol{\sigma_{nc}} = (0.012, 0.015, 0.017, 0.018)^{\top}$. The weighting coefficient representing the effect of non-conforming items on the degradation of QRCs is $\boldsymbol{\zeta} = (0.25, 0.25, 0.25, 0.25)^{\top}$. Additionally, assuming that the IC covariance matrix of stage k is $\boldsymbol{\Sigma}_0^k = \boldsymbol{I}$ and the OC covariance matrix of stage k is $\boldsymbol{\Sigma}_1^k = \boldsymbol{\Sigma}_0^k + \delta \boldsymbol{I}$, where δ represents the shift magnitude. Let

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{\alpha}^1, ..., \boldsymbol{\alpha}^N)^{\top}, \hat{\boldsymbol{\beta}} = (\boldsymbol{\beta}^1, ..., \boldsymbol{\beta}^N)^{\top}, \hat{\omega} = (\omega^1, ..., \omega^N)^{\top},$$
$$\hat{\boldsymbol{\Sigma}}_z = (\boldsymbol{\Sigma}_z^1, ..., \boldsymbol{\Sigma}_z^N)^{\top}, \hat{\boldsymbol{\lambda}} = (\lambda^1, ..., \lambda^N)^{\top}, \hat{\eta} = (\eta^1, ..., \eta^N)^{\top}, \tag{39}$$

and the above parameter values are provided in Table 3.

6.1 Simulation-based optimization results

Next, we simulate the degradation process of QRCs and the variation of KPCs over a finite horizon based on the parameter settings of Table 3. Figure 6 compares the total degradation process of QRCs with inherent degradation process. It can be observed that both the inherent degradation and the total degradation increase continuously. Except for the first stage, the degradation in the other three stages is affected by upstream non-conforming items, which further contributes to the growth of total degradation. This indicates that neglecting the cumulative effect of upstream non-conforming items on QRCs degradation would lead to an underestimation of the actual degradation level of QRCs. In Fig. 7, the variance of random errors $\boldsymbol{\varepsilon}_t^k$ of each stage are further analyzed. The variance of KPCs across stages exhibits a decreasing-then-increasing trend. This suggests that in the early stages, KPCs variations are relatively small and product quality remains stable, whereas in the later stages, the variations become more pronounced and quality deteriorates. Therefore, accurately identifying the transition point of this variance is critical for maintaining the stability of KPCs. It can be observed that as QRCs degrade, when the variance of random errors $\varepsilon_{i,t}^k$ satisfy $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^{\top} < 1$, the variation of KPCs initially decreases and then increases, with relatively small overall variation. However, when $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^{\top} > 1$, the variation of KPCs continuously increases. In this study, we assume the covariance matrix of random errors $\boldsymbol{\varepsilon}_t^k$ as $\boldsymbol{\Sigma}_0^k$ and $\boldsymbol{\Sigma}_1^k$ when $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^{\top} \leqslant 1, \forall i \in \{1, .., q\}$ and $(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)\Sigma_z^k(\boldsymbol{b}_i^k + \boldsymbol{x}_t^{k\top}\boldsymbol{C}_i^k)^{\top} > 1, \exists i \in \{1, .., q\}$, respectively.

According to Algorithm 1, the ETC over a finite horizon is estimated through simulation with 5000 replications. Let $L_D^{\min} = 0$, $L_D^{\max} = 5$, and the control limit set {20, 20.6, 21.2, 21.8, 22.6, 23.5, 24.5, 25.6, 29.5}, the control limit

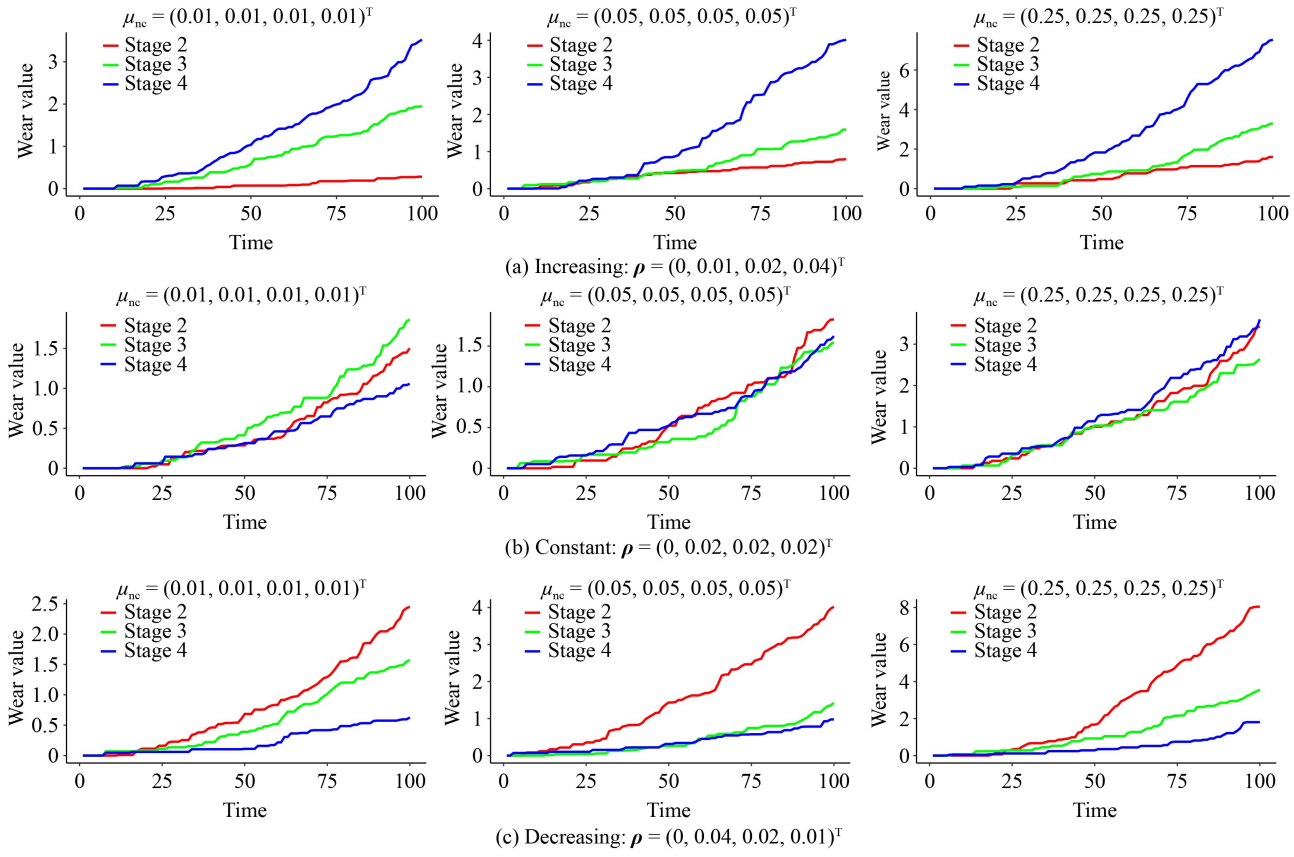


Fig. 5 Trends of $\boldsymbol{\rho}$ across successive stages.

