Uncertainty Representation and Interpretation in Model-based Prognostics Algorithms based on Kalman Filter Estimation

José R. Celaya[1], Abhinav Saxena[2], and Kai Goebel[3]

[1,2] SGT Inc. NASA Ames Research Center, Moffett Field, CA, 94035, USA
jose.r.celaya@nasa.gov
abhinav.saxena@nasa.gov

[3] NASA Ames Research Center, Moffett Field, CA, 94035, USA
kai.goebel@nasa.gov

ABSTRACT

This article discusses several aspects of uncertainty representation and management for model-based prognostics methodologies based on our experience with Kalman Filters when applied to prognostics for electronics components. In particular, it explores the implications of modeling remaining useful life prediction as a stochastic process and how it relates to uncertainty representation, management, and the role of prognostics in decision-making. A distinction between the interpretations of estimated remaining useful life probability density function and the true remaining useful life probability density function is explained and a cautionary argument is provided against mixing interpretations for the two while considering prognostics in making critical decisions.

1. INTRODUCTION

Model-based prognostics methodologies in electronics prognostics have been developed based on Bayesian tracking methods such as Kalman Filter, Extended Kalman Filter, and Particle Filter. The models used in these methodologies are mathematical abstractions of the time evolution of the degradation process and the cornerstone for the estimation of remaining useful life. The Bayesian tracking framework allows for estimation of state of health parameters in prognostics making use of available measurements from the system under consideration. In this framework, health parameters are regarded as random variables for which, in the case of Kalman and Extended Kalman filters, their distribution are regarded as Normal and the estimation process focuses on computing estimates of the expected value and variance as they relate to the mean and variance that fully parametrize the Normal distribution. In addition to the health estimation process, fore-

casting of the health parameters is required up to a future time that results in crossing of the pre-established failure condition threshold. This is ultimately required in order to compute remaining useful life.

Previous work applied to electrolytic capacitor and power MOSFETs (Metal-Oxide Semiconductor Field-Effect Transistor) has focused on implementation of the previously described process and has presented remaining useful life results without any uncertainty measure associated to them (J. R. Celaya et al., 2011; J. Celaya, Saxena, Kulkarni, et al., 2012; J. Celaya et al., 2011; J. Celaya, Kulkarni, et al., 2012). Other work on prognostics based on particle filtering has been presented regarding remaining useful life as a random variable and presenting corresponding uncertainty estimates (Saha et al., 2009; Daigle & Goebel, 2011). This work focuses on reviewing uncertainty representation techniques used in model-based prognostics and on providing an interpretation of uncertainty for the electronics prognostics applications previously presented, and based on Kalman filter approaches for health state estimation.

The Bayesian tracking framework allows for modeling of sources of uncertainty in the measurement process and also on the degradation evolution dynamic model as applied on the application under consideration. This is done in terms of an additive noise in the model, which is regarded as zero mean and normally distributed random variable. This allows for the aggregation of different sources of uncertainty for the health state tracking step. Its implications on the uncertainty estimation for remaining useful life (RUL) including future state forecasting are discussed in this paper.

1.1. Model-based prognostics background

As mentioned earlier, a model-based prognostics methodology based on Bayesian tracking consists of two steps, health state estimation and RUL prediction. The following is a high

level description of the process that will help to provide the appropriate context for the upcoming discussion.

State of health estimation: To initiate the prediction, it is necessary to first establish a starting point, the current state of health. A model-based algorithm employs dynamic models of the physical behavior of the system or component under consideration, along with dynamic degradation models of key parameters that represent the degradation over time. Bayesian tracking algorithms like Kalman filter, extended Kalman filter, and particle filter are among the algorithms typically employed in a model-based prognostics methodology (Daigle & Goebel, 2011; Saha & Goebel, 2009; J. Celaya et al., 2011; J. Celaya, Saxena, Kulkarni, et al., 2012). In such methodologies, dynamic models of the nominal system and degradation models are posed as a discrete state-space system in which the state variable $x(t)$ consists of physical variables, and in some cases, it includes degradation model parameters to be estimated online.

The models consist of a state equation representing the time evolution of the state as shown in Eq. (1a); where $u(t)$ is the system input and $w(t)$ is a zero-mean and normally distributed additive noise representing random model error. In addition, the measurement equation (Eq. (1b)) relates the state variable to measurements of the systems $y(t)$. The term $v(t)$ is a zero-mean and normally distributed additive noise representing the random measurement error. The measurement and model noise normality assumption could be relaxed when using computational Bayesian methods like particle filtering.

$$\dot{x}(t) = f(x(t), u(t)) + w(t) \qquad (1a)$$
$$y(t) = h(x(t)), u(t)) + v(t) \qquad (1b)$$

The state of the system, as it evolves through time, is periodically estimated by the filter as measurements $y(t)$ of key variables become available through the life of the system. This is the health state estimation step of the model-based prognostics algorithm. Typically, in a model-based prognostics method, a Bayesian tracking algorithm attempts to estimate the expected value of the joint probability density function of the state $x(t_p)$. Where t_p is the time at which a remaining useful life prediction is computed using only system observations up to this point in time. Different assumptions about the probability density function are used depending on the filter used.

