CrossMark

ORIGINAL ARTICLE

Integrated Equipment Health Prognosis Considering Crack Initiation Time Uncertainty and Random Shock

Fu-Qiong Zhao[1] · Ming-Jiang Xie[1] · Zhi-Gang Tian[1] · Yong Zeng[2]

Abstract With integrated equipment health prognosis, both physical models and condition monitoring data are utilized to achieve more accurate prediction of equipment remaining useful life (RUL). In this paper, an integrated prognostics method is proposed to account for two important factors which were not considered before, the uncertainty in crack initiation time (CIT) and the shock in the degradation. Prognostics tools are used for RUL prediction starting from the CIT. However, there is uncertainty in CIT due to the limited capability of existing fault detection tools, and such uncertainty has not been explicitly considered in the literature for integrated prognosis. A shock causes a sudden damage increase and creates a jump in the degradation path, which shortens the total lifetime, and it has not been considered before in the integrated prognostics framework either. In the proposed integrated prognostics method, CIT is considered as an uncertain parameter, which is updated using condition monitoring data. To deal with the sudden damage increase and reduction of total lifetime, a virtual gradual degradation path with an earlier CIT is introduced in the proposed method. In this way, the effect of shock is captured through identifying an appropriate CIT. Examples of gear prognostics are given to demonstrate the effectiveness of the proposed method.

✉ Zhi-Gang Tian
  ztian@ualberta.ca

1 Department of Mechanical Engineering, University of Alberta, Edmonton T6G 2G8, Canada

2 Concordia Institute for Information Systems Engineering, Concordia University, Montreal H3G 2W1, Canada

List of Symbols

a	Crack length
m, C	Material parameters in Paris' law
N	Loading cycles
ΔK	Stress intensity factor range
t_0	Crack initiation time
a_0	Initial crack size
ξ	Random vector consisting of material parameters
e	Measurement error
τ	Standard deviation of measurement error
ΔN	Incremental number of loading cycles
a_C	Critical crack size
$f_{prior}(t_0, \xi)$	Prior distribution of CIT and material parameters
a_u^{obs}	Observed crack size at inspection time t_u
$l(a_u^{obs}\vert t_0, \xi)$	Likelihood to observe crack size a_u^{obs} conditional on (t_0, ξ)
$f_{post}(t_0, \xi\vert a_u^{obs})$	Posterior distribution of CIT and material parameters conditional on a_u^{obs}
t_{0_actual}	Actual crack initiation time in the shock degradation
$t_{0\_virtual}$	Virtual crack initiation time in the shock degradation
t_u	Inspection time
$a_{1:s}^{obs}$	Set of observation history $\{a_u^{obs} : u = 1, 2, \cdots, s\}$
l_s	Likelihood to observe a crack history $a_{1:s}^{obs}$

🗘 Springer

$\tilde{f}_{post}$	Adjusted posterior distribution when applying Bayesian inference in shock degradation
δ	Threshold for shock detection
F	Failure time

1 Introduction

Remaining useful life (RUL) prediction of component/system has drawn significant attention from researchers and practitioners because of its great importance for improving reliability and availability of engineering systems. In the context of condition based maintenance/prognostics and health management (CBM/PHM) [1, 2], accurate online prediction on RUL of component/system turns out to be the key for the success of this advanced maintenance technology in system health management. Prognostic algorithms have been proposed in a large number of publications for various industrial applications. Those algorithms are mostly either data-driven [3–7], physically motivated [8–16] or model and data integrated [17–26].

In this paper, an integrated prognosis method is developed to account for both the uncertainty in crack initiation time (CIT) and the shock in the degradation. The introduction of uncertainty in CIT is the key to deal with the discontinuity caused by shock in the degradation path. A crack is typically initiated in a component, propagates due to stress, and eventually will cause component failure. We define CIT as the time instant when the damage is detected and the prognosis starts. Indeed, when the CIT is adjusted, it is the "intercept" with the time axis of degradation path at the initial crack size that is adjusted, which is a newway of adjustment. The combination of adjustment in both "slope" and "intercept" will better characterize the real degradation path given the crack observations. To deal with the discontinuity in the shock degradation, we aim to find a virtual degradation path with an earlier CIT that could achieve the same failure time as the shock degradation.

1.1 Literature Review

Prognostic models are typically categorized into three groups: data-driven methods, physics-based methods and integrated methods.

Data-driven models are built purely on data such as sensor-collected condition monitoring data. They typically require large amount of data to achieve reasonable accuracy. In Ref. [3], a proportional-hazards model with time dependent stochastic covariates was used as lifetime model to predict the failure rate and to optimize the unit replacement policy. The set of failure time and covariates was assumed to be a joint nonhomogeneous Markov process. The maximum likelihood was applied to estimate the model parameters including the coefficients of covariates and the transition probabilities of the Markov process. In Ref. [4], the degradation model was selected to be exponential. Two error terms were considered: one was treated as a multiplicative random variable, and the other was treated as a multiplicative Brownian motion process. A Bayesian procedure was applied to update the parameters in the exponential model. As a research mainstream of data-driven models, various machine learning techniques were investigated in the literature. The authors of Ref. [5] developed neural networks to predict bearing failure time, which aimed to train a relationship between the bearing service time and the corresponding vibration spectrum. The authors of Ref. [6] developed a neural network to predict RUL using both failure and suspension condition monitoring histories. An extended recurrent neural network was proposed to predict the health condition of gears in Ref. [7]. The incorporation of Elman context layer in the proposed networks enhanced its ability to model nonlinear time series. Data-driven models are straightforward to establish given sufficient and well-distributed data. Therefore, data of good quality availability is a prerequisite for data-driven models, which is rare for costly industrial equipment.

In contrast, physics-based models resort to physical laws governing the defect growth, where the values or distributions of the model parameters in the physical models, e.g., material dependent parameters, are kept constant and will not be updated based on condition monitoring information. A commonly used law to describe crack propagation is Paris' law, which was originated in Ref. [16]. Many publications are devoted to devising numerical algorithms to calculate the quantity (e.g., stress intensity factor) needed in Paris' law. Readers can refer to Refs. [8–10] for these approaches in which finite element modeling was discussed for stress analysis near crack tip. In the physics-based prognostic models, component failure time is defined as when the defect size exceeds a critical value. Authors of Ref. [11] investigated several factors that influence the crack growth trajectory in the gear tooth, including backup ratio, initial crack location, fillet geometry, rim/web compliance, gear size and pressure angle. The research work in Ref. [12] took account of the variation of moving load on gear tooth into crack growth prediction by breaking the tooth engagement into multiple steps. Authors in Ref. [13] applied Paris' law to predict fatigue crack growth with utilization of transmission error to estimate the current crack size, which improved predictive accuracy. The service life of gears was divided into crack initiation and crack propagation periods in Ref. [14]. The strain-life method was used to determine the time required for crack initiation, while Paris' law was used to

obtain the time required for the crack to grow from initial crack size to the critical value. Kacprzynksi et al. [15] developed a prognostic tool which predicts gear failure probability by fusing physics-of-failure models and diagnostics information. The results showed variance reduction in failure probability when diagnostics information was present. Physical models can achieve high predictive accuracy if appropriately built. However, it demands comprehensive physics theory and intensive computation to build physical models of high-fidelity. Furthermore, physical models are unavailable for complex systems which limit its usages in real-world applications. In addition, physical models are deterministically used in the above mentioned literature, which means they are unable to address the uncertainty in failure.

