APEM journal
Advances in Production Engineering & Management
Volume 21 | Number 2 | June 2026 | pp 239–253
https://doi.org/10.14743/apem2026.2.571
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

An integrated framework for production lot sizing and predictive maintenance based on Remaining Maintenance Life and Proportional Hazard Models

Zhong, C.W.[a], Lei, L.[a], Zhang, H.[b,*], Lu, M.L.[a,*]

[a]School of Economics and Management, Beijing Institute of Petrochemical Technology, Beijing, P.R. China
[b]Northeast Asian Studies College, Jilin University, Changchun, P.R. China

A B S T R A C T

Equipment reliability is critical to maintaining production efficiency and controlling manufacturing costs. Predictive maintenance (PdM), based on machine condition prediction, can effectively reduce the risk of machine failure; consequently, machine degradation assessment and the prediction of remaining maintenance life (RML) are crucial for maintenance decision-making. Moreover, because equipment condition affects production planning, PdM should be integrated into the traditional economic production quantity (EPQ) model. The main contribution of this study is the introduction of RML as a decision-support metric linking equipment degradation prediction with production lot-sizing decisions, thereby enabling the joint optimization of EPQ and PdM policies. To minimize expected average cost, maintenance decisions are integrated into the production lot-sizing model to determine the optimal production lot size and maintenance policy. This study considers a single-machine production process in which the ARMA method is used to forecast the machine degradation index. Cox's proportional hazard model (PHM) is then employed to estimate machine reliability based on the predicted degradation index. Based on this reliability assessment, remaining maintenance life (RML), rather than the traditional remaining useful life (RUL), is employed to link degradation prediction with the EPQ model and to characterize the machine deterioration process. An integrated EPQ–PdM model is developed to jointly determine the optimal production lot size and maintenance policy while minimizing the expected average cost (EAC) over the production cycle. Finally, a case study of an automotive bumper factory demonstrates the effectiveness of the proposed framework. The results show that the framework reduces EAC by 19.6 % relative to the current production strategy and identifies an optimal maintenance threshold of R_{safe} = 0.4 with six maintenance cycles.

A R T I C L E I N F O

Keywords:
Economic production quantity (EPQ);
Predictive maintenance;
Proportional hazard model (PHM);
Machine degradation;
Remaining maintenance life (RML);
Production lot sizing;
Reliability assessment;
Maintenance optimization;
Condition-based maintenance (CBM);
Integrated production-maintenance optimization

**Corresponding author:*
zhui24@mails.jlu.edu.cn
(Zhang, H.)
lumilin@bipt.edu.cn
(Lu, M.L.)

Article history:
Received 24 September 2025
Revised 22 June 2026
Accepted 25 June 2026

1. Introduction

Currently, many manufacturers are examining various operational strategies designed to reduce manufacturing costs while controlling inventory, thereby increasing profits. Moreover, extensive research has examined an effective decision-making tool for manufacturing processes, namely the economic production quantity (EPQ) model [1].

Within the framework of the traditional economic production quantity model, a fundamental premise is that the manufacturing process operates continuously without breakdowns or malfunctions, which means the traditional EPQ formulation usually excludes explicit consideration of product output control during the modeling process. Rosenblatt and Lee [2] reported that the optimal economic production quantity tends to decrease relative to the classical model when random machine failures are taken into account in the production system. Furthermore, Hariga and Ben-Daya [3] extended the analysis by assuming that system deterioration follows a general probability distribution. In addition, Misra [4] first studied an economic production quantity model that considered component failure, while Rabiu and Majahar Ali [5] extended the classical EPQ model by including Weibull-distributed degradation, dynamic demand, and quality-related defects, aiming to minimize total variable costs in perishable goods production.

A substantial body of research has established that maintenance activities constitute a critical component of production system operation, primarily because they exert a direct impact on maintaining consistent product quality and regulating production output. Since Barlow [6] pioneered the concept of a replacement model with minimal repair in the 1960s, the field of maintenance modeling and analysis has attracted continuous academic interest. Effective application of preventive maintenance (PM) plays a crucial role in preserving the operational health of production systems, thereby ensuring that the manufactured outputs consistently meet required quality specifications. Most PM models assume that the system is "as good as new" after each PM action. Lee and Srinivasan [7] developed an integrated inventory and PM model; however, a more realistic approach describes the failure rate as lying somewhere between "as good as new" and "as old as new". This concept, discussed by authors such as Nakagawa [8] and Pham and Wang [9], is known as imperfect maintenance. In this paper, it is assumed that the reduction in the age of the equipment is proportional to the cost of PM. These changes in the age of the equipment modify the distribution of time required for the system to enter an out-of-control state, thereby directly impacting the volume of nonconforming products.

Beginning in the 1980s, predictive maintenance (PdM) emerged as an important research focus in the field of maintenance, attracting extensive scholarly attention. In predictive maintenance (PdM), often described as condition-based preventive maintenance (PM) [10], the timing and frequency of maintenance are aligned with the current age or health status of the system, thereby producing a schedule that approaches optimality. Prior studies have demonstrated that PdM yields superior outcomes in terms of system reliability, availability, and average performance, while simultaneously incurring the lowest maintenance costs [11]. Furthermore, by forecasting and preventing unplanned equipment failures using observed data, overall maintenance expenses can be further reduced [12, 13]. Considering scheduled preventive replacement policies and stochastic machine breakdowns, Tse and Makis [14] proposed a cost-control model to manage the long-term expected average cost, in which the optimal economic lot size and preventive replacement timing were jointly determined. Recent studies have further advanced condition-based and predictive maintenance in production systems. Bányai [15] examined agile condition-based maintenance from the perspective of production cost efficiency, while BouAbid *et al.* [16] investigated integrated production and maintenance policies under product quality degradation. Zivlak et al. [17] further extended condition-based maintenance optimization to multi-component mass-production systems considering customer satisfaction. Related studies have also emphasized the role of fault prediction, equipment health assessment, and maintenance-policy evaluation in industrial production contexts. Pang [18] applied deep learning to adaptive fault prediction and maintenance in production lines, while He [19] investigated fault prediction in petroleum machinery production using signal-processing-based methods. Jankovič *et al.* [20] investigated machine-learning approaches for production-system condition monitoring and machine-condition prediction. Selech *et al.* [21] analysed reliability assessment and corrective-maintenance costs from an economic perspective, while Shih [22] examined the influence of maintenance policies on production-system performance. These studies indicate increasing attention to production–maintenance integration under degradation, quality, and system-complexity considerations. Nevertheless, the direct incorporation of RML-based predictive maintenance information into the economic production quantity (EPQ) decision structure remains insufficiently explored. This paper is therefore undertaken to contribute toward closing this research gap.

