Chinese Journal of Aeronautics, (2015), 28(4): 1055–1065

Chinese Society of Aeronautics and Astronautics
& Beihang University

Chinese Journal of Aeronautics

cja@buaa.edu.cn
www.sciencedirect.com

Probabilistic inference of fatigue damage propagation with limited and partial information

Huang Min [a], **He Jingjing** [a,*], **Guan Xuefei** [b]

[a] *School of Reliability and System Engineering, Beihang University, Beijing 100191, China*
[b] *Siemens Corporation, Corporate Technology, Princeton, NJ 08540, USA*

Received 30 April 2014; revised 3 November 2014; accepted 29 April 2015
Available online 21 June 2015

KEYWORDS

Fatigue damage propagation;
Maximum relative entropy;
Partial information;
Probability updating;
Uncertainty

Abstract A general method of probabilistic fatigue damage prognostics using limited and partial information is developed. Limited and partial information refers to measurable data that are not enough or cannot directly be used to statistically identify model parameter using traditional regression analysis. In the proposed method, the prior probability distribution of model parameters is derived based on the principle of maximum entropy (MaxEnt) using the limited and partial information as constraints. The posterior distribution is formulated using the principle of maximum relative entropy (MRE) to perform probability updating when new information is available and reduces uncertainty in prognosis results. It is shown that the posterior distribution is equivalent to a Bayesian posterior when the new information used for updating is point measurements. A numerical quadrature interpolating method is used to calculate the asymptotic approximation for the prior distribution. Once the prior is obtained, subsequent measurement data are used to perform updating using Markov chain Monte Carlo (MCMC) simulations. Fatigue crack prognosis problems with experimental data are presented for demonstration and validation.

1. Introduction

Fatigue crack damage of materials exhibits significant uncertainties due to the unstable and stochastic nature of crack propagation mechanism. Accurate deterministic fatigue damage prognosis is difficult to achieve under realistic service conditions. Therefore uncertainty quantification for fatigue damage prognosis using probabilistic methods is usually required to obtain reliable results. Uncertainties in fatigue damage prognostics arise from several sources such as material properties, loading, environmental conditions and the geometry of the cracked-component. Stress intensity factor (SIF)-driven methods are commonly used to model the fatigue crack propagation rate. For example, the classical Paris' equation and its variants.[1–3] To effectively use those models for fatigue crack prognosis, sufficient fatigue testing data are required to identify model parameters. Model prediction may be unreliable when usage condition is very different from the one under

* Corresponding author. Tel.: +86 10 82314879.
E-mail address: hejingjing@buaa.edu.cn (J. He).
Peer review under responsibility of Editorial Committee of CJA.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.cja.2015.06.017
1000-9361

which model parameters are calibrated[4], and updating becomes highly useful and necessary. Probabilistic fatigue damage prognostics using Bayes' rule requires a prior probability density function (PDF) of model parameters identified from a large set of experimental data[5–7] and point measurement data.[8,9] Whether to choose or obtain a prior PDF depends on what information is available. For statistical identification of model parameters, a large set of repeated tests under the same condition is required which is very expensive. Another approach, in the absence of any information, is to construct a homogeneous/uniform probability distribution that assigns to each region of the parameter space a probability proportional to the volume of the region, which is also called non-informative prior in a Bayesian context.[10] However, some non-informative priors cannot be normalized.[11–13] In such cases, methods based on the transformation group[14,15] or reference prior[16,17] can be adopted, but analysis of the specific problem is needed. Limited or partial information, such as the mean value for a specific function involving model parameters, is sometimes available from historical data or field testing. However, there is no formal rule to utilize the partial information to obtain the prior PDF of model parameters or to perform updating in the classical Bayesian framework.

To resolve the above difficulties in probabilistic fatigue damage prognostics, two major extensions are demanded: (1) a reliable initial estimation method for model parameters, which allows for estimating the model parameter PDF in a rational manner under conditions where no enough fatigue testing data are available, and (2) a general updating rule that is capable of handling different types of measurement data. The first aspect is realistic but challenging considering the fact that exhaustively performing fatigue testing for all usage conditions, particularly for unforeseen usage conditions, is not practical. The second aspect is also demanding because not all measurable data are in the form of point measurements which can directly be incorporated for updating using Bayes' rule. The objective of this study is thus not to provide a superior method to the regression method or to argue limitations of Bayes' rule. The objective of the study, however, is to develop a method allowing one to identify the prior PDF of model parameters using limited or partial information for cases where the traditional statistical identification is difficult to apply due to the limited number of data points available for a normal regression, and to formulate an updating rule that is able to handle more versatile data from usage monitoring for uncertainty reduction in prognosis results. Therefore, the underlying assumptions made in this study are: there is no or not enough of testing data for statistical identification of model parameters using normal regression methods and partial information is available for updating. In the study, the probabilistic identification of model parameters given limited or partial information is proposed based on the principal of maximum entropy (MaxEnt), and the updating rule that is capable of handling additional information other than point measurements commonly seen in the classical Bayesian analysis is formulated based on the principal of maximum relative entropy (MRE).

The remainder of the paper is organized as follows. First, the probabilistic identification of model parameters using limited or partial information is derived to obtain the prior PDF of model parameters. To evaluate the prior PDF, a numerical quadrature interpolating method is proposed. Next, the updating rule is formulated according to the principle of MRE for probability updating given additional data such as response measures to reduce the uncertainty in prognostics. The resulting posterior PDF from updating can be evaluated either by analytical solution (if there exists one) or approximation methods such as Markov chain Monte Carlo (MCMC) simulations. Following that, a few fatigue prognosis problems with experiment data are presented to demonstrate and validate the effectiveness of the overall method.

2. Probabilistic model parameter identification with limited or partial information

Under the condition that there is no fatigue testing data available to identify the prior PDF of model parameters using the normal regression method, it is possible that sparse response measurements, such as one crack size measurement from a few cracked components being monitored, are sometimes available. It is not possible to conduct parameter estimation using regression since the number of data points is one and the number of model parameters is larger than one (e.g., two-parameter Paris' equation). The key idea is to treat the mean value of the one response measure associated with each individual target system as a mathematical expectation of the mechanism model output. The expectation value can be considered as a constraint to formulate the prior PDF using the principle of MaxEnt. Given a random variable θ and its probability distribution $p(\theta) \in \mathbf{R}^{+}$, the information entropy[18] of $\theta \in \Theta$ is defined as

$$H(\theta) = -\int_{\Theta} p(\theta) \ln p(\theta) \mathrm{d}\theta \tag{1}$$

The principle of MaxEnt states that the desired probability distribution is the one that maximizes the entropy subject to all constraints.[19] The usual constraints are the mathematical expectations of some functions that involve the variable θ. For example, the first and second order moments of θ, such as $E_{p(\theta)}(\theta)$ and $E_{p(\theta)}(\theta^2)$ or more general $E_{p(\theta)}(f(\theta))$ can serve as the constraints. Here $f(\cdot)$ represents a general real-valued function. The desired prior distribution $p(\theta)$ can be derived using the method of Lagrange multipliers. Given a general expectation constraint $E_{p(\theta)}(f(\theta)) = F$, the Lagrangian Λ reads

$$\Lambda = -\int_{\Theta} p(\theta) \ln p(\theta) \mathrm{d}\theta + \alpha\left(\int_{\Theta} p(\theta) \mathrm{d}\theta - 1\right) + \lambda\left(\int_{\Theta} p(\theta) f(\theta) \mathrm{d}\theta - F\right) \tag{2}$$