Table 3 Parameter setting

Parameter	Value	Parameter	Value
T	100(hour)	h	1(hour)
p	4	q	3
N	4	n	10
δ	0.3	$\boldsymbol{e}$	$(0.1,0.1,0.1,0.1)^T$
$\widehat{\boldsymbol{\alpha}}$	$\begin{bmatrix} 0.122 & 0.252 & 0.192 & 0.182 \\ 0.263 & 0.173 & 0.173 & 0.193 \\ 0.163 & 0.173 & 0.283 & 0.193 \\ 0.149 & 0.159 & 0.169 & 0.279 \end{bmatrix}$	$\hat{\boldsymbol{\beta}}$	$\begin{bmatrix} 4.1 & 3.3 & 4.2 & 3.8 \\ 4.3 & 4.2 & 3.7 & 3.5 \\ 4.2 & 4.1 & 4.4 & 4.6 \\ 4.4 & 4.2 & 3.6 & 3.9 \end{bmatrix}$
$\hat{\boldsymbol{\lambda}}$	$(1.151.251.401.30)^T$	$\hat{\eta}$	$(420,488,426,435)^T$
$\widehat{\boldsymbol{\Sigma}_z}$	$\begin{bmatrix} 0.371 & 0.341 & 0.352 \\ 0.302 & 0.351 & 0.361 \\ 0.390 & 0.381 & 0.357 \\ 0.330 & 0.356 & 0.381 \end{bmatrix}$	$\widehat{\omega}$	$\begin{bmatrix} 0.05 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.05 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.05 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.05 \end{bmatrix}$
c_{in}	10($)	c_{mm}	10($)
c_{pm}, c_r	1000($)	c_{ic}	2($)
c_{oc}	400($)	c_f	500($)

set corresponds to the type I error set {0.01, 0.009, 0.008, 0.007, 0.006, 0.005, 0.004, 0.003, 0.002}. The genetic algorithm with population size n_{pop} = 10 is employed to optimize the ETC. As illustrated in Fig. 8, after 50 generations of iteration, the ETC tends to converge,with its variation amplitude remaining below 400. The variation in convergence results can be attributed to the number of simulation replications and remains within the acceptable range in practice. Specifically, at the 50th generation, ETC_{min} converges to $18402.24, corresponding to the optimal decision variables L_Q^* = 25.6 and L_D^* = 0.52.

To verify the cost-saving effectiveness of the proposed

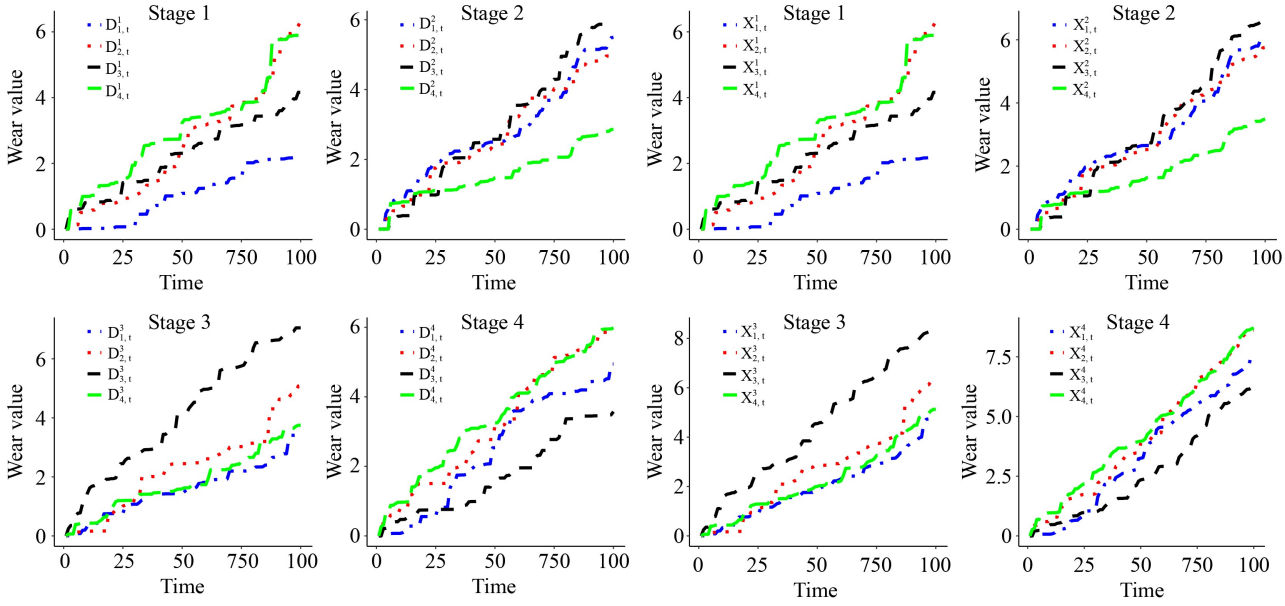


Fig. 6 Comparison of QRCs degradation for each stage.

