Remaininng Useful Life Prediction through Failure Probability Computation for Condition-based Prognostics

Shankar Sankararaman

1 SGT Inc., NASA Ames Research Center, Moffett Field, CA 94035, USA
shankar.sankararaman@nasa.gov

ABSTRACT

The key goal in prognostics is to predict the remaining useful life (RUL) of engineering systems in order to guide different types of decision-making activities such as path planning, fault mitigation, etc. The remaining useful life of an engineering component/system is defined as the first future time-instant in which a set of safety threshold conditions are violated. The violation of these conditions may render the system inoperable or even lead to catastrophic failure. This paper develops a computational methodology to analyze the aforementioned set of safety threshold conditions, calculate the probability of failure, and in turn, proposes a new hypothesis to mathematically connect such probability to the remaining useful life prediction. A significant advantage of the proposed methodology is that it is possible to learn important properties of the remaining useful life, without simulating the system until the occurrence of failure; this feature renders the proposed approach unique in comparison with existing direct-RUL-prediction approaches. The methodology also provides a systematic way of treating the different sources of uncertainty that may arise from imprecisely known future operating conditions, inaccurate state-of-health state estimates, use of imperfect models, etc. The proposed approach is developed using a model-based framework prognostics using principles of probability, and illustrated using a numerical example.

1. MOTIVATION

1.1. Introduction

The prediction of remaining useful life (RUL) is, perhaps, the central component of a prognostics and health management (PHM) system (Vaidya & Rausand, 2011). In order to predict the RUL, it is necessary to thoroughly understand the functioning of the engineering system under consideration, estimate the state-of-health, analyze possible failure modes, predict damage growth using degradation models, and identify the future time-instant at which it is not possible to continue operating the system (Engel, Gilmartin, Bongort, & Hess, 2000). The aforementioned time-instant is referred to as the end-of-life (EOL), and it is possible to check whether EOL has been reached by evaluating a binary threshold constraint (referred to as the EOL-threshold function). Typically, safety constraints and serviceability constraints are used to formulate the EOL-threshold function.

1.2. Uncertainty in Prognostics

An important aspect of prognostics is that future prediction intrinsically needs to account for the various sources of uncertainty that affect the future behavior of the system (Orchard, Kacprzynski, Goebel, Saha, & Vachtsevanos, 2008). As a result, the EOL and RUL predictions become uncertain (Sankararaman & Goebel, 2013); in fact, at any future time-instant, there is a probability that the EOL-threshold is violated/satisfied.

To begin with, it is practically impossible to estimate the state of health because (1) it is rarely possible to measure health directly, and it may be necessary to infer health from system output measurements; and (2) such measurements are obtained from sensors that may not be accurate due to noise, gain, bias, etc. Techniques such as Kalman filtering (Swanson, 2001) and particle filtering (Zio & Peloni, 2011) are used to estimate the state-of-health. Starting from an arbitrary state of health, it is necessary to predict damage growth using a degradation model. Damage growth is a function of usage/operating conditions, loading conditions, etc., all of which may be uncertain, and therefore, render damage growth uncertain. The degradation model used for damage growth prediction may also uncertain, and model uncertainty is represented through uncertain model parameters and model form errors (usually, approximated using process noise). It is important to systematically account for these sources of uncertainty in prognostics, and estimate the overall uncertainty in the RUL prediction.
1.3. Previous Work

Previous work in this context has focused on posing RUL prediction as an uncertainty propagation problem (Sankararaman & Goebel, 2013).

The RUL is expressed as a “black-box function” of all other uncertain quantities; for every realization of these uncertain quantities, the future behavior of the system is simulated until EOL is reached. The aforementioned “black-box function” is a combination of (1) damage degradation model (Luo, Pattipati, Qiao, & Chigusa, 2008), usually expressed as a state-space model (Sun, Zuo, Wang, & Pecht, 2012); and (2) the EOL-threshold. Then, the uncertainty in the RUL prediction is computed by propagating the different sources of uncertainty through the so-called black-box function. Such propagation can be accomplished through a variety of sampling-based (Daigle, Saxena, & Goebel, 2012; Sankararaman, 2015) and analytical methods (Sankararaman, Daigle, & Goebel, 2014). These methods have been applied to a variety of applications such as pumps (Daigle & Goebel, 2013), valves (Daigle & Goebel, 2011), batteries (Chen & Rincon-Mora, 2006), structural crack growth damage prognosis (Farrar & Lieven, 2007), capacitors (Kulkarni, Celaya, Goebel, & Biswas, 2013) etc.