Remaining useful life estimation (prediction): In order to compute remaining useful life, the state-equation (Eq. (1a)) of the model is used to compute the state evolution in a forecasting mode until an end-of-life threshold is reached at time denoted by t_{EOL}. The last state estimate at time t_p in the

health state estimation step is typically used as initial state value for forecasting $x(t)$ up to t_{EOL}. Remaining useful life $R(t_p)$ at time of prediction t_p is defined as

$$R(t_p) = t_{EOL} - t_p, \qquad (2)$$

where t_p is deterministic and known, and t_{EOL} is a random variable function of the failure threshold and the state estimate $x(t_p)$. This function includes the state forecasting step and the identification of when the failure threshold is crossed.

1.2. Ideas explored in this paper

In this paper we explore how the state vector variable should be interpreted during the tracking phase and how it is related to the process of final RUL prediction. This probability interpretation is often overlooked in the literature by interpreting the state vector as the health indicator and a threshold is used on this variable in order to compute EOL (end-of-life) and RUL.

Here, we discuss how the state estimation process is defined in the Bayesian framework. We will, in particular, focus on the output of the estimation process in the Kalman filter framework. Furthermore, we try to interpret the objective of the Kalman filter, whether to estimate $x(t)$ as a random variable or to estimate a parameter of the probability density function of $x(t)$ –such as expected value or variance– or both.

In addition, we will challenge how we usually think about RUL and how it has been interpreted using other, similar, methods. The main objective here is to characterize its impact on uncertainty representation and management. For instance, if RUL is considered as a random variable and we assume that a model-based prognostics framework based on the Kalman filter generates RUL with a particular variance, then it is incorrect to arbitrarily expect, assume, or force the variance to be small. The variance of random variables such as RUL is not under our control as explained in the next section.

These concepts are discussed in the context of prognostics of electronics, particularly, the uncertainty propagation in power MOSFET and capacitor prognostics applications as presented in J. R. Celaya et al. (2011); J. Celaya, Saxena, Kulkarni, et al. (2012) and J. Celaya et al. (2011); J. Celaya, Kulkarni, et al. (2012) respectively. In these applications, uncertainty has not been explicitly considered in the prediction results and this paper is an effort towards augmenting the methods used there with an uncertainty management methodology.

1.3. Background on Uncertainty Management

There are several different types of sources of uncertainty that must be accounted for in a prognostic system formulation. These sources may be categorized into following four categories and accordingly require separate representation and management methods.

1. **Aleatoric or Statistical Uncertainties:** these uncertainties arise from inherent variability in any process and cannot be eliminated. They can be characterized by multiple experimental runs but cannot be reduced by improved methods or measurements. Sampling fluctuations from the characterized probability density function of a source of aleatoric uncertainty can result in different predictions every time. Examples of such uncertainties include manufacturing variations, material properties, etc.

2. **Epistemic or Systematic Uncertainties:** these uncertainties arise due to unknown details that cannot be identified and hence are not incorporated into a process. With improved methods and deeper investigations these uncertainties may be reduced but are rarely eliminated. Modeling uncertainties fall under this category and include modeling errors due to unmodeled phenomena in both system model and the fault propagation model.

3. **Prejudicial Uncertainties:** these uncertainties arise due to the way a process is set up and is expected to change if the process is redesigned. Conceptually these can be considered a type of epistemic uncertainty, except it is possible to control these to a better extent. Examples for these uncertainties include sensor noise, sensor coverage, information loss due to data processing, various approximations and simplifications, numerical errors, etc.

While it is possible to reduce some of these uncertainties, it is not possible or practically beneficial to eliminate them altogether. However, representing them and accounting for them in prognostic outputs is extremely important. Uncertainties in a prognostic estimate directly affect the associated decision making process, which is typically expressed through the concept of risk due to unwanted outcomes. Several PHM approaches quantify risk based on uncertainty quantification in an algorithm's output and incorporate it into a corresponding cost-benefit equation through monetary concepts (Bedford & Cooke, 2001).

1.3.1. Uncertainty management in prognostics

In the context of prognostics and health management uncertainties are talked about from quantification, representation, and management points of view (deNeufville, R., 2004; Hastings & McManus, 2004; Ng & Abramson, 1990; Orchard et al., 2008; Tang et al., 2009). While all three are different processes they are often confused with each other and interchangeably used.

Uncertainty quantification: Deals with characterizing a source of uncertainty so it can be incorporated into models and simulations as correctly as possible. A characterization or quantification step may involve carefully designed experimentation with actual systems observed in realistic and relevant environments. An accurate quantification of uncertainties is considered very challenging as also acknowledged

in Engel (2009). Quantification of uncertainty from various sources in a process has been investigated and a sensitivity analysis conducted to identify which input uncertainty contributes most to the output uncertainty in prognostics for fatigue crack damage (Sankararaman et al., 2011). This allows prioritizing and subsequently focusing on more critical uncertainties instead of all.

