Jurnal Teknik Industri, Vol.28, No. 2, December 2026
ISSN 1411-2485 print / ISSN 2087-7439 online
DOI: 10.9744/jti.28.2.176-196

An Adaptive Predictive–Prescriptive Framework for Intermittent Spare Parts Inventory with Age-Dependent Repair-or-Replace Decisions

Saskia Puspa Kenaka[1,2,3*], Hendra Nunumete[1], Suprayogi[1,2,3]
[1)] Faculty of Industrial Technology, Bandung Institute of Technology, 10 Ganesha Street, Bandung 40132, Indonesia
[2)] Research Group of Industrial System and Techno-Economics, Bandung Institute of Technology, Bandung, Indonesia
[3)] Center for Logistics and Supply Chain Studies, Bandung Institute of Technology, Bandung, Indonesia
Email: saskia@itb.ac.id*, hendra.nunumete@gmail.com, yogi@itb.ac.id
*Corresponding author

Abstract: Critical spare parts with intermittent demand are difficult to manage because failures occur infrequently and irregularly, while stockouts may disrupt operations. Conventional inventory approaches based mainly on historical demand may not adequately represent future requirements because spare part demand is also influenced by component age and maintenance decisions. This study proposes an adaptive predictive–prescriptive framework integrating age-based failure forecasting, repair-or-replace decisions, and a time-varying order-up-to inventory policy. Failure probabilities during the procurement lead time are estimated using a Weibull reliability model and aggregated to obtain gross demand. The predicted failures are then evaluated through an economic repair-or-replace model to generate net demand. Net demand is represented using a state-dependent Markov chain, while the inventory policy is determined using a Greedy Marginal Analysis Heuristic. The framework was evaluated using an Automatic Voltage Regulator as a critical repairable component. The demand was classified as intermittent, with an Average Inter-Demand Interval of 4.0625 and a squared Coefficient of Variation of 0.0986. The repair-or-replace adjustment reduced new spare part requirements by 14.3%, while the resulting policy achieved a 95.1% non-stockout probability and an expected holding cost of 1.4273 per unit per period.

Keywords: Age-based forecasting; intermittent demand; repair-or-replace decision; spare parts inventory; time-varying order-up-to policy; predictive–prescriptive framework

Introduction

Spare parts inventory management directly affects operational continuity and cost efficiency [1], [2]. Insufficient inventory may cause equipment downtime and substantial operational losses [3], as unavailability reduces effective production time and system throughput [4], whereas excessive inventory increases capital and holding costs, particularly for high-value components [5]. This problem becomes more difficult for spare parts with intermittent demand, where demand occurs infrequently and at irregular intervals [6], [7]. Despite their low demand frequency, intermittent spare parts may represent a large proportion of stock-keeping units and inventory value [7]. Therefore, their inventory requirements cannot be managed in the same manner as products with more regular demand.

This problem is relevant to several industries. Giannopoulos et al. [8] discuss intermittent demand forecasting for spare parts in industries such as aviation and automotive. In maritime operations, spare parts supply is also affected by vessel movements, dispersed supply locations, and limited delivery time windows [9]. In automotive manufacturing, Chien et al. [10] study aftermarket spare parts demand forecasting using data from an automotive component manufacturer. These examples show the practical importance of planning spare parts availability when demand is irregular and supply requires time.

For repairable components, spare parts requirements also depend on component age and the maintenance action selected. When failure risk increases with age, a long period without demand does not necessarily indicate a low future requirement. However, not every failure requires a new component because some failed components can be repaired and reused. Repair-or-replace decisions therefore affect both the quantity and timing of replacement demand. Treating all predicted failures as replacement demand may result in excess

Submitted: July 20, 2026; Revised: Oct 16, 2026; Accepted: Nov 3, 2026

inventory and higher holding costs. Conversely, insufficient inventory may prolong downtime when a replacement component must be procured. These conditions motivate the integration of age-based failure forecasting, repair-or-replace decisions, and inventory planning.

The planning problem is to determine how many predicted failures will require new spare parts and how much inventory should be available to meet those requirements. Failure forecasting, maintenance decisions, and inventory planning are connected because not every predicted failure consumes a new spare part. Directly treating all failures as replacement demand may overestimate inventory requirements, while ignoring changes in component condition may produce inventory policies that do not represent the actual operational risk [5], [11]. Therefore, the demand used in the inventory model should first be adjusted according to the maintenance action selected for each predicted failure.

This study proposes an adaptive predictive–prescriptive framework that integrates age-based failure forecasting, repair-or-replace decisions, and a time-varying order-up-to inventory policy. Component failure probabilities during the procurement lead time are estimated using a Weibull reliability model and aggregated as gross demand. The predicted failures are then evaluated using an economic repair-or-replace decision model. Only failures assigned to replacement are converted into net spare part demand. The resulting demand sequence is represented through state-dependent Markov probabilities and used to determine the inventory policy using a Greedy Marginal Analysis Heuristic.

The main contribution of this study is the integration of component age, maintenance decisions, and dynamic inventory control within a single framework for repairable spare parts with intermittent demand. The framework distinguishes potential failure demand from the actual procurement demand generated after maintenance decisions. Therefore, the resulting inventory policy can respond to changes in failure risk and demand intervals while maintaining the required service level at a lower holding cost. The framework is evaluated using an Automatic Voltage Regulator as a critical repairable component in an electricity generation system.

The structure of this paper is as follows. The second section discusses relevant literature related to intermittent demand forecasting, maintenance decisions, and inventory policies. The third section explains the proposed model. The fourth section presents the results and discussion, while the fifth section contains conclusions, study limitations, and future research directions.

Literature Review

This section reviews studies on intermittent demand forecasting, repair-or-replace decisions, and inventory control. The discussion explains how these studies support the proposed framework and identifies the connection addressed in this research.

Intermittent Demand Forecasting

Several forecasting methods have been developed for intermittent demand. The Croston method separately estimates the interval between demands and the size of positive demand [7], [12]. Other studies have applied support vector machines [13], neural networks [14], and ensemble methods [10], [15], [16], [17]. These approaches provide different ways to estimate demand from historical observations. For example, Chien et al. [10] developed an ensemble learning approach for aftermarket spare parts demand forecasting and applied it to an automotive component manufacturer.

Historical demand provides useful information, but it may not fully describe the technical conditions that generate spare parts requirements [5]. A period without demand does not necessarily indicate that the installed components have a low risk of failure. Van der Auweraer and Boute [18] addressed this issue by using service maintenance information to forecast spare parts demand. Their approach links demand estimation to information from maintenance operations, including component age and failure characteristics.

This approach provides a basis for estimating failure-related spare parts requirements. However, for repairable components, the predicted number of failures may differ from the number of new spare parts required. Some failed components can be repaired and reused. Therefore, the maintenance action should be considered when converting predicted failures into replacement demand.

Repair-or-Replace Decision

When a repairable component fails, the maintenance decision determines whether it should be repaired or replaced with a new component [19]. Previous studies have considered both perfect [20], [21] and imperfect repairs [22], [23]. Perfect repair restores the component to a condition equivalent to new, while imperfect repair provides only partial restoration. The repair outcome therefore affects the component's future operation and subsequent maintenance requirements.

Safaei et al. [19] developed a repair-and-replacement policy based on profit per unit time. Their model considers perfect and imperfect repairs, together with periodic preventive maintenance. The economic evaluation compares the benefits and costs associated with the maintenance alternatives.

This study uses that economic evaluation approach under a corrective maintenance setting. The purpose is to determine which predicted failures require new spare parts. A replacement action generates demand for a new component, while a repair action allows the existing component to be reused. The resulting decisions therefore affect both the quantity and timing of demand passed to the inventory model.

Inventory Control for Intermittent Demand

Several studies have examined inventory policies for intermittent demand [24], [25], [26], [27]. An important consideration is whether the probability of demand remains constant or changes with the time since the most recent demand. This distinction affects the amount of stock required in each period.

Prak and Rogetzer [28] developed an inventory model that uses the distributions of demand sizes and intervals to determine time-varying order-up-to levels. The model allows the probability of demand occurrence to change with the time elapsed since the most recent demand. It evaluates holding costs and service performance and uses a Greedy Marginal Analysis Heuristic to determine inventory targets. Their results show that time-varying targets can achieve service requirements at lower costs than fixed order-up-to levels.

This study adopts that inventory approach and applies it to the net demand obtained after failure forecasting and repair-or-replace decisions. Component age is used at the forecasting stage, while the time since the most recent positive demand is used to define the inventory demand state. These variables have different roles within the framework.

Research Gap and Position of This Study

The reviewed studies address different aspects of spare parts planning. Forecasting studies estimate future requirements, maintenance studies evaluate repair-or-replace alternatives, and inventory studies determine the stock needed to meet demand. Table 1 compares selected studies based on their forecasting approach, treatment of component age, maintenance decisions, and inventory policy.

