Ensemble of Models for Fatigue Crack Growth Prognostics

Hoang-Phuong Nguyen, Jie Liu, Enrico Zio

Ensemble of Models for Fatigue Crack Growth Prognostics

Hoang-Phuong Nguyen[1], Jie Liu[2], Enrico Zio[3,4,5]
[1]Chair on System Science and the Energetic Challenge, EDF Foundation, CentraleSupélec, Université Paris-Saclay, 91192 Gif-sur-Yvette, France
[2]School of Reliability and System Engineering, Beihang University, Beijing 10019, China
[3]Department of Energy, Politecnico di Milano, 20156 Milano, Italy
[4]MINES ParisTech / PSL Université Paris, Centre de Recherche sur les Risques et les Crises (CRC), 06904 Sophia Antipolis, France
[5]Department of Nuclear Engineering, Kyung Hee University, Seoul 02447, South Korea

Corresponding author: Jie Liu (e-mail: liujie805@buaa.edu.cn).

ABSTRACT Various models of fatigue crack growth in different scenarios have been proposed in the literature. Here, we propose a general prognostic framework for tracking crack evolution in equipment undergoing fatigue and predicting the Remaining Useful Life (RUL). The main contribution of this work is to integrate Particle Filtering (PF) and a new ensemble model which combines diverse physical degradation models with respect to their accuracy performance in previous time steps, in order to maximize the overall prediction capability. To validate the effectiveness of the proposed framework, a case study concerning multiple fatigue crack growth degradations is extensively investigated.

INDEX TERMS fatigue crack growth, multiple stochastic degradation, prognostics and health management, remaining useful life, particle filter, dynamic ensemble

Nomenclature

Abbreviations

BWWV	Best-Worst Weighted Vote
EOP	End-Of-Process
IMMPF	Interacting Multiple Model Particle Filter
MAPE	Mean Absolute Percentage Error
MC	Monte Carlo
MSE	Mean Square Error
PBM	Physics-Based Model
PDF	Probability Density Function
PF	Particle Filtering
PPI	Prognostic Performance Indicator
RMSE	Root Mean Square Error
RUL	Remaining Useful Life
SMC	Sequential Monte Carlo
SME	Sample Mean Error
SMeE	Sample Median Error
TWEB	Timeliness Weighted Error Bias

Latin symbols

a	constant of polynomial crack growth model
b	constant of curve fitting model
C	material constant
d	width of the specimen undergoing fatigue crack (mm)
f	state transition function
g	measurement function
h(x)	geometric factor
m	material constant
N	number of fatigue load cycles (cycle)
N_M	number of degradation models
N_P	number of particles
N_S	number of units under test
p	probability distribution
q	importance sampling distribution
RUL_t	actual RUL at time t (cycle)
$\hat{RUL}_t^i$	estimated RUL of the *i*th degradation model at time t (cycle)
$\hat{RUL}_t$	estimated RUL of the ensemble at time t (cycle)
t	time (cycle)
T_t^i	estimated failure time of the *i*th degradation model at time t (cycle)
$w_{est}^{i,t}$	previous estimation accuracy-based output weight of the *i*th degradation model in the ensemble at time t

$w_{pre}^{i,t}$	previous prediction accuracy-based output weight of the *i*th degradation model in the ensemble at time t
$w_{overall}^{i,t}$	overall output weight of the *i*th degradation model in the ensemble at time t
$\tilde{w}_{overall}^{i,t}$	normalized overall output weight of the *i*th degradation model in the ensemble at time t
x	degradation state (mm)
x_{th}	failure threshold (mm)
$\hat{x}_t$	estimated degradation state of the ensemble at time t (mm)
$\hat{x}_t^i$	estimated degradation state of the *i*th degradation model at time t (mm)
$\tilde{x}_{t_p:t}^i$	predicted degradation state of the *i*th degradation model at time t with measurements that are available up to time t_p ($t_p < t$) (mm)
z	measurement (mm)

Greek symbols

α	geometric coefficient of fatigue crack
δ_{est}	time horizon for previous estimates considered (cycles)
δ_{pre}	time horizon for previous predictions considered (cycles)
ΔK	stress intensity factor (MPa$\sqrt{m}$)
$\Delta\sigma$	cyclic stress amplitude (MPa)
Δt	time interval (cycle)
ε	weight coefficient of individual degradation model
ν	measurement noise
ω	state noise

I. INTRODUCTION

The rapid development of technology and computer science is bringing opportunities for industrial systems to evolve smarter and faster, but also more complex. In this fast-changing environment, unanticipated risks and failures which may cause large-scale breakdowns with significant losses in both production and economics, have also increased [1]. To cope with this challenging situation, the development of reliability and health management strategies for preventing components and systems from such unexpected failures are urgently required. Specifically, these strategies aim to monitor health conditions of engineering components, predict their Remaining Useful Lives (RULs) and, ultimately, enable optimal maintenance decisions before the breakdown of the components [2], [3]. In practice, the reliability of equipment usually starts decreasing due to gradual degradation, e.g., delamination [4], fatigue crack [5]–[8], corrosion [9], [10], etc., under periodic cyclic loads and eventually leading to failures. Fatigue crack growth is one of the most frequent degradations leading to components and systems failures in several major industries, including energy [6], [11], automotive [7], aerospace [8], etc. Therefore, the demand of prognostic systems for dealing with fatigue crack growth has recently increased.

To address this issue, Physics-Based Models (PBMs), which utilize the physical knowledge of the degradation for constructing a quantitative analytical model of the equipment behavior, have gained significant attention for fatigue crack growth prognostics [12]–[14]. In [13], a failure prognostic scheme for fatigue crack growth prediction was introduced, which employed a stochastic crack growth model and a Bayesian technique to timely update the equipment degradation state from a sequence of monitored measurements. Other Bayesian-based prognostic approach was presented to estimate the stress intensive range of the degradation model in an online manner [14]. The capability of Bayes theorem was fully exploited for updating knowledge about the current degradation state of the target equipment and the unknown parameters in physical models, when a new measurement becomes available.

Among Bayesian-based prognostic techniques, a sequential Monte Carlo (SMC) method, known as Particle Filtering (PF) method, has become very popular due to its capability of effectively handling non-linear systems and non-Gaussian noises. The key idea behind this method is to represent the posterior distribution of the equipment state by a random set of weighted samples, also called *particles*, and then compute the estimated state based on the particles and their associated weights. This methodology has been widely adopted for state estimation and prediction of crack growth [15]–[17], Lithium-ion batteries [18], [19], PEM fuel cells [20], bearings [21], etc.