Uncertainty representation: Next step is the representation of uncertainty, which is, often times, guided by the choice of modeling and simulation frameworks. There are several methods for uncertainty representation that vary in the level of granularity and detail. Some common theories include classical set theory, probability theory, fuzzy set theory, fuzzy measure (plausibility and belief) theory, and rough set (upper and lower approximations) theory. However, in the PHM domain the representation of uncertainties is dominated by probabilistic measures (DeCastro, 2009; Orchard et al., 2008; Saha et al., 2009), which offer a mathematically rigorous approach but assume availability of a statistically sufficient database. Other approaches, such as possibility theory (Fuzzy logic) and Dempster-Shafer theory, can be employed when only scarce or incomplete data are available (Wang, 2011). Furthermore, the choice of type of probability density function affects the quality of prognostic outputs. Several approaches in the literature resorted to assuming Normal probability density functions, however this choice should be guided by the results of the uncertainty characterization and quantification step.

Uncertainty management: The most loosely used term in the PHM literature in the context of uncertainty is that of uncertainty management. Uncertainty management includes two main functions, to incorporate all relevant and/or significant sources of uncertainty into prognostic models and simulations. Therefore, the problem formulation stage itself lays a foundation for an effective uncertainty management. Once all relevant sources of uncertainty are identified and included, the uncertainty propagation is the next component towards effective management. Various measures of uncertainty must be combined in an appropriate manner in the prognostic model as the input variability filters through a complex (possibly non-linear) system model.

If, in a perfect situation, all sources of uncertainties are identified, modeled, and managed correctly, the output probability density function for random variables like RUL or End-of-life (EOL) would match the true spread and would not change from one experiment to another. This is, however, in practice impossible to achieve because no model is perfect and not all sources of uncertainties can be characterized. Furthermore, an exhaustive sampling-based method such as a Monte Carlo simulation would be computationally, prohibitively expensive. This has inspired the development of intelligent sampling based algorithms (DeCastro, 2009; Orchard et al., 2008;

Saha et al., 2009) and mathematical transformations, such as Support vectors (Saha & Goebel, 2008) and Principle Component analysis (Usynin & Hines, 2007), that result in minor approximations but capture most details of the true variability. It may not be possible to identify and accurately characterize all sources of uncertainty and hence use of a sensitivity analysis is recommended to isolate the most important factors (Gu et al., 2007; Sankararaman et al., 2011; Tang et al., 2009). Through effective uncertainty management practices one can at most strive towards bringing the predicted estimate close to the true spread and not arbitrarily reducing the spread of RUL itself. What can be minimized, is the variability in the estimate of a given parameter of interest, not the variability in the parameter of interest itself.

2. REMAINING USEFUL LIFE STOCHASTIC MODELING

Remaining useful life in a prognostics context is defined differently than in a reliability context. In prognostics, it is implied that remaining useful life at time t_p is a condition-based estimation of the usage time left until failure, using measurements of key variables and past usage information up to time t_p. This process typically consists of forecasting the future state of health beyond t_p and identifying when the state of health will cross a failure threshold representative of a functional failure. In addition, RUL in prognostics considers – implicitly or explicitly– future usage conditions. This is not the case in the reliability context. Given the many sources of uncertainty evident from a quick assessment of all that is involved in computing RUL for a system, it is common to consider RUL as a non-deterministic quantity. Furthermore, RUL is also a time evolving process, meaning that RUL at time t_p will be different than RUL for $t \neq t_p$. This can be well illustrated with the use of the alpha-lambda prognostics metric (Saxena et al., 2010) as seen in various publications on prognostics (J. R. Celaya et al., 2011; J. Celaya et al., 2011).

2.1. Remaining useful life as a stochastic process

A random process or stochastic process is defined as a collection of random variables. Following the definition presented in Gross and Harris (1998), a stochastic process is a "mathematic abstraction of an empirical process whose development is governed by probabilistic laws". Furthermore, it is defined as a family of random variables $\{X(t), t \in T\}$ where T is the time range and $X(t)$ is the state of the process at time t. The time range could be discrete or continuous.

A stochastic process is also used in the signal-processing context to represent non-deterministic (stochastic) signals (Oppenheim & Schafer, 1989). From Kalman (1960) we get the following explanation as it relates to filtering: "Intuitively, a random process is simply a set of random variables which are indexed in such a way as to bring the notion of time into the picture".

In several applications, RUL prediction is a process in which periodic computations of RUL are generated through the life of the system under consideration. In our previous work on power MOSFET prognostics (J. R. Celaya et al., 2011), periodic measurements (up to every minute) are available. RUL is computed periodically and can be considered as a random process $R(t)$. In contrast, in our previous work on electrolytic capacitor prognostics (J. Celaya et al., 2011), measurements are not available at regular time intervals. RUL computations are made multiple times whenever a measurement is available. In this case, $R(t)$ can also be considered as a random process but the set T will contain only the times at which RUL was computed.

