First-Order Uncertain Hidden Semi-Markov Process for Failure Prognostics With Scarce Data

JIE LIU, (Member, IEEE)
School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
e-mail: liujie805@buaa.edu.cn

This work was supported in part by the Fundamental Research Funds for the Central Universities.

ABSTRACT Failure prognostics aims at predicting the object equipment’s future degradation trend and derives the remaining useful life with a predefined failure threshold. Hidden semi-Markov process (HSMP) is widely adopted for failure prognostics of degradation process with discrete states. The effective estimation of the holding time distribution on each degradation state is of critical importance for the prediction performance of a HSMP model. The distributions are generally estimated with frequencies-based probabilities given a large amount of degradation data. In practical engineering applications, it is difficult to collect enough data and the data can even be scarce. In such situations, the estimated distributions are no longer reliable. First-order uncertain hidden semi-Markov process (1-UHSMP) based on uncertain statistics is defined in this work. The holding time distributions in 1-UHSMP are described with uncertainty theory and are adaptively updated with conditional uncertainty distributions given observations related to the true degradation states. Analytical expressions are derived for the expected remaining useful life for 1-UHSMP with regular uncertainty distributions, i.e. normal and linear uncertainty distribution. The proposed method can build a degradation model from scarce data and derive adaptively the remaining useful life with associated uncertainty interval. A case study concerning centrifugal pumps in a nuclear power plant is considered to verify the effectiveness of 1-UHSMP.

INDEX TERMS Failure prognostics, hidden semi-Markov process, scarce data, uncertainty theory.

I. INTRODUCTION

Failure prognostics aims at predicting the future degradation evolvement of an object equipment and, thus, its remaining useful life given its intended functions with the desired specifications [1]. Remaining useful life (RUL) is the length of time from the present time to the expected time at which the object equipment will no longer perform its intended function. Effective failure prognostics may increase the maintenance efficiency and equipment availability, thus, reducing the life cycle cost [2].

Many efforts have been devoted to the methodological research. The developed prognostic approaches can be mainly categorized as physics-of-failure (PoF) based and data-driven methods [3]. PoF-based methods model the degradation process by analyzing the inherent physical, chemical and structural interactions of the object equipment. Data-driven methods extract the statistical relations between the monitoring data and the RUL from the collected data. Since equipment are becoming more and more integrated and complex, it is quite difficult and time-demanding for building a precise PoF-based model. With the success of computer science, data-driven prognostic methods have become a major trend in failure prognostics-related research. The commonly adopted data-driven methods include auto-regressive model [4], Bayesian methods [5], Markov process [6], artificial neural networks [7], support vector machines [8], deep learning [9], etc.

Although data-driven methods have achieved satisfactory results in many laboratory applications, much efforts are still needed to bridge the gaps between the theories and industrial applications in failure prognostics. Unlike laboratory applications, a major gap lies between the demand of large amount of degradation data for training an effective data-driven model
and the limited informative degradation data collected from equipment under operation [10]. The reasons for data scarcity are diverse, including limited number of deployed equipment as satellites, conservative maintenance strategies as in nuclear power plants, low signal-to-noise ratio in collected data from equipment as airplanes.

For a degradation process with discrete states, the hidden semi-Markov process (HSMP) is among the most adopted ones for failure prognostics [11]. Figure 1 shows the structure of a first-order HSMP for a degradation process with \( N \) + 1 discrete states. In the figure, \( S_N \) represents the healthy state and \( S_0 \) is the failure state. Along with each degradation state \( S_i \), a holding time distribution \( \Phi_i \) is assigned to describe the possible holding times on the corresponding state, with \( i = 1, \ldots, N \). The lifetime of a new equipment following this degradation process is the sum of the holding time on all possible states \( S_j \) with \( i = 1, \ldots, N \). The time series data \( z_t \) are the observations related to the degradation states of the object equipment. True distributions of holding times on different states are hardly known and are usually approximated with empirical ones based on the collected observations.

The failure prognostics with first-order HSMP can be divided into two parts, i.e. degradation process modelling and RUL prediction. For degradation process modelling, these holding time distributions \( \Phi_i \) are estimated with the degradation data collected from similar equipment. The derived first-order HSMP model describes an average degradation process of this type of equipment. In the second step of RUL prediction for an object equipment, the observations \( z_t \) related to its true degradation states are used to update the estimated holding time distributions \( \Phi_i \), to approximate the true RUL of the object equipment. Results on failure prognostics with first-order HSMP have been reported on hydraulic pumps [12], bearings [6], high-speed milling [13], fuel cells [14], railway turnout systems [15], reciprocating compressors [16], turbofan engines [17], etc.

In these previous works, the holding time distribution on a specific state is estimated with frequencies-based probabilities and updated with Bayes theorem, given a large amount of degradation data. However, it is not always possible to collect enough data, and the data can even be scarce in many industrial applications, such as aerospace, nuclear power plants, etc. For scarce data, frequencies-based probabilities may no longer approximate the true holding time distributions. Considering the scarce data, first-order uncertain hidden semi-Markov process (1-UHSMP) is proposed in this work. A first-order UHSMP is the conventional first-order HSMP with the holding time distributions described by uncertainty distributions in uncertainty theory.

Uncertainty theory capable of tackling statistics with scarce data is firstly proposed in [18]. For distribution estimation, instead of using the statistical frequencies from large amount of data, belief degrees are assigned to possible events in uncertainty theory, describing the chance that these events may occur. The calculation of belief degrees does not rely on a sufficiently large amount of data. Thus, uncertainty theory is more suitable for situations with limited/scarc data for statistical analysis. It has been introduced to various domains, such as reliability analysis [19], [20], risk analysis [21], supply chain [22], accelerated degradation testing [23], data development analysis [24], etc.