On the other hand, the performance of model-based prognostic frameworks for fatigue crack growth largely depends on the choice of the adopted physics-of-failure model [22], [23]. Numerous research on modelling fatigue crack growth have been extensively investigated and developed [5], [24]–[26]. In [24], a comprehensive comparison of stochastic models for fatigue crack growth, including the Markov chain model, the Yang's power law-based model, and a polynomial model, was carried out. The results indicated that each degradation model has its own specific range of applicability, that is, each model is only appropriate to certain degradation processes under certain conditions. To the best knowledge of the authors, there is no general consensus on a prognostic model for fatigue crack growth under different degradation processes. Recently, hybrid and multi-degradation model ensembles have attracted the attention of industrial practitioners and researchers due to their superiority over individual degradation models in terms of higher accuracy and better generalization capability [19], [27]. The fundamental idea of these empirical frameworks is to exploit the diversity of different degradation models, which can offer complementary information about the degradation states to be estimated. In an application of Lithium-ion battery prognostics, an Interacting Multiple Model Particle Filter (IMMPF) has been presented to combine the estimations from three different battery capacity degradation models [27]. The results experimentally indicated that the ensemble approach can yield a promising performance in terms of smaller estimation errors and more accurate predictions than single models.

In this paper, an ensemble-based prognostic approach is presented for predicting the evolution to failure and the RUL of an equipment undergoing fatigue crack growth. To maximize the diversity property of the proposed framework, four stochastic degradation models of fatigue crack growth are considered in this work. Moreover, PF is used to track the crack propagation process with nonlinear and non-Gaussian characteristics and eventually to predict the RUL of the equipment before breakdowns. To further enhance the performance of the proposed framework, a dynamic weighted ensemble strategy is proposed in this paper, based on the previous accuracy performance in degradation state estimation and RUL prediction of each single model in the ensemble. Finally, a set of prognostic performance indicators (PPIs) is employed to validate the prediction capability of the proposed framework.

The rest of this paper is organized as follows. Section 2 introduces the degradation models for fatigue crack and details the proposed prognostic framework. Section 3 describes the illustrative case study and the experimental results of the proposed framework in comparison with individual models. Finally, Section 4 concludes the study.

II. ENSEMBLE-BASED FRAMEWORK FOR FATIGUE CRACK PROGNOSTICS

This section presents the proposed ensemble-based framework for fatigue crack prognostics. Three key issues are addressed: how to select the degradation models for the ensemble; how to use the degradation models for estimating the degradation states and predicting the RUL of the equipment; how to combine the outputs of the individual models for achieving maximum accuracy. Fig. 1 illustrates the flowchart of the proposed prognostic model; more details are given in the following sections.

A. DEGRADATION MODELS FOR FATIGUE CRACK

Diversity is an important aspect to consider in the design of an ensemble modeling framework. To address this issue, four stochastic fatigue crack degradation models are selected for exploiting their diversity in the ensemble: Paris-Erdogan, polynomial, global function-based, and curve fitting models.

1) PARIS-ERDOGAN MODEL

The popular Paris-Erdogan model describes the dynamic evolution of the crack depth x as a function of the load cycle number N as follows [28]:

$$\frac{dx}{dN} = C(\Delta K)^m, \tag{1}$$

where C and m are constants related to the material properties, and ΔK is the Irwin's stress intensity factor defined by [29]:

$$\Delta K = \Delta\sigma\sqrt{\pi x}, \tag{2}$$

where $\Delta\sigma$ is the cyclic stress amplitude. In practice, the statistical variability of the crack growth rate can be addressed by modifying (1) with an intrinsic process stochasticity [30]:

$$\frac{dx}{dN} = e^{\omega} C(\Delta K)^m, \tag{3}$$

where $\omega \sim N(0,\sigma_\omega^2)$ is a white Gaussian noise. For a sufficiently small Δt, the Markov chain state-space model of the degradation state x in (3) can be discretized as follows:

$$x_t = x_{t-1} + e^{\omega} C(\Delta K)^m \Delta t. \tag{4}$$

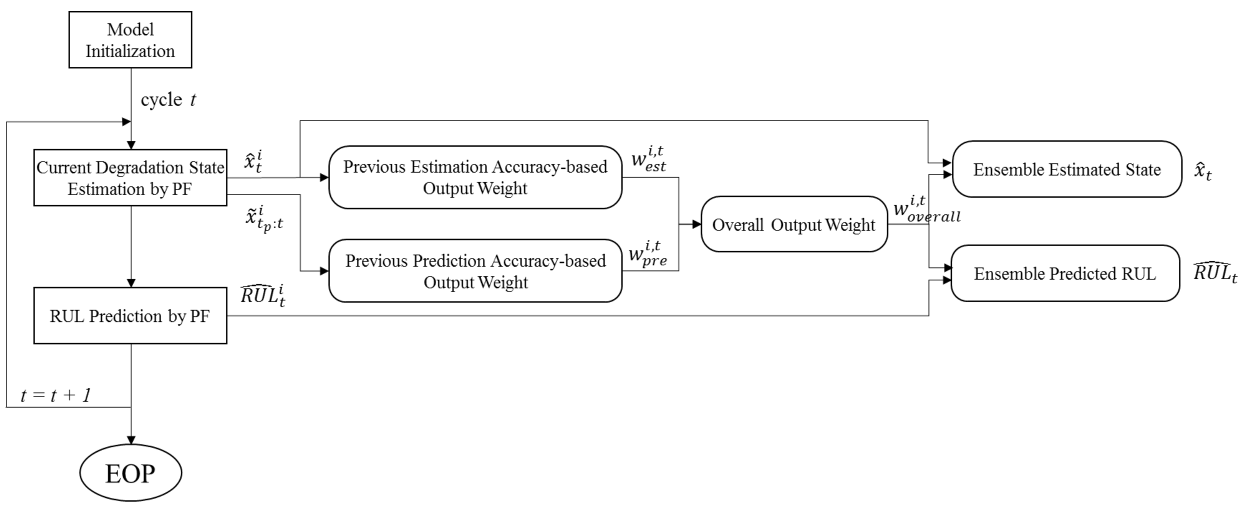


FIGURE 1. Flow diagram of the proposed prognostic framework.

2) POLYNOMIAL MODEL

The polynomial models were first introduced for fatigue crack growth in order to solve the mismatch between the traditional power function-based models, i.e. Paris-Erdogan, and the practical median crack growth curves [24], [31]:

$$\frac{dx}{dN} = e^{\omega}(a_0 + a_1 x + a_2 x^2), \tag{5}$$

where $a_i, i = 0,...,2$ are the second-degree polynomial parameters. Indeed, various works showed that the polynomial models are able to yield the best fit of the linear stage of a degradation process, compared to conventional models [19], [31]. Specifically, the Markov process representation for a polynomial crack growth model can be given as follows:

$$x_t = x_{t-1} + e^{\omega}(a_0 + a_1 x + a_2 x^2)\Delta t. \tag{6}$$

3) GLOBAL FUNCTION

Considering again the Paris-Erdogan model (4) and the fact that fatigue crack growth generally depends not only on material properties but also on equipment geometry, a so-called global function was introduced by reformulating the stress intensity factor [32]:

$$\Delta K = h(x)\Delta\sigma\sqrt{\pi x}, \tag{7}$$

where $h(x)$ denotes the geometric factor of fatigue crack, defined by:

$$h(x) = \alpha_0 + \alpha_1 \frac{x}{d} + \alpha_2 \left(\frac{x}{d}\right)^2 + \alpha_3 \left(\frac{x}{d}\right)^3, \tag{8}$$

where $\alpha_i, i = 0,...,3$ and d are geometric coefficients and the width of the specimen, respectively. The global function-based model for fatigue crack growth can be, then, written as follows:

$$x_t = x_{t-1} + e^{\omega} C(h(x)\Delta\sigma\sqrt{\pi x})^m \Delta t. \tag{9}$$