2.2. Implications on uncertainty management

The definition of RUL as a random variable or random process has many implications on uncertainty management and in the representation of uncertainty in a particular model-based prognostics methodology. If RUL is not modeled within a probability framework, like a fuzzy variable or just a deterministic variable, uncertainty management activities will differ. To illustrate, let us consider a simple point estimate example from basic mathematical statistics (Bain & Engelhardt, 1992).

A parameter estimation example: Let us assume that we can perform a set of run to failure experiments with high level of control, ensuring same usage and operating conditions. In addition, remaining useful life at time t_p is computed by measuring the elapsed time from t_p until failure for all the n samples $(R_1, \ldots, R_n)$ on the set of run to failure experiments. Assuming that these random samples come from a probability density function $f_R(r)$, with expected value $E(R) = \mu$ and variance $Var(R) = \sigma^2$.

Let θ_1 be a parameter estimator of the mean μ of f_R, with expected value $E(\theta_1) = \mu_{\theta_1}$ and variance $V(\theta_1) = \sigma_{\theta_1}^2$. This estimator will be a function of all the sample values and will have a probability density function f_{θ_1}. θ_1 is a point estimate of the random variable R such as the sample mean, the median or some other location statistic. Now, from the uncertainty management perspective in prognostics, it is necessary to judge the ability of the algorithm to properly compute the point estimate of the process, in this case, to properly estimate μ. So it is expected that this estimate θ_1 has the least variability, the least variance possible, therefore making θ_1 less uncertain. As a result, $\sigma_{\theta_1}^2$ should be as small as possible. It is, on the other hand, incorrect to expect the estimation process to reduce σ^2 itself.

This is often misinterpreted for prognostics methodologies base on computational statistics that do not directly focus on a point estimate but on generating an approximation of the distribution of R. Since the variability can be assessed by a measure of spread like the sample standard deviation com-

puted directly from the sample distribution of R, again, this variation should not be arbitrarily decreased by tuning of the algorithm since it is intended to represent the real statistical uncertainty of the process.

The previous discussion applies to RUL predictions without loss of generality as long as they are modeled as random variables, which is typically the case. The concept can be further described considering the sample average $\bar{R}$ as the estimator ($\theta_1 = \bar{R}$). From basic probability theory (Bain & Engelhardt, 1992), one can observe that $\mu_{\theta_1} = \mu$ and the ${\sigma_{\theta_1}}^2 = \sigma^2/n$. This estimator is unbiased, and its variance ${\sigma_{\theta_1}}^2$ can be reduced by increasing sample size. But σ^2 cannot be reduced because it is the inherent variability in the random variable R.

2.3. Implications on how RUL is computed by statistical models

Let us now consider the complete RUL computation algorithm including state estimation and prediction steps, i.e., the prognostics algorithm is a black box estimation of RUL. This statistical model can have different focus in providing estimations of $R(t)$. The following situations (although not exhaustive) are considered here:

1. $R(t)$ could be assumed to be a known random variable with a known probability density or mass function (parametric case). Therefore, the statistical model will focus on providing the best possible estimator of the parameters or key quantities function of the random variable as the expected value and the variance. For instance, if $R(t)$ is presumed Normal, then the statistical model will provide an estimate of the mean and the standard deviation since they fully parametrize the Normal random variable.

2. A computational statistics model could be used to avoid making assumptions about the distribution of $R(t)$ therefore focusing on computing an approximation of the probability density/mass function of $R(t)$. This will be a choice for the cases in which there is no knowledge about the distribution or the non-parametric case is preferred. It will also be the case for when there is no analytical solution tractable for the statistical model structure therefore the use of a computational model, based on Monte Carlo simulation approaches, is needed.

The uncertainty management focus will differ under the two situations described above. In case one, where distribution parameters are estimated, the uncertainty management should focus on properly estimating the spread parameter θ_s of $R(t)$. A spread parameter θ_s could be variance or some other estimator focused on representing the variability of the distribution. This estimator should properly aggregate all the previously identified sources of uncertainty, like measurement, model, future input and environment uncertainty. From the uncertainty management perspective, one should not expect θ_s to be small. Instead, one should expect it to be an accu-

rate representation of the real uncertainty in the real RUL of the system. A similar situation arises in the second case. In this case an approximation of the distribution of $R(t)$ is computed. Its shape and therefore the spread or variability represented by this approximation, should be the real uncertainty of the RUL in the system and should not be made arbitrarily small either by tuning the statistical method to do so or by any other arbitrary transformation to make this approximation more crisp around the location parameter.

2.4. Implications on decision-making

Being able to capture the uncertainty correctly is of paramount importance in prognostics. This might not always be the case for other applications involving parameter estimation. For instance, in a control application, the frequency of the compensation loop is generally high enough to be able to dampen the effects of uncertainty in the parameter estimation process. For prognostics, this will typically not be the case. If the prognostics situation under consideration is used for contingency management, in which safety of operation is at stake; properly estimating the uncertainty of the true RUL is necessary. If the uncertainty estimation is incorrect, then this can lead to risky decision-making, leading to reduced safety and possibly increasing the change of catastrophic failure. A similar argument can be made if prognostics is used in a logistics settings such as condition-based maintenance in manufacturing systems or in military operations.