Facing the scarcity of degradation data, uncertainty theory is integrated with first-order HSMP for failure prognostics in this work. Uncertain graph is proposed in [25] and its characteristics are analyzed in recently published works [26]–[28]. To the knowledge of the author, this is the first time that 1-UHSMP is proposed and adopted for failure prognostics. An update strategy for 1-UHSMP is also given for adaptively adjust the holding time distributions with respect to the observations and, thus, the RUL. First-order UHSMP for failure prognostics takes the same structure as in Figure 1, while the holding time distributions are estimated and updated within the scope of uncertainty theory. An application in nuclear power plants is considered to verify the effectiveness of the proposed model.

The contributions of this work include: 1) the definition of first-order UHSMP; 2) the derivations of predicted RUL and the associated prediction interval with respect to 1-UHSMP; 3) the application of 1-UHSMP in nuclear power plants.

The remaining of the paper is structured as follows. Following a brief review of uncertainty theory, the UHSMP, especially the 1-UHSMP is described in details and the expected value of the predicted RUL is derived for some regular uncertainty distributions. Results on the case study concerning centrifugal pumps in a nuclear power plant are reported, with comparison with conventional first-order HSMP based on probability theory. Some conclusions are drawn in Section 4.

II. METHODOLOGY

A. PRELIMINARIES ON UNCERTAINTY THEORY

In first-order UHSMP, the holding time distribution is estimated with uncertain statistics and updated with the conditional uncertain distribution. To explain the proposed method, a brief introduction to the uncertainty theory is given in this section.

Let \((\Gamma, \mathcal{L})\) be a measurable space and \(\mathcal{M}\) be an uncertain measure on the \(\sigma\)-algebra \(\mathcal{L}\), the following four axioms form the corner stones of the uncertainty theory [18]:

**Axiom 1.** (Normality Axiom) \(\mathcal{M}(\Gamma) = 1\) for the universal set \(\Gamma\).

**Axiom 2.** (Duality Axiom) \(\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1\) for any event \(\Lambda\) in \(\mathcal{L}\).
Axiom 3. (Subadditivity Axiom) For every countable sequence of events \( \Lambda_1, \Lambda_2, \ldots, \) we have
\[
\mathcal{M} \left( \bigcup_{i=1}^{\infty} \Lambda_i \right) \leq \sum_{i=1}^{\infty} \mathcal{M} \{ \Lambda_i \}
\]

Axiom 4. (Product Axiom) Let \((\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)\) be an uncertainty space for \( k = 1, 2, \ldots \). The product uncertain measure \( \mathcal{M} \) is an uncertain measure satisfying
\[
\mathcal{M} \left( \prod_{k=1}^{\infty} \Lambda_k \right) = \bigwedge_{k=1}^{\infty} \mathcal{M}_k \{ \Lambda_k \}
\]
where \( \Lambda_k \) are arbitrarily chosen events from \( \mathcal{L}_k \) for \( k = 1, 2, \ldots \), respectively.

Definition 1: The uncertainty distribution \( \Phi_1 \) of an uncertain variable \( \xi \) is defined by
\[
\Phi_i (x) = \{ \mathcal{M} \{ \xi_i \leq x \} \}
\]
for any real number \( x \) and \( i = 1, \ldots, N \).

In the proposed UHSMP, the uncertain variable \( \xi \) is the holding time on the \( i \)-th state. \( \Phi_i (x) \) represents the uncertainty that the holding time \( \xi_i \) is smaller than \( x \).

Definition 2: An uncertain variable \( \xi \) is called linear if it has a linear uncertainty distribution
\[
\Phi (x) = \begin{cases} 
0, & \text{if } x \leq a \\
\frac{x-a}{b-a}, & \text{if } a < x \leq b \\
1, & \text{if } x > b
\end{cases}
\]
denoted by \( L(a,b) \) where \( a \) and \( b \) are real numbers with \( a < b \).

Definition 3: An uncertain variable \( \xi \) is called normal if it has a normal uncertainty distribution
\[
\Phi (x) = \left( 1 + \exp \left( \frac{\pi (e-x)}{\sqrt{3} \sigma} \right) \right)^{-1}, \quad x \in \mathbb{R}
\]
denoted by \( N(e, \sigma) \) where \( e \) and \( \sigma \) are real numbers with \( \sigma > 0 \). A normal uncertainty distribution is called standard if \( e = 0 \) and \( \sigma = 1 \).

Definition 4: An uncertainty distribution \( \Phi (x) \) is said to be regular if it is a continuous and strictly increasing function with respect to \( t \) at which \( 0 < \Phi ('x) < 1 \), and
\[
\lim_{x \to +\infty} \Phi (x) = 1, \quad \lim_{x \to -\infty} \Phi (x) = 0.
\]

Definition 5: Let \( \xi \) be an uncertain variable with regular uncertainty distribution \( \Phi (x) \). Then the inverse function \( \Phi^{-1}(\alpha) \) is called the inverse uncertainty distribution of \( x \).

Definition 6: Let \( \xi \) be an uncertain variable. Then the expected value of \( \xi \) is defined by
\[
E[\xi] = \int_{0}^{\infty} \mathcal{M}[\xi \geq x] dx - \int_{-\infty}^{0} \mathcal{M}[\xi < x] dx
\]
provided that at least one of the two integrals is finite.

Definition 7: The conditional uncertainty distribution \( \Phi \) of an uncertain variable \( \xi \) given \( A \) is defined by
\[
\Phi (x \mid A) = \mathcal{M} \{ \xi \leq x \mid A \}
\]
provided that \( \mathcal{M} \{ A \} > 0 \).