4) CURVE FITTING FUNCTION

In [32], an empirical crack growth model based on a curve fitting function was presented, which was shown to outperform the conventional models, such as Paris-Erdogan and polynomial models, in terms of higher prediction accuracy and lower computational cost:

$$\frac{dx}{dN} = e^{\omega}\left(\frac{1}{b_1 x^m + b_2}\right), \tag{10}$$

where b_1, b_2 are model constants. The discretized Markov process representation for the model can be given as follows:

$$x_t = x_{t-1} + e^{\omega}\left(\frac{1}{b_1 x^m + b_2}\right)(\Delta K)^m \Delta t. \tag{11}$$

B. DEGRADATION STATE ESTIMATION AND RUL PREDICTION BY PF

In this work, PF is employed to estimate the current degradation state of the equipment and to predict its future evolution until failure. The key idea of PF is based on Bayesian filtering and Monte Carlo (MC) simulation [33]. The basics of the method are recalled in the following sections.

1) CURRENT DEGRADATION STATE ESTIMATION

PF assumes that the state model can be represented as a first-order Markov process, where the current degradation state x_t at time t depends only on its previous state x_{t-1}. The dynamic system process can be described by the following equations:

$$x_t = f_t(x_{t-1}, \omega_{t-1}), \tag{12}$$

$$z_t = g_t(x_t, v_t), \tag{13}$$

where z_t denotes the measurement, ω_t is the state noise sequence, and v_t is the measurement noise sequence at the inspection time $t\,\{t \in \mathrm{N}\}$.

In a Bayesian framework, the system state x_t can be estimated by constructing its posterior probability density function (pdf), $p(x_t \mid z_{1:t})$, via two consecutive steps, namely prediction and update. In the prediction step, the previous state estimation x_{t-1} and the state transition model f_t are utilized to obtain the prior distribution of the system state x_t at current time t via the Chapman-Kolmogorov equation:

$$\begin{aligned} p(x_t \mid z_{1:t-1}) &= \int p(x_t \mid x_{t-1}, z_{1:t-1}) p(x_{t-1} \mid z_{1:t-1}) dx_{t-1}, \\ &= \int p(x_t \mid x_{t-1}) p(x_{t-1} \mid z_{1:t-1}) dx_{t-1} \end{aligned} \tag{14}$$

where $p(x_t \mid x_{t-1})$ is the conditional probability distribution and is defined by the state model in (12). As a new measurement z_t is collected, the required *posterior* distribution of the current state x_t can, then, be obtained by *updating* the prior distribution via Bayes theorem as follows:

$$p(x_t \mid z_{1:t}) = \frac{p(x_t \mid z_{1:t-1}) p(z_t \mid x_t)}{p(z_t \mid z_{1:t-1})}, \tag{15}$$

where $p(z_t \mid x_t)$ is the likelihood function defined by the measurement model in (13) and $p(z_t \mid z_{1:t-1})$ is a normalizing constant given by:

$$p(z_t \mid z_{1:t-1}) = \int p(x_t \mid z_{1:t-1}) p(z_t \mid x_t) dx_t, \tag{16}$$

However, there is usually no analytical solution to (14) and (15) [19]. To address this issue, PF utilizes MC simulation to approximate the true probability distribution with a set of weighted random particles $\{x_t^i, w_t^i, i = 1,.., N_P\}$, where N_P is the total number of particles. In fact, these particles evolve statistically independently of each other, according to the probabilistic state model (12). In this regard, the posterior distribution at time t can be approximated as:

$$p(x_t \mid z_{1:t}) \approx \sum_{i=1}^{N_P} w_t^i \delta\left(x_t - x_t^i\right), \quad (17)$$

where $\delta(.)$ is the Dirac Delta function, often used to represent a discrete distribution as a continuous probability density function $p(x)$:

$$p(x) = \sum_{i=1}^{n} p_i \delta\left(x_t - x_i\right), \quad (18)$$

where $x = \{x_1, ..., x_n\}$ is a discrete distribution with corresponding probabilities $\{p_1, ..., p_n\}$.

In particular, the particle x_t^i is sampled from the importance sampling distribution $q(x_t \mid z_{1:t})$ and its associated weight w_t^i is given by:

$$w_t^i = \frac{p(z_{1:t} \mid x_t^i)p(x_t^i)}{q(x_t^i \mid z_{1:t})}. \quad (19)$$

By setting $q(x_t \mid z_{1:t}) = p(x_t \mid x_{t-1})$ defined in (12), the particle weight w_t^i can be updated with a new collected measurement z_t as follows:

$$w_t^i = w_{t-1}^i p(z_t \mid x_t^i), \quad (20)$$

where $p(z_t \mid x_t^i)$ is the likelihood of measurement z_t given the particle x_t^i. Note that the weights are normalized as $\sum_i w_t^i = 1$.

2) FUTURE DEGRADATION EVOLUTION PREDICTION

Once the posterior distribution $p(x_t \mid z_{1:t})$ of the current degradation state is estimated, it is possible to predict the future degradation evolution and the RUL of the equipment. However, note that there is no available information for estimating the likelihoods of the future degradation states, because future measurements $z_{t+l}, l = 1, ..., T - t$, where T is the time horizon of interest for the analysis, have not been collected yet. The only available information is the dynamic state model (12). Then, the *l*-step ahead posterior distribution $p(x_{t+l} \mid z_{1:t})$ can be written as follows:

$$p(x_{t+l} \mid z_{1:t}) = \int \dots \int \prod_{j=t+1}^{t+l} p(x_j \mid x_{j-1}) p(x_t \mid z_{1:t}) \prod_{j=t}^{t+l-1} dx_j. \quad (21)$$

The numerical evaluation of the integrals in (21) requires significant computational effort. In this paper, an approach presented in [34] is adopted with the assumption that the particle weights do not change from time t to time $t+l$, i.e., $w_t^i = w_{t+1}^i = ... = w_{t+l}^i$. Accordingly, the predicted distribution at time $t+l$ is given by:

$$p(x_{t+l} \mid z_{1:t}) \approx \sum_{i=1}^{N_P} w_t^i \delta(x_{t+l} - x_{t+l}^i), \quad (22)$$

where the particle x_{t+l}^i is obtained by iteratively applying the state model (12) to the corresponding particle of the current state x_t^i.

Finally, the RUL associated to each particle at the present time t can be calculated with reference to the earliest time that the degradation state exceeds the failure threshold x_{th}:

$$R\hat{U}L_t^i = \left\{(T_t^i - 1 - t) \middle| g(x_{T_t^i - 1}, p_t^i, v_t) < x_{th}, g(x_{T_t^i}, p_t^i, v_t) \geq x_{th}\right\} \quad (23)$$

where T_t^i is obtained by simulating the particle evolution via the state model (12). The predicted RUL distribution is, then, given by:

$$p(RUL \mid z_{1:t}, x_i < x_{th}) \approx \sum_{i=1}^{N_P} w_t^i \delta(R\hat{U}L_t - R\hat{U}L_t^i). \quad (24)$$

More details can be found in [35], [36].

