Dynamic Bayesian Network for Aircraft Wing Health Monitoring Digital Twin

Chenzhao Li* and Sankaran Mahadevan†
Vanderbilt University, Nashville, Tennessee 37235
and
You Ling,‡ Sergio Choze,§ and Liping Wang¶
GE Global Research Center, Niskayuna, New York 12309

DOI: 10.2514/1.J055201

Current airframe health monitoring generally relies on deterministic physics models and ground inspections. This paper uses the concept of a dynamic Bayesian network to build a versatile probabilistic model for diagnosis and prognosis in order to realize the digital twin vision, and it illustrates the proposed method by an aircraft wing fatigue crack growth example. The dynamic Bayesian network integrates physics models and various aleatory (random) and epistemic (lack of knowledge) uncertainty sources in crack growth prediction. In diagnosis, the dynamic Bayesian network is used to track the evolution of the time-dependent variables and calibrate the time-independent variables; in prognosis, the dynamic Bayesian network is used for probabilistic prediction of crack growth in the future. This paper also proposes a modification to the dynamic Bayesian network structure, which does not affect the diagnosis results but reduces the time cost significantly by avoiding Bayesian updating with load data. By using a particle filter as the Bayesian inference algorithm for the dynamic Bayesian network, the proposed approach handles both discrete and continuous variables of various distribution types, as well as nonlinear relationships between nodes. Challenges in implementing the particle filter in the dynamic Bayesian network, where 1) both dynamic and static nodes exist and 2) a state variable may have parent nodes across two adjacent networks, are also resolved.

Nomenclature

a	=	crack length after current time step
a_{obs}	=	crack length observation
a^0	=	crack length before current time step
B	=	bolt looseness
F_p	=	shape factor in the plastic zone
M	=	elastic/plastic zone
P	=	load
P_{obs}	=	load observation
S_i	=	first-order sensitivity index
S_i^T	=	total effect sensitivity index
t	=	time step/cycle
$\boldsymbol{X}_t$	=	state variable in time step t
Y_A	=	anchor point position
$\boldsymbol{Z}_t$	=	measurement variable in time step t
Δa	=	crack growth in current time step
ΔK	=	stress intensity factor range
ΔS	=	stress range
$\boldsymbol{\theta}$	=	geometric and material properties
ω	=	weight term in particle filter

I. Introduction

IN DECIDING whether an aircraft is capable of safely performing an upcoming mission, a structural health monitoring (SHM) system is desired to provide the decision maker with the information on damage state of the aircraft, such as the crack length on the wing or the reliability of a replaceable unit. Information based on fleet statistics is not useful in assessing the health and capability of a particular aircraft because the damage state varies from aircraft to aircraft due to the variability in manufacturing, material properties, mission history, pilot variability, etc. The data collected in [1] revealed that, at the same operational hours, some aircraft had twice the fatigue damage rate compared to other aircraft. In summary, a SHM system tailored to each individual aircraft is desirable.

One example of an individualized SHM system is the individual aircraft tracking (IAT) program [2] to track the potential fatigue damage in the major airframe structural components such as the wing. A typical IAT program for the F-16 [3] uses the recorded load history to predict the crack growth and estimate the crack severity index (CSI); then, a comparison between the resultant CSI and a baseline condition will classify the aircraft health into three damage severity levels. The U.S. Air Force has been investigating the extension of IAT to realize the digital twin concept [4] in order to fuse multiple heterogeneous sources of information from models and data to support proactive fleet sustainment decisions. The current IAT system mainly focuses on the variation of load history; other uncertainty sources such as the epistemic uncertainty regarding the true values of geometric or material properties are not considered. (Epistemic uncertainty is caused by lack of knowledge and is reducible, whereas aleatory uncertainty is an irreducible natural variability.) A more comprehensive IAT program integrating various uncertainty sources in crack growth prediction is desirable [5] in order to avoid over- or underestimating the damage prognosis and achieve a balanced decision making by considering safety, performance, and budget.

Therefore, this paper aims to develop a powerful approach for building a probabilistic and prognostic individual aircraft tracking (P2IAT) model. The parameters of the P2IAT model depend on the physics models and collected data that are related to the system of interest [6]. Without loss of generality, this paper takes the crack growth on the leading edge of an aircraft wing as an example, as shown in Fig. 1; but, the underlying concepts can be extended to other airframe structural components or the entire airframe. As explained earlier, this P2IAT model is supposed to integrate various uncertainty sources over the entire life of aircraft wing. In addition, the P2IAT is also desired to achieve the following objectives:

1) Integrate heterogeneous information including test data, mathematical models, expert opinions, etc.

Received 16 March 2016; revision received 18 July 2016; accepted for publication 3 August 2016; published online 13 January 2017.

*Department of Civil and Environmental Engineering.
†Department of Civil and Environmental Engineering; sankaran.mahadevan@vanderbilt.edu (Corresponding Author).
‡Mechanical Engineer; You.Ling@ge.com.
§Senior Engineer; sbut001@gmail.com.
¶Manager of Probabilistics Lab; wangli@ge.com.

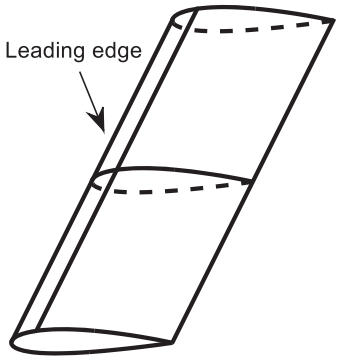


Fig. 1 Aircraft wing and its leading edge.

2) Fly virtually through the same load history as the actual aircraft wing.

3) Reduce the uncertainty in model parameters and track the time-dependent system states using measurement data, i.e., diagnosis.

4) Predict the evolution of damage states if no data are available, i.e., prognosis. An introduction to diagnosis and prognosis can be found in [7].

The concept of Bayesian networks (BNs) is a promising approach to integrate various uncertainty sources and heterogeneous information. In a Bayesian network, random variables are denoted by nodes (vertices) and their dependence relationships are denoted by directed edges; thus, the Bayesian network constitutes a directed acyclic graph to represent the joint distribution of a set of random variables. Regarding various uncertainty sources, the Bayesian network allows different types of random variables, including discrete and continuous variables of different distribution types. Regarding heterogeneous information, the Bayesian network is able to incorporate operational data, laboratory data, reliability data, expert opinion, and mathematical models (physics-based as well as empirical) [8].

The BN explained previously refers to a "static" Bayesian network for a time-independent system. To track a time-dependent system for which the states evolve over time, the Bayesian network is extended to a dynamic Bayesian network (DBN), which can be considered as a series of static BNs (one for each time step), with additional edges connecting the variables in adjacent time steps. Based on the Markov assumption, the states of the current BN depend only on the BN at the previous time step, and this dependence is generally assumed to be independent of the time instant [9]. The ability to track system evolution over time makes the DBN a suitable methodology to build the P2IAT model for diagnosis and prognosis of the aircraft wing.

When data of any child node are obtained, the Bayesian network is updated by Bayesian inference; thus, the uncertainty in the state variables can be reduced. Bayesian inference algorithms for the DBN have been developed in the literature, including the Kalman filter [10], extended Kalman filter [10,11], unscented Kalman filter [12–14], and particle filter [15,16]. The Kalman filter gives exact and analytical updating results [10] for a linear Gaussian DBN, which means the following:

1) The state function and the measurement function are both linear.

2) State variables have a joint Gaussian distribution.

3) All the noise terms are assumed to be independent zero mean Gaussian variables.

If the state function and/or the measurement function are nonlinear, the extended Kalman filter linearizes these functions to the first order, and it gives analytical updating results. The extended Kalman filter requires analytical derivatives to construct the Jacobian matrix, which is not trivial in most applications [12]. In addition, the linearization to the first order may bring high errors in the case of a highly nonlinear system. Detailed discussion on the inaccuracy of the extended Kalman filter can be found in section 2.5 of [17].

Another method to handle the nonlinear relationships in the DBN is the unscented Kalman filter. Both the Kalman filter and extended Kalman filter are purely analytical. In contrast, the unscented Kalman filter uses the method of unscented transform to select several sample points, and it propagates them through the nonlinear functions. The propagation is used to derive analytical updating results with accuracy to the third order, and the computation of the Jacobian matrix is not required [18]. However, the unscented Kalman filter can encounter ill-conditioning problems in the covariance matrix [18].

Although the extended Kalman filter and unscented Kalman filter provide solutions to the nonlinear DBN, they still assume that all the state variables are Gaussian. Due to this limitation, this paper does not select them to solve the problem of interest, which includes discrete variables. A new development on the Kalman filter is to combine the unscented Kalman filter with a particle filter (named the unscented particle filter) [19]. However, this method cannot handle discrete variables either.

The main motivation of this paper is to develop a diagnosis- and prognosis-framework-based DBN that considers 1) both discrete and continuous variables; 2) different distribution types; and 3) linear/nonlinear functional relationships. In contrast to the class of the Kalman filter, the particle filter (PF) is a sampling-based algorithm, where a particle is sampled from the joint distribution of the BN at one time step. The PF is a generic algorithm and fulfills the aforementioned requirements [18,20–22], thus, this paper chooses the PF as the Bayesian inference algorithm for the DBN. Two challenges in the PF are 1) degeneracy, meaning that all but one particle have negligible weights after a few iterations; and 2) sample impoverishment, meaning the loss of sample diversity. These two problems will be discussed and solved in Sec. II.B.

In the rest of the paper, Sec. II discusses the details to implement the particle filter to the DBN. The methods established in Sec. II are applied in Secs. III and IV. Section III analyzes the uncertainty sources in the fatigue crack growth on an aircraft wing and incorporates them into a DBN. Section IV computes and analyzes the results of the diagnosis and prognosis of the aircraft wing.

II. Diagnosis and Prognosis in the DBN

As explained in Sec. I, this paper selected the particle filter as the Bayesian inference algorithm. Here, a particle is a sample from the joint distribution of the BN at one time step. Similar to the Kalman filter, the particle filter at time step t also include two parts: 1) forward propagation, i.e., generating the particles at time step t based on the particles at time step $t-1$ and the conditional probability distribution (CPDs) between adjacent networks; and 2) backward inference, i.e., updating the joint distribution of the network.