Maximizing Λ by $\delta\Lambda/\delta p(\theta) = 0$ to obtain

$$p(\theta) = \frac{1}{Z} \exp(\lambda f(\theta)) \tag{3}$$

where $Z = \int_{\Theta} \exp(\lambda f(\theta)) \mathrm{d}\theta$ is the normalizing constant, and α and λ are Lagrange multipliers. The term λ is calculated by solving

$$\frac{\partial \ln\left(\int_{\Theta} \exp(\lambda f(\theta)) \mathrm{d}\theta\right)}{\partial \lambda} = F \tag{4}$$

The solution also holds true when θ is a vector of variables and $f(\theta)$ is a set of real-valued functions. For polynomial type of functions, such as $f_k(\theta) = \sum_{i=0}^{k} a_i \theta^i$, Eq. (4) has an analytical expression when $k \leqslant 2$. Higher order moments or a more complicated form of function can only be solved by numerical methods.[20,21] As mentioned above, the mechanism model can be used as the function in the constraint $E_{p(\theta)}(f(\theta))=F$. By utilizing the mean value of the first measure associated with a few number of target systems as the expectation of model predictions, the prior PDF can be obtained. Denote the model as M, model parameter as θ, and the computed crack size as $M(\theta)$, the prior PDF can then be expressed as.

$$p(\theta)\propto\exp(\lambda M(\theta)) \tag{5}$$

where λ is computed as

$$\frac{\partial\ln(\int_{\Theta}\exp(\lambda M(\theta))\mathrm{d}\theta)}{\partial\lambda}=\bar{a} \tag{6}$$

where $\bar{a}$ is the mean value of the measurements. For example, given the first crack size measurements from three cracked-components working in a similar condition, $\bar{a}=(a_1+a_2+a_3)/3$. Here a_i represents the first crack size measurement of ith cracked-component. It is noted that the proportional relationship is used in Eq. (5) without the normalizing constant Z as shown in Eq. (3). This is due to the fact that the normalizing constant does not necessarily to be evaluated when MCMC simulations are used to draw samples from the (unnormalized) PDF.

3. Probability updating based on principle of MRE

Once the prior distribution is constructed, the posterior distribution can be calculated using the principle of MRE. Let $p_0(\theta)$ be the prior distribution of the parameter under model M. Probability updating on $p_0(\theta)$ can be performed when new information is available. The information may be a response measure and/or the mathematical expectation of a function of θ. Let $p(x|\theta)$ be the conditional probability distribution of observation $x\in X$ given θ. The joint distribution of x and θ is $p_0(x,\theta)=p_0(\theta)p(x|\theta)$. Let $p(x,\theta)$ be the optimal posterior distribution given new information as constraints. The search space for this optimal distribution is $X\times\Theta$. The relative information entropy (or the equivalent mathematical form of Kullback–Leibler (KL) divergence[22]) of the desired optimal posterior distribution $p(x,\theta)$ with respect to the prior distribution of $p_0(x,\theta)$ is defined as

$$I(p||p_0)=\int_{X\times\Theta}p(x,\theta)\ln\frac{p(x,\theta)}{p_0(x,\theta)}\mathrm{d}x\mathrm{d}\theta \tag{7}$$

The principle of MRE[22,23] states that given new facts, the new distribution $p(x,\theta)$ should be chosen which is as difficult to discriminate from the original distribution $p_0(x,\theta)$ as possible. The method of Lagrange multipliers can be used to obtain $p(x,\theta)$ by minimizing Eq. (7) under given constraints. The direct response measure of the event $x=x'$ can be formulated as the constraint $\int_{\Theta}p(x,\theta)\mathrm{d}\theta=\delta(x-x')$. Combining the normalization constraint $\int_{X\times\Theta}p(x,\theta)\mathrm{d}x\mathrm{d}\theta=1$, the Lagrangian can be expressed as

$$\begin{aligned}\Lambda=&\int_{X\times\Theta}p(x,\theta)\ln\frac{p(x,\theta)}{p_0(x,\theta)}\mathrm{d}x\mathrm{d}\theta+\alpha\left(\int_{X\times\Theta}p(x,\theta)\mathrm{d}x\mathrm{d}\theta-1\right)\\&+\int_X\beta(x)\left(\int_{\Theta}p(x,\theta)\mathrm{d}\theta-\delta(x-x')\right)\mathrm{d}x\end{aligned} \tag{8}$$

The optimal distribution $p(x,\theta)$ is obtained by $\delta\Lambda/\delta p(x,\theta)=0$ as

$$p(x,\theta)=\frac{1}{Z_0}p(x,\theta)\exp(-\beta(x)) \tag{9}$$

where $Z_0=\int_{X\times\Theta}p_0(x,\theta)\exp(-\beta(x))\mathrm{d}x\mathrm{d}\theta$ is the normalization constant. Substituting $p(x,\theta)$ in Eq. (9) into the constraint of $\int_{\Theta}p(x,\theta)\mathrm{d}\theta=\delta(x-x')$, integrating over θ yields

$$p(\theta)\propto p_0(\theta)p(x'|\theta) \tag{10}$$

which is identical to the usual Bayesian posterior distribution for θ given that the response measure x'. The term $p(x'|\theta)$ in Eq. (10) is also referred to as the likelihood function. Denote the deterministic model prediction for event (response measure) x as M. Considering the statistical mechanism modeling uncertainty ε_1 and the measurement uncertainty ε_2, the probabilistic description of x can be expressed as

$$x=M(\theta)+\varepsilon_1+\varepsilon_2 \tag{11}$$

Without loss of generality, terms ε_1 and ε_2 are usually described by independent Gaussian variables with the standard deviations of σ_{ε_1} and σ_{ε_2}, respectively. The likelihood function for n independent response measures is formulated as

$$p(x_1,x_2,\ldots,x_n|\theta)=\frac{1}{(\sqrt{2\pi}\sigma_{\varepsilon})^n}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-M_i(\theta)}{\sigma_{\varepsilon}}\right)^2\right] \tag{12}$$

where $\sigma_{\varepsilon}=\sqrt{\sigma_{\varepsilon_1}^2+\sigma_{\varepsilon_2}^2}$. Substituting Eq. (3) for the prior distribution and Eq. (12) for the likelihood function into Eq. (10), the posterior distribution is now

$$p(\theta)\propto\exp(\lambda f(\theta))\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-M_i(\theta)}{\sigma_{\varepsilon}}\right)^2\right] \tag{13}$$