The comparison shows that these studies provide different parts of the planning process. Service maintenance information supports failure-related demand forecasting. Economic maintenance evaluation supports the choice between repair and replacement. Time-varying inventory policies adjust stock targets according to demand occurrence patterns. The issue addressed in this study is how to connect these steps when a predicted failure does not always require a new spare part.

For repairable components, the maintenance action affects the demand passed to the inventory model. Replacement consumes a new spare part, while repair allows the existing component to be reused. Repair also affects the component's effective life and may change the timing of subsequent failures. Therefore, inventory planning should consider both the predicted failures and the maintenance actions selected.

This study connects these steps through a sequential framework. Component age is first used to estimate failure probability and gross demand. Repair-or-replace decisions then convert the predicted failures into net replacement demand. The resulting demand sequence is used to estimate Markov demand-state probabilities and determine inventory targets. The contribution lies in this connection and the evaluation of how repair-or-

Table 1. Comparison of selected studies and the present study

Study	Demand forecasting approach	Treatment of component age	Repair-or-replace decision	Inventory policy
Chien et al. [10]	Ensemble learning for aftermarket spare parts demand	Focuses on demand data rather than an age-based failure model	Not the main decision addressed	Main focus is demand forecasting
Van der Auweraer and Boute [18]	Forecasting using service maintenance information	Component age and maintenance information support demand estimation	Economic selection between repair and replacement is not the main focus	Evaluates inventory implications of maintenance-based forecasts
Safaei et al. [19]	Main focus is maintenance evaluation rather than spare parts demand forecasting	System lifetime and repair effects are considered	Compares repair and replacement using profit per unit time	State-dependent spare parts inventory targets are not the main focus
Prak and Rogetzer [28]	Demand-size and demand-interval distributions	State represents time since the most recent demand rather than component age	Repair-or-replace selection is not the main focus	Time-varying order-up-to policy
This Study	Age-based failure forecasting followed by net demand estimation	Component age is used to estimate failure probability	Economic evaluation determines which predicted failures require replacement	State-dependent order-up-to policy based on net demand

replace decisions affect spare parts requirements and inventory performance. The framework does not jointly optimize maintenance actions and inventory targets.

Methods

System Description

This research examines spare parts inventory management in a power generation system operating under a corrective maintenance scheme, where maintenance actions are performed after a component failure occurs. The system consists of several generating units, each utilizing several critical components to maintain operational stability and continuity. Failure of any one critical component can cause significant operational disruptions if not promptly addressed and supported by adequate spare parts availability.

Spare parts inventory management in this system faces several key challenges, namely the uncertainty of component failure times, intermittent demand frequency, and relatively long procurement lead times. Under these conditions, inventory approaches that rely solely on historical demand data tend to be ineffective, as past demand patterns do not directly reflect the risk of future component failure.

To address these challenges, this research proposes an integrated modeling framework that explicitly links the system's technical condition to inventory decisions. This framework consists of three main components: (i) age-based component failure forecasting; (ii) repair-or-replace decisions; and (iii) a dynamic inventory model for intermittent demand. Figure 1 presents the overall framework and shows how these components are connected.

Unlike conventional approaches that directly link demand forecasting results to inventory decisions, this study positions maintenance decisions as a crucial intermediate step. Component failure forecasting results are first used to determine the type of corrective action to be taken. Only failures decided upon for replacement result in spare part consumption and contribute to net demand, while failures addressed through repair have no direct impact on inventory levels.

Thus, the proposed model framework ensures that inventory decisions are based not only on statistical demand estimates but also consider the technical mechanisms of maintenance and the operational risks of the system. This approach allows for a more realistic integration of failure forecasting, maintenance decisions, and inventory control, particularly for parts with intermittent demand characteristics and high criticality.

Demand Characteristics and Modeling Assumption

The characteristics of spare part demand must first be analyzed to ensure that the modeling approach used is appropriate for the demand patterns encountered. In the generation system studied, spare part demand arises

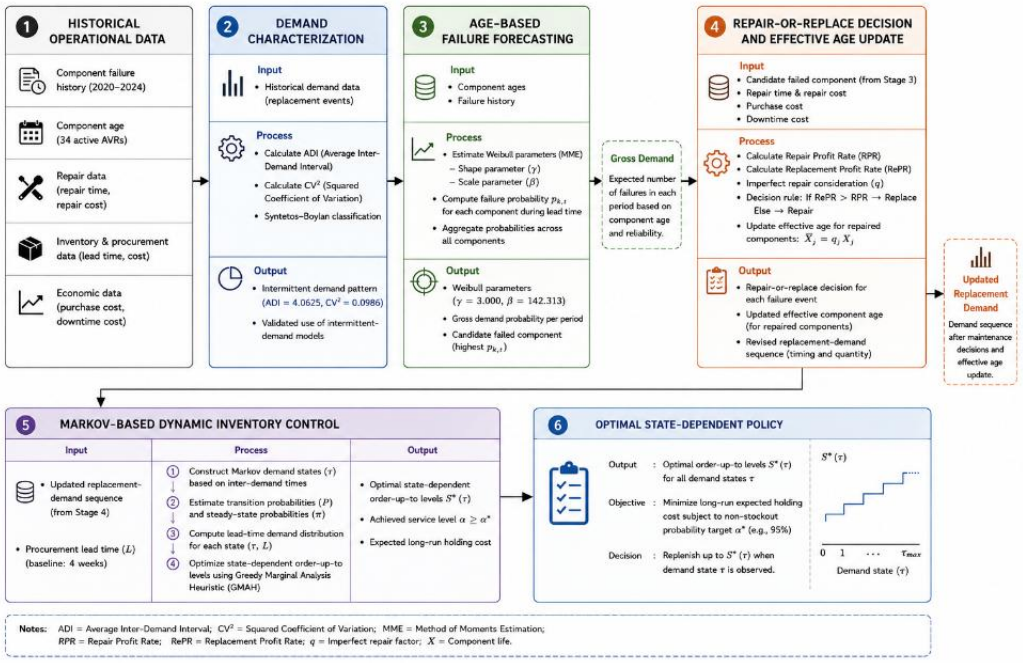


Figure 1. Framework of the proposed model

as a consequence of component failure events, making it irregular and infrequent. This situation requires a specialized approach that differs from conventional time series forecasting.

Demand Characteristics

The characteristics of spare parts demand were analyzed to assess the appropriateness of an intermittent demand modeling approach. Demand classification was conducted using two widely adopted metrics in the literature, namely the Average Inter-Demand Interval (ADI) and the squared Coefficient of Variation (CV^2). ADI represents the average time interval between successive demand occurrences, while CV^2 quantifies the variability of demand size conditional on the occurrence of demand [7], [29].

A combination of high ADI values and low CV^2 values indicates that demand events occur infrequently, while the demand magnitude remains relatively stable when they occur. Such demand behavior is consistent with the definition of intermittent demand and suggests that conventional time-series forecasting methods, which rely on regular demand patterns, are not well suited. Consequently, a probabilistic, event-based modeling approach that explicitly captures both the occurrence probability of demand and the inter-demand intervals is more appropriate.

Modeling Assumption

The model development in this study is based on the following assumptions:

1. The analysis focuses on a single type of critical spare part.
2. Spare part demand is estimated from predicted component failures based on component age. Each predicted failure is evaluated to determine whether repair or replacement is selected. Only predicted failures assigned to replacement generate demand for new spare parts. Additional demand from scheduled preventive replacement is not included.
3. The spare part procurement lead time, L, is assumed to be constant over the analysis horizon.
4. Each failure event decided upon for replacement consumes one new spare part unit.
5. Failure events between units are assumed to be independent for the modeling approach.
6. Inventory decisions are evaluated periodically at discrete time intervals.

The framework predicts future failures and uses these predictions to plan maintenance actions and spare parts requirements in advance. The predicted events represent failures that may occur during the planning horizon, rather than failures that have already been observed. Repair-or-replace decisions are evaluated for these projected events to estimate net replacement demand. This advance planning is consistent with a corrective

maintenance setting because the maintenance alternatives address the predicted failures. The model does not determine a separate preventive replacement time before failure. These assumptions allow for consistent integration between age-based failure modeling, maintenance decisions, and dynamic inventory policies, and align with common practice in reliability-based spare part inventory modeling.

Age-Based Demand Modeling and Demand Estimation

The demand for spare parts in this study is based on the risk of component failure in operating generating units. Given the intermittent nature of demand, forecasting is not based on the time series of spare part consumption, but rather through a reliability-based probabilistic approach that links component age to the probability of failure during the lead time period. This approach aligns with the reliability-based spare parts forecasting and Service Maintenance Information (SMI) framework [18].

Age-Based Component Failure Modeling

Component failures are modeled using the Weibull distribution, which is widely used in reliability analysis due to its flexibility in representing various failure rate patterns, including increasing failure rates due to component aging [30]. The Weibull reliability function is expressed as follows, where x is the component age, β is the scale parameter, and γ is the shape parameter.

$$R(x) = P(T > x) = \exp\left[-\left(\frac{x}{\beta}\right)^{\gamma}\right] \quad (1)$$

A value of $\gamma > 1$ indicates that the failure rate increases with age, consistent with the degradation characteristics of critical components in a generation system. The Weibull parameters are estimated using the Method of Moments Estimation (MME). This method provides a direct link between the first two sample moments of the lifetime data and the theoretical moments of the Weibull distribution. It was selected for its straightforward implementation and transparent calculation of the shape and scale parameters. The purpose is to parameterize the reliability model within the proposed framework, rather than to compare alternative estimation methods.