Recently, integrated prognostics methods [17–26] were developed to achieve real-time RUL prediction during the system operations by combining both condition monitoring (CM) data and physics of failure. The integrated methods usually have an updating process by assimilating observations, during which the uncertainty is expected to shrink so that the confidence increases in the predicted results. The integrated prognostics methods are more advantageous than physics-based methods in that the model parameters are able to be adjusted for a specific component in a specific working condition, and more advantageous than data driven methods in that massive data trending is not necessary. Bayesian framework allows for uncertainty quantification; hence, it is widely used in integrated prognostics methods. A problem of crack propagation in a fuselage panel of aircraft was considered in Ref. [17]. Bayesian inference was used to characterize parameters in Paris' law and error term. The work in Ref. [18] extended the methods in Ref. [17] to consider the correlation between model parameters. In authors' prior work [19], integrated prognostics methods were proposed for gear health prediction and uncertainty quantification. Physical models include Paris' law, fracture mechanics model and one-stage gearbox dynamics model. Through Bayesian updating, the uncertainty in RUL prediction was reduced as crack measurements became available. To increase the efficiency of uncertainty quantification, the authors proposed to use polynomial chaos expansion to accelerate the Bayesian process in prognostics [20]. Furthermore, an approach was devised to deal with time-varying operating conditions to make them applicable in various loading environment [21]. Bayesian framework was also used for bearing health prognostics. The bearing spall propagation was investigated in Ref. [22] where Bayesian inference was applied to reduce prediction uncertainty. In recent years an increasing volume of literature was published that treated prognostic model as a dynamic system mainly because of its natural interface with real-time data. In Ref. [23], a health indicator extracted from vibration signature was fed into Paris' law to update the RUL of bearings. The update was implemented using a Kalman filter. Prognostic models in Refs. [24–26] are established in a particle filtering framework, in which the problem of non-linear state transition and non-Gaussian noise can be tackled.

1.2 Motivations to Consider the Uncertainty in CIT

A large grain of inherent uncertainty imposes major challenges in prognostic methods development. There are various sources for uncertainties, such as micro-structure of material, operating conditions, working environment, measurement as well as human factors. Many research efforts go to addressing how to identify, capture and manage these multiple uncertainties to make the RUL prediction more accurate, precise and reliable. The existing prognostic approaches usually start the prediction at an assumed time instant when a fault at certain severity is detected. This is based on an assumption that the starting point of prognosis is accurate. However, due to the limitations of the fault detection and diagnostic technologies, there is a large variation in the accuracy of fault detection. This variation affects the prediction accuracy accordingly: an early starting point of prognostics will lead to underestimated RUL, and late starting point will lead to overestimated RUL. In authors' prior work [19–21], it was proposed to update the uncertain model parameters in Paris' law to make RUL prediction more accurate by feeding CM data to a Bayesian framework. These model parameters actually determine the "slope" of degradation path in a scale of damage size versus time. What the Bayesian updating process does is adjusting the "slope" of the degradation path to maximize the likelihood of crack observations when an uninformative prior is given. Apart from the "slope", another factor that controls the degradation path is the "intercept" with time axis. Hence, to better characterize the degradation path, it is needed to explicitly consider the uncertainty in CIT, which determines its "intercept" with time axis.

1.3 Motivations to Consider the Shock Degradation

Among existing approaches to real-time prognostics, most of them investigated components/systems undergoing a gradual degradation with time. Such gradual degradation is governed by a dynamic model, either data driven or physics-of-failure based, which usually leads to a continuous degradation. Very few of them have considered the degradation path containing shock, which appears as a sudden jump in the degradation. The shock will cause sudden damage increase and then accelerate the

degradation rate. In reliability engineering, researchers have investigated the ways of modeling shock process in system reliability analysis [27–29]. The purpose of these studies was to investigate reliability properties and/or maintenance strategies considering the shock effects. While in the present study, we look into the RUL prediction in a different perspective to consider the shock in the degradation for a specific unit under monitoring, so that the prognostic solution can be given in real time. This purpose is also the main focus of prognostic algorithm development in CBM/PHM.

1.4 Shock Degradation Prediction Using CIT Adjustment

In the crack propagation problem, model parameters in the Paris' law include material dependent coefficients, which should not be affected by external forcing factors, i.e., overload causing shock. Hence, additional uncertainty source other than model parameters is needed to account for the effect of shock on the degradation prediction. Because sudden damage increment results in a discontinuity in the degradation path, the lifetime is shortened accordingly. Note that, if the slope of degradation path is given as fixed, the degradation path with a shortened lifetime can be considered as equivalent to a gradual degradation path with an earlier CIT. Therefore, the variation in CIT provides a degree of freedom in translational adjustment for the degradation path. With both slope and translational adjustments, the proposed method is expected to reduce the uncertainty in RUL prediction and to capture the effect of shock on the degradation as well. In different applications, the shock occurrence time might be known or unknown. If known, the corrective action can be taken right after the shock occurs; otherwise additional work is needed. In this study, both the two situations will be considered.

The remainder of this paper is organized as follows. Section 2 introduces the integrated prognostics framework that gives a global view of the structure of the proposed method. In Section 3, the Bayesian updating procedure is presented to deal with gradual degradation considering uncertainty in both CIT and material parameters. Section 4 investigates integrated prognostics for degradation with shock. Two cases are considered: a). shock degradation with known shock occurrence time; b) shock degradation with unknown shock occurrence time. The formula to compute RUL is also given in this section. In Section 4, examples are given to show the effectiveness and efficacy of the methods. Section 5 concludes the work.

2 Integrated Prognostics Framework

An integrated prognostics framework is proposed in this section, and we use gears with fatigue crack as an example to present the proposed method. The degradation types considered include gradual degradation and shock degradation. The gradual degradation model was described by Paris' law [16], which was a widely accepted equation describing crack propagation rate. This equation is shown in Eq.(1)

$$
\begin{cases}
\dfrac{\mathrm{d}a}{\mathrm{d}N} = C(\Delta K(a))^m, \\
a(t_0) = a_0,
\end{cases}
\tag{1}
$$

where a is crack size, N is loading cycle, ΔK is the range of stress intensity factor (SIF) obtained using finite element (FE) method, t_0 is CIT, a_0 is initial crack size, m and C represent the material dependent model parameters. This model was based on physics of failure and the model parameters have physical meanings. In linear fracture mechanics theory, stress intensity factors (ΔK) fully describe the stress field distribution near the crack tip area. Pairs' law states that the crack growth rate is a function of ΔK, therefore, it is based on physics of failure. The model parameters can be obtained by regressing the Paris' law in a log-log scale using crack measurements, stress intensity factors, and associated cycles in the fatigue experiment as reported in Ref. [30]. Through the experiment, it turns out that the parameters in Paris' law are material dependent. SIF together with material parameters determines the crack growth rate from a physical point of view as given in Paris' law. Compared to data-driven models, it provides a better predictive capability. In this paper, parameters that are considered uncertain include CIT t_0 and material dependent parameters $\xi = (m, C)$. In addition, measurement error e is also considered, which follows a normal distribution with zero mean and standard deviation τ. The Paris' law is discretized using first-order Euler's rule as in Eq.(2) to obtain the crack size at inspection times.

$$
\begin{cases}
a(i\Delta N) = a((i-1)\Delta N) + (\Delta N)C[\Delta K((a(i-1)\Delta N))]^m, \\
a(t_0) = a_0,
\end{cases}
\tag{2}
$$

where ΔN is incremental loading cycles. The iteration proceeds until the current inspection time t_i is reached. The crack size obtained is denoted as $a_i(t_0, \xi)$. The other type of degradation contains shock. The shock causes a sudden damage accumulation due to external impact, such as a transient overload. In the case of fracture, the phenomenon of shock is a sudden increase in crack size. The degradation containing shock is defined in this paper as shock degradation.