If sufficient experimental data are available, function $f(\theta)$ can be chosen as θ or θ^2 and the prior distribution becomes the commonly used exponential family distribution. For example, given $E_{p(\theta)}(\theta)=\phi_1$ and $E_{p(\theta)}(\theta^2)=\phi_2$ only, the prior distribution $\exp(\lambda f(\theta))$ is the usual normal distribution with the mean value of ϕ_1 and variance of $\phi_2-\phi_1^2$. For realistic system and structures, several issues must be addressed. First, the direct statistical identification of θ may be difficult to apply because the response measures are usually not θ but some indirect variables. Second, the accurate statistical identification of θ may also require a large set of experimental data. Based on those considerations, the response measure prediction model $M(\theta)$ is used for $f(\theta)$ in the prior distribution. The posterior distribution becomes

$$p(\theta)\propto\exp(\lambda M(\theta))\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-M_i(\theta)}{\sigma_{\varepsilon}}\right)^2\right] \tag{14}$$

In many engineering problems, the parameter θ is multi-dimensional and direct evaluation of Eq. (4) for the Lagrange multiplier λ in the prior distribution is difficult. More general numerical methods, such as Monte Carlo simulations are usually used. To illustrate and validate the proposed procedure for building the entropic prior and posterior for fatigue damage prognostics, a practical example with experimental data is presented next.

4. Fatigue damage prognostics with limited or partial information

The fatigue crack damage propagation is one of the major failure modes for many engineering systems and the damage state and the crack growth need to be quantified to avoid catastrophic events. Many fatigue crack growth models are semi-empirical because the underlying mechanisms are either too sophisticated to be modeled exactly or the resulting models are too computationally intensive for practical use.[4] Therefore, updating the parameter using measurement data becomes an effective and efficient way to reduce the uncertainty in prognostics. In practice, parameters in the fatigue crack growth model are usually obtained through standard testing. In the standard testing, the geometry of the specimen has a specific configuration. The parameter value obtained in this manner is not generic to be used for other geometry or loading configurations. Therefore, using the parameter value obtained in one dataset to a different geometry and loading configuration may lead to unreliable results.[4] On the other hand, conducting the testing for the target component of interest may not be practical due to the economic and time constraints.

In this section, a practical fatigue crack growth example is presented based on the above considerations. First, given no testing data, the prior is obtained directly using response measures from a limited number of target components using fatigue crack growth model as the constraint function. The prior can reflect uncertainties associated with the target components as a whole. Then the subsequent measurement data associated with the target component of interest can be used to perform updating. The updating uses the component-specific measurement data and the updated results become more specific as more data are used in the updating process. Eventually, the measurement data diminish the effects from the prior information and become more relevant to the target component. To demonstrate the applicability of the proposed method, three commonly used fatigue crack growth models are included in this example. Two experimental datasets are used to validate the effectiveness of the proposed method.

4.1. Fatigue crack growth models

Three commonly used fatigue crack growth models, namely Paris' model,[1] Forman's model,[2] and McEvily's model,[3] are used here to demonstrate the applicability of the proposed method. The three models are briefly introduced for the completeness of the paper. More details about the models can be found in the referred articles. Paris' model is given as

$$\frac{\mathrm{d}a}{\mathrm{d}N} = c(\Delta K)^m \tag{15}$$

where a is the crack size, N is the number of applied cyclic loads, and c and m are model parameters. Following the convention, parameter $\ln c$ is generally used instead of c when fitting the model parameters. The term $\Delta K = \sqrt{\pi a}\Delta\sigma g(a/w)$ is the range of stress intensity factor during one cycle. The term $\Delta\sigma$ is the range of applied stress during one cycle, $g(a/w)$ the geometric correction term, and w the width of the specimen. Paris' model describes the lg–lg linear region in the $\mathrm{d}a/\mathrm{d}N - \Delta K$ coordinate.

Forman's model is stated as

$$\frac{\mathrm{d}a}{\mathrm{d}N} = \frac{c(\Delta K)^m}{(1-R)K_{\mathrm{cr}} - \Delta K} \tag{16}$$

where R is the load ratio, K_{cr} the fracture toughness of the material, and c and m are two parameters of the Forman's model.

McEvily's model is defined as

$$\frac{\mathrm{d}a}{\mathrm{d}N} = c(\Delta K - \Delta K_{\mathrm{th}})^2\left(1 + \frac{\Delta K}{K_{\mathrm{cr}} - K_{\max}}\right) \tag{17}$$

where ΔK_{th} is the threshold stress intensity range below which cracks either propagate at an extremely low rate or do not propagate at all. Knowledge of ΔK_{th} permits the calculation of permissible crack lengths and applied stresses in order to avoid fatigue crack growth.[4] $K_{\max}$ is the maximum stress intensity in one cyclic load. It should be noted that parameters c and m take different values in different models and they are usually obtained by experimental data via regression analysis. Statistical identification of the parameters usually requires a large number of coupon tests under the same conditions. It is worth mentioning that a PDF can be uniquely determined by the method of maximum a posterior estimation (MPE) with the first fourth moments, and the PDF may not be unique since different PDFs of parameters may lead to the same mean of the crack size when only using the first moment as a constraint. Since choosing constraints is entirely subjectively, it should be made based on the data available and the problem. When only limited crack growth curve is available, e.g., two or three curves, the statistical identification for the variance or higher order moments is not reliable. This scenario is typical for practical problems where the component is usually different from the specimen in the standard testing. The mean crack measurement can be a reasonable approximation for mathematical expectation of the model prediction. Higher order moments can also be applied when they can be reliably estimated. In this study only the most reliable constraint is applied to demonstrate the development of the method; therefore, the constraint is formulated as $E_{p_0(\theta)}(M(\theta)) = \bar{a}$. The prior distribution can be expressed, according to Eq. (3), as

$$p_0(\theta) \propto \exp(\lambda M(\theta)) \tag{18}$$

where θ is the model parameter and $M(\theta)$ is the model prediction. For example, using Paris' model, $\theta = (\ln c, m)$ and $M(\theta)$ is the crack size computed by Eq. (4). $p_0(\theta)$ is the prior distribution of the model parameter and $E_{p_0(\theta)}(M(\theta))$ represents the mathematical expectation of $M(\theta)$ under the distribution of $p_0(\theta)$. $\bar{a}$ is the mean value of the crack size measures. The Lagrange multiplier λ is obtained by solving

$$\frac{\partial \ln(\int_\Theta \lambda M(\theta)\mathrm{d}\theta)}{\partial \lambda} = \bar{a} \tag{19}$$