In a system consisting of multiple generating units, each component may have a different operational age at the beginning of the decision period. To simplify the system-level demand aggregation process, components are grouped by operational age. Let $n_{ij}(t)$ denote the number of components with age j installed on the i-th unit in decision period t.

Probability of Failure During Lead Time

In a spare parts inventory system, ordering decisions must consider the possibility of failure during the lead time of replacement components. Therefore, the required quantity is not the probability that a component has failed to date, but rather the probability that it will fail during the interval $(t, t + L)$, where L represents the lead time. Based on the definition of the reliability function

$$R(x) = P(T > x) \quad (2)$$

Therefore, the probability that a component will fail during the age interval $j < T < j + L$ is

$$P\ (j < T < j + L) = P(T > j) - P(T > j + L) = R(j) - R(j + L) \quad (3)$$

Equation (3) shows that the probability of failure during the lead time is obtained by the difference between the probability of the component surviving at the beginning of the interval and the probability of the component surviving at the end of the interval. Next, by substituting the Weibull reliability function, the probability of component failure during the lead time becomes:

$$\hat{p}_{i,j,t,L} = R(j) - R(j + L) \quad (4)$$

$$\hat{p}_{i,j,t,L} = \exp\left[-\left(\frac{j}{\beta}\right)^{\gamma}\right] - \exp\left[-\left(\frac{j+L}{\beta}\right)^{\gamma}\right] \quad (5)$$

where $\hat{p}_{i,j,t,L}$ is the probability that component age j in the i-th unit will fail before replacement parts are available. This formulation is identical to the approach in the Service Maintenance Information (SMI) model, where failure probabilities are calculated individually for each component age group before being aggregated into system-level demand [18].

Gross Demand Estimation

After the failure probability of each age group is obtained, the next step is to calculate the expected value of the number of failures during the lead time interval. If $n_{ij}(t)$ represents the number of components with age j in the i-th unit in period t and each component has a failure probability of $\hat{p}_{i,j,t,L}$, then the expected contribution to demand from each age group is $n_{ij}(t) \cdot \hat{p}_{i,j,t,L}$. By summing all age groups in all generating units, an estimate of gross demand during the lead time is obtained as follows, where N is the number of operating generators and J is the maximum observed age group.

$$\hat{D}_{t,L}^{gross} = \sum_{i=1}^{N} \sum_{j=1}^{J} n_{ij}(t) \cdot \hat{p}_{i,j,t,L} \quad (6)$$

It should be noted that this equation represents the expected number of component failures, not probabilities. Therefore, the value of $\hat{D}_{t,L}^{gross}$ is not limited to the interval [0,1] and can exceed one if there are many components with a high risk of failure during the lead time.

The value of $\hat{D}_{t,L}^{gross}$ still represents all predicted failure events during the lead time, regardless of the maintenance actions to be taken. In practice, some failed components can be repaired, while others must be replaced. Therefore, the gross demand estimate in this equation is then processed through the repair–replacement decision model in Section 3.4 to obtain the net spare part demand, which forms the basis of the inventory control model.

Corrective Maintenance Decision Model (Repair-Replacement)

A corrective maintenance decision model is used to determine whether each predicted component failure should be addressed through repair or replacement. This formulation is adapted from the probability-based and average-profit repair-replacement policy developed by Safaei et al. [19]. In the original model, preventive and corrective maintenance are considered simultaneously. In this study, the preventive maintenance component is not used because the system under study implements a corrective maintenance policy. Therefore, the evaluation focuses on the economic consequences of repair and replacement actions after a failure occurs.

This decision model serves as a bridge between age-based failure forecasting and inventory policies. The goal is not to optimize the maintenance policy as the final output, but rather to determine which failures result in the need for new spare parts.

Definition of Repair-Replacement Decisions

The two alternatives considered in this study follow the repair-and-replacement framework of [19]. Since the AVR is a repairable component, repair is evaluated as an alternative to replacement for each predicted failure. These alternatives are represented as a binary decision because they have different consequences for new spare parts demand. The model considers the following actions:

1. Repair, where the failed component is repaired and reused without consuming a new spare part.
2. Replacement, where the failed component is replaced with a new component, generating demand for one spare part.

The binary representation is used to distinguish failures that require a new spare part from those that can be addressed through repair. It does not represent all possible maintenance activities. Within this model, repair is assumed not to consume the spare part being studied, while replacement consumes one unit. The economic comparison determines the selected action, and the resulting decisions are used to calculate net spare parts demand.

The maintenance decision variable for the j-th failure is defined as

$$z_j = \begin{cases} 1, & \text{if replacement is selected,} \\ 0, & \text{if repair is selected.} \end{cases} \quad (7)$$

The value of z_j is determined by comparing the economic performance of the two alternative actions.

The conversion of the failure forecast into predicted events and the selection of maintenance actions are two separate steps. First, the age-based forecasting model estimates failure probabilities. In the numerical implementation, an operational threshold is used to generate discrete predicted failure events. These events represent failures anticipated during the planning horizon, rather than failures that have already been observed.

Second, each predicted failure is evaluated by comparing the profit per unit time of repair and replacement using Equations (8) and (9). The binary variable z_j represents the selected maintenance action, not the occurrence of failure. A value of $z_j = 0$ indicates repair, while $z_j = 1$ indicates replacement. The action with the higher profit per unit time is selected according to Equation (10). Therefore, the failure forecast identifies the events to be evaluated, while the economic comparison determines whether each event generates demand for a new spare part.

Profit per Unit Time for Repair Activities

Following the structure of Safaei et al. [19], the performance of repair alternatives is evaluated based on the profit earned per unit of cycle time. After removing the preventive maintenance cost component, which is not used in this study, the profit per unit of time for a repair action is expressed as:

$$TP_j^{repair} = \frac{e_0 \tilde{X}_j - (e_1 + k) Ta_j}{\tilde{X}_j + Ta_j} \tag{8}$$

where TP_j^{repair} denotes the profit per unit time when the j-th failure is addressed through repair, $\tilde{X}_j$ the effective operating life in the j-th maintenance cycle, Ta_j the repair time for the j-th failure, e_0 the system operating revenue per unit time, e_1 the repair cost per unit time, and k the downtime penalty cost per unit time.

The numerator of the equation represents the operating revenue over the component's effective life after deducting repair costs and losses due to downtime. The denominator represents the total cycle duration, i.e., the effective operating life plus the repair time. Thus, TP_j^{repair} is a measure of profitability per unit time, not repair time. The longer the repair time Ta_j, the greater the repair cost and downtime penalty. Therefore, with other parameters held constant, increasing the repair time decreases the attractiveness of the repair alternative.

Profit per Unit Time for Replacement Activities

The replacement alternative is evaluated under a procurement-based reference assumption, where a new component is obtained with procurement lead time L. This reference condition is used to compare replacement with repair before determining the inventory policy. The profit per unit time for replacement is expressed as follows:

$$TP_j^{replace} = \frac{e_0 \tilde{X}_j - (kL + c)}{\tilde{X}_j + L} \tag{9}$$

where $TP_j^{replace}$ denotes the profit per unit time when the j-th failure is addressed through replacement, L the procurement lead time, and c the purchase cost; $\tilde{X}_j$, e_0, and k are as previously defined. The numerator of Equation (9) represents operating income after deducting the purchase cost and downtime penalty. The denominator represents the effective life of the component plus the replacement lead time.

Under this reference assumption, a longer procurement lead time increases the downtime penalty and reduces the profit per unit time of replacement. If a new component is available from stock, the procurement waiting time does not apply to that replacement. Therefore, the term kL in Equation (9) represents the downtime penalty under the reference condition, rather than the actual waiting cost of every replacement under the resulting inventory policy.

Holding cost is considered separately in Section 3.5.5, where inventory levels are determined by minimizing expected holding cost subject to the required non-stockout probability. The framework follows a sequential approach. The maintenance decision first determines net spare part demand, which is then used to determine the inventory policy. The repair-or-replace decision is not recalculated according to the stock available at each predicted failure.

Repair-Replacement Decision Rule

Since TP_j^{repair} and $TP_j^{replace}$ represent profit per unit time, the action with the higher value is chosen. The decision rule is stated as

$$z_j = \begin{cases} 0, & TP_j^{repair} \geq TP_j^{replace}, \\ 1, & TP_j^{replace} > TP_j^{repair}. \end{cases} \quad (10)$$

This rule is consistent with the objective of Safaei et al. [19], which is to choose the action that provides the higher profit per unit time. If the values of both alternatives are equal, repair can be chosen as the decision rule because it does not immediately consume new parts. However, other decision rules can be used if the criticality of the component requires immediate replacement.

