Entropy-based probabilistic fatigue damage prognosis and algorithmic performance comparison

Xuefei Guan[1], Yongming Liu[2], Abhinav Saxena[3], Jose Celaya[3] and Kai Goebel[4]

[1] *Department of Mechanical & Aeronautical Engineering, Clarkson University, Potsdam, NY 13699-5725, USA*
guanx@clarkson.edu

[2] *Departments of Civil & Environmental Engineering, Clarkson University, Potsdam, NY 13699-5710, USA*
yliu@clarkson.edu

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

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

ABSTRACT

In this paper, a maximum entropy-based general framework for probabilistic fatigue damage prognosis is investigated. The proposed methodology is based on an underlying physics-based crack growth model. Various uncertainties from measurements, modeling, and parameter estimations are considered to describe the stochastic process of fatigue damage accumulation. A probabilistic prognosis updating procedure based on the maximum relative entropy concept is proposed to incorporate measurement data. Markov Chain Monte Carlo (MCMC) technique is used to provide the posterior samples for model updating in the maximum entropy approach. Experimental data are used to demonstrate the operation of the proposed probabilistic prognosis methodology. A set of prognostics-based metrics are employed to quantitatively evaluate the prognosis performance and compare the proposed method with the classical Bayesian updating algorithm. In particular, model accuracy, precision and convergence are rigorously evaluated in[*] addition to the qualitative visual comparison. It is shown that the proposed maximum relative entropy methodology has narrower confidence bounds of the remaining life prediction than classical Bayesian updating algorithm.

1 INTRODUCTION

Fatigue damage is a critical issue in many structural and non-structural systems, such as aircrafts, critical civil structures, and electronic components. The estimation of the reliability and remaining useful life (RUL) is important in condition-based maintenance of a system so that unit replacements can be taken in time prior to catastrophic failures. Several physics-based models have been proposed in order to describe the fatigue process and predict the damage propagation; among those, Paris-type crack growth laws are most commonly used (Bourdin *et al.*, 2008). However, due to the stochastic nature of fatigue crack growth, a deterministic model is not capable of quantifying the crack growth subject to various uncertainties associated with the fatigue damage. Uncertainties arising from a number of sources, such as measurement errors, model prediction residuals, and non-optimal parameter estimation, affect the quality of life predictions.

Probabilistic updating methods based on Bayes theorem have been used to evaluate the probability density functions (PDF) of input parameters using response measurements (Madsen, 1997; Zhang and Mahadevan, 2000; Perrin *et al.*, 2007). Methodologies based on maximum entropy, such as Maximum Entropy principle (MaxEnt) and Maximum relative

Entropy (MRE), are alternative approaches for probabilistic updating and have been used in many applications such as statistical mechanics (Caticha and Preuss, 2004; Tseng and Caticha, 2008). The objective of this paper is to develop and apply a prognosis approach based on MRE for probabilistic fatigue damage prognosis. One of the advantages of the proposed MRE approach is that the resulting confidence bounds are narrower compared to the classical Bayesian method, which is beneficial for decision making in a health management setting.

2 MRE UPDATING

2.1 MRE formulation for model updating

The relative information entropy, also referred to as Kullback-Leibler divergence (Kullback and Leibler, 1951), of two PDFs $f_1(\theta)$ and $f_2(\theta)$ is defined as,

$$I(f_1 : f_2) = -\int_\Theta d\theta \cdot f_1(\theta)\log\left(f_1(\theta)/f_2(\theta)\right) \tag{1}$$

where θ is the parameter vector and Θ is the parameter vector space. The axioms of maximum entropy (Skilling, 1988) indicate that the form of Eq. (1) is the unique entropy representation for inductive inference. For a generic inverse problem, the posterior of parameter vector θ is inferred based on the prior information about θ (the prior PDF of θ), the observations of data x (the response measurements), and the known relationship between x and θ (the physics-based models). Let $\mu(x,\theta)$ be a prior joint distribution and $p(x,\theta)$ be a posterior joint distribution. According to the entropy axioms, the selected joint posterior is the one that maximizes the relative entropy $I(p:\mu)$, subject to constraints, such as statistical moments and measures of response variable.

$$I(p:\mu) = -\int dxd\theta \cdot p(x,\theta)\log\left(p(x,\theta)/\mu(x,\theta)\right) \tag{2}$$

In Eq. (2), $\mu(x,\theta) = \mu(\theta)\mu(x \mid \theta)$ contains all prior information. $\mu(x \mid \theta)$ is the likelihood function and $\mu(\theta)$ is the prior PDF. The same relationship applies to the joint posterior $p(x,\theta)$. When new information is available in the form of a constraint, the updating procedure will search in the space of $\mathrm{X}\times\Theta$ for a posterior which maximizes $I(p:\mu)$. Measurements of the response variable x can be used to perform the updating, which is performed in a similar way as the classical Bayesian updating. The benefit of MRE updating is that it can incorporate other information for inference, which cannot be included in the classical Bayesian updating. For example, the expected value of a function of θ or the empirical judgment on the mean value of θ can be used in MRE updating (Giffin and Caticha, 2007). This flexibility of applicable information can pose more constraints on a posterior thus yield a more accurate result given that those constraints are justified. Following the derivation of MRE posterior (Caticha and Giffin, 2006), if a new observation x' is obtained, the posteriors that reflect the fact x is now known to be x' is a constraint such that

$$c_1 : p(x) = \int d\theta \cdot p(x,\theta) = \delta(x - x') \tag{3}$$

Other information in the form of moment constraints, such as the expected value of some function $g(\theta)$, can be formulated as

$$c_2 : \int dxd\theta \cdot p(x,\theta)g(\theta) = \langle g(\theta)\rangle = G \tag{4}$$

The normalization constraint is

$$c_3 = \int dxd\theta \cdot p(x,\theta) = 1 \tag{5}$$

Maximizing Eq. (2), subject to constraints Eqs. (3-5), the posterior can be obtained as (Caticha and Giffin, 2006).