The previous argument can also be made from the opposite end by considering the implications of the decision-making method on how RUL is computed and how uncertainty management is performed. For the last few years, research in prognostics and health management (PHM) has mainly focussed on the prognostics element, which deals with methods to predict RUL. There have been several methodologies published and many more under development for a variety of man-made systems. As a result of the previous effort, prognostics methodologies have been developed in a sort of unbounded or unguided way with respect to how the actual method is going to be used in the decision-making process. This meas that input from the types of decisions that will use the prognostics information and from the overall optimization of system performance have so far not been considered.

The type of decision-making application may dictate the prognostics methodology as well as the types of estimates to be generated (recall cases in Section 2.2.3). Consequently, this will also have an impact on requirements generation. For instance a fleet based optimization of aircraft maintenance operations considers very different decisions as compared to an unmanned aerial vehicle (UAV) mission reconfiguration based on prognostics indication on power train failures. Following the same argument, it is clear that different decision-making methodologies will have different capabili-

ties in terms of handling the prognostics information. For instance, an optimization of a particular decision process might not be able to work with random variables, therefore a point estimate would be provided. This will be different if the optimization itself is able to deal with RUL as a random variable, in this case, the computation distribution function of $R(t)$ or the estimators of the parameters that fully parametrize it would be provided. If the decision-making process, can further use information about how reliable the prognostics information is, then information about a measure of quality of the estimators, which is different than just bias, would be provided.

3. Uncertainty Interpretation and Computation in Model-based Prognostics with Kalman Filter Estimation

Model-based prognostics methodologies for electronics components like electrolytic capacitors (J. Celaya et al., 2011; J. Celaya, Kulkarni, et al., 2012) and power MOSFETs (J. R. Celaya et al., 2011; J. Celaya, Saxena, Kulkarni, et al., 2012) have been previously introduced. The methodologies make use of empirical degradation models and a single precursor to failure parameter to compute RUL. These methodologies rely on accelerated aging experiments to identify degradation behavior and to create time dependent degradation models. The process followed in these methodologies is presented in the block diagram in Figure 1.
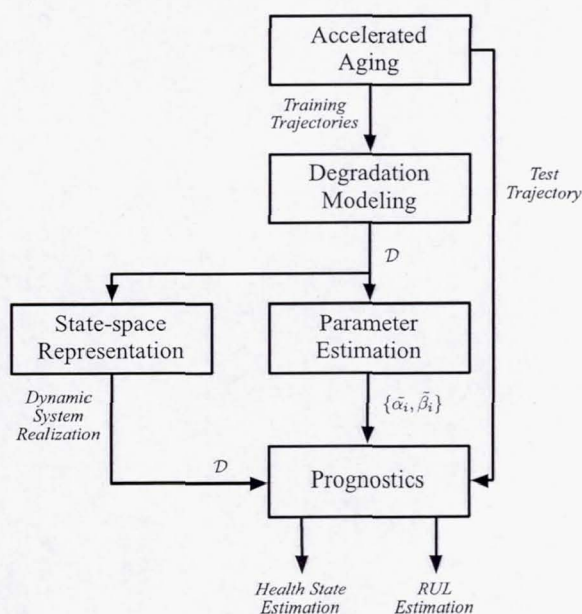

Figure 1. Methodology for electronics component prognostics development.

Accelerated aging tests provided measurements throughout the aging process, including measurements at pristine con-

dition and measurements after failure condition. Empirical degradation models that are based on the observed degradation process during the accelerated aging tests are developed. The objective of the models is to generate a parametrized model of the time-dependent degradation process for these components. The time dependent degradation model is transformed into a discrete-time linear dynamic system in order to be used in a Bayesian tracking setting. The Kalman filter algorithm is used to track the state of health and the degradation model is used to make predictions of remaining useful life once no further measurements are available.

3.1. Prognostics methodology

The methodology consists of the three main steps described below and it is depicted in Figure 2. This is the explanation of what it is considered inside the prognostics block in Figure 1.

This methodology follows from the general concepts of model-based prognostics described in Section 1.1.1. In the electronics component case, the system dynamics consists only of the degradation process dynamics since the prognostics focuses at the component level only.

1. State tracking (Kalman Filter): The state variable in the degradation model $\mathcal{D}$ is a precursor of failure parameter represented by Eq. (3a). When the degradation model uses static parameters (parameters not estimated online by the filter), then the state variable is a scalar quantity and the state evolution equation is scalar. The degradation model is expressed as a discrete time dynamic model in order to estimate the state as new measurements become available. The simplified Kalman filter model set up is given as

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}, \qquad (3a)$$
$$y_k = Hx_k + v_k. \qquad (3b)$$

The output of this step is the optimal state estimate $\hat{x}_p$.

2. Health state forecasting: It is necessary to forecast the state variable once there are no more measurements available at the time of RUL prediction t_p. This is done by evaluating the degradation model (Eq. (3a)) through time using the state estimate $\hat{x}_p$ from the previous step as the initial state value for forecasting.