The economic comparison is conditional on a predicted failure requiring a maintenance action. It does not mean that the failure probability becomes either zero or one. The binary value records the selected action for planning purposes. Only events assigned to replacement contribute to the discrete net demand in Equation (17).

Imperfect Repair and Effective Life Renewal

When a repair is selected, the component does not always return to its good-as-new condition. Therefore, the model incorporates the concept of imperfect repair, which is a condition where the repair only partially restores the component's capability, resulting in a shorter effective life after the repair than before the repair.

Suppose X is the component's initial life. The component's effective life at the j-th cycle after the imperfect repair is expressed as:

$$\widetilde{X_j} = q_j \widetilde{X_{j-1}} \quad (11)$$

with the initial condition:

$$\tilde{X}_0 = X \quad (12)$$

The degree of repair quality can be expressed as

$$q_j = \frac{\tilde{X}_j}{\tilde{X}_{j-1}} \quad (13)$$

The parameter q_j indicates the proportion of the remaining effective life after repair:
$q_j = 1$ indicates a perfect repair in the Safaei et al. [19] representation, because the effective life after repair remains the same as the previous cycle;
$0 < q_j < 1$ indicates an imperfect repair, meaning the effective life after repair is shorter.

This interpretation needs to be distinguished from the virtual age reduction model, where a smaller coefficient value indicates a better repair result. In this formulation, $q_j$multiplies the effective life, so the smaller q_j, the greater the reduction in component life after repair.

If replacement is selected, the failed component is replaced with a new one. Therefore, the new component age is reset to zero in the next forecasting period:

$$a_j^{new} = 0 \quad (14)$$

Conversely, if repair is selected, the updated effective age is carried over to the next forecasting cycle and affects the Weibull failure probability in the following period.

Formation of Net Spare Part Demand

The maintenance decisions are applied to the failure events predicted by the age-based forecasting model. Let F_t denote the set of predicted failure events in period t, and let $z_{r,t}$ denote the decision for event r. The value of $z_{r,t}$ is 1 when replacement is selected and 0 when repair is selected. Each predicted event is assigned to one of these two actions.

Therefore, $1 - z_{r,t}$ equals 1 for a repair action and 0 for a replacement action. Summing this value over all predicted events in F_t gives the number of repair actions in period t:

$$N_t^{\text{repair}} = \sum_{r \in F_t} (1 - z_{r,t}) \quad (15)$$

Similarly, $z_{r,t}$ equals 1 only when replacement is selected. Summing $z_{r,t}$over the same set of predicted events gives the number of replacement actions:

$$N_t^{\text{replacement}} = \sum_{r \in F_t} z_{r,t} \tag{16}$$

Each replacement action requires one new spare part, while repair does not consume the spare part being studied. Therefore, net spare part demand equals the number of replacement actions:

$$D_t^{\text{net}} = N_t^{\text{replacement}} = \sum_{r \in F_t} z_{r,t} \tag{17}$$

Net demand can also be expressed in expected-value form without first converting the failure forecast into discrete events. For each component group, the expected number of failures during the lead time is obtained by multiplying the number of installed components by their failure probability. This is the contribution of that group to gross demand in Equation (6).

To obtain expected net demand, this contribution is multiplied by the replacement decision. If replacement is selected, the expected failures contribute to new spare part demand. If repair is selected, the contribution is zero. Summing these contributions over all component groups gives:

$$\widehat{D}_{t,L}^{\text{net}} = \sum_{i=1}^{N} \sum_{j=1}^{J} n_{ij}(t)\, \hat{p}_{i,j,t,L}\, z_{ij,t} \tag{18}$$

where $n_{ij}(t)$is the number of components of age j installed in generating unit i during period t, $p_{i,j,t,L}$ is the estimated probability of component failure during the lead time, and $z_{ij,t}$ equals 1 if replacement is selected and 0 if repair is selected. N is the number of generating units, and J is the maximum age group considered.

Equation (18) follows from the gross-demand expression and the assumption that each replacement requires one new spare part. The replacement decision is treated as given for each component group at the time of evaluation. Components within the same group are assumed to share the failure probability and replacement decision used in the calculation. The expected contributions can then be added using the linearity of expectation. The result represents an expected quantity and does not need to be an integer.

Equations (17) and (18) describe two representations of net demand. Equation (17) counts replacement requirements after predicted failures have been represented as discrete events. Equation (18) retains the failure forecast in expected-value form and weights each group's expected failures by its replacement decision. The two expressions are alternative representations and are not added together.

In the numerical implementation, the operational threshold is used to construct a discrete sequence of predicted failure events. The repair-or-replace evaluation then assigns a binary maintenance decision to each event. Equation (17) is used to calculate the resulting predicted net demand, which becomes the input to the inventory model. These quantities represent projected replacement requirements rather than observed demand.

Markov-Based Dynamic Inventory Control

The net spare part demand obtained after the repair-or-replace decision is used as the input for the inventory control stage. Although the maintenance decisions change the number and timing of replacement requirements, the resulting demand sequence remains intermittent, with many zero-demand periods and irregular intervals between positive demands. Therefore, an inventory policy with a single fixed target may not adequately represent changes in demand risk over time.

This study adopts the time-varying order-up-to policy developed by Prak and Rogetzer [28]. The policy assigns a different inventory target to each demand state, where the state represents the number of periods elapsed since the most recent positive demand. The demand states are represented using a discrete-time Markov chain because the state in the following period depends only on the current state and whether positive demand occurs in the current period.

Unlike the original time-varying order-up-to model, which uses historical intermittent demand directly, the proposed framework uses the revised demand sequence obtained from age-based failure forecasting and repair-or-replace decisions. Therefore, the Markov states and the resulting inventory targets reflect not only historical demand behavior but also the technical condition of the installed components and the selected maintenance actions.

Demand Model and Markov-State Representation

Demand in each period is represented by two components: the probability that positive demand occurs and the demand size conditional on its occurrence. Let p_τ denote the probability of positive demand when the system is in state τ, and let D^+denote the positive demand size.

The positive demand size is treated as a discrete random variable with probability mass function $f(d)$and cumulative distribution function $F(d)$. The probability mass function of demand in state $\tau$is therefore expressed as:

$$P(D_\tau = d) = \begin{cases} 1 - p_\tau, & d = 0, \\ p_\tau f(d), & d > 0. \end{cases} \tag{19}$$

The state $\tau_t$represents the number of periods elapsed since the most recent positive replacement demand. Its development is determined by the revised demand sequence obtained after the repair-or-replace decisions. The transition of the demand state is defined as:

$$\tau_{t+1} = \begin{cases} 1, & D_t^{net} > 0, \\ \min\{\tau_t + 1, \tau'\}, & D_t^{net} = 0, \end{cases} \tag{20}$$

where $\tau'$is the maximum demand state.

When positive demand occurs, the state returns to 1. When no demand occurs, the state increases by one until reaching the maximum state. For example, the demand sequence

$$1, 0, 0, 0, 1, 0$$

produces the state sequence

$$1, 1, 2, 3, 4, 1$$

where the state is recorded at the beginning of each period, before the demand in that period is observed. This convention is applied consistently in the numerical implementation and the reported results. The Markov property applies because the state in period $t + 1$ can be determined from the state in period $t$and the occurrence of demand in period t, without requiring the complete demand history. For every state τ, an order-up-to level $S_\tau$is determined. The inventory decision vector is written as:

$$\mathrm{S} = (S_1, S_2, \dots, S_{\tau'}). \tag{21}$$

When the system is in state τ_t, the replenishment quantity is:

$$Q_t = \max\{0, S_{\tau_t} - IP_t\} \tag{22}$$

where IP_t is the inventory position before replenishment.

Transition and Steady-State Probabilities

The transition probability describes the movement between demand states. From state $\tau < \tau'$, two transitions are possible:
the state returns to 1 with probability p_τ when positive demand occurs;
the state moves to $\tau + 1$ with probability $1 - p_\tau$ when no demand occurs.

The transition probability is therefore expressed as:

$$P_{\tau r} = \begin{cases} p_\tau, & r = 1, \\ 1 - p_\tau, & r = \tau + 1, \tau < \tau', \\ 1 - p_{\tau'}, & r = \tau = \tau', \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

At the maximum state τ', the chain either returns to state 1 when demand occurs or remains at τ' when no demand occurs. If the model assumes $p_{\tau'} = 1$, the self-transition at the maximum state disappears and positive demand is forced to occur at that state. This study adopts this assumption, which ensures that the Markov chain is irreducible.

Let

$$\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_{\tau'}) \tag{24}$$

denote the steady-state probability vector. The steady-state probabilities satisfy:

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \tag{25}$$

and

$$\sum_{\tau=1}^{\tau'} \pi_\tau = 1 \tag{26}$$

For the sequential state structure used in this study, the steady-state probability can also be written directly as:

$$p_\tau^S = \frac{\prod_{r=1}^{\tau-1}(1-p_r)}{\sum_{\ell=1}^{\tau'}\prod_{r=1}^{\ell-1}(1-p_r)}, \tau = 1, \dots, \tau' \tag{27}$$

A state with a higher steady-state probability is visited more frequently over the long run and therefore has a larger contribution to the expected service level and holding cost. However, states with lower steady-state probabilities may still require a higher inventory target when their demand occurrence probabilities are relatively high.