C. SELECTIVE ENSEMBLE BASED ON PREVIOUS ESTIMATION AND PREDICTION ACCURACIES

With respect to the way of calculating the weights of the models in an ensemble, the existing ensemble methods can generally be divided into three categories: (a) simple vote ensemble [37], where all individual models outputs are given the same weight coefficients in the voting strategy; in this scheme, majority vote is the most popularly used rule; (b) weighted ensemble [27], which combines individual models with different weight coefficients: each individual is assumed to have a different contribution to the performance of the ensemble model; (c) selective ensemble [38], which includes only an optimal subset of models. This latter method has recently attracted increasing interest, due to its capability of significantly reducing the bias and variance in the ensemble estimation [38].

In this section, we present a selective ensemble approach for prognostics of fatigue crack growth based on a best-worst weighted vote (BWWV) strategy [39]. A novel ensemble weight constructed by using both previous estimation and prediction accuracies of each individual model in the population is proposed.

1) PREVIOUS ESTIMATION ACCURACY-BASED OUTPUT WEIGHT CALCULATION

Suppose that we have a sequence of measurements collected until the current time t, $\{z_j, j = 1, ..., t\}$. The degradation states described by the individual models, $\{\hat{x}_j^i, i = 1, ..., N_M, j = 1, ..., t\}$, where N_M is the number of individual models in the population ($N_M = 4$ in this study), can be estimated by using the PF described in Section 2.2. A weight coefficient of the *i*th model, based on the Root Mean Square Error (RMSE) of its previous estimates with respect to the corresponding measurements, can be calculated as follows:

$$\varepsilon_t^i = \sqrt{\frac{1}{\delta_{est}} \sum_{k=t-\delta_{est}}^{t} (z_k - \hat{x}_k^i)^2}, \quad (25)$$

where δ_{est} is the time horizon of previous estimates considered ($\delta_{est} = 50$ load cycles in the case study that follows).

The previous estimation accuracy-based output weight of each single model is, then, obtained based on the BWWV as follows:

$$w_{est}^{i,t} = 1 - \frac{\varepsilon_t^i - \varepsilon_t^{\min}}{\varepsilon_t^{\max} - \varepsilon_t^{\min}}, \tag{26}$$

where $\varepsilon_t^{\min} = \min_i\{\varepsilon_t^i\}$ and $\varepsilon_t^{\max} = \max_i\{\varepsilon_t^i\}$. By using the BWWV strategy, a maximum weight, $w_{est}^t = 1$, is assigned to the model in the ensemble with highest accuracy at the present time t, and a null weight, $w_{est}^t = 0$, is given to the model with least accuracy, which is equivalent to removing the model from the ensemble for the estimation at time t.

2) PREVIOUS PREDICTION ACCURACY-BASED OUTPUT WEIGHT CALCULATION

Due to the fact that there is no available information from observations to predict the future equipment RUL, the prediction accuracy of each model in the ensemble for the previous time steps is used to calculate the corresponding output weight.

We first identify a time instant t_p before the present time t in the time horizon, where $t = t_p + \delta_{pre}$ ($\delta_{pre} = 100$ load cycles in the following case study), as illustrated in Fig. 2. The state prediction $\tilde{x}_{t_p:t}$ (the dashed line) of one model at time t_p is obtained by iteratively applying the system model to the estimated state $\tilde{x}_{t_p}$, which is set to z_{t_p} in this study. We can now calculate the weight coefficient of the ith model, based on the RMSE of its predictions for degradation states between time t_p and t, with respect to the measurements:

$$\varepsilon_t^i = \sqrt{\frac{1}{\delta_{pre}} \sum_{k=t_p}^{t} (z_k - \tilde{x}_k^i)^2}. \tag{27}$$

Subsequently, the previous prediction accuracy-based output weight of each single model is computed as:

$$w_{pre}^{i,t} = 1 - \frac{\varepsilon_t^i - \varepsilon_t^{\min}}{\varepsilon_t^{\max} - \varepsilon_t^{\min}}, \tag{28}$$

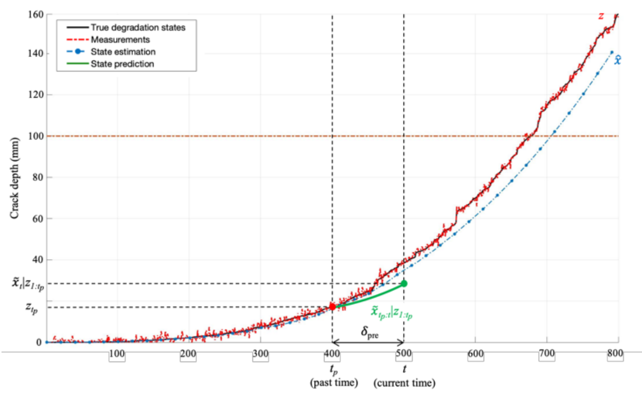


FIGURE 2. Sketch of the previous prediction accuracy-based output weight calculation approach.

3) OUTPUT WEIGHT CALCULATION

Finally, the complete output weight of the ith model in the ensemble at time t is calculated as an average of the previous estimation accuracy-based and the previous prediction accuracy-based weights:

$$w_{overall}^{i,t} = \frac{w_{est}^{i,t} + w_{pre}^{i,t}}{2}. \tag{29}$$

The output weight is, then, normalized as:

$$\tilde{w}_{overall}^{i,t} = \frac{w_{overall}^{i,t}}{\sum_{i}^{N_M} w_{overall}^{i,t}}, \tag{30}$$

Once the output weights for all models are updated, a weighted-sum strategy is used to obtain the degradation state estimation and the RUL prediction of the ensemble as follows:

$$\hat{x}_t = \sum_{i=1}^{N_M} \hat{x}_t^i \times \tilde{w}_{overall}^{i,t}, \tag{31}$$

$$R\hat{U}L_t = \sum_{i=1}^{N_M} R\hat{U}L_t^i \times \tilde{w}_{overall}^{i,t}, \tag{32}$$

where $\hat{x}_t$ and $R\hat{U}L_t$ are the degradation state estimation and the RUL prediction of the proposed ensemble at the present time t, respectively; $R\hat{U}L_t^i$ is the RUL prediction of the ith model in the ensemble.

III. CASE STUDY

A case study of fatigue crack growth is carried out in this work to demonstrate the effectiveness of the proposed method, including crack depth measurements of 100 simulated degradation trajectories, as shown in Fig. 3. The common Paris-Erdogan model in (4) is adopted for describing the evolution of the crack depth with the parameters predefined as follows:

- The model constants are $C = 0.1$ and $m = 1.3$;
- The state and measurement noise variances are

$\sigma_{\omega}^{2}=1.10$ and $\sigma_{v}^{2}=2.25$, respectively.