Theorem 1: Let \( \xi_1, \xi_2, \ldots, \xi_N \) be independent uncertain variables with regular uncertainty distribution \( \Phi_1, \Phi_2, \ldots, \Phi_N \), respectively. If \( f \) is a continuous and strictly increasing function, then
\[
\xi = f(\xi_1, \xi_2, \ldots, \xi_N)
\]
has an inverse uncertainty distribution
\[
\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha), \Phi_2^{-1}(\alpha), \ldots, \Phi_N^{-1}(\alpha)
\]
The sum
\[
\xi = \xi_1 + \xi_2 + \ldots + \xi_N
\]
has an inverse uncertainty distribution
\[
\Psi^{-1}(\alpha) = \Phi_1^{-1}(\alpha) + \Phi_2^{-1}(\alpha) + \ldots + \Phi_N^{-1}(\alpha)
\]

Theorem 2: Let \( \xi_1, \xi_2, \ldots, \xi_N \) be independent uncertain variables with regular uncertainty distribution \( \Phi_1, \Phi_2, \ldots, \Phi_N \), respectively. If \( f \) is a continuous and strictly increasing function with respect to \( x_1, x_2, \ldots, x_m \) and strictly decreasing with respect to \( x_{m+1}, x_{m+2}, \ldots, x_n \), then
\[
\xi = f(\xi_1, \xi_2, \ldots, \xi_N)
\]
has an uncertainty function
\[
\Psi (x) = \min_{f(x_1, x_2, \ldots, x_N) = x} \left( \Phi_1(x_1) \cap \min_{1 \leq i \leq m} (1 - \Phi_1(x_i)) \right).
\]

Theorem 3: Let \( \xi \) be an uncertain variable with regular uncertainty distribution \( \Phi (x) \). Then
\[
E[\xi] = \int_0^1 \Phi^{-1}(\alpha) d\alpha.
\]

Theorem 4: Let \( \xi \) be an uncertain variable with regular uncertainty distribution \( \Phi (x) \), and let \( t \) be a real number with \( \Phi (x) < 1 \). Then the conditional uncertainty distribution of \( \xi \) given \( \xi > t \) is
\[
\Phi (x \mid (t, +\infty)) = \begin{cases} 
0 & \text{if } \Phi (x) \leq \Phi(t) \\
\frac{\Phi (x)}{1 - \Phi(t)} & \text{if } \Phi (t) < \Phi (x) \leq (1 + \Phi (t)) / 2 \\
\frac{\Phi (x) - \Phi(t)}{1 - \Phi(t)} & \text{if } (1 + \Phi (t)) / 2 < \Phi (x)
\end{cases}
\]

Theorem 5: Let \( f \) and \( g \) be comonotonic functions. Then, for any uncertain variable \( \xi \), we have
\[
E[f(\xi) + g(\xi)] = E[f(\xi)] + E[g(\xi)].
\]
\[ \theta_i \text{ in the uncertainty distribution } \Phi_i \text{ of } \xi_i \text{ are estimated by minimizing the following objective function } Q. \]
\[ Q = \min_{\theta_i} \sum_{j=1}^{n} (\Phi_i(x_{ij}) - \alpha_{ij})^2 \]  

By the end of the first step, the estimated uncertainty distributions \( \hat{\Phi}_i \) of all uncertain variables \( \xi_i \) are obtained, with \( i = 1, \ldots, N \).

2) RUL PREDICTION WITH FIRST-ORDER UHSMP

Considering a first-order UHSMP, the second step is to adaptively update the uncertainty distribution of the predicted RUL and to derive the RUL prediction results including the point estimation and its associated prediction interval.

With the observations \( z_{1:t_0} \) until the current time instance \( t_0 \), suppose the object equipment is on the \( k \)-th degradation state and it has been on this state for a time \( T_{t_0}^k \) which can be derived from \( z_{1:t_0} \), the uncertain variable \( \xi \) of the RUL is expressed as the following equation, with \( \xi_i \) being the holding time on the \( i \)-th degradation state, \( i = 1, \ldots, k \).

\[ \hat{\xi} = \sum_{i=1}^{k} \xi_i - T_{t_0}^k \]  

With Theorem 2, the uncertainty distribution \( \hat{\Psi}(x) \) of the uncertain variable \( \hat{\xi} \), is estimated as Equation (9).

\[ \hat{\Psi}(x) = \sup_{\sum_{k=1}^{k} x_k = x} \left( \sup_{0 \leq i \leq k-1} \min_{j \geq T_{t_0}^i} \hat{\Phi}_i(x) \right) \]  

Note that, in this equation, the uncertainty distributions \( \hat{\Phi}_i \) of uncertain variable \( \xi_i, i = 1, \ldots, k - 1 \) can be adopted directly from Section 2.2.1, while the uncertainty distribution \( \hat{\Phi}_k \) of \( \xi_k \) is no longer \( \hat{\Phi}_k \) in Section 2.2.1. And \( \hat{\Phi}_k^* \) is the conditional uncertainty distribution given \( T_{t_0}^k \), i.e. \( \hat{\Phi}_k^*(x) = \Phi_k(x | T_{t_0}^k, +\infty) \).

Since the uncertainty distribution \( \hat{\Phi}_k \) obtained from the first step follows a regular uncertainty distribution, with respect to Theorem 4 the conditional uncertainty distribution is expressed as follows.

\[ \hat{\Phi}_k^*(x) = \hat{\Phi}_k(x | T_{t_0}^k, +\infty) \]

\[ = \begin{cases} 
0 & \text{if } \hat{\Phi}_k(x) \leq \hat{\Phi}_k(T_{t_0}^k) \\
1 - \Phi_k(T_{t_0}^k) \wedge \hat{\Phi}_k(x) & \text{if } \hat{\Phi}_k(T_{t_0}^k) < \hat{\Phi}_k(x) \leq \left( 1 + \hat{\Phi}_k(T_{t_0}^k) \right) / 2 \\
\hat{\Phi}_k(x) - \Phi_k(T_{t_0}^k) / 2 & \text{if } \left( 1 + \hat{\Phi}_k(T_{t_0}^k) \right) / 2 < \hat{\Phi}_k(x) 
\end{cases} \]  

It is obvious that \( \hat{\Phi}_k^*(x) \) is not always a regular uncertainty distribution with the update process in Equation (7), as it may not be a continuous and strictly increasing function with respect to \( x \) at which \( 0 < \hat{\Phi}_k^*(x) < 1 \). However, Theorem 2
requires that all uncertainty distributions should be regular. Thus, Equation (7) should be slightly perturbed such that it becomes regular. The small perturbations are added with respect to two separate cases.