The probabilities p_τ are estimated from the revised demand sequence obtained after the maintenance decisions. For each state, p_τ is calculated as the proportion of observations with positive demand among all observations associated with that state.

Demand Distribution During Lead Time

Inventory requirements depend on the total demand that may occur while a replenishment order is still in transit. Since the probability of demand changes with the state, the demand distribution during the lead time is non-stationary. Therefore, the model must consider both the number of positive demand occurrences during the lead time and the states passed through during that interval.
Let L denote the deterministic procurement lead time, n the number of positive demand occurrences during the lead time, τ_p the state before the lead-time interval, $\tau$the state at the end of the interval, and $f_{L,\tau}^N(n, \tau_p)$ the joint probability that n positive demands occur during the lead time and that the initial state is τ_p, given the final state τ.

The probability $f_{L,\tau}^N(n, \tau_p)$ is obtained by evaluating all feasible binary demand trajectories of length L. A value of one in a trajectory represents a positive demand occurrence, while a value of zero represents a period without demand. For every feasible trajectory, its probability is obtained by multiplying the state-dependent probabilities of demand and no demand along the trajectory.

For a trajectory $b = (b_1, \dots, b_L)$, its probability is calculated as:

$$P(b \mid \tau_p) = \prod_{\ell=1}^{L} \left[p_{\tau_\ell}\right]^{b_\ell} \left[1-p_{\tau_\ell}\right]^{1-b_\ell} \tag{28}$$

where the state is reset to 1after $b_\ell = 1$, and increases by one after $b_\ell = 0$.

The joint probability $f_{L,\tau}^N(n, \tau_p)$ is obtained by summing the probabilities of all feasible trajectories containing exactly n positive demand occurrences and connecting state τ_p to state τ. The number of positive demand occurrences is then converted into total demand quantity. Since each positive period demand is assumed to follow the same probability mass function $f(d)$, the distribution of the sum of $n$positive demands is obtained through the n-fold convolution of f, denoted by $f^{*n}(d)$. Therefore, the joint probability mass function of total lead-time demand is:

$$f_{L,\tau}^D(d, \tau_p) = \sum_{n=0}^{L} f_{L,\tau}^N(n, \tau_p) f^{*n}(d) \tag{29}$$

For $n = 0$, the convolution is defined as:

$$f^{*0}(d) = \begin{cases} 1, & d = 0, \\ 0, & d > 0. \end{cases} \tag{30}$$

Equation (29) provides the probability that total demand during the lead time equals d, while also considering the state before the lead time. This distribution is used to evaluate the service level and expected inventory remaining after replenishment.

Non-Stockout Probability

The service level in this study is measured using the non-stockout probability of Prak and Rogetzer [28]. It represents the long-run probability of non-negative net inventory at the end of a period.

Let τ denote the current state and τ_p the state before the lead time. For lead-time demand d, the remaining net inventory is $S_{\tau_p} - d$. When this quantity is non-negative, current demand can be satisfied if no demand occurs, with probability $1 - p_\tau$, or if positive demand does not exceed the remaining inventory, with probability $p_\tau F(S_{\tau_p} - d)$. Applying the law of total probability gives:

$$\alpha_\tau(\mathrm{S}) = \sum_{\tau_p=1}^{\tau'} \sum_{d=0}^{S_{\tau_p}} \left[(1 - p_\tau) + p_\tau F(S_{\tau_p} - d)\right] f_{L,\tau}^{D}(d, \tau_p). \tag{31}$$

Here, $f_{L,\tau}^{D}(d, \tau_p)$ is the joint probability defined in Equation (29). Cases with $d > S_{\tau_p}$ contribute zero because net inventory is already negative. Averaging over the steady-state probabilities gives:

$$\alpha(\mathrm{S}) = \sum_{\tau=1}^{\tau'} p_\tau^s \ \alpha_\tau(\mathrm{S}) \tag{32}$$

The inventory policy must satisfy:

$$\alpha(\mathrm{S}) \geq \alpha^* \tag{33}$$

where α^* denotes the minimum non-stockout probability determined by the decision maker.

In this study, the formulation evaluates inventory availability for replacement demand obtained after the repair-or-replace decisions. Thus, the service constraint is applied to net spare parts requirements rather than all predicted failures. The service-level target is treated as a hard constraint because the inventory policy is required to provide a specified minimum level of spare parts availability. Holding cost is minimized among policies that satisfy this requirement. A stockout-cost objective alone would allow the achieved service level to depend on the assigned shortage penalty and would not necessarily satisfy the specified target. The constraint does not eliminate stockout risk but limits it through the required non-stockout probability.

Expected Holding Cost and Optimization Model

The expected holding cost represents the long-run cost of inventory remaining after demand has been satisfied. Let h denote the holding cost per unit per period. For a given order-up-to policy S, the expected long-run holding cost is expressed as:

$$H(\mathrm{S}) = h \sum_{\tau=1}^{\tau'} p_\tau^s \sum_{\tau_p=1}^{\tau'} \sum_{d=0}^{S_{\tau_p}} (S_{\tau_p} - d) f_{L,\tau}^{D}(d, \tau_p) \tag{34}$$

The term $S_{\tau_p} - d$ represents the remaining inventory when total lead-time demand is equal to d. Only nonnegative remaining inventory is included in the holding-cost calculation.

The inventory optimization problem is formulated as:

$$\min_{\mathrm{S}} H(\mathrm{S}) \tag{35}$$

subject to:

$$\alpha(\mathrm{S}) \geq \alpha^* \tag{36}$$

$$S_\tau \in \mathbb{Z}_{\geq 0}, \tau = 1, \ldots, \tau' \tag{37}$$

The objective is to determine the minimum-cost vector of state-dependent order-up-to levels while maintaining the required service level. A higher inventory level generally increases the non-stockout probability but also increases the holding cost. Therefore, the model searches for a balance between spare part availability and inventory investment.

Greedy Marginal Analysis Heuristic

The optimization problem contains one inventory decision variable for every demand state. When the number of states is large, evaluating every possible combination of $S_1, \dots, S_{\tau'}$ requires a large computational effort. Therefore, this study uses the Greedy Marginal Analysis Heuristic proposed by Prak and Rogetzer [28].

The heuristic consists of two main stages. The first stage is a forward allocation procedure. Starting from zero inventory in all states, the algorithm evaluates the effect of adding inventory to each state. The additional inventory is assigned to the state that provides the largest increase in service level relative to the increase in expected holding cost.

For each state τ, let e_τ denote the unit vector corresponding to that state. If the current inventory vector is S, the marginal improvement ratio is:

$$\mathrm{Incr}(\tau) = \frac{\alpha(\mathrm{S}_\tau^{new}) - \alpha(\mathrm{S})}{H(\mathrm{S}_\tau^{new}) - H(\mathrm{S})} \tag{38}$$

where S_τ^{new} is the candidate policy after increasing inventory in state τ.

For a state with zero inventory, the first allocation may need to cover the minimum feasible lead-time demand. This minimum value is expressed as:

$$D_{\tau,L}^{min} = \left\lfloor \frac{L+\tau+\tau'-1}{\tau'} \right\rfloor + 1 \tag{39}$$

Therefore, the candidate policy is defined as:

$$\mathrm{S}_\tau^{new} = \begin{cases} \mathrm{S} + D_{\tau,L}^{min}\mathrm{e}_\tau, & S_\tau = 0, \\ \mathrm{S} + \mathrm{e}_\tau, & S_\tau > 0. \end{cases} \tag{40}$$

The state selected in each forward iteration is:

$$\tau^* = \arg\max_\tau \mathrm{Incr}(\tau) \tag{41}$$

The selected inventory addition is applied, and the process continues until:

$$\alpha(\mathrm{S}) \geq \alpha^* \tag{42}$$

After the service-level target is achieved, a backward reduction procedure is performed. This stage evaluates whether one inventory unit can be removed from any state while providing the largest holding-cost reduction relative to the resulting service-level reduction.

The marginal reduction ratio is expressed as:

$$\mathrm{Decr}(\tau) = \frac{H(\mathrm{S}) - H(\mathrm{S}-\mathrm{e}_\tau)}{\alpha(\mathrm{S}) - \alpha(\mathrm{S}-\mathrm{e}_\tau)} \tag{43}$$

The state selected for inventory reduction is:

$$\tau^* = \arg\max_\tau \mathrm{Decr}(\tau) \tag{44}$$

Inventory is reduced until a further reduction causes the service level to fall below the target. The most recent feasible solution is then retained as the final inventory policy.

The complete procedure can be summarized as follows.

