Bayesian Knowledge Fusion in Prognostics and Health Management—A Case Study

Masoud Rabiei [a], Mohammad Modarres [a] and Ali Mohammad-Djafari [b]

[a] *Center for Risk and Reliability, Department of Mechanical Engineering, University of Maryland College Park, MD, 20742, USA*
masrab@umd.edu
modarres@umd.edu

[b] *Laboratoire des Signaux et Systèmes, UMR 8506 CNRS - SUPELEC - Univ Paris Sud 11*
Supelec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE
djafari@lss.supelec.fr

Abstract. In the past few years, a research effort has been in progress at University of Maryland to develop a Bayesian framework based on Physics of Failure (PoF) for risk assessment and fleet management of aging airframes. Despite significant achievements in modelling of crack growth behavior using fracture mechanics, it is still of great interest to find practical techniques for monitoring the crack growth instances using non-destructive inspection and to integrate such inspection results with the fracture mechanics models to improve the predictions. The ultimate goal of this effort is to develop an integrated probabilistic framework for utilizing all of the available information to come up with enhanced (less uncertain) predictions for structural health of the aircraft in future missions. Such information includes material level fatigue models and test data, health monitoring measurements and inspection field data. In this paper, a case study of using Bayesian fusion technique for integrating information from multiple sources in a structural health management problem is presented.

Keywords: Prognostics and Health Management (PHM), Bayesian Knowledge Fusion, Crack Growth Modeling, Non-destructive Inspection (NDI).

INTRODUCTION

Decision making about the remaining useful life of aging airframes is very complicated and should be based upon all information available. Aerospace structures have been traditionally inspected using on-ground nondestructive techniques such as Eddy current and visual inspection. New developments in recent years are heading towards continuous on-line monitoring and damage detection in the structure (Boller, 2001; Fang and Berkovits, 1993). These new technologies, along with high fidelity crack growth simulation models, can be used to provide better estimates of the crack size distribution at given hotspots. Such estimations always suffer from both reducible and non-reducible uncertainties. Such uncertainties play an important role in the risk and reliability calculations and should be characterized properly using probability and statistical techniques.

The authors have previously proposed a probabilistic model (Wang et al., 2009) to assess the reliability of aging airframes by predicting the probability that a fatigue-induced crack will reach an unacceptable length after specified future flight hours. They have also shown in a different paper (Wang et al., 2008) that predictions based on empirical models alone are not sufficient to assure the safety of a mission. The next step towards enhancing the quality of risk predictions is to use non-destructive inspection (NDI) to monitor the crack growth in the structure and to use this additional information to supplement the uncertain model predictions. (Rabiei et al., 2010) have proposed an approach for estimating the crack size using real-time acoustic emission (AE) monitoring. Having multiple independent predictions for the crack size poses a new challenge in fleet management: how could this information get reduced to a unified risk measure that can be used for fleet management purposes?

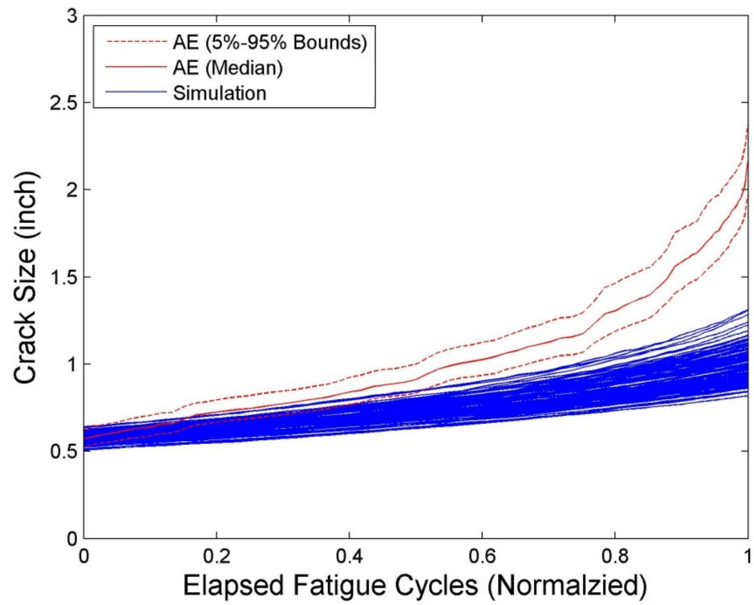

Figure 1 Crack size prediction based on Simulation and AE NDI results

Figure 1 shows the crack size estimation based on simulation as well as AE NDI monitoring in a laboratory experiment. It is clear from the figure that the AE technique is predicting a higher crack growth rate which results in different crack size distribution predictions based on the two techniques. A formal approach for combining these predictions is necessary before this information can be used for fleet management purposes.