4.2. Compact tension (CT) specimen testing data and prognostics

The testing data reported in Ref.[24] includes crack growth trajectories for 2024-T351 aluminum alloy compact tension specimens. The dimensions of the specimens are 50.0 mm wide and 12.0 mm thick, with nominal yield strength and ultimate yield strength of 320 MPa and 462 MPa, respectively. The initial crack length is 18.0 mm. Sinusoidal signals with the maximum force of 4.5 kN and the minimal force of 0.9 kN are used as the input loads. The loading frequency is 15 Hz. In order to reflect the largest uncertainty in specimens, three crack growth curves with the fastest, moderate, and the slowest crack growth rate are chosen from the dataset to represent the only available testing data. The three crack growth curves are shown in Fig. 1, where Specimens 1–3 represent specimens with the fastest, moderate, and slowest crack growth rate, respectively.

Assume the crack size for the three specimens at 10^4 cycles are measured and the results are 18.71 mm, 18.92 mm, and 19.19 mm associated with the slow, medium and fast rate curves, respectively. With this information only (the first response measure for the three specimens), the deterministic values for parameters in Paris' model for the three specimens cannot be obtained since the number of unknowns is larger than the measured points. However, the entropic prior distribution can be constructed according to Eqs. (18) and (19) by

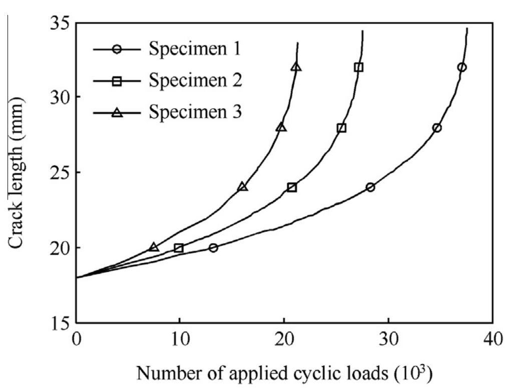

Fig. 1 Three crack growth curves of compact tension specimens (data source is from Ref.[24]).

treating the mean values of the three crack measures as the expectation of Paris' model prediction. Therefore, in Eq. (19), $\bar{a} = (18.71 + 18.92 + 19.19)/3 = 18.94$ mm. In the case where only one measurement point is available for each of those models, the classical statistical and deterministic method for parameter estimation is difficult to apply because the number of unknowns is larger than the minimal required data points. For example, the Paris' model has two parameters $(\ln c, m)$ whereas only one equation $M(\ln c, m) = 18.71$ mm is available.

Solving for the Lagrange multiplier λ in Eq. (19) is an optimization problem in nature. Classical gradient-based optimization algorithms can be directly applied. For problems with a small number of parameters, numerical quadrature is efficient to obtain the integral of $\int_{\Theta} \lambda M(\theta) \mathrm{d}\theta$ in Eq. (19) for a given value of λ. For problems with a large number of parameters, simulation-based method can be used to evaluate this integral. Since the solution of λ (i.e., the Lagrange multiplier) is unique[25], interpolation can also be adopted to reduce the total number of integral evaluations. Denote the value of $E_{p_0(\theta)}(M(\theta))$ associated with a specific value of λ as $\bar{a}_{\lambda}$. Given λ taking a set of different values, the corresponding $\bar{a}_{\lambda}$ can be obtained either by numerical quadrature or simulation-based methods. Therefore, given the actual measured value of $\bar{a} = 18.94$ mm, the desired solution for λ can be interpolated. In this paper, a numerical quadrature interpolating method is used to obtain the Lagrange multiplier λ. Parameters $\ln c$ and m are related to crack growth properties of the material. Theoretically the integration ranges for $\ln c$ and m are $(-\infty, +\infty)$ and $(0, +\infty)$, respectively. Due to practical and empirical considerations of model parameters, $\ln c$ is bounded in the range of $[-35, -5]$ and m is bounded in $[1, 5]$ to improve numerical efficiency. The chosen bounds in this study are conservative enough compared with the observed ones from experimental data. For example, typical metal materials have a value of M around 2.5 to 4 and $\ln c$ around -26 to -20.[26] As long as the chosen range covers most of the probability mass of a variable, the variation of the calculated results due to the change of integration range is small enough and the results can still remain the precision for engineering purposes. To interpolate the Lagrange multiplier λ associated with $\bar{a} = 18.94$ mm, a set of λ_i uniformly takn from $[-0.2, -0.1]$ is used in Eq. (19) to evaluate the corresponding $\bar{a}_{\lambda_i}$. In this study, a_{λ_i} is calculated using the quadrature functions of MATLAB 2008a and results are shown in Table 1. It is

Table 1 Values of $\bar{a}_{\lambda_i}$ for different λ_i calculated using numerical quadrature.

Model	$\bar{a}_{\lambda_i}$					
	$\lambda_i = -0.20$	$\lambda_i = -0.19$	$\lambda_i = -0.18$	$\lambda_i = -0.17$	$\lambda_i = -0.16$	$\lambda_i = -0.15$
Paris	18.432	18.481	18.546	18.630	18.742	18.890
Forman	18.244	18.266	18.294	18.331	18.379	18.443
McEvily	18.265	18.290	18.321	18.361	18.413	18.484
Model	$\bar{a}_{\lambda_i}$					
	$\lambda_i = -0.14$	$\lambda_i = -0.13$	$\lambda_i = -0.12$	$\lambda_i = -0.11$	$\lambda_i = -0.10$	
Paris	19.087	19.349	19.696	20.154	20.755	
Forman	18.528	18.642	18.794	18.988	19.271	
McEvily	18.574	18.697	18.858	19.077	19.372	

expected that the measurement data will dominate the crack growth predictions when more and more data are used for updating. As a result, the influence of the prior PDF becomes weak. Although the three specimens share the same prior PDF, the predicted crack growth trajectory will gradually converge to its physical one with continuous updating.

