IET Collaborative Intelligent Manufacturing

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

A comprehensive framework from real-time prognostics to maintenance decisions

Amit Kumar Jain ⬤ | Maharshi Dhada | Marco Perez Hernandez |
Manuel Herrera ⬤ | Ajith Kumar Parlikad ⬤

Department of Engineering, Institute for
Manufacturing, University of Cambridge, Cambridge,
UK

Correspondence

Amit Kumar Jain, Department of Engineering,
Institute for Manufacturing (IfM), University of
Cambridge, Alan Reece Building, 17 Charles
Babbage Rd, Cambridge CB3 0FS, UK.
Email: amitkjain23@gmail.com

Abstract

Studying the influence of imperfect prognostics information on maintenance decisions is
an underexplored area. To bridge this gap, a new comprehensive maintenance support
system is proposed. First, a survival theory-based prognostics module employing the
Weibull time-to-event recurrent neural network was deployed in which prognostics
competence was enhanced by predicting the parameters of failure distribution. In
conjunction with this, a new predictive maintenance (PdM) planning model was framed
via a trade-off between corrective maintenance and time lost due to PdM. This optimises
maintenance time based on operational and maintenance cost parameters from the his-
torical data. The performance of the proposed framework is demonstrated using an
experimental case study on maintenance planning for cutting tools within a
manufacturing facility. Systematic sensitivity analysis is provided, and the impact of
imperfect prognostics information on maintenance decisions is discussed. Results show
that uncertainty about prediction declines as time goes on, and as uncertainty declines, the
maintenance timing becomes closer to the remaining useful life. This is expected, as the
risk of making a wrong decision decreases over time.

1 | INTRODUCTION

Much of the maintenance undertaken today is either
corrective (replacing an asset after it fails) or preventive
(assuming a certain level of degradation with no input from
the asset itself, and maintaining the asset on a fixed schedule
regardless of whether the asset's condition would require it).
Both strategies are exceedingly inefficient. Thus, predictive
maintenance has attracted substantial research focus over the
last decade [1]. Predictive maintenance involves forecasting
the asset's remaining useful life (RUL) to plan maintenance
activities. For instance, a novel integrated diagnostics and
prognostics system using a support vector machine was
presented in [2, 3] technologically advanced a statistical
approach for prognostics using the estimated linear model.
[4] developed three data-driven prognostics models, namely,
offline, online and semi-offline, for different industrial cir-
cumstances, namely, online and offline condition monitoring.
[5] presented a new product quality-based failure-prediction
module for the manufacturing environment. [6] proposed a
multi-agent system framework for online prognostics. The
same was implemented for collaborative prognostics in [7].
[8] also exhibited accurate life predictions through the
stepwise regression feature subset selection technique. [9]
designed a prognostics approach under unknown initial wear.
[10] proposed a new particle-filtering-based method to pro-
vide a precise time-to-failure estimate. [11] circumvented
losing potentially valuable information about an asset's
degradation process via extended capsule neural networks for
fault prognostics, and in particular, RUL estimation for multi-
dimensional sensor data. In [12], probabilistic estimation of
RUL by three data-driven prognostic methodologies is pre-
sented based on state-of-the-art as well as innovative math-
ematical models, which are gradient-boosted trees, Bayesian
neural networks and non-homogeneous hidden semi-Markov
models. In [13–15], many approaches to RUL assessment
using physics-based and data-driven methods are reviewed. It
is observed that available approaches spotlight the prognos-
tics step and do not consider maintenance decisions, which
are addressed independently.

For instance, [16] proposed a maintenance optimisation model not considering prognostics information that resulted in maintenance shortage or overage. [17] shared out an optimisation model with time-to-failure substance to a Weibull probability distribution. [18] associated maintenance planning for a repairable system. [19] gave a methodology for predictive maintenance that solely considered the diagnostics information. [20] presented a cost-oriented predictive maintenance strategy based on reliability state. [21] analysed the cost implications of various multi-agent system architectures but did not consider the effect of imperfect prognostics on maintenance decisions. [22, 23] dealt with post-prognostics implications but under the assumption that the prognostics data of the asset were available at all times. Nonetheless, neither group proposed an extensive system (from real-time prognostics to maintenance decisions) that researched the effect of imperfect prognostics on maintenance decisions. Consequently, a novel comprehensive maintenance support system is proposed herein.

First, a survival theory-based real-time prognostics module employing the Weibull time-to-event recurrent neural network (WTTE-RNN) was deployed. It is innovatively modelled to project a detailed picture of asset reliability by predicting the asset's probability of failure distribution in contrast to predicting time to failure. This is coupled with a novel predictive maintenance (PdM) planning model that resourcefully balances corrective maintenance (CM) costs and PdM costs to facilitate the determination of the optimal maintenance time that minimises overall system maintenance costs. In addition, we analysed how imperfect prognostic information influences maintenance decisions. Critical insights are underlined, namely, uncertainty about prediction drops as time goes on; as uncertainty drops, maintenance timing becomes closer to RUL because the risk of making a wrong decision is decreasing. Based on this analysis, guidelines are offered for managers to help them improve their chances of making the right maintenance decisions. This study acts as a proof of concept, showing the importance of utilising prognostics information in maintenance planning.