The core idea of "integrated" prognostics method is to combine CM data (e.g., crack sizes observations) and physical models in a way that crack size observations can be utilized to adjust the physical model. After the adjustment, the physical model is expected to predict the RUL better. The adjustment is incurred when a new observation is available. The sequential adjustments form a series of updates, which are triggered at every inspection time. Figure 1 shows the data flow in the integrated prognostics method proposed in this paper. The model update is achieved by updating the distributions of uncertain parameters through Bayesian inference. The posterior distribution of the uncertain parameters is applied in Paris' law to calculate RUL. Meanwhile, the posterior distribution is fed into next iteration as the prior distribution. The first prior distribution in absence of any observations is obtained by least-square regressing the existing historical failure paths [19].

When the updating process is executed, three cases are considered: Case 1: no shock occurs; Case 2: shock occurs at a known time; and Case 3: shock occurs at an unknown time. As shown in Figure 1, when gradual degradation without shock is considered (Case 1), the posterior distribution is directly used as the prior distribution for next iteration. However, when shock is considered (Cases 2 and 3), an adjustment is added for marginal posterior distribution of CIT before it is used as the prior for the next iteration. The purpose of this adjustment is to make sure that a virtual gradual degradation can be identified equivalently with an appropriate earlier CIT $t_{virtual}$. The equivalence is defined in a sense of failure time. Accordingly,

the likelihood function also needs modifications to eliminate the adverse effects brought out by the observations before shock occurs. The details of the three cases will be addressed in Section 3.

3 Integrated Prognostics Method Considering Uncertainty in CIT

The well-known Paris' law, as discussed in Section 2, describes the crack propagation rate using the principle of linear elastic fracture mechanics. It represents the crack propagation rate as a function of SIF and material dependent parameters m and C. Hence, given the applied loading, both m and C are the factors that determine the "slope" of degradation path in a scale of crack size versus time. The priorwork in Refs. [19–21] focused on the adjustment of $\xi = (m, C)$, which represents the "slope" of the degradation path. The motivation to identify CIT as another source of uncertainty is that this uncertainty is difficult to account for by "slope". It is actually a translational movement along the time axis. In another word, the CIT is the "intercept" with the time axis of the degradation path at the initial crack size. By adjusting the "slope" and the "intercept" simultaneously, it is expected to obtain an optimal approximation to the real degradation path.

Denoting the initial crack size as a_0 and the critical crack size as a_C, as shown in Figure 2, the degradation paths are generated by varying CIT and physical model parameters. As a result, the variation in the crack size at a certain inspection time t_i is contributed by both of the

Figure 1 Data flow in the proposed integrated prognostics method
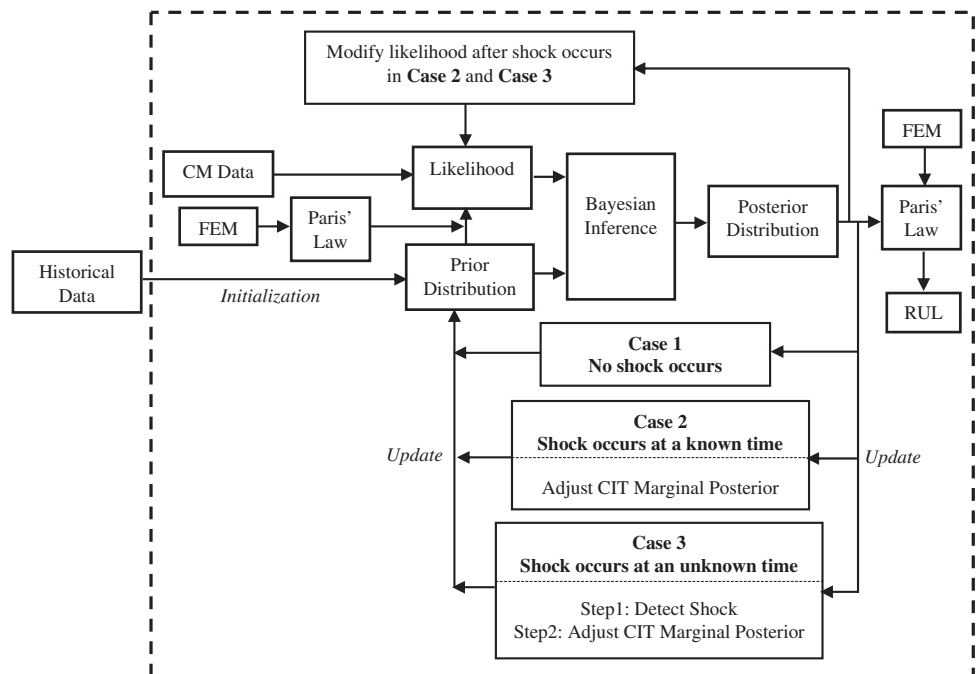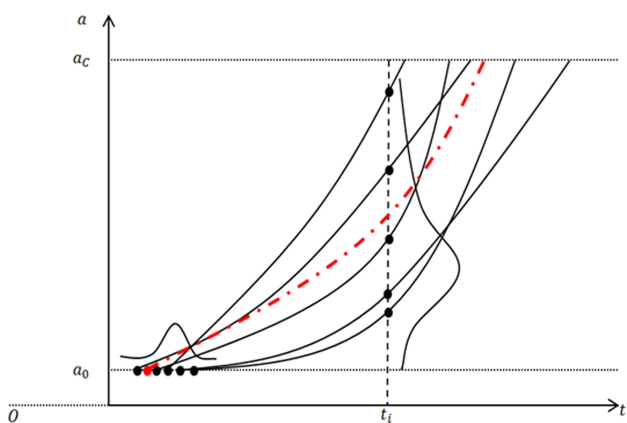
Figure 2 Degradation paths generated by varying CIT and physical model parameters

uncertainty in CIT and the physical model parameters. The path using a red dash line represents the actual degradation path, which has actual values for CIT and the physical model parameters. However, these actual values are not known exactly beforehand. Only some beliefs or assumptions are available based on experience and historical data. These beliefs or assumptions are known as a prior in Bayesian statistics. The objective to use the Bayesian inference is to narrow the prior distribution so that the means will approach actual values and the standard deviations will be reduced. This is achieved by feeding the crack observations into Bayesian inference.

Define a set of random variable, $\xi = (m, C)$, representing the uncertainty from model material parameters. As before, t_0 stands for the CIT. Here, t_0 is also considered as a random variable to account for the uncertainty in CIT. Suppose that several failure histories are available with the information on inspection times and associated crack sizes. Then a prior distribution $f_{prior}(t_0, \xi)$ can be obtained by least-square regression and statistical fitting [19]. If it is assumed that the crack measurement error follows a zero-mean Gaussian distribution with τ as the standard deviation, at a certain inspection time t_u the likelihood to observe a crack size of $a_u^{obs} = a^{obs}(t_u)$ is

$$l\left(a_u^{obs}\big|t_0, \xi\right) = \frac{1}{\sqrt{2\pi}\tau}\exp\left(-\frac{\left(a_u^{obs} - a_u(t_0, \xi)\right)^2}{\tau^2}\right). \quad (3)$$

In the Bayesian inference framework, a posterior distribution $f_{post}(t_0, \xi)$ can be obtained by

$$f_{post}\left(t_0, \xi\big|a_u^{obs}\right) = \frac{l\left(a_u^{obs}\big|t_0, \xi\right)f_{prior}(t_0, \xi)}{\int l\left(a_u^{obs}\big|t_0, \xi\right)f_{prior}(t_0, \xi)\mathrm{d}t_0\mathrm{d}\xi}. \quad (4)$$

The update process is executed at the inspection time when a new observation on crack size is available. The posterior distribution of the current update process will serve as the prior distribution for the next update process.