The crack depths, with a 10^{-4} mm initial length, are recorded every load cycle. The failure threshold is $x_{th}=100$ mm. And the fatigue simulation for each degradation trajectory is performed with a total 800 load cycles.

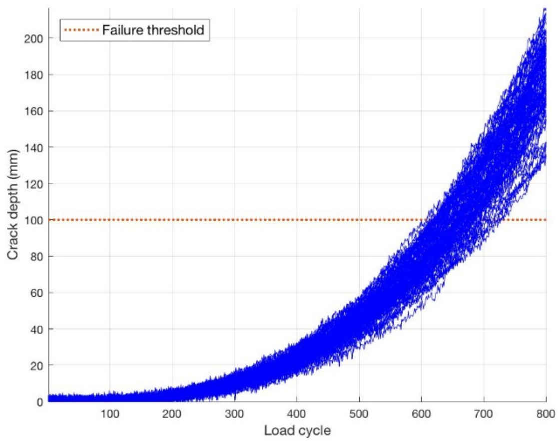


FIGURE 3. **100 fatigue crack growth degradation trajectories.**

1) PERFORMANCE EVALUATION

In this section, the robustness of the proposed ensemble-based prognostic framework is exploited for tracking a fatigue crack growth trajectory and, then, predicting the equipment RUL. The results are compared with four models of fatigue crack growth to validate the improved performance in terms of degradation state estimation and RUL prediction. To evaluate the prognostic framework, five widely used PPIs are considered: a) Timeliness Weighted Error Bias (TWEB); b) Sample Mean Error (SME); c) Mean Absolute Percentage Error (MAPE); d) Mean Square Error (MSE); e) Sample Median Error (SMeE). Details of their definitions are given in Appendix.

When a new measurement is collected, the estimation of the current degradation state for each individual model is also timely updated by using PF as described in Section 2.2. Fig. 4 illustrates the estimation results of four single models over the lifetime of the considered degradation trajectory. The first degradation trajectory from the simulated crack depth dataset described in Section 3.1 is taken. Each model shows a distinctive characteristic in different stages of the degradation evolution of the fatigue crack, which is perfectly suitable for diversity in the proposed ensemble.

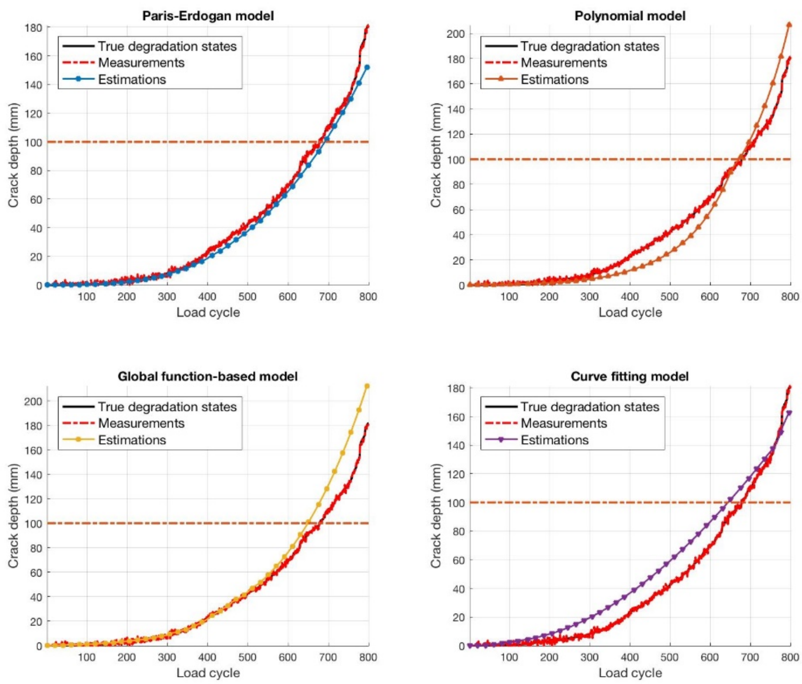


FIGURE 4. **Degradation state estimation for the considered degradation trajectory using individual models.**

Based on the estimations of the individual models, the output weights can be determined and used to update the results of the state estimation and RUL prediction by the proposed ensemble, as shown in Figs. 5 and 6, respectively. As can be seen in Figs. 5 and 6, the individual fatigue crack growth models do not perform very well in the RUL prediction throughout the time horizon considered because of their low accuracy in estimating the current degradation state. In contrast, the proposed approach has a performance which is superior to any individual model throughout the entire life of the equipment, yielding a RUL prediction close to the true RUL.

To further investigate the performance of the proposed method, four different randomly chosen scenarios are considered, whose results are depicted in Figs. 7 and 8. As shown in these figures, the proposed ensemble method consistently exhibits satisfactory performance in estimating the equipment crack growth trend and accurately predicting the RUL. This is due to the proposed prognostic approach which benefits from the diverse accuracy of the individual models by a weighting scheme that can adaptively select the best set of models. Furthermore, in Fig. 8, the confidence intervals show that the RUL prediction accuracy of the proposed method is improved with more available data.

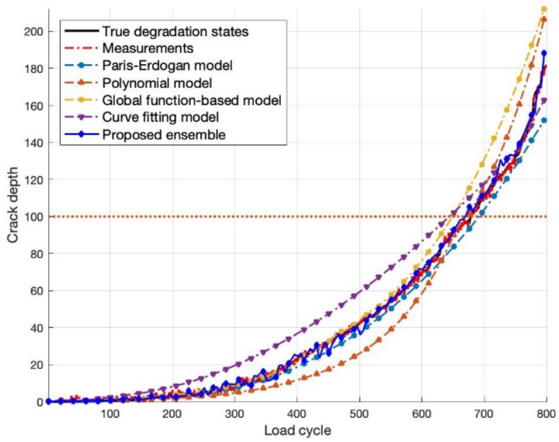


FIGURE 5. Degradation state estimation for the considered degradation trajectory using the proposed ensemble.

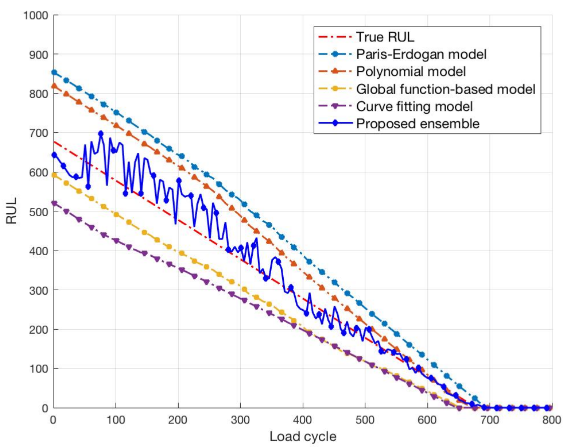


FIGURE 6. RUL prediction for the considered degradation trajectory using the proposed ensemble.

Tables I and II present the average performances in terms of degradation state estimation and RUL prediction, which have been calculated based on 100 crack depth growth scenarios. The results clearly show that the proposed prognostic approach consistently outperforms the individual models for all of the prognostic metrics.