The novel contribution of this paper is in the conceptualisation of a comprehensive maintenance support system satisfying the vital necessities: (a) a real-time prognostics approach that can be extensively realised for several systems; (b) a flexible maintenance-decision model that rapidly evaluates different operational and maintenance costs; (c) consideration of the implications of imperfect prognostics on maintenance decisions to find the right moment for performing maintenance activities. Having such a comprehensive system, a supervisor can design maintenance exercises that are all the more viable for diminishing machine downtimes and improving the production stream. The added contribution lies in the results. The performance of the framework is proven via a case study from a manufacturing environment to distinguish suitability, quality, reliability, robustness, and applicability in a real-world industrial environment. Lastly, the work is complimented with a systematic sensitivity analysis.

The rest of the paper is structured as follows. Section 2 provides details of the methodology involved in developing the comprehensive maintenance support system. Section 3 gives the specifics of the real-life case study and implementation results. Lastly, Section 4 concludes and highlights the key contributions.

2 | METHODOLOGY

2.1 | Prognostics module

In the recent technical literature, a large variety of prognostic applications that estimate time to failure have been reported [13, 14, 24]. For the experiments discussed here, a novel real-time prognostics algorithm, WTTE-RNN, was deployed. WTTE-RNN combines survival theory with recurrent neural networks (RNNs) to model an asset's failure probability distribution. WTTE-RNN assumes the distribution to be Weibull in nature, which is a standard assumption for reliability studies [25]. Inputs to WTTE-RNN are multivariate vectors comprising real-time sensor data, and output are the scale and shape parameters of the Weibull distribution. RNNs have been recommended for time-series predictions such as prognostics on several occasions [26]. Added benefits of WTTE-RNN are the capabilities to analyse both censored and uncensored data and to predict failure distributions rather than a single RUL value. Failure distributions are more informative than single-valued RUL predictions because they model a detailed picture of an asset's future health and therefore enable operators to plan risk-based maintenance activities.

For fulfilment, the relevant mathematics is described here. Because we only use WTTE-RNN for analysing censored data, the mathematics described here corresponds to that part of the WTTE-RNN description only. The reader is advised to refer to [27] for more information. WTTE-RNN relies on a log-likelihood loss function to be optimised by the RNN to predict the shape parameter (θ) and scale parameter (η) of a Weibull probability distribution. The prognostics module analyses a vector of sensor values as an input, representing the asset's current health condition at a given time, and based on this, the RNN estimates (θ, η). The log-likelihood loss function to be maximised by the RNN is

$$\log(L) = \sum_{n=1}^{N} \sum_{t=0}^{T_n} \log \left[\Pr \left(Y_t^n = y_t^n \middle| x_{o:t}^n \right) \right] \qquad (1)$$

The prognostics module tries to maximise the probability of estimated time to failure (Y_t^n) being equivalent to actual time to failure (y_t^n) for an available vector of sensor features (x). The summations ($\sum_{n=1}^{N} \sum_{t=0}^{T_n} \cdot$) are made over every trajectory (N) and over every time step for every trajectory (T_n). The probabilities appearing in Equation (1) are obtained via survival analysis.

For predicting discrete events, the probabilities can be shown as

$$\log(L) = u.\log\left(e^{d(t)} - 1\right) - \Lambda(t + 1)$$

where $\Lambda(t)$ is the cumulative hazard function and $d(t) = \Lambda(t+1) - \Lambda(t)$ is the step cumulative hazard function. For a Weibull distribution for the probability of occurrence of an event with respect to time, the cumulative hazard function is

$$\Lambda(t) = \left(\frac{t}{\eta}\right)^{\theta}$$

The loss function for a Weibull distribution of events therefore becomes

$$log(L_d) = \sum_{n=1}^{N} \sum_{t=0}^{T_n} \left(log\left\{ exp\left[\left(\frac{y_t^n + 1}{\eta_t^n}\right)^{\theta_t^n} - \left(\frac{y_t^n}{\eta_t^n}\right)^{\theta_t^n} \right] - 1 \right\} - \left(\frac{y_t^n + 1}{\eta_t^n}\right)^{\theta_t^n} \right) \tag{2}$$

where η_t^n and θ_t^n are the scale and shape parameters of the Weibull distribution, respectively, and y_t^n is the time to failure at every time step t and trajectory n. In summary, the RNN attempts to find the weights that maximise the $log(L_d)$ described in Equation (2). The output (failure probability distribution) of the prognostics module is coupled with the PdM planning model.

For many assets, data-driven prognostics models overcome the limitations of reliability-based prognostics models. In our method, we have combined the best of both techniques by using a data-driven model to optimise a reliability-theory-based loss function. Concretely, the Weibull loss function (see Equation 2) is inspired by reliability theory. It enables flexible modelling of an asset's probability of failure over its lifetime. However, the Weibull loss function's parameters are estimated using an RNN, which is a commonly used data-driven prognostics technique [28] wherein the RNN learns from the past history of failures and generates the parameters corresponding to the least overall loss.