To circumvent the intractable integration in the Bayesian formula, importance sampling technique is used to obtain samples which follow the posterior distribution. In order to simplify the problem and to emphasize the different effects from "slope" and "intercept", in the following discussions, only uncertainties in m and t_0 are considered while C is treated as a constant.

4 Physical Model Updating Considering Shock in the Degradation

Different from the gradual degradation, shock degradation is another degradation type where the damage is accumulated suddenly leading to a jump and discontinuity in the degradation path. In Figure 3, the blue line depicts a gradual degradation path with CIT t_{0_actual}, and the marks of circle represent crack sizes at inspection times. Assume that a shock happens between the third and the fourth inspection times. The shock will result in a sudden crack increment, denoted by dotted purple line segment. After the shock, the gradual degradation continues at an accelerated rate. This forms a shock degradation path depicted by a magenta line with marks of square. Because of the shock, the lifetime is shortened with an extent denoted by a green dash line segment. The shock degradation is a discontinuous curve. Material dependent model parameters should be static because the material is not changing. It also explains why the two degradation paths are parallel after the shock occurs. Hence, a new uncertainty source should be uncovered to account for the change in lifetime due to shock. By noticing that our interest is to predict the future performance instead of regressing over the past performance, we can assume a virtual crack growth history back propagating from the time when shock occurs. This virtual
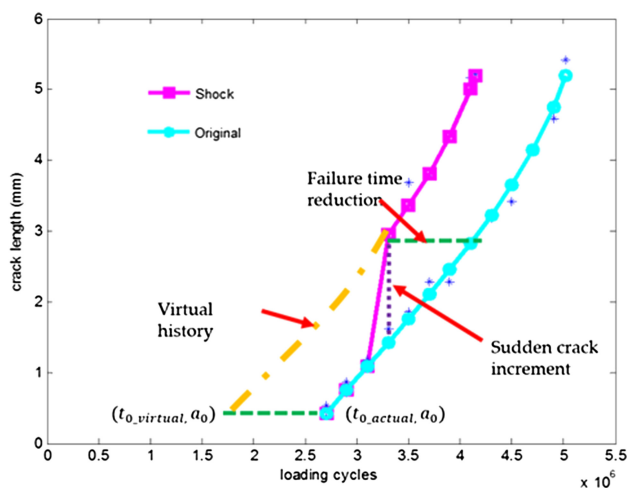

Figure 3 Shock degradation path for illustration purpose

history is indicated in yellow dot-dash line in Figure 3. Therefore, the objective of predicting the degradation after shock occurrence can be achieved by identifying a gradual degradation with a different CIT, to be more specific, an earlier one, denoted by $t_{0\_virtual}$. Based on the foregoing analysis, uncertainty in both CIT and material parameter will be considered to deal with shock degradation. We investigate two cases in the following discussions: shock occurrence time is known and shock occurrence time is unknown.

4.1 Shock Occurrence Time is Known

The update process is executed every time when observation on crack size is available. The posterior distribution in the last iteration will be used as the prior distribution in the next iteration. By assimilating the observations, the variances of uncertainty parameters are expected to be reduced and the mean to approach the actual value. The shrinkage of uncertainty leads to a narrower distribution, which appears to have a peak covering a very small range of parameter values. When we implement the Bayesian inference Eq. (4), importance sampling technique is utilized, where the density is represented by samples from the associated distribution. This discretization of the continuous function excludes the samples with low density, especially when sample size is limited. When the shock occurs, there is a sudden increase in crack size, which results in a noticeable difference between the predicted crack size using degradation model and the observation. In order to attribute this difference to an earlier CIT, it needs a fair extent of adjustment of CIT in its marginal posterior distribution. Hence, the coverage of prior distribution must be large enough to give moderate density for the virtual CIT $t_{0\_virtual}$, otherwise the samples could not carry useful information. Therefore, the posterior distribution before shock occurs is adjusted in a way that its variance of CIT marginal is artificially increased to make sure the vicinity of the targeted virtual CIT $t_{0\_virtual}$ has moderate chance to be sampled.

Let the inspection times be a set $\{t_u : u = 1, 2, \cdots, U\}$. At these inspection times, the observed crack sizes forma set $\{a_u^{obs} : u = 1, 2, \cdots, U\}$ and the uncertain parameters are denoted by $\left\{\left(t_0^{(u)}, \xi^{(u)}\right) : u = 1, 2, \cdots, U\right\}$. Assume that the measurement errors are independent among different inspection times; hence, the likelihood function to observe a crack history up to an inspection time t_s is

$$l_s\left(a_{1:s}^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right) = \prod_{u=1}^{s} l\left(a_u^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right), \qquad (5)$$

where $l\left(a_u^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right)$ was defined in Eq. (3). $a_{1:s}^{obs}$ refers to the set of observation history $\{a_u^{obs} : u = 1, 2, \cdots, s\}$ up to the inspection time t_s. The update process at t_s can be obtained accordingly by

$$f_{post}\left(t_0^{(s-1)}, \xi^{(s-1)}\Big|a_{1:s}^{obs}\right)$$

$$= \frac{l_s\left(a_{1:s}^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right) f_{prior}\left(t_0^{(s-1)}, \xi^{(s-1)}\right)}{\int l_s\left(a_{1:s}^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right) f_{prior}\left(t_0^{(s-1)}, \xi^{(s-1)}\right) dt_0^{(s-1)} d\xi^{(s-1)}}.$$

$$(6)$$

The inspection times divide the whole lifetime into several disjoint intervals. If the degradation is gradually happening without shock, the update of uncertain parameters is to directly assign the posterior distribution of uncertain parameters at current inspection time to the next inspection time. However, if the shock occurs at a known time instant during interval $[t_{v-1}, t_v)$, the posterior distribution will be adjusted before being assigned to the prior distribution for next update. As discussed earlier, the adjustment is to increase the variance of marginal distribution for CIT to cover $t_{0\_virtual}$. Denote the adjusted distribution as $\tilde{f}_{post}\left(t_0^{(v-1)}, \xi^{(v-1)}\right)$. In addition, the likelihood also needs modification. Because of the virtual degradation path, the actual observations on crack sizes before the shock occurs will have adverse effects on the likelihood to observe the crack sizes after the shock. Hence, only the observations after shock will be used to define the likelihood function in Bayesian formula for the updates executed after shock occurrence time. The update process considering the shock occurring at a known time instant thus is modified to be

$$l_s\left(a_{1:s}^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right) = \begin{cases} \prod\limits_{u=1}^{s} l\left(a_u^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right), \text{when } s < v \\ \prod\limits_{u=v}^{s} l\left(a_u^{obs}\Big|t_0^{(s-1)}, \xi^{(s-1)}\right), \text{when } s \geq v \end{cases}$$

$$(7)$$

$$f_{prior}\left(t_0^{(s)}, \xi^{(s)}\right) = \begin{cases} f_{post}\left(t_0^{(s-1)}, \xi^{(s-1)}\right), \text{when } s \neq v \\ \tilde{f}_{post}\left(t_0^{(s-1)}, \xi^{(s-1)}\right), \text{when } s = v \end{cases} \qquad (8)$$

4.2 Shock Occurrence Time is Unknown

If the shock occurs at an unknown time, we could do the proposed adjustment above at every inspection time to ensure that the targeted virtual CIT will be covered within the distribution samples. However, this approach will result in a large amount of uncertainty in RUL prediction, which provides little useful information in maintenance decision making. Hence, an additional step of shock detection is proposed to add into the update process to deal with the case when shock occurrence time is unknown.