The implementation of predictive maintenance (PdM) faces two primary challenges: evaluating the degradation of machine performance and forecasting its remaining useful life (RUL), where accurate degradation assessment is fundamental to reliable RUL estimation. Moreover, as the duration extending from the current operational age of equipment until the end of its functional lifetime, RUL is one of the key factors in condition-based maintenance [23]. In this study, remaining maintenance life (RML) is used in place of RUL because it is better suited to the proposed model. Maintenance decisions are based on RML, and the lengths of successive maintenance cycles are linked through the relationship among RML values introduced by Alley [24] to assess failures in Air Force building systems. Pan *et al*. [25] applied this concept to minimize job tardiness under predictive maintenance.

To evaluate system degradation, we adopt the proportional hazard model (PHM), which was introduced by Cox [26] in 1972. This model has proven effective in applications ranging from age-related studies in medicine to reliability prediction in life-testing experiments. Subsequently, Kumar and Klefsjö [27] provided an extensive survey of the relevant PHM literature. In addition, Makis and Jardine [28] established an optimal maintenance framework based on the PH model, aiming to reduce the long-run expected average cost. Subsequently, Ghasemi *et al*. [29] advanced this line of research by proposing an optimal replacement policy for deteriorating systems within the context of a partially observable PH model. The PH model is commonly employed to characterize system degradation; however, sufficient data are required to support the model. Furthermore, in pursuit of this goal, prior studies have employed a diverse set of predictive methodologies, such as exponential smoothing [30], recurrent neural networks [31] and autoregressive moving average (ARMA) model [32]. Among these methods, the ARMA model demonstrates strong performance in short-term prediction tasks, primarily due to its assumption that the observed time series arises from a linear process [12].

Building on recent studies and industrial needs, this study proposes an integrated EPQ–PdM framework in which Remaining Maintenance Life (RML) is used as an explicit decision-support bridge between equipment degradation prediction and production lot-sizing optimization. Compared with previous EPQ-related production–maintenance models, which mainly consider random failures, imperfect production processes, preventive maintenance, or replacement decisions, the proposed framework directly embeds predictive-maintenance information into the lot-sizing decision structure. Specifically, ARMA-based degradation prediction is linked with PHM-based reliability assessment to estimate RML, and the resulting RML information is then incorporated into the EPQ model to jointly determine the maintenance threshold, the number of maintenance cycles, and the economic production lot size under the objective of minimizing the expected average cost. This positioning distinguishes the proposed framework from existing integrated EPQ–maintenance models by transforming degradation prediction from a separate prognostic output into an operational decision variable for production and maintenance optimization.

The paper is organized as follows. Section 2 introduces the background of the prediction process. The problem is defined in Section 3, together with the necessary notation and assumptions. In Section 4, the mathematical model is developed, while Section 5 describes the solution procedure and provides illustrative examples. Finally, Section 6 summarizes the paper and presents concluding remarks and suggestions for future research.

2. Background knowledge

2.1 Autoregressive moving average (ARMA)

An ARMA (p, q) model for time series y_t is given as follows:

$$y_t = c + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j} + \varepsilon_t \tag{1}$$

where c is a constant, p is the autoregressive order and q is the moving-average order; φ_i represents the autoregressive coefficients, ϕ_j represents the moving-average coefficients, and ε_t is a normal white noise process with zero mean and variance σ^2.

Specifically, when $q = 0$, ARMA (p, q) reduces to a pure autoregressive model AR(p) as follows:

$$y_t = c + \sum_{i=1}^{p} \varphi_i y_{t-i} + \varepsilon_t \tag{2}$$

Similarly, when $p = 0$, a moving-average model MA (q) is obtained as:

$$y_t = c + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j} + \varepsilon_t \tag{3}$$

Box *et al*. introduced a three-phase iterative approach for developing ARMA models in time-series analysis, consisting of model identification, parameter estimation and diagnostic checking, with detailed procedures available in reference [33]. In this study, the Green function is employed to characterize the weight of white noise terms on the predicted values, while the autocorrelation function (ACF), partial autocorrelation function (PACF), and information criteria are jointly used for ARMA model identification and selection. The ARMA model is adopted mainly because the proposed framework focuses on short-term degradation forecasting for a single-machine production process, and the available condition-monitoring data are primarily a univariate vibration time series. Under this setting, ARMA provides a parsimonious, transparent, and computationally efficient forecasting method that can be estimated with limited data and directly linked to the subsequent PHM-based reliability assessment. Although more recent prognostic approaches, such as LSTM, machine learning models, and hybrid prognostic models, may capture more complex nonlinear degradation patterns, they generally require larger training datasets, richer feature inputs, and careful hyperparameter tuning. In the present case study, applying such methods may introduce additional model complexity and potential overfitting without necessarily improving the interpretability of the integrated EPQ–PdM decision framework. Therefore, ARMA is selected as a suitable degradation prediction method for the current RML-based optimization framework, while more advanced data-driven prognostic models can be further examined in future research when larger and richer condition-monitoring datasets are available.

2.2 Proportional hazard model (PHM)

The basic hazard function can be written as:

$$\lambda(t, y_t) = \lambda_0(t)\exp(\alpha \cdot y_t) \tag{4}$$

In this formulation, $\lambda_0(t)$ represents the baseline hazard function, y_t refers to the vector of time-dependent covariates, and α denotes the regression coefficient vector. In this study, y_t are derived from the ARMA (p, q) model rather than through direct observation, while the remaining coefficients are estimated by maximizing the partial likelihood function. To better reflect practical production conditions, $\lambda_0(t)$ is assumed to follow a Weibull distribution, thereby yielding the Weibull proportional hazard model (WPHM), expressed as:

$$\lambda(t, y_t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\exp(\alpha \cdot y_t) \tag{5}$$

Given the failure density function $f(t, y_t) = \lambda(t, y_t) \cdot R(t, y_t)$, the partial likelihood function is expressed as:

$$L(\beta, \eta, \alpha) = \prod_{i=1}^{r} \lambda(t_i, y_{t_i}) \cdot \prod_{j=1}^{N} R\left(t_j, y_{t_j}\right) \tag{6}$$

Parameter-estimation algorithms for PHM have been widely studied. Ripatti and–Palmgren [34] addressed the estimation of regression coefficients by constructing the partial likelihood function and applying the maximum likelihood method. Moreover, Ducrocq and Casella [35] improved the method using Bayesian estimation, and Cortiñas Abrahantes and Burzykowski [36] estimated the parameters using the Laplace transform method. In our study, the Newton-Raphson iteration method is adopted to estimate three parameters: α, β and η.