For $\bar{a} = 18.94$ mm, the corresponding λ is obtained from interpolation as -0.1471, -0.1126, and -0.1158 for Paris' model, Forman's model, and McEvily's model, respectively. The general posterior distribution for multiple response measures can then be written as

$$p(\theta) \propto \exp(\lambda M(\theta)) \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{a_i - M_i(\theta)}{\sigma_\varepsilon}\right)^2\right] \tag{20}$$

where a_i is the subsequent new measurement data on crack length associated with the target component, σ_ε the standard

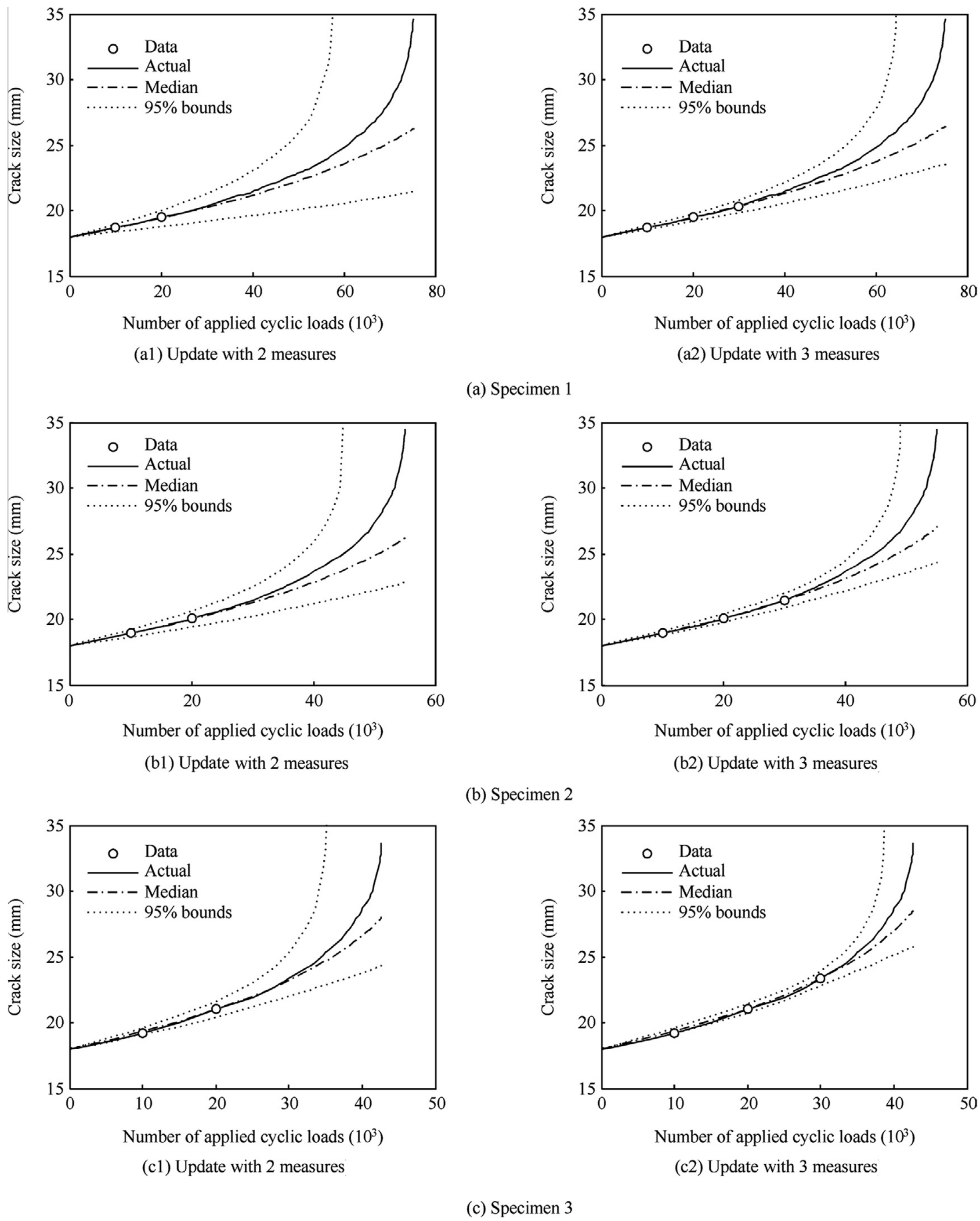

Fig. 2 Fatigue crack growth prognostics update using Paris' model.

deviation of Gaussian likelihood and n the total number of subsequent measures of crack size. For example, using Paris' model, $\theta = (\ln c, m)$ and $\lambda = -0.1471$.

Three measurement data points are arbitrarily chosen to represent the actual measures of crack sizes. Those data are used for updating using the posterior distribution in Eq. (20). Once a new measurement is available, MCMC simulations with the Metropolis–Hastings algorithm[27,28] are employed to draw samples from the posterior distribution. The fatigue crack growth prognostics can readily be evaluated using the resulting MCMC samples. At each updating, 25×10^4 samples are generated. For the purpose of illustration,