**Case 1:** \( \hat{\Phi}_k \left( T_{0}^k \right) \geq 1/3 \)

When \( \hat{\Phi}_k \left( T_{0}^k \right) \geq 1/3, \hat{\Phi}_k^* \left( x \right) = 0.5 \) for \( \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \left( 1 + \hat{\Phi}_k \left( T_{0}^k \right) \right) / 2 \). Equation (7) can be rewritten as the following equation, with \( \epsilon_1 \) and \( \epsilon_2 \) being small positive constants (8), as shown at the bottom of this page.

**Case 2:** \( \hat{\Phi}_k \left( T_{0}^k \right) < 1/3 \).

When \( \hat{\Phi}_k \left( T_{0}^k \right) < 1/3, \hat{\Phi}_k^* \left( x \right) = \frac{\hat{\Phi}_k \left( x \right)}{1 - \hat{\Phi}_k \left( T_{0}^k \right)} \) for \( \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \left( 1 + \hat{\Phi}_k \left( T_{0}^k \right) \right) / 2 \). Equation (7) can be rewritten as the following equation, with \( \epsilon_2 \) being a small positive constant (9), as shown at the bottom of this page.

The prediction results for the current time include the predicted RUL, i.e. the expected value \( E \left[ \xi \right] \), and the associated prediction interval under a given confidence level \( \alpha_0 \).

With respect to **Definition 5, Theorems 3 and 5**, the predicted RUL value \( \hat{y} \) is given in the following equation.

\[
\hat{y} = E \left[ \xi \right] = \sum_{i=0}^{k} E \left( \xi_i \right) - T_{0}^k
\]

The inverse uncertainty distribution \( \hat{\Phi}_k^{n-1} \left( \alpha \right) \) is given in the following two equations for Case 1 with \( \hat{\Phi}_k \left( T_{0}^k \right) \geq 1/3 \) and Case 2 with \( \hat{\Phi}_k \left( T_{0}^k \right) < 1/3 \), respectively (11–12), as shown at the bottom of the next page.

Now, the expected values of the conditional uncertainty distribution for two popular regular uncertainty distributions, i.e. normal and linear uncertainty distribution, are given in the following two theorems.

\[
\hat{\Phi}_k^* \left( x \right) = \begin{cases} 
0 & \text{if } \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) \\
\frac{\hat{\Phi}_k \left( x \right) - \hat{\Phi}_k \left( T_{0}^k \right)}{0.5 - \epsilon_2} & \text{if } \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) + \epsilon_1 \\
0.5 - 1 + \hat{\Phi}_k \left( T_{0}^k \right) - 2 \hat{\Phi}_k \left( x \right) & \text{if } \hat{\Phi}_k \left( T_{0}^k \right) + \epsilon_1 < \hat{\Phi}_k \left( x \right) \leq \frac{1 + \hat{\Phi}_k \left( T_{0}^k \right)}{2}, \epsilon_1 > 0 \text{ and } \epsilon_2 > 0 \\
\hat{\Phi}_k \left( x \right) - \hat{\Phi}_k \left( T_{0}^k \right) & \text{if } \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) + \epsilon_1 \\
1 - \hat{\Phi}_k \left( T_{0}^k \right) & \text{if } \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right)
\end{cases}
\]

\[
\hat{\Phi}_k^* \left( x \right) = \begin{cases} 
0 & \text{if } \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) \\
\frac{2 \hat{\Phi}_k \left( x \right) - \hat{\Phi}_k \left( T_{0}^k \right)}{1 - 3 \hat{\Phi}_k \left( T_{0}^k \right)} \left( 0.5 - \epsilon_2 \right) & \text{if } \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \left( 1 - \hat{\Phi}_k \left( T_{0}^k \right) \right)/2 \\
0.5 - \epsilon_2 & \text{if } \left( 1 - \hat{\Phi}_k \left( T_{0}^k \right) \right)/2 < \hat{\Phi}_k \left( x \right) \leq \frac{1 + \hat{\Phi}_k \left( T_{0}^k \right)}{2}, \epsilon_2 > 0 \\
\hat{\Phi}_k \left( x \right) - \hat{\Phi}_k \left( T_{0}^k \right) & \text{if } \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) \\
\frac{\epsilon_2}{2} - \frac{1 - 2 \hat{\Phi}_k \left( x \right)}{2 \hat{\Phi}_k \left( T_{0}^k \right)} & \text{if } \hat{\Phi}_k \left( T_{0}^k \right)/2 < \hat{\Phi}_k \left( x \right) \leq \frac{1 + \hat{\Phi}_k \left( T_{0}^k \right)}{2} \\
\frac{1 + \hat{\Phi}_k \left( T_{0}^k \right)}{2} & \text{if } \hat{\Phi}_k \left( T_{0}^k \right) < \hat{\Phi}_k \left( x \right) \leq \hat{\Phi}_k \left( T_{0}^k \right) + \epsilon_1
\end{cases}
\]
Case 1: $\Phi(t) \geq 1/3$