A time step of purely forward propagation is defined as a "prognosis step" in this paper; and a time step with both forward propagation and backward inference is defined as a "diagnosis step." The diagnosis step happens if, and only if, any child node is observed. The prognosis step happens in two cases:

1) No observation is available.

2) All the observations are for the root nodes; thus, the samples of other nodes can be obtained by uncertainty propagation with these root nodes fixed at their observations.

In prognosis step, the backward inference part of the particle filter is skipped.

In this section, Sec. II.A gives a brief introduction to particle filter. Section II.B discusses the solution to the degeneracy and sample impoverishment problems in the backward inference part of the particle filter. Section II.C solves the challenges in implementing the particle filter due to the structural complexity of the Bayesian network. Then, Sec. II.D contributes to reducing the computational efforts by modifying the structure of the DBN; thus, inference is not needed if the load is observed but the damage is not. The modified DBN is proved to be equivalent to the original DBN, but it reduces the number of time steps that require Bayesian inference, so that reduces the computational cost.

A. Brief Introduction to Particle Filter

The particle filter is a general algorithm to track the evolution of the state variables in a DBN. In this section, capital letters denote random variables; and lowercase letters denote particles, where the superscript i indicates that it is the ith particle. The subscripts of letters indicate the time step. Thus, the state variables at time step t are

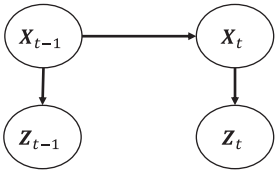


Fig. 2 Simple DBN.

denoted as $\boldsymbol{X}_t$, as shown in the simple DBN of Fig. 2. Assume that the state variables $\boldsymbol{X}_t \in \Re^m$ at time t evolve from the state variables $\boldsymbol{X}_{t-1} \in \Re^m$ according to the state function

$$\boldsymbol{X}_t = f(\boldsymbol{X}_{t-1}, \boldsymbol{v}_{t-1}) \tag{1}$$

where $\boldsymbol{v}_{t-1} \in \Re^m$ is the vector of noise terms in the state function. The measurement $\boldsymbol{Z}_t \in \Re^n$ is obtained according to the measurement function

$$\boldsymbol{Z}_t = h(\boldsymbol{X}_t, \boldsymbol{n}_t) \tag{2}$$

where $\boldsymbol{n}_t \in \Re^n$ is the vector of noise terms in the measurement function.

In the case that the DBN represented by Eqs. (1) and (2) is not a linear Gaussian DBN, several particle filter algorithms are developed to track the evolution of $\boldsymbol{X}_t$ and $\boldsymbol{Z}_t$. The most basic particle filter algorithm is sequential importance sampling (SIS) [15]. SIS considers the full joint posterior distribution at time step t: $p(\boldsymbol{X}_{0:t}|\boldsymbol{Z}_{1:t})$. This distribution is approximated with a weighted set of particles

$$\{\boldsymbol{x}^i_{0:t}, \omega^i_t\}^N_{t=1}$$

These particles approximate the joint posterior distribution $p(\boldsymbol{X}_{0:t}|\boldsymbol{Z}_{1:t})$ by

$$p(\boldsymbol{X}_{0:t}|\boldsymbol{Z}_{1:t}) \approx \sum_{i=1}^{N} \omega^i_t \delta_{x^i_{0:t}} \tag{3}$$

where $\delta_{x^i_{0:t}}$ is a delta function at $\boldsymbol{x}^i_{0:t}$.

At time step t, the ith particle of $\boldsymbol{X}_t$ is denoted as $\boldsymbol{x}^i_t$, and it is sampled based on the current state $\boldsymbol{X}^i_{0:t-1}$ and the observation $\boldsymbol{Z}_{1:t}$ according to a proposal density:

$$\boldsymbol{X}^i_t \sim q(\boldsymbol{X}_t|\boldsymbol{X}^i_{0:t-1}, \boldsymbol{Z}_{1:t}) \tag{4}$$

In other words, the new state $\boldsymbol{X}^i_t$ of the ith particle at time step t is sampled from a distribution that takes the current state $\boldsymbol{X}^i_{0:t-1}$ and the observation $\boldsymbol{Z}_{1:t}$ as parameters.

At time step t, the weight ω^i_t is updated from ω^i_{t-1} by

$$\omega^i_t \propto \omega^i_{t-1} \frac{p(\boldsymbol{Z}_t|\boldsymbol{X}^i_t)p(\boldsymbol{X}^i_t|\boldsymbol{X}^i_{t-1})}{q(\boldsymbol{X}^i_t|\boldsymbol{X}^i_{t-1}, \boldsymbol{Z}_t)} \tag{5}$$

In addition, the initial states $\boldsymbol{X}^i_0$ are sampled from the joint prior distribution of the state variables, and the initial weight ω^i_0 for each particle is $1/N$.

B. Degeneracy, Sample Impoverishment, and Resampling

In practice, iterations of Eqs. (4) and (5) over time step t may lead to a particle degeneracy problem; i.e., only a few particles have significant weights, but most particles have negligible weights. In that case, we are assigning most computational efforts to the particles of nonsignificant contribution to the posterior distribution. This degeneracy problem can be solved by resampling, i.e., generating a new set of N particles based on the particles of $\boldsymbol{X}_t$. The new particles represent the same posterior distribution as the former particles.

The simplest strategy of resampling is generating new resampled particles based on the discrete approximation shown in Eq. (3), and the weight of each new particle is set as $1/N$ again. This resampling is a bootstrapping process of N iterations, and each iteration selects one particle from the current particles with replacement. In an iteration, the probability that a particle is selected is proportional to its weight.

The aforementioned resampling strategy based on Eq. (3) is adopted in a widely used algorithm: the sampling importance resampling (SIR) algorithm [15]. The SIR algorithm 1) takes the state transition distribution $p(\boldsymbol{X}_t|\boldsymbol{X}^i_{t-1})$ as the proposal density distribution $q(\boldsymbol{X}_t|\boldsymbol{X}^i_{0:t-1}, \boldsymbol{Z}_{1:t})$, and 2) conducts resampling at each iteration. Thus, Eqs. (4) and (5) reduce to

$$\boldsymbol{X}^i_t \sim p(\boldsymbol{X}_t|\boldsymbol{X}^i_{t-1}) \tag{6}$$

$$\omega^i_t \propto p(\boldsymbol{Z}_t|\boldsymbol{X}^i_t) \tag{7}$$

Note that resampling based on Eq. (3) is after the calculation of Eqs. (6) and (7) at each time step, where new particles of $\boldsymbol{X}^i_t$ are generated and the weight of each new particle is set as $1/N$.

It is straightforward to implement the SIR algorithm because it only requires sampling from the distribution $p(\boldsymbol{X}_t|\boldsymbol{X}^i_{t-1})$ and evaluating the likelihood $p(\boldsymbol{Z}_t|\boldsymbol{X}^i_t)$. However, the resampling strategy in the SIR brings a new problem of sample impoverishment. In the resampling based on Eq. (3), the particles of higher weights will be duplicated and the particles of lower weights will be discarded. This means the loss of sample diversity and, in the worst case, all the particles will collapse to a single particle [15].

A regularized particle filter (RPF) algorithm was developed in [23] to solve the sample impoverishment problem. Instead of the discrete approximation in Eq. (3), the RPF uses kernel density estimation (KDE) [24] as a continuous approximation to resample the particles, and the resampled particles will have equal weights. Because $\boldsymbol{X}_t$ contains more than one variable, this KDE is multivariate [25]. However, this KDE-based resampling is designed for continuous variables, whereas the DBN in this problem contains both continuous and discrete variables. A new resampling strategy is proposed here by combining these two different resampling strategies:

1) Resample the particles of $\boldsymbol{X}_t$ based on Eq. (3) and assign equal weights $1/N$ to the new particles.

2) For a variable $X \in \boldsymbol{X}_t$, conduct steps 3 and 4 if it is continuous.

3) Resample the particles of X again based on a univariate KDE.

4) Rearrange the new particles of X so that they have the same < magnitude rank, ordinal > pairwise relationship as the particles from step 1. (For example, for the particles of X from step 1, if the largest value has the ordinal number of 100, arrange the largest particle of X from step 3 at ordinal 100 too.)

Note that the step 3 losses the dependence 1) among the continuous variables and 2) between the continuous variables and the discrete variables. Step 4 is used to restore this dependence.

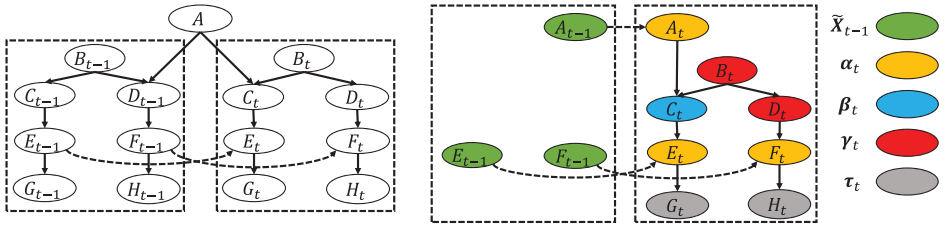
In summary, for a diagnosis step, the particle filter in this paper 1) uses Eqs. (6) and (7) from the SIR algorithm to update the weights of the particles in Bayesian inference; and 2) uses the proposed resampling strategy in the preceding four steps. However, in a prognosis step where backward inference is not required, we only need to implement Eq. (6).

The math of the particle filter was given in this subsection. However, it is still nontrivial to realize the particle filter for a DBN of complex structure. The reasons for this challenge and the solution will be discussed in Sec. II.C.

C. Implementing Particle Filter in DBN

Due to the complexity of the DBN for a realistic system, the implementation of the PF algorithm is nontrivial. There are two challenges in implementing the particle filter algorithm to a complex DBN. First, in addition to dynamic nodes for which the states change over time, static nodes shared by all the time steps are also included. An example of the static node is node A in Fig. 3a. The existence of static nodes violates the prerequisite assumption of the DBN: one separate BN for each step.

Second, the states of some dynamic nodes depend not only on the previous state of these variables but also on some other variables in the current time step. For example, in Fig. 3a, node E_t depends on



a) DBN **b) New particle generation**

Fig. 3 Particle filter for an illustrative DBN.

E_{t-1} and C_t. Thus, at time step t, C_t must be sampled before E_t. This requires us to distinguish the parent nodes of each state variable in $\boldsymbol{X}_t$ to implement Eq. (6).

