Uncertainty Quantification in Remaining Useful Life Prediction using First-Order Reliability Methods

Shankar Sankararaman*, *Member, IEEE*, Matthew J. Daigle**, *Member, IEEE*, Kai Goebel**, *Member, IEEE*.

***Abstract*—In this paper, we investigate the use of first-order reliability methods to quantify the uncertainty in the remaining useful life (RUL) estimate of components used in engineering applications. The prediction of RUL is affected by several sources of uncertainty, and it is important to systematically quantify their combined effect on the RUL prediction in order to aid risk assessment, risk mitigation, and decision-making. While sampling-based algorithms have been conventionally used for quantifying the uncertainty in RUL, analytical approaches are computationally cheaper, and sometimes they are better suited for online decision-making. Exact analytical algorithms may not be available for practical engineering applications, but effective approximations can be made using first-order reliability methods. This paper describes three first-order reliability-based methods for RUL uncertainty quantification: first-order second moment method (FOSM), the first-order reliability method (FORM), and the inverse first-order reliability method (inverse-FORM). The inverse-FORM methodology is particularly useful in the context of online health monitoring, and this method is illustrated using the power system of an unmanned aerial vehicle, where the goal is to predict the end of discharge of a lithium-ion battery.**

NOMENCLATURE

t	Continuous time index
k	Discrete time index
$\mathbf{x}(t)$	State vector
$\boldsymbol{\theta}(t)$	Parameter vector
$\mathbf{u}(t)$	Input vector
$\mathbf{v}(t)$	Process Noise vector
$\mathbf{y}(t)$	Output vector
$\mathbf{n}(t)$	Measurement noise vector
$\Phi(.)$	Standard normal distribution

ABBREVIATIONS

RUL	Remaining Useful Life
FOSM	First-Order Second moment Method
FORM	First-Order Reliability Method
SOC	State Of Charge

I. INTRODUCTION

The need for an accurate, efficient health management system has become exceedingly important in safety-critical and mission-critical engineering systems. The goal of health management is to constantly monitor the performance of these systems, perform diagnosis (fault detection, isolation, and estimation), perform prognosis (predict possible failures in the future and estimate remaining useful life), and aid online decision-making (fault mitigation, fault recovery, mission replanning, etc.). Mathematical models are developed for individual components of the system, and these component models can be integrated to form the overall system. The component-level mathematical models can be constructed either using laws of physics (physics-based models [1]) or using data collected through component-level testing (data driven models [2]), and are used for both system-level diagnostics [3] and prognostics [4]. Conventionally, diagnosis is performed at system or subsystem levels, and prognosis is performed at component levels. However, system-level approaches to prognostics are also available in the literature [5].

Uncertainty assessment and management are important aspects of health management, due to the presence of several unknown factors that affect the operations of the system of interest. Therefore, it is not only important to develop robust algorithms for diagnosis and prognosis, i.e., accurately perform diagnosis and prognosis in the presence of uncertainty, but also important to quantify the amount of confidence in the results of diagnosis and prognosis. This important task can be accomplished by quantifying the uncertainty in fault diagnosis and prognosis. It is also necessary to perform such uncertainty quantification (UQ) online to enable in-flight decision-making capabilities. Sankararaman and Mahadevan [6], [7] developed statistical (both frequentist and Bayesian) approaches to quantify the uncertainty in the three steps of diagnosis (detection, isolation, estimation) in an online health monitoring framework.

Recent research efforts in the domain of health monitoring have focused on prognostics and condition-based maintenance. An important aspect of prognostics is the accurate estimation of remaining useful life (RUL). Bo Sun et al. [8] discuss the benefits of prognostics, and explain how the calculation of RUL is important for technical health determination and life extension [9] in the context of condition-based monitoring [10]. Degradation signals [11], [12] and deterioration models [13] have been used in combination with statistical methods for estimating the remaining useful life in prognostics. Researchers have investigated both model-based approaches [14] and data-driven approaches [15] for prognostics [16] and RUL prediction [17]. These methods have been applied to a variety of applications including mechanical bearings [18], gears [19], lithium-ion batteries [20], etc.

The importance of uncertainty significantly increases in the context of prognosis because the focus is on predicting future behavior, which is far more challenging and uncertain than diagnosis. There have been several attempts in the past to quantify uncertainty in prognosis and RUL estimation. Tang

*SGT Inc., NASA Ames Research Center, Moffett Field, CA 94035, USA
**NASA Ames Research Center, Moffett Field, CA 94035, USA
Manuscript August 30, 2013

et al. [21] discuss the use of Bayesian tracking algorithms for uncertainty quantification and management in prognostics for Integrated Vehicle Health Management (IVHM) systems. The Damage Prognosis Project at Los Alamos National Laboratory [22], [23] exclusively dealt with prognosis and uncertainty quantification applied to structural composites. Several researchers worked on this project and published articles that deal with model development, verification, validation, prediction, and uncertainty quantification; the conclusions of this project have been documented by Farrar et al [24]. Sankararaman et al. [25], [26] quantified the uncertainty in fatigue crack growth prognosis in metals by using finite element models (for structural analysis), crack growth models (to predict future crack growth), and Monte Carlo simulation (for uncertainty quantification). Gu et al. [27] also used Monte Carlo simulation to compute the uncertainty in damage in electronics subjected to random vibration. In some practical applications, Monte Carlo simulation using exhaustive sampling may be computationally expensive, and this challenge has inspired the development of intelligent sampling-based algorithms [14], [28]–[30], and mathematical techniques such as relevance vector machines [31] and principle component analysis [32], that can reasonably approximate the uncertainty in the prognostics. Further, Bayesian [33] and maximum relative entropy methods [34] have also been used for estimating uncertainty in prognostics.

While existing methods for prognostics uncertainty quantification are mostly based on sampling, this paper investigates the use of analytical algorithms for calculating the uncertainty in RUL prediction. Several analytical algorithms such as the first-order second moment method (FOSM), first-order reliability method (FORM), second-order reliability method (SORM), etc. have been used to calculate the reliability of structural systems in the past [35]–[39]. These methods are collectively known as first-order reliability methods [39]; they are based on the explicit definition of performance functions or limit state functions, and are commonly used to calculate the probability of failure of structural systems.

The contribution of this paper is to extend the use of the FOSM [36], [40] and FORM [41], [42] methods for use with state space models, and develop a computational framework to calculate the entire probability distribution of the remaining useful life prediction, instead of simply calculating the probability of failure. A few practical challenges are encountered in this regard, and new statistical approaches are presented in this paper to overcome these challenges. Further, a generic computational framework for uncertainty quantification in online health monitoring and prognostics is developed, and first-order reliability methods are integrated into this framework to aid the prediction of RUL.

There are two major advantages from using first order reliability methods for estimating the uncertainty in RUL prediction. First, these methods require very few prognostic evaluations in comparison with several sampling-based approaches, and therefore are suitable for online prognostic calculations. Second, these methods can produce repeatable (deterministic) calculations, i.e., the exact same PDF for RUL is obtained on every repetition of the algorithm. It is worth noting that the latter feature is an important criterion for existing verification, validation, and certification protocols in the aerospace domain. Therefore, investigating such analytical algorithms allows us to move a step closer towards adopting prediction algorithms (which are inherently stochastic), by meeting the needs of the current certification process.

II. Prognosis, and RUL Calculation

This section formulates the prognosis problem, and explains the estimation of RUL, without considering the effects of uncertainty. Consider an arbitrary time instant t_P at which the remaining useful life needs to be predicted, i.e., t_P denotes the time of the prediction. The architecture for model-based prognosis is first explained, and then it is demonstrated that the remaining useful life can be expressed as a function of other quantities; this functional formulation will aid in uncertainty quantification in the following sections.

A. Architecture for Model-Based Prognostics

As explained by Daigle and Goebel [4], it is important to develop an architecture for model-based prognostics for practical engineering purposes. As seen from Fig. (1), there are three steps that need to be performed: state estimation, state prediction, and remaining useful life computation.

1) State Estimation: The first step of estimating the state at t_P serves as the precursor to prognosis and RUL computation. Consider the state space model, which is used to continuously predict the state of the system, as

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t), \mathbf{v}(t)). \tag{1}$$

The state vector at time t_P, i.e., $\mathbf{x}(t)$ (and the parameters $\boldsymbol{\theta}(t)$, if they are unknown), is (are) estimated using output data collected until t_P. Let $\mathbf{h}$ denote the output equation. Then,

$$\mathbf{y}(t) = \mathbf{h}(t, \mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t), \mathbf{n}(t)). \tag{2}$$

Typically, filtering approaches such as Kalman filtering, particle filtering, etc. may be used for such state estimation [43].

2) State Prediction: Having estimated the state at time t_P, (1) is used to predict the future states of the component or system. This differential equation can be discretized and used to continuously predict $\mathbf{x}(t)$ for all $t > t_P$. Because time-discretization is employed during implementation, time is indicated through the discrete time-index k. There exists a one-to-one mapping between continuous time t and discrete time-index k; for example, if time is discretized every $0.1\ s$, and if continuous time $t = 0$ corresponds to the time-index $k = 1$, then $t = 0.1$ corresponds $k = 2$, $t = 0.2$ corresponds to $k = 3$, and so on. For the purpose of illustration, let t^k denote the time instant corresponding to a generic time-index k, and k^t denote the time-index corresponding to a generic time instant t. Further, let k_P denote the time-index which corresponds to the time instant t_P; therefore, $t_P^k = t_P$, and $k_P^t = k_P$. Using the recursive discretized state space equation, the state at any future time instant can be calculated as $\mathbf{x}(k)$ for all $k > k_P$.

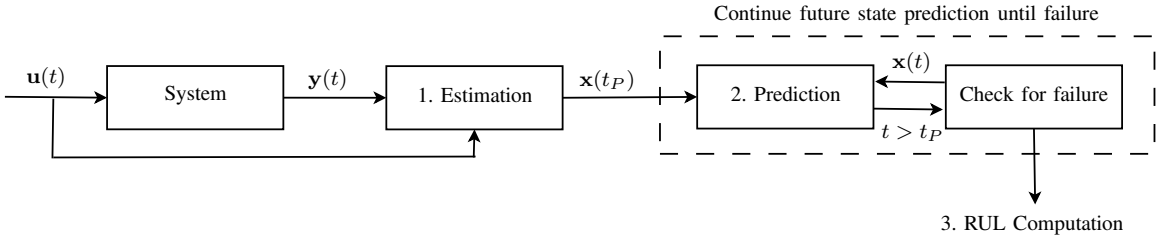

Fig. 1: Model-Based Prognostics Architecture.

3) RUL Computation: RUL computation is concerned with the performance of the component that lies outside a given region of acceptable behavior. The desired performance is expressed through a set of n_c constraints, $C_{EOL} = \{c_i\}_{i=1}^{n_c}$, where $c_i : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_u} \rightarrow \mathbb{B}$ maps a given point in the joint state-parameter space given the current inputs, $(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t))$, to the Boolean domain $\mathbb{B} \triangleq [0, 1]$, where $c_i(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t)) = 1$ if the constraint is satisfied, and 0 otherwise [44].

These individual constraints may be combined into a single *threshold function* $T_{EOL} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_u} \rightarrow \mathbb{B}$, defined as

$$T_{EOL}(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t)) = \begin{cases} 1, & 0 \in \{c_i(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t))\}_{i=1}^{n_c} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

T_{EOL} is equal to 1 when any of the constraints are violated. Then, the End of Life (EOL), denoted by E, at any time instant t_P, is then defined as the earliest time point at which the value of T_{EOL} becomes equal to one. Therefore,

$$E(t_P) \triangleq \inf\{t \in \mathbb{R} : t \geq t_P \wedge T_{EOL}(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t)) = 1\}. \tag{4}$$

The Remaining Useful Life (RUL), denoted by R, at time instant t_P, is expressed as

$$R(t_P) \triangleq E(t_P) - t_P. \tag{5}$$

Note that the output (2) is not used in the prediction stage because EOL and RUL are not defined as functions of the output data ($(\mathbf{y}(t))$).

As stated earlier, time is discretized for the sake of numerical implementation; therefore, $k^{E(t_P)}$ denotes the time-index that corresponds to the end of life. For the sake of simplicity, the RUL can also be defined as a function of the prediction time-index as $R(k_P)$.