Following Theorem 3:

$$\int_0^1 \Phi^{-1}(\alpha) \, d\alpha = f\left(\frac{\epsilon_1}{0.5 - \epsilon_2}, \Phi(t), 0, 0.5 - \epsilon_2\right) + f\left(\frac{1 - \Phi(t) - 2\epsilon_1}{2\epsilon_2}, \frac{1 + \Phi(t)}{2}\right) - \frac{(1 - \Phi(t) - 2\epsilon_1)}{4\epsilon_2}, \frac{0.5 - \epsilon_2}{0.5 - \epsilon_2, 0.5} + f(1 - \Phi(t), \Phi(t), 0.5, 1)$$

$$E(\xi|\xi > t) = \lim_{\epsilon_1 \rightarrow 0, \epsilon_2 \rightarrow 0} \int_0^1 \Phi^{-1}(\alpha) \, d\alpha = e - \frac{\alpha}{2\pi} \ln 1 - \Phi(t) - \frac{\alpha}{2\pi} \ln \frac{1 + \Phi(t)}{2} + \frac{\alpha}{2\pi} \ln \frac{1 + \Phi(t)}{2}$$

Case 2: $\Phi(t) < 1/3$

Following Theorem 3:

$$\int_0^1 \Phi^{-1}(\alpha) \, d\alpha = f\left(\frac{1 - 3\Phi(t)}{0.5 - \epsilon_2}, \Phi(t), 0, 0.5 - \epsilon_2\right) + f\left(\frac{\Phi(t)}{2\epsilon_2}, \frac{1 - (1 - \epsilon_2)\Phi(t)}{2\epsilon_2}, \frac{0.5}{0.5 - \epsilon_2, 0.5} + f(1 - \Phi(t), \Phi(t), 0.5, 1)$$

$$E(\xi|\xi > t) = \lim_{\epsilon_1 \rightarrow 0, \epsilon_2 \rightarrow 0} \int_0^1 \Phi^{-1}(\alpha) \, d\alpha = e + \frac{\alpha}{2\pi} \ln \frac{1 - \Phi(t)}{2} - \frac{\alpha}{2\pi} \ln \Phi(t) \ln \frac{1 + \Phi(t)}{2} + \frac{\alpha}{2\pi} \ln \frac{1 + \Phi(t)}{2} \ln \frac{1 - \Phi(t)}{2} - \frac{\alpha}{2\pi} \ln \frac{1 - \Phi(t)}{2}$$

$$\hat{\Phi}_k^{-1}(\alpha) = \begin{cases} \hat{\Phi}_k^{-1}\left(\frac{\epsilon_1}{0.5 - \epsilon_2} + \hat{\Phi}_k\left(T_{k0}\right)\right) & \text{if } \alpha \leq 0.5 - \epsilon_2 \\ \hat{\Phi}_k^{-1}\left(\frac{1 + \hat{\Phi}_k\left(T_{k0}\right)}{2}\right) - (1 - \hat{\Phi}_k\left(T_{k0}\right) - 2\epsilon_1)(0.5 - \alpha) & \text{if } 0.5 - \epsilon_2 < \alpha \leq 0.5 \\ \hat{\Phi}_k^{-1}(\alpha(1 - \hat{\Phi}_k\left(T_{k0}\right)) + \hat{\Phi}_k\left(T_{k0}\right)) & \text{if } \alpha > 0.5 \end{cases}$$

$$\hat{\Phi}_k^{-1}(\alpha) = \begin{cases} \frac{\alpha}{2\epsilon_2} + 1 - 3\hat{\Phi}_k\left(T_{k0}\right) & \text{if } \alpha \leq 0.5 - \epsilon_2 \\ \frac{1}{2} - (1 - 2\alpha - \epsilon_2)\hat{\Phi}_k\left(T_{k0}\right) & \text{if } 0.5 - \epsilon_2 < \alpha \leq 0.5 \\ \hat{\Phi}_k^{-1}(\alpha(1 - \hat{\Phi}_k\left(T_{k0}\right)) + \hat{\Phi}_k\left(T_{k0}\right)) & \text{if } \alpha > 0.5 \end{cases}$$

$$E(\xi|\xi > t) = \begin{cases} e - \frac{\alpha}{2\pi} \ln \frac{1 - \Phi(t)}{2} - \frac{\alpha}{2\pi} \ln \frac{1 + \Phi(t)}{2} & \text{if } \Phi(t) \geq 1/3 \\ e + \frac{\alpha}{2\pi} \ln \frac{1 - \Phi(t)}{2} & \text{if } \Phi(t) < 1/3. \end{cases}$$
If Φi (x) follows a linear uncertainty distribution \( L(a_i, b_i) \), \( i = 1, \ldots, k \), with Theorem 7, the predicted RUL \( \hat{y} \) is given in the following equation.

\[
E[\xi] = \sum_{i=1}^{k-1} E(\xi_i) - T_k = E(\xi) + \sum_{i=1}^{k-1} \frac{a_i + b_i}{2} - T_k
\]

\[
E(\xi^*_{k}) = \int_0^1 \Phi_{k-1}(\alpha) d\alpha
\]

\[
= \begin{cases} 
  a_k + (b_k-a_k) \left(\frac{5\Phi_k(T_k) + 3}{8}\right) & \text{if } \Phi_k(T_k) \geq 1/3 \\
  a_k + (b_k-a_k) \left(\frac{\Phi_k(T_k) + 2}{4}\right) & \text{if } \Phi_k(T_k) < 1/3 
\end{cases}
\]

(16)

For estimating the prediction interval, we try to find the minimal \( b \) that satisfies the following equation.

\[
\hat{y} - b \geq \hat{y} - b \geq a_0
\]

(17)

Since \( M \hat{y} - b < \hat{y} \leq \hat{y} + b \geq \hat{y} - b \), the prediction interval with a confidence level \( a_0 \) is suggested as \( [\hat{y} - b, \hat{y} + b] \) meaning that we have a chance of \( a_0 \) to cover \( y \) with our confidence interval. And, thus, the estimated value of \( b \) is expressed as

\[
\hat{b} = \min_{\max(0, \Psi(y)-a_0) \leq \alpha \leq \min(1-a_0, \Psi(y))} \max(\hat{y} - \hat{y}, \hat{y} - \hat{y}).
\]