2.3 Remaining maintenance life (RML)

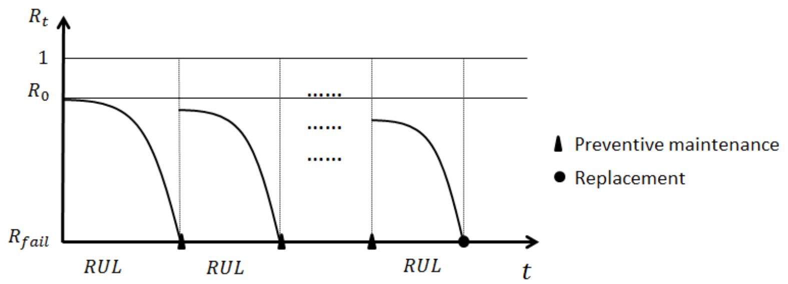
Reliability, rather than the failure rate, is selected to describe the degradation process due to its appropriate mathematical properties. Moreover, reliability can be readily linked to the failure rate obtained from the PH model through the following equation:

$$R(t) = \exp(-\int_0^t \lambda(\xi)d\xi) \tag{7}$$

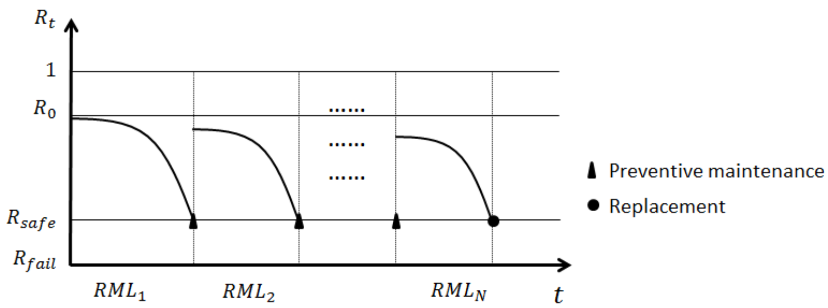
The initial reliability of the machine is R_0, and the reliability will monotonically decrease towar R_{fail}, which indicates the failure threshold of the machine. The time span between the current age and expected service life is the remaining service life (RSL), also referred to as the remaining useful life (RUL; i.e., $RUL = T_{R_{fail}} - T_{R_0}$). To avoid the huge loss of time and other resources caused by the breakdown, a safety threshold R_{safe} will be used to trigger preventive maintenance before breakdown occurs. The time between the current age and a scheduled maintenance or repair action is defined as the remaining maintenance life (RML; i.e. $RML = T_{R_{safe}} - T_{R_0}$).

Through the application of the concept of RML, we can directly connect machine-life prediction to the maintenance strategy; moreover, the time span between RUL and RML may serve as a buffer, which not only helps address unexpected situations but also allows for slight adjustments when the prediction varies. The distinction between RUL and RML is illustrated in Fig. 1.

Considering the degradation characteristics, the decline rate will increase with the machine's age and number of maintenance actions in each preventive maintenance cycle, namely $R_{new} \neq R_0$. Thus, the machine needs more frequent maintenance and the maintenance cycle will be shortened. In order to describe this influence on the system failure rate, Malik introduced the recursion factor k $(1 \leq k \leq \infty)$ [37]. After maintenance is performed at age t, the system returns to the state corresponding to age t/k. In this study, a similar factor δ_i $(0 < \delta_i < 1)$ is used to represent this effect on the machine's age.



a) Illustration of RUL and degradation process



b) Illustration of RML and degradation process

Fig. 1 Comparison between RUL and RML

The RML of each maintenance cycle is expressed as follows:

$$RML_i = \begin{cases} T_{R_{safe}} - T_{R_0}, & i = 1 \\ \delta_{i-1}\left(T_{R_{safe}} - T_{R_{new}}\right), & i = 2 \\ \delta_{i-1} RML_{i-1}, & 3 \leq i \leq N \end{cases} \tag{8}$$

2.4 Sketch of proposed method

Fig. 2 presents the proposed framework for degradation assessment and RML forecasting. The framework is divided into three essential phases, namely *machine performance identification*, *hazard rate estimation* and *RML prediction*. The function of each of these steps is discussed as follows:

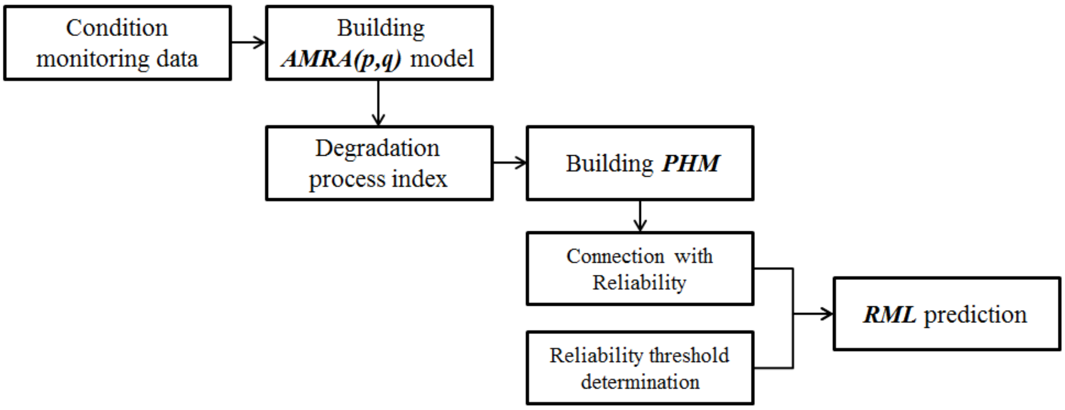


Fig. 2 Sketch of the RML prediction

To build the forecasting model, normal machine-condition data are first acquired through a monitoring process. An ARMA model is then used to represent the degradation process, while the PHM determines the trend in the hazard rate. Finally, a hazard threshold is used to forecast RML.

3. System operation, notation and assumptions

In this study, the production environment is modeled as a single-item system. At the beginning of every cycle, the inventory level is empty, and production takes place at a rate P, whereas the demand is satisfied continuously at a rate D. The analysis considers the inventory function $I(t)$, where the cycle length T_I comprises the total production duration T_P and the subsequent depletion period. For modeling purposes, neither backlog nor stock-out situations are allowed. The production cycle length T_P is measured from the commencement of production activities to the termination of the associated maintenance phase. Within this time span, inventory grows at a net inflow rate of (P-D) during production runs, while during each scheduled maintenance period T_{pr}, inventory is reduced at the demand rate D.