To address this issue, we propose a Bayesian framework for combining crack growth information obtained from independent sources. This problem could be tackled in two different ways: One way is to obtain independent crack size estimates from each of the sources and then try to combine them at any given time. This approach may provide a better instantaneous crack size estimate but does not result in any model improvement. The other approach is to develop a model based on the crack size estimates from the simulation and then update the model parameters whenever new evidence from the NDI measurements becomes available. This approach results in an updated model which can be reused for improved crack size predictions in future applications.

In the remainder of this paper, we first describe the necessity of knowledge fusion in PHM applications and then describe the specific problem of fatigue crack growth prediction in aging airframes. Next we describe the structure of the developed statistical model and then present the Bayesian fusion formulation.

BAYESIAN KNOWLEGE FUSION

The information necessary for developing a structural health diagnostic and prognostic model is obtained from various sources. One can think of this as having a number of ‘experts’ that each provides relevant information about the problem at hand.

In this section, Bayesian fusion technique is presented as a flexible formal framework that can be used to update model parameters in presence of additional information. In this approach the information provided by 'experts' will be used to update the *a priori* knowledge about the quantity of interest. The process of combining information can become challenging specially when dealing with diverse 'types' of information from different experts or when the experts are not equally credible.

The additional information could be provided from various sources and could be of different formats. For example in structural health monitoring and prognostics problems, the important quantity of interest is the instantaneous size of fatigue crack at a given hotspot. Crack size information may be provided from the following sources:

1. Probabilistic crack growth model (Format: PDF of crack size)
2. In-situ (on-board) NDI findings (Format: uncertain crack size estimate that can be represented probabilistically)
3. On-ground NDI inspections (Format: individual crack observations with quantifiable uncertainty)

For fleet management purposes, it is absolutely critical to be able to formally combine these independent sources of information to achieve more accurate and less uncertain predictions about the structural health of the system.

The outcome of the fusion process is a model with updated parameters that can be reused for predictions under similar circumstances. In fact, one of the important advantages of Bayesian approach is the sequential updating property (Eq. (1)) which allows us to constantly update the model parameters with additional information as it becomes available.

$$p(\underline{\lambda} \mid E_1, E_2, ..., E_N) = \frac{p(E_N \mid \underline{\lambda}, E_1, ..., E_{N-1}) ... p(E_1 \mid \underline{\lambda}) P(\underline{\lambda})}{p(E_1, E_2, ..., E_N)} \tag{1}$$

where $\underline{\lambda}$ is the vector of model parameters and E_i*'s* are the evidence available at the i^{th} updating stage.

Problem Definition

Let us assume that at any given time, two independent crack size estimations are available at a given hot spot. The first estimate is obtained from a fracture mechanics-based crack growth simulation (Anderson, 1994) while the second one is an estimate obtained from a non-destructive inspection technique such as AE monitoring (Mix, 2005).

Both methods suffer from various sources of uncertainty; therefore, we use probability density functions to represent their crack size estimates.

Our goal is to build a statistical model for crack size as a function of time based on the available crack size estimates. In reality, the crack growth simulation results are more readily available; so the model parameters are first estimated from this source of information and then, as the crack size estimates from on-board inspection becomes available, the model parameters will be updated to reflect the new observations.

Let D_{SM} and D_{AE} denote the simulation and AE data, respectively. D_{SM} is a set of random samples of the crack size (denoted by Y) obtained via crack growth simulation at various time instances x_j's:

$$D_{SM} = \{(x_j, y_{ij}) \mid i = 1, \cdots, N_s, j = 1, \cdots, N_x\} \tag{2}$$

where N_s is the number of samples of Y and N_x is the number of time instances.

D_{AE} is a set of *uncertain* observations of crack size, obtained from AE measurements, that are represented by random variables $E^{(k)}$'s at various time instances x_k's:

$$D_{AE} = \left\{ D_{AE}^{(k)} \mid k = 1, \cdots, N_e \right\}$$
$$D_{AE}^{(k)} = \left\{ (x_k, E^{(k)}) \mid E^{(k)} \sim f_{E^{(k)}}\left(e^{(k)} \mid \underline{\delta}; x_k\right) \right\} \tag{3}$$

where N_e is the total number of observations and $\underline{\delta}$ is a known vector of parameters for the probability distribution function $f_{E^{(k)}}(e^{(k)})$.