$$p(\theta) \propto \mu(\theta)\mu(x' \mid \theta)e^{\beta\cdot g(\theta)} \tag{6}$$

The coefficient β is determined by,

$$\frac{\partial \ln Z(x',\beta)}{\partial\beta} = G \tag{7}$$

where $Z(x',\beta) = \int d\theta \cdot \mu(\theta)\mu(x' \mid \theta)e^{\beta\cdot g(\theta)}$ is the normalization constant. The right side of Eq. (6) consists of three terms. $\mu(\theta)$ is the parameter prior, $\mu(x' \mid \theta)$ is the likelihood, and $e^{\beta\cdot g(\theta)}$ is the exponential term introduced by moment constraints. Eq. (6) is similar to Bayesian posterior except for the additional exponential term. This equation further indicates that, if no moment constraint is available, i.e., β is zero, MRE updating will be identical to Bayesian updating. In other words, Bayesian updating is a special case of MRE updating. Similar to that of a Bayesian updating problem, the likelihood function is usually constructed using the physics-based model depending on different realistic applications.

2.2 Likelihood function and mechanism model

Considering a general model prediction equation, let d be the observed value of a response variable and y be the prediction value of a model M. If the model is sufficiently accurate to describe the system, the observed value is equal to model prediction value, i.e. $y = d$. However, noises and errors usually exist for both model predictions and measurements. Combining the model prediction error term e and measurement error term ε, the observed value can be expressed as,

$$d = M(x \mid \theta) + e + \varepsilon \tag{8}$$

where $M(x \mid \theta)$ is the model prediction value given parameter vector θ. In general, the probability density functions of the two uncorrelated error terms e and ε can be described using two independent Gaussian distributions with standard deviations of σ_e and σ_ε, respectively. Replacing the two error terms with a total error term $\tau = (e+\varepsilon) \sim Gaussian(0, \sigma_\tau)$, the likelihood function of multiple observation data can be constructed as

$$L(D_N \mid \theta) = \frac{1}{\left(\sqrt{2\pi}\sigma_\tau\right)^n} \exp\left\{ -\sum_{i=1}^{n} \frac{\left[d_i - M(x \mid \theta)\right]^2}{2\sigma_\tau^2} \right\} \tag{9}$$

Substituting Eq. (9) in Eq. (6), an MRE posterior of θ is derived to be

$$p(\theta) \propto \mu(\theta) \frac{1}{\sigma_\tau^{\ n}} \exp\left\{ -\sum_{i=1}^{n} \frac{\left[d_i - M(x \mid \theta)\right]^2}{2\sigma_\tau^2} \right\} e^{\beta g(\theta)} \tag{10}$$

For fatigue damage calculation of $M(x \mid \theta)$, various deterministic models have been proposed to describe the fatigue crack propagation, among which Paris type of laws are commonly used in cycle based fatigue crack growth calculation. In this study, Paris model (Paris and Erdogan, 1963) is used to compute the fatigue crack length for illustration purposes. Let a be the crack length, N be the number of cycles, the Paris' law reads,

$$\frac{da}{dN} = c(\Delta K)^m = c\left(\Delta\sigma \cdot \sqrt{\pi a} \cdot F(a)\right)^m \tag{11}$$

where c and m are material constants, ΔK is the variation of stress intensity factor in one cycle of stress amplitude $\Delta\sigma$, and $F(a)$ is the geometric correction factor. The crack size can be calculated by solving Eq. (11) numerically given c, m, and N. Early studies show that $\log(c)$ follows a normal distribution and m follows a truncated normal distribution (Kotulski, 1998). Assuming $\log(c)$ and m are independent variables and combining Eq. (11) with Eq. (10), the joint posterior can be expressed as

$$\begin{aligned} &p(\log(c), m) \propto \\ &\frac{1}{\sigma_c} \exp\left\{ -\frac{1}{2}\left(\frac{\log(c) - \zeta_c}{\sigma_c}\right)^2 + \beta_c g_c(\log(c)) \right\} \cdot \\ &\frac{1}{\sigma_m} \exp\left\{ -\frac{1}{2}\left(\frac{m - \zeta_m}{\sigma_m}\right)^2 + \beta_m g_m(m) \right\} \cdot \\ &\frac{1}{\sigma_\tau^n} \exp\left\{ -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{d_i - M_i(a_i \mid c, m, N_i)}{\sigma_\tau}\right)^2 \right\} \end{aligned} \tag{12}$$

The probability density function of one parameter can be obtained by integrating over the rest of parameters.

2.3 MCMC simulation

Directly evaluating the PDF of Eq. (12) is difficult because of multi-dimensional integration for normalization. In order to circumvent the direct evaluation of Eq. (12), Markov Chain Monte Carlo sampling technique is used in this study. MCMC was first introduced by (Metropolis *et al.*, 1953) as a method to simulate a discrete-time homogeneous Markov chain. The merit of MCMC is that it overcomes the normalization of Eq. (12) and ensures that the state of the chain converge to the desired distribution after a large number of steps from an arbitrary initial start. The widely used random walk algorithm, Metropolis-Hastings algorithm (Hastings, 1970), is summarized here.

The transition between two successive samples x_t and x_{t+1} is defined by Eq. (13)

$$x_{t+1} = \begin{cases} \tilde{x} \sim q(X \mid x_t) & \textit{with probability } \alpha(x_t, \tilde{x}) \\ x_t & \textit{else} \end{cases} \tag{13}$$

where $q(X \mid x_t)$ is the transition distribution, and $\alpha(x_t, \tilde{x}) = \min(1, r)$ is the acceptance probability. The Metropolis ratio r is defined as,

$$r = \frac{p(\tilde{x})}{p(x_t)} \frac{q(x_t \mid \tilde{x})}{q(\tilde{x} \mid x_t)} \tag{14}$$

where $p(\cdot)$ is the posterior probability representation. In our case, $p(\cdot)$ is computed using Eq. (12). For a symmetric transition distribution of $q(\cdot)$, such as a normal distribution, the property of $q(x_t \mid \tilde{x}) = q(\tilde{x} \mid x_t)$ simplifies Metropolis ratio in Eq. (14) to $r = p(\tilde{x}) / p(x_t)$. In this paper, 100,000 posterior samples of $(\log(c), m)$ are generated with a 5% burn-in period using a Gaussian transition distribution. Additionally, the moment information of these samples is then integrated into the proposed MRE updating procedure.