The solutions to these two challenges are explained using an illustrative DBN shown in Fig. 3. The first challenge can be resolved by separating a static node into two identical nodes. In Fig. 3, the shared static node A is split into A_{t-1} and A_t. Subscripted by $t-1$ and t, A_{t-1} belongs to the BN at time step $t-1$ and A_t belongs to the BN at time step t. An arrow representing the deterministic relationship $A_{t-1} = A_t$ directs from A_{t-1} to A_t so that these two nodes are identical. In summary, this solution fulfills the assumption of one BN for each time step, and it guarantees that the same static node is shared by each time step.

The solution to the second challenge requires several steps in order to realize Eq. (6). The nodes in the BNs at time steps $t-1$ and t are classified into five groups:

1) The first group ($\tilde{\boldsymbol{X}}_{t-1}$) includes the state variables in $\boldsymbol{X}_{t-1}$ with arrows directed to state variables in $\boldsymbol{X}_t$. Among all the nodes in $\boldsymbol{X}_{t-1}$, only $\tilde{\boldsymbol{X}}_{t-1}$ are the parent nodes of the variables in $\boldsymbol{X}_t$; thus, Eq. (6) can be written as $\boldsymbol{X}_t^i \sim p(\boldsymbol{X}_t|\tilde{\boldsymbol{X}}_{t-1}^i)$. $\tilde{\boldsymbol{X}}_{t-1} = \{A_{t-1}, E_{t-1}, F_{t-1}\}$. for the illustrative DBN in Fig. 3.

2) The second group ($\boldsymbol{\alpha}_t$) includes the child nodes of $\tilde{\boldsymbol{X}}_{t-1}$ in the BN at time step t. The sampling of α_t depends on the value of $\tilde{\boldsymbol{X}}_{t-1}$ in the previous BN, where $\boldsymbol{\alpha}_t = \{A_t, E_t, F_t\}$ for the illustrative DBN in Fig. 3.

3) The third group ($\boldsymbol{\beta}_t$) includes the intermediate nodes of α_t. A node in $\boldsymbol{\beta}_t$ has both ancestor and descendant nodes in α_t. A node in $\boldsymbol{\beta}_t$ depends on some nodes in α_t, and a node in α_t can depend on some nodes in $\boldsymbol{\beta}_t$. In Fig. 3, we have $\boldsymbol{\beta}_t = C_t$.

4) The third group ($\boldsymbol{\gamma}_t$) includes the ancestor nodes of α_t or $\boldsymbol{\beta}_t$ in the BN at time step t. No node in $\boldsymbol{\gamma}_t$ is the descendant node of $\tilde{\boldsymbol{X}}_{t-1}$, i.e., the sampling of $\boldsymbol{\gamma}_t$ is independent of the previous BN. The distribution of $\boldsymbol{\gamma}_t$ is denoted as $p(\boldsymbol{\gamma}_t)$. The sampling of $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t$ depends both on $\tilde{\boldsymbol{X}}_{t-1}$ and $\boldsymbol{\gamma}_t$, which can be expressed by a conditional distribution $p(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t|\tilde{\boldsymbol{X}}_{t-1}^i, \boldsymbol{\gamma}_t^i)$. In Fig. 3, we have $\boldsymbol{\gamma}_t = \{B_t, D_t\}$.

5) The fourth group ($\boldsymbol{\tau}_t$) includes the descendant nodes of α_t or $\boldsymbol{\beta}_t$ in the BN at time step t. The sampling of $\boldsymbol{\tau}_t$ depends on α_t or $\boldsymbol{\beta}_t$, i.e., a conditional probability distribution $\boldsymbol{\tau}_t^i \sim p(\boldsymbol{\tau}_t|\boldsymbol{\alpha}_t^i, \boldsymbol{\beta}_t^i)$. In Fig. 3, we have $\boldsymbol{\tau}_t = \{G_t, H_t\}$.

As $\boldsymbol{X}_t$ is denoted as $\{\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t, \boldsymbol{\gamma}_t, \boldsymbol{\tau}_t\}$ based on the aforementioned classification, the sampling of $\boldsymbol{X}_t^i$ in Eq. (6) is realized sequentially by

$$\begin{aligned} \boldsymbol{\gamma}_t^i &\sim p(\boldsymbol{\gamma}_t) \\ \boldsymbol{\alpha}_t^i, \boldsymbol{\beta}_t^i &\sim p\left(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t|\tilde{\boldsymbol{X}}_{t-1}^i, \boldsymbol{\gamma}_t^i\right) \\ \boldsymbol{\tau}_t^i &\sim p(\boldsymbol{\tau}_t|\boldsymbol{\alpha}_t^i, \boldsymbol{\beta}_t^i) \end{aligned} \tag{8}$$

For the illustrative DBN in Fig. 3, to generate new particles

$$\boldsymbol{X}_t^i = \{A_t^i, B_t^i, C_t^i, D_t^i, E_t^i, F_t^i, G_t^i, H_t^i\}$$

based on $\tilde{\boldsymbol{X}}_{t-1}^i = \{A_{t-1}^i, E_{t-1}^i, F_{t-1}^i\}, \boldsymbol{\gamma}_t^i = \{B_t^i, D_t^i\}$ is first sampled by

$$p(\boldsymbol{\gamma}_t) = p(B_t, D_t) = p(B_t)p(D_t|B_t^i)$$

Then, $\boldsymbol{\alpha}_t^i = \{A_t^i, E_t^i, F_t^i\}$ and $\boldsymbol{\beta}_t^i = C_t^i$ are sampled by

$$\begin{aligned} p\left(\boldsymbol{\alpha}_t, \boldsymbol{\beta}_t|\tilde{\boldsymbol{X}}_{t-1}^i, \boldsymbol{\gamma}_t^i\right) &= p(A_t|A_{t-1}^i)p(C_t|A_t^i, B_t^i)p(E_t|E_{t-1}^i, C_t^i) \\ &\times p(F_t|F_{t-1}^i, D_t^i) \end{aligned}$$

Finally, $\boldsymbol{\tau}_t = \{G_t, H_t\}$ is sampled from

$$p(\boldsymbol{\tau}_t|\boldsymbol{\alpha}_t^i, \boldsymbol{\beta}_t^i) = p(G_t|E_t^i)p(H_t|F_t^i)$$

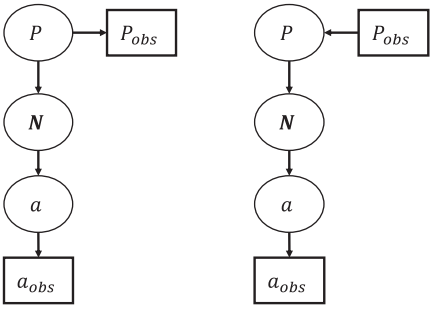
D. Computation Effort Reduction by Modifying the DBN Structure

Computational efficiency is a main concern in implementing the PF. Compared to the analytical algorithms such as the Kalman filter, the PF is more computationally intensive because, in each time step, each particle needs to be processed individually and the number of particles can be numerous, whereas the Kalman filter directly handles the probability distributions of the state variables. In addition, the diagnosis step is more expensive than the prognosis step in implanting the PF. The prognosis step is purely forward uncertainty propagation and only requires particle generation by Eq. (8). In contrast, the diagnosis step requires Bayesian inference, and thus brings extra computation effort for the likelihood $p(\boldsymbol{Z}_t|\boldsymbol{X}_t^i)$, the weight ω_t^i, and the resampling. The computational cost increases as more diagnosis steps are needed.

The damage a in an airframe component is caused by the load P applied on it; thus, the DBN starts from the node of load P and ends at the node of damage a (in crack growth analysis of the aircraft, the damage a is the fatigue crack length).

Generally, P and a are observable. Due to the measurement error, the observed data of P are the realizations of a new random variable P_{obs} and the observed data of a are the realizations of a new random variable a_{obs}. Thus, in the BN node, P directs to node P_{obs}, indicating a measurement model such as $P_{\text{obs}} = P + \epsilon_P$, where ϵ_P is the measurement error; node a directs to node a_{obs}, indicating another measurement model such as $a_{\text{obs}} = a + \epsilon_a$, where ϵ_a is the measurement error. The resultant BN at one time step is shown in Fig. 4a, where $\boldsymbol{N}$ denotes all the other nodes except for P, P_{obs}, a and a_{obs}.

If a data point D_P of the load P is observed, in the BN node, P_{obs} will be fixed at D_P; similarly, node a_{obs} will be fixed at the data point D_a if the damage a is observed. Because neither node of P_{obs} and a_{obs}



a) Original BN **b) Modified BN**

Fig. 4 Original BN and modified BN.

is a root node, the diagnosis of Bayesian inference is needed whenever the data of P and/or a are available. If the full load history is measured, diagnosis of the Bayesian inference is conducted in every time step, even if the crack length data are sparse. This causes tremendous computational cost.

In fact, it can be proved that, under certain assumptions (explained in the following), we can reverse the arrow from P to P_{obs} (i.e., replace $P \to P_{\text{obs}}$ with $P_{\text{obs}} \to P$), and the modified BN shown in Fig. 4b is equivalent to the original one in Fig. 4a. In the modified BN, P_{obs} is a root node so that Bayesian inference is not needed if only the load is observed. In other words, diagnosis is conducted only at limited steps with the crack length observed, which reduces the computational cost significantly. Proof of the equivalence between the original BN and the modified BN is given as follows.

If the load is not observed, the node P_{obs} can be removed in the BN; thus, the original BN and the modified BN are exactly the same. If the load is observed, two cases need to be considered. In the first case, we assume that both the load and the crack length are observed. In the original BN, if the load and crack length data are denoted as D_P and D_a, the posterior distribution over the BN is

$$\begin{aligned} &p(P, \boldsymbol{N}, a|P_{\text{obs}} = D_P, a_{\text{obs}} = D_a) \\ &\quad \propto p(P_{\text{obs}} = D_P, a_{\text{obs}} = D_a|P, \boldsymbol{N}, a)p(P, \boldsymbol{N}, a) \\ &\quad = p(a_{\text{obs}} = D_a|a)p(P_{\text{obs}} = D_P|P)p(P)p(N|P)p(a|\boldsymbol{N}) \end{aligned} \tag{9}$$

where $p(a_{\text{obs}} = D_a|a)p(P_{\text{obs}} = D_P|P)$ is the likelihood function, and $p(P)p(\boldsymbol{N}|P)p(a|\boldsymbol{N})$ is the prior distribution over P, $\boldsymbol{N}$ and a.