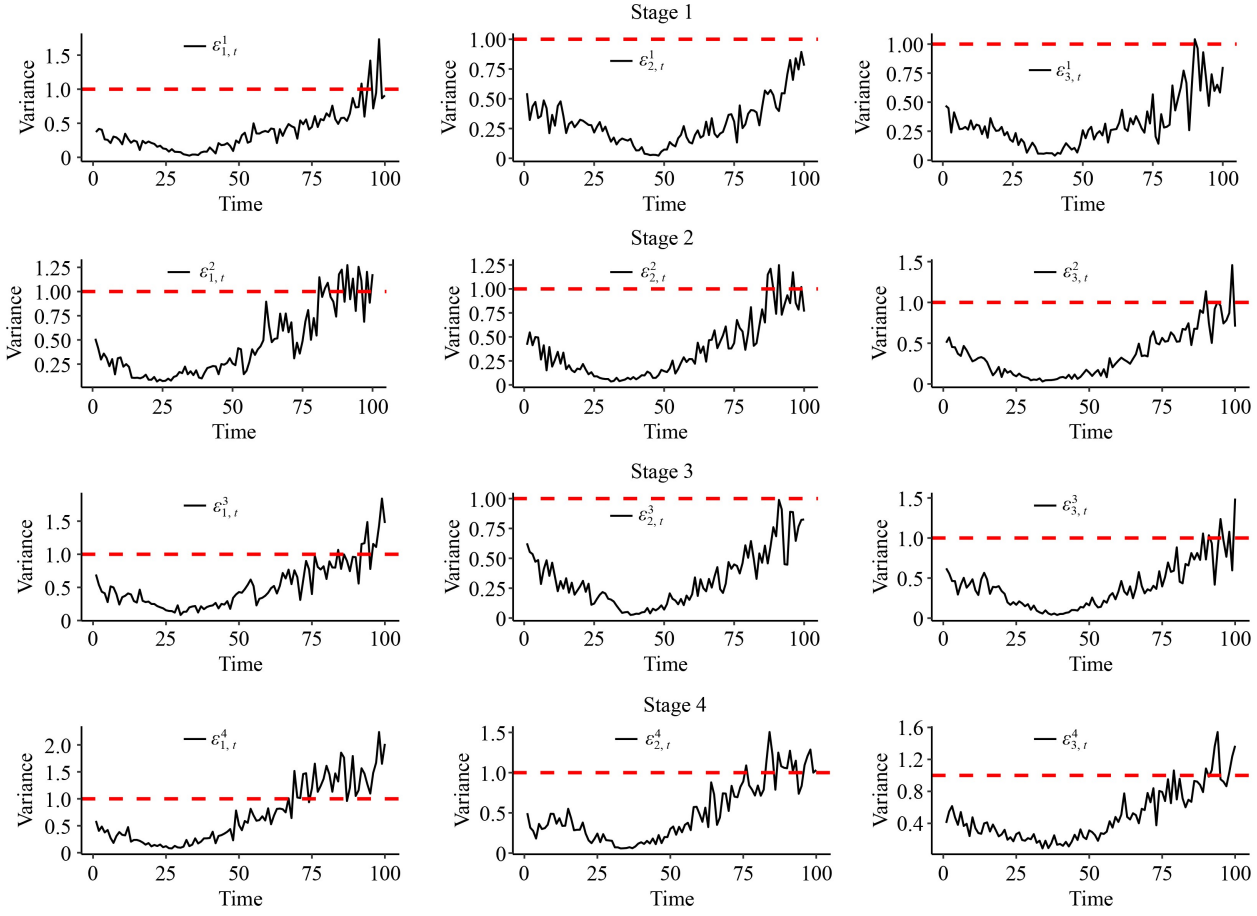


Fig. 7 The variance of random errors for each stage.

strategy, we compare it with tradition single monitoring strategies. Strategy I: The proposed joint monitoring strategy; Strategy II: The strategy which monitors only the QRCs condition; Strategy III: The strategy which

monitors only KPCs. Table 4 presents the optimization results of three strategies and / denotes null. It is evident from Table 4 that Strategy I achieves a lower $ETC_{\min}$ compared to Strategies II and III. Specifically, Strategy I incurs higher maintenance costs due to more frequent maintenances; however, timely maintenance stabilizes product quality, thereby reducing quality-related costs. In contrast, Strategies II and III make maintenance decisions based solely on either the condition of QRCs or the quality of KPCs, overlooking certain critical maintenance opportunities. This oversight prolongs the OC time of the process, resulting in increased quality-related costs.

Subsequently, the real-time monitoring of SMMS is conducted based on the optimal decision variables L_Q^* and L_D^*, the monitoring results are presented in Table 5 and Fig. 9. In Table 5, / denotes null. As shown in Table 5 and Fig. 9, over a finite horizon, four alarms ($M = 4$) occur. The MGLR chart triggers two alarms ($M_{S_1} = 2$); The system PHM triggers two alarms ($M_{S_2} = 2$). Specifically, the MGLR chart triggers alarms at the 86th and the 66th hours at stages 1 and 2, respectively. The system PHM triggers alarms at the 62th and the 76th hours at stages 3 and 4, respectively.

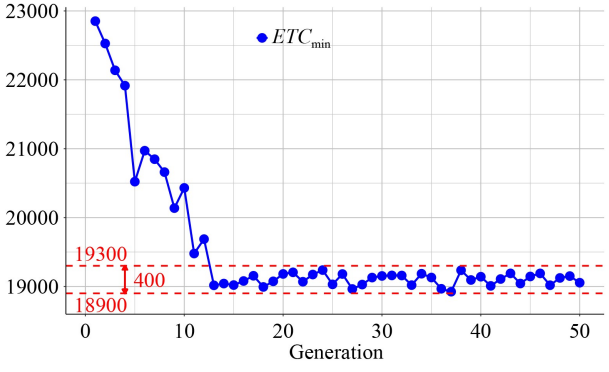


Fig. 8 The minimum fitness obtained at each generation.

6.2 Sensitivity analysis

To further support decision-making, managers may wish to examine the influence of parameter settings on both the optimal decision L^* and the $ETC_{\min}$. Therefore, sensitivity analyses are performed by adjusting each parameter individually within a range of −50% to +50%, while keeping all other parameters constant. The results are shown in the Figs. 10 and 11.

Figure 10 depicts the impact of cost parameters on $ETC_{\min}$. The results show that $ETC_{\min}$ increases with rising values of all cost parameters. In particular, $ETC_{\min}$ exhibits high sensitivity to c_{oc} and c_{pm}, whereas its sensitivity to c_{ic}, c_{mm}, and c_f is comparatively moderate. Figure 11(a) shows the effect of cost parameters on the optimal control limit L_Q^*. The results reveal that L_Q^* decreases with increasing c_{oc} and c_{ic}, but increases with increasing c_{pm}, c_{mm}, and c_f. Among these parameters, the variation of L_Q^* with respect to c_{oc} and c_{pm} is the most pronounced, whereas its response to c_{ic}, c_{mm}, and c_f remains relatively moderate. Figure 11(b) depicts the effect of cost parameters on the hazard threshold L_D^*. The analysis reveals that L_D^* decreases as c_{oc} increases, while it increases with increasing c_{ic}, c_{pm}, c_{mm}, and c_f. Notably, L_D^* exhibits high sensitivity to c_{ic}, c_{oc}, c_{pm}, and c_f, whereas its sensitivity to c_{mm} is relatively moderate. Based on the above analysis, the cost parameters c_{oc} and c_{pm} are the most influential factors affecting both L_Q^* and L_D^*. Their joint impact highlights the necessity of considering these two parameters as the primary drivers in decision-making. Accordingly, managers are advised to dynamically adjust the control limit L_Q^* and the hazard threshold L_D^* in response to variations in c_{oc} and c_{pm}. Such adaptive adjustment not only mitigates quality-

Table 4 The optimal results obtained under the three strategies

Maintenance strategy	Optimal decisions and $ETC_{\min}$	Related costs							
Strategy I	$L_Q^* = 25.6$; $L_D^* = 0.52$; $ETC_{\min} = 18402.2$	$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	$C_I^{S_2}$	$C_M^{S_2}$	$C_Q^{S_2}$	$C_F^{S_2}$
		\$172.2	\$1236.3	\$4416.2	\$97.3	\$744.1	\$2261.3	\$7251.3	\$220.4
		$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	C_I	C_M	C_Q	C_F
		\$83.7	\$1112.5	\$614.6	\$192.4	\$1000	\$4610.1	\$12282.1	\$510.1
Strategy II	$L_D^* = 1.42$; $ETC_{\min} = 20799.2$	$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	$C_I^{S_2}$	$C_M^{S_2}$	$C_Q^{S_2}$	$C_F^{S_2}$
		/	/	/	/	\$820	\$2530	\$9104.23	/
		$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	C_I	C_M	C_Q	C_F
		\$180	\$1713.7	\$6451.3	/	\$1000	\$4243.7	\$15555.5	/
Strategy III	$L_Q^* = 25.6$; $ETC_{\min} = 19360.9$	$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	$C_I^{S_2}$	$C_M^{S_2}$	$C_Q^{S_2}$	$C_F^{S_2}$
		\$887.7	\$2167.2	\$10662.3	\$207.6	/	/	/	/
		$C_I^{S_1}$	$C_M^{S_1}$	$C_Q^{S_1}$	$C_F^{S_1}$	C_I	C_M	C_Q	C_F
		\$112.3	\$1271.4	\$3627.3	\$425.2	\$1000	\$3438.6	\$14289.6	\$632.8