B. RUL: A Dependent Quantity

Thus, it is clear that RUL predicted at time t_P (equivalently, time-index k_P), i.e., $R(t_P)$ (or $R(k_P)$) depends on several factors.

1) Present state estimate ($\mathbf{x}(k_P)$) - using the present state estimate, and the state space (1), the future states ($\mathbf{x}(k_P)$, $\mathbf{x}(k_P + 1)$, $\mathbf{x}(k_P + 2)$, ..., $\mathbf{x}(k^{E(t_P)})$) can be calculated.
2) Future loading ($\mathbf{u}(k_P)$, $\mathbf{u}(k_P + 1)$, $\mathbf{u}(k_P + 2)$, ..., $\mathbf{u}(k^{E(t_P)})$) - these values are needed to calculate the future state values using the state space equations.
3) Parameter values from time-index k_P until time-index $k^{E(k_P)}$ - these values are denoted by $\boldsymbol{\theta}(k_P)$, $\boldsymbol{\theta}(k_P+1)$, ..., $\boldsymbol{\theta}(k^{E(t_P)})$.
4) Process noise - these values are $\mathbf{v}(k_P)$, $\mathbf{v}(k_P + 1)$, $\mathbf{v}(k_P + 2)$, ..., $\mathbf{v}(k^{E(t_P)})$.

For the purpose of RUL prediction, all of the above quantities are functionally independent of each other; hence, RUL becomes a functionally dependent quantity. Let $\boldsymbol{X} = \{X_1, X_2, ...X_i, ...X_n\}$ denote the vector of all the above dependent quantities, where n is the length of the vector $\boldsymbol{X}$, and therefore the number of uncertain quantities that influence the RUL prediction. Then the calculation of RUL (denoted by R) can be expressed in terms of a function, as

$$R = G(\boldsymbol{X}) \tag{6}$$

The above functional relation (6) can be graphically explained, as shown in Fig. (2).

For example, consider the case where the component or system is subjected to uniform loading (characterized by one variable, the amplitude, which remains constant with time), modeled using one parameter (which is invariant with time), and characterized using two states (the state estimates at time t_P, and (1) can be used to predict the state values at any future time instant). Then, excluding the effect of process noise, there are $n = 4$ quantities that affect the RUL prediction. Note that there are $k^{E(t_P)} - k_P + 1$ process noise terms for each state; therefore, the inclusion of process noise increases the value of n, and therefore the dimensionality of the problem. This raises a practical concern that will be addressed later in Section VII. Knowing the values of $\boldsymbol{X}$, it is possible to compute the value of R, using Fig. (2) equivalently represented by (6). In the present paper, the quantities contained in $\boldsymbol{X}$ are considered to be uncertain, and the focus is to compute their combined effect on the RUL prediction. This task can be accomplished by computing the probability distribution of the RUL R. The following section discusses these uncertainties in detail, and illustrates how the RUL computation can be viewed as an uncertainty propagation problem, that can be solved using statistical approaches.

III. Uncertainty in RUL

This section discusses the different sources of uncertainty that affect the RUL prediction, and analyzes the various issues concerned with RUL prediction. In a practical engineering problem, all the quantities contained in $\boldsymbol{X}$ in (6) are uncertain, and therefore, quantifying the uncertainty in RUL is equivalent to propagating the uncertainty in $\boldsymbol{X}$ through G. For the sake of clarity, the various sources of uncertainty in prognostics can be classified into the following categories. While the first two categories of uncertainty (present uncertainty and future

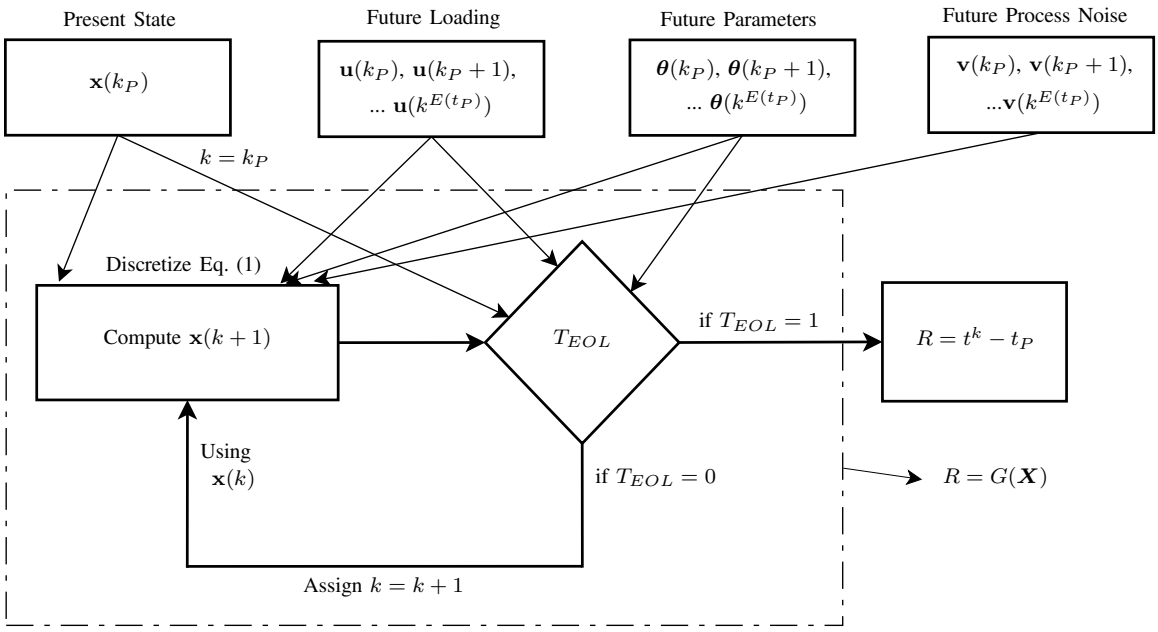

Fig. 2: Definition of G.

uncertainty) are inherent to the prognostics problem, the third category (modeling uncertainty) is inherent to the model-based prognostics architecture, and the fourth category (UQ method uncertainty) is inherent to the choice of uncertainty quantification method.

1) **Present uncertainty**: Prior to prognosis, it is important to be able to precisely estimate the condition-state of the component or system at the time t_P at which RUL needs to be predicted. As explained earlier, this problem is related to state estimation, and commonly addressed using filtering. Output data (usually collected through sensors) is used to estimate the state, and many filtering approaches are able to provide an estimate of the uncertainty in the present state, i.e., $\mathbf{x}(t_P)$. Practically, it is possible to improve the estimate of the states, and thereby reduce the uncertainty, by using better sensors and advanced filtering approaches.
2) **Future uncertainty**: The most important source of uncertainty in the context of prognostics is due to the fact that the future is unknown, i.e. both the loading and operating conditions are not known precisely, and it is important to assess the uncertainty in loading and environmental conditions before performing prognostics. If these quantities were known precisely (without any uncertainty), then there would be no uncertainty regarding the *true* remaining useful life of the component or system. However, this true RUL needs to be estimated using a model that predicts the future evolution of the engineering component or system; the usage of a model imparts additional uncertainty as explained below.
3) **Modeling uncertainty**: This paper uses the model-based prognostics architecture (Fig. (1)) for RUL prediction. Therefore, the response of the component or system to the loading and operating conditions is computed through a physics-based state-space model. Modeling uncertainty represents the difference between the predicted response and the true response (which can neither be known nor measured accurately), and comprises of several parts: model parameters, model form, and process noise. While it may be possible to quantify these terms until $t \leq t_P$, it is practically challenging to know their values at future time instants, i.e, $t > t_P$.
4) **UQ-Method uncertainty**: Even if all the above sources of uncertainty can be quantified accurately, it is necessary to quantify their combined effect on the RUL prediction, and thereby quantify the overall uncertainty in the RUL prediction. This process is equivalent to propagating the uncertainty in $\boldsymbol{X}$ through G, and the resultant uncertainty in RUL is expressed using the probability density function (PDF) or the cumulative distribution function (CDF). To *precisely* estimate the actual PDF (or CDF) of RUL, it is necessary to use an infinite (theoretically) number of samples of $\boldsymbol{X}$; each of these samples is used to calculate R using (6), and the resultant samples of R can be used to estimate both the PDF and CDF. Because it is not practically possible to use such an approach, researchers have developed alternate methods (both sampling-based and analytical). None of the sampling-based approaches can calculate the entire PDF or CDF accurately, because they would depend on the choice and location of samples. Exact analytical approaches are also unavailable in the context of prognostics; even if the state-space models and the EOL-threshold functions are linear, their combination renders G non-linear, and closed-form expressions for uncertainty propagation do not exist. Hence, it is necessary to resort to alternate, approximate UQ-methods. Therefore, it is important to understand that additional uncertainty is imparted by the uncertainty propagation method, and if possible, quantify this uncertainty.

The quantification of UQ-method uncertainty depends on the choice of the method used for uncertainty propagation. Further, each of the existing methods for uncertainty propagation may have a different motive. Monte Carlo sampling (MCS) attempts to approximate the entire probability distribution (PDF or CDF); it can be theoretically proved that calculating all the moments (mean, variance, skewness, kurtsosis, and so on) is equivalent to calculating the entire probability distribution. If a larger number of samples is used in MCS, higher order moments can be calculated, and therefore the entire probability distribution can be approximated with increased accuracy. Adaptive sampling focuses on estimating the cumulative distribution function (CDF) value, i.e., calculating the probability that the RUL is less than a given value. These two sampling methods use random samples for uncertainty quantification, and hence produce different results (albeit similar) on repeating the algorithm, because each repetition draws a different set of random samples. On the other hand, unscented transform sampling deterministically selects samples (i.e., without random sampling) to compute the first two moments (equivalently, mean and variance) of RUL [14], [45]. While it is important to calculate the first two moments of RUL, it is equally important to estimate the tail probabilities of the distribution of RUL, because failure is often caused due to events related to the tails of probability distributions. In other words, in a well-designed system, the probability of failure is very small (usually of the order of 10^{-3} or less), and to compute the value of RUL that corresponds to such low probabilities, it is important to accurately estimate the tail of the probability distribution of RUL. First-order reliability methods can not only quickly calculate tail probabilities but also approximate the entire CDF of RUL. However, these methods are based on repeated, piecewise linear approximation of G, and this approximation may lead to a deviation from the actual RUL distribution. While this paper does not attempt to quantify this deviation, the results from FORM will be verified by comparison against exhaustive sampling (Monte Carlo sampling), thereby demonstrating that the method uncertainty is negligible.

Note that the goal of this paper is simply to investigate the use of first-order reliability methods (FORM) for uncertainty quantification in RUL prediction, and a few simplifying assumptions are made regarding some of the above types of uncertainty. For example, the future loading is assumed to be constant, and the amplitude is chosen to be random. In the context of model uncertainty, process noise is included, but the model parameters are assumed to be known (without uncertainty), and model form uncertainty is not considered. In fact, rigorous methods to quantify the model form uncertainty in prognostics have not been developed. Future work will deal with rigorous uncertainty quantification by including variable amplitude loading, operational, and environmental conditions, model form uncertainty, etc. In this paper, the applicability of FORM-based methods to prognostics alone is discussed, and these methods are presented in the following sections. These methods include the first-order second moment method, the first-order reliability method, and the inverse first-order reliability method. All three methods are computationally cheap (because they require very few prognostic evaluations), and are therefore suitable for online health monitoring. Further, the methods also result in repeatable (deterministic) calculations, i.e., their results do not change on repetition, unlike sampling-based methods. The difference between the three methods is that they compute different quantities; while FOSM estimates the first-order mean and variance of RUL, FORM calculates the cumulative distribution function of RUL, and inverse-FORM calculates the inverse cumulative distribution function of RUL.

