Received: 2 November 2025 | Revised: 27 March 2026 | Accepted: 16 April 2026 | Published online: 19 May 2026

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2026, Vol. 00(00) 1–14
DOI: 10.47852/bonviewJDSIS62028066

BON VIEW PUBLISHING

Reliability Modeling and Predictive Maintenance Integration for an Induced Draft Fan System Using Semi-Markov Process and Machine Learning-Based Predictive Models

Prawar Chaudhary[1] **and Kaushal Kumar**[2,*]

[1] *School of Basic and Applied Sciences, K.R. Mangalam University, India*

[2] *Department of Mechanical Engineering, K.R. Mangalam University, India*

Abstract: The proposed research enhances the dependability and efficiency levels of induced draft fan systems in thermal power plants due to the combination of semi-Markov process reliability modeling and machine learning-based predictive maintenance (PdM) methods. The system in question is a three-fan system that has two working units and one standby unit that runs cold but is not operational to give redundancy. The most important reliability metrics, such as mean time to system failure, system availability, repairman busy period, and predicted downtime, are analyzed analytically with the semi-Markov and regenerative point methods. In order to supplement the analytical reliability model, a PdM framework through machine learning with the help of a Random Forest classifier is created to predict possible failure learning conditions based on historical data on its operation. The model examines the parameters including the failure rate, the repair rate, the downtimes, and the operational capacity to detect the initial signs of system degradation. The proposed structure can greatly ensure the minimization of unplanned breakdowns and maximization of overall system performance due to the possible proactive scheduling of maintenance activities. Hybrid probabilistic reliability analysis and data-driven predictive modeling offer a solution to increasing operational decision-making in thermal power plants' maintenance systems. The suggested methodology shows that there is a prospect of relying on both reliability engineering and machine learning to realize effective real-time monitoring, enhanced availability, and cost-effective maintenance planning of vital industrial infrastructure.

Keywords: reliability engineering, quality control, predictive maintenance, machine learning, semi-Markov process

1. Introduction

In developing countries, the power sector is one of the fastest-growing industries and plays a crucial role in driving overall economic development. As a result, enhancing the efficiency and performance of this sector has become essential, which can be achieved through the adoption of advanced techniques and improved system models [1]. Reliability, in this context, refers to the probability that a system will operate as expected for a specified duration under given conditions. While achieving higher reliability is always desirable, it often leads to increased operational and maintenance costs. Therefore, it is important to strike a balance by improving system reliability in a cost-effective manner [2]. Yu et al. [3] developed a model and analyzed two pricing and warranty scenarios to determine the best product dependability under a free replacement and repair scheme. According to the model, the selling rate fluctuates as the price rises and the warranty period lengthens. The reliability quantities of a two-unit system were determined by Xiao et al. [4], and the optimal preventive maintenance procedures for maximizing the average time to failure and availability of the system were studied analytically. Yadav et al. [5] proposed a model that aids in selecting the best maintenance procedures to ensure the system's most significant overall availability. The optimum failure and repair rates for each subsystem were determined when the primary goal was economic optimization. Bai et al. [6] and Dhanda et al. [7, 8] discussed approaches for assessing the availability and efficacy of continuous production systems. The cost-benefit study of a system with two identical units, one operating and the other held in cold standby, is covered in. When something goes wrong, the system is instantly attended to by a single server. After being repaired, the unit deteriorates. Mellal and Zio [9] performed a stochastic analysis of a two-unit cold standby system. It is assumed that all temporal distributions are random. The regenerating point technique is used to create several system dependability measurements. Sansui et al. [10] and Behboudi et al. [11] considered a model in which human error is corrected using the Gumbel–Hougaard family copula, while the primary and backup units

*Corresponding author: Kaushal Kumar, Department of Mechanical Engineering, K.R. Mangalam University, India. Email: kaushal.kumar@krmangalam.edu.in

are repaired using a general distribution. The system is examined using the Laplace transformation and the additional variable technique. Numerous reliability parameters, like availability, mean time to failure, and profit function, have been examined for the system. Using the regenerating point technique, Juybari et al. [12], Kumar et al. [13], Gitanjali [14], and Koutras [15] utilized the Markov process to measure a variety of system parameters. The parameters defining the system's or dependability unit's characteristics are also assumed to be random variables because life-testing tests require time [16]. The system follows the Weibull distribution, which is the most applied distribution on reliability studies [17], [18], [19], simulating the operation of a cold standby system with two units that receive regular maintenance, inspection, and repair facilities under different weather conditions, with preventative maintenance taking priority over inspection. Replacement should be done with the units if repair is not feasible during the inspection. After a certain period of operation, the operational unit is subjected to preventative maintenance [20]. When the unit is completely broken, the server repairs it.

Yan et al. [21] have applied stochastic models for reliability evaluation of redundant industrial systems. The goal in the paper by Xing et al. [22] is to use a feed-forward neural network approach using MATLAB to determine the cost and factors of dependability of a chocolate manufacturing plant. Yan et al. [21] determined the multi-state probability of its various efficiency levels for a metal sheet production unit. The change in these multi-state probabilities was estimated, in the long run, using the artificial neural network technique. Researchers also assumed and analyzed a system of three units that was being inspected by a unit of correlated unit failure and repair times [23]. Our research serves to add a hybrid model where probabilistic reliability models are combined with AI-based predictive maintenance (PdM), a model that is not extensively used in the current thermal power infrastructure.

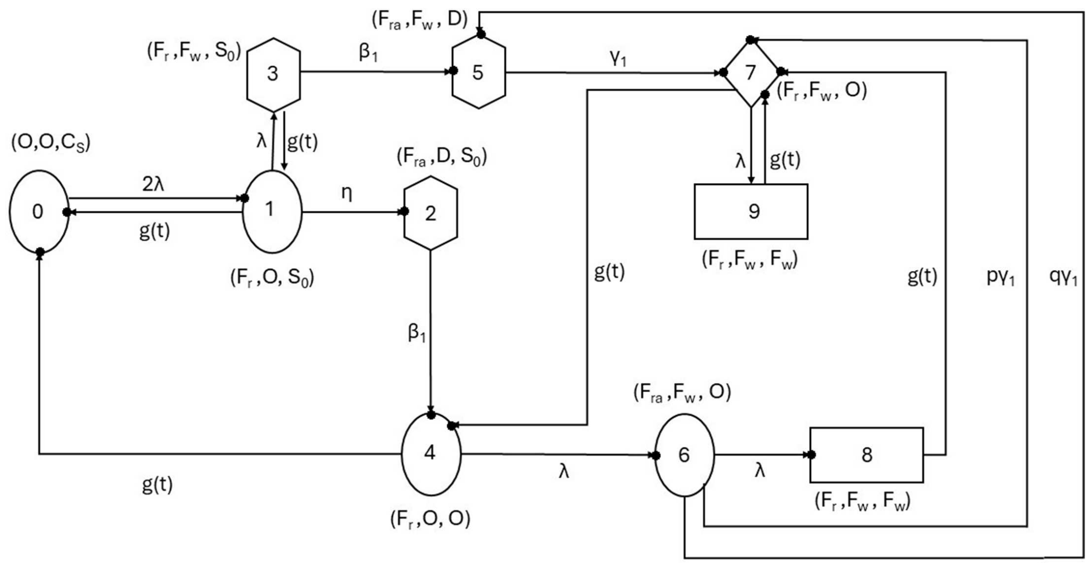
2. System Description

The data was extracted from the local plant, and it is an overall account of the operational activities of a thermal power plant in 1 year (365 days). The main parameters that are important in the reliability and maintenance analysis in each entry include the number of failures reported, daily operating hours, number of repairs done, and total downtime hours. It also monitors the working hours of both professional and ordinary repairmen, which allows them to analyze costs and manpower. Financial metrics are logged daily, including revenue generated at full and reduced capacities, maintenance costs, and losses due to downtime. The Synthetic Minority Oversampling Technique was used to generate additional minority samples.

The model under consideration has three induced draft (ID) fans installed in the boilers of thermal power plants, where repair of failed units is done by an expert repairman (Re) and controlling of parameters, as well as switching over, is done by an ordinary repairman (Ro). Two out of three ID fans operate, and the third operates as a cold standby to offer redundancy. If one of the operating ID fans fails, the cold standby ID fan performs immediately after some activation time. Priority is given to repair over controlling the parameters in the unit. The system may operate at a reduced capacity if parameter optimization is given priority over repair [24–26]. The states 0, 1, 2, 4, 5, 6, 7, and 8 in Figure 1 are called regenerative states; 0, 1, 4, and 6 are up states; 2, 3, and 5 are down states; 9 and 8 are failed states; and state 7 is a reduced-capacity state.

