Development of predictive structural maintenance strategies for aircraft using model-based prognostics

Yiwei Wang

Université Fédérale

Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Institut National des Sciences Appliquées de Toulouse (INSA Toulouse)
Discipline ou spécialité :
Génie Mécanique, Mécanique des Matériaux

Présentée et soutenue par WANG Yiwei
Le 14 Mars 2017

Sujet de la thèse:
Développement de stratégies de maintenance structurales prédictives pour aéronefs utilisant le pronostic à base de modèles

---

JURY:
Bruno CASTANIER, Professeur HDR. Rapporteur. Institut des Sciences et Techniques de l'Ingénieur d'Angers
Christophe BERENGUER, Professeur HDR. Rapporteur. Institut Polytechnique de Grenoble
Pierre VILLON, Professeur HDR. Présent. Université de Technologie de Compiègne Laboratoire Roberval
Laurent SAINTIS, Maître de Conferences. Examinateur. Institut des Sciences et Techniques de l'Ingénieur d'Angers
Johanna SENATORE, Maître de Conferences. Examinateur. Université Paul Sabatier
Christian BES, Professeur HDR. Directeur. Université Paul Sabatier
Nicolas BINAUD, Maître de Conferences, Co-directeur. Université Paul Sabatier
Christian GOGU, Maître de Conferences HDR. Examinateur. Co-encadrant. Université Paul Sabatier

---

Ecole doctorale : Mécanique, Energétique, Génie civil et Procédés (MEGeP)
Unité de recherche : Institut Clément Ader (ICA, CNRS UMR 5312)
Directeur(s) de Thèse :
(Directeur) M. Christian BES, Professeur HDR, Université Paul Sabatier Toulouse
(Co-directeur) M. Nicolas BINAUD, Maître de Conferences, Université Paul Sabatier Toulouse

DEVELOPMENT OF PREDICTIVE STRUCTURAL MAINTENANCE STRATEGIES FOR AIRCRAFT USING MODEL-BASED PROGNOSTICS

By

YIWEI WANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF TOULOUSE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF TOULOUSE

2017

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Christian BES for his continuous support to my PhD study and related research work, for his patience, kindness, and immense knowledge. In addition, I would also like to thank my two co-supervisors, Dr. Christian GOGU and Dr. Nicolas BINAUD for the motivation, and for all the comments and suggestions on all aspects of research provided during my entire length of study. The experience with such an amazing team did shape me to become a better person. I am grateful to them for enlightening me the first glance of research.

I would like to thank the two reviewers of my dissertation, Dr. Bruno CASTANIER and Dr. Christophe BERENGUER. Their insightful comments further improved my work. Besides, my sincere thank also goes to other committee members, Dr. Pierre VILLON, Dr. Laurent SAINTIS and Dr. Johanna SENATORE for their valued suggestions as well as for their hard questions, which incented me to widen and deepen my research work from various perspectives.

I would like to express my special thanks to Dr. Haftka and Dr.Kim, at the University of Florida, Gainesville, USA, for providing me an opportunity to work with them for three months. Their valuable insights and enlightening advice improved my academic level. Many thanks to their help from both aspects of research and life during my time in USA.

I would like to thank all my past and present colleagues and friends in the lab, who are from many different countries, Daniel, Varun, Dominique, Simon, Ernane, Jim, Kahina … ... They gave me a wonderful experience of multi-culture. The warm friendship drives away my loneliness in an alien land. I will not forget the endless discussion on planning hangout, on culture, on politics, on everything in the restaurant when we had lunch together every day. I will always memorize all the beautiful time we had together.

I would like to thank my parents for their understanding and support to my decision of pursuing a PhD degree abroad. As their only child, I know this is not easy for them. A simple thank will never be enough to what you all mean to me.

Last but not least, I would also thank my country, who provides me the financial support to pursue my PhD degree abroad.

ABSTRACT

Aircraft maintenance represents a major economic cost for the aviation industry. Traditionally, the aircraft maintenance is highly regulated based on fixed schedules (thus called scheduled maintenance) in order to ensure safety. The frequency of scheduled maintenance is designed to be very conservative to maintain a desirable level of reliability. Developing efficient maintenance can be an important way for airlines to allow a new profit growth. With the development of sensor technology, structural health monitoring (SHM) system, which employ a sensor network sealing inside aircraft structures to monitor the damage state, are gradually being introduced in the aviation industry. Once it is possible to monitor the structure damage state automatically and continuously by SHM systems, it enables to plan the maintenance activities according to the actual or predicted health state of the aircraft rather than a fixed schedule. This work focus on the fatigue crack propagation in the fuselage panels. The SHM system is assumed to be employed. A model-based prognostics method is developed, which enables to filter the noise of SHM data to estimate the crack size, and to predict the future health state of the panels. This predictive information is integrated into the maintenance decision-making and two types of predictive maintenance are developed. The numerical study shows that the predictive maintenance significantly reduces the maintenance cost by reducing the number of maintenance stop and the repaired panels.

KEYWORDS:

Fatigue crack growth, Predictive maintenance, Model-based prognostics, Extended Kalman filter, First-order Perturbation method

RESUME

La maintenance aéronautique est fortement régulée, notamment à travers l'établissement d'un planning de maintenance obligatoire, permettant de garantir la sureté structurale. La fréquence des arrêts en maintenance est déterminée de manière très conservative en vue d'assurer les exigences de fiabilité. Développer des stratégies de maintenance moins conservatives et plus efficaces peut alors représenter une voie pour une nouvelle croissance des compagnies aériennes. Les systèmes de monitoring embarqué de structures, sont progressivement introduits dans l'industrie aéronautique. Ces développements pourraient alors permettre de nouvelles stratégies de maintenance structurale basées sur la prévision de l'état de santé de chaque élément structural, plutôt que basée sur une maintenance programmée, tel qu'implémentée actuellement. Dans ce cadre général, ce travail se concentre sur le suivi par un système embarqué de la propagation de fissures de fatigue dans les panneaux de fuselage. Une nouvelle méthode de prévision des fissures basée sur des modèles de propagation est développée, qui permet de filtrer le bruit des mesures du système embarqué, identifier la taille actuelle de la fissure et prédire son évolution future et par conséquent la fiabilité des panneaux. Cette approche prédictive est intégrée dans le processus de maintenance structurale aéronautique et deux types de maintenances prédictives sont proposés. L'étude numérique montre que ces stratégies de maintenance prédictive peuvent réduire de manière significative les coûts de maintenance en réduisant le nombre d'arrêts en maintenance et le nombre de réparations inutiles.

MOTS CLES :

Propagation de fissures de fatigue, maintenance prédictive, filtrage de Kalman étendu, méthode de perturbation au premier ordre

CONTENTS

CHAPTER 1 INTRODUCTION

Aircraft fuselage structures are designed with the concept of damage tolerance, wherein small damage are allowed to remain on the aircraft, and damage that otherwise affect the safety of the structure are repaired. To ensure a desirable level of reliability, currently, the aircraft maintenance is highly regulated based on fixed scheduled, known as scheduled maintenance, during which the aircraft is sent to the maintenance hangar and partially disassembled to undergo a series of maintenance activities including both engine and airframe maintenance. Airframe maintenance that deals with non-structural items such as furniture and electronic systems is called non-structural airframe maintenance while the one that concerns the fatigue damage in the structural section such as fuselage panels is called structural airframe maintenance. Structural airframe maintenance is often implemented by techniques such as non-destructive inspection (NDI), general visual inspection (GVI), detailed visual inspection (DVI). The activities of NDI, GVI or DVI are time-consuming, leading the down time of scheduled maintenance up to one month. The frequency of scheduled maintenance for commercial aircraft is designed to maintain a low probability of failure, it is thus very likely that for most aircraft no large crack exists during earlier life of the aircraft. Even so, the intrusive inspections by NDI or DVI still need to be performed to guarantee the absence of critical cracks that otherwise cause fatigue failure.

With progress in sensor technology, data acquisition and storage techniques as well as advanced data processing algorithms, structural health monitoring (SHM) systems are gradually being introduced in the aviation industry. SHM employs a sensor network that is sealed inside aircraft structures for monitoring their damage state. The network may consist of a wide variety of sensor node types, strategically distributed in the structure to maximize data collection. The data is collected and processed using damage detection algorithms and feature extraction methods to obtain the damage state of the structure. Once it is possible to monitor the structure automatically and continuously by an SHM system, more advanced condition-based maintenance (CBM) can be implemented instead of the scheduled maintenance. CBM is defined by the maintenance being subordinated to an event and being triggered when some conditions are satisfied, *e.g.* the system state exceeds some thresholds. For structural airframe maintenance, CBM plans maintenance based

on the actual condition of the aircraft, rather than fixed schedules and inspection routines that might not be necessary, and thereby reduces aircraft's downtimes and save the maintenance cost.

Much attention has been paid to CBM strategies in the literature and more recently to predictive maintenance. CBM and predictive maintenance share some characteristics in common notably that both of them rely on the damage-associated data collected by the SHM system. The difference lies in that in CBM, the maintenance decision-making relies only on current damage level while the predictive maintenance makes use, in addition to current damage information, of the *prognostics index* to support the decision-making. The remaining useful life (RUL) is the most common prognostics index. The RUL-based predictive maintenance strategies aim at updating the maintenance schedule in real time according to the RUL of a component or system, and dynamically predicting the next maintenance time. Generally, the maintenance schedule is determined by the manufacture in concertation with safety authorities. Arbitrarily deciding the maintenance time only depending on the estimated RUL without taking into account the periodic maintenance time may however be too disruptive to the original maintenance plan. The arbitrary triggered maintenance is unexpected and less optimal from the economic point of view due to less notification in advance, *e.g.*, the absence of maintenance crew, the lack of spare parts, etc. In such cases, instead of using the estimated RUL to dynamically decide the next maintenance time, it would be more desirable to predict the probability that a component operates normally before some future time (*e.g.*, before next scheduled maintenance). In other words, use the "*reliability in future time*" as the prognostics index.

The implementation of predictive maintenance relies on prognostics, which focus on predicting the future degradation trend or the RUL of a system given different kinds of information available, *e.g.*, the degradation model, the real time monitoring data collected from the current system, the run-to-fail data from other similar systems, etc. The methods and techniques for implementing prognostics are diverse in terms of how to employ and process the above information. Researchers reviewed various prognostics methods from different point of view, and there is little consensus as to what classifications are the most appropriate for these methods. With the development of prognostics discipline, the opinion converges that the prognostics methods can be categorized broadly into data-driven approaches, model-based approaches. Data-driven methods use observed data to identify the characteristics of damage progress and predict the future state without relying

on any particular physical model. Data-driven methods show the strength when the understanding of the principles of system operation is not comprehensive or when the system is sufficiently complex such that developing an accurate model is prohibitively expensive. However, data-driven methods suffer the bottle neck of expensive and time-consuming collection of fault progression data and unsatisfactory long-term prediction. Model-based prognostics approach can provide superior performance than data-driven when the degradation model is available. The fundamental assumption is that the physical model describing the evolution or degradation of damage is known. Due to the availability of the degradation model, model-based prognostics show advantages in long-term prediction. If the degradation model is accurate then it completes prognostics because the future behavior of damage can be determined by progressing the degradation model in the future time. However, in practice, the degradation model is not complete, mainly due to lack of knowledge to model parameters. Even small uncertainties in model parameters could yield a large uncertainty in predicting the damage growth trend or the RUL. Therefore, the key issue in model-based prognostics is to identify the model parameters. This typically resorts to estimation algorithms. The most widely used estimation algorithms include particle filter, Bayesian method, Extended Kalman filter, and nonlinear least squares method. Once the model parameters as well as their uncertainty are estimated, they can be substituted into the degradation model to predict the distribution of RUL or the distribution of damage size in future time.

In summary, it is necessary to establish a link between prognostics and maintenance strategies so that the prognostics information can be incorporated into the maintenance decision-making, which enables more efficient and cost-saving maintenance strategies. This is the motivation of this thesis. Specifically, this work investigates the benefits that SHM systems bring to the structural airframe maintenance. However, compared to the manual inspections, the accuracy of SHM could be low. The SHM data is generally noisy due to the sensor limitation and harsh working conditions. This becomes a technique issue that preventing SHM system to be popular in practice. One challenge of this work is that how to accurately estimate the damage growth when the measured data contains noise, and in addition, how to use the noisy data to reduce the estimate uncertainty of damage growth model parameters, or in other words, to identify the damage model parameters. Estimating the damage state as well as the damage model parameters refers to as the *state-parameters estimation* problem.