Shock occurrence will cause sudden increase of the crack size. The amount of increase is assumed to be far out of the range of measurement error. A large adjustment of CIT from the last inspection time is expected. The average

of predicted crack size using the distribution of CIT at the last inspection time ought to deviate a lot from the observed crack size at the inspection time right after shock occurs. Therefore, a shock is said to be detected if the amount of such deviation exceeds a predefined threshold δ. The criterion is that, shock occurs during $[t_{v-1}, t_v)$ if

$$\left| \mathbb{E}_{\left(t_0^{(v-1)}, \xi^{(v-1)}\right)} \left[a\left(t_v \middle| t_0^{(v-1)}, \xi^{(v-1)}\right) \right] - a_v^{obs} \right| > \delta. \quad (9)$$

The symbol of $\mathbb{E}$ denotes the operator for expectation. After identifying the shock occurrence time, update process Eqs. (7–8) will apply.

4.3 Remaining Useful Life Prediction Updates

The ultimate goal of updating uncertain parameters is to predict the RUL of the cracked gear more accurately. The RUL is defined as the difference between final failure time and current inspection time. The failure time is the instant when the observed crack size reaches or exceeds a predefined critical value a_C for the first time. With the updated parameter distribution obtained through Bayesian inference, the RUL prediction is given accordingly at the inspection time. Paris' law can be written in its reciprocal form as in Eq. (10),

$$\frac{dN}{da} = \frac{1}{C(\Delta K(a))^m}. \quad (10)$$

Let the current inspection cycle be $N_u(t_0)$ and the crack increment be Δa. The RUL is calculated by discretizing Eq.(10) in the following way,

$$\Delta N_i(\xi) = N_{i+1}(\xi) - N_i(\xi) = \Delta a \left[C\left(\Delta s K\left(a_{i/2}\right)\right)^m \right]^{-1}, \quad (11)$$

where $a_{i/2} = a_0 + \Delta a(N_i + N_{i+1})/2$. The RUL is the summation $\sum_i \Delta N_i(\xi)$ from the current inspection cycle to the cycle when failure occurs. Therefore, the total failure time is expressed as $F(t_0, \xi) = N_u(t_0) + \sum_i \Delta N_i(\xi|t_0)$. This expression shows that the uncertainty in total failure time prediction is determined by the uncertainty in both the CIT and the physical model parameters, (t_0, ξ). Thus more accurate values of (t_0, ξ) will produce more accurate $F(t_0, \xi)$. An update of $F(t_0, \xi)$ will be triggered after an update of (t_0, ξ) to adjust the lifetime prediction.

5 Examples

As shown in Figure 1, there are three cases when Bayesian updates are applied. Case 3 can be considered as a generalization of case 2 in terms of an additional step for crack detection. Hence, it is sufficient to only verify case 3 for case 2 to be valid. In this section, we will show the
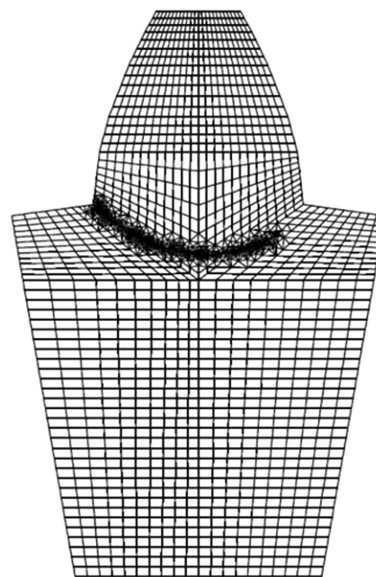

Figure 4 2D FE model for a spur gear tooth [19]

effectiveness of the proposed method for cases 1 and 3: gradual degradation considering CIT and degradation with shock occurring at an unknown time.

A FE model of a 2D spur gear tooth has been built to calculate SIFs when crack is propagating under cyclic fatigue loading, as shown in Figure 4. The geometry and material properties of the gear tooth are tabulate in Table 1. The initial crack size a_0 that can be detected is assumed to be 0.1 mm. The critical crack size a_C is set to be 5.2 mm, which is about 80% of the circular thickness of the tooth. The gear is failed when the measured crack size reaches or exceeds this critical value.

5.1 Case 1: Gradual Degradation Considering CIT

In this case, the prior information for uncertain parameters is assumed as: $m \sim N(1.4, 0.2)$, $t_0 \sim N(2 \times 10^6, 2 \times 10^5)$. The measurement error is $e \sim N(0, 0.15)$. The values of the constants are set to be: $C = 9.12 \times 10^{-11}$, $a_0 = 0.1$ mm, $a_C = 5.2$ mm. The history of SIF is adopted from Ref. [19] in which the input torque is 320 N·m and the effect of dynamic load is considered. The characteristics of two actual degradation paths are shown in Table 2.

We have two test degradation paths, which share the real value of $m = 1.6$, but have different CITs. The characteristics of these two paths are given in Table 2. The two degradation paths are shown in Figures 5 and 6. The crack observations and updated results for the two paths are listed in Tables 3 and 4, respectively. The updated PDFs for distributions of t_0 and m are displayed in Figures 5 and 6 for path #1, and in Figures 7 and 8 for path #2.

The update results show that, as more crack observations are fed into Bayesian inference, the means of m and t_0

Table 1 Material properties, and main geometry parameters [19]

Young's modulus/ Pa	Poisson's ratio	Module/ mm	Diametral pitch/ in^{-1}	Base circle radius/ mm	Outer circle/ mm	Pressure angle/ (°)	Teeth No.
2.07×10^{11}	0.30	3.20	8.00	28.34	33.30	20.00	19

Table 2 Characteristics of two actual degradation paths #1 and #2

Path #	m	t_0 / cycles	Failure time / cycles	Inspection interval / cycles
1	1.6	1.7×10^6	4215500	3×10^5
2	1.6	2.3×10^6	4815500	5×10^5
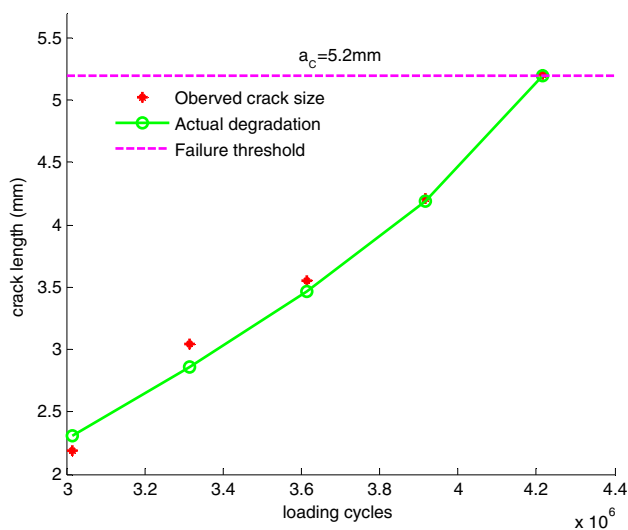
Figure 5 Degradation path #1

Figure 6 Degradation path #2

progressively approach their actual values. Meanwhile, the standard deviations are shrinking, which makes the distribution shape become narrower. This is beneficial for improving the performance of prognostic algorithm since narrower distribution indicates reduced uncertainty, and it is useful in making more accurate and cost-effective maintenance decisions. Based on the updated uncertain parameters, the mean of the failure time approaches its actual value accordingly. The actual values are denoted using star marks in the Figures 7, 8, 9 and 10.