2.4 MRE updating procedure

The proposed MRE updating procedure for fatigue crack growth problem using Paris' law is described in the following steps. When one observation data (d_i, N_i) is available, MRE updating procedure begins with initial values of $\beta_c = \beta_m = 0$ in Eq. (12)

Step 1: Generate a sufficient number of $(\log(c), m)$ samples using MCMC.

Step 2: Calculate the crack size a_i at any interested cycles N and RUL using the samples obtained in step 1. This step is for crack size and RUL prognosis purpose.

Step 3: Calculate β_c and β_m with the statistics of the samples obtained in step 1 for the next updating.

If additional observation data become available, repeat the above steps 1-3. To illustrate this process, a flow chart of MRE updating procedure is shown in Figure 1. In order to exemplify the MRE updating procedure, two application examples are given in the next section.

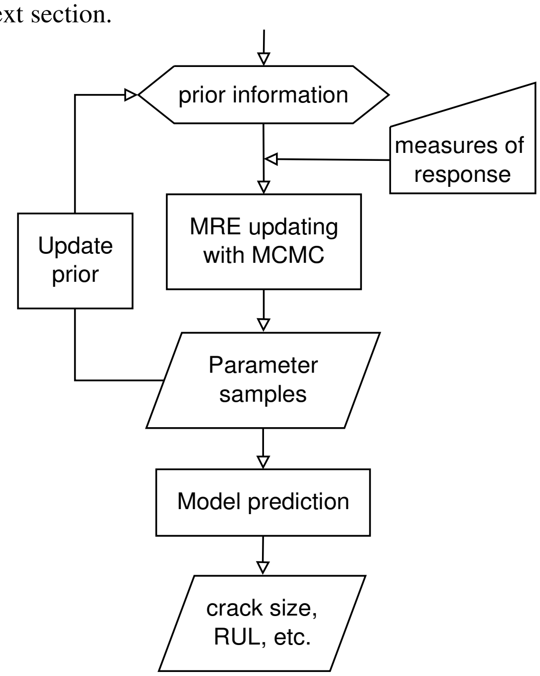

Figure 1: MRE updating procedures

3 APPLICATION EXAMPLES

Two fatigue crack growth experimental datasets are used to demonstrate the proposed MRE updating procedure to show the benefits of this approach.

3.1 Virkler's 2024-T3 aluminum alloy experimental data

An extensive fatigue crack growth data under constant loading for Al 2024-T3 plate specimens with center through cracks was collected in (Virkler *et al.*, 1979). The dataset consists of 68 fatigue crack growth trajectories and each trajectory contains 164 measurement points. All specimens have the same geometry, i.e., an initial crack size $a_i = 9mm$, length $L = 558.8mm$, width $w = 152.4mm$ and thickness $d = 2.54mm$. The loading information is $\Delta\sigma = 48.28MPa$ and stress ratio $R = 0.2$. The geometry correction factor for Virkler's experiments is $F(a) = 1/\sqrt{\cos(\pi a / w)}$. Kotulski (1998) reported the statistical information of the parameters in Paris' law, namely, mean values $\zeta_c = \langle \log(c) \rangle = -26.155$ and $\zeta_m = \langle m \rangle = 2.874$ with standard deviations $\sigma_c = 0.968$ and $\sigma_m = 0.164$, respectively. Assuming the total error term is $\sigma_\tau = 0.1mm$ (see Eq. (8)) and substituting the statistics information into Eq. (12) with $g_c(c) = \log(c)$ and $g_m = m$, the updating procedure can be performed when observation data become available.

One arbitrary crack growth trajectory in Virkler's dataset is selected for fatigue crack length prediction updating from Ostergaard and Hillberry (1983). Five points in the early stage of the crack propagation are randomly chosen to represent the measured values of crack length a obtained from health monitoring system or nondestructive inspection. The observation data points are shown in Table 1.

Table 1: Data used for updating (Virkler's dataset)

Number	Crack size (mm)	Cycle
1	9.7330	21269
2	10.5272	42734
3	11.2557	56392
4	12.1708	73161
5	15.0549	110487

The predictions of MRE updating and Bayesian updating are shown in Figure 2 where MRE updating gives a narrower prognosis interval and a more accurate median prediction compared to classical Bayesian updating. It further justifies that the additional moment constraints posed on posterior yield a more satisfactory prognosis results with narrower confidence bounds. The prior estimation is the point estimation computed using Eq. (11) with prior information, specifically, in this example, $\log(c) = -26.155$ and $m = 2.874$ reported in (Kotulski, 1998).

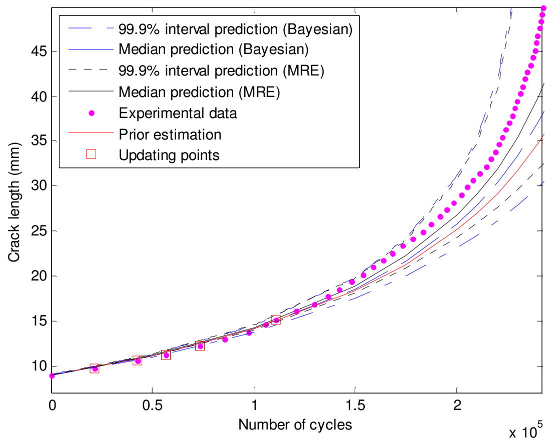

Figure 2: MRE and Bayesian prognosis (Virkler's dataset)