These assumptions define the initial scope of the proposed model. The present framework is developed for a single-machine production system with constant demand and deterministic production rate, which allows the interaction between RML-based maintenance scheduling and EPQ-based lot-sizing decisions to be analyzed in a tractable manner. Under more realistic production conditions, such as variable demand, multiple machines, or stochastic production rates, the model structure would need to be extended. Variable demand would affect the inventory depletion process and may require a dynamic or stochastic demand function. Multiple-machine systems would require equipment-specific degradation states, RML values, and maintenance coordination among machines. Stochastic production rates would further introduce uncertainty into the production cycle length and expected cost function. Therefore, the current model should be interpreted as a baseline framework that can be extended to more complex production environments rather than as a fully generalized model for all manufacturing systems.

The deterioration state of the system is represented by the reliability $R(t)$, which decreases as degradation progresses. By employing observed historical records, the trajectory of $R(t)$ can be estimated using a combination of an ARMA model and a PH model. When $R(t)$ falls to or below the specified reliability threshold R_{safe}, production activities must be suspended and preventive maintenance must be performed. The resulting inventory cycle under predictive maintenance is illustrated in Fig. 3.

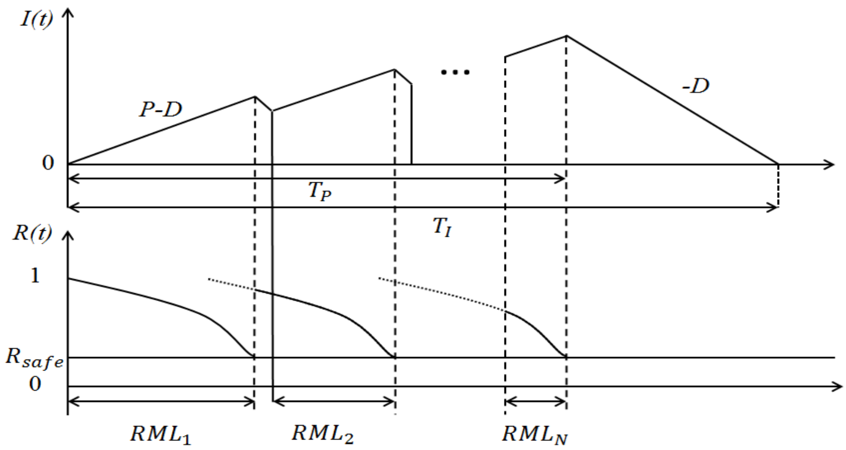


Fig. 3 The inventory cycle of the EPQ model with predictive maintenance

We assume that the system degradation trend is the same in each maintenance cycle, as seen in Fig. 1. The whole production cycle T_P is divided into N maintenance cycles by the maintenance critical value R_{safe}, and the length of each maintenance cycle is denoted as RML_i, $i = 1,2,\dots,N$. Furthermore, preventive maintenance is performed for a duration t_{pr} with the associated cost denoted as c_{pr}. At the conclusion of the final cycle, the equipment is assumed to be replaced, and this activity involves a cost of c_r.

4. Model construction

4.1 Time model

The total production cycle time T can be divided into three parts: the production time T_P, the total maintenance time T_M and the inventory consumption time T_I.

The production time T_P consists of RML_i ($i = 1,2,\dots,N$), and the total maintenance time includes two parts: T_M the preventive maintenance time T_{pr} and the replacement time T_r. Hence, we have:

$$T_P = \sum_{i=1}^{N} RML_i \qquad i = 1,2,\dots,N \tag{9}$$

$$T_M = T_{pr} + T_r = (N-1)\cdot t_{pr} + T_r \tag{10}$$

The inventory consumption time can be represented by the following expression:

$$T_I = \frac{P-D}{D}\sum_{i=1}^{N} RML_i - (N-1)t_{pr} \tag{11}$$

Therefore, the overall duration of the production cycle T can be written as:

$$T = T_P + T_M + T_I = \frac{P}{D}\sum_{i=1}^{N} RML_i + T_r \tag{12}$$

4.2 Operation and maintenance cost

We introduce the operation cost $C_O(i,t) = c_c + i\cdot c_f + t\cdot c_t, i = 1,2,\dots,N$ to improve our model, in which three components are included: c_c is the constant operation cost which will take place in each maintenance cycle, c_f is a constant representing the effect of the number of maintenance actions on the operation cost, and c_t denotes the cost due to the age of the production system.

Considering the operation cost in the i-th maintenance cycle of $\int_0^{RML_i} C_O(i,t)dt$, the total operation cost can be expressed as:

$$C_o = \sum_{i=1}^{N} \int_0^{RML_i} C_O(i,t)dt = \sum_{i=1}^{N} \int_0^{RML_i} (c_c + i\cdot c_f + t\cdot c_t)dt \tag{13}$$

In this model, the overall maintenance cost C_M is defined as the sum of two components: the cost incurred from preventive maintenance operations C_{pr} and the cost associated with system replacement C_r. This relationship can be expressed mathematically as follows:

$$C_M = C_{pr} + C_r = (N-1)c_{pr} + C_r \tag{14}$$

4.3 Inventory cost

Since the holding cost per unit time h is a constant, the total inventory cost is only related to the inventory quantity. As shown in Fig. 4, the whole inventory can be separated into three parts: the first is the production period with an increasing rate of (P-D), the second comprises the solid triangles with a decreasing rate of D lasting t_M, and the last is the depletion process with a decreasing rate of D.

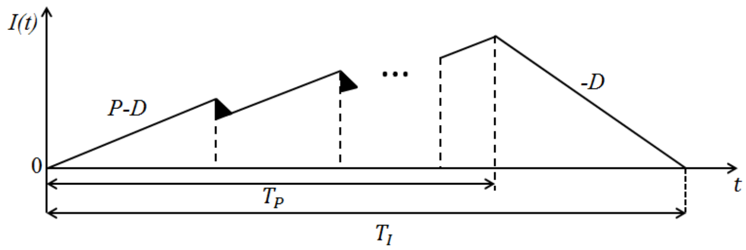


Fig. 4 Illustration of the inventory cost

In the production cycle, the linear equation for the inventory should be:

$$I(t)_i = \begin{cases} (P-D)t & i = 1 \\ (P-D)t - (i-1)P \cdot t_{pr}/(P-D) & i = 2,3,\dots,N \end{cases} \tag{15}$$

Hence, the inventory quantity for each maintenance cycle is:

$$S_{IP(i)} = \int_0^{RML_i} I(t)_i dt \tag{16}$$

Then, the area of the solid triangles is: $S_s = P \cdot t_{pr}{}^2/2$ Considering the total production quantity in the production cycle, the depletion area is $S_c = [(P-D)T_P - (N-1)P \cdot t_{pr}]^2/2D$

Using the area of each part, the total inventory quantity is obtained as:

$$S_I = \sum_{i=1}^{N} [S_{IP}(i) + (i-1)S_s] + S_c \tag{17}$$

Hence, the inventory cost $C_I = h\ S_I$ can be obtained from Eqs. 15-17.