2.2 | Predictive maintenance planning model

We consider an industrial facility consisting of a single asset system with a time-to-failure complying Weibull distribution. In this case, failure is viewed as asset degradation (F_{AD}) because of wear and tear. It is assumed that whenever a failure is observed, corrective replacement (CR) is carried out leading to a CR cost. The unexpected failure of the asset due to degradation can increase risks and safety hazards. Accordingly, predictive replacement (PdR) of the asset is actioned to bring down the probability of asset failure and reduce the risk of an unexpected failure. However, PdR requires additional time and funds. Therefore, PdR optimisation is executed to trade off the failure and PdR cost. To exhibit the benefits of PdM, a cost model is developed by capturing the various costs pertaining to

the industrial operation that are governed via failures and PdM planning. The economic objective is to minimise the expected total cost per unit of time to conduct predictive maintenance ($[ETC]_{(PdM)}$) by choosing the optimal time for PdR (O_{PdR}). Here, $[ETC]_{(PdM)}$ is the proportion of the addition of the expected total cost of CR due to asset degradation ($E[(C_{CR})_{F_{AD}}]$) and PdR ($E[C_{PdR}]$) to the planning period/evaluation time (E_T) for which the analysis is performed. It is written as follows:

$$[ETC]_{(PdM)} = \left(E\big[(C_{CR})_{F_{AD}}\big] + E[C_{PdR}] \right)/E_T \tag{3}$$

Theoretic and numerical models of constitutional costs in $[ETC]_{(PdM)}$ are detailed in the following subsections.

2.2.1 | Corrective replacement cost

Assume the system is stopped during the replacement, and take $(C_{CR})_{F_{AD}}$ as the cost of CR due to asset degradation including the downtime cost. Consequently, the expected cost of CR owing to asset degradation ($E[(C_{CR})_{F_{AD}}]$) is determined as

$$E\big[(C_{CR})_{F_{AD}}\big] = \left\{ A_{CR} \times \left[P_r \times C_{lp} + C_L \right] + C_{FCR} \right\} \times F(E_T)_{\theta,\eta} \tag{4}$$

where $A_{CR} \times [P_r \times C_{lp} + C_L]$ is the downtime cost owing to CR, A_{CR} is mean time to perform the corrective replacement (hours), P_r is the production rate (products/hours), C_{lp} is the cost of lost production (GBP), C_L is the cost of the labour (GBP/hours), C_{FCR} is the fixed cost of corrective replacement (including the cost of asset replacement) and $F(E_T)_{\theta,\eta}$ is the cumulative probability of failure owing to asset degradation for a given evaluation time as a function of a given shape (θ) and scale (η) parameter.

2.2.2 | Predictive replacement cost

In general, the cost of PdR for an asset is modelled to include the downtime cost with the replacement, labour, and asset costs. The replaced asset always has some useful remaining life that is usually not considered in PdR cost [29]. Comprehensive models that can consider the effect of lost remaining life on overall PdR cost will become increasingly critical. Therefore, in our model, the effect of an asset's lost remaining life is modelled in the PdR cost. This prompts the optimum utilisation of asset life. RUL is the residual life of the asset after a certain time. Here, the proposed model captures the real-time RUL information of the asset with the help of the failure probability distribution acquired as an output from the prognostics module. Moreover, we take the cost of lost remaining life ($CLRUL_i$) relative to mean life cost. It is assumed that asset cost is uniformly distributed over the lifetime of the asset [30]; the $CLRUL_i$ is given as

$$CLRUL_i = (C_A/A_L) \times \eta_i \Gamma\left(1 + \frac{1}{\theta_i}\right) \qquad (5)$$

where C_A is the cost of the asset (GBP) and A_L is the mean life of the asset (hours).

The function $\eta_i \Gamma\left(1 + \frac{1}{\theta_i}\right)$ gives the RUL of the asset in hours at a given point of time. Therefore, the total cost per PdR is given as

$$E[C_{PdR}] = \{A_{PdR} \times [P_r \times C_{lp} + C_L] + C_{FPdR} + CLRUL_i\} \qquad (6)$$

where $A_{PdR} \times [P_r \times C_{lp} + C_L]$ is the downtime cost owing to PdR, A_{PdR} is mean time to perform PdR (hours), C_{FPdR} is the fixed cost of PdR (GBP) (including the cost of asset replacement).

For assessing the optimal time for PdR, (O_{PdR}), a balance is created between the cost of the lost remaining life of the asset and maintenance and failure costs. The addition of both the costs, $[ETC]_{(PdM)}$, is calculated for each time step of the operation to be made by the asset, and corresponding to the minimum cost, the optimal PdR cost, along with the optimal time for PdR, is obtained.