For the modified BN, the posterior distribution is

$$\begin{aligned} &p(P, \boldsymbol{N}, a|P_{\text{obs}} = D_P, a_{\text{obs}} = D_a) \\ &\quad \propto p(P_{\text{obs}} = D_P, a_{\text{obs}} = D_a|P, \boldsymbol{N}, a)p(P, \boldsymbol{N}, a) \\ &\quad = p(a_{\text{obs}} = D_a|a)p(P_{\text{obs}} = D_P|P, \boldsymbol{N}, a)p(P, \boldsymbol{N}, a) \\ &\quad = p(a_{\text{obs}} = D_a|a)p(P|P_{\text{obs}} = D_P)p(N|P)p(a|\boldsymbol{N}) \end{aligned} \tag{10}$$

where $p(a_{\text{obs}} = D_a|a)$ is the likelihood function, $p(P|P_{\text{obs}} = D_P)p(N|P)p(a|\boldsymbol{N})$ is the prior distribution over P and $\boldsymbol{N}$, and a is conditioned at $P_{\text{obs}} = D_P$. The posteriors in Eqs. (9) and (10) are equivalent if

$$p(P_{\text{obs}} = D_P|P)p(P) \propto p(P|P_{\text{obs}} = D_P)$$

As the measurement noise is generally assumed to be a zero mean Gaussian distribution, we have $P_{\text{obs}} = P + N(0, \sigma_p)$, which gives $P = P_{\text{obs}} - N(0, \sigma_p)$. Then, it can be proved that

$$\begin{aligned} p(P_{\text{obs}} = D_P|P) &= p(P|P_{\text{obs}} = D_P) \\ &= \frac{1}{\sigma_p\sqrt{2\pi}}\exp\left(-\frac{(P - D_p)^2}{2\sigma_p^2}\right) \end{aligned} \tag{11}$$

Thus, the condition of

$$p(P_{\text{obs}} = D_P|P)p(P) \propto p(P|P_{\text{obs}} = D_P)$$

will be fulfilled if node P in the original BN has a noninformative uniform distribution such that $p(P)$ is a constant.

In the second case, we assume that only the load is observed. The posterior distribution over the original BN is

$$\begin{aligned} p(P, \boldsymbol{N}, a|P_{\text{obs}} = D_P) &\propto p(P_{\text{obs}} = D_P|P, \boldsymbol{N}, a)p(P, \boldsymbol{N}, a) \\ &= p(P_{\text{obs}} = D_P|P)p(P)p(N|P)p(a|\boldsymbol{N}) \end{aligned} \tag{12}$$

The posterior distribution over the modified BN is obtained purely by uncertainty propagation:

$$p(P, \boldsymbol{N}, a|P_{\text{obs}} = D_P) = p(P|P_{\text{obs}} = D_P)p(N|P)p(a|\boldsymbol{N}) \tag{13}$$

Note that Eqs. (12) and (13) are derived based on the Bayes's theorem, and distribution normality is not assumed. Equations (12) and (13) are equivalent under the same assumptions: 1) zero mean Gaussian measurement error for the load; and 2) $p(P)$ is a noninformative uniform prior distribution in the original BN if the load is observed. The first one is a widely used assumption in the literature. The second one requires that load P_t is independent of P_{t-1} if P_t is observed. This is also a reasonable assumption because the observation of the load at time t provides strong information for the true value of the load at time t such that the information from P_{t-1} can be neglected. A time series model giving the CPD of $p(P_t|P_{t-1})$ is still applicable if P_t is not observed.

In summary, this section distinguished the steps of diagnosis and prognosis. The time-consuming diagnosis was required if, and only if, any child node was observed. This section also showed that, under two weak assumptions, the CPD $P \to P_{\text{obs}}$ could be replaced by $P_{\text{obs}} \to P$ such that the number of diagnosis steps could be reduced significantly, i.e., Bayesian updating wais required only if crack inspection data were available.

III. Dynamic Bayesian Network of Crack Growth on Leading Edge of Aircraft Wing

Different fracture mechanics-based fatigue crack growth models have been developed to calculate the propagation of long cracks, including Paris's law [26], modified Paris's law [27], Wheeler's retardation model [28], etc. Generally, these models require computing the stress intensity factor K, for which the finite element analysis (FEA) is widely used. Two techniques of using the FEA to compute the crack growth can be found in the literature:

1) Include the crack geometry into the FEA model and compute the stress intensity factor by the FEA directly; then, calculate the crack growth using a crack growth law and adjust the crack geometry by modifying the input file to the FEA model [29,30].

2) Build a FEA model without the crack geometry and compute the nominal stress at the crack; then, calculate the stress intensity factor using an analytical formula and subsequent crack growth using a crack growth law [31].

Due to the mesh complexity of the FEA model with the crack geometry, the computational cost of technique 2 is significantly smaller than technique 1. Because crack growth prediction under uncertainty requires numerous runs of the FEA model, technique 2 is applied in this paper. Based on technique 2, the rest of this section discusses the uncertainty sources in predicting the fatigue crack growth on the leading edge of an aircraft wing; then, all the uncertainty sources are incorporated into the DBN. Note that the DBN is still applicable regarding technique 1, and this paper selects technique 2 only for higher computational efficiency.

A. Uncertainty Sources in Aircraft Wing Crack Growth Prediction

1. *Uncertainty Sources in the FEA Model and Surrogate Model*

Figure 5 shows the FEA model of the leading edge of an aircraft wing. The spring and beam elements in Fig. 5 simulate the connection between the leading edge and the wing body. The load on the leading

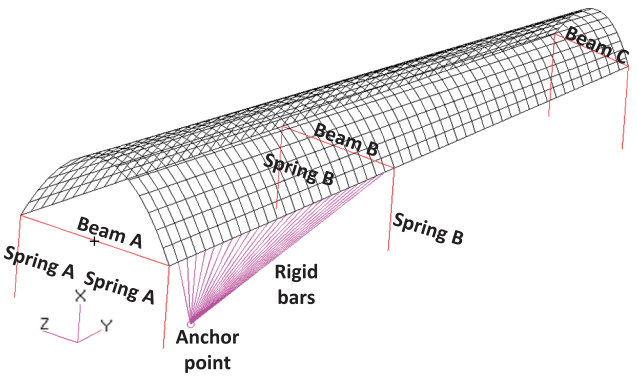


Fig. 5 Leading edge of an aircraft wing.

edge is simulated by connecting the leading edge to an anchor point through rigid bars and applying the load P on the anchor point. A single bolt is assumed to fix the anchor point to the wing body. Fifteen geometric and material parameters are assumed as random variables in the FEA model:

1) For $T_i (i = 1 \text{ to } 7)$, the leading edge is divided into seven sections along the Y axis, and T_i is the thickness of the ith section.

2) For $K_i (i = 1 \text{ to } 4)$, K_1 and K_2 are the stiffnesses of springs A and B along the Y axis; K_3 and K_4 are the stiffnesses of springs A and B along the Z axis.

3) The inboard beam property (IBP) is the area moment of inertia of beam A.

4) The outboard beam property (OBP) is the area moment of inertia of beam B.

5) The taper ratio (TR) measures the rate that the leading-edge width shrinks from the wing root to the wing tip; here, it is defined as the ratio of the beam A length to the beam C length.

6) Y_A is the coordinate of the anchor point along the Y axis; the value of Y_A varies if the bolt is loose.

All the aforementioned parameters, except for Y_A, have deterministic but unknown values, and thus bring epistemic uncertainty. Prior distributions are assigned to them, and the proposed DBN-based P2IAT model seeks to reduce their epistemic uncertainty by Bayesian inference. The value of Y_A changes over time; thus, the proposed P2IAT model needs to track its evolution.

These 15 parameters and the load P are the inputs to the FEA model in Fig. 5, which computes the nominal stress S at the crack. Probabilistic prediction as well as Bayesian inference require many evaluations of the analysis model. To achieve computational efficiency, this paper uses a Gaussian process (GP) surrogate model [32,33] to replace the FEA model. Training points are obtained by repeatedly running the FEA model at different combinations of values (design of experiments points) of the 15 parameters and the load P. At given inputs, the prediction of the GP model is a normal distribution $S \sim N(\mu_{\text{GP}}, \sigma^2_{\text{GP}})$, which represents the surrogate model uncertainty in computing the stress for a given value of the inputs. This also indicates that these 15 parameters and the load P are the parent nodes of stress S in the DBN, and the corresponding conditional probability distribution is given by the GP model prediction $S \sim N(\mu_{\text{GP}}, \sigma^2_{\text{GP}})$.

Not all the 15 parameters are equally important to the crack growth. A global sensitivity analysis (GSA) by Sobol indices [34–37] can be used to assess the contribution of each parameter to the uncertainty in the crack growth. Parameters of low sensitivity can be fixed at their nominal values, thus reducing the computational cost in diagnosis and prognosis. Consider a function $Y = F(\boldsymbol{X})$ where $\boldsymbol{X} = \{X_1, \ldots, X_k\}$ is a vector containing all the stochastic model inputs. The first-order sensitivity index S_i and the total effects sensitivity index S_i^T are defined as follows:

$$S_i = \frac{V(E(Y|X_i))}{V(Y)}, \quad S_i^T = 1 - \frac{V(E(Y|X_{-i}))}{V(Y)} \tag{14}$$

where X_{-i} denotes all the model inputs other than X_i. S_i measures the contribution of X_i by itself, whereas S_i^T measures the overall contribution of X_i by itself plus interactions with other inputs.

2. *Crack Growth Model Uncertainty and Damage State Uncertainty*

Once the nominal stress at the crack is computed using the GP model, the next step is to compute the stress intensity factor and crack growth. Methods to compute the stress intensity factor for different load conditions and crack shapes were summarized in [38]. The validity of these models is generally problem dependent. For the sake of illustration, this paper assumes a mode 1 uniaxial crack; thus, the range of stress intensity factor in one time step is

$$\Delta K = 1.2F\Delta S\sqrt{\pi a^0} \tag{15}$$

where $1.2F$ is the crack shape factor, ΔS is the stress range, and a^0 is the initial crack length in the current time step. Here, F is defined as a multiplier for the shape factor, and the uncertainty in F represents the uncertainty in the shape factor.