5.2 Case 3: Shock Degradation with an Unknown Shock Occurrence Time

5.2.1 Results with the Proposed Method

In this case, the prior distributions for uncertain parameters are given as: $m \sim N(1.6, 0.2)$, $t_0 \sim N(6.5 \times 10^7, 0.5 \times 10^7)$. The measurement error is $e \sim N(0, 0.15)$ while the values of the constants are set to be $C = 9.12 \times 10^{-11}$,
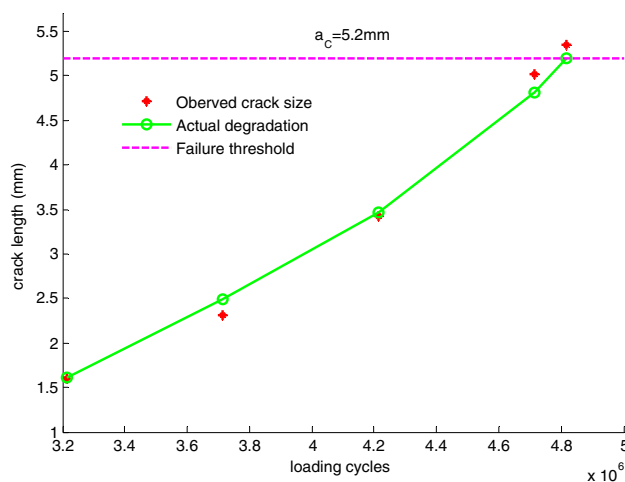
Table 3 Update process for path #1

Update #	Inspection time / cycles	Crack size / mm	Mean of m	Mean of t_0 / cycles	Mean of $F(t_0, m)$ / cycles
Prior			1.4	2×10^6	17159000
1	3015700	2.1900	1.6277	1.9572×10^6	4072300
2	3315700	3.0425	1.6420	2.0105×10^6	3972800
3	3615700	3.5473	1.6275	1.9325×10^6	3991700
4	3915700	4.2072	1.6134	1.8354×10^6	3989700
Actual value			1.6	1.7×10^6	4215500

$a_0 = 0.1$ mm, $a_C = 5.2$ mm. SIF is evaluated using the FE analysis in software FRANC2D, and the computed SIF values versus crack size are shown in Figure 11, in which the input torque is 40 N·m and the effect of dynamic load is not considered. Third-order polynomial is used to fit the

Table 4 Update process for path #2

Update #	Inspection time / cycles	Crack size / mm	Mean of m	Mean of t_0 / cycles	Mean of $F(t_0, m)$ / cycles
Prior			1.4	2×10^6	17159000
1	3215700	1.591	1.5469	1.9827×10^6	5520800
2	3715700	2.3063	1.5545	1.9999×10^6	5358600
3	4215700	3.4307	1.5821	2.1583×10^6	5135000
4	4715700	5.0218	1.6073	2.3604×10^6	5087700
Actual value			1.6	2.3×10^6	4815500

Figure 9 Updated PDF of t_0 for path #2

Figure 7 Updated PDF of t_0 for path #

Figure 10 Updated PDF of m for path #2

Figure 8 Updated PDF of m for path #1

discrete crack size observations to obtain a continuous function for SIF. One actual degradation path, path #3, is used to demonstrate this case, and the characteristics of the degradation path are shown in Table 5. The actual
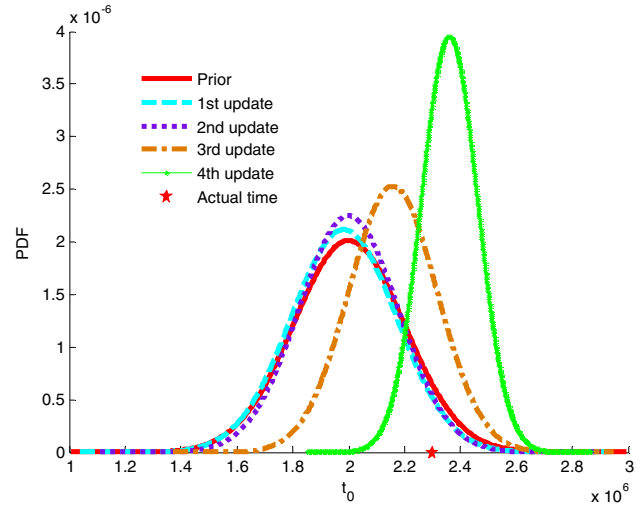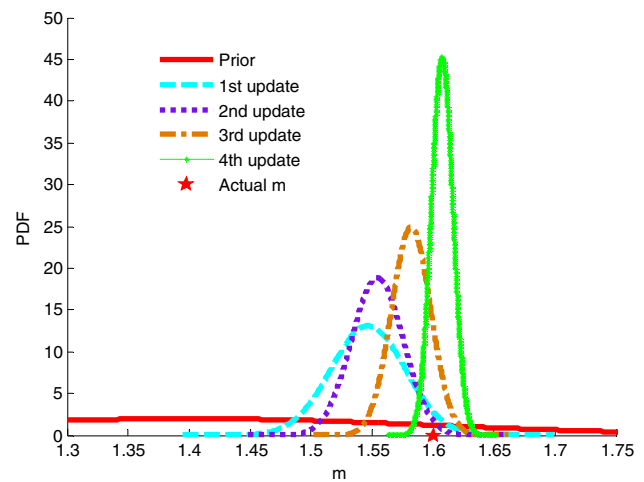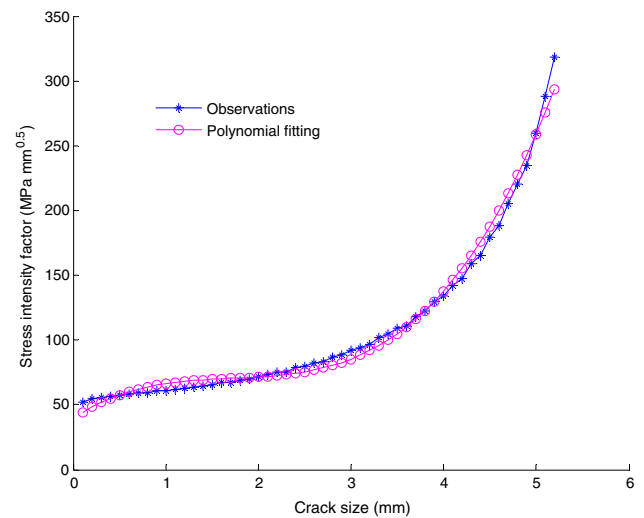

Figure 11 History of stress intensity factor for path #3

Table 5 Characteristics of actual degradation path #3

Path #	m	t_{0_actual} / cycles	$t_{0\_virtual}$ / cycles	Failure time / cycles	Inspection interval / cycles
3	1.4354	7.5×10^7	4.9×10^7	1.474×10^8	1×10^7

Figure 12 Actual shock degradation path #3

Table 6 Update process for path #3

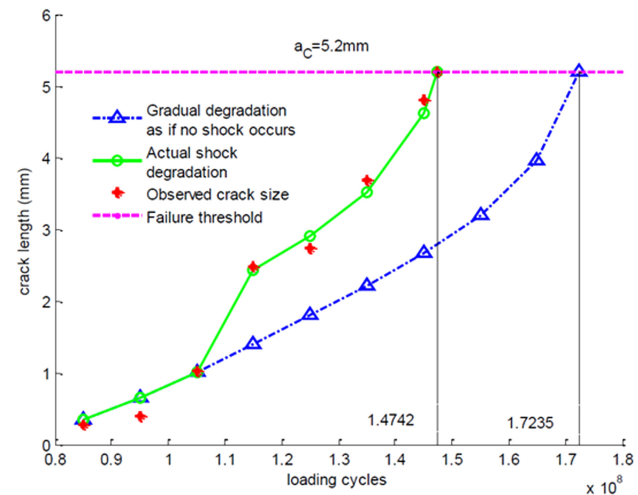| Update #	Inspection time / cycles	Crack size / mm	Mean of m	Mean of t_0 / cycles	Mean of $F(t_0, m)$ / cycles
Prior			1.6	6.5×10^7	1.331×10^8
1	8.5×10^7	0.2775	1.3006	6.7311×10^7	2.639×10^8
2	9.5×10^7	0.3968	1.2382	6.8447×10^7	3.126×10^8
3	10.5×10^7	1.0214	1.3204	6.8454×10^7	2.183×10^8
4	11.5×10^7	2.4759	1.4303	4.7962×10^7	1.480×10^8
5	12.5×10^7	2.7370	1.4233	4.7805×10^7	1.524×10^8
6	13.5×10^7	3.6924	1.4301	4.7735×10^7	1.454×10^8
7	14.5×10^7	4.7976	1.4334	4.8387×10^7	1.465×10^8
Actual value		1.4354		5.0060×10^7	1.474×10^8