3. Remaining life computation: RUL is computed as the time between time of prediction t_p and the time at which the forecasted state crosses the failure threshold value F.

This process is repeated for different values of t_p through the life of the component under consideration.

3.2. Kalman Filter Background

The Kalman filter framework is based on Bayesian parameter estimation. A Bayes estimator allows to estimate parameters
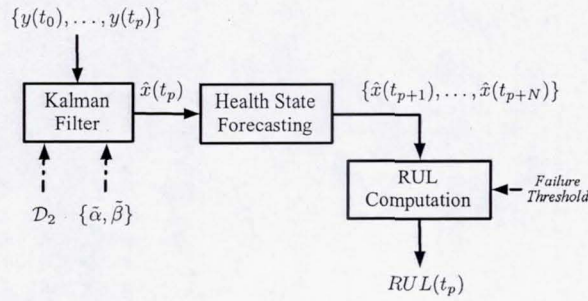
$\{y(t_0), \ldots, y(t_p)\}$

Figure 2. Model-based prognostics methodology

based on prior knowledge about the parameter distribution. In the tracking problem, system measurements serve as a form of prior knowledge, therefore the objective is to estimate the state $x(t)$ conditional to all the previous measurements of the system. The Bayes estimation framework is based on the concepts of risk and loss functions in which the risk is defined as the expected loss (Bain & Engelhardt, 1992). This background information is relevant since it helps to understand the statistical origins of the Kalman filter framework which is the focal point of the discussions in this paper. Based on the seminal paper for the Kalman filter (Kalman, 1960), the optimal state estimate is given as $x^*(t) = E[x(t)|y(t_o), \cdots, y(t)]$. This is the solution that minimizes the risk (expected loss), for a loss function based on the estimation error. Furthermore, the random process for the state and for the process noise are Normal. Additional details on the problem formulation and assumptions are presented in Kalman (1960).

Implications on Kalman filter for prognostics: Considering a scalar implementation of the Kalman filter over discrete-time model as in Eqs. (3). The output of the filter referred to as the optimal state estimate x_k^* is basically given by the conditional state estimate $\hat{x}_k = E[x_k|y_k]$ and the state conditional probability density function is given by,

$$p(x_k|y_k) \sim N(\hat{x}_k, P_k), \qquad (4)$$

where P_k is the filter's estimate of the error variance.

The output of the filter is the estimate of the expected value $\hat{x}_k$, and the estimation error covariance P_k. The state random variable x_p is normally distributed with mean $\hat{x}_k$ and variance P_k.

3.3. Uncertainty propagation in prognostics

Based on the previous discussion regarding the interpretation of the Kalman filter output in terms of probabilities, it can be observed that the health state estimation output is a Normal random variable with known parameters considering the sources of uncertainties derived from modeling error and measurement error.

Uncertainty in the health state estimation step: We assume

here a scalar case for state estimation, like in the case of the capacitor prognostics method where the health indicator is a scalar state variable (J. Celaya et al., 2011). Time index p is defined as the time of RUL prediction t_p, which is also the time of the last available measurement in the state estimation step. The state estimate x_p is a normally distributed random variable with mean $\hat{x}_p$ and variance P_k.

$$x_p \sim N(\hat{x}_p, P_k) \qquad (5)$$

This variable includes the propagation of measurement uncertainty and also model error uncertainty as included in the Kalman filter implementation.

Uncertainty in the health state forecasting step: Forecasting is needed for the state variable to be able to estimate its value at a future time until it crosses a pre-established failure threshold F. The forecasting process is carried out using the state equation (Eq. (3a)) recursively, using the last health state estimate $\hat{x}_p$ as initial value. Let $\tilde{x}_p(l)$ be the l_{th} step ahead forecast starting from x_p. From the uncertainty propagation point of view and focusing on a one step ahead forecasting using Eq. (3a), the forecast value is given by

$$\tilde{x}_p(1) = Ax_p + Bu_p + w_p. \qquad (6)$$

Variables x_p and w_p are Normal and independent with known mean and variance. Following from basic probability theory, the forecast $\tilde{x}_p(1)$ is also Normal. In general, the l_{th} step ahead forecast $\tilde{x}_p(l)$ will have a Normal distribution as well. It should also be noted that $\tilde{x}_p(l)$ is a function of the last state estimate ($\tilde{x}_p(l) = f(x_p)$). Considering the forecast variables as random variables and given the analytical properties of the Normal distribution, the probability density function $f_{\tilde{x}_p(l)}$ can be derived analytically and is given by,

$$\tilde{x}_p(l) \sim N(\mu_l, \sigma_l^2); \qquad (7)$$

where the mean is given by

$$\mu_l = A^l \hat{x}_p + Bu_p + \sum_{i=0}^{l-1} A^i, \qquad (8)$$

and the variance is given by

$$\sigma_l^2 = (A^2)^l P_k + \sum_{i=1}^{l-1} (A^2)^i \sigma_w^2 + \sigma_w^2. \qquad (9)$$

Uncertainty in RUL: Computing the uncertainty in the RUL is more complicated from an analytical point of view. Defining $R(t_p)$ as the remaining useful life,

$$R(t_p) = t_{EOL} - t_p. \qquad (10)$$

The time at end-of-life (t_{EOL}) is a continuous variable which

is computed from the forecast $\tilde{x}_p(l)$.