If repair is prioritized above controlling parameters in a functional state. In that case, the system may still be able to operate at total capacity, which is not necessarily achievable if controlling parameters are given priority over repair. But if repair is prioritized, there may be a drawback of entering a down or failed state and no chance of working at a reduced capacity [27, 28]. However, priority is given to controlling the parameters over repair in the down state. To determine reliability parameters and variation in their cost analysis, the system models were evaluated using the semi-Markov process (SMP) theory and the Regenerative Point Technique (RGPT). A researcher applied SMPs to analyze system availability and maintainability [28]. However, these studies focus primarily on probabilistic reliability modeling and do not incorporate intelligent PdM techniques. Recent studies have explored machine learning approaches for PdM. Many researchers reviewed AI-driven PdM in industrial systems [29–32]. However, most

Figure 1
State transition diagram



studies focus either on statistical reliability modeling or machine learning independently. The integration of stochastic reliability models with AI-based PdM remains limited [33, 34].

Despite significant progress in reliability modeling and PdM, few studies integrate semi-Markov reliability analysis with machine learning-based PdM frameworks for thermal power plant systems. This study addresses this gap by combining stochastic reliability modeling with Random Forest-based failure prediction to improve maintenance decision-making.

2.1. Assumptions of the model

1) The system is initially operating at its full capacity under normal working conditions.
2) At the initial stage, two ID fans are in active operation, while the third fan is maintained as a cold standby unit.
3) All three ID fans are identical in design and performance, and together, they form a parallel system configuration.
4) The duration required to transition the cold standby unit into an active operational state is defined as the activation time.
5) The standby unit is assumed to be failure-free during the switching or activation process.
6) A repair facility is always available in the system. The Ro is responsible for system monitoring, parameter control, and activation of standby units, whereas the Re performs the repair operations.
7) The Ro remains continuously present with the system to ensure smooth operation and immediate response to system requirements.
8) Repair activities are temporarily suspended when the repairman is engaged in controlling system parameters to maintain operational safety.
9) Priority is given to the repair of failed operative units over parameter control tasks to ensure system continuity and performance.
10) The failure times of the system components are assumed to follow an exponential distribution, implying a constant failure rate.
11) After each repair, the system is restored to a condition equivalent to a new system, indicating a perfect repair assumption.

2.2. Mathematical notations

For the reader's convenience, all mathematical symbols used throughout the paper are defined in a glossary (Table 1) prior to the analytical sections. This ensures clarity and ease of understanding of the reliability model.

Table 1
Notations used throughout the model

Notation	Meaning
O	System operating in a fully functional (operative) state
C_S	Cold standby condition of a unit
F	Failed condition of a unit
F_r	Failed unit currently undergoing repair
F_R	Unit continuing in repair from a previous state
F_{ra}	Repair process temporarily suspended (in abeyance)
F_w	Failed unit awaiting repair
S_0	State representing the switching/activation process
λ	Constant failure rate of an operative unit
D	System in a down (non-operational) state
η	Transition rate from operational (up) state to down state when only one unit is active
γ_1	Rate corresponding to the allowable time for parameter adjustment to enable reduced-capacity operation
β_1	Activation rate of the standby unit
p	Probability that system parameters are successfully adjusted to achieve reduced-capacity operation
q	Probability that parameter adjustment fails, preventing reduced-capacity operation
g(t)	Probability density function of repair time
G(t)	Cumulative distribution function of repair time
C_0	Revenue earned per unit time when the system operates at full capacity
C_1	Revenue earned per unit time under reduced-capacity operation
C_2	Cost incurred per unit time when the repairman is engaged in repair activities
C_3	Cost incurred per unit time when the system remains non-operational
C_4	Loss incurred per unit time due to reduced power generation
C_5	Payment rate per unit time to the expert repairman
C_6	Payment rate per unit time to the ordinary repairman
AF_0	Probability that the system operates at full capacity given it starts in state 0 at time ($t = 0$)
AR_0	Probability that the system operates at reduced capacity given it starts in state 0 at time ($t = 0$)
B_0	Total busy period of the repairman dedicated to repair work
DT_0	Expected downtime; proportion of total time during which the system remains non-operational

3. Problem Formulation

The reliability behavior of the ID fan system is developed to a stochastic modeling framework on the basis of SMP and regenerative point technique (RPT). The system comprises three identical ID fans, installed in a 2-out-of-3 setup with two fans running at the same time and the third piece of equipment in the cold standby position to offer redundancy. System performance is determined by various stochastic events such as component failures, repairs done by skilled repairmen, parameter control operations done by common repairmen, and activation of the standby unit when an operating unit malfunctions. The system is modeled as a finite set of operating states that describe full-capacity operation, reduced-capacity operation, repair conditions, waiting states, and total system failure. The transition between these states will be at probabilistic rates such as the failure rate of operating units, the rate of repair, activation rate of standby, and the likelihood of parameter revision that will allow the system to run at reduced capacity. This analysis is especially appropriate to semi-Markov modeling, in that the semi-Markov model permits state transitions to be based on arbitrary time distributions as opposed to the memoryless assumption in Markov models. The regenerative qualities of a given state of a system are used to derive a series of recursive equations assessing important system reliability metrics like mean time to system failure (MTSF), ability of the system to operate at full capacity and reduced capacity, repairman busy period, anticipated system downtime, and profitability of the system. These expressions of analysis provide a comprehensive basis for evaluating the performance of the system as well as assimilation of PdM plans based on operational data.

3.1. Transition probabilities

In order to mathematically model the dynamics of the system, the operational states that were identified in the section above are modeled using a state transition structure where the semi-Markov construct is used. The states are associated with a certain working mode of the ID fan system, including full-capacity work, reduced-capacity work, repair work, or malfunction. The passages between these states are caused by stochastic processes such as component failures, repair termination, parameter changes, and standby activation of units. Such transitions are described by transition probability functions based on the failure probability, repair probability, activation probability, and parameter control probability. These transition probabilities and the sojourn times in each regenerative state can then be used to describe the system stochastically. The Mathematical relationships used to obtain reliability measures are MTSF, System Availability, Busy Period of Repairman and System Downtime.

The epochs of entry into states 0, 1, 3, 5, and 6 are regeneration points, and thus, these states are regenerative states.

Various transition probabilities are given below:

$q_{01}(t) = 2\lambda\ e^{-2\lambda t}$

$q_{57}(t) = \gamma_1 e^{-Y1t}$

$q_{10}(t) = e^{-(\lambda+\eta)t}$g(t)

$q_{65}(t) = (q\gamma_1)\ e^{-(\lambda+\gamma_1)t}$

$q_{12}(t) = \eta\ e^{-(\lambda+\eta)t}\overline{G(t)}$

$q_{67}(t) = (p\gamma_1)\ e^{-(\lambda+\gamma 1)t}$

$q^3_{1,5}(t) = (-(\lambda+\eta)t \copyright \beta_1\ e^{-\beta_1 t})\overline{G(t)}$

$q_{68}(t) = (\lambda)\ e^{-(\lambda+\gamma 1)t}$

$q^3_{1,1}(t) = (\lambda e^{-(\lambda+\eta)t} \copyright e^{-\beta 1t})\ g(t)$

$q_{74}(t) = e^{-\lambda t}\ \underline{g(t)}\text{q}_{24}(t) = \beta_1\ e^{-\beta 1t}$

$q_{79}(t) = e^{-\lambda t}\overline{G(t)}$

$\text{q}_{40}(\text{t}) = e^{-\lambda t}g(t)$

$q^9_{7,7}(t) = (\lambda e^{-\lambda t} \copyright 1)\ g(t)$

$q_{46}(t) = (\lambda)e^{-\lambda t}\overline{G(t)}$

$q_{87}(t) = g(t)$

3.2. Mean sojourn times

The mean sojourn times (μ_i) in the regenerative state are given by

$\mu_0 = \frac{1}{2\lambda}$

$\mu_1 = g*(\beta_1)\frac{-1}{\beta_1} + g*(\lambda+\eta)\frac{1}{\lambda+\eta} + \frac{\lambda+\eta-\beta_1}{(\lambda+\eta)\beta_1}$

$\mu_2 = \frac{1}{\beta 1}$

$\mu_3 = -g^{*}{}'(0)$

$\mu_4 = \frac{1-g\times(\lambda)}{\lambda}$

$\mu_5 = \frac{1}{\gamma 1}$

$\mu_6 = \frac{1}{\lambda+\gamma 1}$

$\mu_7 = \frac{1-g\times(\lambda)}{\lambda}$

$\mu_8 = -g^{*}{}'(0)$

$\mu_9 = -g^{*}{}'(0)$

The mean lifetime by the model for transiting into any state *j* counted from the time of arrival point of state *i*. Therefore, the relation we have,

$\text{m}_{01} = \mu_0$

$\text{m}_{10} + \text{m}_{12} + \text{m}^3_{1,1} + \text{m}^3_{1,5} = k$

$\text{m}_{24} = \mu_2$

$\text{m}_{40} + \text{m}_{46} = \mu_4$

$\text{m}_{65} + \text{m}_{67} + \text{m}_{68} = \mu_6$

$\text{m}_{57} = \mu_5$

$\text{m}_{74} + \text{m}_{79} = \mu_7$

$\text{m}_{74} + \text{m}^9_{7,7} = -g^{*}{}'(0) = \mu_9$

$\text{m}_{87} = \mu_8$

3.3. Mean time to system failure

It is the time up to which the model functions properly without failing. The Stieltjes convolution was used to derive system reliability transitions, as it accommodates time-dependent state changes in SMPs.