Table 5 The monitoring results of joint CBM strategy

Scenario	S_1				S_2			
Stage	Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4
MGLR chart	86th hour	66th hour	/	/	/	/	/	/
System PHM	/	/	/	/	/	/	62th hour	76th hour
Number of alarms		$M_{S_1}^1=1, M_{S_1}^2=1$					$M_{S_1}^3=0, M_{S_1}^4=0$	
		$M_{S_2}^1=0, M_{S_2}^2=0$					$M_{S_2}^3=1, M_{S_2}^4=1$	
				$M_{S_1}=2, M_{S_2}=2, M=4$				

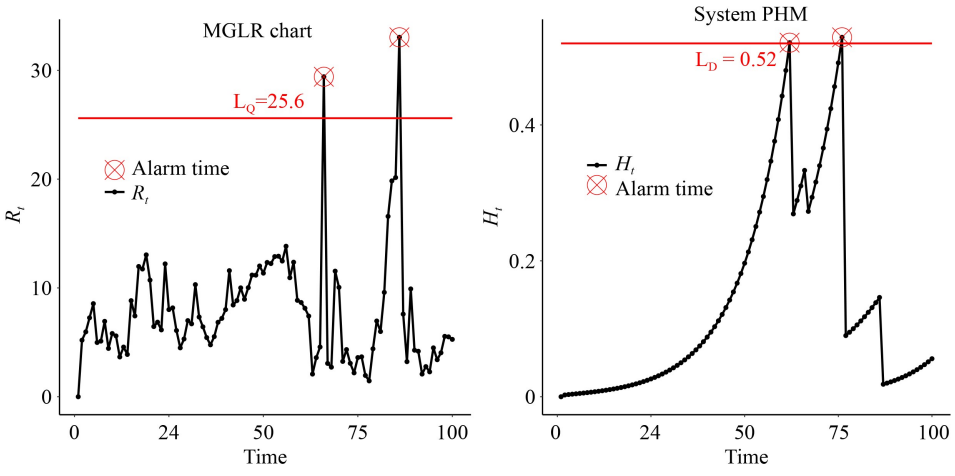


Fig. 9 Alarm result of proposed strategy for finite horizon.

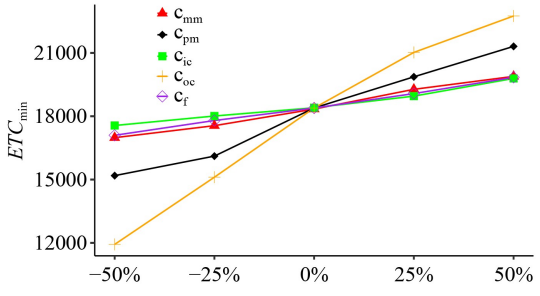


Fig. 10 Influence of cost parameters on $\boldsymbol{ETC}_{\min}$.

related losses stemming from process instability but also minimizes maintenance costs associated with preventive interventions, thereby supporting more effective optimal decision-making.

In Fig. 12, we analyze how the sample size n, shift magnitude δ, age reduction coefficient e, and non-conforming rate $\boldsymbol{\rho}$ affect the $ETC_{\min}$. The Fig. 12 considers three levels for each parameter: $n=5,20,40$, $\delta=0.1,0.5,1$, $e=0,0.5,1$, and $\boldsymbol{\rho}=(0,0.005,0.001,0.002)^{\top}$, $(0,0.01,0.02,0.04)^{\top}$,$(0,0.02,0.04,0.08)^{\top}$. **(1)** As the shift magnitude δ becomes larger, $ETC_{\min}$ also declines. This is because the monitoring system is more responsive to larger shifts, which facilitates quicker identification of instability. Therefore, in systems prone to large shifts, joint monitoring can be particularly cost-effective. **(2)** An

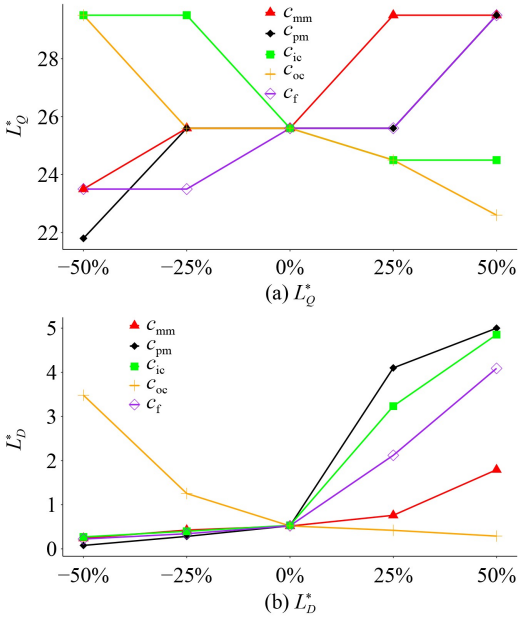


Fig. 11 Influence of cost parameters on decision variables.

increase in the age reduction coefficient e leads to a modest decline in $ETC_{\min}$. Although a higher e reflects faster deterioration after maintenance, it also increases the frequency of hazard warnings. This prompts more frequent PM actions, which shortens periods of instability

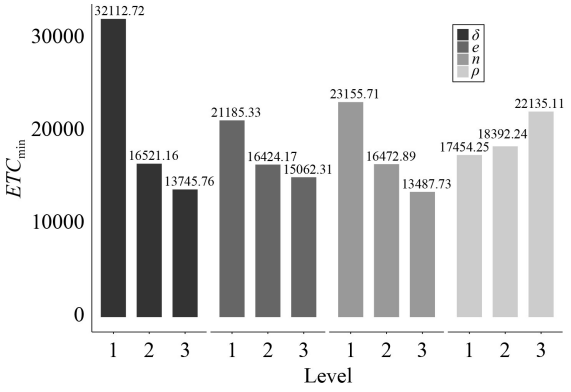


Fig. 12 Influence of parameters on $ETC_{\min}$.

and reduces the overall cost. Hence, managers should account for the machine recovery rate when designing maintenance strategies. **(3)** As the sample size n increases, $ETC_{\min}$ decreases steadily. A larger sample size improves the sensitivity of the monitoring scheme, allowing earlier detection of process instability and reducing quality-related and maintenance costs. This suggests that managers should consider increasing the sample size when cost and operational constraints allow. **(4)** Conversely, an increase in the non-conforming rate $\boldsymbol{\rho}$ leads to a rise in $ETC_{\min}$. A higher $\boldsymbol{\rho}$ elevates the probability of producing non-conforming items within each inspection interval, thereby accelerating the degradation of downstream components. To mitigate the expected total cost, it is imperative for managers to strengthen upstream quality control measures aimed at reducing the non-conforming rate. Overall, $ETC_{\min}$ is jointly affected by n, δ, e, and $\boldsymbol{\rho}$. Among these, n and δ have the strongest impact, while e and $\boldsymbol{\rho}$ play moderate roles. The joint effects of sample size n and shift magnitude δ on $ETC_{\min}$ indicate that a larger n improves monitoring sensitivity, while a larger δ enables faster detection of process instability. Considering both together, the optimal strategy is to choose a moderate-to-large n and set the control limit L_Q according to the expected δ. In addition, by accounting for the non-conforming rate $\boldsymbol{\rho}$ and selecting an appropriate age reduction coefficient e, managers can determine the optimal system hazard rate threshold L_D, achieving $ETC_{\min}$ and timely interventions.