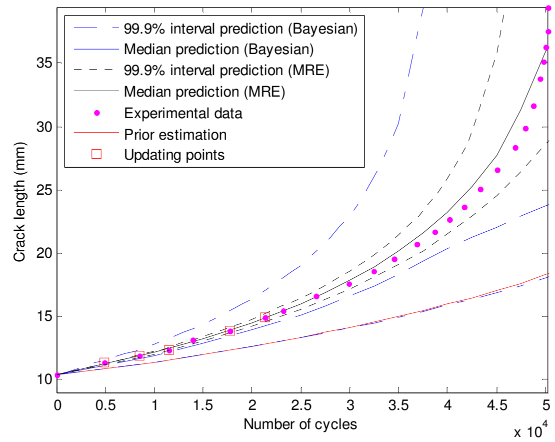

Figure 3: MRE and Bayesian prognosis (McMaster's dataset)

3.2 McMaster's 2024-T351 aluminum alloy experimental data

A large set of 2024-T351 aluminum alloy experimental data under constant and variable loading conditions was obtained in (McMaster and Smith, 1999) The experimental data of center-cracked specimens with length $L = 250mm$, width $w = 100mm$ and thickness $t = 6mm$ under constant loading $\Delta\sigma = 65.7MPa$ and stress ratio $R = 0.1$ are used in this paper. The priors of parameters are obtained by $\log(da/dN) \sim \log(\Delta K)$ regression using the experimental data. Five points shown in Table 2 are chosen arbitrarily as sensor measurements monitoring the process and are used for updating.

Table 2: Data used for updating (McMaster's dataset)

Number	Crack size (mm)	Cycle
1	11.3611	4875
2	11.9282	8475
3	12.3254	11550
4	13.8563	17775
5	14.8771	21375

The predictions of MRE and Bayesian updatings are shown in Figure 3, where interval predictions obtained by MRE updating is much narrower than that by Bayesian updating. In addition, the trend of median prediction of MRE updating is more accurate than that of Bayesian updating.

MRE updating shows the advantages over Bayesian updating in two application examples visually probably because of the additional statistical moment constraints of MCMC samples added to posteriors. To quantify the performance, prognosis metrics need to be considered to provide a rigorous comparison between MRE updating and Bayesian updating as shown below.

4 MODEL COMPARISON BASED ON PROGNOSIS

The Various metrics are available to quantify the performance of prognosis algorithms (Saxena *et al.*, 2008). In this section, classical error based statistical measures and several prognosis metrics are applied to quantify the prediction performance of application examples in the previous section.

4.1 Statistical metrics

Metrics, such as mean squared error (MSE), mean absolute percentage error (MAPE), average bias, sample standard deviation (STD), and their variations are widely used in medicine and finance fields where large datasets are available for statistical data analysis (Saxena *et al.*, 2008). The results for those classical metrics shown in Table 3 and Table 4, (rows 1-4) are computed using the prediction residuals (the difference between actual RUL and predicted RUL) obtained after the fifth updating. The proposed MRE approach shows its advantages over Bayesian method in all cases.

4.2 Prognosis metrics

The metrics mentioned in Section 4.1 are general purpose metrics and not specifically designed for prognosis. In (Saxena *et al.*, 2008) authors proposed several metrics, such as Prognostic Horizon (PH), Alpha-Lambda (α-λ) Performance, Relative Accuracy (RA), Cumulative Relative Accuracy (CRA), and Convergence that were designed specifically for prognosis. These metrics help assess how well prediction estimates improve over time as more measurement data become available. For readers' reference, we present brief definitions of these metrics here.

1. Prognostic Horizon is defined as the length of time before end-of-life (EoL) when an algorithm starts predicts within specified accuracy limits. These limits are specified as $\pm\alpha\%$ of the true EoL.

2. ***α-λ*** **Accuracy** determines whether predictions from an algorithm are within $\pm\alpha\%$ accuracy of the true RUL at a given time instant, specified by the parameter λ. For instance a $\lambda = 0.5$ would specify midway between the first time a prediction is made and EoL.
3. Relative Accuracy quantifies the percent accuracy w.r.t. actual RUL at a given time (specified by λ). It's an accuracy measure normalized by RUL, signifying that predictions closer to EoL should be more accurate and precise.
4. Cumulative Relative Accuracy is a weighted average of RAs computed at different time instances. Weights can be assigned to the predictions based on how critical it becomes as EoL approaches, and hence the accuracy of the predictions.
5. Convergence quantifies the rate at which any performance metric of interest improves to reach its desired value as time passes by.

For more description, implementation details and application examples on these metrics reader is suggested to refer to (Saxena *et al.*, 2009). In general, these metrics were designed to capture the time varying aspects of prognostics. As more data become available prognostic estimates get revised. It is, therefore, important to track how well an algorithm performs as time passes by as opposed to evaluating performance at one specific time instant only. Further, these metrics also incorporate the notion of increased criticality as EoL approaches, which capture the notion that a successful prognosis algorithm should improve as the system approaches its EoL.

In this paper we compare the two approaches based on Bayesian and MRE updating. In addition to evaluating performance based on prognosis metrics, we also include some classical statistical metrics. For this purpose, in our approach we include an additional updating point from the end of time series to establish EoL and compute the RUL curves. Results obtained from this evaluation exercise are presented next.

Performance results for Virkler's dataset

The results for PH, α-λ accuracy, and RA metrics comparing Bayesian and the proposed MRE updating algorithms are shown in Figure 4, Figure 5, and Figure 6, respectively. For computing CRA (see Table 3), the starting point is cycle zero because the specimens have initial cracks. We evaluated RA at 20, 40, 60, and 80% of EoL and did not use weighting factors. This assumes that relative accuracy is equally weighted at all time instants. Though this may not always be preferable, a simplistic evaluation was carried out to observe the natural behavior of the algorithm itself.