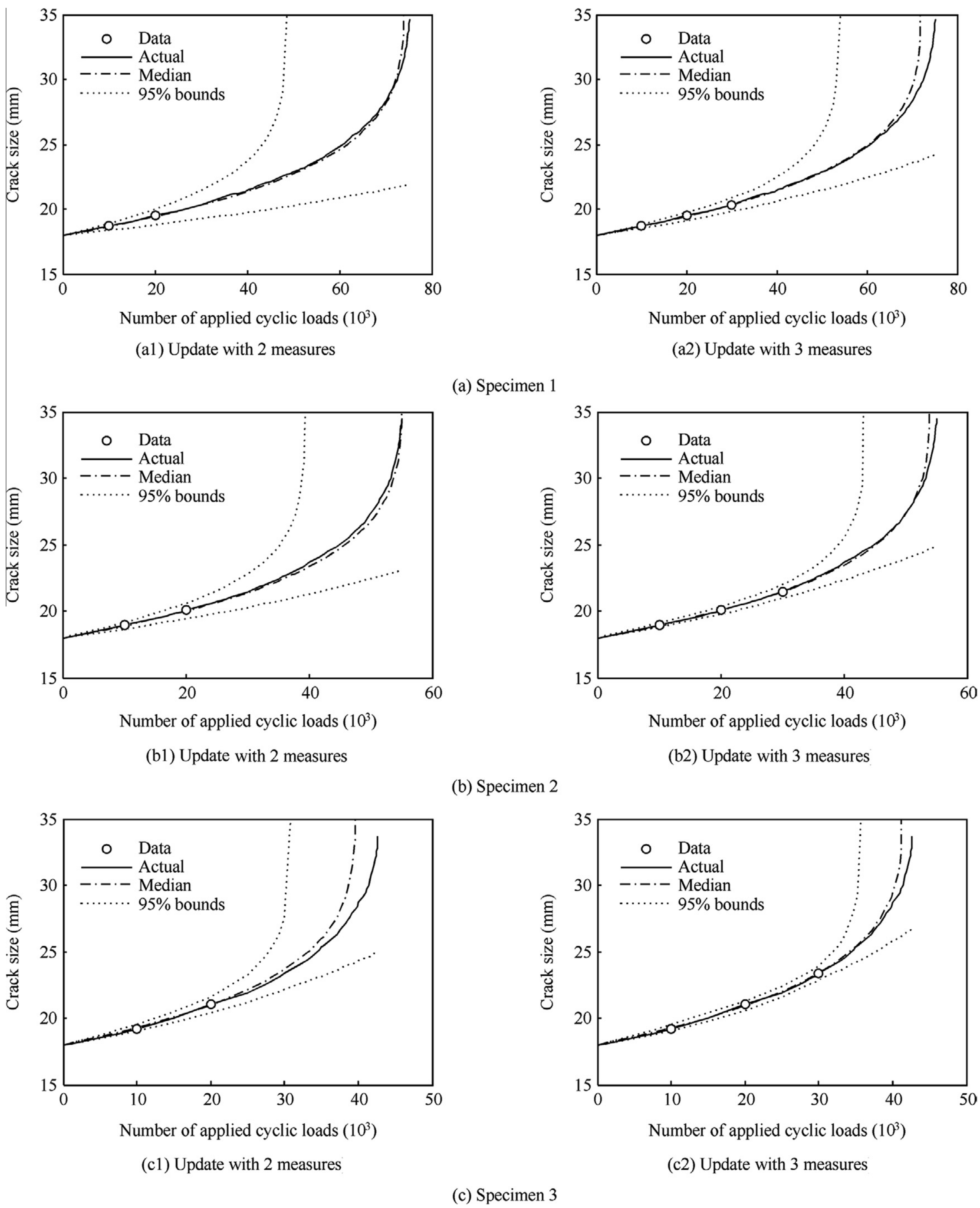

Fig. 3 Fatigue crack growth prognostics update using Forman's model.

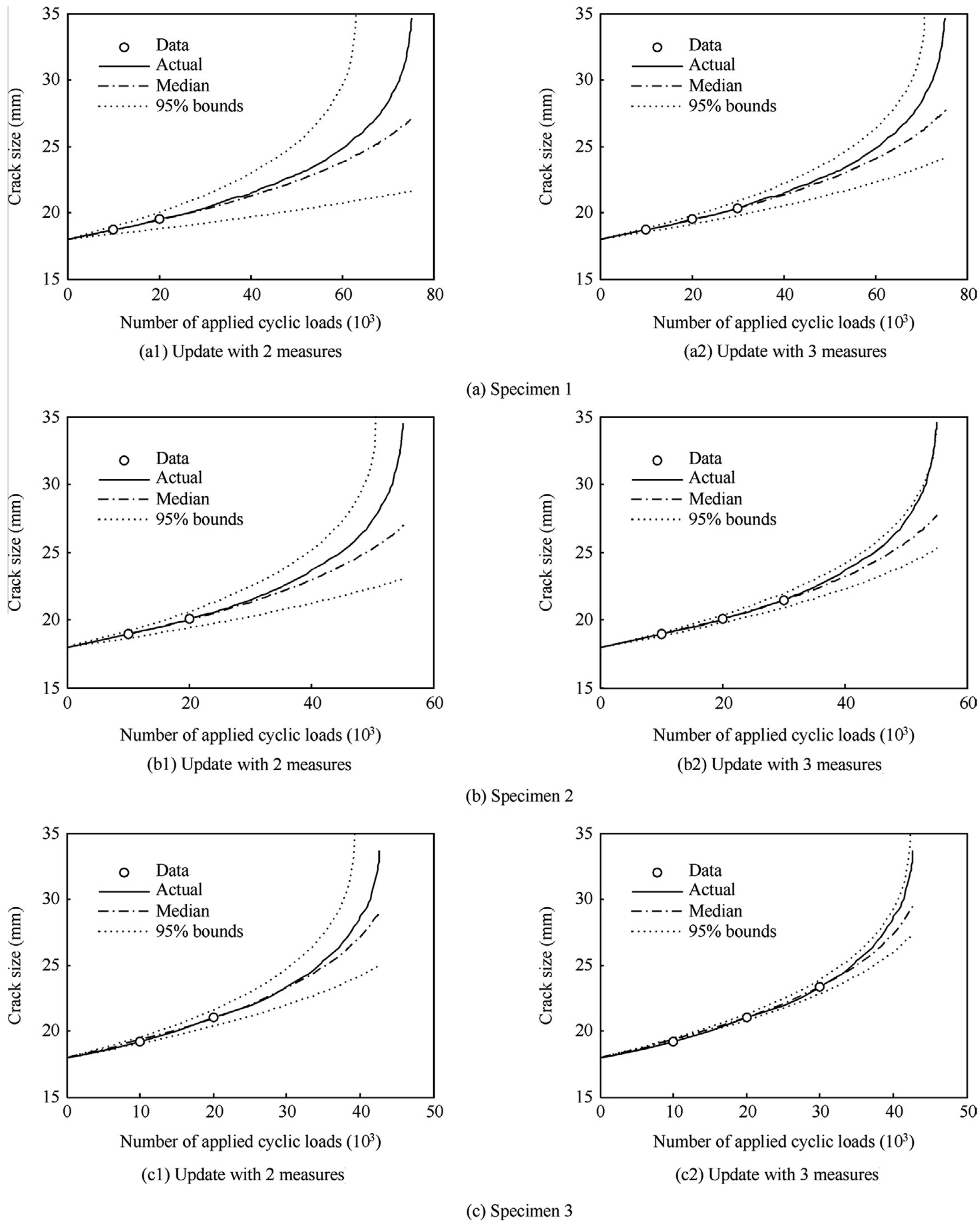

Fig. 4 Fatigue crack growth prognostics update using McEvily's model.

term σ_ε is explicitly set to 0.30 mm for general fatigue prognostic problems.[7] Alternatively, the actual value for σ_ε can be estimated using historical data or the calibration uncertainty of the measurement equipment. The results of crack growth prognostics update using Paris' model are presented in Fig. 2. Median predictions and 95% bounds predictions are shown. Although the three components have the same prior distribution, the final prognosis results for each of the components are different. Forman's model and McEvily's model produce results given in Figs. 3 and 4, respectively. As more and more measures are used for updating, the data become dominant and they will diminish the effect from the prior eventually. From a practical point of view, it is shown that reasonable prognostic results can be obtained using the prior based on partial information.

In realistic applications, the model is generally sophisticated and the prior distribution with an exponential model with no analytical solution might be slow for numerical evaluation of the posterior distribution. On the other hand, if physical justification can be made in the form of the parameter distribution, it would be convenient to transform the samples to a particular type of distribution. For example, in the application example shown above, it is appropriate to consider $\ln c$ as a normal variable and m as a truncated normal variable. MCMC samples after the first updating can be used to fit the distributions. In fact, the parameters in Paris' model are considered as normal variables in Ref.[29] Fitting MCMC samples into an analytical distribution may introduce additional uncertainties and in some cases it might be risky for prognostics and decision-making.