TABLE I
PERFORMANCE COMPARISON IN TERMS OF MSE OF DEGRADATION STATE ESTIMATIONS

	Paris-Erdogan	Poly-nomial	Global function	Curve fitting	Proposed ensemble
Avg. MSE (std)	117.72 (102.68)	166.30 (80.39)	138.64 (74.91)	102.90 (69.38)	8.85 (5.04)

TABLE II
PERFORMANCE COMPARISON IN TERMS OF PPIS OF RUL PREDICTIONS

	TWEB	SME	MAPE	MSE	SMeE
Paris-Erdogan	0.09	115.25	0.62	18.28×10^3	114.63
Polynomial	0.07	85.68	0.37	11.56×10^3	85.43
Global function	0.02	45.79	0.20	3.11×10^3	45.86
Curve fitting	0.03	65.18	0.23	7.01×10^3	64.18
Proposed ensemble	0.01	29.41	0.16	3.03×10^3	31.81

IV. CONCLUSIONS

In this paper, a prognostic modelling framework for fatigue crack growth is proposed. The main original contribution of the work is to combine the PF and a new adaptive ensemble approach, which integrates models of diverse accuracies in previous estimations and predictions for maximizing the generalized prediction performance. The proposed framework is, then, applied to track the degradation evolution and predict the equipment RUL. Various prognostic metrics are employed to evaluate the prediction performance. The results indicate that the proposed ensemble-based prognostic framework outperforms conventional models and is a powerful tool for prognostics of fatigue crack growth.

A limitation of the study is the lack of a real application for validation. Even though several simulation tests were performed to prove the effectiveness of the proposed approach in terms of different PPIs, a real case study of fatigue crack growth is still needed. Further research on addressing this issue with practical applications of fatigue crack can be considered in future work.

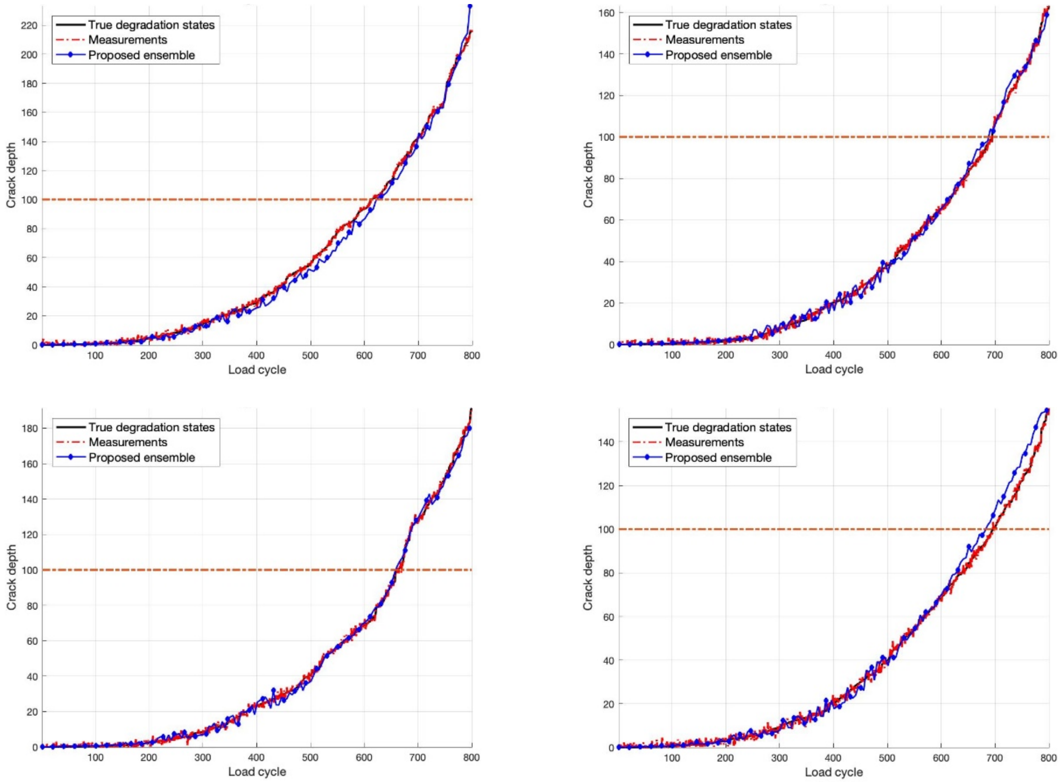


FIGURE 7. **Degradation state estimation using the proposed ensemble with different available measurements.**

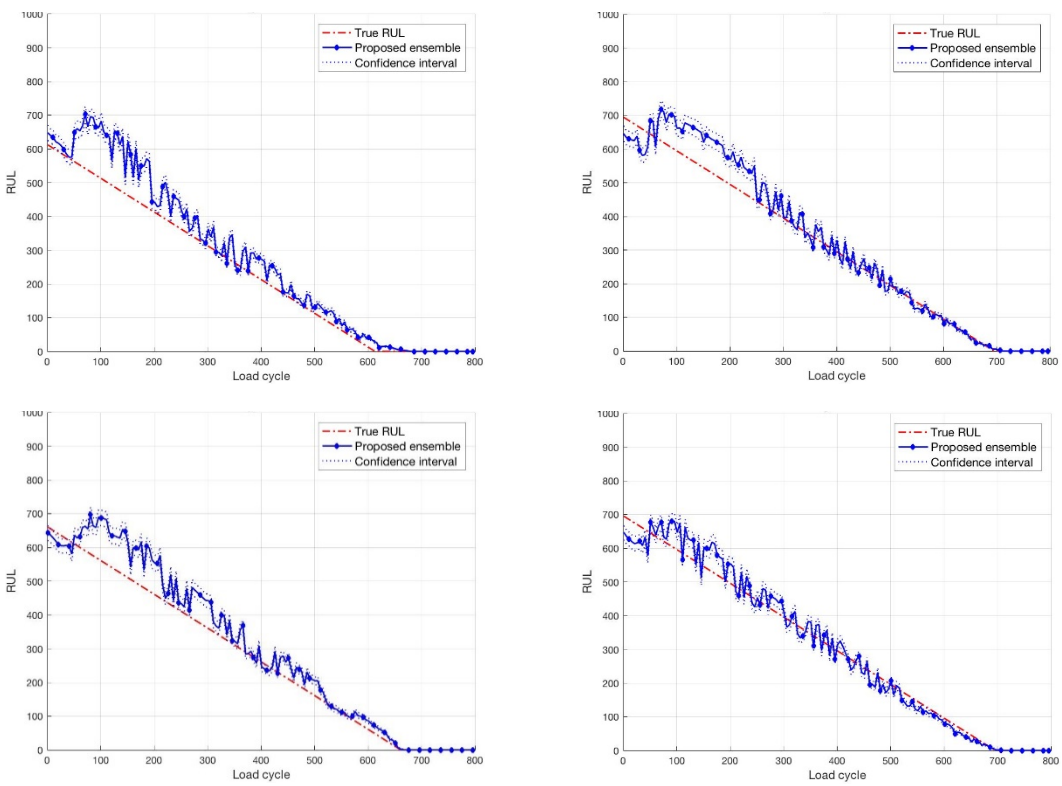


FIGURE 8. **RUL prediction using the proposed ensemble with different available measurements**