IV. First-Order Second Moment Method

The FOSM approach [36], [40], [46], as the name suggests, is a simple approximation of R using first-order Taylor series expansion. The first two moments, i.e. mean ($\mu_{\boldsymbol{X}}$) and variance ($\sigma_{\boldsymbol{X}}$) of $\boldsymbol{X}$, are used to approximate the first two moments of R. Consider the first-order Taylor series expansion of $R = G(\boldsymbol{X})$ around $\mu_{\boldsymbol{X}}$, as

$$R = G(\mu_{\boldsymbol{X}}) + \sum_{i=1}^{i=n} (X_i - \mu_{X_i}) \left(\frac{\partial G}{\partial X_i} \right)_{\mu_{\boldsymbol{X}}}. \tag{7}$$

Note that, now, R is a linear function of $\boldsymbol{X}$ with the partial derivatives as coefficients, and therefore it is straightforward to approximate the mean and variance of R, as

$$\mu_R = G(\mu_{\boldsymbol{X}}) \tag{8}$$

$$\sigma_R^2 = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \left(\frac{\partial G}{\partial X_i} \right)_{\mu_{\boldsymbol{X}}} \left(\frac{\partial G}{\partial X_j} \right)_{\mu_{\boldsymbol{X}}} Cov(X_i, X_j). \tag{9}$$

When the inputs to G are uncorrelated, the expression for variance in (9) simplifies to

$$\sigma_R^2 = \sum_{i=1}^{i=n} \left(\frac{\partial G}{\partial X_i} \right)_{\mu_{\boldsymbol{X}}}^2 \sigma_{X_i}^2. \tag{10}$$

Note that the first-order second moment approach can only compute the mean and variance of the remaining useful life (R) prediction. It is not directly useful for estimating the type of probability distribution of R, and cannot be used to calculate tail probabilities. However, if each of the inputs X_i were to be normally distributed (and s-independent of each other), then it can be easily proved that their linear combination is also a normally distributed variable [39]. This observation is used to develop a general methodology to calculate the entire CDF of the Remaining Useful Life (R) prediction in the following stages.

1) **FORM** - The FOSM approach linearized G at $\mu_{\boldsymbol{X}}$. By altering the location (point) of linearization, the FOSM approach is extended to the FORM [41], [47] approach.
   a) Normally Distributed Variables - First, consider the case when the variables ($\boldsymbol{X}$) are normally distributed. Even in this case, the FOSM method yields accurate results only when G is linear. When G is non-linear, it is not appropriate to always calculate the gradient of G at the mean $\mu_{\boldsymbol{X}}$. This leads to the obvious question: why linearize at the mean? The location of linearization is chosen analytically,

and then used to calculate the CDF of RUL at a particular $R = r$, i.e., $F_R(R = r) = P(R \leq r)$. (Note that an upper case letter refers to the name of a random variable, whereas the corresponding lower case letter refers to its realization.) It will be illustrated that the point of linearization varies with the choice of r. By repeating the entire process for different choices of r, the entire CDF can be calculated. This method, known as the first-order reliability method, is discussed in Section V-A.

b) Non-normally Distributed Variables - Then, this FORM approach is extended to non-normally distributed variables in Section V-B.

2) **Inverse FORM** - While Section V deals with calculating the CDF value for a given realization of R, Section VI considers the inverse problem, i.e. calculating the realization of R that corresponds to a given CDF value. This method, popularly known as the inverse-FORM approach [42], can be used to calculate probability bounds on the remaining useful life prediction, and hence is useful for online decision-making.

V. First-Order Reliability Method

The first-order reliability method (FORM) was originally developed by structural engineers to estimate the reliability of structural systems [35]. The main contribution of this paper is to extend this method to state space models and health monitoring with the goal of computing the probability distribution of RUL. Though the estimation of reliability and the calculation of CDF are statistically equivalent, the FORM method is described purely from an uncertainty propagation perspective, without reference to reliability calculation.

A. Normal Variables

In this subsection, assume that all the uncertain variables follow Gaussian distributions, i.e., $X_i \sim \text{N}(\mu_{X_i}, \sigma_{X_i})$ for all i. When Kalman filtering-based approaches (such as extended Kalman filtering, unscented Kalman filtering, etc.) are used for state estimation, $\mathbf{x}(t_P)$ is normally distributed. It is also common to choose normal distributions for quantities that represent process noise. In addition, this paper considers constant amplitude loading, and the amplitude may also be chosen to be normally distributed. Therefore, the FORM method is first explained for normal variables in this subsection.

Consider (6), which expresses the RUL as a function of the various sources of uncertainty. The goal of the FORM approach is to calculate the CDF value, i.e. $F_R(r) = P(R \leq r)$, given r, a realization of the random variable R. FORM achieves this goal by approximating the non-linear equation $R = G(\boldsymbol{X})$ using a linear equation to easily compute the CDF of R. The linear equation is constructed using a Taylor series approximation around the so-called point of linearization. The difference in FORM (with respect to FOSM) is that the point of linearization varies from one choice of r to another. The identification of the point of linearization is the most important component of FORM.

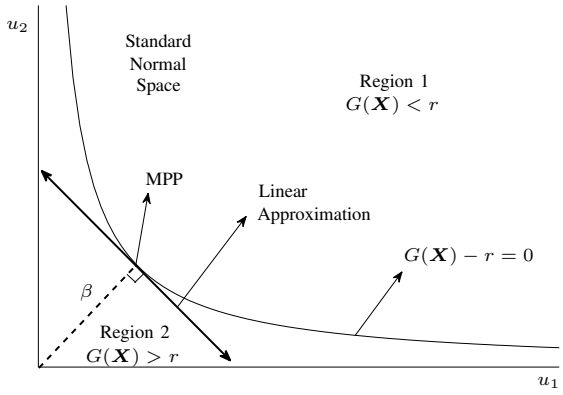

Fig. 3: Most Probable Point Estimation.

To calculate $F_R(r) = P(R \leq r)$, consider the contours of the function $R - r = G(\boldsymbol{X}) - r$; in particular, consider the curve described by $G(\boldsymbol{x}) - r = 0$, where $\boldsymbol{x}$ is a realization of $\boldsymbol{X}$. This curve differentiates the multidimensional space into two regions: one region where $R < r$, and another region where $R > r$, as shown in Fig. (3). (This curve of demarcation is popularly called the limit state in reliability analysis [39]. The equation corresponding to the limit state divides the space into two zones: a zone of failure, and a zone of safety. This terminology is not used in this paper because FORM is used here for uncertainty propagation rather than reliability calculation. For details of the original method, refer to Rackwitz and Fiessler [35], Sitar et al. [48], Haldar and Mahadevan [39], etc.)

Any point lying on the curve of demarcation would satisfy the equation $R = G(\boldsymbol{X})$. Because this curve serves as the demarcation between the two zones given by $R > r$ and $R < r$, and it is of interest to calculate the probability $P(R \leq r)$, it is intuitive that it is important to identify a linear function which closely resembles the contour $G(\boldsymbol{x}) - r = 0$. Hence, the point of linearization must lie on this curve of demarcation; in other words, the point of linearization must satisfy $G(\boldsymbol{x}) - r = 0$. This is clearly different from the FOSM approach, where the mean $\mu_{\boldsymbol{X}}$ was chosen as the point of linearization; for an arbitrary choice of r, it is obvious that the mean $\mu_{\boldsymbol{X}}$ will not satisfy this equation.

Therefore, the point of linearization should be located on the curve of demarcation. However, there are infinite points that satisfy this criterion, and it is important to select the appropriate one. Each of these infinite points has a likelihood of occurrence, and intuitively the *point of maximum likelihood* is chosen as the point of linearization. This likelihood can be calculated using the probability density function of the underlying random variables. For a single normal random variable X with mean μ and standard deviation σ, the PDF is given by

$$f_X(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right]. \tag{11}$$

For example, when $\mu = 10$, and $\sigma = 1$, $x = 10$ is 1.65 times more likely to occur than $x = 9$. The maximum value of the

likelihood function occurs at $x = \mu$; therefore, the farther x is away from the mean μ, the lower the likelihood of occurrence of x. As explained earlier, the mean cannot be chosen as the point of linearization because $G(\mu) - r \neq 0$. Therefore, if there is a single input variable, the point of linearization is chosen in such a way that it satisfies the equation $G(x) - r = 0$, and the value of $||x - \mu||$ is minimum.

However, in a general uncertainty propagation problem, the input to G is a vector, i.e., $\boldsymbol{X} = \{X_1, X_2, ... X_i, ...X_n\}$, and each X_i has its own mean μ_{X_i} and standard deviation σ_{X_i}. The objective is to identify the point of maximum likelihood, which can be calculated by maximizing the joint probability density function of all the input random variables. If the variables are *s*-independent, then the joint density function of $\boldsymbol{X}$ is expressed as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{i=n} \frac{1}{\sigma_i \sqrt{2\pi}} exp\Big[- \frac{(x_i - \mu_i)^2}{2\sigma_i^2} \Big]. \tag{12}$$

It can be verified (by taking the logarithm) that the maximizer of the above function simultaneously minimizes

$$\beta = \sqrt{\sum_{i=1}^{i=n} \Big(\frac{x_i - \mu_i}{\sigma_i} \Big)^2}. \tag{13}$$

Equation (13) can be rewritten as

$$\beta = \sqrt{\sum_{i=1}^{i=n} u_i^2}, \tag{14}$$

where

$$u_i = \frac{x_i - \mu_i}{\sigma_i}. \tag{15}$$

If the above computation were performed for every realization x_i of the random variable X_i, then the corresponding u_i would be realizations of the standard normal variable U_i, i.e., $U_i \sim N(0, 1)$. Therefore, (15) is referred to as the standard normal transformation. In the space of standard normal variables, maximizing the likelihood of occurrence is equivalent to minimizing (14), which implies that the point of linearization is that point on the curve of demarcation whose distance (measured in the standard normal space) from the origin is minimum. Because the point of linearization has the maximum likelihood of occurrence, it is popularly known as the Most Probable Point (MPP), as indicated in Fig. (3). The Most Probable Point is closest to the origin (in the standard normal space), and from (14) and (15), it can be seen that the distance of MPP from the origin is exactly equal to β. In fact, it can be proved that

$$P(R \leq r) \approx \Phi(-\beta) \tag{16}$$

where $\Phi(.)$ represents the standard normal cumulative distribution function. To prove (16), consider the first order Taylor's series expansion of $Z = G(\boldsymbol{X}) - r$ by linearizing around the MPP. Then, Z is approximated to be a normal random variable (denoted by $\tilde{Z}$) with mean $\mu_{\tilde{Z}}$, and standard deviation $\sigma_{\tilde{Z}}$. The CDF value of Z measured at $Z = 0$ is equal to $P(R \leq r)$, and is calculated in terms of $\mu_{\tilde{Z}}$ and $\sigma_{\tilde{Z}}$ as

$$P(Z \leq 0) = P(R \leq r) \approx \Phi(-\frac{\mu_{\tilde{Z}}}{\sigma_{\tilde{Z}}}). \tag{17}$$

Using this Taylor series expansion, it can be proved that [39]

$$\beta = \sqrt{\sum_{i=1}^{i=n} u_i^2} = \frac{\mu_{\tilde{Z}}}{\sigma_{\tilde{Z}}}. \tag{18}$$

Note that an approximation symbol is used in (16) and (17) because only the first-order term in the Taylor series is considered in (7).

Therefore, the problem of calculating the CDF reduces to identifying the MPP on the curve of demarcation. This problem can be posed as a constrained optimization problem, as

$$\begin{gathered} \underset{\boldsymbol{u}}{\text{Minimize}}\ \boldsymbol{u}^T \boldsymbol{u} \\ \text{s.t. } G(\boldsymbol{x}) = r \\ \boldsymbol{u} = \{u_1, u_2 ... u_i, ... u_n\} \\ u_i = (x_i - \mu_i)/\sigma_i \\ (i = 1 \text{ to } n) \end{gathered} \tag{19}$$

The above optimization problem can be solved using the Rackwitz-Fiessler [49] algorithm, an iterative procedure, as follows.