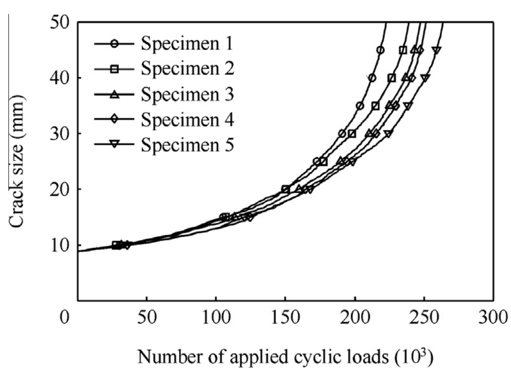

Fig. 5 Five crack growth curves of center-through cracked specimens (data source is from Ref.[30]).

4.3. Compact tension specimen testing data and prognostics

The center-through cracked specimen testing data for aluminum 2024-T3 materials provided by Ref.[30] are used. The dataset consists of 68 sample trajectories, each containing 164 measurement points. The entire specimen has the same geometry, i.e., an initial crack size of 9 mm, length $L = 558.8$ mm, width $w = 152.4$ mm and thickness $t = 2.54$ mm. The stress range during each test is constant $\Delta\sigma = 48.28$ MPa and the stress ratio is $R = 0.2$. The failure criterion is that the crack size equals 49.8 mm. Five specimens from the dataset are arbitrarily chosen to represent the target components and they are shown in Fig. 5.

Assume the crack size for the five specimens at 10^5 cycles are measured. The crack lengths are 14.662 mm, 14.380 mm, 13.862 mm, 12.983 mm, and 12.998 mm for Specimens 1–5, respectively. The mean value of the model output at 10^5 cycles is $\bar{a} = 13.777$ mm. Following the same procedure as shown in the compact tension specimen dataset, the values of Lagrange multiplier λ are obtained as -0.0316 mm, -0.0244 mm, and -0.0251 mm for Paris' model, Forman's model and McEvily's model, respectively. The detailed values for interpolating λ_i are shown in Table 2.

Four data points are chosen to represent the subsequent measurements to perform updating. For illustration purposes, only Paris' model results are presented here. Other models follow the same procedure as Paris' model. The results for fatigue crack growth associated with the five specimens are shown in Fig. 6. Although the five specimens share the same entropic prior distribution, the subsequent fatigue crack growth curves are quite different. As more measurement points are used to perform updating, the effect of prior gradually reduces. The 95% bounds of the crack growth curves also narrow, indicating the uncertainty in prognostics reduces. In addition, the median prediction of the crack growth curves also becomes closer to the actual crack growth curve as more data points are used for updating.

Table 2 Values of $\bar{a}_{\lambda_i}$ for different λ_i calculated using numerical quadrature.

Model	$\bar{a}_{\lambda_i}$					
	$\lambda_i = -0.040$	$\lambda_i = -0.038$	$\lambda_i = -0.036$	$\lambda_i = -0.034$	$\lambda_i = -0.032$	$\lambda_i = -0.030$
Paris	11.218	11.599	12.096	12.742	13.583	14.676
Forman	10.062	10.222	10.431	10.706	11.065	11.538
McEvily	10.185	10.358	10.584	10.880	11.267	11.775
Model	$\bar{a}_{\lambda_i}$					
	$\lambda_i = -0.028$	$\lambda_i = -0.026$	$\lambda_i = -0.024$	$\lambda_i = -0.022$	$\lambda_i = -0.020$	
Paris	16.093	17.921	20.263	23.238	26.976	
Forman	12.157	12.968	14.027	15.407	17.193	
McEvily	12.439	13.308	14.442	15.914	17.817	

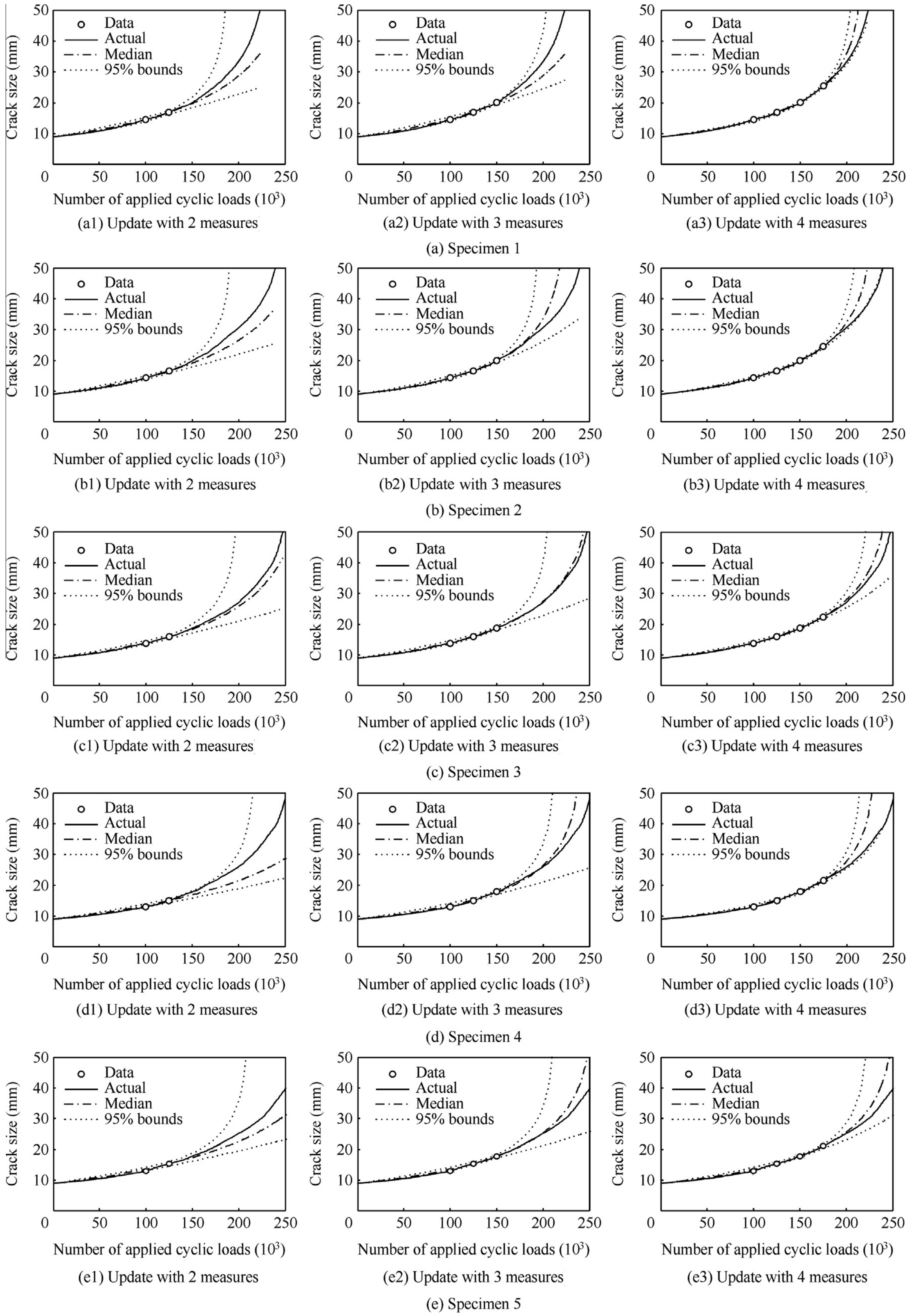

Fig. 6 Crack growth prognostics update using Paris' model.