(18)

With Equation (2), the inverse uncertainty distribution \( \hat{y}^{-1} \) of the RUL in Equation (18) with respect to a belief degree \( \alpha \) is expressed as the sum of the inverse uncertainty distributions of \( \hat{y}^{-1}_k \) and \( \hat{y}^{-1} \), as shown in the following equation.

\[
\hat{y}^{-1} = \hat{y}^{-1}(\alpha) = \max\{0, \Psi(y)-a_0\} \leq \alpha \leq \min\{1-a_0, \Psi(y)\} \max(\hat{y} - \hat{y}, \hat{y} - \hat{y}),
\]

III. CASE STUDY
A. DESCRIPTION OF THE PROBLEM
In nuclear power plants, increasing the inherent safety features is of critical importance and unexpected accidents are not acceptable. The passive residual heat removal system (PRHRS) removes the core decay heat by natural circulation in the case of emergency conditions and long-term cooling for repairing or refueling [29]. The layout of a passive residual heat removal system is shown in Figure 3. The circulation of the cooling liquid is controlled by the pneumatic valves driven by centrifugal pumps. In case of emergency, the normal function of the centrifugal pumps is highly important for the safety of a nuclear power plant. Effective failure prognostics may provide valuable information for proper maintenance actions. However, limited degradation data are collected concerning the degradation process of centrifugal pumps. The main reasons include: i) PRHRS as a standby system does not work for most of the time and limited degradation may occur; ii) the undertaken conservative maintenance strategy barely tolerates the failure of a centrifugal pump; iii) for the safety of personnel and the plant, the restricted access to PRHRS makes it hard to acquire the health condition data.

The degradation process of a centrifugal pump includes four different states, i.e. \( S_3, S_2, S_1, S_0 \), with \( S_3 \) being the healthy state and \( S_0 \) being the failure state, as shown in Figure 4 [30]. Uncertain variable \( \xi_i \) represent the holding time on state \( S_i \), and it follows a normal uncertainty distribution \( \Phi_i \) with unknown parameters \( \{e_i, \sigma_i\} \). The holding times on different states are independent. Assume that the maintenance is carried out with equal time interval. The observations \( z_t \) are collected during each scheduled maintenance and may indicate the current state of the pump and, thus, the cumulated

\[
E(\xi_k) = \begin{cases} 
  e_k + \frac{2 \pi \sqrt{3}}{\sigma_k \sqrt{3}} \left[ \frac{1 - \hat{\Phi}_k(T_{i0})}{2} \right] - \frac{2 \pi \sqrt{3}}{\sigma_k \sqrt{3}} \left[ \frac{1 + \hat{\Phi}_k(T_{i0})}{2} \right] & \text{if } \hat{\Phi}_k(T_{i0}) \geq \frac{1}{3} \\
  e_k + \frac{2 \pi \sqrt{3}}{\sigma_k \sqrt{3}} \left[ \frac{1 - \hat{\Phi}_k(T_{i0})}{2} \right] - \frac{2 \pi \sqrt{3}}{\sigma_k \sqrt{3}} \left[ \frac{1 + \hat{\Phi}_k(T_{i0})}{2} \right] & \text{if } \hat{\Phi}_k(T_{i0}) < \frac{1}{3}.
\end{cases}
\]
FIGURE 4. Degradation process of a centrifugal pump.

TABLE 1. Holding times (with arbitrary unit) of failed pumps on different states.

| PUMP NO. | $\xi_3$ | $\xi_2$ | $\xi_1$ |
|----------|---------|---------|---------|
| Training pumps | 1 | 18 | 13 | 8 |
| 2 | 22 | 15 | 5 |
| 3 | 24 | 11 | 7 |
| Test pumps | 4 | 21 | 14 | 6 |
| 5 | 17 | 16 | 9 |

TABLE 2. Estimated values of unknown parameters in $\varphi_i$.

| Degradation state | $\hat{\gamma}_i$ | $\hat{\sigma}_i$ | Error |
|-------------------|----------------|----------------|-------|
| $S_3$ | 21.5187 | 4.1346 | 0.0058 |
| $S_2$ | 13.0000 | 2.6875 | 1.40e-22 |
| $S_1$ | 6.7593 | 2.0673 | 0.0058 |

time $T_i^k$ that the pump has been on the current state $S_i$ can be derived. In the case study the observations $z_t$ and cumulated time $T_i^k$ are supposed to be precisely known.

The degradation data (holding times on different states) are only collected from five failed pumps, as shown in Table 1. Randomly, the degradation data on the first three pumps form the training data and the last two pumps are the test ones. The test pumps are supposed to be under operation, and their degradation states can only be observed during the scheduled maintenance. The objective is to adaptively predict the RUL of the test pumps with the observations $z_{1:t_0}$ until the current time $t_0$.

B. EXPERIMENTAL RESULTS

For failure prognostics, the first step is to model the degradation process with the holding times on different states of training pumps. Following Section 2.2.1, the degradation model can be obtained by estimating the unknown parameters in the corresponding uncertainty distributions. Given that the uncertain variable representing the holding time on a degradation state follows a normal uncertainty distribution, the unknown parameters are estimated with respect to least squares principles as in Equation (4). The estimated values are shown in Table 2.

The obtained degradation model gives a general trend of the degradation process. For failure prognostics of a specific pump, the uncertainty distributions for holding times should be updated with respect to the observations and, thus, the prediction results. The uncertainty distribution and the prediction results are, separately, updated with Equation (3) and Equations (9) and (15). The prediction results for pumps number 4 and 5 are shown, separately, in Figures 5 and 6.