1.4. Proposed Approach

In general, the computation of the black-box function may be computationally intensive since it requires simulation until EOL is reached. The present paper explores an alternative method, where RUL can be predicted without simulating the system until EOL. This can be achieved by evaluating the likelihood of system failure, and the relationship between such likelihood and the RUL can be mathematically proved. The likelihood of system failure can be calculated analytically, based on methods developed by researchers in the field of “model-based reliability analysis”. It is important not to confuse this terminology with reliability-based life-prediction or testing-based life-prediction (Saxena, Sankararaman, & Goebel, 2014) that focus on fleet-wide prognostics; the present paper deals only with condition-based prognostics by focusing on the operation of one particular unit at the component-level or system-level, as the case maybe (Sankararaman, 2015). The proposed method is applicable for all scenarios where the state-of-health of the system is monotonically decreasing (there may be some special scenarios such as crack closure where the state-of-health improves, and such cases are not considered in this paper).

1.5. Organization of the Paper

The rest of this paper is organized as follows. Section 2 presents the model-based framework for prognostics and mathematically defines the remaining useful life of an engineering system. Section 3 develops the new computational approach that connects the prediction of remaining useful life with failure probability calculation. Section 4 illustrates the proposed methodology using a numerical example, and finally, Section 5 concludes the paper.

2. Remaining Useful Life in Prognostics

This section describes a framework for model-based prognostics, and mathematically defines the remaining useful life, which in turn is based on the definition of end-of-life threshold function.

Suppose that it is desired to perform prognostics and predict the RUL at a generic time-instant $t_P$. Consider the architecture shown in Fig. 1, where the whole problem of prognostics can be considered to consist of the following three sub-problems:

1. Present state estimation
2. Future state prediction
3. RUL computation

2.1. 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 that is used to continuously predict the state of the system, as:

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

where $x(t) \in \mathbb{R}^{n_x}$ is the state vector, $\theta(t) \in \mathbb{R}^{n_\theta}$ is the parameter vector, $u(t) \in \mathbb{R}^{n_u}$ is the input vector, $v(t) \in \mathbb{R}^{n_v}$ is the process noise vector, and $f$ is the state equation.

As stated earlier, the state of the system uniquely defines the amount of damage in the system.

The state vector at time $t_P$, i.e., $x(t)$ (and the parameters $\theta(t)$, if they are unknown) is (are) estimated using output data collected until $t_P$. Let $y(t) \in \mathbb{R}^{n_y}$, $u(t) \in \mathbb{R}^{n_u}$, and $h$ denote the output vector, measurement noise vector, and output.
equation respectively. Then,

\[ y(t) = h(t, x(t), \theta(t), u(t), n(t)) \quad (2) \]

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

### 2.2. Future State Prediction

Having estimated the state at time \( t_p \), the next step is to predict the future states of the component/system. Note that, since the focus is predicting future, no data is available, and it is necessary to completely rely and use Eq. 1 for this purpose. This differential equation can be discretized and used to predict the states at any future time instant \( t > t_p \), as a function of the states at time \( t_p \).

### 2.3. 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, \( C_{\text{EOL}} = \{ c_i \}_{i=1}^{n_c} \), 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, \( (x(t), \theta(t), u(t)) \), to the Boolean domain \( \mathbb{B} \triangleq [0, 1] \), where \( c_i(x(t), \theta(t), u(t)) = 1 \) if the constraint is satisfied, and 0 otherwise (Daigle & Goebel, 2013).

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

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

For the sake of simplicity, the above equation can be simply written as:

\[ T_{\text{EOL}}(t) = T_{\text{EOL}}(x(t), \theta(t), u(t)) \quad (4) \]

The above notation is valid and simpler to use because all the arguments of \( T_{\text{EOL}} \) in Eq. 3 can be calculated as a function of \( t \), and hence, \( T_{\text{EOL}}(t) \) can be easily represented as a function of only \( t \).