APPENDIX

Detailed definitions of the PPIs

Formula	Description
1. Timeliness weighted error bias (TWEB) $$TWEB = \frac{1}{N_S}\sum_{j=1}^{N_S}\phi\left(\sum_{t=1}^{T_i}\gamma_{j,t}\frac{R\hat{U}L_{j,t}-RUL_{j,t}}{T_j}\right)$$ $$\phi(y)=\begin{cases}\exp\left(\frac{\lvert y\rvert}{e_1}\right)-1, & for\ y<0\\ & e_1>e_2>0\\ \exp\left(\frac{\lvert y\rvert}{e_2}\right)-1, & for\ y\geq 0\end{cases}$$	Measure the weighted prediction error over the lifetime T_j by using a penalty function $\phi(y)$ and a weighting function $\gamma_{j,t}$. $\gamma_{j,t}$ is defined as a Gaussian kernel function with a mean value T_j and a standard deviation $0.5T_j$. The optimal value for TWEB is 0, which indicates that the predicted RUL is centered on the true one. Higher values of TWEB indicate a great discrepancy between the predicted RUL and the true one.
2. Sample mean error (SME) $$SME = \left\lvert\frac{1}{N_S}\sum_{j=1}^{N_S}\frac{1}{T_j}\sum_{t=1}^{T_j}\left(R\hat{U}L_{j,t}-RUL_{j,t}\right)\right\rvert$$	Calculate the average errors of all sample points during the lifetime T_j. The optimal value for SME is 0, which indicates that the average errors of all samples is 0, that is, the predicted RUL is centered on the true one. Higher values of SME indicate a great discrepancy between the predicted RUL and the true one.
3. Mean absolute percentage error (MAPE) $$MAPE = \frac{1}{N_S}\sum_{j=1}^{N_S}\frac{1}{T_j}\sum_{t=1}^{T_j}\left\lvert\frac{R\hat{U}L_{j,t}-RUL_{j,t}}{RUL_{j,t}}\right\rvert$$	Measure the average absolute percentage error of all samples throughout the lifetime T_j. The optimal value for MAPE is 0, which indicates a negligible error for all samples during their lifetime. Higher values of MAPE indicate a great discrepancy between the predicted RUL and the true one.
4. Mean square error (MSE) $$MSE = \frac{1}{N_S}\sum_{j=1}^{N_S}\frac{1}{T_j}\sum_{t=1}^{T_j}\left(R\hat{U}L_{j,t}-RUL_{j,t}\right)^2$$	Take into account the average quadratic error of the predicted RUL of all samples during the lifetime T_j. The optimal value for MSE is 0, which indicates that the predicted RUL is equal to the true one for all samples. Higher values of MSE indicate high errors in the predicted RUL.
5. Sample median error (SMeE) $$SME = \left\lvert\underset{j=1,\ldots,N_S}{median}\left(\frac{1}{T_j}\sum_{t=1}^{T_j}\left(R\hat{U}L_{j,t}-RUL_{j,t}\right)\right)\right\rvert$$	Exploit the absolute median of average errors of all samples over the lifetime T_j. The optimal value for SMeE is 0, which indicates that the median error of all samples is zero. Higher values of SMeE indicate that most predicted RULs are wrong.

REFERENCES

[1] E. Zio, "Challenges in the vulnerability and risk analysis of critical infrastructures," *Reliab. Eng. Syst. Saf.*, vol. 152, pp. 137–150, 2016.

[2] E. Zio, "Some Challenges and Opportunities in Reliability Engineering," *IEEE Trans. Reliab.*, vol. PP, no. 99, pp. 1769–1782, 2016.

[3] E. Zio and M. Compare, "Evaluating maintenance policies by quantitative modeling and analysis," *Reliab. Eng. Syst. Saf.*, vol. 109, pp. 53–65, Jan. 2013.

[4] H. Nykyforchyn, O. Zvirko, O. Tsyrulnyk, and N. Kret, "Analysis and mechanical properties characterization of operated gas main elbow with hydrogen assisted large-scale delamination," *Engineering Failure Analysis*, 2017.

[5] J. N. Yang and S. D. Manning, "Stochastic crack growth analysis methodologies for metallic structures," *Eng. Fract. Mech.*, vol. 37, no. 5, pp. 1105–1124, Jan. 1990.

[6] T. Prusek, E. Moleiro, F. Oukacine, A. Adobes, M. Jaeger, and M. Grandotto, "Deposit models for tube support plate flow blockage in Steam Generators," *Nucl. Eng. Des.*, vol. 262, pp. 418–428, Sep. 2013.

[7] H. K. Yang, V. Doquet, and Z. F. Zhang, "Fatigue crack growth in two TWIP steels with different stacking fault energies," *Int. J. Fatigue*, vol. 98, pp. 247–258, 2017.

[8] M. A. Haile, J. C. Riddick, and A. H. Assefa, "Robust Particle Filters for Fatigue Crack Growth Estimation in Rotorcraft Structures," *IEEE Transactions on Reliability*, vol. 65, no. 3. pp. 1438–1448, 2016.

[9] F. Li, Y. Qu, and J. Wang, "Bond life degradation of steel strand and concrete under combined corrosion and fatigue," *Eng. Fail. Anal.*, vol. 80, pp. 186–196, 2017.

[10] A. Vazdirvanidis, G. Pantazopoulos, and A. Rikos, "Corrosion investigation of stainless steel water pump components," *Engineering Failure Analysis*, 2016.

[11] G. Yang, V. Pointeau, E. Tevissen, and A. Chagnes, "A review on clogging of recirculating steam generators in Pressurized-Water Reactors," *Prog. Nucl. Energy*, vol. 97, pp. 182–196, 2017.

[12] M. E. Orchard and G. J. Vachtsevanos, "A particle-filtering approach for on-line fault diagnosis and failure prognosis," *Trans. Inst. Meas. Control*, vol. 31, no. 3–4, pp. 221–246, 2009.

[13] F. Cadini, E. Zio, and D. Avram, "Model-based Monte Carlo state estimation for condition-based component replacement," *Reliab. Eng. Syst. Saf.*, vol. 94, no. 3, pp. 752–758, Mar. 2009.

[14] B. A. Zárate, J. M. Caicedo, J. Yu, and P. Ziehl, "Bayesian model updating and prognosis of fatigue crack growth," *Eng. Struct.*, vol. 45, pp. 53–61, Dec. 2012.

[15] P. Baraldi, M. Compare, S. Sauco, and E. Zio, "Ensemble neural network-based particle filtering for prognostics," *Mech. Syst. Signal Process.*, vol. 41, no. 1–2, pp. 288–300, 2013.

[16] M. E. Orchard, F. A. Tobar, and G. J. Vachtsevanos, "Outer Feedback Correction Loops in Particle Filtering-based Prognostic Algorithms: Statistical Performance Comparison," *Stud. Informatics Control*, vol. 18, no. 4, pp. 295–304, 2009.