$\Phi_0(t) = \text{Q}_{01}(t) \circledS \Phi_1(t)$

$\Phi_1(t) = \text{Q}_{10}(t) \circledS \Phi_0(t) + \text{Q}_{12}(t) \circledS \Phi_2(t) + \text{Q}^3_{1,1}(t) \circledS \Phi_1(t) + \text{Q}^3_{1,5}(t) \circledS \Phi_5(t)$

$\Phi_2(t) = \text{Q}_{24}(t) \circledS \Phi_4(t)$

$\Phi_4(t) = \text{Q}_{40}(t) \circledS \Phi_0(t) + \text{Q}_{46}(t) \circledS \Phi_6(t)$

$\Phi_5(t) = \text{Q}_{57}(t) \circledS \Phi_7(t)$

$\Phi_6(t) = \text{Q}_{65}(t) \circledS \Phi_5(t) + \text{Q}_{67}(t) \circledS \Phi_7(t) + \text{Q}_{68}(t)$

$\Phi_7(t) = \text{Q}_{79}(t) + \text{Q}_{74}(t) \circledS \Phi_4(t)$

Taking the Laplace–Stieltjes transform and using Cramer's rule

$$\text{MTSF} = \lim_{s\to 0} \frac{1-\frac{N_0(s)}{D_0(s)}}{s} = \frac{N_1}{D_1}$$

$N_1 = D_0'(0) - N_0'(0)$

$$
\begin{aligned}
&= m_2 [p_{12} [1 - p_{46} p_{74} (1 - p_{68})]] \\
&+ m_0 [[1 - p_{1,1}^3] [1 - p_{46} p_{74} (1 - p_{68})] \\
&+ k [1 - p_{46} p_{74} (1 - p_{68})] \\
&+ m_7 [p_{1,5}^3 + p_{12} p_{46} (1 - p_{68})] \\
&+ m_4 \left[p_{12} + p_{1,5}^3 p_{74}\right] \\
&+ m_5 \left[p_{1,5}^3 + p_{12} p_{46} p_{65} - p_{1,5}^3 p_{46} p_{67} p_{74}\right] \\
&+ m_6 \left[p_{12} p_{46} + p_{46} p_{74} p_{1,5}^3\right]
\end{aligned}
\quad (1)
$$

$$
\begin{aligned}
D_0(0) = D_1 = p_{12} p_{46} [p_{68} p_{74} + p_{79}] \\
+ p_{1,5}^3 [p_{79} + p_{46} p_{68} p_{74}]
\end{aligned}
\quad (2)
$$

3.4. Availability analysis

Availability represents the probability that the system remains operational at a given time. For systems modeled using SMPs, availability can be derived using RPTs. The availability function $AF_i(t)$ denotes the probability that the system operates at full capacity at time t, given that it entered state i at time $t = 0$. Using the regenerative property of states, the availability equations can be derived as follows.

3.4.1. Availability analysis at full capacity

Suppose $\mathrm{AF}_i(t)$ be the probability that the model works at full capacity at time t given the model has entered state i at $t = 0$, which is regenerative.

$\mathrm{AF}_0(t) = \mathrm{M}_0(t) + q_{01}(t)© \mathrm{AF}_1(t)$

$\mathrm{AF}_1(t) = \mathrm{M}_1(t) + q_{10}(t)© \mathrm{AF}_0(t) + q_{12}(t)© \mathrm{AF}_2(t)$
$\quad + q_{1,1}^3 (t)© \mathrm{AF}_1(t) + q_{1,5}^3 (t)© \mathrm{AF}_5(t)$

$\mathrm{AF}_2(t) = q_{24}(t)© \mathrm{AF}_4(t)$

$\mathrm{AF}_4(t) = q_{40} (t)© \mathrm{AF}_0(t) + q_{46}(t)© \mathrm{AF}_6(t)$

$\mathrm{AF}_5(t) = q_{57}(t)© \mathrm{AF}_7(t)$

$\mathrm{AF}_6(t) = \mathrm{M}_6(t) + q_{68}(t)© \mathrm{AF}_8(t) + q_{67}(t)© \mathrm{AF}_7(t)$
$\quad + q_{65}(t)©\mathrm{AF}_5(t)$

$\mathrm{AF}_7(t) = q_{7,7}^9 (t)© \mathrm{AF}_7(t) + q_{74}(t)© \mathrm{AF}_4(t)$

$\mathrm{AF}_8(t) = q_{87}(t)© \mathrm{AF}_7(t)$

Where $\mathrm{M}_0(\mathrm{t}) = e^{-(2\lambda)t}$ $\mathrm{M}_1(\mathrm{t}) = [e^{-(\beta_1)t} - e^{-(\lambda+\eta)t}] \overline{G(t)}$
$\mathrm{M}_2(\mathrm{t}) = e^{-(\lambda+\gamma_1)t}$

Taking the Laplace transform and using Cramer's rule

$\mathrm{AF}_0 = \lim_{s\to 0} s \frac{N_2(s)}{D_2(s)} = \frac{N_2}{D_2}$ where

$$
\begin{aligned}
N_2 = \mu_0 \left[1 - p_{1,1}^3\right] p_{74} p_{40} + \mu_1 p_{74} p_{40} \\
+ \mu_6 p_{74} p_{46} (p_{12} + p_{1,5}^3)
\end{aligned}
\quad (3)
$$

$$
\begin{aligned}
D_2 = {} & \mu_5 [p_{74} [p_{46} p_{65} p_{1,5}^3 + p_{1,5}^3 p_{40} + p_{12} p_{46} p_{65}]] \\
& + \mu_8 p_{74} p_{68} p_{46} p_{12} + p_{1,5}^3 + k[p_{74} p_{40}] + \mu_0 p_{74} p_{40} [1 - p_{1,1}^3] \\
& + \mu_2 p_{12} p_{74} p_{40} + \mu_2 [p_{74} [p_{1,5}^3 + p_{12}]] + \mu_7 [p_{1,5}^3 + p_{12} p_{46}] \\
& + \mu_6 [p_{74} p_{46} [p_{1,5}^3 + p_{12}]]
\end{aligned}
\quad (4)
$$

3.4.2. Availability analysis at reduced capacity

$\mathrm{AR}_0(t) = q_{01}(t)© \mathrm{AR}_1(t)$

$\mathrm{AR}_1(t) = q_{10}(t)© \mathrm{AR}_0(t) + q_{12}(t)© \mathrm{AR}_2(t)$
$\quad + q_{1,1}^3 (t)© \mathrm{AR}_1(t) + q_{1,5}^3 (t)© \mathrm{AR}_5(t)$

$\mathrm{AR}_2(t) = q_{24}(t)© \mathrm{AR}_4(t)$

$\mathrm{AR}_4(t) = \mathrm{M}_4(t) + q_{40}(t)© \mathrm{AR}_0(t) + q_{46}(t)© \mathrm{AR}_6(t)$

$\mathrm{AR}_5(t) = q_{57}(t)© \mathrm{AR}_7(t)$

$\mathrm{AR}_6(t) = q_{68}(t)© \mathrm{AR}_8(t) + q_{65}(t)© \mathrm{AR}_5(t) + q_{67}(t)© \mathrm{AR}_7(t)$

$\mathrm{AR}_7(t) = \mathrm{M}_7(t) + q_{7,7}^9 (t)© \mathrm{AR}_7(t) + q_{74}(t)© \mathrm{AR}_4(t)$

$\mathrm{AR}_8(t) = q_{87}(t)© \mathrm{AR}_7(t)$

Where $M_4(t) = e^{-(\lambda)t}\overline{G(t)}$ $M_7(t) = e^{-(\lambda)t}\overline{G(t)}$

Taking the Laplace transform and using Cramer's rule, we have

$\mathrm{AR}_0 = \lim_{s\to 0} s \frac{N_3(s)}{D_2(s)} = \frac{N_3}{D_2}$ where

$$
N_3 = \mu_4 p_{74} [p_{12} + p_{1,5}^3] + \mu_7 [p_{12} p_{46} + p_{1,5}^3] \quad (5)
$$

3.5. Analysis of busy span of repairmen

The probability that the Repairman I is occupied with the repairment at time t, given that the model has entered state i at $t = 0$, which is regenerative.