In these two figures, for a specific time instance $t_0$ on the horizontal axis, the observations $z_{1:t_0}$ (dot in the figures) until $t_0$ are known, indicating the degradation states for the past time. The holding time on the current state until $t_0$, i.e. $T_i^k$ is obtained from the observations. For example, in Figure 5, at time instance $t_0 = 30$, the pump 4 is on the state $S_2$ and $T_4^k = 9$. Thus, the holding time of pump 4 on $S_2$ is at least $T_4^k = 9$. The uncertainty distribution of the holding time on $S_2$ is updated with the previous conditional uncertainty distribution $\Phi_2 (x | (9, +\infty))$. The predicted RUL (i.e. 11.02) and its associated prediction interval with 90% confidence level (i.e. [3.93 18.11]) for pump 4 at $t_0 = 30$ are derived and they are marked as five-pointed star and vertical segment.
line in the figures, respectively. The diamond shows the true RUL for this instance (i.e., 11).

These figures show that satisfactory results have been obtained for both test pumps. It is observed that relatively better results are achieved for pump 4, in comparison with pump 5. This is because the holding times of pump 4 on different degradation states are close to the expected value of the estimated uncertainty distribution given in Table 2.

Figures 6 shows clearly the effectiveness of the adaptive prediction update process with conditional uncertainty distribution in 1-UHSMP. Since the true holding times of pump 5 on $S_2$ and $S_1$ are all higher than those in the training pumps, the predicted RUL becomes lower than the true RUL from time 18. However, the proposed method seems realize this problem with the new observations and successfully approximates again the true RUL at the end.

C. COMPARISON WITH FIRST-ORDER HSMP

Traditional first-order HSMP with frequency-based probability estimation is considered as benchmark method. The holding time distribution is also updated with the observations with Bayes theorem. Except 1-UHSMP with normal uncertainty distribution, 1-UHSMP with linear uncertainty distribution is also considered in the comparison. They are noted as 1-UHSMP\textsubscript{normal} and 1-UHSMP\textsubscript{linear}, respectively. First-order UHSMP and HSMP are compared with respect to the mean relative squared error (MRSE) and the coverage percentage of the prediction interval on the true RUL. The prediction results of HSMP for test pumps are shown in Figures 7 and 8, respectively, and the results of these two methods are compared in Table 3.

| Pump | MRSE  | Coverage |
|------|-------|----------|
| Pump 4 | 0.0501 | 100.00%  |
| 1-UHSMP\textsubscript{normal} | 0.0621 | 100.00%  |
| Pump 5 | 0.0286 | 100.00%  |
| 1-UHSMP\textsubscript{linear} | 0.0775 | 75.61%   |
| HSMP | 0.0383 | 100.00%  |
|       | 0.0616 | 69.05%   |

From Table 3, one may observe that, with respect to MRSE, these three methods give all acceptable results on both pumps, with HSMP slightly better on pump 5 and 1-UHSMP\textsubscript{linear} better on pump 4. The proposed method gives higher coverage on the true RUL for pump 5. And with respect to coverage, UHSMP with normal uncertainty distribution achieves better results than that with linear uncertainty distribution, since linear uncertainty distribution is bounded with upper and lower bound.

In this case study, it can be concluded that for fixed training pumps, 1) first-order HSMP and UHSMP both perform well on the test pump (i.e., pump 4) with holding times within the range of those in training pumps; 2) UHSMP performs better than HSMP for test pump (pump 5) with holding times out of the range. The latter one is caused by the fact that HSMP based on probability theory may overfit the limited training data and be misled by the limited data in the distribution estimation. 3) For scarce data, 1-UHSMP with normal uncertainty distribution is preferred than that with linear uncertainty distribution.

Thus, 1-UHSMP is recommended for degradation process modelling when the training data is not enough to approximate the true holding time distributions with probability theory.

IV. CONCLUSIONS

For tackling failure prognostics with scarce data, the first-order uncertain hidden semi-Markov process denoted as 1-UHSMP is proposed for degradation processes with discrete states. The holding time distributions are estimated with uncertain statistics and the proposed model can adaptively update the prediction results with the conditional
uncertainty distribution considering observations related to the true degradation state of the object equipment. Analytical expressions are derived for the predicted RUL with two popular regular uncertainty distributions. A case study on the centrifugal pump in nuclear power plants is carried out. Experimental results show the effectiveness of the proposed method. Comparison with the conventional HSMP proves the benefits of 1-UHSMP in failure prognostics with limited degradation data.

The proposed model can be further improved in the future on the following to directions. 1) The imprecise observations can be integrated in the proposed model, since the precise observations assumption is not always guaranteed. 2) The proposed semi-Markov process can be extended to cases with high orders, since some degradation process exhibits more complex and diverse degradation process.

REFERENCES

[1] L. Liao and F. Kottig, “Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction,” IEEE Trans. Rel., vol. 63, no. 1, pp. 191–207, Mar. 2014.

[2] F. Camci, “System maintenance scheduling with prognostics information using genetic algorithm,” IEEE Trans. Rel., vol. 58, no. 3, pp. 539–552, Sep. 2009.

[3] K. Javed, R. Gouriveau, and N. Zerhouni, “State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels,” Mech. Syst. Signal Process., vol. 94, pp. 214–236, Sep. 2017.

[4] Y. Zhao, L. Ye, P. Pinson, Y. Tang, and P. Lu, “Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting,” IEEE Trans. Power Syst., vol. 33, no. 5, pp. 5029–5040, Sep. 2018.

[5] J. Guo, Z. Li, and M. Pecht, “A Bayesian approach for li-ion battery capacity fade modeling and cycles to failure prognostics,” J. Power Sources, vol. 281, pp. 173–184, May 2015.

[6] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, “A data-driven failure prognostics method based on mixture of gaussians hidden Markov models,” IEEE Trans. Rel., vol. 61, no. 2, pp. 491–503, Jun. 2012.