Require: service level target α*, achieved service function
: for all τ do $D_{\tau,L}^{min} = \left\lfloor \frac{L+\tau+\tau'-1}{\tau'} \right\rfloor + 1$
: end for
: Set $S \equiv (S_1, \dots, S_{\tau'}) = (0, 0, \dots, 0)$
: while $\alpha(S) < \alpha *$ do
: for all τ do
: if $S_\tau = 0$ $then$ $S^{new} = S + e_\tau D_{\tau,L}^{min}$
: else $S^{new} = S + e_\tau$
: end if
: Incr $(\tau) = (\alpha(S^{new}) - \alpha(S))/(H(S^{new}) - H(S))$
: end for
: $\tau *=$ argmax $(Incr)$
: if $S_{\tau*} = 0$ then $S = S + e_{\tau*} D_{\tau,L}^{min}$

13 : else $S = S + e_{\tau*}$
14 : end if
15 : end while
16 : while $\alpha(S) > \alpha *$ do
17 : for all τ do
18 : $S^{new} = S - e_{\tau}$
19 : Decr $(\tau) = (H(S) - H(S^{new}))/(\alpha(S) - \alpha(S^{new}))$
20 : end for
21 : $\tau *=$ argmax $(Decr)$
22 : $S = S - e_{\tau*}$
23 : end while
24 : $S = S + e_{\tau*}$
25 : return S

The output of the heuristic is the state-dependent order-up-to policy:

$$S^* = (S_1^*, S_2^*, \dots, S_{\tau'}^*) \quad (45)$$

which achieves the required service level with the minimum expected holding cost obtained by the heuristic.

Integration with the Proposed Framework

The inventory model is the final stage of the proposed predictive–prescriptive framework. The age-based forecasting stage estimates the potential component failures, while the repair-or-replace stage determines the appropriate corrective action for each predicted failure. Components assigned to repair are updated using the imperfect repair model, whereas components assigned to replacement generate spare part demand. As a result, the revised replacement-demand sequence is used to estimate the state-dependent demand probabilities p_τ, the positive demand-size distribution $f(d)$, and the steady-state probabilities π_τ. These quantities are subsequently used to construct the lead-time demand distribution and determine the state-dependent order-up-to levels. Consequently, changes in component age, failure risk, maintenance decisions, and the updated effective age after repair can influence the resulting inventory policy. This integration follows the model sequence presented in Figure 1.

The Weibull failure probability and the Markov state-dependent demand probability have different roles in the framework. The Weibull model uses component age to forecast failures. The repair-or-replace decisions then determine which predicted failures generate demand for new spare parts. The resulting net-demand sequence is used to estimate the Markov state-dependent demand probabilities, where each state represents the time elapsed since the most recent positive demand. These probabilities are used to construct the lead-time demand distribution and determine inventory levels. Thus, the two probabilities are used sequentially and are not added or multiplied together.

This integration distinguishes the proposed model from the original time-varying order-up-to policy. The mathematical structure of the inventory model is maintained, but the demand input is generated from the technical condition of the installed components and the maintenance decisions. As a result, the inventory policy does not depend only on historical demand observations but also responds to the predicted operational risk of the system.

Results and Discussions

This section presents the results of the proposed framework using the Automatic Voltage Regulator (AVR) as the case-study component. The analysis begins with the demand characteristics and age-dependent failure behavior, followed by the adjustment of gross demand through the repair-or-replace decision and its effect on the resulting inventory policy.

The results are evaluated based on the estimated spare part demand, the achieved non-stockout probability, and the expected inventory holding cost. The discussion focuses on how the integration of component-age information and maintenance decisions changes the inventory requirement compared with a policy that directly treats predicted failures as replacement demand.

Table 2. Experimental setting

Parameter	Value
Case-study component	Automatic Voltage Regulator
Historical data period	2020–2024
Active installed components	34 units
Planning horizon	52 weeks
Baseline procurement lead time	4 weeks
Target non-stockout probability	0.95

Table 3. Characteristics of demand

Measure	Result	Threshold	Interpretation
ADI	4.0625	>1.32	Long intervals between demands
(CV^2)	0.0986	<0.49	Relatively stable positive demand size
Classification	Intermittent	–	Infrequent demand with stable size

Case Study and Experimental Setting

The proposed framework was applied to an electricity generation system in Indonesia. The system consists of several diesel generating units operating at different locations. The Automatic Voltage Regulator (AVR) was selected as the case-study component because it is a critical and repairable component. Its failure may interrupt generating-unit operation when a replacement component is not immediately available.

The analysis used 77 historical AVR failure records from 2020 to 2024, together with machine and component age data for 34 installed units. All historical failures were addressed through replacement. This study evaluates repair as an alternative, so that not every predicted failure requires a new spare part. The failure records and age data supported reliability parameterization and age-based failure forecasting.

The case study considers AVR requirements under a corrective maintenance policy. The proposed framework uses component ages and estimated failure probabilities to predict future failure events. These predicted events are then evaluated to plan repair-or-replace actions and estimate the resulting spare parts requirements. Therefore, inventory planning is based on anticipated failures rather than waiting for actual failures to be recorded.

The model was evaluated using weekly periods over a planning horizon of 52 weeks. The baseline procurement lead time was four weeks, while the required non-stockout probability was set at 0.95. The Weibull parameters were estimated using the Method of Moments Estimation, and the inventory policy was determined using the Greedy Marginal Analysis Heuristic. The main experimental setting is summarized in Table 2.

This experimental setting allows the proposed framework to be evaluated under actual intermittent demand conditions. The use of components with different operating ages is particularly important because the demand forecast is generated from the technical condition of the installed base rather than only from historical spare part transactions.

Demand Characteristics and Reliability Results

The historical AVR demand was first analyzed using the Average Inter-Demand Interval (ADI) and the squared coefficient of variation (CV^2). The calculated ADI was 4.0625, while the (CV^2) value was 0.0986. The ADI value was higher than the threshold of 1.32, whereas the (CV^2) value was lower than 0.49. Based on the Syntetos–Boylan classification [29], the AVR demand was categorized as intermittent, as summarized in Table 3.

The results show that AVR demand occurs infrequently, but the quantity is relatively stable when positive demand occurs. Under this condition, historical demand alone may not sufficiently represent future requirements because periods without demand do not necessarily indicate that the future failure risk is low. The risk may increase when several installed components enter older operating ages, although no recent spare part demand has been recorded.

The Weibull parameters were then estimated from the historical failure-age data using the Method of Moments Estimation. The estimation resulted in a shape parameter of $\gamma = 3$ and a scale parameter of $\beta = 142.313$weeks. Since $\gamma > 1$, the AVR components exhibit an increasing failure rate, indicating a wear-out failure pattern.

Table 4. Effect of repair-or-replace decisions on spare part demand

Measure	Result
Gross failure-driven demand	14 units
Repair decisions	3 events
Net procurement demand	12 units
Effective demand reduction	2 units
Relative demand reduction	14.3%

The age-dependent failure pattern differentiates the proposed approach from a demand model that assumes the same risk for all operating components. In the proposed framework, older components provide a higher contribution to expected demand than recently installed components. Therefore, changes in the age composition of the installed base may affect the demand estimate even when the historical demand pattern does not substantially change.

From Gross Demand to Net Demand

The estimated Weibull parameters were used to calculate the probability of failure for each of the 34 installed AVR components according to its current operating age. The component-level probabilities were then aggregated for each weekly period. An operational threshold of 0.45 was used to convert the aggregated forecast into discrete predicted demand events. The threshold was selected based on operational considerations in the case-study setting, taking into account AVR criticality and the need to anticipate spare parts requirements during the procurement lead time. It represents a practical planning assumption rather than a statistically optimized value and may require adjustment for other operating conditions. These events were subsequently evaluated using the repair-or-replace model to determine whether new spare parts were required. Under this threshold, the age-based model generated a cumulative gross demand of 14 units over the 52-week planning horizon.

The gross demand represents the failure-driven requirement before the maintenance decision is considered. It should not be directly interpreted as the procurement requirement because a failed AVR may still be repaired and returned to operation. Therefore, each predicted failure was evaluated using the repair-or-replace decision model.

The economic evaluation identified three failure events for which repair was preferred to immediate replacement. These decisions occurred in weeks 8, 13, and 30. Under the imperfect repair assumption, the three repaired components were estimated to remain operational for an additional 58, 35, and 78 weeks, respectively.

After the effective component ages were updated and the subsequent failure probabilities were recalculated, the cumulative net demand was 12 units. Thus, the repair-or-replace adjustment reduced the estimated procurement requirement by two units, or approximately 14.3%, as summarized in Table 4.

The reduction in net demand is not obtained by simply subtracting the three repair decisions from gross demand. Repair changes the effective age of the component and consequently changes its probability of failure in the remaining planning periods. Therefore, a repair may remove an immediate replacement requirement but may still generate demand at a later period when the repaired component fails again.

This result shows that the repair-or-replace decision affects both the quantity and timing of spare part demand. The quantity effect determines the total number of new components required, while the timing effect changes the demand intervals used by the inventory model. This distinction is important for intermittent demand because inventory availability depends not only on cumulative demand but also on when the demand occurs relative to the procurement lead time.