$B_0(t) = q_{01}(t) © B_1(t)$

$B_1(t) = W_1(t) + q_{10}(t) © B_0(t) + q_{12}(t) © B_2(t)$
$\quad + q_{1,1}^3 (t) © B_1(t) + q_{1,5}^3 (t) © B_5(t)$

$B_2(t) = W_2(t) + q_{24}(t) © B_4(t)$

$B_4(t) = W_4(t) + q_{40}(t) © B_0(t) + q_{46}(t) © B_6(t)$

$B_5(t) = W_5(t) + q_{57}(t) © B_7(t)$

$B_6(t) = W_6(t) + q_{68}(t) © B_8(t) + q_{65}(t) © B_5(t)$
$\quad + q_{67}(t) © B_7(t)$

$B_7(t) = W_7(t) + q_{7,7}^9 (t) © B_7(t) + q_{74}(t) © B_4(t)$

$B_8(t) = W_8(t) + q_{87}(t) © B_7(t)$

Where $W_1(t) = [e^{-(\beta 1)t} - e^{-(\lambda+\eta)t}] \overline{G(t)}$,

$W_2(t) = e^{-\beta 1 t}$,

$W_4(t) = e^{-(\lambda)t}\overline{G(t)}$,

$W_5(t) = e^{-\gamma 1 t}$

$W_6(t) = e^{-(\lambda+\gamma 1)t}$,

$W_7(t) = e^{-(\lambda)t}\overline{G(t)}$,

$W_8(t) = \overline{G(t)}$

$\mathrm{B}_0 = \lim_{s\to 0} s \frac{N_4(s)}{D_2(s)} = \frac{N_4}{D_2}$

$$
\begin{aligned}
N_4 = {} & \mu_1 p_{74} p_{40} + \mu_2 p_{12} p_{74} p_{40} + \mu_4 p_{74} (p_{12} + p_{1,5}^3) \\
& + \mu_5 (p_{1,5}^3 p_{74} p_{40} + p_{74} p_{65} p_{46} (p_{12} + p_{1,5}^3)) \\
& + \mu_6 p_{74} p_{46} (p_{12} + p_{1,5}^3) + \mu_7 (p_{12} p_{46} + p_{1,5}^3) \\
& + \mu_8 p_{74} p_{46} p_{68} (p_{12} + p_{1,5}^3)
\end{aligned}
\quad (6)
$$

3.6. Anticipated downtime of the system

It is the probability that the model is down at time t, given that the model has entered the state i at $t = 0$, which is regenerative.

$DT_0(t) = q_{01}(t) © DT_1(t)$

$DT_1(t) = q_{10}(t) © DT_0(t) + q_{12}(t) © DT_2(t)$
$\quad + q^3_{1,1}\ (t) © DT_1(t) + q^3_{1,5}\ (t) © DT_5(t)$

$DT_2(t) = V_2(t) + q_{24}(t) © DT_4(t)$

$DT_4(t) = q_{40}(t) © DT_0(t) + q_{46}(t) © DT_6(t)$

$DT_5(t) = V_5(t) + q_{57}(t) © DT_7(t)$

$DT_6(t) = q_{68}(t) © DT_8(t) + q_{65}(t) © DT_5(t)$
$\quad + q_{67}(t) © DT_7(t)$

$DT_7(t) = q^9_{7,7}(t) © DT_7(t) + q_{74}(t) © DT_4(t)$

$DT_8(t) = q_{87}(t) © DT_7(t)$

Where $W_2(t) = e^{-\beta 1t}$ $W_5(t) = e^{-\gamma 1t}$

Taking the Laplace transform and using Cramer's rule, we have

$DT_0 = \lim_{s \to 0} s\ \frac{N_5(s)}{D_2(s)} = \frac{N_5}{D_2}$ where

$$N_5 = \mu_2 p_{12} p_{74} p_{40} + \mu_5 (p^3_{1,5}\ p_{74} p_{40} + p_{74} p_{65} p_{46} (p_{12} + p^3_{1,5}) \tag{7}$$

3.7. Analysis of profit

The anticipated profit per unit time of the model:

$$P = C_0\,(AF_0) + C_1\,(AR_0) - C_2\,(B_0) - C_3\,(DT_0) - C_4 - C_5 - C_6 \tag{8}$$

Where $C_1 < C_0$

Particular Case

Consider, $g(t) = \alpha e^{-\alpha t}$ $g_1{}^*{}'(0) = -\frac{1}{\alpha}$

The values of various measurements of system effectiveness are obtained in Table 2 using the values estimated from the data collected, that is, $p = 0.8$, $q = 0.2$, 1 $\beta = 0.3$, $\alpha = 0.1$, $\eta = 0.5$, $\gamma_1 = 0.6$, $\gamma_2 = 0.6$, $\lambda = 0.0001$, $C_0 = 100000$, $C_1 = 60000$, $C_2 = 20000$, $C_3 = 40000$, $C_4 = 30000$, $C_5 = 500$, $C_6 = 41.6667$

The dataset consists of a 1-year log from a thermal plant supplemented with synthetic data points created using distribution estimates to address class imbalance.

3.8. Graphical analysis

Figure 2 depicts the behavior of Mean Time To Failure (M.T.T.F) with respect to failure rate. It shows that when the system is having failures frequently or when the failure rate is high, the MTSF goes down; that is, the system is failing more often. The high value of MTSF obtained from the model arises from the redundancy structure of the system and the assumed exponential failure distribution with a relatively low failure rate (λ = 0.0001). In reliability theory, such values represent theoretical expectations under idealized stochastic assumptions rather than physical operational lifespan. The redundancy of two operational units with one cold standby significantly increases system survival probability. To ensure realism, sensitivity analysis with varying failure rates is presented in Figure 2, demonstrating the expected decline in MTSF as the failure rate increases.

Figure 2
Mean time to failure with respect to failure rate (λ)

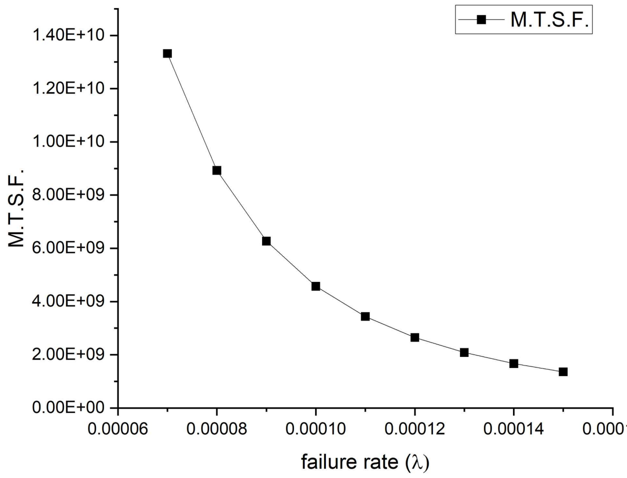


Figure 3 is a graph plotted between the availability of the system working at full capacity and the failure rate, which indicates that with an increase in failure rate, the availability of the system running at full capacity decreases.

Figure 4 shows the profit generated by the system with failure rate, and it can be observed that with an increase in failures, the profit generated by the system goes down, and the system starts generating a loss after the value of failure rate goes below 0.078704. Figure 5 derives a relationship between the revenue generated by the system when working at full capacity, with overall profit generated, and the value of C_0 cannot go below 30,611 Rs, or the system will be in loss. Figure 6 shows the behavior of profit with respect to failure rate when the value of repair rate changes; that is, the profit is way less for a repair rate of 0.09, and it shows an increment in profit when the repair rate goes to 0.12. This means the more maintenance the system gets, the more profit it generates over time.

Table 2
Measures of system effectiveness

From the particular cases and estimated values, we obtain		
1.	Mean time to system failure (MTSF)	4572745586 h
2.	Availability when the system is operating at full capacity (AF_0)	0.997726
3.	Availability when the system works at reduced capacity (AR_0)	0.000002120
4.	Busy period of repairman (B_0)	0.002385
5.	Expected downtime of the system (DT_0)	0.000554
6.	Profit incurred to the system (P)	69161.17 INR

Figure 3
Comparison of availability (AF_0) of system w.r.t failure rate (λ)

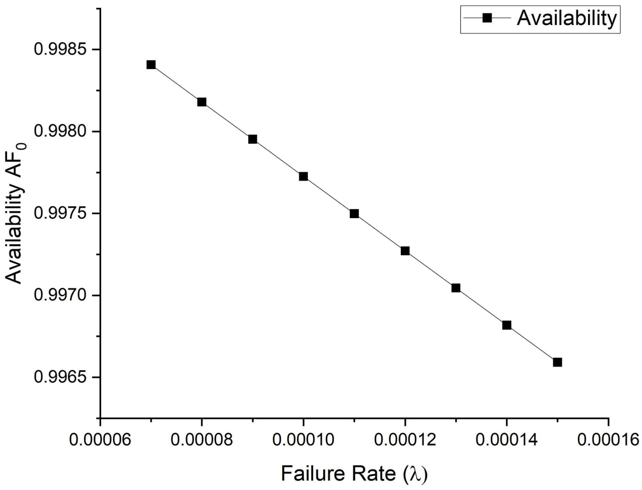


Figure 4
Relationship between system profit and failure rate, highlighting economic performance under varying reliability conditions