1) Initialize counter $j = 0$ and start with an initial guess for the MPP, i.e., $\boldsymbol{x^j} = \{x_1^j, x_2^j, ... \ x_i^j, ...x_n^j\}$, a column vector.
2) Transform into standard normal space and calculate $\boldsymbol{u^j} = \{u_1^j, u_2^j, ... \ u_i^j, ...u_n^j\}$ using (15), a column vector.
3) Compute the gradient vector in the standard normal space, i.e., $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ...\alpha_n\}$, another column vector where
$$\alpha_i = \frac{\partial G}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\partial x_i}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \sigma_i \tag{20}$$
4) In the iterative procedure, the next point u^{j+1} is calculated using a Newton-Raphson type equation, as
$$\boldsymbol{u^{j+1}} = \frac{1}{||\boldsymbol{\alpha}||} [\boldsymbol{\alpha}^T \boldsymbol{u^j} - G(\boldsymbol{x^j})] \frac{\boldsymbol{\alpha}}{||\boldsymbol{\alpha}||} \tag{21}$$
5) Transform back into the original space, i.e., compute $\boldsymbol{x^{j+1}}$, and continue at Step 3 until the iterative procedure converges. Using tolerance limits δ_1 and δ_2, convergence can be verified if two criteria are satisfied: (i) the point lies on the curve of demarcation, i.e., $|G(\boldsymbol{x^j}) - r| \leq \delta_1$; and (ii) the solution does not change between two iterations, i.e., $|\boldsymbol{x^{j+1}} - \boldsymbol{x^j}| \leq \delta_2$.

The above described iterative procedure usually converges within 4 or 5 iterations. Sometimes, due to the lack of a convex objective function, it may be challenging to calculate the MPP. This may happen when the distribution of RUL is multi-modal. In such cases, it may be necessary to use alternative methods to calculate the uncertainty in RUL; this will be considered in future work.

In this paper, it is assumed that the optimal MPP can be calculated by using the above algorithm for optimization. In each iteration of the above algorithm, note that the transformation to the standard normal space is straightforward only when the variables are originally normal. Therefore, the method needs to be modified to account for non-normal variables, as explained in the next subsection.

B. Extension to Non-normal Variables

Now consider the case where the inputs X_i ($i = 1$ to n) have arbitrary probability distributions given by their CDFs as $F_{X_i}(x_i)(i = 1 \text{ to } n)$. Because X_i is not normally distributed, (15) cannot be used for a standard normal transformation. There, it is necessary to calculate u_i from a given x_i meaningfully, so that u_i represents a realization of the standard normal variable. One simple transformation is based on the probability integral transform concept [50] as

$$u_i = \Phi^{-1}(F_{X_i}(X_i = x_i)) \tag{22}$$

where $\Phi^{-1}(.)$ refers to the inverse of the standard normal distribution function [39]. Now, the calculation of the gradient in the standard normal space is different from (20), and can be derived directly using (22), as follows. First, decompose (22) into two parts as

$$v_i = F_{X_i}(X_i = x_i) \tag{23}$$

$$u_i = \Phi^{-1}(v_i). \tag{24}$$

Then, each element of the gradient vector $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ...\alpha_n\}$ can be calculated as

$$\alpha_i = \frac{\partial G}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\partial x_i}{\partial v_i} \times \frac{\partial v_i}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\phi(u_i)}{f_{X_i}(x_i)} \tag{25}$$

where $\phi(.)$ refers to the standard normal density function, and $f_{X_i}(x_i)$ is the PDF of the i^{th} input variable X_i.

In addition to the above procedure, there are also other transformation techniques. For example, a two-parameter transformation [39] procedure estimates the mean μ_i and standard deviation σ_i of an equivalent normal distribution by solving two simultaneous equations in μ_i and $sigma_i$. These two simultaneous equations equate the original CDF and PDF values of X_i to the respective CDF and PDF values of the equivalent normal distribution with mean μ_i and σ_i. Then, (15) can be used to calculate u_i from x_i. Note that the mean μ_i and standard deviation σ_i are dependent on the value of x_i. Similarly, Chen and Lind [51] proposed a three-parameter transformation procedure by introducing a third parameter, a scale factor which is estimated by matching the slope of the probability density function in addition to the PDF and CDF values. Further, when the inputs are correlated or *s*-dependent, it is necessary to transform them to the *uncorrelated* standard normal space. Haldar and Mahadevan [39] describe methods for such transformations. It must be noted that any transformation must be accompanied by suitably computing derivatives in the standard normal space, and (20) must be appropriately replaced.

Note that the above FORM procedure calculates the CDF value at a particular value of RUL. It answers a question: what is the probability that the RUL is smaller than a given number? To obtain the entire CDF, the whole procedure is repeated with multiple values of RUL. Sometimes, it may not be possible to identify values of RUL to calculate the entire CDF because the spread of the distribution may not be known in advance. So, the next section discusses the inverse-FORM procedure, which answers a different question: what is the value of RUL that corresponds to a given probability level? In other words, what is the α-percentile (e.g., 5%, 95%, etc.) value of RUL? By repeating this procedure for one lower percentile and one upper percentile value, the probability bounds on RUL can be calculated.

VI. Inverse First-Order Reliability Method

Given β or $\lambda = \Phi(-\beta)$, the inverse-FORM approach can be used to calculate r such that $F_R(r) = P(R \leq r) = \lambda$. The theory behind inverse-FORM is exactly the same as FORM, and the algorithm discussed in Section V is modified so that the CDF value can be specified and r can be calculated. The various steps involved in the iterative procedure for inverse-FORM are outlined below.

1) Initialize counter $j = 0$, and start with an initial guess for the MPP, i.e., $\boldsymbol{x^j} = \{x_1^j, x_2^j, ... \ x_i^j, ...x_n^j\}$.
2) Transform into the standard normal space, and calculate $\boldsymbol{u^j} = \{u_1^j, u_2^j, ... \ u_i^j, ...u_n^j\}$.
3) Compute the gradient vector in the standard normal space, i.e., $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ...\alpha_n\}$, as explained in Section V-A.
4) In the iterative procedure, the next point u^{j+1} is calculated as
$$\boldsymbol{u^{j+1}} = -\frac{\boldsymbol{\alpha}}{|\boldsymbol{\alpha}|}\beta. \tag{26}$$
5) Transform back into the original space, i.e., compute $\boldsymbol{x^{j+1}}$, and continue starting from Step 3 until the iterative procedure converges. Using tolerance limits δ_1 and δ_2, convergence can be verified if two criteria are satisfied: (i) the point lies on the curve of demarcation, i.e., $|G(\boldsymbol{x^j}) - r| \leq \delta_1$; and (2) the solution does not change between two iterations, i.e., $|\boldsymbol{x^{j+1}} - \boldsymbol{x^j}| \leq \delta_2$.

Similar to the Rackwitz-Fiessler algorithm, the above iterative procedure also convergences within 4 or 5 iterations, and therefore is suitable for quick calculations in an online heath monitoring context.

As stated earlier, the inverse-FORM procedure is useful to calculate probability bounds. For example, by repeating the above algorithm for $\lambda = 0.05$ and $\lambda = 0.95$, it is possible to estimate the 90% probability bounds on the remaining useful life. In fact, the entire CDF can be constructed by repeating the analysis for several values of λ $(0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99)$.

The advantage of inverse-FORM is that the number of prognostic evaluations is extremely small compared to sampling-based approaches. For the sake of illustration, consider that the above iterative algorithm converges in 5 iterations. Each iteration requires the computation of the gradient vector; if there are n inputs, then $n + 1$ computations are necessary for each evaluation: one for each derivative, and one for the evaluation of G. Hence, the total number of prognostic evaluations is equal to $5(n+1)$. If the analysis is repeated for 2 values of λ (to compute the probability bounds on RUL), then it takes $10(k + 1)$ prognostic evaluations, which is more computationally efficient than sampling-based techniques, and therefore suitable for online decision-making.

Therefore, this paper recommends the use of the inverse-FORM approach for RUL uncertainty quantification; while

FOSM is not accurate (due to the choice of point of linearization), FORM requires the algorithm to choose values of RUL for which the CDF needs to be calculated, and this choice may not be obvious in practical applications. However, the discussion on FOSM and FORM preceded the discussion on inverse-FORM so as to motivate and develop the concept of MPP, and culminate in the development of an algorithm for inverse-FORM. In fact, the numerical example in Section VIII will be solved using inverse-FORM; however, prior to that, it is necessary to address certain practical challenges in the implementation of the above described algorithm.

VII. Practical Challenges

The previous sections discussed the mathematical theory behind the application of FORM-based methods for uncertainty quantification in RUL prediction. This section identifies certain practical challenges which may be encountered during the implementation of FORM-based methods for online health monitoring, and proposes statistical methods to overcome them.

1) **Implicitness of** G: The FORM method and the inverse-FORM can be applied in the presence of an explicit function $R = G(\boldsymbol{X})$. However, from Section II, it may be seen that G is actually implicit, because the number of terms in $\boldsymbol{X}$ depends on R. In this paper, by choosing constant amplitude loading, $\mathbf{u}(t)$ becomes invariant with time, and can be replaced by $\mathbf{u}$. On the other hand, the parameter values $\boldsymbol{\theta}(t)$ are pre-determined (without uncertainty), and hence don't have to be included for uncertainty propagation in $\boldsymbol{X}$. However, the process noise terms ($\mathbf{V}(k) = [\mathbf{v}(k_P), \mathbf{v}(k_P+1), \mathbf{v}(k_P+2), ..., \mathbf{v}(k^{E(t_P)})]$) still need to be included in $\boldsymbol{X}$. It is clear that the number of terms to include depends on R, and hence renders G implicit. Note that the number of terms is a function of the time-discretization chosen to solve the differential equation in (1).
2) **Curse of dimensionality**: Even if the number of process noise terms to be included may be known, there are computational difficulties, because R (i.e., the remaining useful life) may be on the order of hundreds or thousands (or more), and therefore the length of the vector $\boldsymbol{X}$ will be of the same order. Therefore, during the implementation of FORM or inverse-FORM, it is necessary to compute the gradient of G, and this computation will involve several hundreds or thousands (or more) of evaluations of G. Thereby, the computational effort involved becomes comparable with sampling-based approaches, and hence may not be suitable for online health monitoring.

This paper proposes a new likelihood-based method that overcomes both of the above challenges. First, suppose that the time-variant process noise is replaced with a time-invariant constant value denoted by $\mathbf{v}^E$. In other words,

$$\mathbf{v}(k) \coloneqq \mathbf{v}^E \quad \forall \quad k \in [k_P, k^{E(t_P)}] \tag{27}$$

The above equation means that the same realization of process noise will be used for prediction at every future time instant. This equation does not imply that a time-varying process, i.e., a random vector, can be replaced by a time-invariant constant, i.e., a random variable. Instead, the goal is to estimate an *appropriate probability distribution* for $\mathbf{v}^E$ that can capture the effect of the true time-varying process noise on the RUL distribution. In other words, the goal is to choose a suitable probability distribution for $\mathbf{v}^E$, so that the effect of propagating this distribution through G is equivalent to propagating the original distribution of $\mathbf{V}(k)$ through G. Because this probability distribution for $\mathbf{v}^E$ and the probability distribution of the actual time-varying process noise lead to the same distribution of the RUL, $\mathbf{v}^E$ is referred to as the equivalent time-invariant process noise, and its probability distribution is referred to as the equivalent time-invariant process noise distribution.

Further, this equivalence can be numerically verified using Monte Carlo sampling. First, generate multiple samples of $\mathbf{V}(k)$ along with the other sources of uncertainty, and calculate corresponding samples of $R(k_P)$ (using the algorithm in Fig. (2)) that may be used to construct the probability distribution of RUL. Alternatively, select multiple samples of $\mathbf{v}^E$, along with the other sources of uncertainty, use this constant value of process noise at every time instant in the algorithm in Fig. (2) for every sample, and calculate the corresponding samples of $R(k_P)$ that may be used to construct an alternative, equivalent distribution for RUL. The challenge is to estimate the distribution of $\mathbf{v}^E$ from which samples may be drawn so that the equivalent probability distribution of RUL (estimated using multiple samples of equivalent time-invariant process noise $\mathbf{v}^E$) is exactly the same as the original probability distribution of RUL (estimated using multiple samples of the original process noise $\mathbf{V}(k)$.