Inventory Policy Performance

The comparison uses the same state-dependent order-up-to inventory model for both gross and net demand. Its purpose is to examine the effect of incorporating repair-or-replace decisions into spare parts requirements. Therefore, the results show the effect of the maintenance adjustment within this inventory model and do not establish superiority over a conventional fixed base-stock policy.

The gross- and net-demand sequences were separately used to determine the state-dependent order-up-to policies. For the gross-demand sequence, the resulting policy was $\boldsymbol{S} = [2,2,3,3,3,3]$, with an achieved non-

Table 5. Inventory performance before and after the repair-or-replace adjustment

Performance measure	Gross-demand policy	Net-demand policy	Change
Demand states	6	10	–
Order-up-to policy, $\boldsymbol{S}$	$[2, 2, 3, 3, 3, 3]$	$[2, 1, 2, 3, 3, 3, 2, 3, 3, 3]$	–
Non-stockout probability	0.9544	0.9510	−0.0034
Expected holding cost, H	1.5147	1.4273	−5.77%
Holding cost over 52 weeks	Rp251.25 million	Rp225.00 million	−10.45%

Table 6. Sensitivity to procurement lead time

Lead time	State-dependent order-up-to policy, S	Expected holding cost, H	Change from baseline
2 weeks	$[2, 0, 0, 1, 2, 2, 0, 1, 2, 2]$	0.7584	−46.9%
3 weeks	$[2, 1, 1, 2, 2, 3, 1, 2, 3, 3]$	1.1874	−16.8%
4 weeks	$[2, 1, 2, 3, 3, 3, 2, 3, 3, 3]$	1.4273	Baseline
5 weeks	$[2, 2, 3, 3, 3, 3, 3, 3, 4, 3]$	1.5529	+8.8%
6 weeks	$[3, 3, 3, 3, 3, 3, 3, 3, 4, 3]$	1.6613	+16.4%

stockout probability of 0.9544 and an expected normalized holding cost of 1.5147. After the repair-or-replace adjustment, the resulting policy became $\boldsymbol{S} = [2,1,2,3,3,3,2,3,3,3]$, with a non-stockout probability of 0.9510 and an expected holding cost of 1.4273. Both policies satisfied the minimum target of 0.95.

The net-demand policy produced a slightly lower non-stockout probability than the gross-demand policy. However, the resulting value remained above the required target. This indicates that the gross-demand policy contained an additional inventory buffer that provided only a limited service improvement but increased the inventory cost.

The repair-or-replace adjustment also changed the demand-state structure from six to ten states. This occurred because the repair decisions extended some intervals between positive demands. Consequently, the proposed policy did not only reduce the total demand but also redistributed the target inventory levels according to the revised demand intervals.
The expected normalized holding cost decreased by approximately 5.77%, from 1.5147 to 1.4273. When the state-dependent inventory levels were evaluated over the complete 52-week sequence using the case-study holding-cost parameter, the total holding cost decreased from Rp251.25 million to Rp225.00 million. This represents a potential annual saving of Rp26.25 million or 10.45%.

The difference between the reduction in the expected H value and the reduction in the 52-week holding cost results from the measures used. The H value represents the long-run expected holding cost based on the steady-state distribution, whereas the 52-week value is accumulated according to the state sequence occurring during the planning horizon. Both results nevertheless show that using net demand can reduce inventory expenditure while maintaining the required service level.

More importantly, the cost reduction was not obtained by uniformly decreasing the inventory level in all periods. The net-demand policy maintained higher stock in states with greater demand risk and reduced the target stock in states with lower demand probability. Therefore, the benefit of the proposed framework comes from reallocating inventory according to the updated demand risk rather than only reducing the total quantity of inventory.

Sensitivity Analysis and Managerial Implications

The sensitivity analysis focused on procurement lead time and the target non-stockout probability. These parameters were selected because lead time determines the duration of exposure to failure before replenishment arrives, while the service target determines the amount of protection required against uncertain demand.

Sensitivity to Procurement Lead Time

The baseline procurement lead time was four weeks. The lead time was varied from two to six weeks while maintaining the service target at 0.95. The results are presented in Table 6.

The expected holding cost increased consistently with lead time. Shortening the lead time from four to two weeks reduced H by approximately 46.9%, whereas extending it to six weeks increased H by approximately 16.4%. A longer lead time requires more inventory because a greater number of components remain exposed to failure while the replenishment order is in transit.

Table 7. Sensitivity to the target non-stockout probability

Service target	State-dependent order-up-to policy, S	Expected holding cost, H	Change from baseline
0.80	$[0, 0, 1, 2, 2, 2, 2, 2, 2, 1]$	0.4840	−66.1%
0.85	$[1, 1, 1, 2, 2, 2, 1, 2, 2, 1]$	0.6924	−51.5%
0.90	$[2, 1, 2, 2, 2, 2, 2, 2, 2, 1]$	0.9968	−30.2%
0.95	$[2, 1, 2, 3, 3, 3, 2, 3, 3, 3]$	1.4273	Baseline
0.99	$[4, 3, 2, 3, 4, 4, 3, 4, 4, 5]$	2.4859	+74.2%

The results also show that reducing lead time may provide a larger cost improvement than making small adjustments to the inventory policy alone. Therefore, inventory managers should not treat lead time as a fixed external parameter. Supplier selection, order processing, transportation arrangements, and component standardization may provide opportunities to reduce the required safety stock.

Sensitivity to the Target Non-Stockout Probability

The target non-stockout probability was varied between 0.80 and 0.99, with 0.95 used as the baseline. The resulting inventory policies and expected holding costs are presented in Table 7.
The relationship between the service target and holding cost was non-linear. Increasing the target from 0.95 to 0.99 required a 74.2% increase in expected holding cost, although the service target increased by only four percentage points. This occurs because the inventory policy must cover demand outcomes with increasingly lower probabilities when the target approaches one.

Conversely, reducing the target from 0.95 to 0.90 decreased the expected holding cost by approximately 30.2%. However, this reduction should not automatically be considered preferable because an AVR stockout may interrupt generating-unit operation. The appropriate target should therefore be selected by considering the consequence of stockout in addition to inventory cost.
The baseline target of 0.95 provides a practical balance in the case study. It can be achieved with substantially less inventory than the 0.99 target, while providing a higher level of protection than the 0.90 scenario. Nevertheless, the sensitivity results indicate that the service target should be differentiated according to component criticality rather than applying the same value to all spare parts.

Conclusion

This study developed an adaptive age-based predictive inventory framework for spare parts with intermittent demand. The proposed framework integrates component age and failure probability, repair-or-replace decisions, and inventory control within a single decision process. The framework was applied to an Automatic Voltage Regulator (AVR) as the case-study component. The results show that spare part demand should not be determined only from historical consumption, particularly when demand is mainly generated by component failures and maintenance decisions.

The age-based failure analysis provides an estimate of the expected gross demand during the lead time. This demand is then adjusted based on the repair-or-replace decision to obtain the net demand that must be fulfilled by the inventory system. The adjustment prevents every component failure from being directly interpreted as a requirement for a new spare part. Consequently, the resulting inventory policy can maintain the required service level while avoiding unnecessary inventory and holding costs.

The main contribution of this study is the integration of reliability information, maintenance decisions, and adaptive inventory control for intermittent spare part demand. However, the implementation was limited to one critical component and a specific maintenance environment. The demand model focuses on predicted failures under a corrective maintenance policy. In settings where preventive or condition-based replacement is also performed before failure, the timing and spare parts requirements of those actions need to be included. Component ages and subsequent failure forecasts must also be updated because an early replacement changes future failure risk. Such an extension would avoid omitting planned replacement demand or counting a planned replacement and the failure it prevents as separate requirements.

The framework also evaluates maintenance actions and inventory decisions sequentially. The replacement evaluation uses a procurement-based reference condition and does not account for the stock available at each predicted failure. Future research could jointly determine maintenance actions and inventory levels, allowing the replacement waiting time to depend on stock availability and considering the associated holding costs.

Future research could extend the framework to multiple spare parts. For each spare part, the model would require its own failure parameters, component ages, repair and replacement costs, procurement lead time, and service-level target. If demand and available resources are independent across spare parts, the forecasting, maintenance decision, and inventory calculations could be performed separately for each item.

When spare parts share a purchasing budget, storage space, or repair capacity, the extension would require joint inventory and maintenance decisions under these shared constraints. Dependence between component failures and differences in spare part criticality would also need to be considered where relevant. These extensions would allow the framework to address a wider range of maintenance and inventory settings. Further research could also consider uncertain repair times and procurement lead times and longer observation periods. A simulation-based comparison with a conventional fixed base-stock policy under the same maintenance assumptions, demand scenarios, lead times, and service targets would provide a broader evaluation of inventory performance.

References

[1] T. B. Affonso, S. V. Conceição, L. R. Muniz, J. F. de Freitas Almeida, and J. C. de Lima, "A new hybrid forecasting method for spare part inventory management using heuristics and bootstrapping," *Decision Analytics Journal*, vol. 10, p. 100415, Mar. 2024, doi: https://doi.org/10.1016/j.dajour.2024.100415.