7 Conclusions

This paper considers a CBM strategy that simultaneously monitors two data sources in the SMMS. First, the inherent degradation processes of QRCs are modeled using the Gamma process. To capture the influence of KPCs on QRCs degradation, the impact of upstream non-conforming items is incorporated. This degradation is modeled by a truncated normal distribution, from which the cumulative effect of all non-conforming items is derived. By integrating this with the inherent degradation, the total degradation of QRCs is obtained. The degradation of QRCs is further incorporated as a covariate into the constructed PHM. Meanwhile, the MGLR statistic is employed to monitor variations in the KPCs. By simultaneously considering the PHM and the MGLR control chart, a joint monitoring scheme is proposed. Based on the monitoring results, a diagnostic strategy is designed to identify the stages requiring maintenance, while taking into account the effects of different PM actions on the degradation of the QRCs. Subsequently, the possible alarm scenarios over a finite horizon are enumerated, and the corresponding cost formulations are derived. Finally, simulations are conducted to replicate the alarm scenarios and estimate the ETC. To minimize the ETC, a simulation-based genetic algorithm is employed to optimize the control limits and hazard rate thresholds.

In the numerical analysis, an example involving a four-stage serial scroll machining process is implemented, and we simulate the degradation process of QRCs and the variation of KPCs. Compared to existing approaches, the proposed joint CBM strategy achieves a lower total cost. Furthermore, we conducted a comprehensive sensitivity analysis to examine the influence of cost parameters on $ETC_{\min}$, L_Q^*, and L_D^*. The results indicate that c_{oc} and c_{pm} exert the most significant influence on these metrics. Based on the sensitivity analysis, managers should prioritize reducing the OC cost c_{oc} and the PM cost c_{pm} by negotiating more favorable service agreements and optimizing maintenance schedules. Additionally, we analyzed the effects of various process parameters on $ETC_{\min}$. The results reveal that a larger sample size n, greater shift magnitude δ, higher age reduction coefficient e generally lead to a lower $ETC_{\min}$, and a larger non-conforming rate $\boldsymbol{\rho}$ leads to a larger $ETC_{\min}$.

Future work can move in two directions. First, this study only considers changes in variance of random errors, while mean shifts may also occur and jointly affect product quality; extending the framework to monitor both mean and variance of random errors would allow earlier and more accurate detection of quality problems. Second, the current approach reflects a reactive intervention, whereas predictive maintenance offers a proactive way to reduce downtime and quality risks; combining data on QRC degradation and product quality for joint predictive maintenance in SMMS would be a promising direction to improve reliability and quality.

Acknowledgments The authors would like to express their sincere gratitude to the National Natural Science Foundation of China for providing financial support for this research under Grants No. 72032005, 72231005, and 72401008. The authors also appreciate the valuable comments and suggestions from colleagues and reviewers, which have greatly helped improve the quality of this work.

Competing Interests The authors declare that they have no competing interests.

Appendixes

Appendix A Proof of Eq. (7)

According to Eq. (4), we have

$$\begin{aligned}\mathrm{E}[\exp(\omega^k \boldsymbol{X}_t^k)] &= \mathrm{E}[\exp(\textstyle\sum_{l=1}^{p}\omega_l^k X_{l,t}^k)] \\ &= \mathrm{E}[\exp(\textstyle\sum_{l=1}^{p}\omega_l^k (D_{l,t}^k + Q_{l,t}^k))].\end{aligned} \tag{A.1}$$

Considering the independence of $D_{l,t}^k$ and $Q_{l,t}^k$, we can get

$$\begin{aligned}\mathrm{E}[\exp(\textstyle\sum_{l=1}^{p}\omega_l^k (D_{l,t}^k + Q_{l,t}^k))] &= \prod_{l=1}^{p}\mathrm{E}[\exp(\omega_l^k (D_{l,t}^k + Q_{l,t}^k))] \\ &= \prod_{l=1}^{p}\mathrm{E}[\exp(\omega_l^k D_{l,t}^k)]\,\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)].\end{aligned} \tag{A.2}$$

Next, we prove the $\mathrm{E}[\exp(\omega_l^k D_{l,t}^k)]$ and $\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)]$ respectively. Since $\omega_l^k D_{l,t}^k \sim \mathrm{Ga}(\alpha_l^k t, \omega_l^k \beta_l^k)$ and according to the moment generating function of the Gamma distribution, we can get

$$\mathrm{E}[\exp(\omega_l^k D_{l,t}^k)] = \int_0^{\infty}\exp\left(\omega_l^k x\right) f_{D_l^k}(x)\,\mathrm{d}x = (1-\omega_l^k\beta_l^k)^{-\alpha_l^k t}, \tag{A.3}$$

where $f_{D_l^k}(x)$ is the probability density function of D_l^k. Since $Q_{l,t}^k = \zeta_l \sum_{j=1}^{N_t^k} s(\theta_j^k)$ is a compound Poisson process, we have

$$\begin{aligned}&\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)] = \mathrm{E}[\exp(\zeta_l\omega_l^k\textstyle\sum_{j=1}^{N_t^k} s(\theta_j^k))] \\ &= \textstyle\sum_{n=0}^{\infty} P(N_t^k = n)\mathrm{E}[\exp(\zeta_l\omega_l^k\sum_{j=1}^{n} s(\theta_j^k))].\end{aligned} \tag{A.4}$$