Such distribution of the equivalent time-invariant process noise can be estimated by computing the likelihood of $\mathbf{v}^E$ such that (27) is satisfied. In other words, any value of $\mathbf{v}^E$ has an associated probability with which (27) is satisfied; the likelihood of $\mathbf{v}^E$ is proportional to this probability. The probability distribution of the true process noise can be used to calculate this likelihood as

$$L(\mathbf{v}^E | R(k_P)) \propto \prod_{k=k_P}^{k=k^{E(t_P)}} f_{\mathbf{V}(k)}(\mathbf{V}(k) = \mathbf{v}^E) \tag{28}$$

where $f_{\mathbf{V}(k)}(\mathbf{v}(k))$ is the probability density function of the true process noise $\mathbf{V}(k)$. Also note that the likelihood function is conditioned on the RUL, and written as $R(k_P)$. Further, the above equation assumes that the process noise values at two different times are *s*-independent of each other. If any statistical dependence is unknown, then it can be easily included in the above equation by conditioning appropriately. Having calculated the likelihood, the PDF of $\mathbf{v}^E$ can be calculated as [52]

$$f_{\mathbf{V}^E}(\mathbf{v}^E | R(k_P)) = \frac{L(\mathbf{v}^E | R(k_P))}{\int_D L(\mathbf{v}^E | R(k_P)) d\mathbf{v}^E}. \tag{29}$$

In (29), the domain of integration D is chosen such that $\mathbf{v}^E \in D$ iff $L(\mathbf{v}^E) \neq 0$. Note that the definition of $\mathbf{v}^E$ and its probability distribution are related only to the uncertainty

in $R(k_P)$, and not to the Most Probable Point in the FORM method. The only condition is that propagating the uncertainty in $\mathbf{v}^E$ through G needs to be equivalent to propagating the uncertainty in $\mathbf{V}(k)$ through G by leading to the same distribution for RUL.

Now the function G is not implicit because $\mathbf{v}^E$ is simply a time-invariant constant s-independent of $R(k_P)$. However, the PDF of $\mathbf{v}^E$ is conditionally s-dependent on $R(k_P)$, and this s-dependence can be expressed explicitly. In other words, for a given realization of $\mathbf{v}^E$ (along with a realization of the present state and future loading conditions), the value of $R(k_P)$ can be computed using the definition of G in Fig. (2) by substituting $\mathbf{v}(k) = \mathbf{v}^E$ for all k. The calculated value of $R(k_P)$ can then be used to calculate the PDF $f_{\mathbf{V}^E}(\mathbf{v}^E|R(k_P))$. For implementation through FORM or inverse-FORM, instead of the true process noise, the equivalent time-invariant process noise is considered to be a part of $\boldsymbol{X}$ in $R = G(\boldsymbol{X})$. Therefore, first an initial value of $\mathbf{v}^E$ is assumed (along with initial assumptions for the present state and future loading conditions), and for this initial assumption, both $R(k_P)$ and $f_{\mathbf{V}^E}(\mathbf{v}^E|R(k_P))$ can be calculated. Then, this PDF is transformed to the standard normal space using (22), and the gradient in the normal space can be computed using (25). Now, the iterative algorithms for FORM, and inverse-FORM (in Section V, and Section VI respectively) can then be used to update the value of $\mathbf{v}^E$, along with the other elements of $\boldsymbol{X}$; when the algorithm converges, the optimal value of $\mathbf{v}^E$, and the corresponding value of $R(k_P)$, can be estimated.

In this section, the time-variant process noise has been replaced by a time-invariant equivalent, and the probability distribution of the latter has been chosen analytically based on the probability distribution of the former. This equivalent time-invariant distribution has been applied only to process noise in this paper. This concept has the potential to be applied to time-variant model parameters, and variable-amplitude loading conditions also; this possibility will be explored in future research.

VIII. Case Study: Lithium-Ion Battery

This section presents a numerical example consisting of a lithium-ion battery that powers an unmanned aerial vehicle [53] at NASA Langley Research Center. This unmanned aerial vehicle is being used as a test-bed for prognostics and decision-making at NASA Ames Research Center, and NASA Langley Research Center. The goal is to estimate the end of discharge of the battery, which is indicative of the remaining flight time; therefore, the end of discharge is synonymous with the end of life, for the sake of this numerical example. Though this numerical example deals with a battery model, the proposed uncertainty quantification methodology is general, and can be applicable to different types of models in several engineering domains.

A. Description of the Model

The battery model, extended from that used by Daigle et al. [14] for prognosis, is similar to the models presented by Chen and Rincon-Mora [54]. The model is based on an electrical circuit equivalent to that shown in Fig. (4), where the large capacitance C_b holds the charge q_b of the battery. The nonlinear C_b captures the open-circuit potential, and concentration overpotential. The R_{sp}-C_{sp} pair captures the major nonlinear voltage drop due to surface overpotential, R_s captures the so-called Ohmic drop, and R_p models the parasitic resistance that accounts for self-discharge. This empirical battery model is sufficient to capture the major dynamics of the battery, but ignores temperature effects and other minor battery processes. The governing equations for the battery model are presented in continuous time below. The implementation of the proposed methodology considers a discrete-time version with a discrete time-step of 1 second.

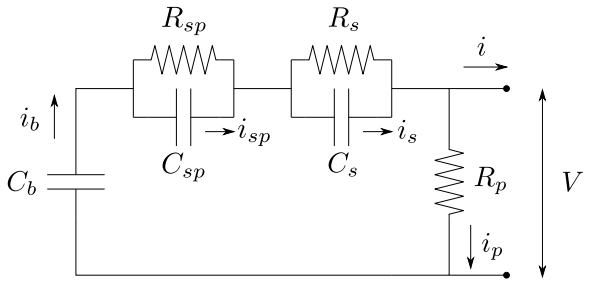

Fig. 4: Battery equivalent circuit.

The state-of-charge, SOC, is computed as

$$SOC = 1 - \frac{q_{max} - q_b}{C_{max}}, \tag{30}$$

where q_b is the current charge in the battery (related to C_b), q_{max} is the maximum possible charge, and C_{max} is the maximum possible capacity. The resistance related to surface overpotential is a nonlinear function of SOC:

$$R_{sp} = R_{sp_0} + R_{sp_1} \exp\left(R_{sp_2}(1 - SOC)\right), \tag{31}$$

where R_{sp_0}, R_{sp_1}, and R_{sp_2} are empirical parameters. The resistance, and hence the voltage drop, increases exponentially as SOC decreases.

Voltage drops across the individual circuit elements are given by

$$V_b = \frac{q_b}{C_b}, \tag{32}$$

$$V_{sp} = \frac{q_{sp}}{C_{sp}}, \tag{33}$$

$$V_s = \frac{q_s}{C_s}, \tag{34}$$

$$V_p = V_b - V_{sp} - V_s, \tag{35}$$

where q_{sp} is the charge associated with the capacitance C_{sp}, and q_s is the charge associated with C_s. The voltage V_b is also the open-circuit voltage of the battery, which is a nonlinear function of SOC [54]. This is captured by expressing C_b as a third-order polynomial function of SOC:

$$C_b = C_{b_0} + C_{b_1} SOC + C_{b_2} SOC^2 + C_{b_3} SOC^3. \tag{36}$$

The terminal voltage of the battery is

$$V = V_b - V_{sp} - V_s. \tag{37}$$

TABLE I: Battery Model Parameters

Parameter	Value	Unit
C_{b_0}	19.80	Farad (F)
C_{b_1}	1745.00	Farad (F)
C_{b_2}	-1.50	Farad (F)
C_{b_3}	-200.20	Farad (F)
R_s	0.0067	Ohm (Ω)
C_s	115.28	Farad (F)
R_p	1×10^4	Ohm (Ω)
C_{sp}	316.69	Farad (F)
R_{sp_0}	0.0272	Ohm (Ω)
R_{sp_1}	1.087×10^{-16}	Ohm (Ω)
R_{sp_2}	34.64	(No unit)
q_{max}	3.11×10^4	Coulomb (C)
C_{max}	30807	Coulomb (C)
V_{EOD}	16	Volt (V)

Currents associated with the individual circuit elements are given by

$$i_p = \frac{V_p}{R_p}, \tag{38}$$

$$i_b = i_p + i, \tag{39}$$

$$i_{sp} = i_b - \frac{V_{sp}}{R_{sp}}, \tag{40}$$

$$i_s = i_b - \frac{V_s}{R_s}. \tag{41}$$

The charges are then governed by

$$\dot{q}_b = -i_b, \tag{42}$$

$$\dot{q}_{sp} = i_{sp}, \tag{43}$$

$$\dot{q}_s = i_s. \tag{44}$$

It is of interest to predict the end-of-discharge as defined by a voltage threshold V_{EOD}. So, C_{EOL} consists of only one constraint:

$$c_1 : V > V_{EOD}. \tag{45}$$

All voltages are measured in Volts, resistances are measured in Ohms, charges are measured in Coulombs, and capacitances are measured in Coulombs per Volt (or Farads). Note that C_{b_0}, C_{b_1}, C_{b_2}, and C_{b_3} are simply fitting parameters in (36), and do not have physical meaning.

Though the proposed methods can account for parameter uncertainty, the parameters of this battery model are assumed to be deterministic, and are shown in Table I. If these parameters are estimated to be uncertain, they can be represented using probability distributions and included in the uncertainty quantification procedure, as indicated earlier in Fig. (2), i.e., by treating the uncertain parameters as a part of $\boldsymbol{X}$ in $R = G(\boldsymbol{X})$.

The following sections deal with uncertainty quantification in RUL prediction using the inverse-FORM approach. While three methods (FOSM, FORM, and inverse-FORM) were discussed in this paper, it is intuitive that the inverse-FORM is most suitable for RUL uncertainty quantification. As explained earlier in Section IV, FOSM is not accurate because it chooses the mean of uncertain variables as the point of linearization; however, it mathematically motivates the FORM method and the inverse-FORM method. While both the FORM and inverse-FORM approaches can be useful in prognostics, FORM can be used only to calculate the probability that the RUL is smaller than or greater than a given value. This value needs to be selected by the algorithm, and this choice may become arbitrary. Instead of specifying an arbitrary value of RUL (which may not be known in a practical scenario), inverse-FORM is used to quantify those values of RUL that correspond to different probability levels such as 0.01, 0.1, 0.5, 0.90, 0.99, etc.

B. Sources of Uncertainty

The different sources of uncertainty considered in this case study are listed below.

1) **Loading Uncertainty** - Saha et al. [53] quantified the uncertainty in the loading for a battery used to power an unmanned aerial vehicle. Several flight segments were identified, and the loading was quantified in each segment. At present, the inverse-FORM methodology has not been extended to consider variable amplitude loading. Hence, a constant amplitude input, which is taken to be the average current value across multiple flight segments, is considered in this example; however, the constant amplitude is chosen to be random. In particular, the constant amplitude (in Amps) is considered to be normally distributed ($N(35, 5)$), and this distribution is truncated at a specified lower bound (5.0) and upper bound (80) pair. Future research will rigorously consider the inclusion of variable amplitude loading; the equivalent time-invariant distribution concept may be investigated for this purpose.
2) **State Uncertainty** - Typically, the state estimation, which is an inverse problem, is addressed using a filtering technique that can continuously estimate the uncertainty in the state when measurements are continuously available as a function of time. In this paper, the state estimation is not explicitly carried out. The state values are assumed to be available, and the uncertainty in the states is predetermined based on the authors' past experiences with the use of filtering techniques for the above described problem. There are three state variables in this example: charge in C_b, charge in C_{sp}, and charge in C_s. At any time instant, they are assumed to be normally distributed with a specified mean (shown in Figs. (5) through (7)); for example, the mean of the initial values of the three states are set as [3.1×10^4, 0, 0]. For the purpose of illustration, three different values of CoV (Coefficient of variation, defined as the ratio between standard deviation and mean) are considered; those values are 0.05, 0.10, and 0.20. The analysis is repeated for each CoV value.
3) **Process Noise Uncertainty** - At any time instant, there are three states, and hence three process noise terms. All the three process noise terms are assumed to have zero mean, and variances equal to 1, 1×10^{-4}, and 1×10^{-6} respectively. For the sake of illustration, it is assumed that the three process noise terms are s-independent, and further these process noise values at two different time instants are also s-independent of each other.

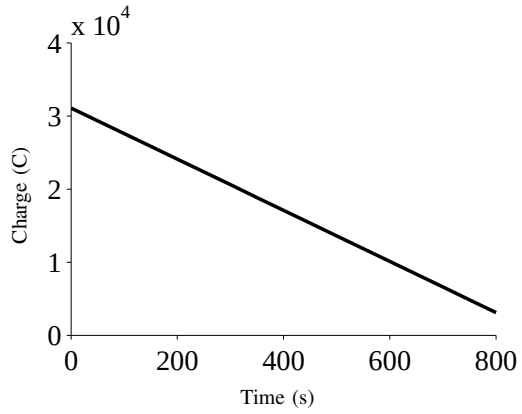

Fig. 5: State No. 1: Charge in C_b.