3 | EXPERIMENTAL CASE STUDY

In this section, the proposed methodology is verified on the cutting tool degradation data set from reliability and prognostics repository provided by Industrial and Systems Engineering, IIT Indore, India. This data set is generated for prognostics and health management studies [31]. The data are provided by a testing platform furnished with a CNC milling machine and sensors viz.dynamometers, etc. The primary objective is to provide real-life historical data at different operating conditions for a population of identical cutting tools with cutting force sensor data that characterise the degradation of tools along with their entire operational life. In the present study, the data considered comprises six identical cutting tools operated at a fixed operating condition. Here, cutting force signal in feed direction is measured for every 0.07 h' of operation for each tool until complete failure. As the pre-processing step, four statistical features viz.average cutting force, root mean squared value, signal power, and maximum force level are extracted from cutting force signals for all time steps (0.07 h'). These four features, along with time, are used to represent the degradation of the tools.

3.1 | Performance assessment of the prognostics module

An exhaustive performance investigation of the prognostics module is executed to distinguish the suitability, quality, reliability, robustness, and applicability in a real-world industrial

TABLE 1 Implementation results for both test IDs

		Value	
	Performance metrics	Test ID I	Test ID II
Suitability	Mean absolute error	1.14	0.41
Stability	Precision	1.37	0.51
Quality	Correlation coefficient	0.81	0.91
	Spearman's rho	0.86	0.88
Reliability	Root mean squared error	0.01	0.15

environment. Consequently, we divided the data into two subsets, training data (four tools) and testing data (two tools). All four trajectories in training data correspond to the same failure type (breakage) and operating conditions. The trajectories are of different lengths and comprise the same number of sensor features. Moreover, the noise associated with the sensor values is random and can be filtered using a moving average. Therefore, the trajectories are cleaned (rolling average with window size 10), and the values are normalised. Finally, we obtain a training data set of four run-to-failures with six features corresponding to each time step.

We demonstrate training WTTE-RNN with one long short-term memory (LSTM) layer. The architecture for the RNN is 15*10*20*10*5, with the 15-neuron layer being the LSTM layer. Two sets of experiments (for two assets from testing data, test IDs I and II) were performed to analyse performance. Table 1 gives the details of implementation results for both test IDs. Here, to gauge suitability, mean absolute error (MAE) is computed. In this analysis, MAE measures how close the module's predictions of the RUL are to the actual RUL. The MAE value of 1.14 and 0.41 from the prognostics module displays that the predicted RUL is close to the actual RUL, demonstrating the suitability of the module in real life. Stability is measured by precision in terms of the dispersion of the prediction error around its mean. Values as low as 1.37 and 0.51 show that the predicted RULs have higher stability. Next, the quality of prediction is scaled based on the correlation coefficient and Spearman's rho. These measures disclose the strength of the association between predicted and actual RUL. Values near 1 for both test IDs show that the predicted RULs have a strong positive relationship with the actual RULs, showcasing the high quality of predictions from the module. Moreover, the high reliability of these predictions is evident from lower values for the root mean squared error. Lastly, robustness is evaluated by plotting each performance output of the prognostics module, as shown in Figures 1 and 2. The perception from this figure is that each actual and predicted RUL is near the other, showcasing that the prognostics module is robust in predicting the RUL of the asset. These implementation results guarantee a proficient predictive maintenance framework dependent on a timely cautioning of upcoming failures.

(Note: Errors exist between the predicted and actual RUL (as seen in Figures 1 and 2). Concerning their influence, as
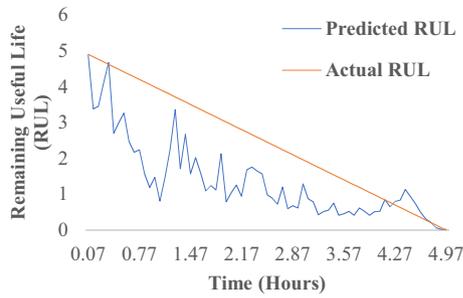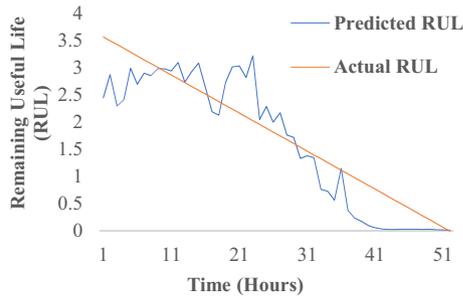
FIGURE 1 Results for test ID I

FIGURE 2 Results for test ID II

TABLE 2 Parameters utilised in the case study

Parameter	Production rate P_r (Product/h)	Cost of lost production C_{lp} (GBP/product)	Cost of the Labour C_L (GBP/h)
Value	8	60	100

uncertainty in prediction decreases, the optimal maintenance timing gets closer to the RUL. Section 3.2.1 analyses the implications of imperfect prognostics information on maintenance decisions in detail.)