[2] S. Sareminia and F. Amini, "A reliable and ensemble forecasting model for slow-moving and repairable spare parts: Data mining approach," *Comput. Ind.*, vol. 145, p. 103827, Feb. 2023, doi: https://doi.org/10.1016/j.compind.2022.103827.

[3] Q. Hu, J. E. Boylan, H. Chen, and A. Labib, "OR in spare parts management: A review," *Eur. J. Oper. Res.*, vol. 266, no. 2, pp. 395–414, Apr. 2018, doi: https://dooi.org/10.1016/j.ejor.2017.07.058.

[4] N. Y. Salsabila, N. Siswanto, E. Widodo, and O. A. Rochmadhan, "Throughput analysis on a multi-state manufacturing system by considering availability," *Jurnal Teknik Industri: Jurnal Keilmuan dan Aplikasi Teknik Industri*, vol. 21, no. 2, pp. 69–78, 2019, doi: https://doi.org/10.9744/jti.21.2.69-78.

[5] S. Zhu, W. van Jaarsveld, and R. Dekker, "Spare parts inventory control based on maintenance planning," *Reliab. Eng. Syst. Saf.*, vol. 193, p. 106600, Jan. 2020, doi: https://doi.org/10.1016/j.ress.2019.106600.

[6] X. Tian, H. Wang, and E. E, "Forecasting intermittent demand for inventory management by retailers: A new approach," *Journal of Retailing and Consumer Services*, vol. 62, p. 102662, Sep. 2021, doi: https://doi.org/10.1016/j.jretconser.2021.102662.

[7] Ç. Pinçe, L. Turrini, and J. Meissner, "Intermittent demand forecasting for spare parts: A Critical review," *Omega.*, vol. 105, p. 102513, Dec. 2021, doi: https://doi.org/10.1016/j.omega.2021.102513.

[8] P. G. Giannopoulos, T. K. Dasaklis, I. Tsantilis, and C. Patsakis, "Machine learning algorithms in intermittent demand forecasting: a review," *Int. J. Prod. Res.*, pp. 1–43, Oct. 2025, doi: https://doi.org/10.1080/00207543.2025.2578701.

[9] M. Mouschoutzi and S. T. Ponis, "A comprehensive literature review on spare parts logistics management in the maritime industry," *Asian Journal of Shipping and Logistics*, vol. 38, no. 2, pp. 71–83, Jun. 2022, doi: https://doi.org/10.1016/j.ajsl.2021.12.003.

[10] C.-F. Chien, C.-C. Ku, and Y.-Y. Lu, "Ensemble learning for demand forecast of After-Market spare parts to empower data-driven value chain and an empirical study," *Comput. Ind. Eng.*, vol. 185, p. 109670, Nov. 2023, doi: https://doi.org/10.1016/j.cie.2023.109670.

[11] T. E. Goltsos, A. A. Syntetos, C. H. Glock, and G. Ioannou, "Inventory – forecasting: Mind the gap," *Eur. J. Oper. Res.*, vol. 299, no. 2, pp. 397–419, Jun. 2022, doi: https://doi.org/10.1016/j.ejor.2021.07.040.

[12] M. Hasni, M. S. Aguir, M. Z. Babai, and Z. Jemai, "Spare parts demand forecasting: a review on bootstrapping methods," *Int. J. Prod. Res.*, vol. 57, no. 15–16, pp. 4791–4804, 2019, doi: https://doi.org/10.1080/00207543.2018.1424375.

[13] P. Jiang, Y. Huang, and X. Liu, "Intermittent demand forecasting for spare parts in the heavy-duty vehicle industry: A support vector machine model," *Int. J. Prod. Res.*, vol. 59, no. 24, pp. 7423–7440, 2021, doi: https://doi.org/10.1080/00207543.2020.1842936.

[14] F. Lolli, R. Gamberini, A. Regattieri, E. Balugani, T. Gatos, and S. Gucci, "Single-hidden layer neural networks for forecasting intermittent demand," *Int. J. Prod. Econ.*, vol. 183, pp. 116–128, Jan. 2017, doi: https://doi.org/10.1016/j.ijpe.2016.10.021.

[15] X. Zhuang, Y. Yu, and A. Chen, "A combined forecasting method for intermittent demand using the automotive aftermarket data," *Data Science and Management*, vol. 5, no. 2, pp. 43–56, Jun. 2022, doi: https://doi.org/10.1016/j.dsm.2022.04.001.

[16] J. M. Rožanec, B. Fortuna, and D. Mladenić, "Reframing demand forecasting: A two-fold approach for lumpy and intermittent demand," *Sustainability*, vol. 14, no. 15, p. 9295, Jul. 2022, doi: https://doi.org/10.3390/su14159295.

[17] S. P. Kenaka, A. Cakravastia, A. Ma'ruf, and R. T. Cahyono, "enhancing intermittent spare part demand forecasting: A novel ensemble approach with focal loss and SMOTE," *Logistics*, vol. 9, no. 1, Mar. 2025, doi: https://doi.org/10.3390/logistics9010025.

[18] S. Van der Auweraer and R. Boute, "Forecasting spare part demand using service maintenance information," *Int. J. Prod. Econ.*, vol. 213, pp. 138–149, Jul. 2019, doi: https://doi.org/10.1016/J.IJPE.2019.03.015.

[19] F. Safaei, J. Ahmadi, and N. Balakrishnan, "A repair and replacement policy for repairable systems based on probability and mean of profits," *Reliab. Eng. Syst. Saf.*, vol. 183, pp. 143–152, Mar. 2019, doi: https://doi.org/10.1016/j.ress.2018.11.012.

[20] S. H. Sheu, T. H. Liu, Z. G. Zhang, and H. N. Tsai, "The generalized age maintenance policies with random working times," *Reliab. Eng. Syst. Saf.*, vol. 169, pp. 503–514, Jan. 2018, doi: https://doi.org/10.1016/j.ress.2017.09.003.

[21] Y.-H. Chien and Z. G. Zhang, "Optimal replacement policy based on number of failures for a system with multiple attempt minimal repairs," *Reliab. Eng. Syst. Saf.*, vol. 251, p. 110366, Nov. 2024, doi: https://doi.org/10.1016/j.ress.2024.110366.

[22] S. Tekin, N. O. Bakır, and B. Keleş, "Maintenance policies for protection systems with imperfect inspection and imperfect repair," *Reliab. Eng. Syst. Saf.*, vol. 257, p. 110798, May 2025, doi: https://doi.org/10.1016/j.ress.2025.110798.

[23] I. Nafisah, M. Shrahili, N. Alotaibi, and P. Scarf, "Virtual series-system models of imperfect repair," *Reliab. Eng. Syst. Saf.*, vol. 188, pp. 604–613, 2019, doi: https://doi.org/10.1016/j.ress.2019.03.046.

[24] Y. Ye, Y. Lu, P. Robinson, and A. Narayanan, "An empirical Bayes approach to incorporating demand intermittency and irregularity into inventory control," *Eur. J. Oper. Res.*, vol. 303, no. 1, pp. 255–272, Nov. 2022, doi: https://doi.org/10.1016/j.ejor.2022.02.033.

[25] B. Rostami-Tabar and S. M. Disney, "On the order-up-to policy with intermittent integer demand and logically consistent forecasts," *Int. J. Prod. Econ.*, vol. 257, p. 108763, Mar. 2023, doi: https://doi.org/10.1016/j.ijpe.2022.108763.

[26] C. L. P. Pennings, J. van Dalen, and E. A. van der Laan, "Exploiting elapsed time for managing intermittent demand for spare parts," *Eur. J. Oper. Res.*, vol. 258, no. 3, pp. 958–969, 2017, doi: https://doi.org/10.1016/j.ejor.2016.09.017.

[27] M. Z. Babai, H. Chen, A. A. Syntetos, and D. Lengu, "A compound-Poisson Bayesian approach for spare parts inventory forecasting," *Int. J. Prod. Econ.*, vol. 232, p. 107954, Feb. 2021, doi: https://doi.org/10.1016/j.ijpe.2020.107954.

[28] D. Prak and P. Rogetzer, "Timing intermittent demand with time-varying order-up-to levels," *Eur. J. Oper. Res.*, vol. 303, no. 3, pp. 1126–1136, Dec. 2022, doi: https://doi.org/10.1016/j.ejor.2022.03.019.

[29] A. A. Syntetos, J. E. Boylan, and J. D. Croston, "On the categorization of demand patterns," *Journal of the Operational Research Society*, vol. 56, no. 5, pp. 495–503, May 2005, doi: https://doi.org/10.1057/palgrave.jors.2601841.

[30] V. Bafandegan Emroozi, M. Kazemi, and M. Doostparast, "Enhancing industrial maintenance planning: Optimization of human error reduction and spare parts management," *Operations Research Perspectives*, vol. 14, p. 100336, Jun. 2025, doi: https://doi.org/10.1016/j.orp.2025.100336.