Next, for the sake of illustration, this paper uses the Paris law to compute the crack growth Δa in each time step:

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C\Delta K^m \tag{16}$$

where C and m are the Paris law parameters obtained from material component experiments; $\mathrm{d}a/\mathrm{d}N$ is the crack growth rate, and its magnitude is equal to the predicted crack growth Δa in one time step. The crack length after the current time step is $a = a^0 + \Delta a$. C and m are assumed to be known constants in this paper to keep the focus on other parameters that provide particular challenges to the DBN that are addressed in this paper; C and m can be easily treated as aleatory or epistemic uncertain quantities and are included in the DBN as needed.

The uncertainty sources in the crack growth prediction are the uncertainties in the parameters of Eq. (15), which are affected by the damage state, and the uncertainties regarding the parameters of Eq. (2) (ignored in this paper). In this paper, two damage states are considered:

1) The first state is bolt looseness B. For the sake of illustration, assume that bolts are used to fix the anchor point to the wing body. Assume that all the bolts are collectively represented by one notional bolt with equivalent properties. Whether the bolt becomes loose depends on its resistance R and the current load P. A higher P or a lower R leads to a higher probability of a loose bolt ($B = 1$). The bolt will stay loose once it becomes loose ($B = 0$). The loose bolt causes uncertainty in the anchor point position Y_A, thus affecting the nominal stress S at the crack location. In addition, Eq. (17) is assumed to simulate the degradation of the bolt resistance with the time step t. In Eq. (17), R_0 is the initial bolt resistance; k is the degradation coefficient, and it has a negative value so that $R(t)$ decreases with t:

$$R(t) = R_0 \exp(kt) \tag{17}$$

2) The second state is a crack tip in the elastic zone vs the plastic zone M. The aircraft wing is mostly elastic ($M = 0$); it is assumed that randomly located plastic zones ($M = 1$) can be caused by accidents such as a dropped hammer; the crack is assumed to start at the elastic zone, and there is a finite probability that the crack grows into a plastic zone in any time step; and the crack is assumed to stay in the plastic zone once it reaches it. The following are assumed:

a) The shape factor multiplier in the elastic zone F_e has a known deterministic value obtained from material coupon experiments.

b) The plastic zone retards the crack growth; thus, the multiplier F_p in the plastic zone is smaller than F_e.

c) F_p has a deterministic but unknown value, i.e., epistemic uncertainty. This damage state M can be represented by expanding Eq. (15) as follows:

$$\Delta K = \begin{cases} 1.2F_e\Delta\sigma\sqrt{\pi a^0} & \text{if } M = 0 \\ 1.2F_p\Delta\sigma\sqrt{\pi a^0} & \text{if } M = 1 \end{cases} \tag{18}$$

The preceding damage states bring two new uncertainty sources: 1) whether the damage states have occurred; and 2) uncertainty in the value of F_p. The proposed DBN-based P2IAT model is beneficial in tracking the damage states and quantifying the uncertainty in F_p. In addition, the damage states are discrete variables, thus requiring a DBN that can handle both discrete and continuous variables.

3. *Load Uncertainty*

The uncertainty in load P depends on specific cases. In case 1, the load history at each time step is measured by sensors in the aircraft wing. The measured load history can be used to simulate the flight, diagnose damage states, and compute the crack length after the flight. Techniques to measure the load history include flight parameter-based loads monitoring and strain-gauge-based loads monitoring [39]. In this case,

the uncertainty in the load history is the measurement error. The numerical example in Sec. IV assumes the measurement error as a zero mean Gaussian noise, i.e., $\epsilon_p \sim N(0, \sigma_p^2)$.

In case 2, the P2IAT model is used to simulate the future load time history and predict the crack growth. To model this time series input based on the observed load history in earlier flights and capture the uncertainty in the future loading, two types of time domain methods have been developed: time-step counting methods, and random process methods. The time-step counting methods [40] discretize the time series into k levels and extract a counting matrix from the data. The counting matrix is used to generate the load history stochastically. In contrast, one of the random process methods (e.g., the autoregressive moving average [41] model) assumes that the input in the current time step is a linear function of 1) its past p values; and 2) the current and past q values of noise terms. Both types of methods can be used in the P2IAT model. The values of its orders p and q are first identified by matching the theoretical autocorrelation function to the sample autocorrelation function computed from the observed time series data. The Ljung–Box Q statistic [42] can be used to measure the adequacy of the matching. In case 2, the load uncertainty includes the natural variability in the time series input and epistemic uncertainty due to limited information in building the model.

Different conditions in cases 1 and 2 affect the DBN structure. Let P_{t-1} and P_t denote the loads at times $t-1$ and t, and let $P_{\text{obs}_{t-1}}$ and P_{obs_t} denote their observations at time $t-1$ and t, respectively. In the BN of case 1, P_{t-1} and P_t are directly connected to the nodes $P_{\text{obs}_{t-1}}$ and P_{obs_t}, respectively, giving the conditional probability distributions of $P_{\text{obs}_{t-1}} \sim N(P_{t-1}, \sigma_p^2)$ and $P_{\text{obs}_t} \sim N(P_t, \sigma_p^2)$. In the BN of case 2, the value of P_{t-1} affects the value of P_t; thus, an arrow of CPD defined by the time series model is used to connect them. The node P_{obs_t} or $P_{\text{obs}_{t-1}}$ is not necessary because the load is not observed. A hybrid case is also possible, i.e., both cases 1 and 2 occur in the DBN. The DBN structures of case 1, case 2, and the hybrid case are shown in Fig. 6.

4. *Crack Length Data Uncertainty*

The crack length data are assumed to be available from on-ground inspection, which brings two uncertainty sources: measurement error and data sparsity. Similar to the load uncertainty, the measurement error in the crack length data depends on the accuracy of the inspection technique, and it is generally assumed to have a zero mean Gaussian distribution $\epsilon_a \sim N(0, \sigma_a^2)$. The proposed methodology can also handle other distributions of measurement error.

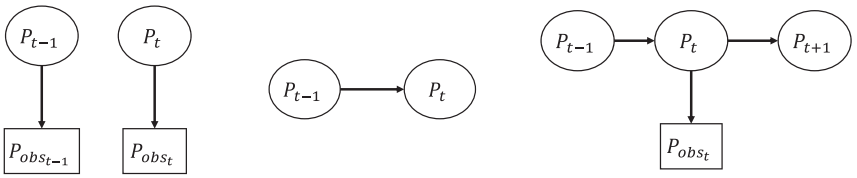
Crack length data are rarely available for every time step. Even if one data point is obtained after each mission and applied in the DBN for diagnosis and prognosis, the crack length data are missing during the mission; thus, data uncertainty is introduced by data sparseness.

In summary, two data sources are available for the P2IAT model of aircraft wing: load history data, and crack inspection data. The availability of these data can be quite flexible: load history data can be available at all time steps (case 1 in Fig. 6), no time step (case 2 in Fig. 6), and limited time steps (case 3 in Fig. 6); whereas crack inspection data are only available at sparse time steps. The DBN has the capacity of both Bayesian inference (diagnosis) and uncertainty propagation (prognosis).

B. Construction of the Dynamic Bayesian Network

As shown in Fig. 7, the uncertainty sources identified in Sec. III.A are represented by nodes in the DBN; nodes are connected by arrows that represent conditional probability distributions or deterministic functional relations. The subscript $t-1$ or t denotes the time step, and the symbols in Fig. 7 are explained in Table 1.

In Fig. 7, an elliptical node is a stochastic node, meaning the variable is stochastic for given values of parent nodes; thus, the



a) Case 1 **b) Case 2** **c) Hybrid case**

Fig. 6 DBN structure for loading history uncertainty.

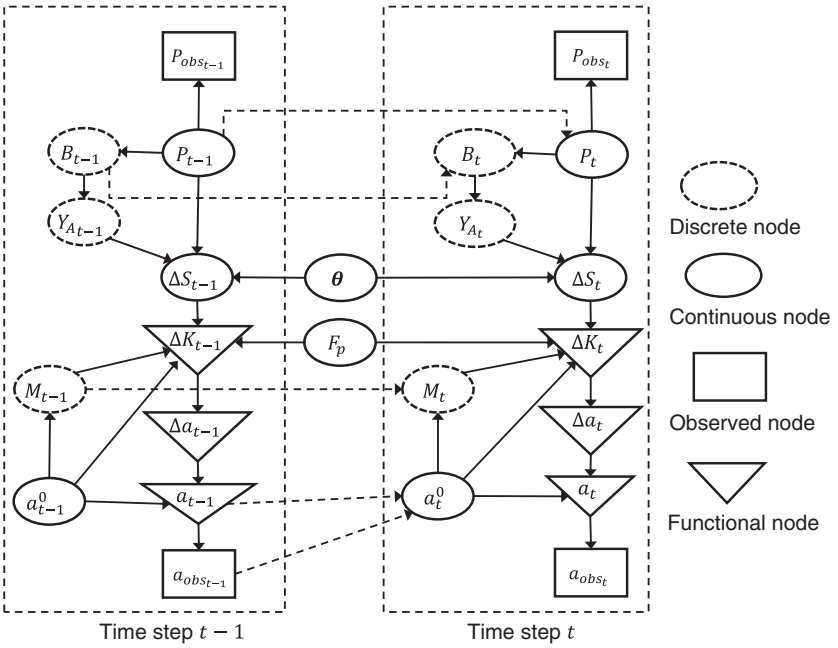


Fig. 7 Dynamic Bayesian network for crack growth.

Table 1 Nomenclature for the DBN

Value	Parameter
P_{obs}	Load observation
P	Load
B	Bolt looseness
Y_A	Anchor point position
ΔS	Stress range
M	Elastic/plastic zone
a^0	Crack length before current time step
ΔK	Stress intensity factor range
Δa	Crack growth in current time step
a	Crack length after current time step
a_{obs}	Crack length observation
$\boldsymbol{\theta}$	Geometric and material properties
F_p	Shape factor in the plastic zone

arrows toward it represent a CPD. A triangular node is a functional node, meaning the variable is the result of deterministic calculation for given values of parent nodes; thus, the arrow toward it represents a deterministic function. In addition, elliptical nodes with solid lines represent continuous variables, whereas elliptical nodes with dashed lines represent discrete variables. The rectangular nodes represent observed variables (e.g., load and crack length). In addition, solid arrows are used within a BN slice, and dashed arrows connect the nodes across different time steps.