Once the state-parameter estimation problem is addressed, the estimated states and the corresponding uncertainties can be used to infer the remaining useful life of the system. This is typically referred to as the *prognostics problem*. The second challenge in this work is to develop a robust, accurate and efficient prognostics method, which could predict the remaining useful life or the future damage behavior of the system. Moreover, it is necessary to establish a link between the developed prognostics method and the maintenance planning. The third challenge is to develop some predictive maintenance strategies, which incorporate the prognostics into the maintenance planning. The developed predictive maintenance strategies are expected to be able to improve the traditional scheduled maintenance and ultimately reduce the maintenance cost.

In this work, we focus on the prognostics of fatigue damage in fuselage panels and the predictive maintenance of structural airframe. Fatigue damage is one of the major failure modes of aircraft structures. Especially repeated pressurization and depressurization during take-off and landing cause many loading and unloading cycles which can lead to fatigue cracks in the fuselage panels. For this application, the model-based prognostics method is adopted since the fatigue damage models for metal materials have been well researched. In this work, the well-known Paris' law is employed to model the fatigue crack propagation. The two Paris' law parameters are assumed unknown. This makes sense since due to lack of knowledge, the true value of the Paris' law parameters are uncertain. This kind of uncertainty can be reduced by collecting relevant data based on embedded structural health monitoring (SHM) sensors. Filtering methods are then needed to be employed to estimate the crack size and Paris' law parameters from the noisy SHM data. Based on the estimated crack size and parameters, the evolution of crack size distribution can then be predicted and the reliability of the fuselage at a future time can be calculated. On the basis of the fuselage reliability at a future time, predictive maintenance strategies can be developed in order to reduce costs of the corresponding structural airframe maintenance. In summary, the objectives of this work are listed as follows.

(1) Study the behavior of fatigue crack propagation

(2) Investigate the techniques that are able to incorporate the noisy SHM data into the fatigue propagation model to reduce the uncertainty of the unknown model parameters as well as to give more precise estimate of the fatigue crack size.

(3) Develop a model-based prognostics method that is able to predict the future damage state of

individual component.

(4) Based on the model-based prognostics method, develop predictive maintenance strategies. Quantify the cost effectiveness of the predictive maintenance over the traditional scheduled maintenance. The outlines of the dissertation are as follows. CHAPTER 2 gives the general introduction to structural health monitoring, evolution of maintenance strategies and prognostics methods, with emphasis on predictive maintenance and model-based prognostics algorithms. The literature regarding the above subjects in the past decade is reviewed. Based on the literature review, the motivation and research objectives of this thesis are detailed.

CHAPTER 3 focuses firstly on the approaches for addressing the state-parameter estimation problem. The theory of Bayesian filter is presented followed by elaborations of three widely used filters, Extended Kalman filter (EKF), Unscented Kalman filter (UKF) and particle filter (PF). Then the Paris' law is employed as the fatigue damage propagation model. The damage propagation model and the noisy SHM data are modeled under filtering framework as a nonlinear filtering problem. The damage state and the unknown parameters of Paris' law are estimated by the above mentioned three filtering methods. The performances of the three filters are then compared.

CHAPTER 4 proposed a model-based prognostics method that couples Extended Kalman filter and a first-order perturbation method. The developed prognostics method is composed of two sequential steps. In the first step, the EKF is used to estimate the damage state and Paris' law parameters. Based on the information given by EKF, *i.e.*, the mean and covariance of estimate, in the second step, a first-order perturbation method is developed to predict the damage growth in future time. By using the first-order perturbation method, the first and second moments of damage state distribution in any future flight cycle can be computed analytically.

CHAPTER 5 focuses on developing the predictive maintenance. Based on the model-based prognostics method that proposed in CHAPTER 4, two types of predictive maintenance strategies are developed. Both of them employ predictive information in maintenance planning. The main difference lies in the repair policy in which the predictive information is used. By comparing with a series of other maintenance, the proposed two types of predictive maintenance show their strength: under the premise of ensuring safety, they request fewer maintenance stops and repair fewer panels, thus save maintenance cost.

CHAPTER 6 gives a full summary, draws conclusions and open future perspectives.

CHAPTER 2 LITERATURE REVIEW

2.1 Structural health monitoring system

Structural health monitoring (SHM) is a type of infrastructure monitoring that employs a sensor network to monitor the physical integrity of a structure. The network may consist of a wide variety of sensor node types, strategically distributed, about the structure to maximize data collection. Lynch [1] gave a summary review of wireless sensors and sensor networks for structuring health monitoring. Network data is collected by the end user and statistically analyzed, using specialized damage detection algorithms and feature extraction, to determine if the structure is performing as intended, or if it has become damaged over time. By observing the data trends provided by an SHM system, the end-user can detect minor structural changes that would otherwise not be discernible by traditional inspection techniques. SHM data analysis has proved crucial in determining the overall safety of structural elements in bridges, tunnels and dams, as well as providing for early warning systems that can alert system operators prior to a catastrophic failure. It is also being applied to aeronautical and automotive monitoring, especially on structural components such as airframes, fuselage panels, landing gears, rotor blades and bulkheads for both fixed wing and rotor wing aircraft. Diamanti et al. [2] investigated the SHM techniques in application of aircraft composite structures. A review of currently used inspection methods was presented and some examples were described where Lamb wave based scanning techniques have been used to identify internal damage in multi-layered composite structures. A complete SHM system mainly consists of hardware and software parts.

Hardware:

(1) Sensor Network: A variety of sensor types (strain, temperature, pressure, vibration, acceleration, etc.) and may include a combination of both wired and wireless sensor nodes, and energy harvesting power generation.

(2) Data acquisition, storage and transfer systems: Embedded hardware modules and software systems that collect, store, process and transmit sensor data to the required end point.

Software:

(1) Data Management System

Data management system refers to personal computers or cloud computing with an interface that allows the end user to manipulate data and apply various statistical analysis techniques. Data management capabilities may include:

- Viewing sensor node data
- Sending commands to the SHM network
- Statistically analyzing data
- Assessing condition of structure
- Assessing condition of a structural model
- Predicting remaining service life of structure

(2) Statistical analysis algorithm

Since sensor nodes themselves cannot determine if actual structural damage has occurred, the data collected by the node array must be converted into meaningful information via statistical analysis. By properly analyzing the SHM data, "damage sensitive features" can be extracted to determine the current state of a structure's health using machine learning and artificial intelligence techniques. Such algorithms include:

Supervised Learning (Baseline): Algorithms that are applied when SHM data is available from both the undamaged, and damaged, parts of a structure. Group classification and regression analysis are examples of algorithms applied in this situation.

Unsupervised Learning (Non-Baseline): Algorithms that are applied when undamaged structure data is not available, for comparison to damaged structure data. Outlier, or novelty detection, is the primary class of algorithms applied in this situation.

In this work, the objective is to monitor the fatigue damage on the aircraft's panels. Thus the techniques for fatigue damage detection are of particular interest. The most widely used technique for damage detecting in SHM system is the application of acoustic-ultrasonic and guided ultrasonic waves. These techniques are ultrasonic waves introduced to a structure at one point and sensed at a different location. It appears that Lamb waves are the most widely used guided wave for damage detection and monitoring. Generation and sensing of Lamb waves can be accomplished using various transducers. These transducers include: wedge comb or phased array transducers, piezoelectric (PZT) ceramics, rubbers and paints, Smart Layers, interdigital transducers, optical

fiber sensors and micro-electro mechanical systems (MEMS). Some researches on application of Lamb wave for aircraft structure damage detection are detailed below.

Zhao et al. [3] employed the guided wave PZT sensor/actuator network technique to monitoring the defect on an aircraft wing. An E-2 surveillance aircraft wing section was cut into pieces for experimental studies. A relatively sparse PZT array with a diameter of one foot was bounded to the inner surface of the wing to generate and receive ultrasonic guided wave. Some simulated defects like loose rivets, rivet-hole cracks and material loss on the wing panel were studied. To interpret the signals collected from the PZT disc array, a correlation analysis based algorithm called RAPID (reconstruction algorithm for probabilistic inspection of defects) was developed for defect detection, localization and growth monitoring. Ihn et al. [4] developed a piezoelectric based built-in diagnostic technique for monitoring fatigue crack growth in metallic structures. The technique uses diagnostic signals, generated from nearby piezoelectric actuators built into the structures, to detect crack growth. Ihn et al. [5] continued the work based on their previous work that a built-in diagnostic technique for monitoring hidden fatigue crack growth in aircraft structures has been developed. In this study, the proposed diagnostic technique was applied to monitor fatigue crack growth in riveted fuselage joints and a cracked metallic plate repaired with a bonded composite path. A complete built-in diagnostic system for the tests was developed, including a sensor network, hardware and the diagnostic software. The damage index successfully detected both crack growth and debones damage for the structures considered. Leong [6, 7] used a commercial laser vibrometer, designed for vibration/modal analysis can be used for crack detection in metallic structures. The study involves a simple fatigue test in order to initiate and grow a crack. Lamb waves generated by one bonded piezoelectric transducer were sensed using a multi-point scanning laser vibrometer. Besides the technique talked above, there are some other techniques for monitoring the fatigue damage. Ignatovich et al. [8] developed a fatigue sensor made of alclad aluminium alloy D16AT. In the developed sensor the necessary raising of sensitivity is achieved by the redistribution of stress due to the corresponding distribution of stiffness along the length of the sensor.

The application of SHM systems in commercial aircraft is still at the research stage and its widespread application to airlines has a long way to go. Tests have been done however during the last decades by airlines as well as research centers. For example, very recently, seven of the Boeing Co.’s 737 narrow body aircraft operated by Delta Air Lines have been outfitted with Comparative

Vacuum Monitoring sensors for crack detection in a program that aims to obtain approval for SHM as an alternative inspection method. Part of a broader SHM initiative at the Airworthiness Assurance Nondestructive Inspection Validation Center operated by Sandia National Labs for the US Federal Aviation Administration (FAA). In addition, the major aircraft OEMs as well as operators, regulators and technology suppliers have been striving for years to standardize SHM integration and certification requirements and to mature system for widespread use.

2.2 Maintenance strategy evolution and predictive maintenance

The evolvement of maintenance strategy has experienced four different types, *i.e.*, corrective maintenance (CM), preventive maintenance (PvM), condition-based maintenance (CBM) and predictive maintenance (PdM). The evolution is illustrated in Figure 2-1.

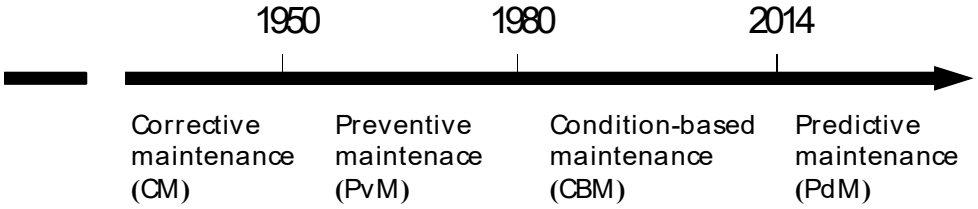

Figure 2-1 Evolution of maintenance strategies

The CM means that the system is allowed to operate until a breakdown or an obvious fault, thus CM is also referred to as run-to-fail strategy. Due to the nature of unpredictable system failure, it is difficult to anticipate when the manpower and spare parts are needed for repairing the system, leading to the difficulty of efficient planning of staff and resources. The cost associated with this strategy can be significant. These costs include production cost, downtime cost, in addition to inventory cost since the maintenance team need to hold spare parts to accommodate the sudden breakdown. Therefore, the CM is unsuitable to the cases where the system breakdown cause huge economic lost or even safety risk. On the other hand, CM has its advantages. It is easy to implement due to its simplicity and needs very low planning requirements since it does not need to be scheduled in advanced. The CM makes sense for the cases where the maintenance cost after system breakdown is less than that of performing other maintenance strategies.