Figure 4 compares the prediction horizon for the two algorithms with 10% error bound around EoL value. Using the strict definition for PH as laid out in (Saxena *et al.*, 2009) we observed that MRE yields a larger PH. The plot of PH performance in Figure 4 shows that MRE prediction enters the 90% accuracy zone at the fourth updating, while Bayesian prediction enters the zone at the fifth updating showing that MRE is slightly better than Bayesian. In general it indicates that, for engineering practice, the proposed MRE can give an informative prediction at an earlier stage of the whole lifecycle.

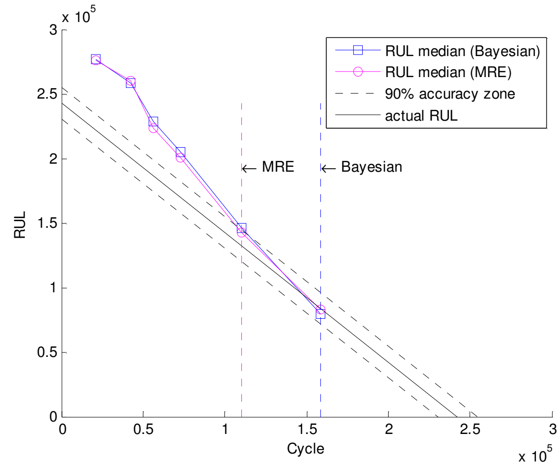

Figure 4: Performance comparison for PH at α=0.1 (10% error bound) on Virkler's dataset

Figure 5 compares the accuracy at $\lambda = 0.4$ and as we can see both algorithms fail to satisfy 10% error bounds. But MRE performs marginally better than Bayesian method at this particular time instant. We chose $\lambda = 0.4$ arbitrarily, however, depending on a specific application it may be desirable to know how different algorithms compare at a given time instant (λ) that may hold specific significance related to time criticality due to the approaching failure.

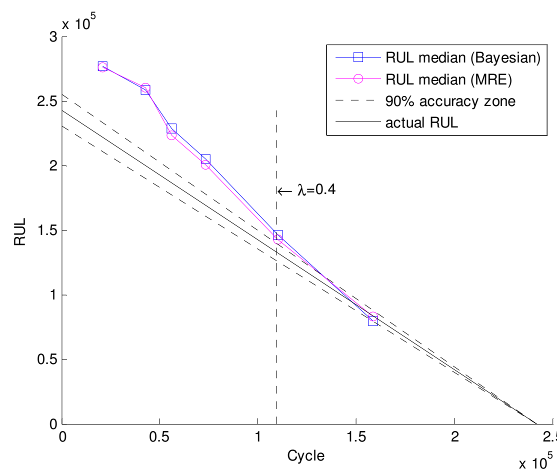

Figure 5: Performance comparison for α-λ accuracy (α=0.1 and λ=0.4) on Virkler's dataset

$RA_{\lambda=0.4}$ value indicates that the proposed MRE approach achieves better relative accuracy. It can be shown in Figure 6 that in general, the prediction curve obtained by MRE approach is slightly closer to the actual RUL than that by Bayesian approach. This of course is dependent on the choice of λ. For instance, for some other time instances Bayesian might have performed better.

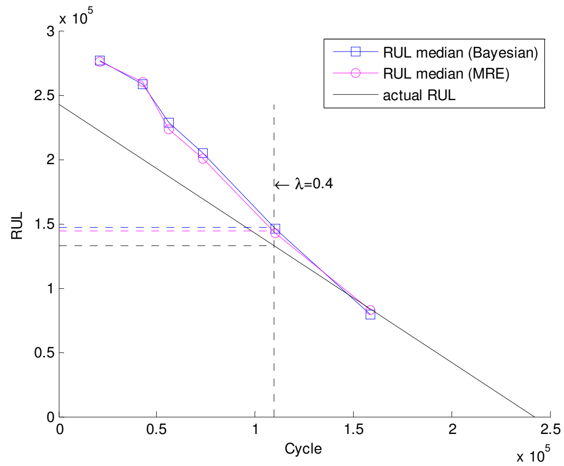

Figure 6: Performance comparison for RA at λ=0.4 on Virkler's dataset

Other prognostics metrics like RCA and convergence were computed and are given in Table 3 along with conventional statistical metrics. Looking at the table one can see that on Virkler's dataset MRE performs better than Bayesian approach under all performance measures. One must note that although classical metrics conclude the same as the new prognostics metrics, they do not take into account the time varying nature of the prognostics and hence may not always be useful in practice.

Table 3 Comparison of metrics between MRE and Bayesian approaches (Virkler's dataset)

Metric	MRE	Bayesian
MAPE	8.66	10.93
Average Bias (cycles)	10956.27	14051.92
STD (cycles)	7628.77	9115.78
MSE(cycle2)	178.23×10^6	280.5×10^6
RAλ=0.4	0.92	0.89
CRA λ=0.4	0.89	0.87
Convergence	74365.72	77349.24

Performance results for McMaster's dataset

Next, we perform a similar analysis for the McMaster's dataset. The results for PH, α-λ accuracy, and RA metrics comparing Bayesian and MRE updating are shown in Figure 7, Figure 8 and Figure 9, respectively. The rest of the metrics are included in Table 4. Looking at these results, the general conclusion about the superior performance of the MRE procedure from Virkler's dataset is further strengthened. The MRE's superior performance over Bayesian approach is attributed to the ability to incorporate additional knowledge about the system using additional constraints.

Table 4 Comparisons of metrics between MRE and Bayesian approaches (McMaster's dataset)

Metric	MRE	Bayesian
MAPE	4.06	22.53
Average Bias (cycles)	418.76	4561.93
STD (cycles)	1413.53	6888.38
MSE (cycle^2)	2.17×10^6	68.26×10^6
RA $_{\lambda=0.4}$	0.99	0.86
CRA $_{\lambda=0.4}$	0.95	0.87
Convergence	13757.94	22175.16

For this dataset these metrics clearly distinguish the two approaches and show better outcomes from the MRE method. For example, the α-λ and RA performance metrics shown in Figure 8 and Figure 9, respectively, present clear visual comparisons.