In Fig. 7, node $\boldsymbol{\theta}$ represents all 15 geometric and material properties (except for Y_A) of the aircraft wing. Each property should be a node in the DBN connected to ΔS. They are depicted as a single node to save space.

Another special node in the DBN is a^0. For the BN in any time step, prior distributions are assigned to all the root nodes first; then, uncertainty propagation or Bayesian inference will be conducted. Except for time step 1, where prior distributions are defined by users, BNs at other time steps obtain the prior distributions by propagating the posterior distributions of previous time steps through the arrows connecting adjacent BNs. But, for a_t^0, the following is true:

1) If the crack length is not observed at time step $t-1$, its prior distribution is the predicted distribution of a_{t-1}, which means a deterministic functional relationship $a_{t-1} = a_t^0$; thus, a_{t-1} directs to a_t^0 in Fig. 7.

2) If the crack length is observed at time step $t-1$, its prior distribution for time step t should be defined using this data point. Let $D_{a_{t-1}}$ denote the observed data point value. This paper defines the prior distribution of a_t^0 as $N(D_{a_{t-1}}, \sigma_a)$; thus, $a_{\text{obs}_{t-1}}$ also directs to a_t^0 in Fig. 7.

Once the DBN is constructed, the diagnosis and prognosis are the next steps in the health monitoring of the aircraft wing. This was explained in Sec. III.

IV. Results and Analysis

A numerical example of crack growth on the leading edge of an aircraft wing is used to illustrate all the concepts explained in earlier sections. The structure of the aircraft wing was explained in Sec. III.A.1. A time series input of 10,000 steps is applied at the anchor point. The FEA result in Fig. 8 shows that, under the geometric and material property uncertainty and load uncertainty, the location of maximum stress is always around node 389. Thus, we assume that a crack of 0.0588 in. is initialized at node 389 and grows under the time series loading at the anchor point. A GP surrogate model predicting the stress at node 389 is built to replace the FEA model.

It is assumed that the time series input is observed at each step and that the measurement error is a zero mean Gaussian variable $N(0, 0.002^2)$. The observed load history is shown in Fig. 9. Furthermore, crack length data are assumed to be observed only at time steps 2000, 4000, 5600, 6400, and 6800.

As explained in Sec. III.A.1, the aircraft wing contains 15 stochastic geometric parameters and one stochastic crack growth model parameter F_p. Except for the anchor point position Y_A, all these parameters are static root nodes in Fig. 7.

Global sensitivity analysis results for the elastic zone ($M = 0$) and plastic zone ($M = 1$) are shown in Table 2. In the elastic zone, T_4 is the only significant parameter; in the plastic zone, T_4 and F_p are both significant. The sensitivity index of Y_A is small, indicating that Y_A and its only parent node bolt looseness B can be fixed at nominal values and the crack length data cannot track the evolution of B effectively. In this paper, we retain the nodes of B and Y_A in the DBN to quantitatively prove this proposition. The parameters in Table 2, except for T_4, F_p, and Y_A, are fixed at their nominal values.

The deterministic relationships (represented by the arrows to deterministic nodes in the DBN) were discussed in Sec. III.A. The conditional probability distribution for the continuous node ΔS_t is a Gaussian distribution

$$N(\mu_{\text{GP}}(Y_{A_t}, P_t, \boldsymbol{\theta}), \sigma_{\text{GP}}(Y_{A_t}, P_t, \boldsymbol{\theta}))$$

obtained by the GP surrogate model. Then, the DBN for the crack growth is constructed as in Fig. 7 and used for diagnosis and prognosis.

In this example, because the load P is observed at each time step, which provides strong evidence on the true value of P, the CPD of $p(P_t|P_{t-1})$ can be neglected; thus, the arrow from P_{t-1} to P_t can be removed in the DBN of Fig. 7. The prior distribution of node P at each time step is assumed to be a uniform distribution $U(P_l, P_u)$, where P_l and P_u are lower and upper bounds based on expert opinion. With these assumptions, the method of replacing $P \to P_{\text{obs}}$ by $P_{\text{obs}} \to P$ in Sec. II.D is applied to improve the computational efficiency. This paper uses 10^4 particles in the computation of this example, and the overall time cost is $T = 11{,}109$ s, including 1) 11,102 s ≈ 3.1 h spent on forward propagation of 10^4 time steps; and 2) 7 s spent on updating. If the method in Sec. II.D is not used so that each time step requires updating, the time spent on updating will be 17,500 s and the overall time cost will be 28,602 s ≈ 7.9 h. In other words, the proposed method in Sec. II.D reduces the time cost by 61%.

The resistance of the bolt decreases with time, as shown in Eq. (17). Here, we assume that the initial resistance of the bolt is $R_0 = 0.275$; the resistance $R(t)$ reduces to $0.9R_0$ after 10^4 time steps so that the degradation coefficient is $k = -1.0536 \times 10^{-5}$. The conditional probability tables for the discrete nodes B, Y_A, and M are assumed, as shown in Tables 3–5 for the sake of illustration.

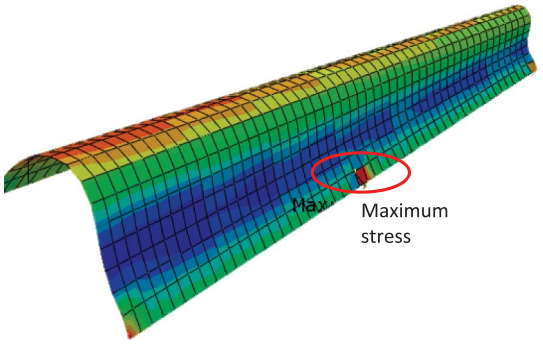


Fig. 8 Maximum stress in the aircraft wing.

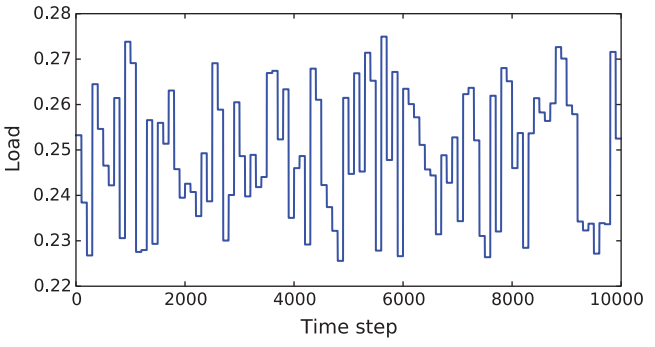


Fig. 9 Load history observation.

Table 2 GSA results

	Elastic zone		Plastic zone	
Parameters	First-order index	Total effects index	First-order index	Total effects index
T_1	0.000	0.000	0.000	0.000
T_2	0.000	0.000	0.000	0.000
T_3	0.012	0.025	0.001	0.002
T_4	0.875	0.902	0.104	0.277
T_5	0.002	0.010	0.000	0.000
T_6	0.001	0.003	0.000	0.000
T_7	0.000	0.000	0.000	0.000
IBP	0.000	0.000	0.000	0.000
OBP	0.000	0.001	0.000	0.000
K_1	0.000	0.000	0.000	0.000
K_2	0.001	0.002	0.000	0.000
K_3	0.000	0.000	0.000	0.000
K_4	0.000	0.000	0.000	0.001
TR	0.000	0.000	0.000	0.000
Y_A	0.022	0.031	0.002	0.004
F_p	— —	— —	0.696	0.865

Table 3 Conditional probability table of B_t

		$B_{t-1} = 0$		
$p(B_t \mid P_t, B_{t-1})$	$B_{t-1} = 1$	$P_t < 0.85R(t)$	$0.85R(t) < P_t < 0.95R(t)$	$P_t > 0.95R(t)$
$B_t = 0$	0	1	0.975	0.95
$B_t = 1$	1	0	0.025	0.05

Table 4 Conditional probability table of Y_{A_t}

$p(Y_{A_t} \mid B_t)$	$Y_{A_t} = 0$	$Y_{A_t} = 9.935 \pm 0.5$	$Y_{A_t} = 9.935 \pm 1.0$	$Y_{A_t} = 9.935 \pm 1.5$	$Y_{A_t} = 9.935 \pm 2.0$
$B_t = 0$	1.0	0.0	0.0	0.0	0.0
$B_t = 1$	0.0	0.25	0.125	0.075	0.05

Table 5 Conditional probability table of $M_{t-1} = 0$

	$M_{t-1} = 0$				
$p(M_t \mid a_t, M_{t-1})$	$M_{t-1} = 0$	$a_t < 0.1$	$0.1 < a_t < 0.12$	$0.12 < a_t < 0.15$	$0.15 < a_t$
$M_t = 0$	0	1	0.8	0.6	0.5
$M_t = 1$	1	0	0.2	0.4	0.5

True values are assumed for the model parameters in Table 2. Synthetic data for the observed crack length at time steps 2000, 4000, 5600, 6400, and 6800 are generated using the load history in Fig. 9 and the assumed true values. Using these data, the objectives of this numerical example are to calibrate the static variables T_4 and F_p, track the evolution of the time-dependent damage state variables B and M, and predict the crack length in the future. The results of these calculations are shown in Figs. 10–14.

Figures 10 and 11 show the updating of T_4 and F_p at each time step of inspection. Due to its high sensitivity in both the elastic zone and plastic zone, the uncertainty of T_4 is reduced significantly just after inspection 1 at $t = 2000$ of the crack length. In contrast, F_p is not updated at inspection 1 at $t = 2000$. The reason is that the crack tip has not reached the plastic zone at $t = 2000$ (shown in Fig. 12); thus, the obtained data do not contain information on the parameter F_p of the plastic zone. The uncertainty in F_p is reduced using the data from later inspections, where the crack tip has reached the plastic zone.