Statistical Model Development

Hierarchical models (also called *multilevel models*) are used whenever information is available on several different levels in a problem. The hierarchy of these models can better reflect the complexity of a given problem than a simple model with a non-hierarchical cloud of parameters. In this approach, to capture the complicated structure of data and information, the prior is structured using a series of conditional distributions called *hierarchical stages* of the prior distribution (Gelman et al., 2003).

Given evidence E and parameters $\underline{\lambda}$, a simple Bayesian model consists of *a priori* probability $p(\underline{\lambda})$ and the likelihood $p(E|\underline{\lambda})$ that are used to compute a posterior probability $p(\underline{\lambda}|E) \propto p(E|\underline{\lambda})\, p(\underline{\lambda})$.

A Bayesian hierarchical model is defined when the prior distribution on $\underline{\lambda}$ itself depends on other parameters $\underline{\phi}$ that are not present in the definition of likelihood. Therefore, the prior $p(\underline{\lambda})$ must be replaced by a prior $p(\underline{\lambda}|\underline{\phi})$, and a prior $p(\underline{\phi}|\underline{\psi})$ with parameters $\underline{\psi}$ on the newly introduced parameters $\underline{\phi}$ is required. This will result in the posterior probability below:

$$\begin{aligned} p(\underline{\lambda} \mid E, \underline{\phi}, \underline{\psi}) &\propto p(E \mid \underline{\lambda}, \underline{\phi}, \underline{\psi})\, p(\underline{\lambda} \mid \underline{\phi}, \underline{\psi})\, p(\underline{\phi} \mid \underline{\psi}) \\ &\propto p(E \mid \underline{\lambda})\, p(\underline{\lambda} \mid \underline{\phi})\, p(\underline{\phi} \mid \underline{\psi}) \end{aligned} \tag{4}$$

The latter simplification holds because the likelihood does not usually directly depend on $\underline{\phi}, \underline{\psi}$; the effect of $\underline{\phi}, \underline{\psi}$ on the likelihood is only through $\underline{\lambda}$. This process may be repeated for more hierarchical stages but in this case, $\underline{\psi}$ is considered to be a vector of constant parameters.

The prior distribution in this model formulation is characterized by a two-level hierarchy: $p(\underline{\lambda}|\underline{\phi})$ is the first level, and $p(\underline{\phi}|\underline{\psi})$ is the second level. Prior distributions of the higher levels are called *hyperpriors* and the corresponding parameters are known as *hyperparameters*. In Eq. (4), $p(\underline{\phi}|\underline{\psi})$ is the hyperprior and $\underline{\psi}$ are the hyperparameters of the prior parameters $\underline{\phi}$.

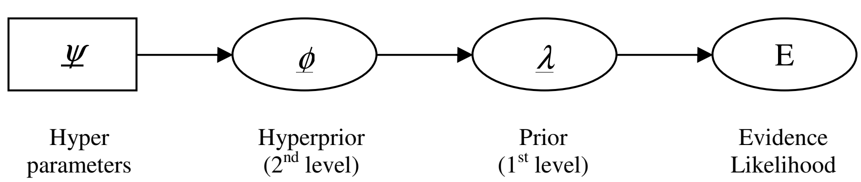

Figure 2 Graphical representation of a two-stage Bayesian hierarchical model. Square nodes refer to constant parameters; oval nodes refer to stochastic components of the model.

A useful tool for representing hierarchical Bayesian models is the *Directed Acyclic Graph* (DAG) (Figure 2). In this diagram, each prior is represented as a separate node pointing to the nodes that depend on it. Oval nodes represent stochastic components of the model, whereas square nodes refer to constant parameters (Ntzoufras, I., 2009).

The first level in the hierarchy is defined by the likelihood relationship. The definition of the likelihood in our model is presented in Eq. (5) where Y is assumed to have a Lognormal distribution defined as follows:

$$Y \sim \text{Lognormal}(\mu, \tau)$$
$$\mu = \log(\theta_1 X + \theta_2) \tag{5}$$

where Y is the inverse of the crack size and X is the number of elapsed fatigue cycles corresponding to Y. We assume in this model that the median of Y is a linear function of X.