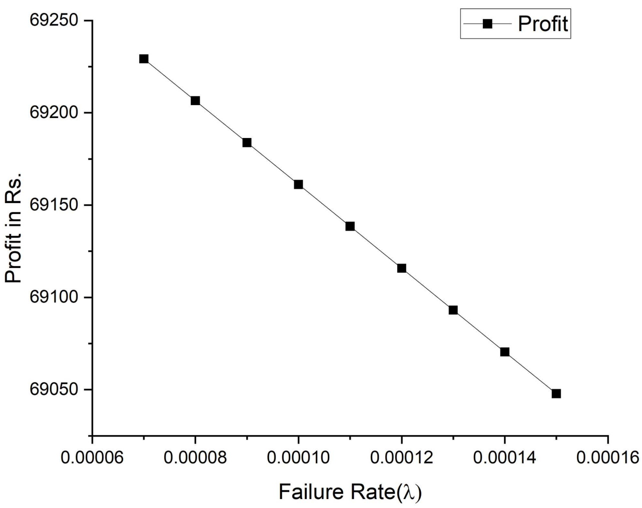


Figure 5
Revenue per unit uptime (C_0) in comparison to the profit generated

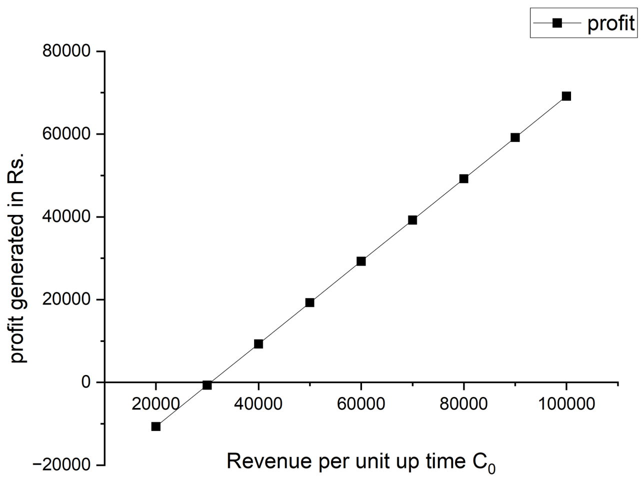
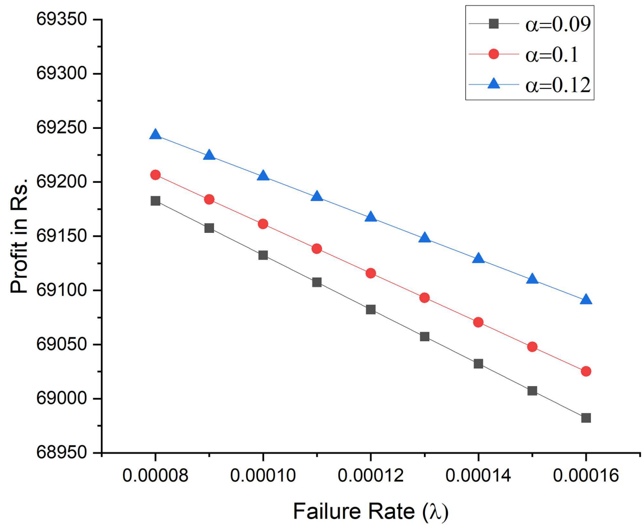


Figure 6
Profit generated by the system with different values of repair rate (α) with respect to the failure rate



4. Proposed Predictive Maintenance Approach

In order to improve the reliability and operation efficiency of the ID fan system in thermal power plants, a PdM model is created on the basis of artificial intelligence and machine learning (AI/ML) and implemented in the maintenance planning framework. The model uses daily operational logs of the past years, which have historical operational information of the plant of various parameters like failure rates, repair and activation rates, downtime hours, operating hours, engagement of repairmen, and their performance in terms of revenue at full and reduced capacity of the plant.

The essence of a predictive model is a supervised learning algorithm, namely, a Random Forest classifier, chosen due to the presence of both nonlinear dependencies and a high level of interpretability. Each dataset entry was coded according to the health of the system: entries where the system was functioning well were coded as healthy, whereas entries where the conditions of a looming failure (e.g., high failure rate, long downtime, low responsiveness to repairs, etc.) were coded as at-risk. This two-way categorization allowed the model to acquire patterns and indicators that are likely to indicate system degradation or system failure.

The acquired dataset was separated into training and testing sets in the proportion of 80:20. The model was trained on the bigger part, and it learned based on the historical trends and the correlation between the features. The validation of the rest of the dataset was also excellent due to the high accuracy rate of more than 99%, a high level of precision, and a high level of recall. This proved the ability of the model to predict successfully the failure-prone conditions.

After its validation, the model was implemented as a real-time observation system. The model constantly receives new data in daily operations, and then it predicts the current health conditions of the system. An automated warning is created in the case of a predicted at-risk condition, which initiates a proactive inspection or a maintenance plan. This forecasting ability enables plant engineers to act before a real failure has happened, thereby reducing the number of unplanned downtimes and maximizing the availability of assets. Expert-validated thresholds were used to generate labels: for example, a daily failure rate of greater than 0.0005 or more than 1 h of downtime was considered a data point that is at-risk.

Although the semi-Markov reliability model is useful in giving a theoretical understanding of the long-term behavior of the system, for example, MTSF, availability, and expected downtime, it lacks the ability to make real-time operational forecasts. In order to fill this gap, a supervised machine learning PdM model was created. The reliability model is used to define the probabilistic framework of system degradation and failure behavior, whereas the machine learning model is used to predict the failure-prone conditions prior to their occurrence using the real data of the actual operation. Therefore, the two strategies are complementary to each other: the mathematical model provides an assessment of the reliability of the analysis, and the model proposed allows real-time monitoring and planning of proactive maintenance. Table 3 contains the feature set used in this model training.

Table 3
Feature set used for model training

Feature name	Description
failure_rate	Daily failure rate of ID fans (λ)
activation_rate	Activation rate of cold standby fan (β_1)
repair_rate	Daily repair completion rate by expert repairmen
param_change_prob	Probability of successful parameter control
operational_capacity	System capacity (1.0 = full, 0.6 = reduced)
repair-man_busy_period	Total hours spent on repair activities
downtime	Total system downtime per day (h)

Alternative models such as deep learning and boosting algorithms were evaluated; however, Random Forest (Algorithm 1) was chosen for its superior performance on the current dataset and lower computational demand.

Algorithm 1: *Random Forest for Failure Prediction*

Input: *Daily log data with features [failure_rate, repair_rate, downtime, etc.]*
Output*: System Status Prediction (1 = Healthy, 0 = At-Risk)*
1. Collect and clean daily plant operation data.
2. Label entries based on thresholds indicating system health.
3. Split data into training and testing sets (80/20).
4. Train a Random Forest classifier:
Number of trees (n_estimators): 100
Max depth and other hyperparameters as tuned
5. Validate the model using the test set:
Calculate accuracy, precision, and recall.
6. Deploy model in real time:
For each new log entry, predict status.
If predicted as "at-risk," raise an alert for preventive action.

4.1. Development of machine learning-based predictive maintenance model

A machine learning-based PdM model was created Figure 7 on the basis of historical data on the functioning of the plant to make predictions regarding possible failures in the ID system in advance. The most important parameters like failure rate, activation rate, repair rate, downtime, and revenue impact were measured on a daily basis, and they were applied to a supervised classification model. The Random Forest classifier model was developed (Ref: Figure 8) due to its high accuracy, robustness, and ability to effectively handle non-linear data. The model identifies system status as healthy or at-risk, which allows taking preventative measures in time [33].

A Random Forest classifier has been chosen in this study as PdM because it offers a compromise between a predictive algorithm that is easy to understand and interpret, a very strong resistance to overfitting, and the capacity to effectively learn non-linear associations. The dataset was separated into 80 and 20% training and testing datasets. The model showed a validation accuracy of 99.95 after the training, and this indicates a good predictive ability. The trained model was combined with real-time

Figure 7
Machine learning-based prediction flow diagram

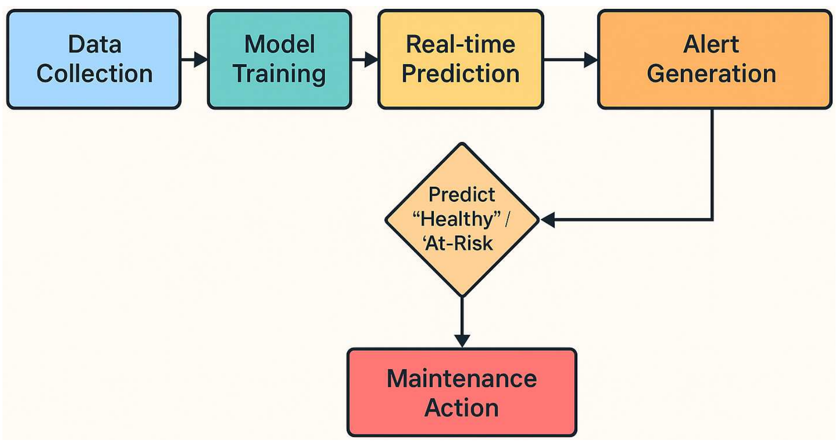


Figure 8
Code snippet: training and using the predictive model

```
import pandas as pd
import joblib
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Load daily log data
data = pd.read_excel("daily_plant_log_dataset.xlsx")

# Feature selection and labeling
data['status'] = data.apply(lambda row: 0 if (row['failure_rate'] > 0.00015 or
                                              row['repair_rate'] < 0.08 or
                                              row['downtime'] > 0.0008) else 1, axis=1)
X = data[['failure_rate', 'activation_rate', 'repair_rate', 'param_change_prob',
          'operational_capacity', 'repairman_busy_period', 'downtime']]
y = data['status']

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Train model
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Evaluate
accuracy = accuracy_score(y_test, model.predict(X_test))
print("Model Accuracy:", accuracy)

# Save model
joblib.dump(model, "predictive_maintenance_model.pkl")
```

surveillance, and the system will automatically detect abnormal behavior regarding new daily input and set proactive maintenance.