Figure 12 shows the inferred evolution of damage state M. Recall that $M = 0$ indicates that the crack tip is in the elastic zone, whereas $M = 1$ indicates that the crack tip is in the plastic zone. Because the two states of the discrete variable M are zero and one, the mean value of the inferred M is $0 \times p(M = 0) + 1 \times p(M = 1)$, i.e., equal to the probability $p(M = 1)$. This probability $p(M = 1)$ increases before inspection 1 due to the assumed conditional probability distribution $p(M_t \mid a_t, M_{t-1})$ and reaches around 0.1 at $t = 2000$. Then, the network was updated by the crack length data from inspection 1 and $p(M = 1)$ corrected to zero. The unobserved true value of M is still zero at $t = 2000$; thus, this correction is valid. This reduced

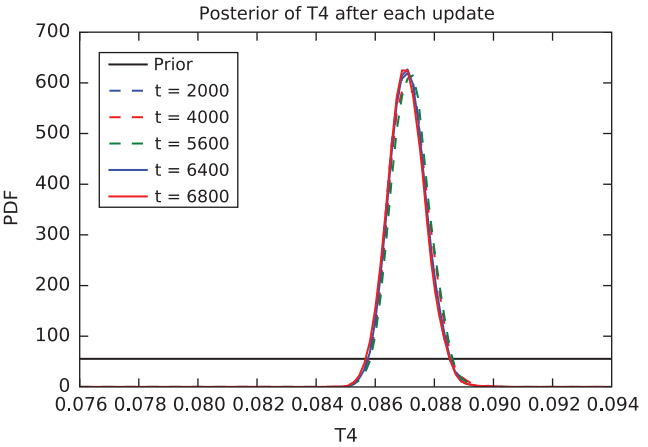


Fig. 10 Updating of T_4 (PDF, probability density function).

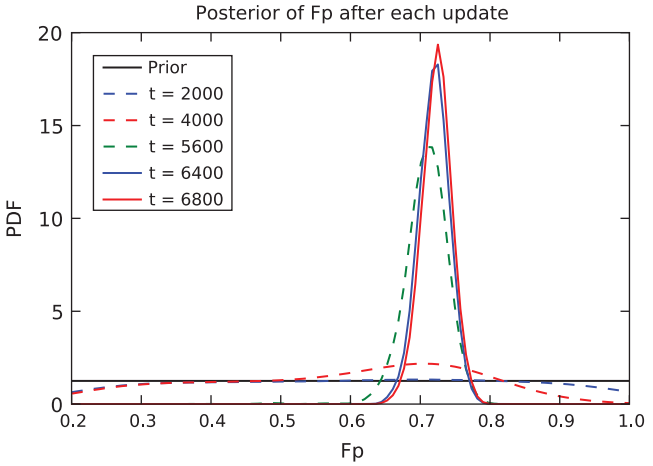


Fig. 11 Updating of F_p (PDF, probability density function).

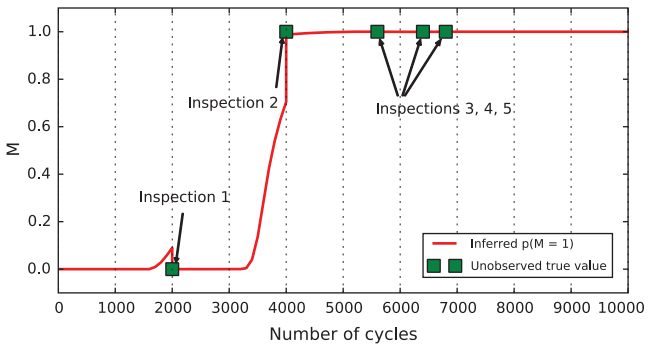


Fig. 12 Tracking damage state M (crack tip in elastic vs plastic zones).

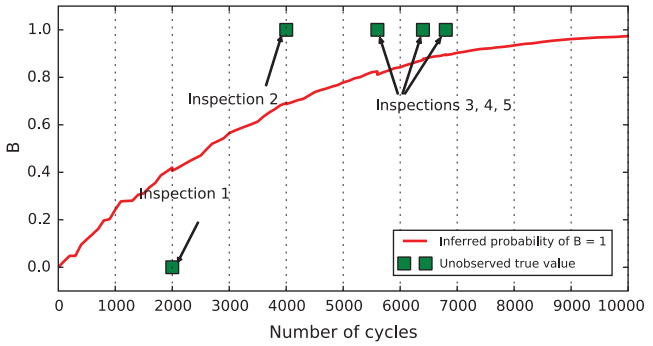


Fig. 13 Tracking damage state B (bolt loosening).

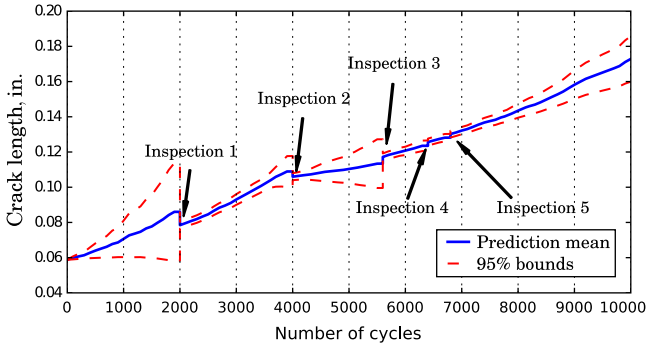


Fig. 14 Diagnosis and prognosis of crack length.

$p(M = 1)$ also reduces the effect of the uncertainty in F_p, and thus reduces the uncertainty in the crack length prediction, as shown in Fig. 14. A similar correction also occurs at inspection 2, where $(M = 1)$ is corrected from 0.7 to 1.0. In inspections 3, 4, and 5, Fig. 12 shows a probability of $p(M = 1) = 1$, i.e., the crack has reached the plastic zone.

Figure 13 shows the inferred evolution of damage state B, where $B = 1$ indicates a loose bolt and $B = 0$ indicates a tight bolt. The probability $p(B = 1)$, which is equal to the mean value of inferred B, increases before any inspection due to the assumed conditional probability distribution $p(B_t|P_t, B_{t-1})$. The curve fluctuates due to the randomness in load P_t. However, $p(B = 1)$ is not corrected significantly by the crack length data in the inspection. This can be explained by the GSA results in Table 2. The sensitivity of Y_A with respect to the crack length is negligible, meaning that, as the only parent node of Y_A, the bolt looseness B also has negligible influence on the crack growth, and is therefore not updated significantly.

Figure 14 shows the diagnosis and prognosis of the crack length. The uncertainty in the crack length is reduced to measurement error at each inspection, and it grows between inspections. The uncertainty grows fast (wide 95% bounds) before the first inspection because the uncertainty propagation is based on prior distributions of T_4. The uncertainty grows slower between inspections 1 and 2 due to 1) significantly reduced uncertainty in T_4 at the first inspection, as shown in Fig. 10; and 2) low probability that the crack has reached the plastic zone, as shown in Fig. 12, i.e., low probability that the uncertainty in F_p is introduced. The uncertainty grows fast between inspections 2 and 3 because the crack has reached the plastic zone so that the uncertainty in F_p is introduced; and the data from inspection 2 barely reduce the uncertainty in F_p, as shown in Fig. 11. The uncertainty grows slower again after inspection 3 because the uncertainty in F_p has been reduced by the observation data at inspection 3, as shown in Fig. 11.

Note that model validation [43–46] is not included in current examples, but future work may need to consider it.

V. Conclusions

Various uncertainty sources affect the health state diagnosis and prognosis of aircraft components. This paper establishes a framework for probabilistic health diagnosis and prognosis using a dynamic Bayesian network. This framework is versatile due to the following characteristics: 1) incorporates various aleatory and epistemic uncertainty sources; 2) handles both discrete and continuous variables; 3) allows the continuous variables to have any distribution type; and 4) allows nonlinear functional relationships.

A particle filter is used as the Bayesian inference algorithm for the nonlinear and non-Gaussian DBN. The implementation of the particle filter for this dynamic Bayesian network (DBN) is nontrivial due to 1) the existence of static nodes, which are time-independent variables shared by all the Bayesian networks; and 2) state variables that may have parent nodes across two adjacent Bayesian networks. Therefore, this paper classifies the nodes in adjacent Bayesian networks into five groups to facilitate generating new particles based on the particle in the Bayesian network in the previous time instant. The generated new particles are used in Bayesian updating and help to realize the diagnosis.

Prognosis requires no Bayesian inference, and thus is computationally less demanding than diagnosis. In the case that the load is observed at each time step, theoretically Bayesian updating of the DBN is required at each time step, which implies large computational cost. This paper shows that the DBN can be modified under reasonable assumptions about the measurement error and load observation; as a result, the number of time steps requiring Bayesian updating of the DBN is reduced significantly, thus providing substantial savings in computational effort (61% saving in the numerical example).

All the preceding concepts are illustrated by a numerical example of fatigue crack growth on the leading edge of an aircraft wing. The results for this example show that the proposed framework has the capabilities to 1) track the evolution of time-dependent state variables (diagnosis); 2) reduce the uncertainty in time-independent state variables (diagnosis); and 3) probabilistically predict the crack growth in the future (prognosis).

Quantification of the prognosis uncertainty is necessary to assist decision making under uncertainty. The U.S. Air Force is currently investigating this methodology within the Airframe Digital Twin program for demonstration with U.S. Air Force legacy aircraft using full-scale experimental tests. The demonstrated P2IAT methodology is expected to be used for both legacy and new aircraft to reduce maintenance costs.

Acknowledgments

The research reported in this paper is supported by funds from the US Air Force Research Laboratory in Wright–Patterson Air Force Base, Ohio, under contract no. FA8650-14-D-2443 (Airframe Digital Twin Spiral 1 with Project Monitors: Douglas Henderson, Pamela Coburn, and Eric Teugel). This support is gratefully acknowledged. The authors are grateful to Guowei Cai and Saideep Nannapaneni at Vanderbilt University for assistance with computations and coding. The authors also thank Joshua Mullins at Sandia National Laboratories, Chen Liang at Ford Motor Company, and Youngwon Shin at GE Global Research for valuable discussions.

References

[1] Aktepe, B., and Molent, L., "Management of Airframe Fatigue Through Individual Aircraft Loads Monitoring Programs," *8th International Aerospace Congress*, ICMS Pty Ltd., Victoria, Australia, 1999, p. 5.