3.2 | Predictive maintenance framework

We consider a production facility representing the cutting tool as a single-component machine producing mild steel plates. All gathered operational cost parameters from the historical data are mentioned in Table 2. The cost of the cutting tool (C_A) utilised in the process is 3000 GBP. The mean life of the tool is computed by historical failure data and given as 3.90 h. As per maintenance history, the mean times to conduct corrective (A_{CR}) and predictive (A_{PdR}) replacement tasks are 0.6 h each. The fixed cost of corrective (C_{FCR}) and predictive (C_{FPdR}) replacement is 3000 GBP separately.

3.2.1 | Implications of imperfect prognostics information on maintenance decisions

To analyse the implications of imperfect prognostics information on maintenance decisions, we evaluate the $[ETC]_{(PdM)}$ at every 10th time step of the test assets. Here, for each test asset ID, we run the prognostics module at a different time step ($T_1, T_2 \dots T_n$), predict the probability distribution (θ and η) and estimate the variance. These parameters (θ and η) are fed to the predictive maintenance planning model to obtain the

TABLE 3 Implementation results for both test assets

	Time (T_i) h	Predicted η (h)	Predicted θ	Variance	RUL (h)	Optimal time for replacement O_{PdR} (h)
Test asset ID I	$T_1 = 0.07$	5.38	3.96	1.90	4.87	4.08
	$T_2 = 0.70$	2.40	3.98	0.37	2.17	1.36
	$T_3 = 1.40$	2.95	3.98	0.57	2.67	1.65
	$T_4 = 2.10$	1.39	3.97	0.13	1.26	0.41
	$T_5 = 2.80$	0.66	3.95	0.03	0.60	0.17
	$T_6 = 3.50$	0.46	3.93	0.01	0.41	0.15
	$T_7 = 4.20$	0.72	3.96	0.03	0.65	0.17
	$T_8 = 4.90$	0.02	1.65	0.00	0.02	0
Test asset ID II	$T_1 = 0.07$	2.70	3.98	0.48	2.45	1.72
	$T_2 = 0.70$	3.29	3.98	0.71	2.98	2.04
	$T_3 = 1.40$	3.33	3.98	0.73	3.02	2.04
	$T_4 = 2.10$	1.48	3.97	0.14	1.34	0.63
	$T_5 = 2.80$	0.11	3.50	0.00	0.10	0.01
	$T_6 = 3.50$	0.03	1.97	0.00	0.02	0

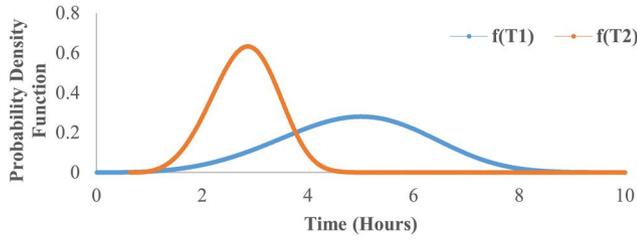Abbreviation: RUL, remaining useful life.

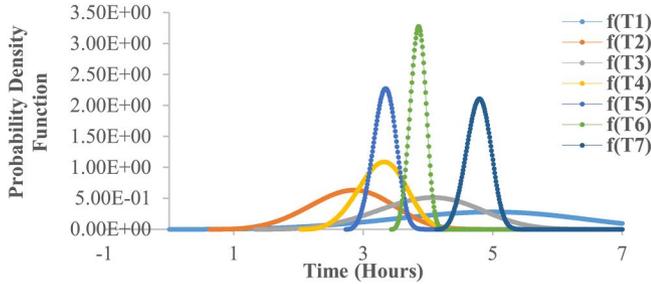**FIGURE 3** Probability density function plot at T_1 and T_2 for ID I

FIGURE 4 Probability density function plot for T_1 to T_7 for ID I
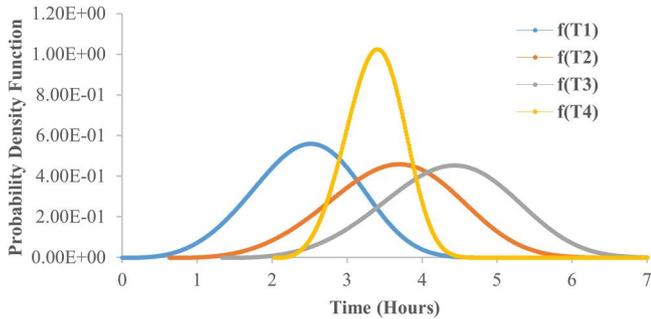

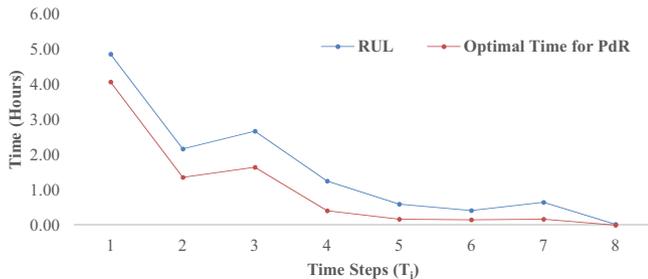**FIGURE 5** Probability density function plot for T_1 to T_4 for ID II

FIGURE 6 Remaining useful life (RUL) versus optimal time for predictive replacement for T_1 to T_8 for ID I

optimal time for predictive replacement (O_{PdR}). Table 3 presents the detailed results for both test assets.