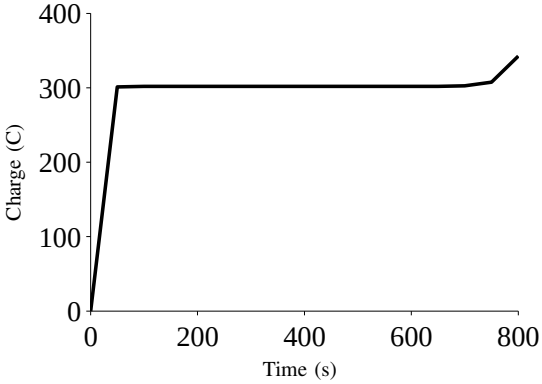

Fig. 6: State No. 2: Charge in C_{sp}.

Therefore, loading uncertainty is represented by one variable, present state uncertainty is represented using three variables, and each state variable is associated with a corresponding process noise term. As each of the time-varying process noise quantities is replaced with time-invariant process noise, there are 7 uncertain quantities. Hence, the dimension of $\boldsymbol{X}$ in $R = G(\boldsymbol{X})$ is equal to 7. Given a realization of each uncertain variable, the corresponding realization of R can be calculated using the following iterative algorithm. Note that this algorithm is simply a specific case of the more general algorithm in Fig. (2).

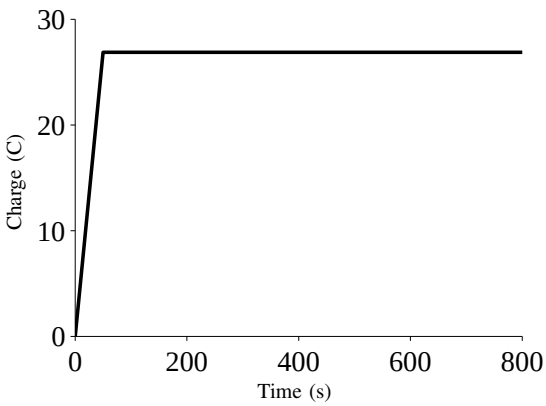

Fig. 7: State No. 3: Charge in C_s.

1) Given a prediction time instant t_P, calculate the corresponding time-index k_P.
2) Also given are realizations of uncertain variables, i.e., loading ($\mathbf{u}$); present state estimate ($\mathbf{x}(k_P)$; a vector consisting of 1) charge in C_b, 2) charge in C_{sp}, and 3) charge in C_s at time-index k_P); and process noise ($\mathbf{V}(k_P)$). Note that $\mathbf{V}(k_P)$ is a matrix of the three process noise terms at multiple time-instants until the end of discharge. When the equivalent time-invariant process noise $\mathbf{v}^E$ is used, the same value of process noise is used at all time-instants. Note that the equivalent process noise, i.e., $\mathbf{v}^E$ is a vector consisting of the three process noise terms that correspond to the three state variables.
3) Set $k = k_P$.
4) For a given time-index k, discretize (30) through (44) using the Euler approach with a sampling time of 1 second to calculate all the state values, currents, and voltages at the next time-index $k+1$. At this step, use the process noise realization $\mathbf{v}(k)$. Because the time-invariant equivalent process noise is used, $\mathbf{v}^E$ will be used at every time-index.
5) Check if constraint c_1 is violated, i.e., if the condition $V > V_{EOD}$ is violated. If violated, then $T_{EOL} = 1$; otherwise, $T_{EOL} = 0$.
6) If $T_{EOL} = 0$, then it is necessary to continue further prediction, and therefore set $k = k + 1$. Go to Step 4.
7) If $T_{EOL} = 1$, then the end of life is reached, and k corresponds to the time-index of end of life. Therefore t^k corresponds to the time of the end of life, and $R = t^k - t_P$.

C. Verification of Equivalent Time-Invariant Process Noise

For the given specification of process noise, it can be easily shown that, if the true distribution of the process noise follows a normal distribution with mean 0, and standard deviation σ, then the equivalent time-invariant process noise follows a normal distribution with mean 0, and standard deviation $\frac{\sigma}{\sqrt{R}}$, where R is the remaining useful life prediction calculated using G. Before the equivalent process noise is directly used in the first-order reliability method, its equivalence is numerically demonstrated through Monte Carlo sampling.

To verify the mathematical accuracy of the proposed time-invariant process noise approach, it is necessary to compare two probability distributions of RUL that are obtained by using the 1) true process noise, and 2) the equivalent time-invariant process noise. To make a fair comparison, the other sources of uncertainty (loading and present state) are assumed to be absent; in other words, deterministic values (corresponding to the mean) are chosen for $\mathbf{u}$, and $\mathbf{x}(k_P)$. These two distributions of RUL are calculated at $t_P = 0$, as follows.

1) **Distribution of RUL using actual process noise** - Select one sample of an entire trajectory of the actual process noise. While generating this trajectory, different

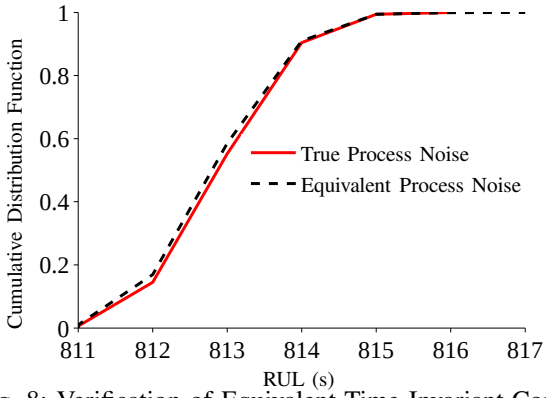

Fig. 8: Verification of Equivalent Time-Invariant Concept.

values of the process noise are sampled at every time-instant. For this chosen sample, and given values of $\mathbf{u}$ and $\mathbf{x}(k_P)$, use the above described algorithm to calculate a corresponding sample of $R(k_P)$. Repeat the entire procedure for multiple samples of uncertain quantities, and obtain corresponding samples of $R(k_P)$. Using these samples, the CDF of the RUL can be constructed, as shown using a solid line in Fig. (8).

2) **Distribution of RUL using equivalent time-invariant process noise** - Select one sample of the equivalent time-invariant process noise based on its distribution, and use it as the process noise value at every time-instant, thereby resulting in a constant trajectory for the entire process noise. Using the given values of $\mathbf{u}$ and $\mathbf{x}(k_P)$, the above described algorithm can again be used to calculate a corresponding sample of $R(k_P)$. Finally, repeat for multiple samples of $\mathbf{v}^E$, and calculate multiple samples of $R(k_P)$. Using these samples, the CDF of the RUL can be constructed, as shown using a dotted line in Fig. (8).

As seen from Fig. (8), the distribution resulting from the true process noise and the distribution resulting from the equivalent time-invariant process noise compare well with each other; the minor difference can be attributed due to errors which occur due to sampling. Thus, the concept of equivalent time-invariant process noise which was theoretically established using (28) has been numerically verified in Fig. (8). Here on, the time-invariant equivalent process noise can be used in lieu of the true process noise for uncertainty quantification along with the other uncertain variables. The results of uncertainty quantification using inverse-FORM are presented in the following subsection.

D. Uncertainty Quantification in RUL

The inverse-FORM method is directly used to compute the RUL value corresponding to $F_R(r_1) = 0.01$, $F_R(r_2) = 0.5$, and $F_R(r_3) = 0.99$, for each CoV. While r_1 and r_3 correspond to the 98% probability bounds of RUL, r_2 corresponds the median of RUL. The bounds and the mean are continuously calculated until $t_P = 800$ seconds when failure seems to be imminent, and the inverse-FORM calculation is performed every 50 seconds. The results of RUL calculation (r_1, r_2, and r_3, in seconds) are tabulated in Table II, and graphically shown in Figs. (9) through (11).

TABLE II: Results of Uncertainty Quantification in RUL

t_P (sec)	CoV=0.05	CoV=0.10	CoV=0.15
0	$r_1 = 534$ $r_2 = 813$ $r_3 = 1278$	$r_1 = 501$ $r_2 = 813$ $r_3 = 1327$	$r_1 = 444$ $r_2 = 813$ $r_3 = 1396$
200	$r_1 = 388$ $r_2 = 613$ $r_3 = 975$	$r_1 = 365$ $r_2 = 613$ $r_3 = 1012$	$r_1 = 324$ $r_2 = 613$ $r_3 = 1066$
400	$r_1 = 240$ $r_2 = 413$ $r_3 = 671$	$r_1 = 227$ $r_2 = 413$ $r_3 = 697$	$r_1 = 202$ $r_2 = 413$ $r_3 = 736$
600	$r_1 = 93$ $r_2 = 213$ $r_3 = 368$	$r_1 = 87$ $r_2 = 213$ $r_3 = 383$	$r_1 = 77$ $r_2 = 213$ $r_3 = 406$
T=800	$r_1 = 1$ $r_2 = 13$ $r_3 = 64$	$r_1 = 1$ $r_2 = 13$ $r_3 = 68$	$r_1 = 1$ $r_2 = 13$ $r_3 = 71$

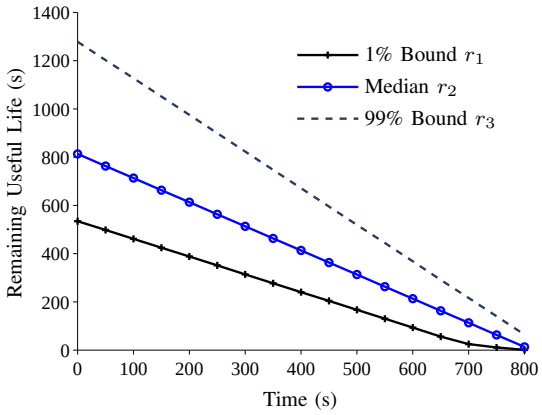

Fig. 9: 98% Bounds and Median of RUL (CoV = 0.05).

It is seen from the results that the uncertainty in the RUL is high initially, and then gradually decreases until failure is imminent. Initially, the uncertainty in RUL is high because it is necessary to predict at a farther time instant; future loading, and the associated uncertainty, need to be considered for a longer period of time. However, at a latter time instant, future loading needs to be assumed for a reduced period of time, and hence the uncertainty in the RUL decreases. In fact, any good prognostic algorithm should depict this behavior, i.e., the prediction of RUL at a later time instant must have lower uncertainty than the prediction at an earlier time instant. Further, the larger the coefficient of variation in the assumed state estimates, the larger the uncertainty in RUL; this behavior is observed at every time instant, and is consistent with intuition because a larger uncertainty in the state estimate implies that the corresponding RUL prediction uncertainty will also be high.

E. Verification using Monte Carlo Sampling

To verify the above performed uncertainty quantification, Monte Carlo sampling (MCS) was performed using 1000 samples. It is computationally infeasible to perform MCS at every time instant considered. Therefore, MCS is performed for five time instants, starting from $t_P = 0$ seconds until $t_P = 800$ seconds, in steps of 200 seconds. At each

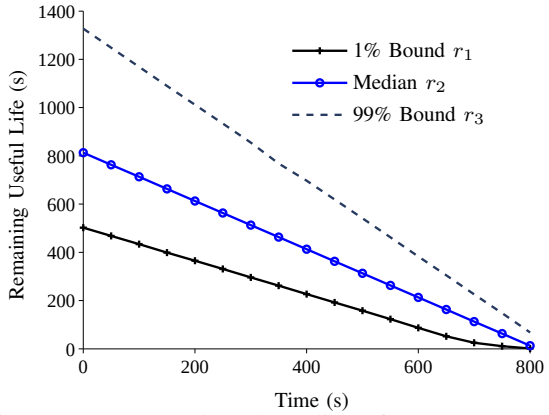

Fig. 10: 98% Bounds and Median of RUL (CoV = 0.10).