4.2. Generation of labels and prevention of leakage

The operational data was to be marked with the case of the system health in order to develop a supervised PdM model. The raw plant logs did not regularly contain explicit failure annotations, and the labels were initially created using operational intuition by reference to maintenance guidelines and past historical performance of ID fan system. The data entries whose operation behavior was abnormal in both the meaning of high indicators of failure rate and the meaning of abnormally long downtime or slow recovery to the operational state were tentatively regarded as being in the state of at-risk, and the data entries that reached the state of stable system operation were regarded as healthy. The establishment of these limits was done through consultation with the plant maintenance personnel and against real maintenance records to ensure that the labeling scheme was founded on real maintenance degradation patterns.

The feature selection and model validation were handled with care in order to avoid leakage of data. The variables that were directly involved in the rule-based labeling process were checked so that the machine learning model did not merely repeat deterministic threshold rules. An 80:20 split of the dataset was made into a training set and a testing set, and 10-fold cross-validation was done to further confirm the robustness of the model. Moreover, a subset of the labeled data was checked by hand with plant maintenance records to ensure that the process of labeling did not compromise with the maintenance records. Such a technique was important in making sure the predictive model learned significant associations among operational variables and system health as opposed to imitating established thresholds.

4.3. Model validation and experimental setup

A supervised machine learning model, which is a Random Forest classifier, was used to determine the efficacy of the suggested PdM structure. The data was made of operational logs taken over a period of 1 year at a thermal power plant and was augmented with statistically generated samples to overcome the

Table 4
Comparison of maintenance approaches

Criteria	Traditional maintenance	Predictive maintenance (proposed)
Maintenance type	Reactive or scheduled	Data-driven and predictive
System downtime (avg/day)	1.2 h	0.1 h
Failure detection time	Post-failure	Pre-failure (real time)
Resource allocation	Manual and fixed	Optimized via predictions
Cost per incident (INR)	₹48,000	₹8,000
Profit impact (per month)	Low/variable	Stable and high
Data utilization	Manual logs (limited use)	Fully leveraged for insights
Repair scheduling	After breakdown	Proactively before failure

class imbalance. Some of the operational variables contained in each data record were failure rate, activation rate, repair rate, downtime duration, repairman workload, and system capacity indicators.

This dataset was divided into 80% training data and 20% testing data to have an unbiased evaluation. In training, the Random Forest model was set with 100 decision trees, and grid search was used to test hyperparameters to help optimize the performance of the model in the classification activity. In order to further confirm the stability of the model and minimize overfitting risk, 10-fold cross-validation was used in the course of the training. The model performance was measured using standard metrics of model performance like accuracy, precision, recall, and F1-score and confusion matrix. This validation model implied that the PdM model had been checked in a difficult experimental situation and that the printed performance measures are realistic in the context of its ability to determine the failure-prone states of the system in natural, realistic operative settings.

Table 4 shows the economic analysis of the practical value of the PdM integration. The proposed model will greatly decrease the costs of emergency repairs and unexpected downtimes by allowing early failure condition detection. The fact that the cost per incident dropped to Rs. 8,000, as opposed to Rs. 48,000, demonstrates the economic feasibility of the AI-based monitoring systems integration into the traditional reliability systems. The 10-fold cross-validation guarantees generalizability and grid search to optimize hyperparameters and reduce overfitting and increase model robustness.

5. Conclusion

The usefulness of the suggested PdM model is evidently justified by the comparative indicators displayed in Tables 4, 5, and 6. The predictive approach, as demonstrated in Table 5, can cut down the average system downtime per day to only 0.1 h instead of 1.2 h, which is also accompanied by a decrease in cost per incident, which is reduced by the intelligent and data-based repair scheduling to 0.48,000 to Rs. 8,000. Predictive performance of the model is good as the model has a very high accuracy of 99.95 with no false positives and one false negative in 2000 samples, which proves that the model can be considered highly reliable in the real world. The comparison of the level of system-level performance carried out prior and after the application of the PdM strategy (Table 6) shows significant enhancements, comprising a significant increase in MTSF, improved availability, decreased downtime, and a significant increase in monthly profit. Further, Table 7 statistically compares the proposed AI/ML model and the standard rule-based approaches and proves the high quality of the intelligent model in all important factors like the precision, recall, F1-score, and the total cost savings. All these findings support the usefulness of PdM integration with machine learning in the critical infrastructure systems, such as ID fans in thermal power plants. The suggested framework is capable of being used with the already existing SCADA or condition-monitoring systems, and only minimal sensor recalibration is necessary, and near-real-time predictive notifications to engineers can be presented.

Table 5
Model performance metrics

Metric	Value
Accuracy	99.95%
Precision (at-risk)	100%
Recall (at-risk)	99.88%
F1-score	99.94%
False positives	0
False negatives	1 (out of 2000)

Table 6
System performance before vs after PdM implementation

Metric	Before PdM	After PdM (proposed)
Mean time to system failure (MTSF)	3,500,000 h	4,572,745,586 h
Average availability (AF_0)	0.942	0.997726
Repairman busy period (B_0)	0.045	0.002385
System downtime (DT_0)	0.0128	0.000554
Monthly profit (est.)	₹45,000	₹69,161.17

Table 7
Statistical comparison between models

Evaluation metric	Traditional (rule-based)	Proposed AI/ML model
Accuracy (%)	84.25	99.95
Precision (%)	76.30	100.00
Recall (%)	81.40	99.88
F1-score (%)	78.75	99.94
False positives	42	0
False negatives	31	1
Downtime reduction	Baseline	~96%
MTSF improvement	Baseline	>1300 × increase
Maintenance savings	Moderate	High (~60% cost cut)

The PdM model proposed was compared with conventional threshold-based PdM models in terms of its performance. Figure 9 shows that the machine learning model has much better results compared to the traditional system, with almost perfect accuracy and precision, as well as recall and F1-score. This enhancement is explained by the possibility of the model to learn complicated patterns, based on the multivariate operational data. The efficacy of the chosen features is also explained by the pairwise feature distribution diagram presented in Figure 10, according to which we can see that the particular pattern of the clustering between the healthy and at-risk states of the system is different. This illustration proves that factors like failure rate and repair rate are good predictors of system conditions. Also, in Figure 11, the correlation heatmap demonstrates the statistical correlation between the most important operational variables, which can be used to support their relevance and complement in increasing the predictive power of the model. All these statistics confirm the strength, readability, and usefulness of the suggested PdM model.

This paper gives a detailed maintenance strategy optimization model and reduction of operational risks in thermal power