The variance of the distribution gives details about uncertainty in prediction (higher variance means higher prediction uncertainty). At the initial stage of the operation when there is very little information about the condition of the asset,

the variance will be higher, implying higher uncertainty in prediction. As time goes on and we obtain more information about the asset's health, the variance will be reduced, implying lower uncertainty in prediction. Along those lines, Figure 3 shows the probability density function (PDF) for test asset ID I at the first and second time steps. It can be observed that at T_1, the PDF is very wide, with a variance of 1.90. This is because the prediction of θ and η at time step T_1 has high uncertainty because it is the first point of prediction, and there is a lack of information in terms of asset condition. However, as time goes on and we obtain more information about asset condition, uncertainty in predicted θ and η is reduced. This is evident as the PDF at T_2 is narrowed, with 80.53% less variance than in time step T_1, indicating a drop in uncertainty. This becomes clearer from Figures 4 and 5, which show the PDF for all time steps for test asset IDs I and II. This implies that uncertainty about prediction drops as time goes on. On the other hand, Figures 6 and 7 show the predicted RUL and optimal time for replacement (O_{PdR}). It can be observed that at time step T_1, the optimal time for PdR is far from the predicted RUL, although at time step T_5, the optimal time for PdR is very close to the predicted RUL. This is again because of the reduced uncertainty of the prediction. This implies that as uncertainty goes down, the maintenance timing becomes closer to the RUL because the risk of making a wrong decision is decreasing. Accordingly, the guideline from this for the operational planner is to not stick to the initial optimal maintenance plan but to instead dynamically update the predictive maintenance plan as time goes on to make maintenance decision at the right time.

3.2.2 | Sensitivity analysis

In the exercise, the approximation of appropriate process and cost parameters are subject to inaccuracies, so it is vital to distinguish the influences of errors on the quality of the output attained. Accordingly, a systematic sensitivity analysis utilising essential model parameters is carried out, as can be seen in Table 4. The base level utilised is the same as in the case study along with four other levels of these parameters at $\pm 10\%$ and $\pm 20\%$ of the base value. The ranges of the optimal parameter and obtained cost are presented in Tables 4 and 5. Figure 8 shows that $[ETC]_{(PdM)}$ is more sensitive to the fixed cost of PdR and less susceptible to the mean time to perform the PdR etc. Thus, the estimation of the fixed cost of a PdR should be carried out accurately.

4 | CONCLUSION AND FUTURE WORK

This paper formulates a comprehensive framework from real-time prognostics to maintenance decisions. The intention was to render manufactures with a complete maintenance support system to instantaneously prevent asset performance degradation and unexpected failures. The following are the key offerings of this study:

F I G U R E 7 Remaining useful life (RUL) versus optimal time for predictive replacement for T_1 to T_6 for ID II
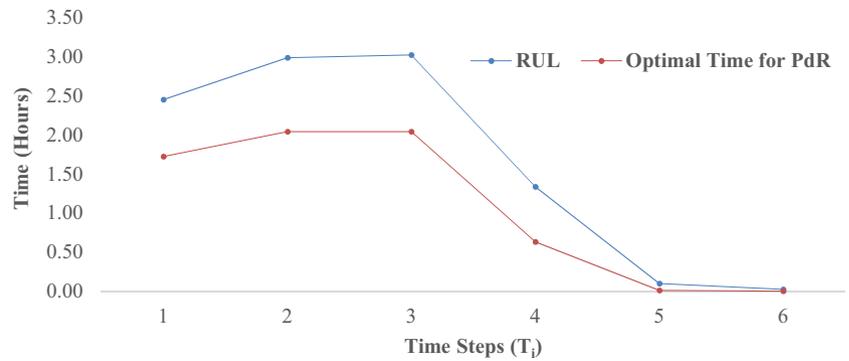

T A B L E 4 Sensitivity analysis

Parameter	Base Level	−20%	−10%	+10%	+20%	Expected total cost per unit time of carrying out predictive maintenance ($[ETC]_{(PdM)}$)					Range
						Base level	−20%	−10%	+10%	+20%	
A_{PdR}	0.6	0.48	0.54	0.66	0.72	1008.86	994.57	1001.71	1016.00	1023.52	994.57-1023.52
C_{lp}	60	48	54	66	72	1008.86	993.69	1001.28	1016.38	1023.90	993.69-1023.90
C_L	100	80	90	110	120	1008.86	1005.73	1007.28	1010.42	1011.99	1005.73-1011.99
C_{FPdR}	3000	2400	2700	3300	3600	1008.86	885.64	947.25	1070.46	1132.06	885.64-1132.06
C_A	3000	2400	2700	3300	3600	1008.86	975.52	994.91	1020.00	1026.06	975.52-1026.06
A_{CR}	0.6	0.48	0.54	0.66	0.72	1008.86	1004.87	1006.89	1010.85	1012.76	1004.87-1012.76
C_{FCR}	3000	2400	2700	3300	3600	1008.86	965.61	990.01	1024.54	1038.05	965.61-1038.05