\( T_{\text{EOL}} \) is equal to 1 when EOL-threshold constraint is 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 \( T_{\text{EOL}} \) becomes equal to one. Therefore,

\[ E(t_p) \triangleq \inf \{ t \in \mathbb{R} : t \geq t_p \land T_{\text{EOL}}(t) = 1 \}. \quad (5) \]

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

\[ R(t_p) \triangleq E(t_p) - t_p. \quad (6) \]

Note that the output equation (Eq. 2) or output data \( y(t) \) is not used in the prediction stage, and EOL and RUL are dependent only on the state estimates at time \( t_p \); though these state estimates are obtained using the output data, the output data is not used for EOL/RUL calculation after state estimation.

For the purpose of implementation, \( f \) in Eq. 1 is transformed into the corresponding discrete-time version. Discrete time is indexed by \( k \), and there is a one-to-one relation between \( t \) and \( k \) depending on the discretization level. While the time at which prediction needs to be performed is denoted at \( t_p \), the corresponding index is denoted by \( k_p \). Similar let \( k_e \) denote the time index that corresponds to the end of life. Thus, it is clear that RUL predicted at time \( t_p \), i.e., \( R(t_p) \) depends on

1. Present state estimate \( (x(k_p)) \); using the present state estimate and the state space equations in Eq. 1, the future states \( (x(k_p), x(k_p+1), x(k_p+2), \ldots, x(k_e)) \) can be calculated.
2. Future loading \( (u(k_p), u(k_p+1), u(k_p+2), \ldots, u(k_e)) \); 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 \( k_e \) (denoted by \( \theta(k_p), \theta(k_p+1), \ldots, \theta(k_e) \)).
4. Process noise \( (v(k_p), v(k_p+1), v(k_p+2), \ldots, v(k_e)) \).

For the purpose of RUL prediction, all of the above quantities are independent quantities and hence, RUL becomes a dependent quantity. Let \( X = \{ X_1, X_2, \ldots X_i, \ldots X_n \} \) denote the vector of all the above dependent quantities, where \( n \) is the length of the vector \( X \), 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:

\[ R = G(X) \quad (7) \]

The above functional relation in Eq. 7 can be graphically explained, as shown in Fig. 2. Knowing the values of \( X \), it is possible to compute the corresponding value of \( R \), using Fig. 2 that is equivalently represented by Eq. 7. The quantities contained in \( X \) are uncertain, and the focus in prognostics to compute their combined effect on the RUL prediction, and thereby compute the probability distribution of \( R \). The problem of estimating the uncertainty in \( R \) is equivalent to propagating the uncertainty in \( X \) through \( G \), and researchers have investigated different types of methods for this purpose. The most commonly used methodology for this purpose is Monte Carlo sampling, using which multiple realizations of \( R \) can be obtained from multiple realizations of uncertain quantities. This approach is referred to as the direct-RUL-prediction approach in this paper, and this is not pursued. An alternative statistical methodology is developed and the direct-RUL-prediction approach will be used as a benchmark to compare results from the newly proposed methodology.
3. Failure Probability and RUL

The proposed methodology is based on evaluating the likelihood of satisfying the EOL-threshold constraint, and relating this likelihood to the RUL prediction.

Consider the time of prediction $t_p$, all time-instants $t > t_p$. As per Eq. 3, the system is said to be safe and operable so long as “$T_{EOL}(t) = 0$”, and the first future time-instant $t_E$ at which $T_{EOL}(t_E)$ becomes equal to one is said to be equal to the EOL. For the sake of illustration and terminology description, assume that failure corresponds to $T_{EOL} = 1$. Consider the following probability that calculates the likelihood of failure:

$$P_f(t|t_p) = P(T_{EOL}(t) = 1|t_p)$$  \hspace{1cm} (8)

Note that the above probability is calculated at the time of prediction $t_p$, but for a generic future time-instant $t$.

If it can be assumed the amount of damage/fault is non-decreasing, then it can be easily visualized that the function $P_f(t|t_p)$ is non-decreasing with respect to time $t$. While this does seem to be a reasonable assumption, it is not universally true. For example, in structural damage prognosis, crack closure is a widely studied phenomena, and can lead to an “improvement in the health state” (since crack closure can result in an increased stiffness). The rest of this paper only focuses on scenarios in which the amount of damage/fault is non-decreasing, and therefore, $P_f(t|t_p)$ is non-decreasing with respect to time $t$.