4.4 Objective function

From the preceding analysis, the overall expected cost is shown to consist of the set-up cost, total maintenance cost, operation cost and the total inventory cost. Because the production set-up cost is assumed to occur only at the beginning of the cycle, it is represented by the constant cost C_s. Therefore, the expected average cost, denoted by EAC, is formulated as follows:

$$EAC = \frac{C_s + C_o + C_M + C_I}{T} \tag{18}$$

The optimization problem is therefore to determine the optimal reliability threshold R_{safe}^* and the optimal number of maintenance cycles N^* that minimize EAC.

The procedures of the search algorithm are shown below:

- Estimate the upper bound of the number of maintenance cycles N_{UP} from historical records or system conditions.
- Initialize $N = 1$ and EAC^* to a sufficiently large value.
- For a given interval [R_1, R_2], set R_{safe} = R_1 and calculate $EAC(R_{safe}, N)$. Increase R_{safe} by 0.2 within the interval. If $EAC^* > EAC(R_{safe}, N)$, set $EAC^* = EAC(R_{safe}, N)$, record R_{safe}^* and continue the search; otherwise, stop the search and record EAC^*.

- For $N = N + 1$, when $EAC^{*} > EAC(R_{safe}, N)$, set $EAC^{*} = EAC(R_{safe}, N)$, and continue the search until $N = N_{UP}$; otherwise, stop the search and record EAC^{*}.
- The EAC^{*} obtained after the search is the minimum expected average cost, with the corresponding (R_{safe}^{*}, N^{*}) being the optimal solution.

5. Case study

5.1 Data acquisition and parameter estimation

This section employs a case study of a bumper production plant to evaluate the effectiveness of the integrated EPQ framework. The bumper factory is one of the largest suppliers to GM in China, and we focus on the module injection process, which is the most important process in the whole production cycle.

In monitoring machine condition, various signal sources can be applied, such as vibration, sound, pressure, and temperature measurements. Of these indicators, vibration is most widely adopted, given its convenience in data collection and efficiency in subsequent processing. The data were recorded from July 2013 to January 2014, during continuous production for 8 hours per day. A portion of the recorded vibration data is shown in Fig. 5.

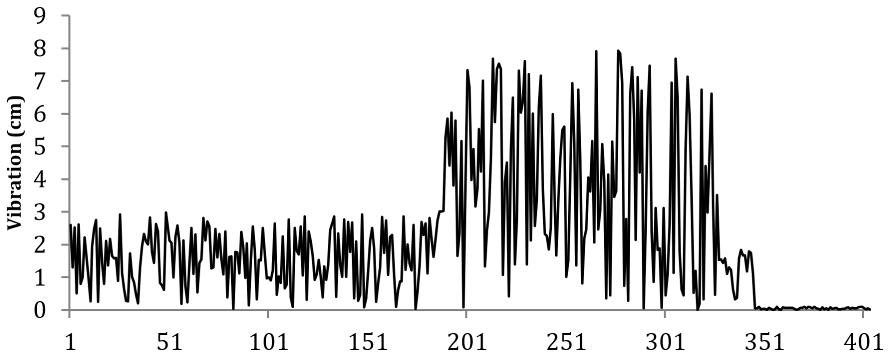


Fig. 5 Part of the vibration data from the injection process

An ARMA (2, 2) model is used for prediction, as discussed in Section 2.1. Based on the collected condition data, the autocorrelation function and partial autocorrelation function are used to select an appropriate ARMA model, while the model parameters are determined using Schwarz's Bayesian information criterion, as described in Section 2. The estimated parameters are listed in Table 1.

Table 1 Parameters for ARMA model

Parameters	φ_1	φ_2	θ_1	θ_2
Estimated value	-0.0241	0.175	-0.932	-0.114

The degradation process index is:

$$\widehat{H}_t = -0.0241 \cdot H_{t-1} + 0.175 \cdot H_{t-2} - 0.932 \cdot \varepsilon_{t-1} - 0.114 \cdot \varepsilon_{t-2} \tag{19}$$

The degradation process index will be used as the input to the PH-model. It should be noted that the present case study mainly aims to demonstrate the integration of degradation prediction, PHM-based reliability assessment, RML estimation, and EPQ-based production lot-sizing within a unified decision framework. Therefore, the forecasting component is evaluated through model identification and parameter estimation procedures rather than through a separate out-of-sample forecasting comparison. Although prediction performance metrics such as RMSE, MAE, and MAPE would provide a more direct assessment of degradation and RML prediction accuracy, the primary objective of this study is to develop and validate an integrated EPQ–PdM decision framework rather than to benchmark alternative prognostic algorithms. In the proposed framework, the forecasting component serves as an enabling module for reliability assessment, RML estimation, and production–maintenance optimization. A comprehensive evaluation of forecasting performance through RMSE, MAE, MAPE, and systematic out-of-sample validation would require dedicated training–validation procedures and larger condition-monitoring datasets. Such an investigation is beyond the intended scope of the present study but represents an important direction for future research.

According to Section 2, the Newton-Raphson iteration method has been employed to solve the parameter estimation problem by maximizing the partial likelihood function without specifying the baseline hazard function. In addition, the PH model in this case can be calculated as follows:

$$\lambda(t, H_t) = \frac{2.36}{4501.3}\left(\frac{t}{4501.3}\right)^{1.36} \exp(0.559 \cdot H_t) \tag{20}$$

To evaluate the machine's remaining maintenance life (RML), it is essential to define a reliability threshold, denoted as R_{safe}, which serves as the critical decision variable. To limit repeated repairs instead of necessary replacement, the upper bound on the number of repairs is set to $N_{UP} = 10$. The remaining parameters are listed in Table 2.

For different cycle lengths, the optimal maintenance cycle length, number of maintenance actions and economic production batch size can thus be determined. This is because the cycle length has a linear relationship with the prediction length, while the batch size also has a linear relationship with the production cycle length.

Table 2 Parameter values for the integrated model

Parameter type	Parameter Value
Batch parameter	$P = 20, D = 12$
Maintenance cost	$c_{pr} = 7500, C_r = 15000$
Production cost	$c_c = 5, c_f = 2, c_t = 0.01$
Recursion factor	$\delta = 0.9$
Batch cost	$h = 0.002, C_S = 30000$

5.2 Result analysis

Fig. 6 shows how EAC changes under different maintenance strategies. When $R_{safe} = 0.4$ and the number of maintenance cycles $N^* = 6$, the minimum expected average cost is 96.745, and the economic production quantity is 4052 (i.e. $EAC = 96.745$, $Q^* = 4052$). Compared with the current production scenario of $EAC = 120.4$, $Q = 3500$ (data from October 2015), the optimal strategy can reduce the EAC by (120.4-96.745)/120.4 = 19.6 %, thereby offering a smoother production process with a higher production quantity.