[2] Lee, H., Cho, H., and Park, S., "Review of the F-16 Individual Aircraft Tracking Program," *Journal of Aircraft*, Vol. 49, No. 5, 2012, pp. 1398–1405. doi:10.2514/1.C031692

[3] Lee, H., Park, S., and Kim, H., "Estimation of Aircraft Structural Fatigue Life Using the Crack Severity Index Methodology," *Journal of Aircraft*, Vol. 47, No. 5, 2010, pp. 1672–1678. doi:10.2514/1.C000250

[4] Glaessgen, E., and Stargel, D., "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles," *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, AIAA Paper 2012-1818, 2012, pp. 1–14. doi:10.2514/6.2012-1818

[5] Bond, R., Underwood, S., Adams, D. E., and Cummins, J. J., "Structural Health Monitoring-Based Methodologies for Managing Uncertainty in Aircraft Structural Life Assessment," *Structural Health Monitoring*, Vol. 13, No. 6, 2014, pp. 621–628. doi:10.1177/1475921714553733

[6] "Airframe Digital Twin Spiral 1," U.S. Air Force Research Lab., Broad Agency Announcement BAA-13-03-RQKP, Wright–Patterson AFB, OH, 2013.

[7] Jardine, A. K. S., Lin, D., and Banjevic, D., "A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance," *Mechanical Systems and Signal Processing*, Vol. 20, No. 7, 2006, pp. 1483–1510. doi:10.1016/j.ymssp.2005.09.012

[8] Bartram, G., and Mahadevan, S., "Integration of Heterogeneous Information in SHM Models," *Structural Control Health Monitoring*, Vol. 21, No. 3, 2014, pp. 403–422. doi:10.1002/stc.v21.3

[9] Friedman, N., Murphy, K., and Russell, S., "Learning the Structure of Dynamic Probabilistic Networks," *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publ. Inc., San Francisco, CA, 1998, pp. 139–147.

[10] Welch, G., Bishop, G., and Hill, C., *An Introduction to the Kalman Filter*, Chapel Hill, NC, 2000.

[11] Khorasgani, H., Biswas, G., and Sankararaman, S., "Methodologies for System-Level Remaining Useful Life Prediction," *Reliability Engineering and System Safety*, Vol. 154, Oct. 2016, pp. 8–18. doi:10.1016/j.ress.2016.05.006

[12] Julier, S. J., and Uhlmann, J. K., "A New Extension of the Kalman Filter to Nonlinear Systems," *Proceedings of the 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, edited by Kadar, I., 1997, p. 182.

[13] Fan, J., Yung, K., and Pecht, M., "Prognostics of Chromaticity State for Phosphor-Converted White Light Emitting Diodes Using an Unscented Kalman Filter Approach," *IEEE Transactions on Device and Materials Reliability*, Vol. 14, No. 1, 2014, pp. 564–573. doi:10.1109/TDMR.2013.2283508

[14] Al-Hussein, A., and Haldar, A., "Unscented Kalman Filter with Unknown Input and Weighted Global Iteration for Health Assessment of Large Structural Systems," *Structural Control Health Monitoring*, Vol. 23, No. 1, 2016, pp. 156–175. doi:10.1002/stc.1764

[15] Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T., "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, 2002, pp. 174–188. doi:10.1109/78.978374

[16] Sun, J., Zuo, H., Wang, W., and Pecht, M. G., "Application of a State Space Modeling Technique to System Prognostics Based on a Health Index for Condition-Based Maintenance," *Mechanical Systems and Signal Processing*, Vol. 28, April 2012, pp. 585–596. doi:10.1016/j.ymssp.2011.09.029

[17] van der Merwe, R., "Sigma-Point Kalman Filters for Probabilitic Inference in Dynamic State-Space Models," Ph.D. Dissertation, Oregon Health and Science Univ., Portland, OR, 2004.

[18] Mihaylova, L., Boel, R., and Hegyi, A., "Freeway Traffic Estimation Within Particle Filtering Framework," *Automatica*, Vol. 43, No. 2, 2007, pp. 290–300. doi:10.1016/j.automatica.2006.08.023

[19] Miao, Q., Xie, L., Cui, H., Liang, W., and Pecht, M., "Remaining Useful Life Prediction of Lithium-Ion Battery with Unscented Particle Filter Technique," *Microelectronics Reliability*, Vol. 53, No. 6, 2013, pp. 805–810. doi:10.1016/j.microrel.2012.12.004

[20] Roychoudhury, I., Biswas, G., and Koutsoukos, X., "Comprehensive Diagnosis of Continuous Systems Using Dynamic Bayes Nets," *Proceedings of the 19th International Workshop on Principles of Diagnosis*, 2008, pp. 151–158.

[21] Chen, H., and Chang, K. C., "K-Nearest Neighbor Particle Filters for Dynamic Hybrid Bayesian Networks," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 44, No. 3, 2008, pp. 1091–1101. doi:10.1109/TAES.2008.4655366

[22] Jianzhong, S., Hongfu, Z., and Pecht, M. G., "Advances in Sequential Monte Carlo Methods for Joint State and Parameter Estimation Applied to Prognostics," *Prognostics and System Health Management Conference*, IEEE Publ., Piscataway, NJ, 2011, pp. 1–7. doi:10.1109/PHM.2011.5939505

[23] Musso, C., Oudjane, N., and Le Gland, F., "Improving Regularised Particle Filters," *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001, pp. 247–271.

[24] Rosenblatt, M., "Remarks on Some Nonparametric Estimates of a Density Function," *Annals of Mathematical Statistics*, Vol. 27, No. 3, 1956, pp. 832–837. doi:10.1214/aoms/1177728190

[25] Duong, T., and Hazelton, M. L., "Cross-Validation Bandwidth Matrices for Multivariate Kernel Density Estimation," *Scandinavian Journal of Statistics*, Vol. 32, No. 3, 2005, pp. 485–506. doi:10.1111/sjos.2005.32.issue-3

[26] Paris, P. C., and Erdogan, F., "A Critical Analysis of Crack Propagation Laws," *Journal of Basic Engineering*, Vol. 85, No. 4, 1963, pp. 528–533. doi:10.1115/1.3656900

[27] Donahue, R. J., Clark, H. M., Atanmo, P., Kumble, R., and McEvily, A. J., "Crack Opening Displacement and the Rate of Fatigue Crack Growth," *International Journal of Fracture Mechanics*, Vol. 8, No. 2, 1972, pp. 209–219. doi:10.1007/BF00703882

[28] Yuen, B., and Taheri, F., "Proposed Modifications to the Wheeler Retardation Model for Multiple Overloading Fatigue Life Prediction," *International Journal of Fatigue*, Vol. 28, No. 12, 2006, pp. 1803–1819. doi:10.1016/j.ijfatigue.2005.12.007

[29] Shi, P., and Mahadevan, S., "Corrosion Fatigue and Multiple Site Damage Reliability Analysis," *International Journal of Fatigue*, Vol. 25, No. 6, 2003, pp. 457–469. doi:10.1016/S0142-1123(03)00020-3

[30] Ling, Y., and Mahadevan, S., "Integration of Structural Health Monitoring and Fatigue Damage Prognosis," *Mechanical Systems and Signal Processing*, Vol. 28, April 2012, pp. 89–104. doi:10.1016/j.ymssp.2011.10.001

[31] de Jesus, A. M. P., and Pereira, R. M. G., "FEM Analysis of Riveted Connections Aiming Fatigue and Fracture Assessments," *Proceedings of the Iberian Conference on Fracture and Structural Integrity*, Emerald, Porto, Portugal, 2010, p. 22.

[32] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, 2006, p. 152.

[33] Xu, P., Su, X., Mahadevan, S., Li, C., and Deng, Y., "A Non-Parametric Method to Determine Basic Probability Assignment for Classification Problems," *Applied Intelligence*, Vol. 41, No. 3, 2014, pp. 681–693. doi:10.1007/s10489-014-0546-9

[34] Sobol', I. M., "Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates," *Mathematics and Computers in Simulation*, Vol. 55, Nos. 1–3, 2001, pp. 271–280. doi:10.1016/S0378-4754(00)00270-6

[35] Li, C., and Mahadevan, S., "Relative Contributions of Aleatory and Epistemic Uncertainty Sources in Time Series Prediction," *International Journal of Fatigue*, Vol. 82, Jan. 2016, pp. 474–486. doi:10.1016/j.ijfatigue.2015.09.002

[36] Li, C., and Mahadevan, S., "An Efficient Modularized Sample-Based Method to Estimate the First-Order Sobol' Index," *Reliability Engineering and System Safety*, Vol. 153, Sept. 2016, pp. 110–121. doi:10.1016/j.ress.2016.04.012

[37] Li, C., and Mahadevan, S., "Global Sensitivity Analysis for System Response Prediction Using Auxiliary Variable Method," *17th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2015-0661, 2015. doi:10.2514/6.2015-0661

[38] Liu, Y., and Mahadevan, S., "Threshold Stress Intensity Factor and Crack Growth Rate Prediction Under Mixed-Mode Loading," *Engineering Fracture Mechanics*, Vol. 74, No. 3, 2007, pp. 332–345. doi:10.1016/j.engfracmech.2006.06.003

[39] Staszewski, W., Boller, C., and Tomlinson, G. R., *Health Monitoring of Aerospace Structures: Smart Sensor Technologies and Signal Processing*, Wiley, Hoboken, NJ, 2004, p. 201.

[40] "Standard Practices for Cycle Counting in Fatigue Analysis," ASTM International STD E1049-85, West Conshohocken, PA, 2005.

[41] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C., *Time Series Analysis: Forecasting and Control*, Wiley, Hoboken, NJ, 1976, p. 195. doi:10.1002/9781118619193

[42] Ljung, G. M., and Box, G. E. P., "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, Vol. 65, No. 2, 1978, pp. 297–303. doi:10.1093/biomet/65.2.297

[43] Oberkampf, W. L., and Barone, M. F., "Measures of Agreement Between Computation and Experiment: Validation Metrics," *Journal of Computational Physics*, Vol. 217, No. 1, 2006, pp. 5–36. doi:10.1016/j.jcp.2006.03.037

[44] Li, C., and Mahadevan, S., "Role of Calibration, Validation, and Relevance in Multi-Level Uncertainty Integration," *Reliability Engineering and System Safety*, Vol. 148, April 2016, pp. 32–43. doi:10.1016/j.ress.2015.11.013

[45] Li, C., and Mahadevan, S., "Uncertainty Quantification and Output Prediction in Multi-Level Problems," *16th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2014-0124, 2014. doi:10.2514/6.2014-0124

[46] Li, C., and Mahadevan, S., "Robust Test Resource Allocation using Global Sensitivity Analysis," *18th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2016-0952, 2016. doi:10.2514/6.2016-0952

R. K. Kapania
Associate Editor