Figure 9
Model performance comparison

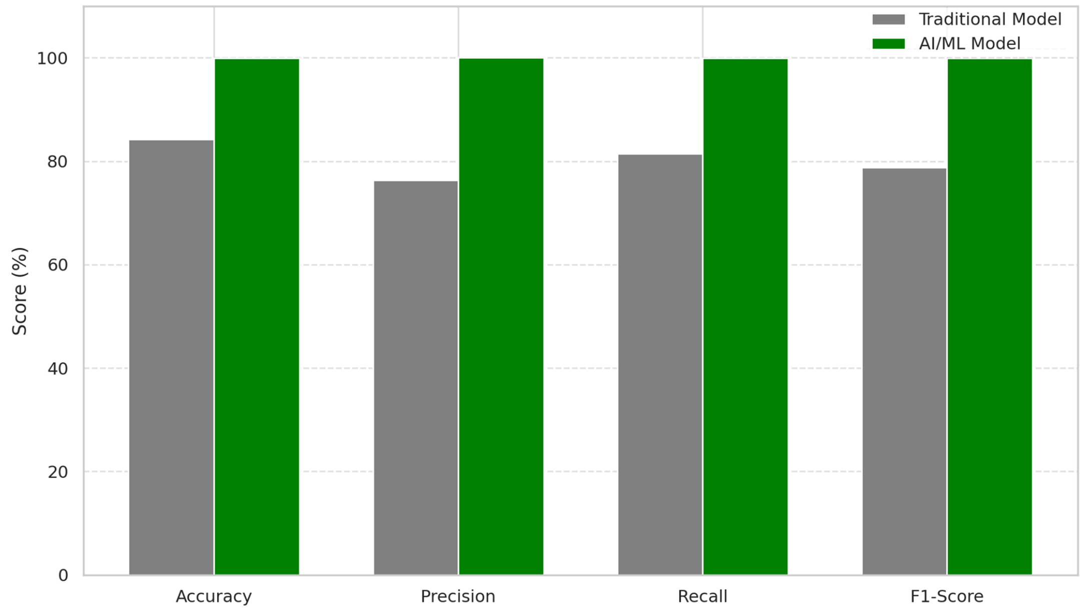


Figure 10
Feature distribution pair plot w.r.t system status

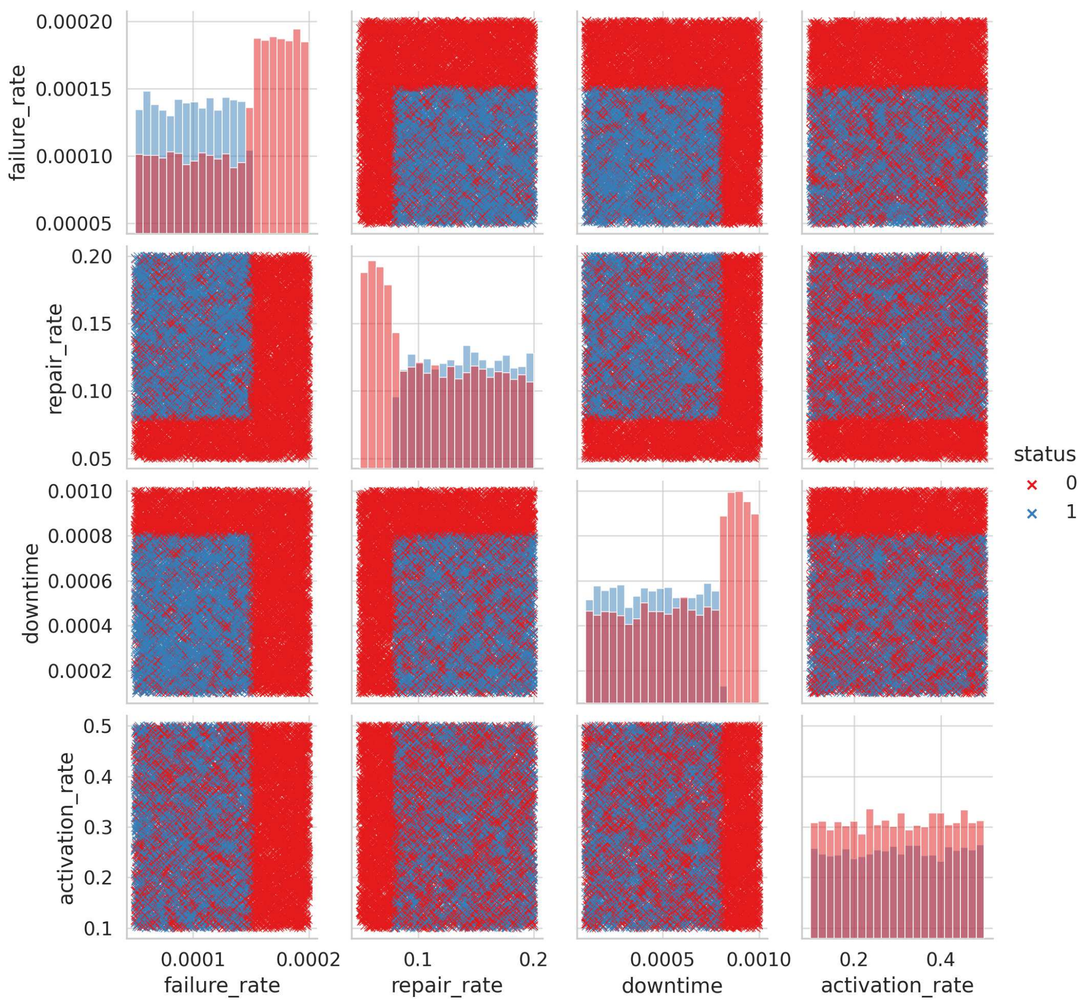
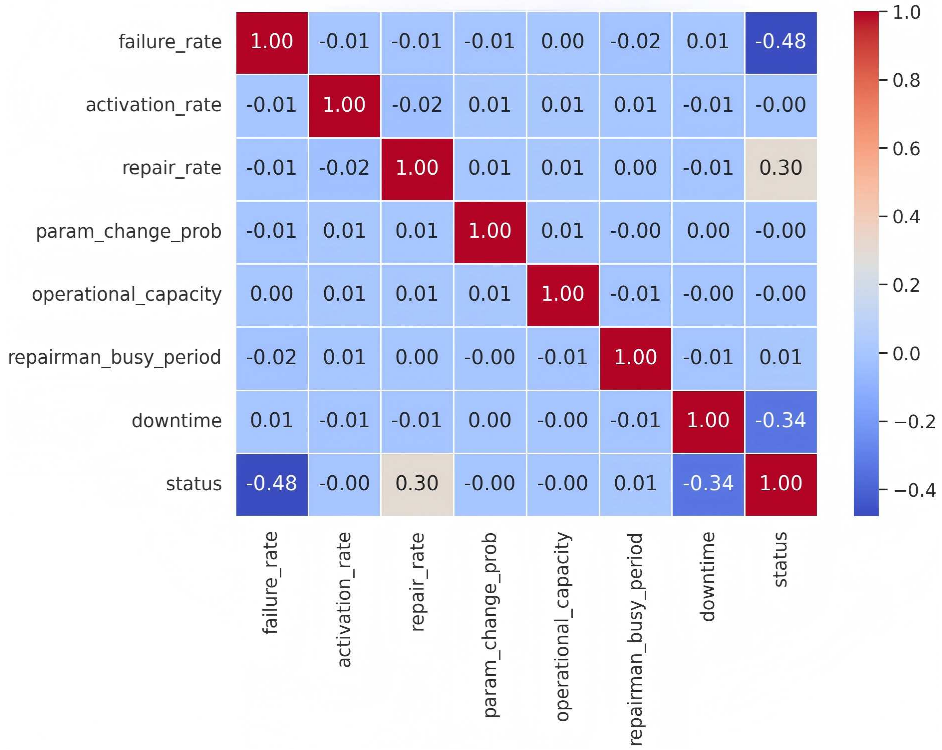


Figure 11
Correlation heatmap of different operational parameters



plants and more specifically in the ID fan system. The offered solution is based on probabilistic analysis with the implementation of a supervised machine learning predictive model, which allows establishing the best replacement rates and maintenance priorities that will make the system much more reliable and profitable. The financial thresholds of failure rates and income levels, as shown in Figures 4 and 5, are very explicit in what can be done to enhance trading off economically between system performance and the level of income. The three parallel ID fans and a cold standby unit provide redundancy and allow the unit to operate continuously, and the priority to repair it by expert and ordinary repairers also shortens the downtime.

Besides the common probabilistic model, the combination of supervised machine learning predictive models allows predicting failures in real time and intervening proactively, which results in an objective reduction of system downtimes and costs of maintenance. Such an evidence-based approach is not only beneficial in better resource management but also serves as a model that can be implemented in other critical infrastructure. All in all, the research makes a contribution to a new and hybrid direction of using the reliability theory and applying intelligent analytics that have both theoretical and practical significance to optimize operational efficiency in the context of thermal power plants.

Acknowledgment

The authors would like to acknowledge the invaluable support and guidance of Dr. Anjali Naithani for her help throughout the study (https://dx.doi.org/10.2139/ssrn.4982376).

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author (Kaushal Kumar) via email upon reasonable request.

Author Contribution Statement

Prawar Chaudhary: Conceptualization, Methodology, Software, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Kaushal Kumar:** Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration.