Let $\tilde{x}_p(j)$ be the first forecast value to cross the failure threshold F. An interpolation between $\tilde{x}_p(j)$ and $\tilde{x}_p(j-1)$ is used to compute t_{EOL}. Considering that the forecasts are functions of x_p, RUL is also a function of x_p.

$$R(t_p) = g(x_p). \tag{11}$$

From the random variable uncertainty propagation point of view, $R(t_p)$ is a function g of a normally distributed random variable, therefore, it is also a random variable. It is nevertheless difficult to derive its probability density function analytically. There is also no information that suggests that $R(t_p)$ will be Normal. The probability density function of $R(t_p)$ can be approximated using computational statistics methods. This can be done by taking N samples from x_p and computing $R(t_p)$ for each sample. An histogram can be built from the N computed $R(t_p)$ values and a density estimation method could be used to generate the approximation of the probability density function.

3.4. Discussion

From the analytical results presented for the first two steps of the prognostics process (Section 3.3.3), it can be observed that the variance will be larger after the forecasting process. In addition, there is no evidence to suggest that $R(t_p)$ will be Normal and further investigation is needed to explore its dependance on the forecasting process, like number of steps ahead forecasts and step length. It is also clear, that simply defining the variance of $R(t_p)$ as P_k or σ_l^2 is not an accurate representation of the uncertainty in the process.

The model-based methodology for electronics prognostics based on the Kalman filter is able to capture additive degradation model error uncertainty and additive measurement uncertainty. In order for the approximation of the probability density function of $R(t_p)$ to be a true representation of the system uncertainty, the variances of the measurement noise and modeling noise should be properly estimated. If considered as tuning parameters, then the generated uncertainty in $R(t_p)$ will not be representative of the real process.

4. CONCLUSION

This article presented a discussion on uncertainty representation and management for model-based prognostics methodologies based on the Bayesian tracking framework and specifically for a Kalman filter application to electronics components. In particular, it explores the implication of modeling remaining useful life prediction as a stochastic process and how it relates to remaining useful life computation by statistical models, to uncertainty representation and management, and to the role of prognostics in decision-making. A discussion on how uncertainty propagates from the health state

estimation process through the health state forecasting process is provided. Remaining useful life computation steps under uncertainty are presented and analytical results on uncertainty quantification are provided under a simplified scenario. A proper propagation of uncertainty through the RUL prediction step as well as its correct interpretation are key to developing decision-making methodologies that make use of the remaining useful life prediction estimates and their corresponding uncertainties in order to make actionable choices that will optimize reliability, operations or safety in view of the prognostics information.

NOMENCLATURE

R	Remaining useful life random variable
t_p	Time of remaining useful life prediction
$R(t_p)$	Remaining useful life prediction at time t_p
$\hat{x}_k$	Optimal state estimator from Kalman filter
$\tilde{x}_k(l)$	l_{th} step ahead forecast from x_k
t_{EOL}	Time at end-of-life
$x(t)$	Scalar continuous state variable for filter model
$\mathbf{x}(t)$	Vector continuous state variable for filter model
x_k	Scalar discrete-time state variable for filter model
F	Failure threshold
$N(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2

REFERENCES

Bain, L. J., & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics*. Duxbury.

Bedford, T., & Cooke, R. M. (2001). *Probabilistic risk analysis: Foundations and methods*. Cambridge University Press.

Celaya, J., Kulkarni, C., Biswas, G., & Goebel, K. (2011, September 25-29). A model-based prognostics methodology for electrolytic capacitors based on electrical overstress accelerated aging. In *Proceedings of annual conference of the phm society*. Montreal, Canada.

Celaya, J., Kulkarni, C., Saha, S., Biswas, G., & Goebel, K. (2012, January). Accelerated aging in electrolytic ca-

pacitors for prognostics. In *2012 proceedings - annual reliability and maintainability symposium (rams)* (p. 1 -6). doi: 10.1109/RAMS.2012.6175486

Celaya, J., Saxena, A., & Goebel, K. (2012, June 19-21). A discussion on uncertainty representation and interpretation in model-based prognostics algorithms based on kalman filter estimation applied to prognostics of electronics components. In *Infotech@aerospace 2012 online conference proceedings.* Garden Grove, California.

Celaya, J., Saxena, A., Kulkarni, C., Saha, S., & Goebel, K. (2012, January). Prognostics approach for power mosfet under thermal-stress aging. In *2012 proceedings - annual reliability and maintainability symposium (rams)* (p. 1 -6). doi: 10.1109/RAMS.2012.6175487

Celaya, J. R., Saxena, A., Saha, S., & Goebel, K. (2011, September 25-29). Prognostics of power mosfets under thermal stress accelerated aging using data-driven and model-based methodologies. In *Proceedings of annual conference of the phm society.* Montreal, Canada.