Figure 3 shows the hierarchy of the nodes in the model. We already discussed the first level of the hierarchy which is the likelihood relationship. The next level in the hierarchy is the prior of the model parameters: $\underline{\lambda} = (\theta_1, \theta_2, \tau)$. The prior should reflect the correlation between parameters. θ_1 and θ_2 are the slope and intercept of a line and are expected to be correlated with each other but they are both uncorrelated with τ. The precision parameter τ follows a Gamma distribution with shape parameter α and inverse scale parameter β.

$$\tau \sim \text{Gamma}(\alpha, \beta) \tag{6}$$

Gamma distribution is a popular choice for modeling precision of Normal distributions; it represents positive random variables and for large shape parameters, it resembles the bell shape of a Normal distribution.

Correlated random variables θ_1 and θ_2 form the random vector $\underline{\theta}$. A multivariate Normal distribution is used to represent $\underline{\theta}$ as follows:

$$\underline{\theta} \sim \text{Normal}(M_\theta, \Sigma_\theta) \tag{7}$$

where M_θ is the vector of the means and Σ_θ is the precision matrix of the multivariate Normal distribution.

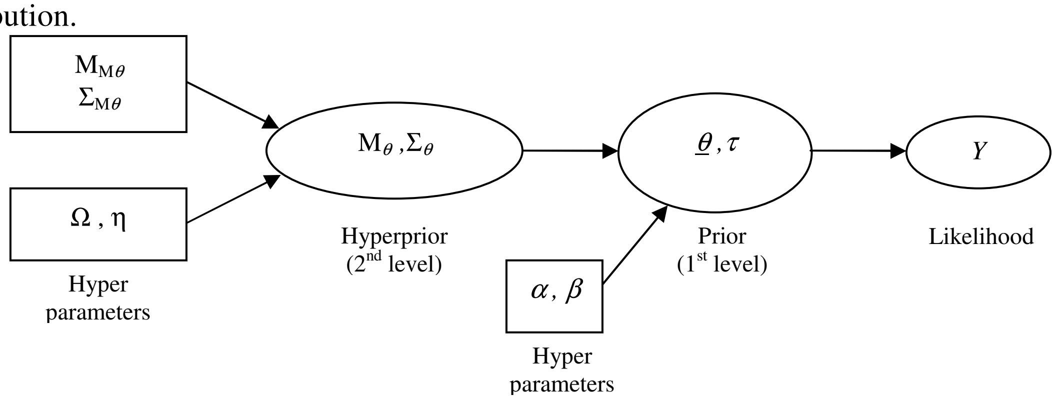

Figure 3 Hierarchy and dependence of the nodes in the fusion model

Most Bayesian models in engineering applications would stop at this level in the hierarchy and constant values are assigned to the parameters that define the prior distributions. But in this fusion problem, we need an extra level in the hierarchy that will allow us to better represent the available information.

The next level in the hierarchy is defining the hyperprior. To do that we choose a multivariate Normal distribution to define a prior for M_θ and use the Wishart distribution to define a prior for the precision matrix Σ_θ.

$$M_\theta \sim \text{Normal}(M_{M_\theta}, \Sigma_{M_\theta})$$
$$\Sigma_\theta \sim \text{Wishart}(\Omega, \eta) \quad (8)$$

where M_{M_θ} is the vector of the means and Σ_{M_θ} is the precision matrix of the multivariate Normal distribution of M_θ.

The Wishart distribution has two parameters: Ω which is the scale matrix and η which is the parameter indicating the degrees of freedom.

The Wishart distribution is a multivariate generalization of the Gamma distribution (Gelman et al., 2003). It can be used to describe positive-definite random matrices (matrix-valued random variables). The Wishart is widely used as the prior distribution for the covariance (or precision) matrix in multivariate Normal modeling. The integral of the Wishart distribution is finite only if η is greater than or equal to the rank of the random matrix. A non-informative prior can be obtained as η takes smaller values. In a bivariate Normal distribution the rank of covariance matrix is 2, therefore $\eta \geq 2$ should hold. Values of η closer to 2 result in a more diffused (non-informative) prior distribution.