The hypothesis proposed in this paper, is that $P_f(t|t_p)$ is exactly equal to the “probability that the end of life is less than or equal to time $t$”. In mathematical terms:

$$P_f(t|t_p) = P(T_E \leq t|t_p)$$  \hspace{1cm} (9)

Note that the right hand side of the above equation is exactly equal to the cumulative distribution function of the End-of-Life.

Recall from Section 2 that the End-of-Life is an uncertain quantity and needs to be expressed using a probability distribution. If $T_E$ denotes the random variable, and $t_E$ an instance of this variables, then the probability density function (PDF) and the cumulative distribution function (CDF) of this variable are denoted by $f_{T_E}(t_E)$ and $F_{T_E}(t_E)$ respectively. Therefore, Eq. 9 can be extended as:

$$P_f(t|t_p) = P(T_E \leq t|t_p) = F_{T_E}(t_E)$$  \hspace{1cm} (10)

In order to prove the above hypothesis, consider “$N$” different, random system paths starting from the time of prediction $t_p$. Consider a generic future time-instant $t$, by which “$m$” paths have already reached the end-of-life. Therefore, by mere definition of the CDF, it can be written that:

$$P(T_E \leq t|t_p) = F_{T_E}(t_E) = \frac{m}{N}$$  \hspace{1cm} (11)

At the particular time $t$, there are now a total of “$N$” states out of which “$m$” states “fall in” the zone of failure. Therefore, it also follows that:

$$P_f(t|t_p) = \frac{m}{N}$$  \hspace{1cm} (12)

Therefore, by comparing Eq. 11, and Eq. 12, the proposed hypothesis is therefore proved.
This hypothesis provides a fundamentally different way of calculating EOL and therefore, the RUL. Once EOL is obtained, RUL can be easily calculated using Eq. 6. The major advantage of the proposed methodology is that, in order to calculate $P_f(t|t_p)$, it is not necessary to simulate the system until failure; nevertheless, this probability, through Eq. 10, can provide the cumulative distribution function of the EOL. The CDF value of EOL is critical in assigning credible-intervals for the EOL, which are useful for decision-making, and hence, it is believed that the proposed methodology will be of immense value in this regard.

The computation of failure probability has been discussed by several researchers, particular in the field of model-based reliability analysis. The most simplest method (simple to build and code, but expensive to implement) is Monte Carlo sampling (Robert & Casella, 2004). There are several advanced sampling methods such as importance sampling (Glynn & Iglehart, 1989), adaptive sampling (Bucher, 1988), stratified sampling (Caflisch, 1998), etc., which can improve upon the efficiency of basic Monte Carlo sampling. Alternatively, there are also analytical methods developed by structural reliability engineers; these include the first-order reliability methods (Haldar & Mahadevan, 2000), second-order reliability method (Der Kiureghian, Lin, & Hwang, 1987), etc. The focus of the present paper is not on testing the applicability of this methods, but on developing and proving the hypothesis in Eq. 9, as applicable to condition-based prognostics. As this hypothesis has been mathemtically proved in this section, it is illustrated using a numerical example in the following section.

4. Numerical Example

In order to illustrate the proposed methodology, consider the problem of crack growth prognosis in a simple plate. This plate is subjected to cyclic, uniform uniaxial stress ($S$), and Paris’ law is used for crack growth analysis. Paris law calculates the increment in crack size per cycle of loading, in terms of crack growth parameters ($C$ and $n$), threshold stress intensity factor ($\Delta K_{th}$), and load stress intensity factor ($\Delta K$):

$$\frac{da}{dN} = C(\Delta K)^n(1 - \frac{\Delta K}{\Delta K_{th}})^p$$  \hspace{1cm} (13)

The stress intensity factor ($\Delta K$), for the sake of illustration, is assumed to be available in closed form, as:

$$\Delta K = S\sqrt{\pi a}$$  \hspace{1cm} (14)

For the sake of this numerical example, “the 7075-T6” aluminum alloy is considered. The quantities $S \sim LN(100, 40)$, $C \sim LN(6.54 \times 10^{-13}, 4.0 \times 10^{-13})$, and $\Delta K_{th} \sim LN(5.66 \times 10^6, 0.268 \times 10^6)$ are chosen to be lognormal random variables. The quantities in parentheses above indicate the mean and standard deviation of the random variables (all numerical values are in SI units). The exponents “$n$” and “$p$” are assumed to 3.89 and 0.75 (no unit) respectively.