Figure 13 Update PDF of m for path #3

degradation path #3 is depicted in Figure 12. The marks of circle represent the actual crack sizes, which are unknown. The marks of star represent the observed crack sizes, which deviate from the corresponding actual values due to measurement error. Shock occurs between Observations #3 and #4, and a sudden increase in crack size can be seen in Figure 12.

The shock could be detected between update #3 and #4 using the criterion proposed in Section 3. The horizontal boarder between these two inspection points in Table 6 separates the updated results into two parts: history before shock occurrence and history after shock occurrence. From Table 6, it is observed a sudden transition of changing pattern for mean of t_0 after the shock is detected. Before the shock occurs, the mean of t_0 tends to increase to reach $t_{0\_actual} = 7.5 \times 10^7$ cycles, which is the actual CIT of the shock degradation path. Accordingly, the mean of failure time $F(t_0, m)$ tends to approach 1.7235×10^8, which is the failure time of a gradual degradation as if no shock occurs, as depicted by dot-dash line with triangle marks in Figure 12. However, after the shock occurs, the mean of t_0 inverses its changing direction to approach $t_{0\_virtual} = 5.006 \times 10^7$ cycles, which is the CIT of the virtual gradual degradation path. Compared to the actual shock degradation, this virtual gradual degradation path compensates the reduction of lifetime due to shock occurrence by an earlier CIT. As a material dependent parameter, there is no such
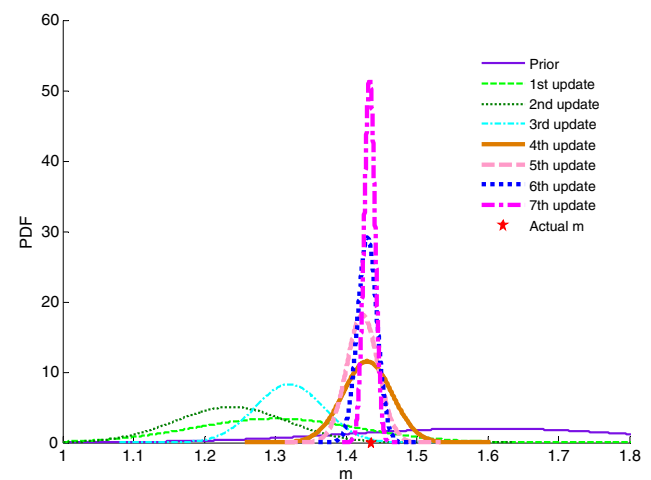
transition in m because its value is dependent on material property, which should not be changed by shock occurrence. Based on the accurate information on parameters t_0 and m, the predicted mean of failure time $F(t_0, m)$ successfully approach the actual failure time in a shock degradation, which is 1.4742×10^8, as depicted by solid line with marks of circle in Figure 12.

Figure 13 shows the updates of PDF for m, in which the bold line indicates the updates after the shock occurs. It can be seen that, as more observations are available, the mean of m approach the actual value. Also the variance of PDF is reduced progressively which provides more precise information. Figure 14 shows the updates of PDF for t_0, in which two neighbourhoods are apparent to observe. One is in the vicinity of $t_{0\_actual} = 7.5 \times 10^7$, denoted using a diamond mark, and the other is in the vicinity of $t_{0\_virtual} = 5.006 \times 10^7$ denoted using a star mark. As discussed
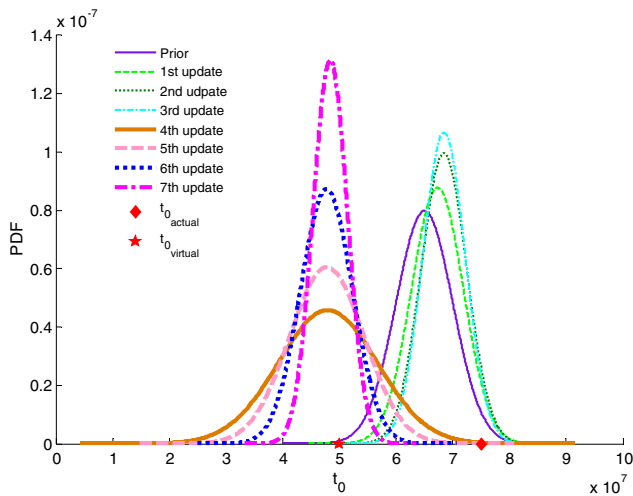
Figure 14 Update PDF of t_0 for path #3

before, this phenomenon explains the adjustment of t_0 due to the shock occurrence, and the mean of t_0 approaches $t_{0\_virtual}$ after the shock occurs.

5.2.2 Comparative Studies

The proposed method is compared with the physics-based method and a data driven method in this section. The example of shock degradation with an unknown shock occurrence time, considered in Section 5.1.1, is used. The results are shown in Figure 15, where the blue line represents the crack length predicted using the physics-based method, and the orange line represents the results obtained using the integrated method. For the physics-based method, the model parameters are kept constant and will not be updated. The model parameters used are obtained from the prior distribution of initial time t and model parameter m.
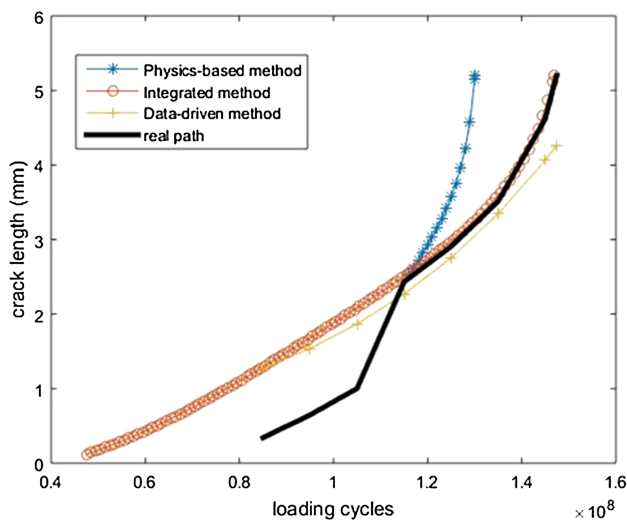

Figure 15 Comparison of degradation paths with different methods

Here, the initial time is 6.5×10^7, and the model parameter $m = 1.6$. The prediction results at inspection point 4 after the shock occurs are shown in the Figure 14.

The integrated method is also compared with a popular type of data driven method based on Bayesian inference and generic degradation models [4]. The degradation model adopted is an exponential degradation model, with the following form:

$$L(t) = a + bt + \varepsilon, \tag{12}$$

where t is the inspection cycle, $L(t)$ is the logarithm of the measurement of the crack length. The prior distributions of the three parameters a, b, ε are $a \sim N(0, 1^2)$, $b \sim N\left(10^{-7}, (0.3 \times 10^{-7})^2\right)$, $\varepsilon \sim N(0, 0.1^2)$ respectively. The results are represented by the yellow line in Figure 15.