5. Conclusions

The paper presents a method of fatigue crack damage prognostics using limited or partial information. The proposed method formulates the prior using the fatigue crack growth model and the first response measurement data as constraint. When only point measurement data are used in updating, the updating rule is equivalent to Bayes' rule. The overall method is demonstrated using fatigue problems with experimental data. The results suggest that the proposed procedure provides a feasible way of conducting prognostics using limited or partial information. Based on the current investigation, two conclusions are drawn.

(1) It is feasible to construct a prior distribution for probabilistic inference using limited or partial information obtained directly from the target components. This is achieved by using the mechanism model as the constraint function. The average value of the first response measures from the target components is treated as the mathematical expectation for the model prediction.
(2) Realistic prognostics may be performed using simple models with limited known information on model parameters. The key is to transform the limited information or data into the prior and perform continuous updating to reduce the uncertainty in prognostics. It should be noted that the proposed method is to provide a rational approach for probabilistic fatigue prognostics with limited or partial information. For regular cases where enough data are available for classical statistical analysis, conventional methods are preferred and the complicated numerical evaluations of the proposed method can be avoided. More efficient algorithms will be investigated in the future to reduce the computational demands of the proposed method.

Acknowledgements

Authors would like to thank editor and anonymous reviewers for their constructive comments.

References

1. Paris P, Erdogan F. A critical analysis of crack propagation. *J Basic Eng* 1963;**85**(4):528–34.
2. Forman RG, Kearney V, Engle R. Numerical analysis of crack propagation in cyclic-loaded structures. *J Fluids Eng* 1967;**89**(3):459–63.
3. McEvily A. On the quantitative analysis of fatigue crack propagation. *ASTM STP* 1983;**811**:283–312.
4. Janssen M, Zuidema J, Wanhill R. Fracture mechanics. 2nd. New York: Taylor & Francis Group; 2004. p. 1–129.
5. Frost N, Pook L, Denton K. A fracture mechanics analysis of fatigue crack growth data for various materials. *Eng Fract Mech* 1971;**3**(2):109–26.
6. Clark Jr WG, Hudak Jr SJ. Variability in fatigue crack growth rate testing. *ASTM J Test Eval* 1975;**3**(6):454–76.
7. Guan X, Jha R, Liu Y. Model selection, updating, and averaging for probabilistic fatigue damage prognosis. *Struct Saf* 2011;**33**(3):242–9.
8. Beck JL, Au SK. Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *J Eng Mech* 2002;**128**(4):380–91.
9. Guan X, He J, Jha R, Liu Y. An efficient analytical Bayesian method for reliability and system response updating based on Laplace and inverse first-order reliability computations. *Reliab Eng Syst Saf* 2012;**97**(1):1–13.
10. Mosegaard K, Tarantola A. Probabilistic approach to inverse problems. *Int Geophys Series* 2002;**81**(A):37–68.
11. Gregory P. Bayesian logical data analysis for the physical sciences: a comparative approach with mathematica® support. New York: Cambridge University Press; 2005. p. 41–69.
12. Jaynes ET. Marginalization and prior probabilities. In: Rosenkrantz RD, editor. *Papers on probability, statistics and statistical physics*. Amsterdam: Springer Netherlands; 1989. p. 340–70.
13. Carlin BP, Louis TA. Bayes and empirical Bayes methods for data analysis. *Stat Comput* 1997;**7**(2):153–4.
14. Jaynes ET. Prior probabilities. *IEEE Trans Syst Sci Cybernetics* 1968;**4**(3):227–41.
15. Jaynes ET. Probability theory: the logic of science. New York: Cambridge University Press; 2003. p. 133–74.
16. Yang R, Berger JO. Estimation of a covariance matrix using the reference prior. *Ann Stat* 1994;**22**(3):1195–211.
17. Yang Y. Invariance of the reference prior under reparametrization. *Test* 1995;**4**(1):83–94.
18. Ayers R. Information, entropy, and progress. New York: AIP Press; 1994. p. 1–301.
19. Jaynes ET. Information theory and statistical mechanics: II. *Phys Rev* 1957;**108**(2):171.
20. Abramov RV. An improved algorithm for the multidimensional moment-constrained maximum entropy problem. *J Comput Phys* 2007;**226**(1):621–44.
21. Abramov RV. The multidimensional maximum entropy moment problem: a review of numerical methods. *Commun Math Sci* 2010;**8**(2):377–92.
22. Kullback S, Leibler RA. On information and sufficiency. *Ann Math Stat* 1951;**22**(1):79–86.
23. Caticha A, Giffin A. Updating probabilities. The 26th international workshop on Bayesian inference and maximum entropy methods; 2006 July 8–13; Pairs, France. 2006. p. 31–42.
24. Wu W, Ni C. A study of stochastic fatigue crack growth modeling through experimental data. *Probab Eng Mech* 2003;**18**(2):107–18.
25. Wachsmuth G. On LICQ and the uniqueness of Lagrange multipliers. *Oper Res Lett* 2013;**41**(1):78–80.
26. Carpinteri A, Paggi M. Are the Paris' law parameters dependent on each other? *Fract Struct Integrity* 2007;**2**:10–6.
27. Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys* 1953;**21**(6):1087.
28. Hastings W. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970;**57**(1):97–109.
29. Kotulski ZA. On efficiency of identification of a stochastic crack propagation model based on Virkler experimental data. *Arch Mech* 1998;**50**(5):829–47.
30. Virkler D, Hillberry B, Goel P. The statistical nature of fatigue crack propagation. *J Eng Mater Technol* 1979;**101**(2):148–53.

Huang Min is an associate professor in School of Reliability and System Engineering at Beihang University. Her area of research includes process reliability and aero-engine reliability.

He Jingjing is an associate professor in School of Reliability and System Engineering at Beihang University. Her area of research includes structural health monitoring and fatigue damage diagnosis and prognosis.