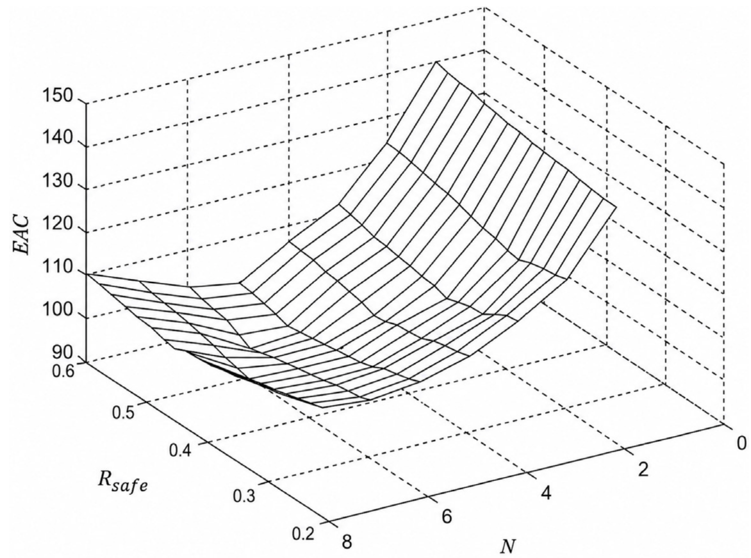


Fig. 6 The values of EAC for different maintenance strategies

According to the presented results, it can also be concluded that:

- When the number of maintenance cycles is $N = 1$, the expected average cost remains the highest across the considered range of R_{safe}. As the number of maintenance cycles increases, the expected average cost declines sharply, indicating the necessity of preventive maintenance in batch production;

- When the number of maintenance cycles varies between 5 and 7, the corresponding long-term expected costs tend to be equal despite changes in maintenance frequency, and they remain at the lowest level. This shows that maintenance should be performed before the equipment health status reaches its worst condition, thus proving the effectiveness of maintenance.

5.3 Discussion of main factors

In this section, several main factors will be discussed to determine their influence on the objective function.

Remaining maintenance life

Fig. 7 shows each maintenance cycle corresponding to the optimal solution, with each cycle length decreasing from $\mathrm{RML}_1 = 43.24$ h to $\mathrm{RML}_6 = 25.53$ h due to the recursion factor ($\delta = 0.9$). This change also demonstrates the system degradation pattern. Moreover, knowing the maintenance cycle lengths RML_i can support inventory control and production scheduling.

For the purpose of examining the robustness of the optimal solution under alternative degradation trajectories, the recursion factor δ is varied from 0.9 to 0.6, with the corresponding optimal maintenance strategies shown in Table 3. Since RML is derived from the predicted degradation process and PHM-based reliability estimation, changes in δ can be interpreted as scenario-based perturbations of the degradation trajectory and the resulting RML sequence. Therefore, this analysis provides an indirect robustness discussion on how the optimal maintenance policy responds to possible deviations in degradation and RML estimation.

As shown in Table 3, when δ decreases from 0.9 to 0.6, the optimal number of maintenance cycles changes only from 6 to 7, while the optimal reliability threshold decreases from $R_{\text{safe}} = 0.40$ to $R_{\text{safe}} = 0.34$. The expected average cost increases from 96.745 to 102.631, corresponding to an increase of approximately 6.1 % relative to the baseline case. These results indicate that the optimal solution does not change abruptly under the considered degradation-trajectory perturbations, and the proposed framework remains relatively stable within this scenario range. It is also worth noting that δ is influenced not only by the degradation rate but also by the maintenance level. Therefore, the relationship between preventive maintenance time, maintenance effectiveness, and δ should be further investigated in future research. However, this analysis should be regarded as a scenario-based robustness discussion rather than a full prediction-error sensitivity test. A more formal robustness analysis could be conducted in future studies by explicitly imposing positive and negative errors on RML or reliability estimates and examining their effects on R_{safe}, N, Q, and *EAC*.

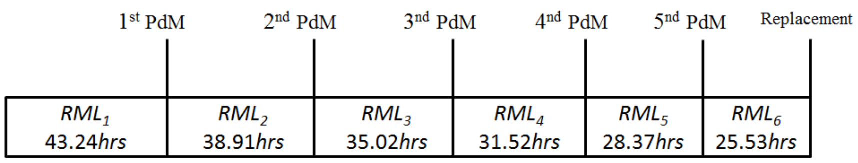


Fig. 7 Illustration of RML

Table 3 Optimal solution for different recursion factors

δ	N^*	$R_{safe}{}^*$	EAC^*
0.9	6	0.4	96.745
0.85	6	0.36	97.106
0.8	7	0.36	97.293
0.75	7	0.34	99.15
0.7	7	0.34	98.684
0.65	7	0.34	100.512
0.6	7	0.34	102.631

Operation cost

The total operation cost comprises three components: the constant operation cost, the cost related to the number of maintenance actions, and the cost related to the production age. All three components are positively related to the expected average cost. The optimal solution changes only slightly with each component of the operation cost, indicating that operation cost may not be a major factor in determining the maintenance strategy or production lot size. The corresponding results are summarized in Table 4.

Table 4 Optimal solutions under different maintenance costs

	Constant operation cost: c_C				Cost related to maintenance times: c_f			
	1	3	5	7	2	4	6	8
R_{safe}^*	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.38
N^*	6	6	6	6	6	6	6	6
EAC	96.056	96.44	96.745	97.104	96.745	96.912	97.13	97.348

Sensitivity comparison for different parameters

To evaluate the relative effects of multiple parameters on the long-term expected average cost and economic production quantity, a numerical sensitivity analysis is conducted under the same conditions, with each parameter varied stepwise by a ratio of 0.2. The resulting EAC sensitivity comparison is shown in Fig. 8.

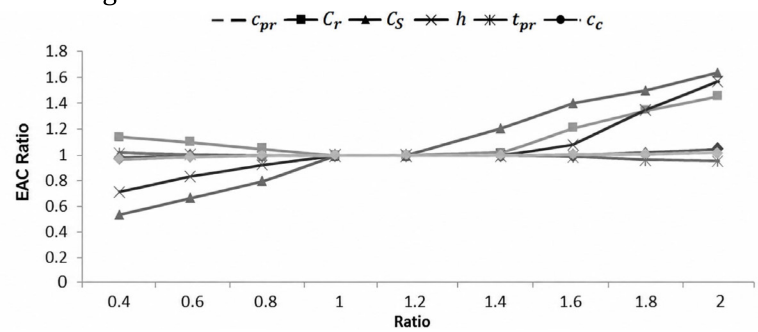


Fig. 8 Sensitivity comparison of EAC for different parameters

As shown in Fig. 8, the following conclusions can be drawn:

- The three most influential factors are the replacement cost C_r, set-up cost C_S and the holding cost rate h, which are positively related to EAC; moreover, the influence sequence will be $C_S > h > C_r$ when the parameters decrease while $C_S > C_r > h$ when they increase. Therefore, the enterprise should focus on controlling these three elements in actual production according to different trends of parameters. Other costs, such as the PM cost c_{pr} and the operation cost, may have little effect on EAC. The enterprise may therefore relax the corresponding constraints appropriately.
- The maintenance time t_{pc} has little effect on the model and is negatively related to EAC. This is because a short maintenance time will increase the inventory holding cost while shortening the cycle length.