The research regarding to PvM could be traced back to 1960s. McCall [9] and Sherif and Smith [10] used the term of “preventive maintenance model” which assumed that the state of equipment subject to stochastic failure is always known with certainty [11]. PvM performs the maintenance

regularly based on a pre-determined schedule [12, 13], thus also known as time-based maintenance or scheduled maintenance. Planning is the biggest advantage of PvM over CM. Through planning, all the cost of unplanned strategy including production lost, higher costs for parts and shipping, as well as time lost corresponding to emergencies can be reduced. Any required parts, supplies and manpower can be gathered in advance to minimize the time taken for maintenance. Safety risk is also decreased significantly since the system breaks down much less often through implementing PvM. On the other hand, PvM has its own shortcomings. The maintenance schedule of PvM in industry is usually implemented through either experiences or recommendations of original equipment manufacturer (OEM) rather than the information of actual system state. The PvM intervals based on OEM recommendations may not be optimal because actual operating conditions may be very different from those considered by the OEM [14]. In order to maintain a desirable level of system reliability, PvM often has a high maintenance frequency and a fixed maintenance interval. In some cases where the damage mechanism of system is dynamic and the degradation rate changes over time, the fixed schedule is too inflexible to adapt the degradation rate. For example, in the aviation industry, the fatigue damage in the fuselage panel is one of the most important failure modes. The fatigue crack in the panel propagates exponentially: very slowly during the first thousand flight cycles and then faster and faster. For the short-range commercial aircraft such as A320, the maintenance schedule is planned such that the first maintenance is done after 20000^{th} flight cycles and the subsequent maintenance is every 4000 cycles until the aircraft's end of life. The schedule is designed conservativly to maintain safety, which results in a fact that no important damage is detected and no maintenance is required during the earlier life of the aircraft. Even so, the aircraft still needs to undergo the routine inspection through the time-consuming technologies, increasing the downtime and the maintenance cost.

Condition-based maintenance (CBM) was introduced in 1975. CBM monitors the health state of systems continuously or as frequently as needed and triggers maintenance when some predefined conditions are satisfied. To the author's best knowledge, Baldin was the first researchers who began to explore the condition-based maintenance and published a number of papers from mid-1970s to mid-1980s [15-19]. Probably because of limits of monitoring techniques, the development of CBM was restricted during 1990s and not many research papers could be found in that period. With the great progress of condition monitoring techniques, data acquisition and

processing techniques, the research of CBM grows fast since 2000 especially in the last 5 years.

Although CBM takes advantage of the known state of the system, the threshold-based maintenance is not always an optimal solution. Determining the thresholds that guarantees safety under different kinds of uncertainties can be a tough work. It requires the experience of reliability experts as well as some robust and accurate algorithm, which might be computational costly. In addition, the pure CBM, *i.e.*, planning the maintenance only based on the actual health state of the system and trigging maintenance anytime when needed without taking into account the preventative maintenance, leads to unscheduled maintenance, which is costlier due to less notification in advance. Recently, a lot of attention moves to predictive maintenance (PdM). In many literatures, CBM refers to as PdM. Indeed, both CBM and PdM rely on the condition monitoring to the system, thus rely on sensor technique, data acquisition and storage technique etc. to know the damage state of the system. However, after being aware of the health state, the policy of planning maintenance scheme in CBM and PdM is different. Zhou [20] presented the concept of condition-based predictive maintenance that integrates the prediction tools into CBM to provide the assessment and prediction of the system hazard rate based on the collected information through continuous monitoring, with the aim of determining the required maintenance action prior to any predicted failure. This concept is very close to current idea of PdM and could be seen as a transition from CBM to PdM. According to [21, 22], the big difference between CBM and PdM is that CBM only uses current system state information to make the maintenance policy while by contrast, PdM makes use, in addition to current degradation information, of predictive information in the form of remaining useful life or the predicted damage index distribution to optimally schedule maintenance actions. With PdM, it is possible to predict the future degradation trajectory of the system thus to predict the possible time when the monitored damage index will reach or exceed a threshold. In that case, the staff could plan the maintenance actions in advance. There are a lot of literatures that study the predictive maintenance from different aspects. The following paragraphs will review the research paper of PdM from policy level, condition-monitoring level and system level.

From policy level, Horenbeek [21]] presented a dynamic predictive maintenance strategy that is able to update the maintenance schedule as the new information on the degradation and remaining useful life of the system become available. Traore [22] considered a complex and dynamic system.

The proposed Failure Mode Effect and Criticity Analysis (FMECA) method was used to identify the key component whose failure caused a long downtime of the system. An architecture build of supervision and prognostics was developed for predictive maintenance. Nguyen et al. [23] also considered a multi-component system with complex structure. A predictive maintenance policy with multi-level, the system level and the component level, was developed. The decision rules at the system level was to address if preventive maintenance was needed based on the predictive reliability of the system and the ones at component level aimed at identifying optimally a group of several components to be maintained. Curcuru et al. [24] proposed a sensor-driven predictive maintenance policy that determined the time at which the decision must be taken and the date for the starting of the maintenance activity. The degradation process was modelled using an autoregressive model with drift. The monitoring system was considered imperfect and the influence of this uncertainty on the performance of the proposed predictive maintenance was analyzed. Supported by huge volume of historic data in combination of failure data, maintenance action data, inspection schedule data, train type data and weather data, Li et al. [25] developed machine-learning approaches to predict impending failures and alarms of critical rail car components to avoid service interruptions and to increase the rail network velocity. Bansal [26] proposed a reverse algorithm to improve the estimation of the machine system parameters from the motion current signature based on non-linear time series techniques, thus improving the predictive maintenance system. Ferrerio et al. [27] applied Bayesian network as the technique for prognostics and the developed prognostics tool was integrated in a framework aiming at making final decision on airplane maintenance actions. The proposed framework was applied to predict brake wear on plane as a use case. Tan et al. [28] developed a framework for predictive maintenance-based scheduling of multi-state system, which can give the maintenance schedules based on the failure time of the overall system as estimated from its degradation trends. Berenguer et al. [29] proposed a predictive maintenance policy for a continuously deteriorating system with two failure mechanisms, an excessive deterioration and a shock. An approach combining statistical process control (SPC) and condition-based maintenance (CBM) is proposed to optimize the maintenance policy. SPC was used to monitor the stress covariate and CBM was used to inspect and replace the system according to the observed deterioration level. Kaiser [30] presented a sensor-updated degradation-based predictive maintenance policy that combined the component's real-time

degradation signals, acquired through condition monitoring, with degradation and reliability characteristics of the component's population to predict and update the residual life distribution. With the aid of the developed stopping rule, the maintenance routine was scheduled based on the most recently updated residual life. Zhao et al. [31] developed a predictive maintenance policy based on process data, which utilized the results of probabilistic fault prediction that revealing the evolvement of the system's degradation. The proposed two-stage probabilistic fault prediction method was based on principle component analysis (PCA) and Bayesian Auto-regression (BAR) model. For the off-line stage, the PCA was performed to the historic normal process data to get system's monitoring index, *i.e.*, the combined index. In the on-line stage, BAR model was used to get the prediction density of the transformed index in the future time.

Since the condition-monitoring techniques and the system development of predictive maintenance is not the main concern of this paper, the literature of these two aspects will be briefly discussed. The most common condition monitoring techniques used in predictive maintenance include infrared thermography, ultrasound analysis, acoustic analysis, oil analysis, motor current signature analysis, vibration analysis, etc. Huda et al. [32] investigated the application of infrared thermography technology as predictive/preventive maintenance to identify the presence of thermal defect in electrical equipment. The proposed technique used multiayered perceptron network, statistical features and discriminant analysis classifier to characterize the thermal status of hotspot into 'defect' and 'no defect' categories. Orhan et al. [33] studied the vibration monitoring for defect diagnosis of rolling element bearing as a predictive maintenance tool. Hashemian et al. [34] investigated the function of wireless sensors during the implementation of predictive maintenance to the High Flux Isotope Reactor at the Oak Ridge National Laboratory. Carnero et al. [35] proposed a model to help select the diagnostic techniques and the instrumentation in a predictive maintenance program. From the system development level, Efthymiou et al. [36] proposed an integrated predictive maintenance platform consisting of three parts. The first part, data acquisition and analysis, was responsible for data extraction and processing. The second part, knowledge management, focused on maintenance knowledge modelling and representation. The third part, maintenance dashboard, provided advisory capabilities on maintenance planning. Chen et al. [37] proposed an aircraft maintenance decision system based on real-time condition monitoring in which the data link is used to transmit the aircraft condition monitoring information in real time to

the maintenance information process terminal on ground station to complete the automatic diagnostic of aircraft equipment or parts. Garcia et al. [38] developed a software application called Intelligence System for Predictive Maintenance (abbreviated as SIMAP). SIMAP incorporated the real-time information from different sensors and other information sources to early detect possible anomalies and thus to optimize and dynamically adapt maintenance calendar. Bansal et al. [39] proposed a preventive maintenance system based on a neural network approach. Carnero [40] proposed an evaluation system aiming at carrying out the decision making in relation to the feasibility of setting up a predictive maintenance. Lee et al. [41] integrated industrial big data analytics and cyber-physical systems for future maintenance & service innovation. The authors proposed a systematic architecture for applying cyber-physical systems in manufacturing industry.

2.3 Prognostics methods

As mentioned preciously, prognostic is the prerequisite of the predictive maintenance and it is necessary to establish a link between predictive maintenance and prognostics methods. Prognostics originally come from the medical field and now it has developed into an independent engineering discipline. Prognostic focuses on predicting the time at which the objective system or component can no longer perform its desired function. The predicted time is then called remaining useful life (RUL), which is an important concept for decision making. Motivated by reducing the system down time and thus saving maintenance cost, prognostics received increasing attention and have been widely employed in many industry domains. These domains are mainly heavy industry including manufacture, heavy vehicles, mining, power generation, railway, aerospace and defense etc., where the system failure will lead to huge economic losses or even catastrophe.

There are various methodologies and techniques in prognostics. Many researchers reviewed these prognostics methods from different points of view. Unfortunately, there is little consensus among reviews of the prognostics field as to what classifications are most appropriate for grouping remaining useful life prediction models. With the development of prognostics discipline, the opinion for the classification of prognostics techniques among researchers converges, that is, the prognostics approaches can be categorized broadly into data-driven approaches, model-based approaches, and hybrid approaches, each of which has sub-classifications. A family of prognostics

approaches is illustrated in Figure 2-2.

Some review papers will be firstly investigated in this paragraph. Sikorska [42] and An [43] gave an overall review of prognostics methods. Although following a different classification, the two above mentioned reviews gave comprehensive summaries that include almost all prognostics approaches mentioned in Figure 2-2 with investigation on the strength and weakness of each type of approach, which provided good references for young researchers first considering options for remaining useful life prediction. [44-46] reviewed the prognostics methods in application of machinery systems. Si [47] has focused on the statistical data driven approaches in application of remaining useful life prediction. The authors classified the observed condition monitoring (CM) data into direct CM and indirect CM data. Based on this classification, they classified the statistical methods into models based on direct CM data and models based on indirect CM data. The former one include regression-based model, wiener process model, Gamma process model and Markovian-based model, among which, the regression, Wiener process and Gamma process models are used to address the continuous process and the Markovian-based model is used to process the discrete space. Models based on indirect CM data is classified into 3 types which are filtering type of models, covariate-based hazard model and hidden Markov model (HMM). For filtering type of models, the most commonly used are Kalman filter approach and stochastic filter methods such as particle filter and multiple-model filter. An [30] studied and discussed the difference in estimating model parameters among different prognostics methods: the particles filter, the overall Bayesian methods and the incremental Bayesian methods, all of which are based on the same theoretical foundation, Bayesian inference. Jouin et al. [48] focused on the particle filter branch among the numerous prognostics methods. The authors argued that the particle filter methods had ever become a state-of-the-art technique for prognostics and around 50 papers dealing with prognostics based on particle filter can be found in the literature in very recent years. This review paper introduced the background and a short review about the particle filter in a general context and then analyzed the particle filter in prognostics applications. Finally, it highlighted the remaining issues and challenges as well as gave solution proposals.