Daigle, M. J., & Goebel, K. (2011). A model-based prognostics approach applied to pneumatic valves. *International Journal of Prognostics and Health Management, 2 (2)*(008).

DeCastro, J. A. (2009). Exact nonlinear filtering and prediction in process model-based prognostics. In *Annual conference of the prognostics and health management society.* San Diego, CA..

deNeufville, R. (2004). Uncertainty management for engineering systems planning and design. In *Engineering systems symposium mit.* Cambridge, MA..

Engel, S. J. (2009). PHM engineering perspectives, challenges and 'crossing the valley of death'. In *Annual conference of the prognostics and health managament society.* San Diego, CA..

Gross, D., & Harris, C. M. (1998). *Fundamentals of queueing theory* (3rd ed.). John Wiley & Sons Inc.

Gu, J., Barker, D., & Pecht, M. (2007). Uncertainty assessment of prognostics of electronics subject to random vibration. In *Artificial intelligence for prognostics: Papers from the aaai fall symposium.*

Hastings, D., & McManus, H. (2004). A framework for understanding uncertainty and its mitigation and exploitation in complex systems. In *Engineering systems symposium mit* (p. 19). Cambridge MA..

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME– Journal of Basic Engineering, 82 (Series D)*, 35-45.

Ng, K.-C., & Abramson, B. (1990). Uncertainty management in expert systems. *IEEE Expert Systems, 20.*

Oppenheim, A. V., & Schafer, R. W. (1989). *Discrete-time signal processing.* Prentice Hall.

Orchard, M., Kacprzynski, G., Goebel, K., Saha, B., & Vachtsevanos, G. (2008, oct.). Advances in uncertainty representation and management for particle filtering applied to prognostics. In *Prognostics and health management, 2008. phm 2008. international conference on* (p. 1 -6). doi: 10.1109/PHM.2008.4711433

Saha, B., & Goebel, K. (2008, march). Uncertainty management for diagnostics and prognostics of batteries using bayesian techniques. In *Aerospace conference, 2008 ieee* (p. 1 -8). doi: 10.1109/AERO.2008.4526631

Saha, B., & Goebel, K. (2009). Modeling li-ion battery capacity depletion in a particle filtering framework. In *Proceedings of annual conference of the phm society.*

Saha, B., Goebel, K., Poll, S., & Christophersen, J. (2009, feb.). Prognostics methods for battery health monitoring using a bayesian framework. *IEEE Transactions on Instrumentation and Measurement, 58*(2), 291 -296. doi: 10.1109/TIM.2008.2005965

Sankararaman, S., Ling, Y., Shantz, C., & Mahadevan, S. (2011). Uncertainty quantification in fatigue crack growth prognosis. *International Journal of Prognostics and Health Management, 2-1*(1).

Saxena, A., Celaya, J. R., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and Health Management, 1 (1)*(001).

Tang, L., Kacprzynski, G., Goebel, K., & Vachtsevanos, G. (2009, march). Methodologies for uncertainty management in prognostics. In *Aerospace conference, 2009 ieee* (p. 1 -12). doi: 10.1109/AERO.2009.4839668

Usynin, A., & Hines, J. W. (2007). Uncertainty management in shock models applied to prognostic problems. In *Artificial intelligence for prognostics: Papers from the aaai fall symposium.*

Wang, H.-f. (2011, January). Decision of prognostics and health management under uncertainty. *International Journal of Computer Applications, 13*(4), 1–5. (Published by Foundation of Computer Science)

BIOGRAPHIES

José R. Celaya is a research scientist with SGT Inc. at the Prognostics Center of Excellence, NASA Ames Research Center. He received a Ph.D. degree in Decision Sciences and Engineering Systems in 2008, a M. E. degree in Operations Research and Statistics in 2008, a M. S. degree in Electrical Engineering in 2003, all from Rensselaer Polytechnic Institute, Troy New York; and a B. S. in Cybernetics Engineering in 2001 from CETYS University, Meéxico.

Abhinav Saxena is a Research Scientist with Stinger Ghaffarian Technologies at the Prognostics Center of Excellence NASA Ames Research Center, Moffet Field CA. His research focus lies in developing and evaluating prognostic algorithms for engineering systems using soft computing techniques. He is a PhD in Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta. He earned his B. Tech

in 2001 from Indian Institute of Technology (IIT) Delhi, and Masters Degree in 2003 from Georgia Tech. Abhinav has been a GM manufacturing scholar and is also a member of IEEE, AAAI and ASME.

Kai Goebel received the degree of Diplom-Ingenieur from the Technische University Munchen, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively. Dr. Goebel is a senior scientist at NASA Ames Research Center where he is the deputy area lead for Discovery and Systems Health in the Intelligent Systems division. In addition, he directs the Prognostics Center of Excellence and he is the Associate Principal Investigator for Prognostics of NASAs Integrated Vehicle Health Management Program. He worked at General Electrics Corporate Research Center in Niskayuna, NY from 1997 to 2006 as a senior research scientist. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He holds fifteen patents and has published more than 200 papers in the area of systems health management.