Transferability of the proposed framework

Although the empirical validation is based on a single bumper production case, the proposed framework is not limited to this specific manufacturing setting. Its transferability mainly depends on whether the target production system satisfies several practical conditions: the production process contains a key machine or bottleneck piece of equipment whose degradation can be monitored, the degradation trajectory can be represented by condition-monitoring signals, the reliability level can be estimated through a PHM-type model, and the production and maintenance costs can be quantified for EPQ-based optimization. Under these conditions, the proposed ARMA–PHM–RML–EPQ framework can be adapted to other mass-production or continuous-production environments, such as machining, injection molding, chemical production, pharmaceutical manufacturing, and steel production. However, the model parameters, degradation indicators, cost structure, and maintenance threshold should be recalibrated according to the specific equipment, production process, and data characteristics of each application scenario. Therefore, the case study should be regarded as an illustrative validation of the framework rather than a universal empirical proof. Further validation using multiple manufacturing cases would strengthen the generalizability and practical applicability of the proposed approach.

6. Conclusion and future research

This paper develops an integrated EPQ–PdM framework for a degraded production system by linking ARMA-based degradation prediction, PHM-based reliability assessment, and RML-based maintenance decision-making. The proposed framework extends the traditional EPQ model by incorporating predictive maintenance information into production lot-sizing decisions, so that the maintenance threshold, maintenance cycle arrangement, and economic production quantity can be jointly optimized with the objective of minimizing the expected average cost.

The broader scientific contribution of this study lies in using Remaining Maintenance Life (RML) as a decision-support bridge between reliability engineering and production planning. In the proposed framework, RML is not treated only as a prognostic output, but is transformed into an operational decision variable that connects equipment health assessment, reliability estimation, maintenance scheduling, inventory evolution, and production lot-sizing optimization. In this sense, the RML-based framework provides a structured way to translate degradation and reliability information into production and maintenance decisions. The case study of an automotive bumper production process further illustrates the practical effectiveness of the proposed framework, showing that the optimized strategy can reduce the expected average cost while supporting a smoother production process and a higher economic production quantity. Although the empirical validation is based on a single case study, the proposed framework is not limited to this specific production setting. It may be transferred to other manufacturing environments when several practical conditions are satisfied: the production system contains key or bottleneck equipment with observable degradation signals; the degradation process can be linked to reliability assessment; and the production, inventory, and maintenance costs can be quantified for EPQ-based optimization. For large-scale production enterprises, such as chemical, pharmaceutical, steel, machining, and injection-molding manufacturers, the framework may provide a useful decision-support structure. However, the degradation indicators, PHM parameters, cost structure, maintenance thresholds, and production parameters should be recalibrated according to the specific equipment and process characteristics of each application scenario. Therefore, the present case study should be regarded as an illustrative validation of the framework rather than a universal empirical proof.

Future research could further extend the proposed framework to more complex and realistic manufacturing environments. First, variable demand could be incorporated by replacing the constant demand rate with time-dependent or stochastic demand, so that inventory dynamics and lot-sizing decisions can better reflect market fluctuations. Second, the current single-machine setting could be extended to multi-machine or interconnected production systems, where each machine may have its own degradation trajectory, RML sequence, maintenance threshold, and contribution to system-level production capacity. Third, stochastic production rates could be considered by modeling production output as an uncertain process, which would require reformulating the production cycle length and expected average cost function. In addition, rework, quality deterioration, more advanced long-horizon prognostic techniques, and the impact of forecasting uncertainty and prediction errors on maintenance decisions and production lot-sizing performance could be incorporated to further improve the robustness and practical applicability of the integrated EPQ–PdM framework.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The research was funded by the National Natural Science Foundation of China (71973051 and 71703009).

References

[1] Silver, E.A., Peterson, R. (1985). *Decision systems for inventory management and production planning,* 2nd edition, John Wiley & Sons, New York, USA.

[2] Rosenblatt, M.J., Lee, H.L. (1986). Economic production cycles with imperfect production processes, *IIE Transactions*, Vol. 18, No. 1, 48-55, doi: 10.1080/07408178608975329.

[3] Hariga, M., Ben-Daya, M. (1998). Note: The economic manufacturing lot-sizing problem with imperfect production processes: Bounds and optimal solutions, *Naval Research Logistics (NRL)*, Vol. 45, No. 4, 423-433, doi: 10.1002/(SICI)1520-6750(199806)45:4<423::AID-NAV8>3.0.CO;2-7.

[4] Misra, R.B. (1975). Optimum production lot size model for a system with deteriorating inventory, *International Journal of Production Research*, Vol. 13, No. 5, 495-505, doi: 10.1080/00207547508943019.

[5] Rabiu, S., Majahar Ali, M.K. (2024). An enhanced EPQ model: Incorporating Weibull distribution deterioration, defective items, and dynamic demands for sustainable production optimization, *Ain Shams Engineering Journal*, Vol. 15, No. 7, Article No. 102824, doi: 10.1016/j.asej.2024.102824.

[6] Barlow, R.E., Proschan, F. (1996). *Mathematical theory of reliability,* Society for Industrial and Applied Mathematics, Philadelphia, USA, doi: 10.1137/1.9781611971194.

[7] Lee, H.S., Srinivasan, M.M. (2001). A production/Inventory policy for an unreliable machine, In: Rahim, M.A., Ben-Daya, M. (eds.), *Integrated models in production planning, inventory, quality, and maintenance*, Springer, Boston, USA, 79-94, doi: 10.1007/978-1-4615-1635-4_4.

[8] Nakagawa, T. (1986). Periodic and sequential preventive maintenance policies, *Journal of Applied Probability*, Vol. 23, No. 2, 536-542, doi: 10.2307/3214197.

[9] Pham, H., Wang, H. (1996). Imperfect maintenance, *European Journal of Operational Research*, Vol. 94, No. 3, 425-438, doi: 10.1016/S0377-2217(96)00099-9.