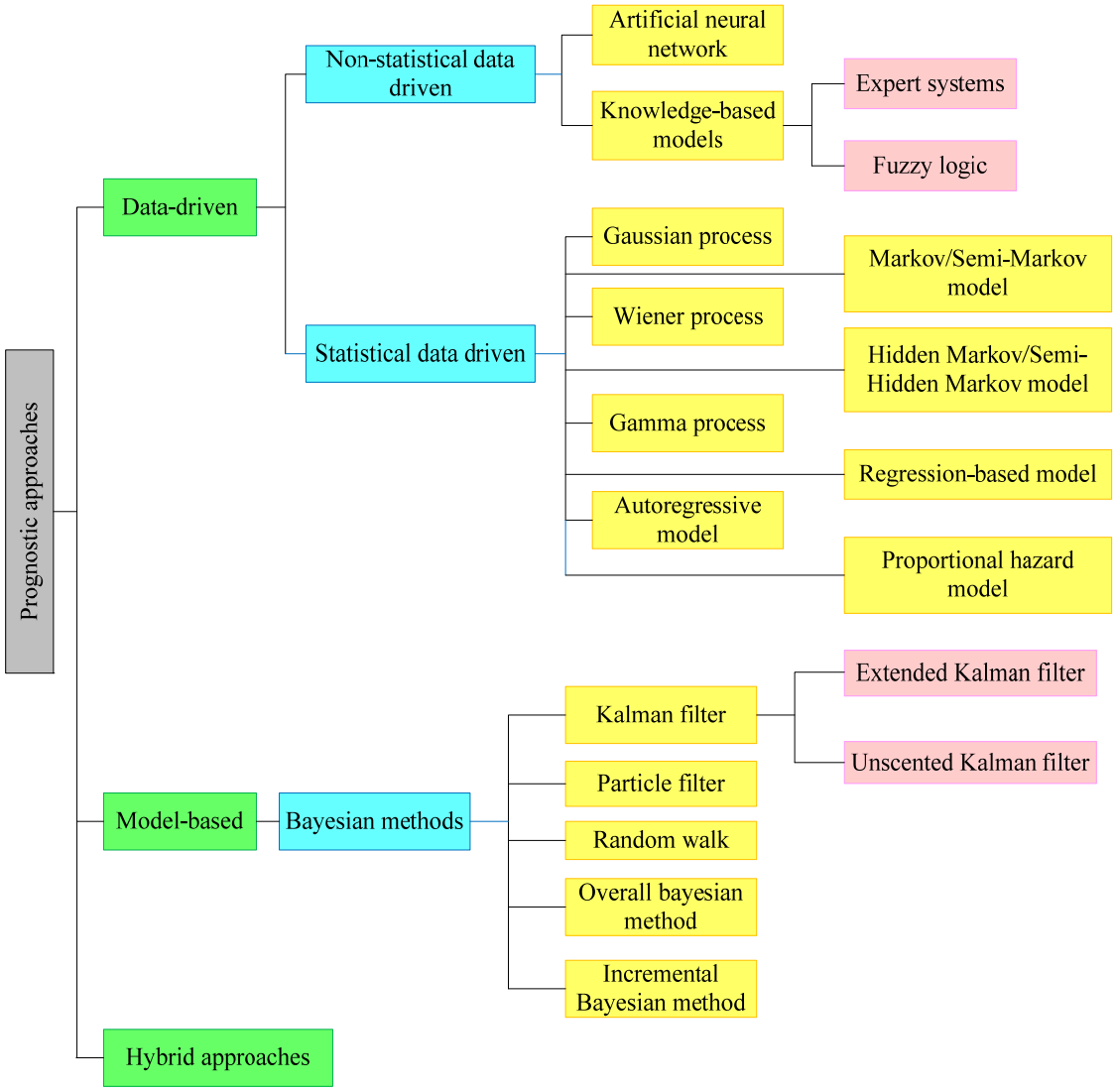

Figure 2-2 Family of prognostics approaches

The application of the above-mentioned prognostics methods will be studied in the following parts in terms of the order of data-driven, model-based and hybrid approaches.

Data-driven methods use information from observed data to identify the characteristics of damage progress and predict the future state without using any particular physical model. Modelling the evolution of the degradation process relies exclusively on process history data. The mathematical models describing the behavior of the degradation process are determined based on training data that are obtained under various usage conditions. Hu et al. [49] proposed an ensemble data-driven approach that combines multiple member algorithms with a weighted-sum formulation. The presented method has overcome the shortcomings of the traditional method: the selected standalone algorithm may not be robust and it wastes resources for constructing the algorithms that are discarded. Mosallam et al. [50] proposed a two phases data driven approach for RUL prediction.

In the offline phase, the method finds variables conveying the information about the degradation behavior and the selected variables are constructed to represent the degradation as a function of time. The constructed function is saved in the offline database. In the online phase, based on the saved data, *k-nearest* neighbors (k-NN) classifier is used as a RUL predictor. Son et al. [51] presented a probabilistic method that combined wiener process with Principal Component Analysis to model the deterioration of the components and to estimate the RUL on a case study.

Model-based approaches attempt to make use of the physical understanding of the system, *i.e.*, the physical damage model of the system is assumed available. The damage index is not observable directly but can be measured by some devices. The devices give a series of measurement data of some physical quantities that related to the system's damage index. The relation between the measured physical quantities and the damage index is described by a measurement model. Eq. (2-1) and Eq. (2-2) show the physical model and the measurement model in general case. $\mathbf{x}_k$ is the state vector containing the relevant information required to describe the system under investigation, or, in prognostics context, the damage index. $\mathbf{z}_k$ is a measurement vector representing the noisy observation that are related to state vector, generally having the dimension equal or lower than the state vector. f is the physical model, which is a possible non-linear function. $\boldsymbol{\theta}$is the model parameters. h is the measurement model, generally a possible non-linear vector function. w and v are the possible process noise and measurement noise. The physical model combined with measured data is employed to predict the future behavior of degradation and to predict the RUL. The behavior of the damage model depends on the model parameters. Once the damage model is determined, model parameter identification becomes the most important issue. The Bayesian family approaches are regarded as powerful tools for parameter identification problem. Some examples using Bayesian method to implement the prognostics are found in [52-55]. Particle filter is the most popular one. Particle filter is a kind of sequential Monte Carlo method that provides a numerical solution for Bayesian inference. [56-58] are some cases using particle filter as the prognostics method in application of fatigue damage model and Lithium-ion battery degradation model.

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \boldsymbol{\theta}_{k-1}, w_{k-1}) \tag{2-1}$$

$$\mathbf{z}_k = h(\mathbf{x}_k, v_k) \tag{2-2}$$

Both data-driven and model-based prognostics methods have their strength and weakness.

Model based approaches can implement long-term prediction accurately. But in practice, it may be difficult to develop an accurate damage model especially in some complex systems where the degradation mechanism is not fully understood. In addition, in many real cases, due to the inaccessibility (*e.g.* piping of deep water offshore well drilling plants) or the measurement difficulty (the crack depth of bearings on rotating machinery during online operation) , a measurement model is hard or impossible to get. Data-driven methods, on the other hand, do not rely on the physical model but on the historic data, such as run-to-fail data, which in practice is hard to collet. Hybrid approaches are presented to leverage the strength from both data-driven approaches as well as model-based approaches. In fact, it is rare that the fielded approaches are completely either purely data-driven or purely model-based. Baraldi [59] presented a hybrid method that combines particle filter with neural network to address the non-available measurement model problem in model-based approach. In the proposed method, a dataset containing pairs of "state-measurement" is available. The dataset is used to train a bagged ensemble of Artificial Neural Networks, which is embedded in the particle filter as empirical measurement model. Chen et al. [60] proposed an integrated RUL prediction method using adaptive neuro-fuzzy inference systems (ANFIS) and high-order particle filtering. The ANFIS is trained and integrated in a high-order particle filter as a model describing the fault progression while the particle filter is used to estimate the current state and carry out the *p-step-ahead* prediction. Yu [60] developed a prognostics system aiming to integrate the strength of both model-based and data-driven approaches. The proposed system first obtains a failure risk probability-based health indication during the operation of the machine by deriving a Bayesian-inference probability based on a baseline self-organizing map. Then based on the obtained health indication values in time series flow, a high-order Markov process is applied to model the machine health propagation.

2.4 A survey of five most widely used model-based prognostics methods

As mentioned in Section 2.3, in model-based prognostics, the model parameters are generally unknown and need to be identified along with the damage state from the noisy measurements. Once the model parameters are identified (usually with some uncertainty), they are substituted into the damage model to predict future behavior of damage and thereby to obtain prognostics information

such as the distribution of remaining useful life, or the evolution of damage distribution. This process can be implemented either through Monte Carlo sampling method or derived in an analytical form [61, 62]. It can be seen that in model-based problems, state-parameter estimation is the most important issue, which typically resorts to some estimation algorithms. Five state-parameter estimation algorithms appeared most commonly in literature include Bayesian method, particle filter, Extended Kalman filter, Unscented Kalman filter and nonlinear least squares method are introduced. The first four methods are based on Bayesian inference, while the last one is a conventional regression method. It is worth mentioning that the estimation methods can also be classified in terms of estimation criteria. The most commonly used estimate criterions for Bayes estimator include maximum a posteriori (MAP) and minimum mean square error (MMSE) criterion while the non-Bayesian criterion involves least squares criterion, maximum likelihood criterion, etc. BM and PF employ the maximum a posteriori criterion while EKF is a minimum mean square error estimator. The advantages of the above-mentioned algorithms over point estimation methods such as maximum likelihood estimate lies in their capability of estimating the uncertainty at the same time while giving the expected value of the identified parameters.

2.4.1 Extended Kalman filter

The EKF is commonly used for recursive nonlinear state-parameter identification due to its excellent filtering properties. A central and vital operation performed in EKF is the propagation of a Gaussian random variable through the system dynamics [63]. In the EKF, the state dynamics model is expanded as a Taylor series around the prior mean of state variables. By ignoring second and higher order terms, the state prediction propagates analytically through the nonlinear system equation whilst the state error covariance propagates through a separate first-order linearization of the nonlinear system [64]. Due to this linear approximation, EKF introduces errors from the second order in the true posterior mean and covariance of the transformed state variable, which may lead to sub-optimal results when dealing with significant nonlinearities. However, despite these approximations, from a practical application perspective, EKF algorithm remains a powerful tool in the nonlinear system state estimation domain and has been successfully used in various engineering state-parameter identification problems [65-68]. D'Alfonso et al. [69] used the EKF to estimate the

position and orientation of a mobile robot, called Khepera III, a battery-powered mobile robot with two independent driving wheels. 5 ultrasonic sensors for medium-range detection (from 25mm to 4m) were equipped in the robot. The data collected from these sensors were used to estimate the position and the orientation of the robot. Kim et al. [70] employed the EKF to estimate the hydrodynamic coefficients (specifically the 15 linear damping coefficients) of the AUV-SNUUV I, an autonomous underwater vehicle developed by Seoul National University. During the experiment, the input/output data were measured for SNUUV I in towing tank. The measured output signals were stored in onboard computer during experiment and these were transferred to the host computer through WLAN (Wireless Local Area Network). The measured data were used to estimate the hydrodynamic coefficients. The results showed that after the initial transition period, all the 15 linear damping coefficients converge to some stable values. Bressel et al. [71] used the EKF to estimate the state health of proton exchange membrane fuel cell (PEMFC). The considered PEMFC is a commercially available five cells stack with an area of 100 cm^2 and with a nominal current of 60 A. This stack was operated under a constant load (60A) for 1500 h. The stack voltage was collected as the measurement data, from which the health state of the PEMFC was estimated. The experiments showed that the EKF offered good results in the health state estimation of the PEMFC.