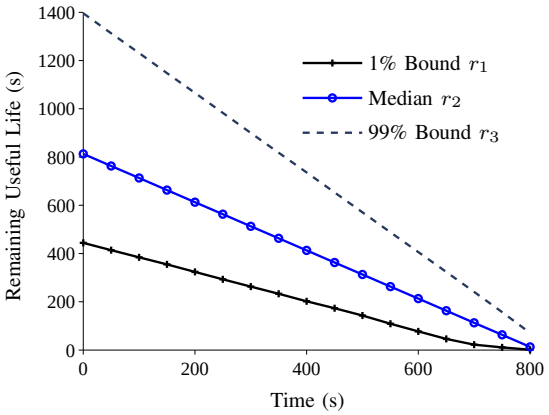

Fig. 11: 98% Bounds and Median of RUL (CoV = 0.15).

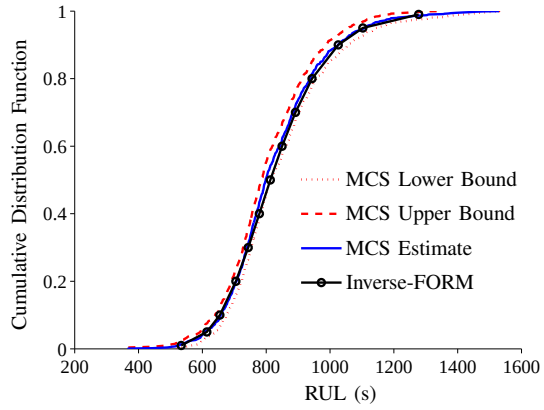

Fig. 12: Inverse-FORM vs. MCS ($t_P = 0$, CoV $= 0.05$).

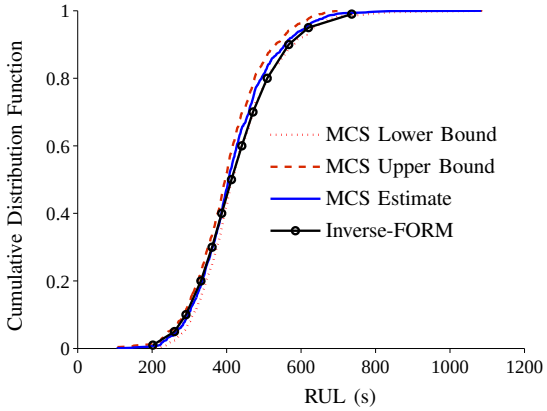

Fig. 13: Inverse-FORM vs. MCS ($t_P = 400$, $CoV = 0.15$).

time instant, the CDF of the RUL was computed by repeating the inverse-FORM for 13 different λ values ($\lambda = 0.01, 0.05, 0.1, 0.2, ...0.8, 0.9, 0.95, 0.99$). The comparison of inverse-FORM and MCS at $t_P = 0$ is shown in Fig. (12); in this illustration, the CoV of the state estimates was chosen to be equal to 0.05. Note that the uncertainty bounds due to the use of a limited number of samples for MCS are also shown.

The probability distribution resulting from inverse-FORM lies within the Monte Carlo bounds, as seen in Fig. (12), thus verifying the uncertainty quantification procedure. In fact, the maximum difference between the inverse-FORM solution and the Monte Carlo solution was found to be less than 0.5% (it must be noted that the Monte Carlo solution is not exact due to the use of limited samples, and hence the 90% Monte Carlo bounds have also been provided in Fig. (12)). Further, the inverse-FORM procedure needed only $13 \times 4 \times 7 = 364$ prognostic evaluations, because 13 λ values and 7 variables (3 state quantities, 3 process noise terms, and one loading term, to represent uncertain quantities) were used, and 4 iterations were needed for convergence whenever the iterative inverse-FORM algorithm was used. In contrast, 1000 prognostic evaluations were used in MCS. Further, a similar comparison was performed at different time steps, and by considering other CoV values. For instance, the comparison of the results from MCS and inverse-FORM at $t_P = 400$ s for CoV $= 0.15$ is shown in Fig. (13).

The agreement of inverse-FORM with MCS is also evident in Fig. (13). In fact, it was observed that the inverse-FORM method performs well in comparison with Monte Carlo sampling, both in terms of accuracy (the maximum difference was found to be less than 1%), and computational cost (approximately one-third of the cost for MCS).

F. Effect of Assuming the Mean of State Variables

Because the focus of this paper is on prognosis, it was stated earlier that the estimation of states is not explicitly carried out in this paper, and therefore the mean of the state values were assumed to be available as a function of time. It is important to note that these assumed values impact the final PDF of the RUL. Further, the estimated state values may not be monotonic, smooth functions as seen in Figs. (5) through (7). To demonstrate the effect of realistic state estimation results, the noisy (simulated) state values shown in Figs. (14) through (16) are assumed to be the mean of the states, and prognosis calculations are repeated to obtain the CDF of the RUL. For the sake of illustration, the CoV for the states is assumed to be 5%, and the results of uncertainty quantification are shown in Fig. (17). The statistics for process noise and loading uncertainty are assumed to be the same as before.

Similar to the previous studies, the results of uncertainty quantification were verified at multiple time instants, by

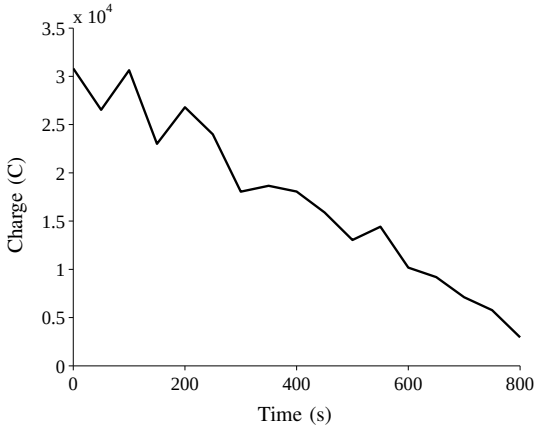

Fig. 14: State No. 1: Charge in C_b.

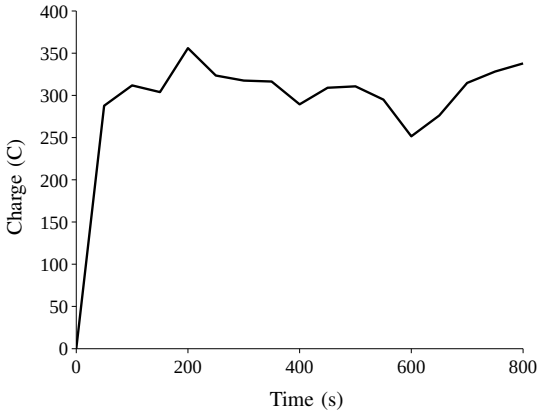

Fig. 15: State No. 2: Charge in C_{sp}.

comparing the results from the inverse-FORM approach and the Monte Carlo solution. The solution through the former approach was found to be within 1% of the solution using the latter approach.

IX. Conclusions

This paper investigated the use of the first-order reliability methods (FORM) to compute the uncertainty in the remaining useful life prediction of components used in engineering systems. The computation of remaining useful life is important in the context of component-level and system-level prognosis, and hence directly used in online operations and decision-making. Conventionally, sampling-based algorithms have been used for quantifying the uncertainty in prognostic calculations, and may require several thousands of system evaluations to quantify the uncertainty in remaining useful life (RUL) with reasonable accuracy. On the contrary, the FORM-based approaches are analytical, compute the uncertainty using a few system evaluations, and therefore are suitable for online prognosis. Further, FORM-based methods are invariant on repetition (as against sampling methods like Monte Carlo), and hence may be preferred for system certification purposes.

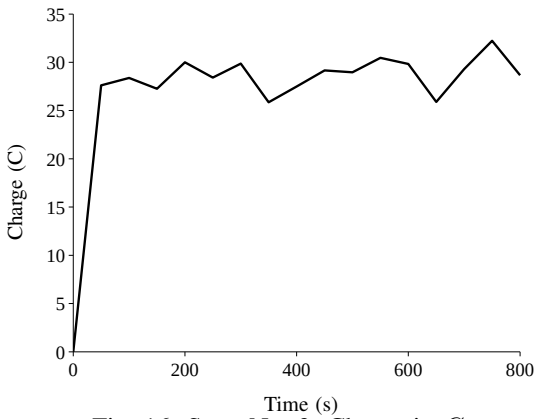

Fig. 16: State No. 3: Charge in C_s.

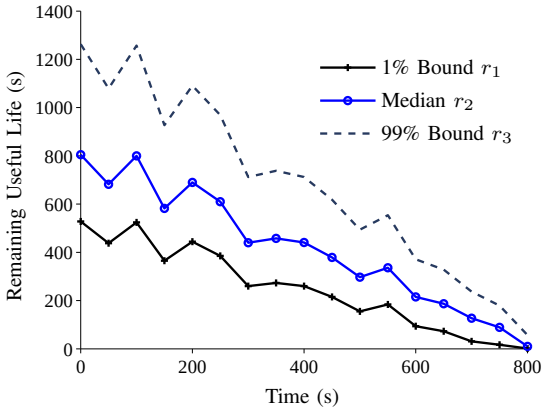

Fig. 17: 98% Bounds and Median of RUL (CoV = 0.05).

While FORM-based approaches have been commonly used in the reliability analysis of structural systems, this paper applied these methods to predict the remaining useful life of engineering components and systems using state-space models. We discussed three FORM-based methods in detail: the First-Order Second Moment method (FOSM), the First-Order Reliability Method (FORM), and the Inverse First-Order Reliability Method (inverse-FORM). While the FOSM method can be used to approximate the first two moments of the RUL prediction, the FORM and inverse-FORM approaches can be used to calculate the entire probability distribution of the RUL. In particular, the inverse-FORM method was preferred because it can calculate those values of RUL that correspond to specified probability levels, which can be useful in online decision-making. Further, several practical challenges involved in the application of FORM-based methods to RUL prediction were discussed, and new computational approaches were proposed to overcome these challenges. Finally, a numerical example was presented, and the uncertainty in remaining useful life of a lithium-ion battery was quantified by accounting for the different sources of uncertainty.

Future work needs to address several issues. First, practical systems are commonly subjected to different types of variable amplitude loading profiles such as block loading, random processes (Gaussian processes, Markov processes, etc.), and therefore, the proposed methods for uncertainty quantification need to be extended to consider variable amplitude loading. The assumption of constant amplitude loading implies

that loading uncertainty is described using a single random variable, whereas variable amplitude loading profiles need to be described using multiple random variables, which not only increases the dimensionality of the problem, but also affects the uncertainty in the RUL prediction. Sometimes the RUL distribution may become multi-modal, and first-order reliability methods may not be able to capture such probability distributions; this needs to be investigated in future work. Second, the extension of the proposed equivalent time-invariant distribution concept to variable amplitude loading may be investigated for this purpose. Third, sensitivity analysis needs to be performed so that the contributions of the different sources of uncertainty to the overall uncertainty in RUL can be quantified. Fourth, this paper did not consider the effect of model form uncertainty on prognosis; future research needs to quantify model form uncertainty, and develop a method to rigorously account for model uncertainty in prognosis and RUL calculations instead of simply lumping all the model uncertainty into the process noise terms. Finally, it is also necessary to quantify the robustness of the proposed approach, by estimating the sensitivity of the RUL bounds to relaxing the various assumptions in the present paper, and thereby investigate the applicability of the methodology to practical engineering systems.

ACKNOWLEDGMENT

The work reported herein was in part funded by the NASA System-wide Satefy Assurance Technologies (SSAT) project under the Aviation Safety (AvSafe) Program of the Aeronautics Research Mission Directorate (ARMD), and by the NASA Automated Cryogenic Loading Operations (ACLO) project under the Office of the Chief Technologist (OCT) of Advanced Exploration Systems (AES).

REFERENCES

[1] J. Luo, K. Pattipati, L. Qiao, and S. Chigusa, "Model-based prognostic techniques applied to a suspension system," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 38, no. 5, pp. 1156–1168, 2008.

[2] K. Goebel, B. Saha, and A. Saxena, "A comparison of three data-driven techniques for prognostics," in *62nd Meeting of the Society For Machinery Failure Prevention Technology (MFPT)*, 2008, pp. 119–131.

[3] M. Daigle, M. Foygel, and V. Smelyanskiy, "Model-based diagnostics for propellant loading systems," in *Aerospace Conference, 2011 IEEE*. IEEE, 2011, pp. 1–11.