According to **Assumption 1**, the sequence $\{s(\theta_j^k)\}_{j\geqslant 1}$ is mutually independent, we have $\mathrm{E}[\exp(\zeta_l\omega_l^k\sum_{j=1}^{n} s(\theta_j^k))] = \mathrm{E}[\exp(\zeta_l\omega_l^k s(\theta_j^k))]^n$. It is worth mentioning that this study assumes that $s(\theta_j^k)$ follows the truncated normal distribution, and the following derivations for other distributions are similar. Let $M_l^k = \mathrm{E}[\exp(\zeta_l\omega_l^k s(\theta_j^k))]$, according to the moment generating function of the truncated normal distribution on $[0,\infty]$, we have

$$\begin{aligned}M_l^k &= \int_0^{\infty}\exp\left(\zeta_l\omega_l^k x\right) f_{Q_l^k}(x)\,\mathrm{d}x \\ &= \exp\left(\zeta_l\omega_l^k\mu_{nc}^k + \frac{1}{2}\left(\zeta_l\omega_l^k\sigma_{nc}^k\right)^2\right)\frac{\Phi\left(\dfrac{\mu_{nc}^k}{\sigma_{nc}^k} + \zeta_l\omega_l^k\sigma_{nc}^k\right)}{\Phi\left(\dfrac{\mu_{nc}^k}{\sigma_{nc}^k}\right)},\end{aligned} \tag{A.5}$$

where $f_{Q_l^k}(x)$ is the probability density function of $s(\theta_j^k)$. Since N_t^k is a Poisson process, we have $P(N_t^k = n) = \dfrac{(\lambda t)^n}{n!}\exp(-\lambda t)$ and

$$\begin{aligned}\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)] &= \textstyle\sum_{n=0}^{\infty}\dfrac{(\lambda t)^n}{n!}\exp(-\lambda t)(M_l^k)^n \\ &= \textstyle\sum_{n=0}^{\infty}\dfrac{(\lambda M_l^k t)^n}{n!}\exp(-\lambda t).\end{aligned} \tag{A.6}$$

According to the Taylor series expansion of the exponential function, we have $\sum_{n=0}^{\infty}\dfrac{(\lambda M_l^k t)^n}{n!} = \exp(\lambda M_l^k t)$. Consequently, $\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)] = \exp(\lambda(M_l^k - 1)t)$. Combing the Eqs. (A.3) and (A.6), the Eq. (7) is given by

$$\begin{aligned}\mathrm{h}_{\mathrm{t}}^{\mathrm{k}} &= \frac{\lambda^k}{\eta^k}\left(\frac{t}{\eta^k}\right)^{\lambda^k-1}\prod_{\mathrm{l}=1}^{\mathrm{p}}\mathrm{E}[\exp(\omega_l^k D_{l,t}^k)]\mathrm{E}[\exp(\omega_l^k Q_{l,t}^k)] \\ &= \frac{\lambda^k t^{\lambda^k-1}}{\eta^{k\lambda^k}}\prod_{\mathrm{l}=1}^{\mathrm{p}}(1-\omega_l^k\beta_l^k)^{-\alpha_l^k t}\exp(\lambda(M_l^k-1)t),\end{aligned} \tag{A.7}$$

where $M_l^k = \exp\left(\zeta_l\omega_l^k\mu_{nc}^k + \dfrac{1}{2}(\zeta_l\omega_l^k\sigma_{nc}^k)^2\right)\dfrac{\Phi\left(\dfrac{\mu_{nc}^k}{\sigma_{nc}^k} + \zeta_l\omega_l^k\sigma_{nc}^k\right)}{\Phi\left(\dfrac{\mu_{nc}^k}{\sigma_{nc}^k}\right)}$. This proves Eq. (7).

Appendix B The proof of Eq. (14)

At time t, a series of observations $\{\boldsymbol{\varepsilon}_{1,j}^k,\ldots,\boldsymbol{\varepsilon}_{t,j}^k, j=1,\ldots,n\}$ has been collected, where $\boldsymbol{\varepsilon}_{t,j}^k = (\varepsilon_{1,t}^k,\ldots,\varepsilon_{q,t}^k)^{\top}$ denotes the vector of q random errors observed at the jth sampling of stage k. Considering a test statistic R_t^k of stage k at time t, which is given by

$$\begin{aligned}R_t^k &= \max_{1\leqslant\tau^k\leqslant t}[l_{\tau^k,t}(\boldsymbol{\Sigma}_0^k,\boldsymbol{\Sigma}_1^k) - l_{\infty,t}(\boldsymbol{\Sigma}_0^k)] \\ &= \max_{1\leqslant\tau^k\leqslant t}[\textstyle\sum_{i=1}^{t}\sum_{j=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_0^k,\boldsymbol{\Sigma}_1^k) \\ &\quad - \textstyle\sum_{i=1}^{t}\sum_{j=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_0^k)] \\ &= \max_{1\leqslant\tau^k\leqslant t}[\textstyle\sum_{i=\tau^k+1}^{t}\sum_{j=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_1^k) \\ &\quad - \textstyle\sum_{i=\tau^k+1}^{t}\sum_{j=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_0^k)],\end{aligned} \tag{A.8}$$

where $G(\cdot)$ is the pdf of $\boldsymbol{\varepsilon}_{i,j}^k$ and τ^k is the change point from $\boldsymbol{\Sigma}_0^k$ to $\boldsymbol{\Sigma}_1^k$. Let

$$\begin{aligned}l_{\tau^k}(\boldsymbol{\Sigma}_1^k) &= \textstyle\sum_{j=\tau^k+1}^{t}\sum_{i=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_1^k), \\ l_{\tau^k}(\boldsymbol{\Sigma}_0^k) &= \textstyle\sum_{j=\tau^k+1}^{t}\sum_{i=1}^{n}\log G(\boldsymbol{\varepsilon}_{i,j}^k|\boldsymbol{\Sigma}_0^k).\end{aligned} \tag{A.9}$$

Based on Eq. (10), $l_{\tau^k}(\boldsymbol{\Sigma}_0^k) = -\dfrac{n(t-\tau^k)}{2}[\log|\boldsymbol{\Sigma}_0^k| + \mathrm{tr}(\boldsymbol{\Sigma}_0^{k-1}\boldsymbol{S}_t^k) + N\log(2\pi)]$ and $l_{\tau^k}(\boldsymbol{\Sigma}_1^k) = -\dfrac{n(t-\tau^k)}{2}[\log|\boldsymbol{\Sigma}_1^k| + \mathrm{tr}(\boldsymbol{\Sigma}_1^{k-1}\boldsymbol{S}_t^k) + N\log(2\pi)]$, where $\boldsymbol{S}_t^k = \sum_{j=1}^{n}\boldsymbol{\varepsilon}_{t,j}^{k\,T}\boldsymbol{\varepsilon}_{t,j}^k/n$. Taking $l_{\tau^k}(\boldsymbol{\Sigma}_1^k)$ and $l_{\tau^k}(\boldsymbol{\Sigma}_0)$ into Eq. (A.8), the R_t^k is given by

$$\begin{aligned}R_t^k &= \max_{1\leqslant\tau^k\leqslant t}\frac{-n(t-\tau^k)}{2}[\log|\boldsymbol{\Sigma}_1^k| - \log|\boldsymbol{\Sigma}_0^k| \\ &\quad + \mathrm{tr}(\boldsymbol{\Sigma}_1^{k-1}\boldsymbol{\Sigma}_1^k) - \mathrm{tr}(\boldsymbol{\Sigma}_0^{k-1}\boldsymbol{\Sigma}_1^k)] \\ &= \max_{1\leqslant\tau^k\leqslant t}\frac{n(t-\tau^k)}{2}[\mathrm{tr}(\boldsymbol{\Sigma}_0^{k-1}\boldsymbol{\Sigma}_1^k) - \log|\boldsymbol{\Sigma}_0^{k-1}\boldsymbol{\Sigma}_1^k| - N].\end{aligned}$$

According to (Jing et al., 2022) and (Anderson, 1958),

$\mathrm{tr}(\boldsymbol{\Sigma}_0^{k^{-1}}\boldsymbol{\Sigma}_1^k) - \log|\boldsymbol{\Sigma}_0^{k^{-1}}\boldsymbol{\Sigma}_1^k| - N$ can be approximated as $\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_0^{k^{-1}}\boldsymbol{\Sigma}_1^k - \boldsymbol{I})^2$. Thus, the R_t^k is given by

$$R_t^k = \max_{1\leqslant\tau^k\leqslant t} \frac{n(t-\tau^k)}{4}\mathrm{tr}(\boldsymbol{\Sigma}_0^{k^{-1}}\boldsymbol{\Sigma}_1^k - \boldsymbol{I})^2, \qquad (A.11)$$

where $\boldsymbol{\Sigma}_1^k = \sum_{i=\tau^k+1}^{t}\sum_{j=1}^{n}\boldsymbol{\varepsilon}_{i,j}^{k\top}\boldsymbol{\varepsilon}_{i,j}^k/n(t-\tau^k)$. This proves Eq. (14).