2.4.2 Unscented Kalman filter

The UKF has been proposed by Julier and Uhlmann [64] in 1990s as a theoretically improved alternative to EKF for calculating the statistics of a random variable which undergoes a nonlinear transformation[72]. In UKF, a set of points (called sigma points) are chosen according to a specific selection algorithm so that the mean and covariance of these sigma points are equal to the state posterior mean and state error covariance. Each of these points propagate in turn through the nonlinear system equation to yield a set of transformed points of which the statistics are calculated as the prior estimates of the transformed state variable. Julier and Uhlmann demonstrated theoretically that compared with EKF, the state value estimated by UKF agreed with the true value up to the third order and errors were introduced in the fourth and higher order terms. Both EKF and UKF predicted the state error covariance up to the second order, but the absolute errors in the

fourth and higher order terms for UKF were smaller than those of EKF. The UKF has received great attention since it was presented and several studies compared it with EKF. VanDyke et al. [73] applied UKF to implement spacecraft attitude state-parameter estimation and compared the results with EKF. They argued that their UKF consistently outperform the EKF. Crassidis et al. [74] also considered UKF for a spacecraft attitude estimation problem and claimed that for realistic conditions, especially when large initialization errors were present, UKF performed better than EKF. Qu et al. [75] used UKF for process monitoring and parameter estimation and argued that UKF outperformed EKF when severe nonlinearities exist and the measurement noise levels were high. However, although UKF has been proved to be a theoretical better approach to EKF, some recent research indicate that UKF often shows only a slight improvement and remains comparable in performance to EKF in some cases especially when the system is only moderately nonlinear. D'Alfonso et al. [69] employed both the EKF and UKF to estimate the position and orientation of a mobile robot. They ran their experiments 20 times to evaluate the performance of the EKF and the UKF, and found that the two filters performed comparably well in constructing the position of the robot and they attributed this to the fact that the nonlinearities in the model were not bad enough to highlight any substantial difference. Chowdhary and Jategaonkar [66] employed UKF and EKF to implement aerodynamics parameter estimation from flight data. Their results indicated that with a nonlinear model, no great difference between the numerical values of parameters was seen and UKF showed little improvement in time to convergence as compared to EKF. Wendel et al. [76, 77] compared EKF and sigma point Kalman filter for nonlinear problem of tightly coupled GPS/INS integration and pointed out that, due to the fact that the higher-order transformation terms were negligible in the GPS/INS integration, these two methods offered an identical performance, which inferred that a modification of existing EKF-based navigation systems may not result in significant performance improvements. Qu et al. [75] applied UKF to a nonlinear process and drew similar conclusion: both UKF and EKF performed comparably well and the difference in the results achieved by these two estimators was minor when the system did not exhibit a strong extent of nonlinearity. Therefore, some researchers hold a conservative attitude for replacing EKF with UKF in practical application.

2.4.3 Particle filter

Since its introduction in 1990s, particle filter has become a very popular class of algorithms to solve numerically the problem of estimating the posterior distribution of state variables through Bayesian inference. PF is greatly considered in many domains such as target tracking, signal processing or prognostics. Its application in health management and prognostics is focused in this study. Particle-filter based prognostics have been applied in many engineering applications such as fatigue crack growth [58, 59, 78], Li-ion batteries [57, 79-81], PEM fuel cells [82-84] machine tools [85]. Jouin et al. [48] has given a literature review of particle-filter-based prognostics, in which the authors elaborated the theory of particle filter and summarized different type of filters in the particle filter family. The drawbacks such as particle degeneracy and particle impoverishment problems were discussed and solutions appeared in literature have been summarized.

The strength of PF lies in its ability for dealing with non-linear non-Gaussian dynamic systems, which characterizes the distribution of the state of the hidden Markov model at the present time, given the information provided by all of the observations received up to the present time. The key idea is to represent the required posterior density function by a set of random particles (or samples) and their weights. Note that technically, the parameters are seen as additional state variables and artificially appended onto the true state vector to form an augmented state-parameter vector so that the particles of state-parameter are processed by PF simultaneously. As the measurements arrive sequentially, the weight of each particle is adjusted by comparing the particles and the new arrived measurement, *i.e.*, the particles that have a higher similarity with the measurement will be assigned a higher weight. PF is composed of two steps at each iteration process: (1) prediction – the particles at previous time step propagate through the damage model to form a collect of particles at current time, which is seen as the priori distribution at current step (2) update – adjust each particle's weight according to how close the particle is to the current measurement, which is quantified by the likelihood function. After a few iterations of prediction-update steps, only a few particles will have negligible weight, known as the degeneracy phenomenon, which implies that a large computational effort is devoted to updating particles whose contribution to the approximated posterior distribution is almost zero. The degeneracy problem typically resorts to resampling methods, whose basic idea is to duplicate the high weighted particles while eliminating the low ones. Conventional resampling algorithms include multinomial, residual, stratified and systematic resampling. Guo et al. [86] compares the PF prognostics results with the above 4 resampling approaches and concludes that

systematic and stratifies resampling show the best results with a slight advantage to systemic resampling which is proven theoretically superior. Some novel resampling methods involve support vector regression-based resampling [87] and monotonic resampling [88]. Note that resampling is not necessary to be implemented at each iteration since it can easily cause the problem of particle impoverishment, *i.e.*, loss of diversity among the particles. This arises due to the fact that in the resampling stage, samples are drawn from a discrete distribution rather than a continuous one. The particles with high weight are duplicated too many times while the ones with small weight are discarded. After iterations, the presentation of the a posteriori density is approximated by only a few high weight particles. If this problem is not address properly, it may lead to "particle collapse", a severe case of sample impoverishment that all particles occupy the same point, giving a poor representation of posterior distribution. Many efforts have been made to address the particle impoverishment [89-94], according to which, the possible ways to address impoverishment include kernel smoothing method, MCMC move method and regularized particle filter. The detail of particle filter will be elaborated in the following sections rather than here since it is difficult to start with particle filter with only basic conceptions and notations.

2.4.4 Bayesian method

Among different ways of applying Bayesian inference, the Bayesian method means here that instead of updating the parameter distribution recursively with measurement data that arrives sequentially, all measurements are processed simultaneously [95, 96]. The joint posterior distribution of parameters at current time step is computed through a single equation, in which the likelihood function at each past measurement prior to current time are multiplied together along with the priori distribution of parameters. Once the expression of the posterior distribution is obtained, a sampling method can be used to draw samples from this distribution, known as target distribution. The Markov Chain Monte Carlo (MCMC) method [97] is usually used. It starts from an arbitrary initial sample (old sample) and a new sample is drawn from a proposal distribution centered at the old sample. The two samples are compared with each other based on an acceptance criterion to decide either the new sample is accepted or the old one is reselected. This process is repeated as many times as necessary until a sufficient number of samples are obtained. The MCMC

sampling results are affected by the proposal distribution and the initial start point. The more similar the proposal distribution is to the target one, the better the samples characterize the target distribution. Also, if the start point is set with a value that differs largely from the possibility from the target distribution, then many iterations (samples) would be required to converge to the target distribution.

2.4.5 Nonlinear least squares method

The non-linear least squares method is the form of least squares analysis used to fit a set of observations with a model that is non-linear with unknown parameters. Detail information about nonlinear least squares method refers to [98-100]. In prognostics context, when the damage model is a non-linear combination of unknown model parameters, NLS can be used to estimate the parameters such that the damage model governed by the estimated parameters fits best the given measurement data in the least squares sense, that is, the sum of the squares of the residuals, defined as the difference between the measurement data and the model prediction, is minimized. This process can be implemented using an iterative process based on some optimization techniques such as Levenberg-Marquardt algorithm [101] that combines the gradient descent and the Gauss-Newton method. NLS processes the measurement data in a batch way, *i.e.*, all measurement data precede to current time are used to identify the model parameters.

In prognostics, the estimated parameters depend on noisy measurement data, and the optimized parameters can be changed when a different set of measurement are used, meaning that the uncertainty in measurements is propagated into the uncertainty in parameters. Given Gaussian white measurement noise, NLS is able to give the uncertainty level in estimated parameters in the form of explicit expressions of covariance. Once the model parameters as well as the covariance are estimated by NLS, the damage growth behavior can be predicted based on the samples of estimated parameters.

CHAPTER 3 STATE-PARAMETER IDENTIFICATION OF DAMAGE GROWTH PARAMETERS USING FILTERING METHODS

3.1 Introduction

In this chapter, the nonlinear state-parameter identification problem by using filter methods is investigated. The theory of Bayesian filter is firstly presented followed by elaborations of three widely-used filters, Extended Kalman filter (EKF), Unscented Kalman filter (UKF) and particle filter (PF). The performances of the three filters in the context of fatigue damage growth are then compared. A simple Paris' law model is used to model the fatigue damage growth. This chapter is organized as follows: Section 3.2 presents the theory foundation of Bayesian inference as well as the theories of these three filters in the general case. Section 3.3 from the Paris' law, a stochastic nonstationary crack growth model is derived. This model is well suited to apply above nonlinear filtering techniques as a way to estimate the unknown parameters. Section 3.4 gives the application of these three filters to the specific fatigue damage model. Numerical experiments are implemented in Section 3.5 and concluding remarks are given in Section 3.6.

3.2 Bayesian estimation

For real system, the true system state can hardly be observed directly. What can be obtained is a series of measurements related to the system state. Due to the work environment or the limitation of the measurement devices, the measurement data can be easily contaminated by the random noise. This noisy data will be used to estimate the system state. This is the so-called system estimation problem. In statistical theory, Bayesian estimate has been a powerful tool for dealing with system state estimation problem especially the non-linear and non-Gaussian system.

Bayesian estimate is based on Bayesian statistic, in which the system state vector $\mathbf{x}$ is regarded as a random variable with some distribution. The observed measurements containing some noise are believed to be used to infer the state distribution. According to Bayesian statistic, some prior distribution, denoted as $p(\mathbf{x})$ is firstly assumed to $\mathbf{x}$. This prior distribution is then updated to a posterior distribution $p(\mathbf{x}|\mathbf{z})$ in the light of new measurement data. The transition from prior to

posterior is implemented by Bayesian theorem.

Suppose a stochastic, discretized, system, which is described by the following state-space model:

$$\begin{cases} \mathbf{x}_k = F(\mathbf{x}_{k-1}) + \mathbf{W}_k \\ \mathbf{z}_k = H(\mathbf{x}_k) + \mathbf{V}_k \end{cases}$$

F is the system transition function, $\mathbf{z}_k$ is the measurement at time step k. H is the measurement function, $\boldsymbol{W}$ and $\boldsymbol{V}$ are the process and measurement noise respectively. The system S contains two stochastic processes, the state transition process and the measurement process. The state transition process, $\{\mathbf{x}_k\}_{k=1,2,\ldots}$, can be regarded as a Markov chain with initial distribution $p(\mathbf{x}_0)$ and the transition probability from one state $\mathbf{x}_{k-1}$ to another state $\mathbf{x}_k$ is denoted by $f(\mathbf{x}_k|\mathbf{x}_{k-1})$. The measurement sequence $\{\mathbf{z}_{1:k}\}$ depends conditionally on the states and the condition distribution is $h(\mathbf{z}_k|\mathbf{x}_k)$. $h(\mathbf{z}_k|\mathbf{x}_k)$ is also known as likelihood function. The system state could be completely characterized by the state initial distribution $p(\mathbf{x}_0)$, state transition probability $f(\mathbf{x}_k|\mathbf{x}_{k-1})$ and the likelihood function $h(\mathbf{z}_k|\mathbf{x}_k)$. The system state estimation problem in terms of Bayesian inference is thus defined as: given a sequence of measurements $\{\mathbf{z}_{1:k}\}=\{\mathbf{z}_1,\mathbf{z}_2,\ldots\mathbf{z}_k\}$, find the optimal estimate of the current state $\mathbf{x}_k|\mathbf{z}_{1:k}$. The posterior distribution of $\mathbf{x}_k|\mathbf{z}_{1:k}$ is calculated recursively by two steps:

(1) Prediction (Chapman-Kolmogorov function)

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \tag{3-1}$$

(2) Update (Bayesian inference)

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1})} \tag{3-2}$$

Eq.(3-1) and Eq.(3-2) are called Bayesian filter functions. However, Eq.(3-2) is only a conceptual solution and in general case, the explicit solution of the functions cannot be obtained. Therefore, a whole family of filtering tools is developed to find the approximate numerical solutions. The choice between theses filters depends on the dynamics of the system and the shape of the noise distribution. If the system is a linear Gaussian system, then the Kalman filter is an effective way to find the numerical solutions and is extended as the so called Extended Kalman filter (EKF) in order to deal with the non-linear Gaussian system. Furthermore, in the frame of EKF, Julier etc. presented the Unscented Kalman filter (UKF) in 1990s, which argues outperform EKF in

terms of precision. However, neither of Kalman filter nor EKF nor UKF could solve Bayesian filter problem in case of the non-linear, non-Gaussian system. In 1990s, particle filter is presented, which gives the numerical solution of Bayesian filter for the non-linear non-Gaussian system. In the reminder parts of this chapter, the basic Kalman filter, EKF, UKF and particle filter will be elaborated, followed by a numerical application in context of fatigue damage propagation.