[4] M. Daigle and K. Goebel, "A model-based prognostics approach applied to pneumatic valves," *International Journal of Prognostics and Health Management*, vol. 2, no. 2, 2011.

[5] M. Daigle, A. Bregon, and I. Roychoudhury, "A distributed approach to system-level prognostics," in *Annual Conference of the Prognostics and Health Management Society*, sep 2012, pp. 71–82.

[6] S. Sankararaman and S. Mahadevan, "Uncertainty quantification in structural damage diagnosis," *Structural Control and Health Monitoring*, vol. 18, no. 8, pp. 807–824, 2011.

[7] ——, "Bayesian methodology for diagnosis uncertainty quantification and health monitoring," *Structural Control and Health Monitoring*, vol. In Press, 2011.

[8] B. Sun, S. Zeng, R. Kang, and M. G. Pecht, "Benefits and challenges of system prognostics," *Reliability, IEEE Transactions on*, vol. 61, no. 2, pp. 323–335, 2012.

[9] P. Vaidya and M. Rausand, "Remaining useful life, technical health, and life extension," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 225, no. 2, pp. 219–231, 2011.

[10] M. Wei, M. Chen, and D. Zhou, "Multi-sensor information based remaining useful life prediction with anticipated performance," *Reliability, IEEE Transactions on*, vol. 62, no. 1, pp. 183–198, 2013.

[11] N. Chen and K. L. Tsui, "Condition monitoring and remaining useful life prediction using degradation signals: revisited," *IIE Transactions*, vol. 45, no. 9, pp. 939–952, 2013.

[12] X.-S. Si, W. Wang, C.-H. Hu, D.-H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," *Reliability, IEEE Transactions on*, vol. 61, no. 1, pp. 50–67, 2012.

[13] K. Le Son, M. Fouladirad, A. Barros, E. Levrat, and B. Iung, "Remaining useful life estimation based on stochastic deterioration models: A comparative study," *Reliability Engineering & System Safety*, 2012.

[14] M. Daigle, A. Saxena, and K. Goebel, "An efficient deterministic approach to model-based prediction uncertainty estimation," in *Annual Conference of the Prognostics and Health Management Society*, sep 2012, pp. 326–335.

[15] C. Hu, B. D. Youn, P. Wang, and J. Taek Yoon, "Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life," *Reliability Engineering & System Safety*, vol. 103, pp. 120–135, 2012.

[16] P. Baraldi, F. Mangili, and E. Zio, "Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data," *Reliability Engineering & System Safety*, 2012.

[17] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation–a review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011.

[18] K. Medjaher, D. A. Tobon-Mejia, and N. Zerhouni, "Remaining useful life estimation of critical components with application to bearings," *Reliability, IEEE Transactions on*, vol. 61, no. 2, pp. 292–302, 2012.

[19] F. Zhao, Z. Tian, and Y. Zeng, "Uncertainty quantification in gear remaining useful life prediction through an integrated prognostics method," *Reliability, IEEE Transactions on*, vol. 62, no. 1, pp. 146–159, 2013.

[20] S. Sankararaman, M. Daigle, A. Saxena, and K. Goebel, "Analytical algorithms to quantify the uncertainty in remaining useful life prediction," in *Aerospace Conference, 2013 IEEE*. IEEE, 2013, pp. 1–11.

[21] L. Tang, G. J. Kacprzynski, K. Goebel, and G. Vachtsevanos, "Methodologies for uncertainty management in prognostics," in *Aerospace conference, 2009 IEEE*. IEEE, 2009, pp. 1–12.

[22] C. R. Farrar, N. A. Lieven, and M. T. Bement, "An introduction to damage prognosis," *Damage Prognosis: For Aerospace, Civil and Mechanical Systems*, pp. 1–12, 2005.

[23] D. J. Inman, C. R. Farrar, V. L. Junior, and V. S. Junior, *Damage prognosis for aerospace, civil and mechanical systems*. Wiley, 2005.

[24] C. Farrar and N. Lieven, "Damage prognosis: the future of structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 623–632, 2007.

[25] S. Sankararaman, Y. Ling, C. Shantz, and S. Mahadevan, "Uncertainty quantification in fatigue damage prognosis," in *Annual Conference of the Prognostics and Health Management Society*, 2009.

[26] ——, "Uncertainty quantification in fatigue crack growth prognosis," *International Journal of Prognostics and Health Management*, vol. 2, no. 1, 2011.

[27] J. Gu, D. Barker, and M. Pecht, "Uncertainty assessment of prognostics of electronics subject to random vibration," in *AAAI fall symposium on artificial intelligence for prognostics*, 2007, pp. 50–57.

[28] J. A. DeCastro, "Exact nonlinear filtering and prediction in process model-based prognostics," in *Annual Conference of the Prognostics and Health Management Society*, San Diego, CA., 2009.

[29] M. Orchard, G. Kacprzynski, K. Goebel, B. Saha, and G. Vachtsevanos, "Advances in uncertainty representation and management for particle filtering applied to prognostics," in *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, oct. 2008, pp. 1 –6.

[30] B. Saha, K. Goebel, S. Poll, and J. Christophersen, "Prognostics methods for battery health monitoring using a bayesian framework," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 2, pp. 291 –296, feb. 2009.

[31] B. Saha and K. Goebel, "Uncertainty management for diagnostics and prognostics of batteries using bayesian techniques," in *Aerospace Conference, 2008 IEEE*, March 2008, pp. 1 –8.

[32] A. Usynin and J. W. Hines, "Uncertainty management in shock models applied to prognostic problems," in *Artificial Intelligence For Prognostics: Papers From The AAAI Fall Symposium*, no. FS-07-02, 2007.

[33] H. Liao and Z. Tian, "A framework for predicting the remaining useful life of a single unit under time-varying operating conditions," *IIE Transactions*, vol. 45, no. 9, pp. 964–980, 2013.
[34] X. Guan, Y. Liu, R. Jha, A. Saxena, J. Celaya, and K. Geobel, "Comparison of two probabilistic fatigue damage assessment approaches using prognostic performance metrics," *International Journal of Prognostics and Health Management*, vol. 1, p. 005, 2011.
[35] R. Rackwitz and B. Flessler, "Structural reliability under combined random load sequences," *Computers & Structures*, vol. 9, no. 5, pp. 489–494, 1978.
[36] T. W. Lee and B. M. Kwak, "A reliability-based optimal design using advanced first order second moment method," *Journal of Structural Mechanics*, vol. 15, no. 4, pp. 523–542, 1987.
[37] A. Der Kiureghian, H.-Z. Lin, and S.-J. Hwang, "Second-order reliability approximations," *Journal of Engineering Mechanics*, vol. 113, no. 8, pp. 1208–1225, 1987.
[38] Y.-G. Zhao and T. Ono, "A general procedure for first/second-order reliability method (form/sorm)," *Structural Safety*, vol. 21, no. 2, pp. 95–112, 1999.
[39] A. Haldar and S. Mahadevan, *Probability, reliability, and statistical methods in engineering design*. John Wiley & Sons, Inc., 2000.
[40] G. Cederbaum, I. Elishakoff, and L. Librescu, "Reliability of laminated plates via the first-order second-moment method," *Composite structures*, vol. 15, no. 2, pp. 161–167, 1990.
[41] M. Hohenbichler and R. Rackwitz, "First-order concepts in system reliability," *Structural safety*, vol. 1, no. 3, pp. 177–188, 1983.
[42] A. Der Kiureghian, Y. Zhang, and C. Li, "Inverse reliability problem," *Journal of engineering mechanics*, vol. 120, p. 1154, 1994.
[43] M. Daigle, B. Saha, and K. Goebel, "A comparison of filter-based approaches for model-based prognostics," in *Aerospace Conference, 2012 IEEE*. IEEE, 2012, pp. 1–10.
[44] M. Daigle and K. Goebel, "Model-based prognostics with concurrent damage progression processes," *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, vol. PP, doi=10.1109/TSMCA.2012.2207109, pp. 1–12, 2013.
[45] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, mar 2004.
[46] K. Dolinski, "First-order second-moment approximation in reliability of structural systems: critical review and alternative approach," *Structural Safety*, vol. 1, no. 3, pp. 211–231, 1983.
[47] M. Hohenbichler, S. Gollwitzer, W. Kruse, and R. Rackwitz, "New light on first-and second-order reliability methods," *Structural Safety*, vol. 4, no. 4, pp. 267–284, 1987.
[48] N. Sitar, J. D. Cawlfield, and A. Der Kiureghian, "First-order reliability approach to stochastic analysis of subsurface flow and contaminant transport," *Water Resources Research*, vol. 23, no. 5, pp. 794–804, 1987.
[49] B. Fiessler, R. Rackwitz, and H. Neumann, "Quadratic limit states in structural reliability," *Journal of the Engineering Mechanics Division*, vol. 105, no. 4, pp. 661–676, 1979.
[50] J. E. Angus, "The probability integral transform and related results," *SIAM review*, vol. 36, no. 4, pp. 652–654, 1994.
[51] X. Chen and N. Lind, "Fast probability integration by three-parameter normal tail approximation," *Structural Safety*, vol. 1, no. 4, pp. 269–276, 1983.
[52] S. Sankararaman and S. Mahadevan, "Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data," *Reliability Engineering & System Safety*, vol. 96, no. 7, pp. 814 – 824, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832011000111
[53] B. Saha, C. C. Quach, and K. Goebel, "Optimizing battery life for electric uavs using a bayesian framework," in *Aerospace Conference, 2012 IEEE*. IEEE, 2012.
[54] M. Chen and G. A. Rincon-Mora, "Accurate electrical battery model capable of predicting runtime and I-V performance," *IEEE Transactions on Energy Conversion*, vol. 21, no. 2, pp. 504 – 511, jun 2006.

BIOGRAPHIES

Shankar Sankararaman received his B.S. degree in Civil Engineering from the Indian Institute of Technology, Madras in India in 2007; and later, obtained his Ph.D. in Civil Engineering from Vanderbilt University, Nashville, Tennessee, U.S.A. in 2012. His research focuses on the various aspects of uncertainty quantification, integration, and management in different types of aerospace, mechanical, and civil engineering systems. His research interests include probabilistic methods, risk and reliability analysis, Bayesian networks, system health monitoring, diagnosis and prognosis, decision-making under uncertainty, treatment of epistemic uncertainty, and multidisciplinary analysis. He is a member of the Non-Deterministic Approaches (NDA) technical committee at the American Institute of Aeronautics (AIAA), the Probabilistic Methods Technical Committee (PMC) at the American Society of Civil Engineers (ASCE), the Institute of Electrical and Electronics Engineers (IEEE), and the Prognostics and Health Management (PHM) Society. Currently, Shankar is a researcher at NASA Ames Research Center, Moffett Field, CA, where he develops algorithms for uncertainty assessment and management in the context of system health monitoring, prognostics, and decision-making.

Matthew Daigle received the B.S. degree in Computer Science, and Computer and Systems Engineering from Rensselaer Polytechnic Institute, Troy, NY, in 2004; and the M.S., and Ph.D. degrees in Computer Science from Vanderbilt University, Nashville, TN, in 2006, and 2008, respectively. From September 2004 to May 2008, he was a Graduate Research Assistant with the Institute for Software Integrated Systems and Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN. During the summers of 2006 and 2007, he was an intern with Mission Critical Technologies, Inc., at NASA Ames Research Center. From June 2008 to December 2011, he was an Associate Scientist with the University of California, Santa Cruz, at NASA Ames Research Center. Since January 2012, he has been with NASA Ames Research Center as a Research Computer Scientist. His current research interests include physics-based modeling, model-based diagnosis and prognosis, simulation, and hybrid systems. Dr. Daigle is a member of the Prognostics and Health Management Society and the IEEE.

Kai Goebel is the Deputy Area Lead for Discovery and Systems Health at NASA Ames where he also directs the Prognostics Center of Excellence. After receiving the Ph.D. from the University of California at Berkeley in 1996, Dr. Goebel worked at General Electric's Corporate Research Center in Niskayuna, NY from 1997 to 2006 as a senior research scientist before joining NASA. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion; and his interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He holds 17 patents, and has published more than 250 papers in the area of systems health management.