An important challenge in crack growth analysis, is that the initial crack size is not known. A rigorous approach to fatigue life would account for crack growth starting from the actual initial flaw, accounting for imperfections, voids and non-metallic inclusions. This procedure is cumbersome because small crack growth propagation is anomalous in nature and hence not completely understood. On the other hand, there are several crack growth models (Paris law (Pugno, Ciavarella, Cornetti, & Carpinteri, 2006), AFGRW (Harter, 1999), etc.) in the long crack regime which are used to study long crack growth behavior. Equivalent initial flaw size is a fictitious quantity which was introduced to bypass small crack growth calculations and make direct use of a long crack growth law in order to make fatigue life prediction; the EIFS must be chosen in such a way that when the long crack growth law is used with EIFS as the initial value, it yields crack growth results that match with observed crack growth data (Liu & Mahadevan, 2009). Since EIFS is fictitious and hence, not measurable, it needs to be estimated based on observed data on crack size. There have been several studies on how to estimate the equivalent initial flaw size (EIFS), and the value reported by Liu and Mahadevan (Liu & Mahadevan, 2009) is used in the analysis below; EIFS is assumed to follow a lognormal distribution with mean and standard deviation equal to 0.23mm and 0.05mm respectively. If average values are assumed for all the quantities (i.e., without any uncertainty in them), then the crack growth behavior can be obtained as shown in Fig. 3.

The failure threshold limit is set to be the time-instant when the crack size exceeds 0.75mm. It can be easily seen that, after this crack size, the rate of increase in crack size is phenomenally high.

Including the different sources of uncertainty, the probability density function of RUL is first computed using the direct-RUL-prediction approach. Monte Carlo sampling is used for the purpose of illustration, and the resulting probability density function is shown in Fig. 4. For this particular analysis,
1000 Monte Carlo samples were used, and the resultant RUL samples were converted into a probability density function using principles of kernel density estimation. Note that Fig. 4 is calculated using “G” in Eq. 2, whereas the proposed approach only calculates $P(T_{EOL}(t) = 1)$, which is equivalent to $P(a > 0.75)$, where “a” represents the crack size in this numerical example. As per the proposed approach, $P_f(t|t_0)$ is plotted as a function of time $t$ (where the time of prediction is denoted by $t_0$, the initial time), as shown in Fig. 5. Here, Monte Carlo analysis is used for the calculation of failure probability (note that Monte Carlo analysis was previously used for direct-RUL-prediction, and required simulation until failure), only for the purpose of illustration; other advanced failure probability computation methods will be investigated in future.

In order to demonstrate the proposed methodology, and illustrate the comparison in Eq. 9, it is necessary to compare (1) the cumulative distribution function corresponding to the PDF in Fig. 4; and (2) the “failure probability versus time” plot in Fig. 5. This comparison is shown in Fig. 6. Note that, in this example, the time of prediction is considered as “$t_p = 0$”, and therefore the EOL and RUL are identical.

It can be seen from Fig. 6 that the two plots compare very well, thereby illustrating the hypothesis proposed in Eq. 9. While the direct-RUL-prediction approach requires simulating every sample until failure to obtain the prediction of EOL, the proposed approach based on failure probability computation does not require this. In fact, the proposed approach requires a much easier computation in comparison with the direct-RUL-prediction approach, and also provides a systematic way of handling uncertainty, analogous to the direct-RUL-prediction approach.

5. Conclusion

This paper presented a new computational methodology for remaining useful life estimation in prognostics. While RUL has been traditionally calculated by forecasting the state-of-health degradation, the proposed approach calculates the likelihood of system failure and mathematically connects such likelihood to the remaining useful life. The major advantage of the proposed approach is that it is possible to learn about the properties of RUL even without simulating until failure or the end-of-life. The proposed approach was developed in the context of model-based, condition-based prognostics, using fundamentals of probabilistic analysis. This method is applicable to scenarios where the degradation is monotonically decreasing (there may be some scenarios such as crack closure where the state of health improves). Finally, the method was illustrated using a simple numerical example, consisting of component-level remaining useful life prediction.

It is further necessary to investigate the applicability of the proposed approach to complicated EOL-threshold functions and extend this approach to larger, practical engineering systems and applications.

Acknowledgment

The work reported herein was in part funded by the NASA System-wide Safety 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


**Biography**

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 and Astronautics (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.