The parameters at different levels of the model can be summarized as follows:

$\underline{\lambda} = \{\underline{\theta}, \tau\}$ (First level prior)

$\underline{\phi} = \{ M_\theta, \Sigma_\theta \}$ (Second level prior)

$\underline{\Psi} = \{M_{M\theta}, \Sigma_{M\theta}, \Omega, \eta, \alpha, \beta\}$ (Vector of hyperparameters)

So far the structure of the Bayesian hierarchical model is described. The next step is to use the available data to infer the model parameters.

Bayesian Inference formulation

In the previous section we introduced the data as well as our modeling approach. In this section we describe the application of Bayesian inference technique to calculate the model parameters given the available data. More specifically, the objective is to infer $\underline{\lambda}$ and $\underline{\phi}$ given the data D_{SM} and D_{AE}, i.e. compute $p(\underline{\lambda}, \underline{\phi} | D_{SM}, D_{AE})$. D_{SM} which is the simulation data is easier to obtain and hence is used first to infer the model parameters. Later as the AE NDI system monitors the crack and provides crack size estimations, the model parameters will be updated to reflect the additional information. The challenge with the AE data is that it is not in a traditional crack observations format; instead, the AE system provides 'uncertain' estimates of crack size that are represented by random variables $E^{(k)}$. The other factor that complicates the calculations is the fact that the AE crack evidence at various time steps are correlated. Here we assume that $E^{(k)}$ is conditionally independent of $E^{(k-2)}, \ldots, E^{(1)}$, i.e.:

$$p\left(E^{(k)} | E^{(k-1)}, E^{(k-2)}, \cdots, E^{(1)}\right) = p\left(E^{(k)} | E^{(k-1)}\right) \quad (9)$$

This is a reasonable assumption because $E^{(k)}$'s form a Markov chain in which the evidence at each stage is only dependent on the evidence at previous stage.

$$p(\underline{\lambda},\underline{\phi} \mid D_{SM}, D_{AE}) = p(\underline{\lambda},\underline{\phi} \mid D_{SM}, D_{AE}^{(1)},\ldots,D_{AE}^{(N_e)}) = p(\underline{\lambda},\underline{\phi} \mid D_{SM}, E^{(1)},\ldots,E^{(N_e)})$$
$$= \int_{e^{(1)}} f_{E^{(1)}}(e^{(1)}) p(\underline{\lambda},\underline{\phi} \mid D_{SM}, e^{(1)}, E^{(2)},\ldots,E^{(N_e)}) de^{(1)}$$
$$= \int_{e^{(1)}} \int_{e^{(2)}} f_{E^{(1)}}(e^{(1)}).f_{E^{(2)}|E^{(1)}}(e^{(2)} \mid e^{(1)}).p(\underline{\lambda},\underline{\phi} \mid D_{SM}, e^{(1)}, e^{(2)}, E^{(3)},\ldots,E^{(N_e)}).de^{(2)}.de^{(1)}$$
$$\vdots$$
$$= \int_{e^{(1)},\ldots,e^{(N_e)}} f_{E^{(1)}}(e^{(1)}).\prod_{k=2}^{N_e}\left(f_{E^{(k)}|E^{(k-1)}}(e^{(k)} \mid e^{(k-1)}).p(\underline{\lambda},\underline{\phi} \mid D_{SM}, e^{(1)}, e^{(2)},\ldots,e^{(N_e)}).de^{(k)}\right) \tag{10}$$

The correlation between AE data at subsequent time instances is captured in the conditional PDF terms that appear in the above equations. Now using the Bayes' rule,

$$p(\underline{\lambda},\underline{\phi} \mid D_{SM}, e^{(1)}, e^{(2)},\ldots,e^{(N_e)}) \propto p(D_{SM}, e^{(1)}, e^{(2)},\ldots,e^{(N_e)} \mid \underline{\lambda},\underline{\phi}) p(\underline{\lambda},\underline{\phi})$$
$$= p(D_{SM} \mid \underline{\lambda},\underline{\phi}).p(e^{(1)} \mid \underline{\lambda},\underline{\phi}).\ldots.p(e^{(N_e)} \mid \underline{\lambda},\underline{\phi}).p(\underline{\lambda},\underline{\phi}) \tag{11}$$

The latter simplification is based on the independence of D_{SM} and D_{AE} and conditional independence of *$E^{(k)}$'s* given the model parameters, i.e.

$$p\left(E^{(k_1)} | E^{(k_2)}, \underline{\lambda}, \underline{\varphi}\right) = p\left(E^{(k_1)} | \underline{\lambda}, \underline{\varphi}\right), \forall k_1, k_2 = 1,\ldots,N_e \tag{12}$$