[17] L. Tang, J. Decastro, G. Kacprzynski, K. Goebel, and G.

Vachtsevanos, "Filtering and prediction techniques for model-based prognosis and uncertainty management," in *2010 Prognostics and System Health Management Conference, PHM '10*, 2010.

[18] Q. Miao, L. Xie, H. Cui, W. Liang, and M. Pecht, "Remaining useful life prediction of lithium-ion battery with unscented particle filter technique," *Microelectron. Reliab.*, vol. 53, no. 6, pp. 805–810, 2013.

[19] Y. Xing, E. W. M. Ma, K.-L. Tsui, and M. Pecht, "An ensemble model for predicting the remaining useful performance of lithium-ion batteries," *Microelectron. Reliab.*, vol. 53, no. 6, pp. 811–820, Jun. 2013.

[20] M. Jouin, R. Gouriveau, D. Hissel, M. C. Péra, and N. Zerhouni, "Prognostics of PEM fuel cell in a particle filtering framework," *Int. J. Hydrogen Energy*, vol. 39, no. 1, pp. 481–494, 2014.

[21] B. Zhang, C. Sconyers, C. Byington, R. Patrick, M. E. Orchard, and G. Vachtsevanos, "A probabilistic fault detection approach: Application to bearing fault detection," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 2011–2018, 2011.

[22] J. Bogdanoff and F. Kozin, "Probabilistic models of cumulative damage," *Wiley, New York*, Oct. 1985.

[23] W.-F. Wu, "On the Markov approximation of fatigue crack growth," *Probabilistic Eng. Mech.*, vol. 1, no. 4, pp. 224–233, Dec. 1986.

[24] W. F. Wu and C. C. Ni, "Probabilistic models of fatigue crack propagation and their experimental verification," *Probabilistic Eng. Mech.*, vol. 19, no. 3, pp. 247–257, Jul. 2004.

[25] K. Sobczyk and B. F. (Billie F. Spencer, *Random fatigue : from data to theory*. Academic Press, 1992.

[26] M. M. Rocha and G. I. Schuëller, "A probabilistic criterion for evaluating the goodness of fatigue crack growth models," *Eng. Fract. Mech.*, vol. 53, no. 5, pp. 707–731, Mar. 1996.

[27] X. Su, S. Wang, M. Pecht, L. Zhao, and Z. Ye, "Interacting multiple model particle filter for prognostics of lithium-ion batteries," *Microelectron. Reliab.*, vol. 70, pp. 59–69, 2017.

[28] P. Paris and F. Erdogan, "A Critical Analysis of Crack Propagation Laws," *J. Basic Eng.*, vol. 85, no. 4, p. 528, Dec. 1963.

[29] G. R. Irwin, "Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate," *J. Appl. Mech.*, 1957.

[30] E. Myötyri, U. Pulkkinen, and K. Simola, "Application of stochastic filtering for lifetime prediction," *Reliab. Eng. Syst. Saf.*, vol. 91, no. 2, pp. 200–208, Feb. 2006.

[31] C. C. Ni, "Formulation of a Polynomial Stochastic Fatigue Crack Growth Model," *Adv. Mater. Res.*, vol. 909, pp. 467–471, 2014.

[32] H. Salimi, M. Pourgol-Mohammad, and S. Kiad, "Assessment of Stochastic Fatigue Failures Based On Deterministic Functions," in *13th International conference on Probabilistic Safety Assessment and Management (PSAM 13)*, 2016.

[33] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, 2002.

[34] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[35] P. Baraldi, F. Cadini, F. Mangili, and E. Zio, "Model-based and data-driven prognostics under different available information," *Probabilistic Eng. Mech.*, vol. 32, pp. 66–79, 2013.

[36] E. Zio and G. Peloni, "Particle filtering prognostic estimation of the remaining useful life of nonlinear components," *Reliab. Eng. Syst. Saf.*, vol. 96, no. 3, pp. 403–409, 2011.

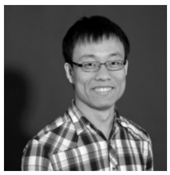
[37] D. W. Opitz and J. W. Shavlik, "Actively Searching for an Effective Neural Network Ensemble," *Conn. Sci.*, vol. 8, no. 3–4, pp. 337–354, Dec. 1996.

[38] Z.-H. Zhou, J. Wu, and W. Tang, "Ensembling neural networks: Many could be better than all," *Artif. Intell.*, vol. 137, no. 1–2, pp. 239–263, May 2002.

[39] F. Moreno-Seco, J. Iñesta, P. De León, and L. Micó, "Comparison of classifier fusion methods for classification in pattern recognition tasks," *Lect. Notes Comput. Sci.*, vol. 4109, pp. 705–713, 2006.

HOANG-PHUONG NGUYEN received the B.Sc. and M.Sc. degree in Computer Engineering from University of Technology, Ho Chi Minh city, Vietnam in 2011 and University of Ulsan, South Korea in 2015, respectively.

Since 2016, he has been working toward the Ph.D. degree at the Chair on System Science and the Energetic Challenge (SSEC), CentraleSupélec, France under the supervision of Prof. Enrico Zio. His current research focuses on prognostics and health management (PHM) methodologies including fault detection and diagnostics, model-based and data-driven prediction models for machine degradation assessment.



JIE LIU received the B.Sc. degree in Mechanical Engineering and the M.Sc. degree in Physics from Beihang University, Beijing, China, in 2009 and 2012, respectively. He received the Ph.D. degree in CentraleSupélec, Chatenay-Malabry, France.

From 2015 to 2017, he was a post-doc researcher in Chair System Science and Energetic Challenges, EDF Foundation, CentraleSupélec, Chatenay-Malabry, France. He works currently at School of Reliability and Systems Engineering of Beihang University, Beijing, China. He has published more than 20 articles. His research interests concern fault detection, diagnostics, prognostics, dynamic reliability assessment.

ENRICO ZIO received the B.Sc. and Ph.D. degree in nuclear engineering from Politecnico di Milano, Italy, in 1991 and 1995, respectively; the M.Sc. degree in mechanical engineering from UCLA, USA in 1995; and a Ph.D. in nuclear engineering from MIT in 1998. He is the Director of the Chair on Systems Science and the Energetic Challenge, EDF Foundation, CentraleSupélec, Chatenay-Malabry, France, and the full Professor of Computational Methods for Safety and Risk Analysis in Politecnico di Milano, Italy. He is President of Advanced Reliability, Availability and Maintainability of Industries and Services (ARAMIS) ltd., and Chairman of the European Safety and Reliability Association, ESRA. His research topics are: analysis of the reliability, safety and security of complex systems under stationary and dynamic conditions, particularly by Monte Carlo simulation methods; development of soft computing techniques (neural networks, support vector machines, fuzzy and neuro-fuzzy logic systems, genetic algorithms, differential evolution) for safety, reliability and maintenance applications, system monitoring, fault diagnosis and prognosis, and optimal design. He is co-author of five international books and more than 200 papers on international journals, and serves as referee of more than 20 international journals.