[10] van Dinter, R., Tekinerdogan, B., Catal, C. (2022). Predictive maintenance using digital twins: A systematic literature review, *Information and Software Technology*, Vol. 151, Article No. 107008, doi: 10.1016/j.infsof.2022.107008.

[11] 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*, Vol. 210, Article No. 107560, doi: 10.1016/j.ress.2021.107560.

[12] Zhang, L., Wang, X., He, S., Mao, X., Li, B., Liu, H. (2024). Multi-models associated with process information-driven process autonomous digital twin for multi-variety production of intelligent machines, *Science China Technological Sciences*, Vol. 67, No. 12, 3825-3842, doi: 10.1007/s11431-024-2778-0.

[13] Wang, S., Guo, Y., Huang, S., Liu, D., Tang, P., Zhang, L. (2024). A spatial-temporal feature fusion network for order remaining completion time prediction in discrete manufacturing workshop, *International Journal of Production Research*, Vol. 62, No. 10, 3638-3653, doi: 10.1080/00207543.2023.2245487.

[14] Tse, E., Makis, V. (1994). Optimization of the lot size and the time to replacement in a production system subject to random failure, In: *Proceedings of the Third International Conference on Automation Technology*, Taipei, Taiwan, Vol. 4, 163-169.

[15] Bányai, Á. (2023). Impact of agile, condition-based maintenance strategy on cost efficiency of production systems, *Advances in Production Engineering & Management*, Vol. 18, No. 3, 317-326, doi: 10.14743/apem2023.3.475.

[16] BouAbid, H., Dhouib, K., Gharbi, A. (2024). Integrated production and maintenance policy for manufacturing systems prone to products' quality degradation, *Advances in Production Engineering & Management*, Vol. 19, No. 4, 512-526, doi: 10.14743/apem2024.4.521.

[17] Zivlak, N., Dong, M., Lalic, B., Yun, F.Z., Liu, Q.M. (2026). Condition-based maintenance optimization for multi-component systems in mass production considering customer satisfaction, *Advances in Production Engineering & Management*, Vol. 21, No. 1, 89-100, doi: 10.14743/apem2026.1.563.

[18] Pang, J.L. (2023). Adaptive fault prediction and maintenance in production lines using deep learning, *International Journal of Simulation Modelling*, Vol. 22, No. 4, 734-745, doi: 10.2507/IJSIMM22-4-CO20.

[19] He, D.X. (2024). Fault prediction in high-efficiency petroleum machinery production, *International Journal of Simulation Modelling*, Vol. 23, No. 1, 184-195, doi: 10.2507/IJSIMM23-1-CO5.

[20] Jankovič, D., Šimic, M., Herakovič, N. (2024). A comparative study of machine learning regression models for production systems condition monitoring, *Advances in Production Engineering & Management*, Vol. 19, No. 1, 78-92, doi: 10.14743/apem2024.1.494.

[21] Selech, J., Matijosius, J., Kilikevicius, A., Marinkovic, D. (2025). Reliability analysis and the costs of corrective maintenance for a component of a fleet of trams, *Tehnički Vjesnik – Technical Gazette,* Vol. 32, No. 1, 205-216, doi: 10.17559/TV-20241121002144.

[22] Shih, Y.C. (2024). Impacts of maintenance policies on fractal layout performance, *International Journal of Simulation Modelling*, Vol. 23, No. 3, 401-411, doi: 10.2507/IJSIMM23-3-688.

[23] Wang, Y., Li, X., Chen, J., Liu, Y. (2022). A condition-based maintenance policy for multi-component systems subject to stochastic and economic dependencies, *Reliability Engineering & System Safety*, Vol. 219, Article No. 108174, doi: 10.1016/j.ress.2021.108174.

[24] Alley, S.L. (2015). *A probabilistic assessment of failure for Air Force building systems,* Master's thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, USA, AFIT-ENV-MS-15-M-196, DTIC Accession No. ADA615418.

[25] Pan, E., Liao, W., Xi, L. (2012). A joint model of production scheduling and predictive maintenance for minimizing job tardiness, *The International Journal of Advanced Manufacturing Technology*, Vol. 60, 1049-1061, doi: 10.1007/s00170-011-3652-4.

[26] Cox, D.R. (1972). Regression models and life-tables, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34, No. 2, 187-202, doi: 10.1111/j.2517-6161.1972.tb00899.x.

[27] Kumar, D., Klefsjö, B. (1994). Proportional hazards model: A review, *Reliability Engineering & System Safety*, Vol. 44, No. 2, 177-188, doi: 10.1016/0951-8320(94)90010-8.

[28] Makis, V., Jardine, A.K.S. (1992). Optimal replacement in the proportional hazards model, *INFOR: Information Systems and Operational Research*, Vol. 30, No. 1, 172-183, doi: 10.1080/03155986.1992.11732183.
[29] Ghasemi, A., Yacout, S., Ouali, M.S. (2007). Optimal condition based maintenance with imperfect information and the proportional hazards model, *International Journal of Production Research*, Vol. 45, No. 4, 989-1012, doi: 10.1080/00207540600596882.
[30] Cansu, T., Bas, E., Egrioglu, E., Akkan, T. (2024). Intuitionistic fuzzy time series forecasting method based on dendrite neuron model and exponential smoothing, *Granular Computing*, Vol. 9, Article No. 49, doi: 10.1007/s41066-024-00474-6.
[31] Hesamian, G., Johannssen, A., Chukhrova, N. (2024). A neural network-based ARMA model for fuzzy time series data, *Computational and Applied Mathematics*, Vol. 43, Article No. 445, doi: 10.1007/s40314-024-02950-w.
[32] Weiss, C.H., Zhu, F., Hoshiyar, A. (2022). Softplus INGARCH models, *Statistica Sinica*, Vol. 32, No. 2, 1099-1120, doi: 10.5705/ss.202020.0353.
[33] Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M. (2015). *Time series analysis: Forecasting and control,* 5th edition, Hoboken, New Jersey, USA.
[34] Ripatti, S., Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood, *Biometrics*, Vol. 56, No. 4, 1016-1022, doi: 10.1111/j.0006-341X.2000.01016.x.
[35] Ducrocq, V., Casella, G. (1996). A Bayesian analysis of mixed survival models, *Genetics Selection Evolution*, Vol. 28, No. 6, 505-529, doi: 10.1186/1297-9686-28-6-505.
[36] Cortiñas Abrahantes, J., Burzykowski, T. (2005). A version of the EM algorithm for proportional hazard model with random effects, *Biometrical Journal*, Vol. 47, No. 6, 847-862, doi: 10.1002/bimj.200410141.
[37] Malik, M.A.K. (1979). Reliable preventive maintenance scheduling, *AIIE Transactions*, Vol. 11, No. 3, 221-228, doi: 10.1080/05695557908974463.