T A B L E 5 Range of optimal design parameters acquired through sensitivity analysis

Design parameter	Optimal time for replacement O_{PdR} (h)
Range	3.59–4.86

a) For the prognostics phase, a new sensor-based prognostics module was modelled employing the WTTE-RNN in which prognostics competence was enhanced by predicting the parameters of failure distribution despite a single time to failure, and therefore, the offered approach delivers a superior response to real-world requirements.

b) For the post-prognostics phase, a new predictive maintenance planning model was framed through a trade-off between CM and lost remaining life due to predictive maintenance, thus allowing rapid optimisation of time for maintenance via all gathered operational and maintenance cost parameters.

c) The model's performance is highlighted via a case study from a manufacturing environment complimented with a systematic sensitivity analysis. The influence of the imperfect prognostics information on maintenance decisions is debated by showcasing interesting insights. Namely, uncertainty about prediction drops as time goes on, and as uncertainty is reduced, the maintenance timing becomes closer to the RUL because the risk of making a wrong decision decreases over time.

In essence, the model is an entire cognitive operation from conducting prognostics to making maintenance decisions. Such complete models integrating monitoring characteristics, prognostics, and maintenance assessment can give rise to fruitful discussions.

In the future, this framework will be extended to consider a fleet of assets for fleet-level prognostics and predictive maintenance planning. The service operations of a fleet of assets can be controlled with a multi-agent system. In this case, our model can be deployed along with these agents so they calculate health indicators using real-time data collected from the controlled assets. Using collaborative prognostics, it has been previously shown that agents engaging in information exchange with others, controlling assets with similar conditions of degradation behaviour, can improve the accuracy of predictions.

One other direction for future work is to generate updated maintenance plans and schedules, based on our proposed model, with an accurate real-time view of asset condition at each time of evaluation. As a desirable characteristic in these agent-based control systems is to minimise resource consumption and communication overhead, another research direction is to analyse the impact on the multi-agent control system of using a reduced number of features and aggregated statistics as in the proposed prognostics model.
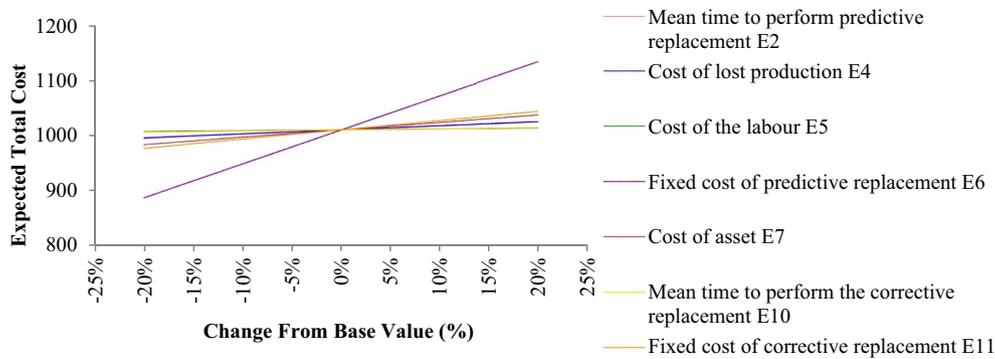
ACKNOWLEDGEMENTS

FIGURE 8 Expected total cost versus percentage change in model parameters. E2, E4, E5, E6, E7, E10, E11 corresponds to the cells in the model

ORCID

Amit Kumar Jain https://orcid.org/0000-0001-7939-7810
Manuel Herrera https://orcid.org/0000-0001-9662-0017
Ajith Kumar Parlikad https://orcid.org/0000-0001-6214-1739