3.2.1 Kalman filter

Kalman filter is an algorithm that uses a series of measurements observed over time, containing statistical noise and other uncertainties, and produces estimates of unknown variables that tend to be more precise than those based only on a measurement or based only on the system model.

Table 3-1 Linear stochastic system model and measurement model

Model	Discrete Time	Equation Number
System state transition function	$\mathbf{x}_k = \Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1}$	(3-3)
Measurement function	$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k$	(3-4)
System process noise	$E(\mathbf{w}_k) = 0$	(3-5)
	$E(\mathbf{w}_k\mathbf{w}_i^T) = \Delta(k-i)Q_k$	(3-6)
Measurement noise	$E(\mathbf{v}_k) = 0$	(3-7)
	$E(\mathbf{v}_k\mathbf{v}_i^T) = \Delta(k-i)R_k$	(3-8)

Consider the stochastic dynamic system given in Table 3-1. Suppose that a measurement is made at time k and this measurement will be used to update the estimate of state $\mathbf{x}$ of the stochastic system at time k. It is assumed that the measurement is linearly related to the state by an equation (Eq.(3-4)), where $\boldsymbol{H}$ is the measurement matrix and v_k is the measurement noise. Kalman filter is a linear estimator, that gives the updated estimate, or the posteriori estimate, denoted by $\hat{\mathbf{x}}_{k|k}$, as a linear function of the a priori estimate $\hat{\mathbf{x}}_{k|k-1}$ and the measurement $\mathbf{z}_k$:

$$\hat{\mathbf{x}}_{k|k} = \mathbf{K}_k^1\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{z}_k \tag{3-9}$$

The objective is to seek the matrices such that the new estimate will satisfy the orthogonality principle, whose condition can be written in the form

$$E([\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]\mathbf{z}_i^{\mathbf{T}}) = 0, i = 1,2,..k-1, \tag{3-10}$$

$$E([\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]\mathbf{z}_k^{\mathbf{T}}) = 0 \tag{3-11}$$

The derivation process of finding $\mathbf{K}_k^1$ is not presented here and only the final expression of $\mathbf{K}_k^1$ is

given. Readers could refer to [102] for details. The choice of $\boldsymbol{K}_k^1$ causes Eq.(3-9) to satisfy a portion of the orthogonality condition given by Eq.(3-11).

$$\boldsymbol{K}_k^1 = \boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k \tag{3-12}$$

Substitute the formula for $\mathbf{K}_k^1$ into Eq.(3-9), one can obtain

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}) \tag{3-13}$$

Define the errors:

$$\widetilde{\mathbf{x}}_{k|k} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k} \tag{3-14}$$

$$\widetilde{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \tag{3-15}$$

$$\widetilde{\mathbf{z}}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k = \mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} \tag{3-16}$$

$\widetilde{\mathbf{x}}_{k|k}$ and $\widetilde{\mathbf{x}}_{k|k-1}$ are the estimation errors after and before updates, respectively. $\widetilde{\mathbf{z}}_k$ is the error between the true measurement $\mathbf{z}_k$ and the estimated measurement $\hat{\mathbf{z}}_k$. $\widetilde{\mathbf{z}}$is also called innovation. By definition, the priori error covariance is

$$\begin{aligned}
\mathbf{P}_{k|k-1} &= E(\widetilde{\mathbf{x}}_{k|k-1}\widetilde{\mathbf{x}}_{k|k-1}^{\ T}) \\
&= E((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T) \\
&= E((\Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} - \Phi_{k-1}\hat{\mathbf{x}}_{k-1|k-1})(\Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} - \Phi_{k-1}\hat{\mathbf{x}}_{k-1|k-1})^T) \\
&= E((\Phi_{k-1}\widetilde{\mathbf{x}}_{k-1|k-1} + \mathbf{w}_{k-1})(\Phi_{k-1}\widetilde{\mathbf{x}}_{k-1|k-1} + \mathbf{w}_{k-1})^T) \\
&= \Phi_{k-1}E(\widetilde{\mathbf{x}}_{k-1|k-1}\widetilde{\mathbf{x}}_{k-1|k-1}^{\ T})\Phi_{k-1}^{\ T} + \mathbf{Q}_{k-1} \\
&= \Phi_{k-1}\mathbf{P}_{k-1|k-1}\Phi_{k-1} + \mathbf{Q}_{k-1}
\end{aligned} \tag{3-17}$$

The error covariance matrix of measurement is defined as

$$\begin{aligned}
\mathbf{P}_{\widetilde{\mathbf{z}}\widetilde{\mathbf{z}},k} &= E(\widetilde{\mathbf{z}}_k\widetilde{\mathbf{z}}_k^{\ T}) \\
&= E((\mathbf{z}_k - \hat{\mathbf{z}}_k)(\mathbf{z}_k - \hat{\mathbf{z}}_k)^T) \\
&= E((\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_{k-1} - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_{k-1} - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})^T) \\
&= E((\mathbf{H}_k\widetilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_{k-1})(\mathbf{H}_k\widetilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_{k-1})^T) \\
&= \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{\ T} + \mathbf{R}_{k-1}
\end{aligned} \tag{3-18}$$

The cross error covariance matrix of state and measurement is defined as

$$
\begin{aligned}
\boldsymbol{P}_{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{z}},k} &= E(\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{z}}_k{}^T) \\
&= E(\tilde{\boldsymbol{x}}_{k|k-1}(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k)^T) \\
&= E(\tilde{\boldsymbol{x}}_{k|k-1}(\boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1})^T) \\
&= E(\tilde{\boldsymbol{x}}_{k|k}(\boldsymbol{H}_k\tilde{\boldsymbol{x}}_{k|k-1} + \boldsymbol{v}_k)^T) \\
&\quad E(\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{x}}_{k|k-1}{}^T)\boldsymbol{H}_k \\
&= \mathbf{P}_{k|k-1}\mathbf{H}_k
\end{aligned} \tag{3-19}
$$

After these definitions, we now present the derivation of the expression of Kalman gain $\mathbf{K}_k$. The $\hat{\mathbf{x}}_{k|k}$ depends linearly on $\mathbf{x}_k$, which depends linearly on $\mathbf{z}_k$. Therefore, from Eq.(3-11)

$$E([\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]\hat{\mathbf{z}}_k^{\mathbf{T}}) = 0 \tag{3-20}$$

and also (by subtracting Eq.(3-11) from Eq.(3-20))

$$E([\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]\tilde{\mathbf{z}}_k^{\mathbf{T}}) = 0 \tag{3-21}$$

Substitute for $\mathbf{x}_k$, $\hat{\mathbf{x}}_{k|k}$ and $\tilde{\mathbf{z}}_k$ from Eq.(3-3), Eq.(3-9) and Eq.(3-16) respectively. Then Eq.(3-21) becomes

$$E(((\Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1}) - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}))(\mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} - \mathbf{z}_k)^T) = 0$$

$$
\begin{aligned}
0 &= E((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k) + \mathbf{K}_k\mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})^T) \\
&= E((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\mathbf{H}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) - \mathbf{K}_k\mathbf{v}_k)(\mathbf{H}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k)^T) \\
&= E((\tilde{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\mathbf{H}_k\tilde{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\mathbf{v}_k)(\mathbf{H}_k\tilde{\mathbf{x}}_{k|k-1} + \mathbf{v}_k)^T) \\
&= E(\tilde{\mathbf{x}}_{k|k-1}\tilde{\mathbf{x}}_{k|k-1}{}^T\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{H}_k\tilde{\mathbf{x}}_{k|k-1}\tilde{\mathbf{x}}_{k|k-1}{}^T\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{v}_k\mathbf{v}_k{}^T) \\
&= E(\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{R}_k) \\
&\quad E((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{R}_k)
\end{aligned}
$$

It satisfies the equation

$$(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T - \mathbf{K}_k\mathbf{R}_k = 0 \tag{3-22}$$

And therefore, the Kalman gain $\mathbf{K}_k$ can be expressed as

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k{}^T[\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k{}^T + \mathbf{R}_k]^{-1} \tag{3-23}$$

Consider Eq.(3-18) and Eq.(3-19), $\mathbf{K}_k$ can also be written as

$$\mathbf{K}_k = \mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{z}},k}\mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}},k}{}^{-1} \tag{3-24}$$

which indicates that the Kalman gain is equivalent to the cross error covariance matrix $\mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{z}},k}$ times

the inverse of the measurement error covariance matrix $\mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}},k}$. This conclusion will be used in the unscented Kalman filter that will be discussed in Section 3.2.3.

Now seek to express $\mathbf{P}_{k|k}$ in terms of $\mathbf{P}_{k|k-1}$.

$$\mathbf{P}_{k|k} = E(\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T) \tag{3-25}$$

Subtract $\mathbf{x}_k$ from both side of Eq.(3-13), one can obtain

$$\tilde{\mathbf{x}}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\tilde{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{v}_k \tag{3-26}$$

By substituting Eq.(3-26) into Eq.(3-25) and noting that $E(\tilde{\mathbf{x}}_{k|k-1}\mathbf{v}_k) = 0$, one obtains

$$\begin{aligned}\mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\ &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{H}_k\mathbf{K}_k^T - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1} + \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T\mathbf{K}_k^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{H}_k^T\mathbf{K}_k^T + \mathbf{K}_k(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)\mathbf{K}_k^T\end{aligned}$$

Consider Eq.(3-23) that $\mathbf{K}_k(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k) = \mathbf{P}_{k|k-1}\mathbf{H}_k^T$, then Eq.(3-26) becomes

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} \tag{3-27}$$

Now, given an initial state value $\hat{\mathbf{x}}_{0|0}$ and initial error covariance matrix $\boldsymbol{P}_{0|0}$, the state and covariance can be calculated recursively. The calculation includes two steps, prediction and measurement update. The algorithm of the Kalman filter is summarized in Table 3-2. Furthermore, Figure 3-1 illustrates intrusively the dynamic process of Kalman filter.

Table 3-2 Algorithm of Kalman filter

1. Initialization
For k=0, give the initial mean $\hat{\mathbf{x}}_{0|0}$ and the covariance $\mathrm{P}_{0|0}$
2. for k=1,2,…
(a) Prediction:
$\hat{\mathbf{x}}_{k|k-1} = \Phi_{k-1}\hat{\mathbf{x}}_{k-1|k-1}$
$\mathbf{P}_{k|k-1} = \Phi_{k-1}\mathbf{P}_{k-1|k-1}{\Phi_{k-1}}^T + \mathbf{Q}_{k-1}$
(b) Measurement update:
$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$
$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})$
$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}$

3.2.2 Extended Kalman filter

Many dynamic systems and sensors are not linear in real cases, but they are not far from it. Following the considerable success enjoyed by linear estimation methods on linear problems,

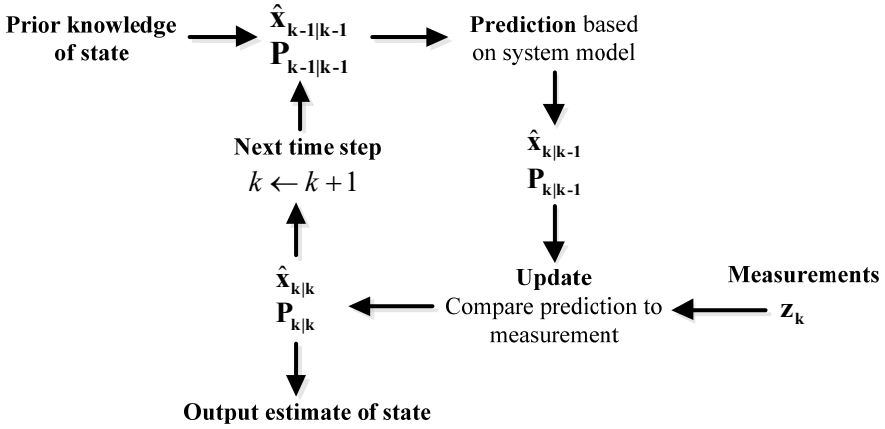

Figure 3-1 Illustration for dynamic process of Kalman filter

people tried to extend it toward nonlinear systems through some approximation ways. A straightforward idea is to use the Taylor expansion to expend the system state transition function and the measurement function around the estimated optimal value.