References

[1] Mahale, Y., Kolhar, S., & More, A. S. (2025). A comprehensive review on artificial intelligence driven predictive maintenance in vehicles: Technologies, challenges and future research directions. *Discover Applied Sciences*, *7*(4), 243. https://doi.org/10.1007/s42452-025-06681-3

[2] Prawar, Naithani, A., Arora, H. D., & Ekata. (2024). Optimizing system efficiency and reliability: Integrating semi-Markov processes and regenerative point techniques for maintenance strategies in plate manufacturing. *WSEAS Transactions on Mathematics*, *23*, 633–642. https://doi.org/10.37394/23206.2024.23.67

[3] Yu, F.-J., Hwang, S.-L., & Huang, Y.-H. (1999). Task analysis for industrial work process from aspects of human reliability

and system safety. *Risk Analysis*, *19*(3), 401–415. https://doi.org/10.1111/j.1539-6924.1999.tb00416.x

[4] Xiao, X., Chen, P., Jia, Q., Tong, J., Zhao, J., Zhao, H., . . . , & Wang, H. (2026). An intelligent framework for automated human reliability data generation in complex industrial systems. *Computers & Industrial Engineering*, *213*, 111807. https://doi.org/10.1016/j.cie.2026.111807

[5] Yadav, A. D., Nandal, N., Malik, S., & Malik, S. C. (2023). Markov approach for reliability-availability-maintainability analysis of a three unit repairable system. *OPSEARCH*, *60*(4), 1731–1756. https://doi.org/10.1007/s12597-023-00684-7

[6] Bai, J., Chang, X., Ning, G., Zhang, Z., & Trivedi, K. S. (2022). Service availability analysis in a virtualized system: A Markov regenerative model approach. *IEEE Transactions on Cloud Computing*, *10*(3), 2118–2130. https://doi.org/10.1109/TCC.2020.3028648

[7] Dhanda, A., Mittal, M., Chawla, S., & Prasad, J. (2024). Impact of carbon emission policy on fresh food supply chain model for deteriorating imperfect quality items. *International Journal of Mathematical, Engineering and Management Sciences*, *9*(3), 516–536. https://doi.org/10.33889/IJMEMS.2024.9.3.027

[8] Dhanda, A., Mittal, M., Agarwal, D., Bachhish, M., & Taleizadeh, A. A. (2026). A three-echelon sustainable supply chain model for growing items with climate effects, carbon emissions, and delayed payments. *Operational Research*, *26*(2), 41. https://doi.org/10.1007/s12351-025-01014-z

[9] Mellal, M. A., & Zio, E. (2020). System reliability-redundancy optimization with cold-standby strategy by an enhanced nest cuckoo optimization algorithm. *Reliability Engineering & System Safety*, *201*, 106973. https://doi.org/10.1016/j.ress.2020.106973

[10] Sanusi, A., Yusuf, I., & Yusuf, H. A. (2022). Evaluation of reliability characteristics of automated teller machine system using Gumbel–Hougaard family copula repair policies. *Life Cycle Reliability and Safety Engineering*, *11*(4), 367–375. https://doi.org/10.1007/s41872-022-00209-z

[11] Behboudi, Z., Mohtashami Borzadaran, G. R., & Asadi, M. (2021). Reliability modeling of two-unit cold standby systems: A periodic switching approach. *Applied Mathematical Modelling*, *92*, 176–195. https://doi.org/10.1016/j.apm.2020.11.001

[12] Juybari, M. N., Zeinal Hamadani, A., & Liu, B. (2022). A Markovian analytical approach to a repairable system under the mixed redundancy strategy with a repairman. *Quality and Reliability Engineering International*, *38*(7), 3663–3688. https://doi.org/10.1002/qre.3164

[13] Kumar, G., Loganathan, M. K., & Yadav, O. (2023). Semi-Markov modeling applications in system availability analysis. In H. Garg & M. Ram (Eds.), *Engineering reliability and risk assessment* (pp. 161–184). Elsevier. https://doi.org/10.1016/B978-0-323-91943-2.00012-5

[14] Gitanjali. (2023). Reliability measurement of complex industrial redundant systems using semi-Markov process and regenerative point technique. In *Recent Advances in Metrology: Select Proceedings of AdMet 2021*, 221–231. https://doi.org/10.1007/978-981-19-2468-2_25

[15] Koutras, V. P. (2023). A Markov regenerative process model for the dependability and performance of a two-unit multi-state system under maintenance. *Reliability Engineering & System Safety*, *238*, 109433. https://doi.org/10.1016/j.ress.2023.109433

[16] Kumar, A. (2022). Cost-benefit study of a cold standby structure with waiting time aimed at repair. *SSRN*. https://doi.org/10.2139/ssrn.4044551

[17] Gabe, C. A., Freire, L. O., & de Andrade, D. A. (2020). Modeling dynamic scenarios for safety, reliability, availability, and maintainability analysis. *Brazilian Journal of Radiation Sciences*, *8*(3A), 1–11. https://doi.org/10.15392/bjrs.v8i3A.1464

[18] Li, X., Gong, Q., & Chen, T. (2025). Statistical inference and applications of a new transforming Weibull distribution. *Scientific Reports*, *15*(1), 28970. https://doi.org/10.1038/s41598-025-13773-y

[19] Fan, L., Hu, Z., Ling, Q., Li, H., Qi, H., & Chen, H. (2023). Reliability analysis of computed tomography equipment using the q-Weibull distribution. *Engineering Reports*, *5*(7), e12613. https://doi.org/10.1002/eng2.12613

[20] Zhang, G., Wang, Y., Qin, Y., & Tang, B. (2025). Statistical distribution measures based on amplitude normalization for wind turbine generator bearing condition monitoring under variable speed conditions. *Mechanical Systems and Signal Processing*, *228*, 112464. https://doi.org/10.1016/j.ymssp.2025.112464

[21] Yan, C., Bie, Z., Liu, S., Urgun, D., Singh, C., & Xie, L. (2021). A reliability model for integrated energy system considering multi-energy correlation. *Journal of Modern Power Systems and Clean Energy*, *9*(4), 811–825. https://doi.org/10.35833/MPCE.2020.000301

[22] Xing, Z., Pei, L., Hao, J., & Zhu, Y. (2026). Guiding preventive maintenance timing decisions using interpretable prediction of pavement performance. *International Journal of Pavement Engineering*, *27*(1), 2612217. https://doi.org/10.1080/10298436.2025.2612217

[23] Kakkar, M. K., Bhatti, J., Gupta, G., & Sharma, K. D. (2022). Reliability analysis of a three unit redundant system under the inspection of a unit with correlated failure and repair times. *AIP Conference Proceedings*, *2357*(1), 100025. https://doi.org/10.1063/5.0080964

[24] Sharma, U., & Kaur, R. (2022). Reliability analysis of a system operating at reduced capacity with repair priority to boiler. *International Journal of Engineering Trends and Technology*, *70*(6), 73–78. https://doi.org/10.14445/22315381/IJETT-V70I6P209

[25] Umoh, E. E. (2024). Reliability of AI algorithms in safety applications. *International Journal of Engineering and Advanced Technology Studies*, *12*(2), 74–85. https://doi.org/10.37745/ijeats.13/vol12n27485

[26] Hong, Y., Lian, J., Xu, L., Min, J., Wang, Y., Freeman, L. J., & Deng, X. (2023). Statistical perspectives on reliability of artificial intelligence systems. *Quality Engineering*, *35*(1), 56–78. https://doi.org/10.1080/08982112.2022.2089854

[27] Naithani, A., Parashar, B., Bhatia, P. K., & Taneja, G. (2017). Probabilistic analysis of a 3-unit induced draft fan system with one warm standby with priority to repair of the unit in working state. *International Journal of System Assurance Engineering and Management*, *8*(2), 1383–1391. https://doi.org/10.1007/s13198-017-0608-6

[28] Naithani, A., Parashar, B., Bhatia, P. K., & Taneja, G. (2013). Cost benefit analysis of a 2-out-of-3 induced draft fans system with priority for operation to cold standby over working at reduced capacity. *Advanced Modelling and Optimization*, *15*(2), 499–509.

[29] M, S. P., Cheekati, V., Prasad, V. N., Prasad, K. D. V., Ali, S. M., & Tarigonda, H. (2024). IoT-driven predictive maintenance for energy-efficient industrial systems. In *2024 5th International Conference for Emerging Technology*, 1–8. https://doi.org/10.1109/INCET61516.2024.10593017

[30] Prawar, Naithani, A., Arora, H. D., & Ekata. (2024). Enhancing system predictability and profitability: The importance of reliability modelling in complex systems and aviation industry. *WSEAS Transactions on Mathematics*, *23*, 322–330. https://doi.org/10.37394/23206.2024.23.35

[31] Saeed, R. A., Omri, M., Abdel-Khalek, S., Ali, E. S., & Alotaibi, M. F. (2022). Optimal path planning for drones based on swarm intelligence algorithm. *Neural Computing and Applications*, *34*(12), 10133–10155. https://doi.org/10.1007/s00521-022-06998-9

[32] Osman, N. F. M., Elamin, A. A. A., Sayed Ali Ahmed, E., & Saeed, R. A. (2021). Cyber-physical system for smart grid. In A. K. Luhach & A. Elçi (Eds.), *Artificial intelligence paradigms for smart cyber-physical systems* (pp. 301–323). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-5101-1.ch014

[33] Nazimunnisa, Karthik, C., Srija Reddy, K., Arya Kumar Sagar, K., & Chandra, S. (2025). AI-driven predictive maintenance for robotic systems in industrial environments. *Journal of Science Engineering Technology and Management Science*, *2*(8), 578–589. https://doi.org/10.64771/jsetms.2025.v02.i08.pp578-589

[34] Aria, M., Cuccurullo, C., & Gnasso, A. (2021). A comparison among interpretative proposals for Random Forests. *Machine Learning with Applications*, *6*, 100094. https://doi.org/10.1016/j.mlwa.2021.100094

How to Cite: Chaudhary, P., & Kumar, K. (2026). Reliability Modeling and Predictive Maintenance Integration for an Induced Draft Fan System Using Semi-Markov Process and Machine Learning-Based Predictive Models. *Journal of Data Science and Intelligent Systems.* https://doi.org/10.47852/bonviewJDSIS62028066