REFERENCES

1. Sherwin, D.: A review of overall models for maintenance management. J. Qual. Mainten. Eng. 6(3), 138–164 (2000)
2. Jain, A.K., Lad, B.K.: A novel integrated tool condition monitoring system. J. Intell. Manuf. 30(3), 1423–36 (2019)
3. Gašperin, M., et al.: Model-based prognostics of gear health using stochastic dynamical models. Mech. Syst. Signal. Process. 25(2), 537–48 (2011)
4. Jain, A.K., Lad, B.K.: Data driven models for prognostics of high speed milling cutters. Int. J. Perform. Eng. 12(1), 3–11 (2016)
5. Jain, A.K., Lad, B.K.: Quality control based tool condition monitoring. In: Annual Conference of the Prognostics and Health Management Society, 1–10 (2015)
6. Bakliwal, K., et al.: A Multi Agent System architecture to implement Collaborative Learning for social industrial assets. IFAC-PapersOnLine. 51(11), 1237–42 (2018)
7. Palau, A.S., et al.: An Industrial Multi Agent System for real-time distributed collaborative prognostics. Eng. Appl. Artif. Intell. 85, 590–606 (2019)
8. Jain, A.K., Lad, B.K.: Predicting remaining useful life of high speed milling cutters based on artificial neural network. In: IEEE International Conference on Robotics, Automation, Control and Embedded Systems, Chennai, pp. 1–5 (2015). https://doi.org/10.1109/RACE.2015.7097283
9. Jain, A.K., Kundu, P., Lad, B.K.: Prediction of remaining useful life of anAircraft engine under unknown initial wear, In 5th International & 26th All India Manufacturing Technology, Design and Research Conference, IIT Guwahati, pp. 494 (1-5) (2014) https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1056.1761&rep=rep1&type=pdf
10. Yang, C., et al.: Particle filtering-based methods for time to failure estimation with a real-world prognostic application. Appl. Intell. 48(8), 2516–2526 (2018)
11. Ruiz-Tagle Palazuelos, A., Droguett, E.L., Pascual, R.: A novel deep capsule neural network for remaining useful life estimation. Proc. Inst. Mech. Eng. O J. Risk Reliab. 234(1), 151–67 (2020)
12. Eleftheroglou, N., et al.: Intelligent data-driven prognostic methodologies for the real-time remaining useful life until the end-of-discharge estimation of the Lithium-Polymer batteries of unmanned aerial vehicles with uncertainty quantification. Appl. Energy. 254, 113677 (2019)
13. Hu, Y., et al.: Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method. Chin. J. Mech. Eng. 32(1), 15 (2019)
14. Baur, M., Albertelli, P., Monno, M.: A review of prognostics and health management of machine tools. Int. J. Adv. Manuf. Technol. 107, 2843–63 (2020)
15. Rezamand, M., et al.: Critical Wind Turbine Components Prognostics: A Comprehensive Review. IEEE Trans. Instrum. Meas. 69(12), 9306–9328 (2020). https://doi.org/10.1109/tim.2020.3030165
16. Xiao, L., et al.: Joint optimization of production scheduling and machine group preventive maintenance. Reliab. Eng. Syst. Saf. 146, 68–78 (2016)
17. Wang, S., Liu, M.: A branch and bound algorithm for single-machine production scheduling integrated with preventive maintenance planning. Int. J. Prod. Res. 51(3), 847–68 (2013)
18. Chen, X., Xiao, L., Zhang, X.: A production scheduling problem considering random failure and imperfect preventive maintenance. Proc. Inst. Mech. Eng. O J. Risk Reliab. 229(1), 26–35 (2015)
19. Freitag, S.C., et al.: Power transformer risk management: predictive methodology based on reliability centered maintenance. In: 2018 Simposio Brasileiro de Sistemas Eletricos, 1–6.SBSE) (2018)
20. He, Y., et al.: Cost-oriented predictive maintenance based on mission reliability state for cyber manufacturing systems. Adv. Mech. Eng. 10(1), 1687814017751467 (2018)
21. Palau, A.S., Dhada, M.H., Parlikad, A.K.: Multi-agent system architectures for collaborative prognostics. J. Intell. Manuf. 30(8), 2999–3013 (2019)
22. Chrétien, S., et al.: Post-prognostics decision for optimising the commitment of fuel cell systems. arXiv preprint arXiv:1710.08812. (2017)
23. Skima, H., et al.: Post-prognostics decision making in distributed MEMS-based systems. J. Intell. Manuf. 30(3), 1125–36 (2019)
24. Lei, Y., et al.: Machinery health prognostics: a systematic review from data acquisition to RUL prediction. Mech. Syst. Signal Process. 104, 799–834 (2018)
25. Ebeling, C.E.: An introduction to reliability and maintainability engineering. Tata McGraw-Hill Education, New Delhi (2004)
26. Heimes, F.O.: Recurrent neural networks for remaining useful life estimation. In: 2008 International Conference on Prognostics and Health Management, pp. 1–6. IEEE, Gothenburg (2008). Oct 6

27. Martinsson, E.: Wtte-rnn: Weibull time to event recurrent neural network. Doctoral dissertation. Chalmers University Of Technology (2016)

28. Baptista, M., Prendinger, H., Henriques, E.: Prognostics in aeronautics with deep recurrent neural networks. PHM Society European Conference. 5(1), 11–11 (2020)

29. Nguyen, K.T., Medjaher, K.: A new dynamic predictive maintenance framework using deep learning for failure prognostics. Reliab. Eng. Syst. Saf. 188, 251–62 (2019)

30. Jain, A.K., Lad, B.K.: Dynamic optimization of process quality control and maintenance planning. IEEE Trans. Reliab. 66(2), 502–17 (2017)

31. Jain, A.K., Lad, B.K.: Data-centric real-time integration of diagnostics, prognostics, and process quality control: realizing intelligent manufacturing. Doctoral dissertation. Discipline of Mechanical Engineering, IIT, Indore (2018)