According to the results in the figure, the results by the proposed integrated method agree well with the real path after the shock. In contrast, the crack length predicted by the physics-based method has large discrepancy with the measured data since the model parameters are not adjusted based on the observed data. The proposed integrated method performs better due to the more accurate estimating of the model parameters. The data-driven method did not perform as well as the integrated method either, due to its inability to capture the shock occurrence and utilize the real degradation mechanism.

6 Conclusions

Prognostics tools are used for RUL prediction starting from the crack initiation time. However there is uncertainty in CIT due to the limited capability of existing fault detection tools, and such uncertainty has not been explicitly considered in the literature for integrated prognosis. A shock causes a sudden damage increase and creates a jump in the degradation path, which shortens the total lifetime, and it has not been considered before either in the integrated prognostics framework. This paper proposes an integrated prognostics method considering these two important factors, the uncertainty in CIT and the shock in equipment degradation. In the proposed integrated prognostics method, CIT is considered as an uncertain parameter, which is updated using CM data. To deal with the sudden damage increase and reduction of total lifetime, a virtual gradual degradation path with an earlier CIT is utilized, and the effect of shock is captured through identifying an appropriate CIT. Examples demonstrate effectiveness of the proposed method in predicting RUL considering shock and uncertainty in CIT.

References

1. A K S Jardine, D Lin, D Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 2006, 20(7): 1483–1510.

2. G J Vachtsevanos, F Lewis, A Hess, et al. *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. Hoboken: Wiley, 2006.

3. D Banjevic, A K S Jardine, V Makis, et al. A control-limit policy and software for condition-based maintenance optimization. *INFOR: Information Systems and Operational Research*, 2001, 39(1): 32–50.

4. N Z Gebraeel, M A Lawley, R Li, et al. Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions*, 2005, 37(6): 543–557.

5. N Z Gebraeel, M A Lawley. A neural network degradation model for computing and updating residual life distributions. *IEEE Transactions on Automation Science and Engineering*, 2008, 5(1): 154–163.

6. Z Tian, L Wong, N Safaei. A neural network approach for remaining useful life prediction utilizing both failure and suspension histories. *Mechanical Systemsand Signal Processing*, 2010, 24(5): 1542–1555.

7. Z Tian, M J Zuo. Health condition prediction of gears using a recurrent neural network approach. *IEEE Transactions on Reliability*, 2010, 59(4): 700–705.

8. H Liebowitz, E T Moyer. Finite element methods in fracture mechanics. *Computers & Structures*, 1989, 31(1): 1–9.

9. R D Henshell, K G Shaw. Crack tip finite elements are unnecessary. *International Journal for Numerical Methods in Engineering*, 1975, 9(3): 495–507.

10. R S Barsoum. On the use of isoparametric finite elements in linear fracture mechanics. *International Journal for Numerical Methods in Engineering*, 1976, 10(1): 25–37.

11. D G Lewicki, R Ballarini. Gear crack propagation investigations. *Lubrication Science*, 1998, 5(2): 157–172.

12. D G Lewicki, R F Handschuh, L E Spievak, et al. Consideration of moving tooth load in gear crack propagation predictions. *Journal of Mechanical Design*, 2001, 123(1): 118–124.

13. C J Li, H Lee. Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics. *Mechanical Systems and Signal Processing*, 2005, 19(4): 836–846.

14. S Glodež, M Šraml, J Kramberger. A computational model for determination of service life of gears. *International Journal of Fatigue*, 2002, 24(10): 1013–1020.

15. G J Kacprzynski, A Sarlashkar, M J Roemer, et al. Predicting remaining life by fusing the physics of failure modeling with diagnostics. *JOM Journal of the Minerals, Metals and Materials Society*, 2004, 56(3): 29–35.

16. P C Paris, F Erdogan. A critical analysis of crack propagation laws. *Journal of Basic Engineering*, 1963, 85(4): 528–533.

17. A Coppe, R T Haftka, N H Kim, et al. Uncertainty reduction of damage growth properties using structural health monitoring. *Journal of Aircraft*, 2010, 47(6): 2030–2038.

18. D An, J-H Choi, N H Kim. Identification of correlated damage parameters under noise and bias using Bayesian inference. *Structural Health Monitoring*, 2012, 11(3): 293–303.

19. F Zhao, Z Tian, Y Zeng. Uncertainty quantification in gear remaining useful life prediction through an integrated prognostics method. *IEEE Transactions on Reliability*, 2013, 62(1): 146–159.

20. F Zhao, Z Tian, Y Zeng. A stochastic collocation approach for efficient integrated gear health prognosis. *Mechanical Systems and Signal Processing*, 2013, 39(1–2): 372–387.

21. F Zhao, Z Tian, E Bechhoefer, et al. An integrated prognostics method under time-varying operating conditions. *IEEE Transactions on Reliability*, 2015, 64(2): 673–686.

22. S Marble, B P Morton. Predicting the remaining life of propulsion system bearings. *Aerospace Conference*, 2006: 8–pp.

23. E Bechhoefer. A method for generalized prognostics of a component using Paris law. *Annual Forum Proceedings-American Helicopter Society*, 2008, 64: 1460.

24. M E Orchard, G J Vachtsevanos. A particle filtering approach for on-line failure prognosis in a planetary carrier plate. *International Journal of Fuzzy Logic and Intelligent Systems*, 2007, 7(4): 221–227.

25. M E Orchard and G J Vachtsevanos. A particle filtering-based framework for real-time fault diagnosis and failure prognosis in a turbine engine. *Control & Automation, 2007. MED'07. Mediterranean Conference on*, 2007: 1–6.

26. E Zio and G Peloni. Particle filtering prognostic estimation of the remaining useful life of nonlinear components. *Reliability Engineering & System Safety*, 2011, 96(3): 403–409.

27. Z Wang, H-Z Huang, Y Li, et al. An approach to reliability assessment under degradation and shock process. *IEEE Transactions on Reliability*, 2011, 60(4): 852–863.

28. F Mallor, J Santos. Classification of shock models in system reliability. *Monografias del Semin. Matem. Garcıa de Galdeano*, 2003, 27: 405–412.

29. L Jiang, Q Feng, D W Coit. Reliability and maintenance modeling for dependent competing failure processes with shifting failure thresholds. *IEEE Transactions on Reliability*, 2012, 61(4): 932–948.

30. ASTM International. *Standard test method for measurement of fatigue crack growth rates*. ASTM International, 2011.

Fu-Qiong Zhao, received her PhD degree from the *Department of Mechanical Engineering, University of Alberta*, in 2015. Her research is in the field of prognostics, physical modeling, and reliability engineering. E-mail: fuqiong@ualberta.ca

Ming-Jiang Xie, born in 1992, is currently a Ph.D candidate at *University of Alberta, Canada*. His research is focused on prognostics and pipeline integrity management. E-mail: mingjian@ualberta.ca

Zhi-Gang Tian, born in 1978, is currently an Associate Professor at the *University of Alberta, Canada*. He received his PhD degree from the *Universtiy of Alberta, Canada*, in 2007. His research is focused on prognostics, pipeline integrity management, reliability, condition based maintenance, renewable energy systems, etc. Tel: +1-7804922822; E-mail: ztian@ualberta.ca

Yong Zeng, is currently a Professor in the *Concordia Institute for Information Systems Engineering at Concordia University*, Montreal, Canada. He is the Canada Research Chair in Design Science (2004–2014). He received his Ph.D degree in design engineering from the *University of Calgary* in 2001. His research is focused on the modeling and computer support of creative design activities.