Based on the fact that the hierarchical model is parameterized so that $\underline{\phi}$ does not appear in the likelihood functions and by using the rules of conditional probability one can write:

$$p(D_{SM} \mid \underline{\lambda},\underline{\phi}).p(e^{(1)} \mid \underline{\lambda},\underline{\phi}).\ldots.p(e^{(N_e)} \mid \underline{\lambda},\underline{\phi}).p(\underline{\lambda},\underline{\phi}) =$$
$$p(D_{SM} \mid \underline{\lambda}).p(e^{(1)} \mid \underline{\lambda}).\ldots.p(e^{(N_e)} \mid \underline{\lambda}).p(\underline{\lambda} \mid \underline{\phi}).p(\underline{\phi};\underline{\psi}) \tag{13}$$

At this level, each of the likelihood terms, $p(. \mid \underline{\lambda})$, can be easily calculated as follows,

$$p(D_{SM} \mid \underline{\lambda}) = \prod_{j=1}^{Nx}\prod_{i=1}^{Ns} f_Y(y_{ij} \mid \underline{\lambda}; x_j) \tag{14}$$

$$p(e^{(k)} \mid \underline{\lambda}) = f_{Y_k}(e^{(k)} \mid \underline{\lambda}; x_k) \tag{15}$$

The integrals in Eq. (10) should be solved numerically but this could become computationally tedious especially for large datasets. The other option is to use WinBUGS software package (Cowles, 2004) for MCMC simulation to find the posterior distribution of parameters but this is also complicated due to the correlation between the AE data. Efforts are still under way to find a computationally efficient approach to solve this problem.

CONCLUSION

A case study of using Bayesian fusion technique for integrating information from multiple sources in a structural health management problem was presented in this paper. Specifically, the goal was to develop a statistical model for crack size as a function of elapsed fatigue cycles using data from two independent sources of information, i.e. fracture mechanics simulation and AE

NDI monitoring. The simulation data was first used to find the model parameters and then, as crack size estimates from AE became available, the model parameters were updated in light of the new evidence. The mathematical formulation of the problem as well as the setup of the Bayesian inference solution was given. The presented approach addresses the uncertainty in the evidence and takes into account the correlation between AE observations. The resulting equations should be solved numerically. Efforts are under way to use MCMC simulation technique to provide an efficient computational solution to this problem.

REFERENCES

(Anderson, 1994) Anderson, T. L. *Fracture Mechanics: Fundamentals and Applications, Second Edition*. 2nd ed. CRC, December 16, 1994.

(Boller, 2001) Boller, C. Ways and options for aircraft structural health management. *Smart materials and structures* 10, no. 3: 432-440, 2001.

(Cowles, 2004) Cowles, M.K. Review of WinBUGS 1.4. *The American Statistician*, 58(4), 330–336, 2004.

(Fang and Berkovits, 1993) Fang, D., and A. Berkovits. Fatigue damage mechanisms on the basis of acoustic emission measurements. In *Novel experimental techniques in fracture mechanics: presented at the 1993 ASME Winter Annual Meeting New Orleans, Louisiana November 28-December 3, 1993*, 219. American Society of Mechanical Engineers, 1993.

(Gelman et al., 2003) Gelman, A. et al., *Bayesian Data Analysis, Second Edition* 2nd ed., Chapman & Hall, 2003.

(Mix, 2005) Mix, P. E. *Introduction to nondestructive testing: a training guide*. Wiley-Interscience, 2005.

(Ntzoufras, I., 2009) Ntzoufras, I. *Bayesian Modeling Using WinBUGS* 1st ed., Wiley, 2009.

(Rabiei et al., 2010) M. Rabiei, M. Modarres, and P. Hoffman. Towards Real-Time Quantification of Fatigue Damage in Metallic Structures. In *Aircraft Airworthiness & Sustainment 2010*. Austin TX, May, 2010.

(Wang et al., 2008) Wang, X., M. Rabiei, M. Modarres, and P. Hoffman. A probability-based individual aircraft tracking approach for airframe integrity. In *Aging Aircraft 2008*. Phoenix AZ, April, 2008.

(Wang et al., 2009) Wang, X., M. Rabiei, J. Hurtado, M. Modarres, and P. Hoffman. A probabilistic-based airframe integrity management model. *Reliability Engineering & System Safety* 94, no. 5 (May): 932-941, 2009.