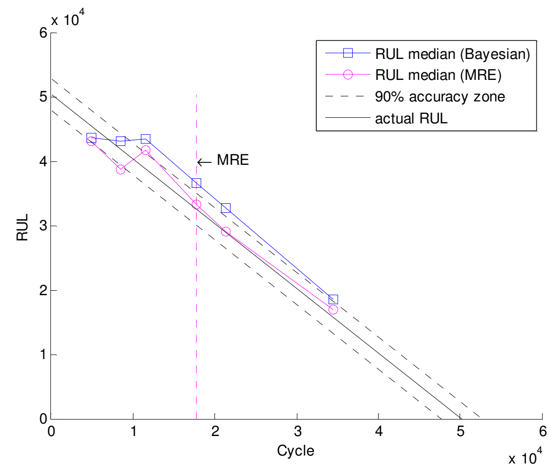

Figure 7: Performance comparison for PH at α=0.1 (10% error bound) on McMaster's dataset

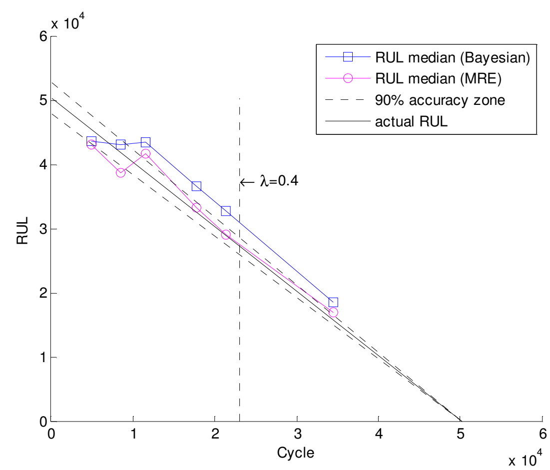

Figure 8: Performance comparison for α-λ accuracy (α=0.1 and λ=0.4) on McMaster's dataset

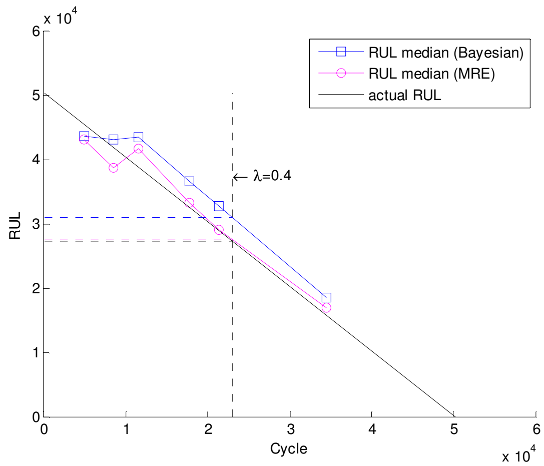

Figure 9: Comparison of RA performance (McMaster's dataset)

5 DISCUSSION

As observed in the previous section, there are a few aspects where these metrics can be further enhanced to improve performance evaluation. For instance, whereas it was straight forward to establish different PHs in these examples as per the original definition, it must be noted that the performance was established based on point estimates. However, the methods examined here produce PDFs. Depending which method is chosen to characterize the PDF, this may result in different performance outcomes. For instance, it is possible that PHs reverse order if mean or any other estimator is used instead of median. Furthermore, a point estimate representation of a PDF neglects to take advantage of the rich information a PDF affords. Performance metrics that incorporate the PDF are not currently available and need to be developed to allow a fair comparison of algorithms.

The significant difference between the PHs for the two algorithms may also be an artifact of the frequency at which these algorithms make a prediction.

We also observed that in a probabilistic prognosis updating scheme, the selection of priors may produce different prognosis results and affect the performance. Consequently, different updating methods may exhibit different robustness with inappropriate priors. Next, we discuss some of these issues as they relate to prognosis metrics.

5.1 Convergence metric

The convergence metric computes a value to quantify how fast prognostic estimates improve and converge towards the ground truth. A metric like convergence is meaningful only if an algorithm improves with time and passes various criteria defined by other prognostic metrics. For example, the convergence in terms of RA using Virkler's data (Figure 10) shows a monotonic decreasing trend after the second update. But for McMaster's dataset (Figure 11) the RE curve for MRE shows a non-monotonic trend and Bayesian curve shows a diverging trend. This suggests that a metric like convergence will not make complete sense if the algorithms do not show improvements with time and hence more fine tunings of the algorithms are required.

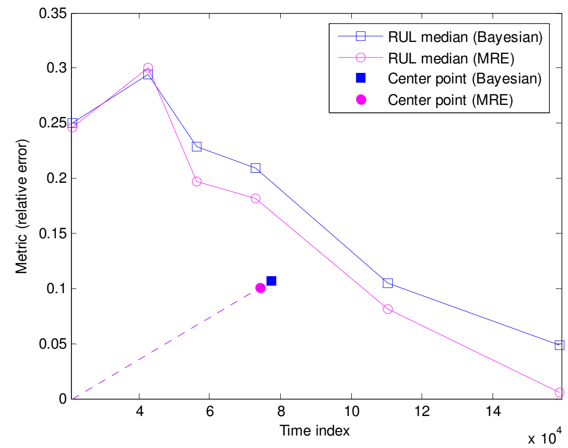

Figure 10: Comparison of convergence performance (Virkler's dataset)

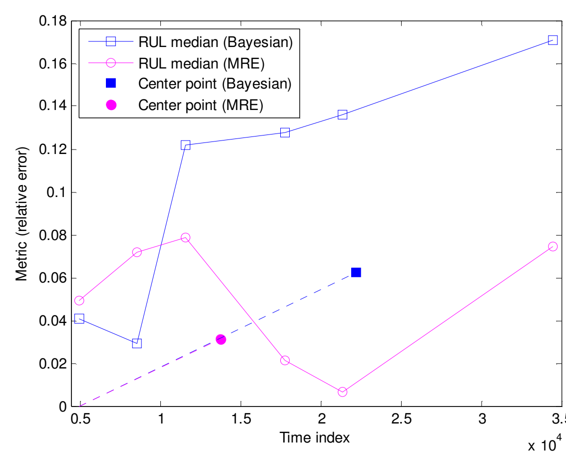

Figure 11: Comparison of convergence performance (McMaster's dataset)