A central and vital operation performed in EKF is the propagation of a random variable through the system dynamics [63]. In the EKF, the state dynamics model is expanded as a Taylor series around the a posteriori mean of state variable of previous time step. By ignoring second and higher order terms, the state prediction propagates analytically through the nonlinear system equation whilst the state error covariance propagates through a separate first-order linearization of the nonlinear system. Due to this linear approximation, EKF introduces errors from the second order in the true posterior mean and covariance of the transformed state variable, which may lead to sub-optimal results when dealing with significant nonlinearities. However, despite these approximations, from a practical application perspective, EKF algorithm remains a powerful tool in the nonlinear system state estimation domain and has been successfully used in various engineering state-parameter identification problems. Consider the nonlinear stochastic system given by Table 3-3, in which $\mathbf{F}_k$ and $\mathbf{H}_k$ are nonlinear functions.

Table 3-3 Nonlinear stochastic system model and measurement model

Model	Discrete Time	Equation Number
System state transition function	$\mathbf{x}_k = \mathbf{F}_k(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}$	(3-28)
Measurement function	$\mathbf{z}_k = \mathbf{H}_k(\mathbf{x}_k) + \mathbf{v}_k$	(3-29)
System process noise	$E(\mathbf{w}_k) = 0$	(3-30)
	$E(\mathbf{w}_k \mathbf{w}_i^T) = \Delta(k-i)Q_k$	(3-31)
Measurement noise	$E(\mathbf{v}_k) = 0$	(3-32)
	$E(\mathbf{v}_k \mathbf{v}_i^T) = \Delta(k-i)R_k$	(3-33)

Expend the state transition function and the measurement function around the optimal value

$\hat{\mathbf{x}}_{k-1|k-1}$ and neglect the second the higher order terms, Eq.(3-28) and Eq.(3-29) have been approximated by

$$\mathbf{x}_k = \mathbf{F}_k(\hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{\Phi}_{k-1}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{w}_{k-1} \tag{3-34}$$

$$\mathbf{z}_k = \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{C}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k \tag{3-35}$$

in which $\mathbf{\Phi}_{k-1}$ and $\boldsymbol{C}_k$ are the Jacobian matrix of system model and measurement model, respectively.

$$\mathbf{\Phi}_{k-1} = \left.\frac{\partial \mathbf{F}_k(\mathbf{x})}{\partial \mathbf{x}}\right|\mathbf{x} = \hat{\mathbf{x}}_{k-1|k-1} \tag{3-36}$$

$$\mathbf{C}_k = \left.\frac{\partial \mathbf{H}_k(\mathbf{x})}{\partial \mathbf{x}}\right|\mathbf{x} = \hat{\mathbf{x}}_{k|k-1} \tag{3-37}$$

Then the linearized non-linear Gaussian system is

$$\mathbf{x}_k = \mathbf{\Phi}_{k-1}\mathbf{x}_{k-1} + \mathbf{F}_k(\hat{\mathbf{x}}_{k-1|k-1}) - \mathbf{\Phi}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{w}_{k-1} \tag{3-38}$$

$$\mathbf{z}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1} + \mathbf{v}_k \tag{3-39}$$

in which $\mathbf{F}_k(\hat{\mathbf{x}}_{k-1|k-1}) - \mathbf{\Phi}_{k-1}\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}$ are deterministic given $\hat{\mathbf{x}}_{k-1|k-1}$ and $\hat{\mathbf{x}}_{k|k-1}$. Note that in EKF, only the state errors propagate through the separate linear system whilst the state prediction propagates through the nonlinear equations. The algorithm of EKF is summarized in Table 3-4.

Table 3-4 Algorithm of Extended Kalman filter

1. Initialization

For k=0, give the initial mean $\hat{\mathbf{x}}_{0|0}$ and the covariance $\mathrm{P}_{0|0}$

2. for k=1,2,…

(a) Prediction:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k(\hat{\mathbf{x}}_{k-1|k-1})$$

$$\mathbf{\Phi}_{k-1} = \left.\frac{\partial \mathbf{F}_k(\mathbf{x})}{\partial \mathbf{x}}\right|\mathbf{x} = \hat{\mathbf{x}}_{k-1|k-1}$$

$$\mathbf{P}_{k|k-1} = \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^{T} + \mathbf{Q}_{k-1}$$

(b) Measurement update:

$$\mathbf{C}_k = \left.\frac{\partial \mathbf{H}_k(\mathbf{x})}{\partial \mathbf{x}}\right|\mathbf{x} = \hat{\mathbf{x}}_{k|k-1}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}_k^{\mathbf{T}}(\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^{\mathbf{T}} + \mathbf{R}_k)^{-1}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}))$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{C}_k\mathbf{P}_{k|k-1}$$

3.2.3 Unscented Kalman filter

The UKF is based on the idea that it is easier to approximate a probability distribution than to approximate an arbitrary nonlinear transformation. The algorithm is based on propagating carefully selected finite set of points, called sigma points, through the system nonlinear dynamics, and then approximating the first two moments of the distribution (mean and covariance) through a suitable method; such as weighted sample mean and covariance calculations. Studies on the theoretical framework of the UKF algorithm can be found in [64, 72]. For the nonlinear system presented in Table 3-3 the procedure of propagating the system state from time step *k-1* to time step *k* through UKF is explained below.

(1).The filter is initialized with the estimated state mean $\hat{\mathbf{x}}_{0|0}$ and state covariance matrix $\boldsymbol{P}_{0|0}$.

(2).*2n+1* points, called sigma points ($\chi_i, i = 0,1,...,2n$), are calculated based on the state posterior mean and error covariance matrix, where *n* is the dimension of the state vector. Supposing at time step *k-1*, the posterior mean and error covariance $\hat{\mathbf{x}}_{k-1|k-1}$ are and $\mathbf{P}_{k-1|k-1}$. The sigma points are calculated as

$$\boldsymbol{\chi}_{0,k-1} = \hat{\mathbf{x}}_{k-1|k-1} \tag{3-40}$$

$$\boldsymbol{\chi}_{i,k-1} = \boldsymbol{\sigma}_{i,k-1} + \hat{\mathbf{x}}_{k-1|k-1} \qquad i = 1,2,...2n \tag{3-41}$$

where

$$\boldsymbol{\sigma}_{i,k-1} = (\sqrt{(n+\kappa)\mathbf{P}_{k-1|k-1}})_i \quad for\ i = 1,2,...n \tag{3-42}$$

$$\boldsymbol{\sigma}_{i+n,k-1} = -(\sqrt{(n+\kappa)\mathbf{P}_{k-1|k-1}})_i \quad for\ i = 1,2,...n \tag{3-43}$$

Note that index $(\bullet)_i$ denotes the *i*-th column of the matrix $(\bullet)$. The matrix's square root can be calculated by using a lower triangular Cholesky factorization method that prevents the negative covariance matrix [103]. κ is a scaling parameter that we will detail later. It provides an extra degree of freedom to "fine tune" the higher order moments of the approximation, and can be used to reduce the overall prediction error. The generated sigma points assure that

$$\sum_{i=1}^{2n} W_i \boldsymbol{\chi}_{i,k-1} = \hat{\mathbf{x}}_{k-1|k-1} \tag{3-44}$$

$$\sum_{i=1}^{2n} W_i (\boldsymbol{\chi}_{i,k-1} - \hat{\mathbf{x}}_{k-1|k-1})(\boldsymbol{\chi}_{i,k-1} - \hat{\mathbf{x}}_{k-1|k-1})^T = \mathbf{P}_{k-1|k-1} \tag{3-45}$$

in which W is the weight of each point that can be calculated as:

$$W_0 = \kappa/(n+\kappa) \tag{3-46}$$

$$W_i = 1/(2(n+\kappa)) \qquad i = 1,2,...2n \tag{3-47}$$

(3).Each sigma point propagates from k-1to k through the nonlinear system equation:

$$\boldsymbol{\chi}_{i,k} = \mathbf{F}_k(\boldsymbol{\chi}_{i,k-1}) \tag{3-48}$$

(4).The statistics of the propagated sigma points are calculated as the a priori mean and error covariance of the state at time step k

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2n} W_i \boldsymbol{\chi}_{i,k} \tag{3-49}$$

$$\mathbf{P}_{k|k-1} = \sum_{i=0}^{2n} W_i (\boldsymbol{\chi}_{i,k} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\chi}_{i,k} - \hat{\mathbf{x}}_{k|k-1})^T \tag{3-50}$$

(5). Each sigma point $\boldsymbol{\chi}_{i,k}$ transforms through the measurement equation to generate an estimated measurement $\boldsymbol{\gamma}_{i,k}$.

$$\boldsymbol{\gamma}_{i,k} = \mathbf{H}_k(\boldsymbol{\chi}_{i,k}) \tag{3-51}$$

(6). The mean and covariance of the estimated measurement are calculated based on the statistics of the transformed points:

$$\hat{\mathbf{z}}_k = \sum_{i=0}^{2n} W_i \boldsymbol{\gamma}_{i,k} \tag{3-52}$$

$$\mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}},k} = \sum_{i=0}^{2n} W_i (\boldsymbol{\gamma}_{i,k} - \hat{\mathbf{z}}_k)(\boldsymbol{\gamma}_{i,k} - \hat{\mathbf{z}}_k)^T \tag{3-53}$$

(7).The cross-correlation covariance of state-measurement can be obtained as:

$$\mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{z}},k} = \sum_{i=0}^{2n} W_i (\boldsymbol{\chi}_{i,k} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\gamma}_{i,k} - \hat{\mathbf{z}}_k)^T \tag{3-54}$$

(8). The Kalman gain matrix is approximated from the cross-correlation covariance and the estimated measurement covariance as:

$$\mathbf{K}_k = \mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{z}},k} \mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}},k}^{-1} \tag{3-55}$$

(9). The posterior estimates of state mean and error covariance are:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_k) \tag{3-56}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} + \mathbf{K}_k \mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}},k} \mathbf{K}_k^T \tag{3-57}$$

Different aspects that need to be considered during the implementation of the above procedures are summarized below:

(1).The scaling parameter κ in Eq.(3-42) affects the scaling of the fourth and higher order moments of the distribution of σ_i. It is often given by the following equation $\kappa = \alpha^2(n+\lambda) - n$ where n is the dimension of the state vector. α controls the spread extent of the sigma points around the mean and is usually set to a small positive value (*e.g.*,1e-3). λ is the secondary scaling parameter usually set to be 0. The value of κ is crucial for the UKF estimator. An inappropriate choice of κ will cause the estimator divergence.

(2).In the algorithm of UKF, the generation of sigma points depends on the calculation of the square root of the covariance matrix. Since the orthogonal or symmetric matrix square roots are numerically sensitive and computationally expensive to find, in practical, more efficient and stable methods such as the Cholesky decomposition are generally recommended to calculate the matrix square root.

(3).In general UKF, the process noise should be incorporated into the state vector, which makes the dimension of state vector augment from n to $n+q$, where q is the dimension of process noise vector. However, the process noise is assumed to be zero-mean or additive in practice, which does not require augmenting the state vector with the noise variables, thus decreasing again the number of points required to be propagated through the nonlinear system from $2(n+q)$ to $2n$.

3.2.4 Particle filter

In this part the algorithm of particle filter will be derived. The problems coming across in this process, particle degeneracy and particle impoverishment, as well as their solutions will also be presented.

3.2.4.1 Basic Monte Carlo method

As discussed in Section 3.2, Bayesian inference in non-linear non-Gaussian dynamic models relies on the posterior distribution $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$. In many scenarios, it is not possible to compute

these distributions in closed-form. Monte Carlo method is introduced to approximate these distributions; that is numerical schemes in which the distributions of interests are approximated by a large collection of N random samples termed particles [103].

Suppose that N independent random particles are sampled from the posterior distribution $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$; that is $\mathbf{x}_k^i \sim p(\mathbf{x}_k \mid \mathbf{z}_{1:k}), i = 1,2,...N$, then the Monte Carlo method approximates the distribution $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$:

$$\hat{p}(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{1}{N}\sum_{i=1}^{N}\delta(\mathbf{x}_k - \mathbf{x}_k^i) \tag{3-58}$$

in which δ is the dirac delta function. For system state estimation problem, the objective is to obtain the expectation value of the state. Once the posteriori distribution of state is approximated, the expectation value, given by

$$E(f(\mathbf{x}_k)) = \int f(\mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \tag{3-59}$$

can be approximated by

$$\begin{aligned}\hat{E}(f(\mathbf{x}_k)) &= \int f(\mathbf{x}_k)\hat{p}(\mathbf{x}_k \mid \mathbf{z}_{1:k})d\mathbf{x}_k \\ &= \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_k^i)\end{aligned} \tag{3-60}$$

Eq.(3-60) indicates that the expectation value of the system state can be estimated by the average of each particle's state value. However, one problem exists in Monte Carlo approach is that if $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ is very complex, then it is difficult or even impossible to sample from $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$. To address this problem, the importance sampling is introduced.