[7] D. Laredo, Z. Chen, O. Schütze, and J.-Q. Sun, “A neural network-evolutionary computational framework for remaining useful life estimation of mechanical systems,” Neural Netw., vol. 116, pp. 178–187, Aug. 2019.

[8] R. Khelif, B. Chebel-Morello, S. Malinowski, E. Laajili, F. Fnaiech, and N. Zerhouni, “Direct remaining useful life estimation based on support vector regression,” IEEE Trans. Ind. Electron., vol. 64, no. 3, pp. 2276–2285, Mar. 2017.

[9] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, “Remaining useful life estimation of engineered systems using vanilla LSTM neural networks,” Neurocomputing, vol. 275, pp. 167–179, Jan. 2018.

[10] X. Li, X. Yang, Y. Yang, I. Bennett, and D. Mba, “An intelligent diagnostic and prognostic framework for large-scale rotating machinery in the presence of scarce failure data,” Struct. Health Monitor., Oct. 2019, Art. no. 147592171988401.

[11] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, “Particle filter-based prognostics: Review, discussion and perspectives,” Mech. Syst. Signal Process., vol. 72–73, pp. 2–31, May 2016.

[12] Y. Peng and M. Dong, “A prognosis method using age-dependent hidden semi-Markov model for equipment health prediction,” Mech. Syst. Signal Process., vol. 25, no. 1, pp. 237–252, Jan. 2011.

[13] K. Zhu and T. Liu, “Online tool wear monitoring via hidden semi-Markov model with dependent durations,” IEEE Trans. Ind. Informat., vol. 14, no. 1, pp. 69–78, Jan. 2018.

[14] X. Wu and Q. Ye, “Fault diagnosis and prognostic of solid oxide fuel cells,” J. Power Sources, vol. 321, pp. 47–56, Jul. 2016.

[15] O. F. Eker, F. Camci, A. Guçu, H. Ylboga, M. Sevkli, and S. Baskan, “A simple state-based prognostic model for railway turnout systems,” IEEE Trans. Ind. Electron., vol. 58, no. 5, pp. 1718–1726, May 2011.

[16] T. Loutas, N. Eleftheroglou, G. Georgoulas, P. Loukopoulos, D. Mba, and I. Bennett, “Valve failure prognostics in reciprocating compressors utilizing temperature measurements, PCA-based data fusion, and probabilistic algorithms,” IEEE Trans. Ind. Electron., vol. 67, no. 6, pp. 5022–5029, Jun. 2020.

[17] R. Moghaddas and M. J. Zuo, “An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process,” Rel. Eng. Syst. Saf., vol. 124, pp. 92–104, Apr. 2014.

[18] D. B. Liu, “Uncertainty theory,” in Uncertainty Theory, 2007.

[19] Q. Zhang, R. Kang, and M. Wen, “Belief reliability for uncertain random systems,” IEEE Trans. Fuzzy Syst., vol. 26, no. 6, pp. 3605–3614, Dec. 2018.

[20] L. Hu, D. Yue, and G. Zhao, “Reliability assessment of random uncertain multi-state systems,” IEEE Access, vol. 7, pp. 22781–22789, 2019.

[21] B. Liu, “Uncertain risk analysis and uncertain reliability analysis,” J. Uncertain Syst., vol. 4, no. 3, pp. 163–170, Aug. 2010.

[22] M. Huang, L. Ren, L. H. Lee, X. Wang, H. Kuang, and H. Shi, “Model and algorithm for 4PLRP with uncertain delivery time,” Inf. Sci., vol. 330, pp. 211–225, Feb. 2016.

[23] X.-Y. Li, J.-P. Wu, L. Liu, M.-L. Wen, and R. Kang, “Modeling accelerated degradation data based on the uncertain process,” IEEE Trans. Fuzzy Syst., vol. 27, no. 8, pp. 1532–1542, Aug. 2019.

[24] M. Wen, Q. Zhang, R. Kang, and Y. Yang, “Some new ranking criteria in data envelopment analysis under uncertain environment,” Comput. Ind. Eng., vol. 110, pp. 498–504, Aug. 2017.

[25] X. Guo and Y. Gao, “Connectedness index of uncertain graph,” Int. J. Uncertainty, Fuzziness Knowl.-Based Syst., vol. 21, no. 1, pp. 127–137, Feb. 2013.

[26] Y. Gao and Z. Qin, “On computing the edge-connectivity of an uncertain graph,” IEEE Trans. Fuzzy Syst., vol. 24, no. 4, pp. 981–991, Aug. 2016.

[27] B. Zhang and J. Peng, “Euler index in uncertain graph,” Appl. Math. Comput., vol. 218, no. 20, pp. 10279–10288, Jun. 2012.

[28] H. Li and H. Zhang, “On the edge-connectivity of an uncertain random graph,” IEEE Access, vol. 8, pp. 59126–59134, 2020.

[29] Y.-J. Chung, S.-H. Yang, H.-C. Kim, and S.-Q. Zee, “Thermal hydraulic calculation in a passive residual heat removal system of the SMART-P plant for forced and natural convection conditions,” Nucl. Eng. Design, vol. 232, no. 3, pp. 277–288, Aug. 2004.

[30] M. J. Daigle and K. Goebel, “A model-based prognostics approach applied to pneumatic valves,” Int. J. Progn. Heal. Manag., vol. 2, no. 2, pp. 84–99, Oct. 2011.

JIE LIU (Member, IEEE) received the B.Sc. degree in mechanical engineering and the M.Sc. degree in physics from Beihang University, in 2009 and 2012, respectively, and the Ph.D. degree in industrial engineering from CentraleSupelec, in 2015. He is currently an Associate Professor at Beihang University. His current research interests include diagnostics, prognostics, and health management with data-driven methods.