References

Alaswad S, Xiang Y (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliability Engineering & System Safety, 157: 54–63

Anderson T W (1958). An Introduction to Multivariate Statistical Analysis (Vol. 2). Wiley New York

Bahrami H, Niaki S T A, Khedmati M (2021). Monitoring multivariate profiles in multistage processes. Communications in Statistics. Simulation and Computation, 50(11): 3436–3464

Bahria N, Chelbi A, Bouchriha H, Dridi I H (2019). Integrated production, statistical process control, and maintenance policy for unreliable manufacturing systems. International Journal of Production Research, 57(8): 2548–2570

Bera S, Mukherjee I (2016). A multistage and multiple response optimization approach for serial manufacturing system. European Journal of Operational Research, 248(2): 444–452

Boumallessa Z, Chouikhi H, Elleuch M, Bentaher H (2023). Modeling and optimizing the maintenance schedule using dynamic quality and machine condition monitors in an unreliable single production system. Reliability Engineering & System Safety, 235: 109216

Cao Y, Liu S, Fang Z, Dong W (2020). Modeling ageing effects in the context of continuous degradation and random shock. Computers & Industrial Engineering, 145: 106539

Cao Y, Wang P, Xv W, Dong W (2025). Joint optimization of quality-based multi-level maintenance and buffer stock within multi-specification and small-batch production. Frontiers of Engineering Management, 1–20

Cassady C R, Bowden R O, Liew L, Pohl E A (2000). Combining preventive maintenance and statistical process control: A preliminary investigation. IIE Transactions, 32(6): 471–478

Chen Y, Jin J (2005). Quality-reliability chain modeling for system-reliability analysis of complex manufacturing processes. IEEE Transactions on Reliability, 54(3): 475–488

Cheng G, Li L (2020). Joint optimization of production, quality control and maintenance for serial-parallel multistage production systems. Reliability Engineering & System Safety, 204: 107146

Cheng G Q, Zhou B H, Li L (2018). Integrated production, quality control and condition-based maintenance for imperfect production systems. Reliability Engineering & System Safety, 175: 251–264

Ebadi M, Ahmadi-Javid A (2020). Control charts for monitoring multistage service processes with optimal queue performance. Communications in Statistics. Simulation and Computation, 49(9): 2472–2484

Feng T, He Y, Shi R, Li J, Dai W, Yu S (2025). Functional risk modeling and selective maintenance optimization approach for multi-stage manufacturing system considering operational robustness. Reliability Engineering & System Safety, 256: 110775

Hadian S M, Farughi H, Rasay H (2021). Joint planning of maintenance, buffer stock and quality control for unreliable, imperfect manufacturing systems. Computers & Industrial Engineering, 157: 107304

Han X, Wang Z, Xie M, He Y, Li Y, Wang W (2021). Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence. Reliability Engineering & System Safety, 210: 107560

He Y, Liu F, Cui J, Han X, Zhao Y, Chen Z, Zhou D, Zhang A (2019). Reliability-oriented design of integrated model of preventive maintenance and quality control policy with time-between-events control chart. Computers & Industrial Engineering, 129: 228–238

Hu J, Chen P (2020). Predictive maintenance of systems subject to hard failure based on proportional hazards model. Reliability Engineering & System Safety, 196: 106707

Huang S, Yang J, Xie M (2020). A double-sampling SPM scheme for simultaneously monitoring of location and scale shifts and its joint design with maintenance strategies. Journal of Manufacturing Systems, 54: 94–102

Jafari L, Naderkhani F, Makis V (2018). Joint optimization of maintenance policy and inspection interval for a multi-unit series system using proportional hazards model. Journal of the Operational Research Society, 69(1): 36–48

Jin R, Deng X, Chen X, Zhu L, Zhang J (2019). Dynamic quality-process model in consideration of equipment degradation. Journal of Quality Technology, 51(3): 217–229

Jing H, Li J, Bai K (2022). Directional monitoring and diagnosis for covariance matrices. Journal of Applied Statistics, 49(6): 1449–1464

Kampitsis D, Panagiotidou S (2022). A Bayesian condition-based maintenance and monitoring policy with variable sampling intervals. Reliability Engineering & System Safety, 218: 108159

Khatab A, Diallo C, Aghezzaf E H, Venkatadri U (2019). Integrated production quality and condition-based maintenance optimisation for a stochastically deteriorating manufacturing system. International Journal of Production Research, 57(8): 2480–2497

Kong D, Balakrishnan N, Cui L (2017). Two-phase degradation process model with abrupt jump at change point governed by Wiener process. IEEE Transactions on Reliability, 66(4): 1345–1360

Li S, Yang Z, He J, Li G, Yang H, Liu T, Li J (2023). A novel maintenance strategy for manufacturing system considering working schedule and imperfect maintenance. Computers & Industrial Engineering, 185: 109656

Li Y, Peng S, Li Y, Jiang W (2020). A review of condition-based maintenance: Its prognostic and operational aspects. Frontiers of Engineering Management, 7(3): 323–334

Liu H C, Liu R, Gu X, Yang M (2023). From total quality management to Quality 4.0: A systematic literature review and future research agenda. Frontiers of Engineering Management, 10(2): 191–205

Liu P, Ma Y, Zhang C (2025). Stream of variation modeling and monitoring for heterogeneous profiles in multi-stage manufacturing

processes. IISE Transactions, 57(11): 1263–1276

Lu B, Luo Y (2024). A dynamic condition-based maintenance policy for heterogeneous-wearing tools with considering product quality deterioration. International Journal of Production Research, 62(19): 7096–7113

Lu B, Zhou X (2017). Opportunistic preventive maintenance scheduling for serial-parallel multistage manufacturing systems with multiple streams of deterioration. Reliability Engineering & System Safety, 168: 116–127

Lu B, Zhou X (2019). Quality and reliability oriented maintenance for multistage manufacturing systems subject to condition monitoring. Journal of Manufacturing Systems, 52: 76–85

Lv X, Shi L, He Y, He Z, Lin D K (2024). Joint optimization of production, maintenance, and quality control considering the product quality variance of a degraded system. Frontiers of Engineering Management, 11(3): 413–429