5.2 Robustness metric

From the examples explored here, we found that the selection of a prior PDF is critical for a meaningful prognosis using probabilistic updating schemes such as Bayesian and MRE. An inaccurate prior may render a poor prediction of RUL. For example, when the prior prediction (red solid line) shown in Figure 12 is very different from the actual distribution, the Bayesian predictions lead to inaccurate estimates with very wide confidence bounds. However, the proposed MRE updating methodology performs well while using the same inaccurate prior. On the other hand, starting with a relatively accurate prior prediction, both MRE and

Bayesian give similar predictions as shown in Figure 13. This calls for a robustness metric that can quantify such effects. Such metrics may also be defined with respect to other factors like initial conditions, training data size, and prior knowledge.

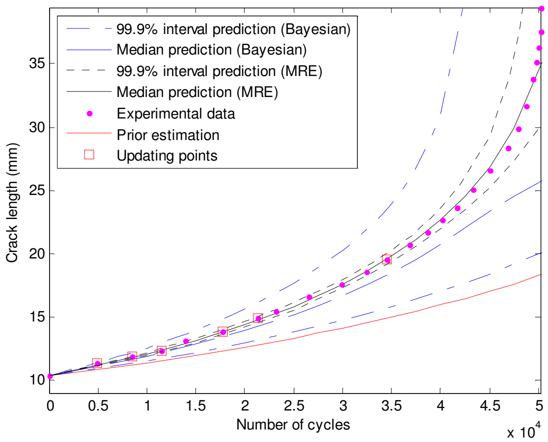

Figure 12: MRE and Bayesian prognosis with an inaccurate prior (McMaster's dataset)

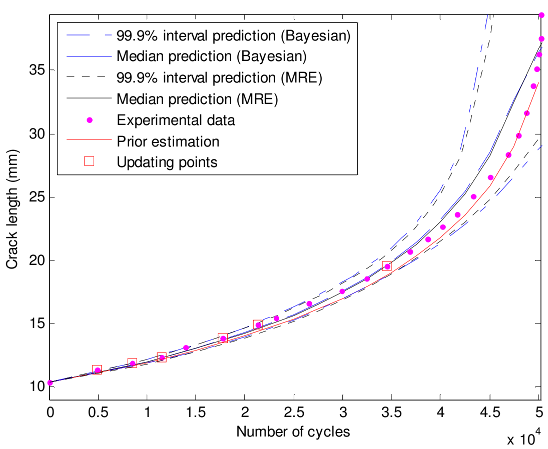

Figure 13: MRE and Bayesian prognosis with an accurate prior (McMaster's dataset)

5.3 RUL distribution

RUL prediction distributions are obtained by using a probabilistic framework for prognosis. Using only a central tendency point estimator like the sample mean or median will ignore useful information regarding the spread of the distribution of the RUL predictions. To tackle such situations we suggest plotting confidence intervals around median prediction. This will enhance the amount of visual information that most of these metrics tend to convey. As an illustration, we provide confidence intervals for predictions obtained for McMaster's dataset. The plot in Figure 14 shows distributions of RUL with a 0.95 confidence interval. Beyond the information already deduced, it also shows that MRE gives a narrower interval prediction and can be trusted more than the Bayesian approach.

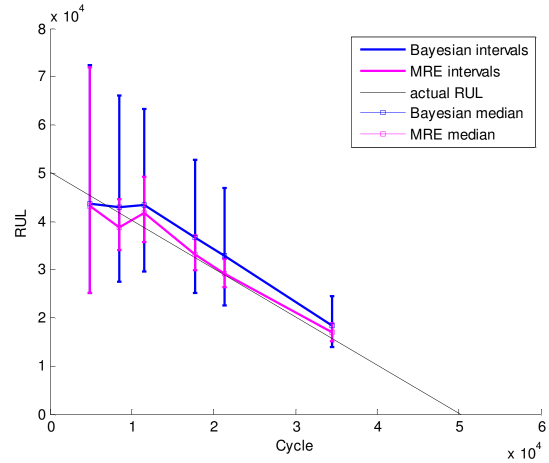

Figure 14: RUL median and 95% confidence intervals

6 CONCLUSION

A general framework for probabilistic prognosis using maximum entropy approach, MRE, is proposed in this paper to include all available information and uncertainties for RUL prediction. Prognosis metrics were used for model comparison and performance evaluation. Several conclusions can be drawn based on the results in the current investigation:

- The proposed MRE updating approach results in more accurate and precise prediction compared with the classical Bayesian method.
- The classical Bayesian method is a special case of the proposed MRE approach and MRE approach is more flexible to include additional information for inference, which cannot be handled by the classical Bayesian method.
-The proposed prognosis metrics can be successfully used for algorithm comparison and can give quantitative values in model (algorithm) performance evaluation.
- It is important to realize when to apply these metrics to arrive at meaningful interpretations. For instance, use of the convergence metric makes sense only when the algorithm predictions converge (get better) with time.
- Metrics to evaluate the robustness or sensitivity of models with respect to prior information, training data size, model initial parameters, noise levels, etc. are also desired. Further studies are needed to establish such concepts.

To make further comparison between different updating and prognosis approaches, more data points and even the whole dataset can be used as observation data to see with enough measures of response whether MRE and Bayesian give similar prognosis results and show convergence. Though in practice it is more desirable to get an early stage accurate prognosis, it is necessary to explore the characteristics of different updating algorithms using experimental data as we showed in previous sections.

ACKNOWLEDGMENT

The research reported in this paper was supported by funds from NASA Ames Research Center, Contract No. NNX09AY54A. The support is gratefully acknowledged.