Majdouline I, Dellagi S, Mifdal L, Kibbou E, Moufki A (2022). Integrated production-maintenance strategy considering quality constraints in dry machining. International Journal of Production Research, 60(9): 2850–2864

Mehrafrooz Z, Noorossana R (2011). An integrated model based on statistical process control and maintenance. Computers & Industrial Engineering, 61(4): 1245–1255

Mohamed-Larbi R, Daoud A K (2024). Condition-based maintenance optimisation for multi-component systems using mean residual life. International Journal of Production Research, 62(13): 4831–4855

Montgomery D C (2019). Introduction to Statistical Quality Control. John wiley & sons. ISBN: 9781119646453

Naderkhani F, Makis V (2016). Economic design of multivariate Bayesian control chart with two sampling intervals. International Journal of Production Economics, 174: 29–42

Odom G J, Newhart K B, Cath T Y, Hering A S (2018). Multistate multivariate statistical process control. Applied Stochastic Models in Business and Industry, 34(6): 880–892

Olde Keizer M C A, Flapper S D P, Teunter R H (2017). Condition-based maintenance policies for systems with multiple dependent components: A review. European Journal of Operational Research, 261(2): 405–420

Panagiotidou S, Tagaras G (2010). Statistical process control and condition-based maintenance: a meaningful relationship through data sharing. Production and Operations Management, 19(2): 156–171

Park C, Lee J (2022). Monitoring profiles in multistage processes using the multivariate multiple regression model. Quality and Reliability Engineering International, 38(7): 3437–3450

Pedersen T I, Liu X, Vatn J (2023). Maintenance optimization of a system subject to two-stage degradation, hard failure, and imperfect repair. Reliability Engineering & System Safety, 237: 109313

Qiu P (2013). Introduction to statistical process control. Boca Raton, FL: Chapman Hall/CRC

Rasay H, Taghipour S, Sharifi M (2022). An integrated maintenance and statistical process control model for a deteriorating production process. Reliability Engineering & System Safety, 228: 108774

Sabri-Laghaie K, Fathi M, Zio E, Mazhar M (2022). A novel reliability monitoring scheme based on the monitoring of manufacturing quality error rates. Reliability Engineering & System Safety, 217: 108065

Salmasnia A, Abdzadeh B, Namdar M (2017). A joint design of production run length, maintenance policy and control chart with multiple assignable causes. Journal of Manufacturing Systems, 42: 44–56

Shang Y, Tsung F, Zou C (2013). Statistical process control for multistage processes with binary outputs. IIE Transactions, 45(9): 1008–1023

Shi H, Zhang J, Zio E, Zhao X (2023). Opportunistic maintenance policies for multi-machine production systems with quality and availability improvement. Reliability Engineering & System Safety, 234: 109183

Shi J (2023). In-process quality improvement: Concepts, methodologies, and applications. IISE Transactions, 55(1): 2–21

Shi L, Lv X, He Y, He Z (2024). Optimising production, maintenance, and quality control for imperfect manufacturing systems considering timely replenishment. International Journal of Production Research, 62(10): 3504–3525

Shi L, Lv X, Wang Y, He Y, He Z (2025). Joint optimization strategy based on a three-state manufacturing system considering dynamic buffers. International Journal of Production Economics, 283: 109587

Shojaee M, Noori S, Jafarian-Namin S, Johannssen A (2024). Integration of production–maintenance planning and monitoring simple linear profiles via Hotelling's T^2. Computers & Industrial Engineering, 188: 109864

Taguchi G (1986). Introduction to quality engineering: Designing quality into products and processes

Tasias K A (2022). Integrated quality, maintenance and production model for multivariate processes: A Bayesian approach. Journal of Manufacturing Systems, 63: 35–51

Tasias K A (2024). Simultaneous optimization of inventory, maintenance, and quality for production systems subject to multiple mean and variance shifts. Communications in Statistics-Theory and Methods, 53(9): 3078–3101.

Vlok P, Coetzee J, Banjevic D, Jardine A K, Makis V (2002). Optimal component replacement decisions using vibration monitoring and the proportional-hazards model. Journal of the Operational Research Society, 53(2): 193–202

Wan Q, Chen L, Zhu M (2023). A reliability-oriented integration model of production control, adaptive quality control policy and maintenance planning for continuous flow processes. Computers & Industrial Engineering, 176: 108985

Wang J, Luo L, Mu G, Ma Y, Ni C (2025). Joint optimization of quality control and maintenance policy for a production system with quality-dependent failures. Expert Systems with Applications, 272: 126800

Wang L, Lu Z, Ren Y (2020). Joint production control and maintenance policy for a serial system with quality deterioration and stochastic demand. Reliability Engineering & System Safety, 199: 106918

Wang Y, Xia T, Xu Y, Si G, Wang D, Pan E, Xi L (2024). Quality-centered production and maintenance scheduling for multi-machine manufacturing systems under variable operating conditions. Reliability Engineering & System Safety, 250: 110264

Wu C J, Hamada M S (2011). Experiments: Planning, Analysis, and

Optimization. John Wiley & Sons. ISBN: 9781119786891

Xiang Y (2013). Joint optimization of X control chart and preventive maintenance policies: A discrete-time Markov chain approach. European Journal of Operational Research, 229(2): 382–390

Ye Z, Cai Z, Zhou F, Zhao J, Zhang P (2019). Reliability analysis for series manufacturing system with imperfect inspection considering the interaction between quality and degradation. Reliability Engineering & System Safety, 189: 345–356

Yeganeh A, Johannssen A, Chukhrova N, Rasouli M (2024). Monitoring multistage healthcare processes using state space models and a machine learning based framework. Artificial Intelligence in Medicine, 151: 102826

Yeung T G, Cassady C R, Schneider K (2007). Simultaneous optimization of [Xbar] control chart and age-based preventive maintenance policies under an economic objective. IIE Transactions, 40(2): 147–159

Zhang J, Chen X, An Y, Zhang L, Shi H, Xu W (2025). Joint optimization of preventive maintenance and product quality improvement policies for deteriorating manufacturing systems with quality-reliability dependency. Reliability Engineering & System Safety, 255: 110661

Zhang Y, Yan M, Ke C (2024). Joint design of VSI Tr control chart and equipment maintenance in high quality process. Journal of Process Control, 136: 103177

Zhao X, Chen P, Tang L C (2025). Condition-based maintenance via Markov decision processes: A review. Frontiers of Engineering Management, 12(2): 330–342

Zheng M, Lin J, Xia T, Liu Y, Pan E (2023). Joint condition-based maintenance and spare provisioning policy for a K-out-of-N system with failures during inspection intervals. European Journal of Operational Research, 308(3): 1220–1232

Zheng R, Zhou Y (2021). Comparison of three preventive maintenance warranty policies for products deteriorating with age and a time-varying covariate. Reliability Engineering & System Safety, 213: 107676

Zou C, Tsung F (2008). Directional MEWMA schemes for multistage process monitoring and diagnosis. Journal of Quality Technology, 40(4): 407–427