NOMENCLATURE

- $I(\cdot)$ Relative information entropy
- $\mu(\cdot)$ Prior PDF
- $p(\cdot)$ Posterior PDF
- $L(\cdot)$ Likelihood function
- $M(\cdot)$ Model prediction of crack length
- $F(\cdot)$ Geometric correction factor
- N_i Number of cycles
- d_i Actual crack length

REFERENCES

(Kullback and Leibler, 1951) S. Kullback and R. A. Leibler (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, vol. 22, pp.79-86.

(Metropolis *et al.*, 1953) N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, vol. 21(6), pp.1087-1092.

(Paris and Erdogan, 1963) P. Paris and F. Erdogan (1963). A critical analysis of crack propagation laws. *Trans. The ASME, J. Basic Eng*. pp. 528-534.

(Hastings, 1970) W. K. Hastings (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, vol. 57(1), pp. 97-109.

(Virkler *et al.*, 1979) D. A. Virkler, B. M. Hillberry, and P. K. Goel (1979). The statistical nature of fatigue crack propagation, *ASME Journal of Engineering Materials and Technology*, vol. 101, pp.148-153.

(Ostergaard and Hillberry, 1983) D. F. Ostergaard and B. M. Hillberry (1983). Characterization of the variablility in fatigue crack propagation data, In J. M. Bloom and J. C. Ekvall (Eds.), *Probabilistic Fracture Mechanics and Fatigue Methods: Application for Structural Design and Maintenance*, ASTM STP 798, pp. 97-115.

(Skilling, 1988) J. Skilling (1988). The axioms of maximum entropy. In G.J. Erickson and C.R. Smith (Eds.), *Maximum Entropy and Bayesian Methods, in Science and Engineering, volume 1*. pp 173-187, Netherlands: Kluwer Academic Publishers.

(Madsen, 1997) H. Madsen (1997). Stochastic modeling of fatigue crack growth and inspection. In C.G. Soares (Ed.), *Probabilistic methods for structure design, collection Solid Mechanics and its applications*, pp. 59-83. Netherlands: Kluwer Academic Publishers.

(Kotulski, 1998) Z. A. Kotulski (1998). On efficiency of identification of a stochastic crack propagation model based on Virkler experimental data. *Archives of Mechanics*, vol. 50(5), pp. 829-847.

(McMaster and Smith, 1999) F. J. McMaster and D. J. Smith (1999). Effect of load excursions and specimen thickness on crack closure measurements, *Advances in Fatigue Crack Closure Measurements and Analysis: Second Volume, ASTM STP 1343*, pp 246-264. R. C. McClung and J. C. Newman, Jr., Eds., West Conshohocken, Pennsylvania: American Society for Testing and Materials.

(Zhang and Mahadevan, 2000) R. Zhang and S. Mahadevan (2000). Model uncertainty and Bayesian updating in reliability-based inspection. *Structure Safety*, vol. 22(2), pp.145-160.

(Perrin *et. al.*, 2007) F. Perrin, B. Sudret, and M. Pendola (2007). Bayesian updating of mechanical models – Application in fracture mechanics. *In 18ème Congrès Français de Mécanique*, CFM 2007, Grenoble.

(Caticha and Preuss, 2004) A. Caticha and R. Preuss (2004). Maximum entropy and Bayesian data analysis: Entropic prior distributions. *Physical Review E*. vol. 70(4), 046127.

(Caticha and Giffin, 2006) A. Caticha and A. Giffin (2006). Updating probabilities. In A.M. Djafari (Ed.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering, AIP Conference Proceedings*, 31: 872.

(Giffin and Caticha, 2007) A. Giffin and A. Caticha (2007). Updating probabilities with data and moments. In K. Knuth (Ed.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, AIP Conference Proceedings, 954:74

(Tseng and Caticha, 2008) C. Tseng and A. Caticha (2008). Using relative entropy to find optimal approximations: An application to simple fluids. *Physica A*, vol. 387, pp. 6759-6770.

(Bourdin *et. al.*, 2008) B. Bourdin, G. A. Francfort and J. J. Marigo (2008). *The Variational Approach to Fracture*, Netherlands: Springer.

(Saxena *et al.*, 2008) A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and M. Schwabacher (2008). *Metrics for evaluating performance of prognostic techniques.* International Conference on

Prognostics and Health Management, PHM 2008. 6-9 Oct. 2008, pp: 1-17.

(Saxena *et al.*, 2009) A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel (2009). *Evaluating algorithm performance metrics tailored for prognostics.* IEEE Aerospace conference, 7-14 March 2009, pp. 1-13.

Xuefei Guan is a graduate research assistant in department of mechanical engineering at Clarkson University. He received his B.S. degree in Reliability Engineering and M.S. degree in Aeronautical Engineering from Beihang University in China in 2005 and 2008, respectively. Prior his study at Clarkson, he worked as a software engineer in Thomson Broadband (Beijing) R&D Co., LTD. His research interests are probabilistic prognosis, Bayes and entropy based methods and applications, Markov Chain Monte Carlo simulation, and system reliability.

Yongming Liu is an assistant Professor in the department of civil and environmental engineering. His research interests include fatigue and fracture analysis of metals and composite materials, probabilistic methods, computational mechanics, and risk management. He completed his PhD at Vanderbilt University, and obtained his Bachelors' and Masters' degrees from Tongji University in China. Dr. Liu is a member of ASCE and AIAA and serves on several technical committees on probabilistic methods and advanced materials.

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.

Jose R. Celaya is a staff scientist with Stinger Ghaffarian Technologies 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, Mexico.

Kai Goebel received the degree of Diplom-Ingenieur from the Technische Universität München, 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 leads the Diagnostics & Prognostics groups in the Intelligent Systems division. In addition, he directs the Prognostics Center of Excellence and he is the Associate Principal Investigator for Prognostics of NASA's Integrated Vehicle Health Management Program. He worked at General Electric's 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 eleven patents and has published more than 100 papers in the area of systems health management.