3.2.4.2 Importance sampling

Importance sampling is a fundamental Monte Carlo method. The idea behind this method is that since it is difficult to sample from $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$, then a proposed density, $q(\mathbf{x}_k \mid \mathbf{z}_{1:k})$, from which it is easier to sample is introduced. $q(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ is called importance density. After the introduction of $q(\mathbf{x}_k \mid \mathbf{z}_{1:k})$, Eq.(3-59) could be written as

$$
\begin{aligned}
E(f(\mathbf{x}_k)) &= \int f(\mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k \\
&= \int f(\mathbf{x}_k) \frac{p(\mathbf{x}_k \mid \mathbf{z}_{1:k})}{q(\mathbf{x}_k \mid \mathbf{z}_{1:k})} q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k \\
&= \int f(\mathbf{x}_k) \frac{p(\mathbf{z}_{1:k} \mid \mathbf{x}_k) p(\mathbf{x}_k)}{q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) p(\mathbf{z}_{1:k})} q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k \\
&= \int f(\mathbf{x}_k) \frac{w_k(\mathbf{x}_k)}{p(\mathbf{z}_{1:k})} q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k
\end{aligned} \tag{3-61}
$$

in which

$$
w_k(\mathbf{x}_k) = \frac{p(\mathbf{z}_{1:k} \mid \mathbf{x}_k) p(\mathbf{x}_k)}{q(\mathbf{x}_k \mid \mathbf{z}_{1:k})} \propto \frac{p(\mathbf{x}_k \mid \mathbf{z}_{1:k})}{q(\mathbf{x}_k \mid \mathbf{z}_{1:k})} \tag{3-62}
$$

Given that

$$
p(\mathbf{z}_{1:k}) = \int p(\mathbf{z}_{1:k} \mid \mathbf{x}_k) p(\mathbf{x}_k) d\mathbf{x}_k
$$

Eq.(3-61) can be further written as:

$$
\begin{aligned}
E(f(\mathbf{x}_k)) &= \frac{1}{p(\mathbf{z}_{1:k})} \int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k \\
&= \frac{\int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k}{\int p(\mathbf{z}_{1:k} \mid \mathbf{x}_k) p(\mathbf{x}_k) d\mathbf{x}_k} \\
&= \frac{\int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k}{\int w_k(\mathbf{x}_k) q(\mathbf{x}_k \mid \mathbf{z}_{1:k}) d\mathbf{x}_k} \\
&= \frac{E_{q(\mathbf{x}_k \mid \mathbf{z}_{1:k})}(f(\mathbf{x}_k) \bullet w_k(\mathbf{x}_k))}{E_{q(\mathbf{x}_k \mid \mathbf{z}_{1:k})}[w_k(\mathbf{x}_k)]}
\end{aligned} \tag{3-63}
$$

The expectation computation problem in Eq.(3-63) could be approximated by the basic Monte Carlo method introduced in 0. Specifically, sample N particles from the proposed importance density, *i.e.*, $\{\mathbf{x}_k^{(i)}\} \sim q(\mathbf{x}_k \mid \mathbf{z}_{1:k})$, and the approximation of $E(f(\mathbf{x}_k))$, denoted by $\hat{E}(f(\mathbf{x}_k))$, is calculated by

$$
\begin{aligned}
\hat{E}(f(\mathbf{x}_k)) &= \frac{\frac{1}{N} \sum_{i=1}^{N} f(\mathbf{x}_k^{(i)}) \bullet w_k(\mathbf{x}_k^{(i)})}{\frac{1}{N} \sum_{i=1}^{N} w_k(\mathbf{x}_k^{(i)})} \\
&= \sum_{i=1}^{N} \tilde{w}_k(\mathbf{x}_k^{(i)}) f(\mathbf{x}_k^{(i)})
\end{aligned} \tag{3-64}
$$

in which

$$\tilde{w}_k(\mathbf{x}_k^{(i)}) = \frac{w_k(\mathbf{x}_k^{(i)})}{\sum_{i=1}^{N} w_k(\mathbf{x}_k^{(i)})} \tag{3-65}$$

Eq.(3-65) is the normalized weight of particle *i*. It can be seen that in Eq.(3-64), the estimated state expectation is no longer the average state value of each particle, but a form of weighted sum. Each particle has its own weight. The greater the weight is, the more confidence one has on the corresponding particle. Note that in Eq.(3-60), the particles come from the distribution $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ while in Eq.(3-64) the particles are sampled from the importance density $q(\mathbf{x}_k \mid \mathbf{z}_{1:k})$

In Eq.(3-62), the calculation of the weight w_k depends on all the measurements proceed *k*, *i.e.*, $z_{1:k}$. The computational complexity increases with the time evolving. In order to admit a fixed computational cost in each time step, it is necessary to derive a recursive form of calculating w_k, which is elaborated in the following. This solution involves selecting an importance density function that could be decomposed as the following form:

$$q(\mathbf{x}_{0:k} \mid \mathbf{z}_{1:k}) = q(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) q(\mathbf{x}_{0:k-1} \mid \mathbf{z}_{1:k-1}) \tag{3-66}$$

The recursive form of the posteriori probability density function is presented as

$$p(\mathbf{x}_{0:k} \mid \mathbf{z}_{1:k}) = p(\mathbf{x}_{0:k-1} \mid \mathbf{z}_{1:k-1}) \frac{p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{x}_{k-1})}{p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1})}$$
$$\propto p(\mathbf{x}_{0:k-1} \mid \mathbf{z}_{1:k-1}) p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$$

The weight in Eq.(3-62) is then derived as

$$\begin{aligned} w_k^{(i)} &\propto \frac{p(\mathbf{x}_{0:k}^{(i)} \mid \mathbf{z}_{1:k})}{q(\mathbf{x}_{0:k}^{(i)} \mid \mathbf{z}_{1:k})} \\ &= \frac{p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{z}_{1:k-1}) p(\mathbf{z}_k \mid \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{z}_{1:k-1}) q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})} \\ &= w_{k-1}^{(i)} \frac{p(\mathbf{z}_k \mid \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})} \\ &= w_{k-1}^{(i)} \frac{p(\mathbf{z}_k \mid \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)} \end{aligned} \tag{3-67}$$

Note that the last step during the above derivation is due to the Markov process: the current state is only related to the previous state. This gives $q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k}) = q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$. **The most common choice for the importance density is the priori probability density function;** that is $q(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k) = p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})$. In that case, the Eq.(3-67) is reduced to

$$w_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{z}_k \mid \mathbf{x}_k^{(i)}) \tag{3-68}$$

Note that the weight in Eq.(3-67) is non-normalized. It should be normalized according to Eq.(3-65) before it can be used to calculate the expectation in Eq.(3-64). The basic particle filter is summarized in Table 3-5.

Table 3-5 Basic algorithm of particle filter

(1)Draw N particles from the initial state distribution

$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), i = 1,2,..N$

(2)For k=1,2,…

(a) Draw particles: $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})$

(b) According to Eq.(3-67), calculate the weight of each particle recursively.

(c) Normalize the weight according to Eq.(3-65)

(d) Calculate the estimated state according to Eq.(3-64)

End for

3.2.4.3 Particle degeneracy and resampling

After a few iterations, all but one particle will have negligible weight. It has been shown that the variance of the importance weights increases over time, and thus, it is impossible to avoid the degeneracy phenomenon. This degeneracy implies that a large computational effort is devoted to updating particles whose contribution to the approximation to $p(x_k \mid z_{1:k})$ is almost zero. Suitable measures of degeneracy could be [89]:

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N} (w_k^{(i)})^2} \tag{3-69}$$

After getting the weight $w_k^{(i)}$, calculate $\hat{N}_{eff}$. If $\hat{N}_{eff}$ is smaller than a threshold denoted by N_T, then resample the particles $\{w_k^{(i)}, \mathbf{x}_k^{(i)}\}$ to obtain N i.i.d samples $\{1/N, \mathbf{x}_k^{(i)*}\}$. $\mathbf{x}_k^{(i)*}$ represents the resampled particles and after resampling, the weight of each particle becomes *1/N*. This improved particle filter is called Sequential Importance Resampling (SIR) particle filter. SIR particle filter is the most prevalent particle filter. The idea of SIR is to redistribute the particles according to their individual weight. Particles with high weights are replicated and particles with low weights are discarded, as shown in Figure 3-2 [79].

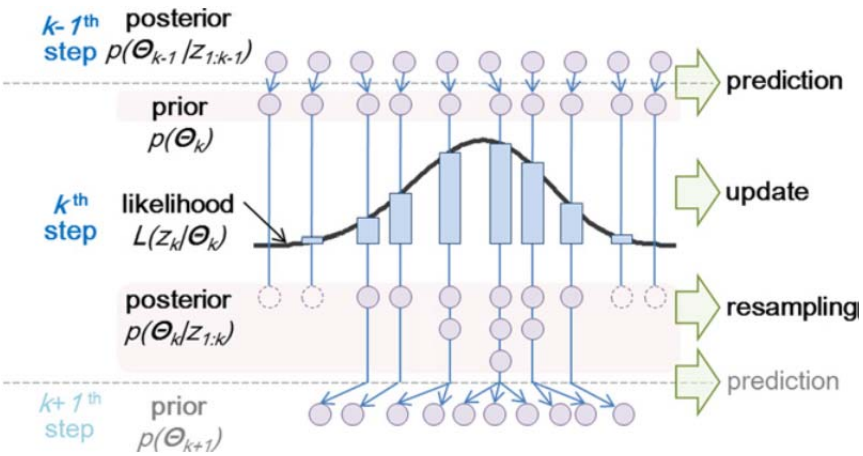

Figure 3-2 Illustration of particle filtering process [79]

Among several methods, the inverse CDF method is used for resampling, shown as Figure 3-3. a CDF is constructed from the likelihood function. Then, a random value is generated from U [0, 1], which becomes a CDF value. Finally, a sample of the parameter having the CDF value is found, which is marked by a rectangle in Figure 3-3.

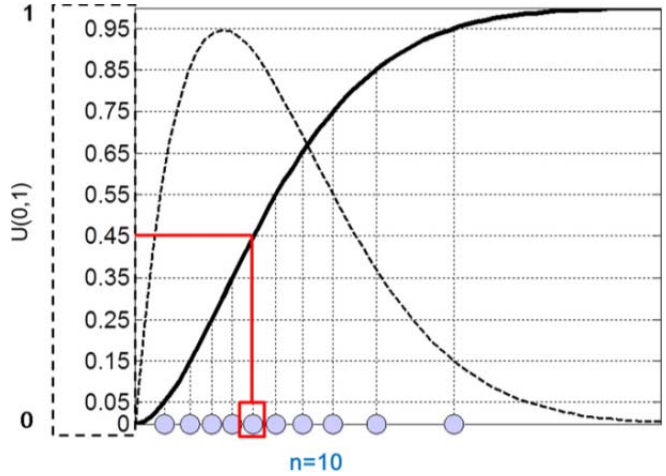

Figure 3-3 Illustration of resampling method [79]

Table 3-6 SIR particle filter

(1)Draw N particles from the initial state distribution

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), i = 1,2,..N$$

(2)For k=1,2,…

(a) Draw particles: $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})_k$

(b) According to Eq.(3-67), calculate the weight of each particle.

(c) Normalize the weight according to Eq.(3-65).

(d) If $\hat{N}_{eff}$ <N_T, do resample, obtained the resampled particles and their weight $\{\mathbf{x}_k^{(i)*}, 1/N\}$

(e) Calculate the estimated state according to Eq.(3-64)

End for

3.2.4.4 Particle impoverishment and its potential solutions

Resampling is prevalent in particle filter to address the degeneracy problem. However, it introduces a new problem, the problem of loss of diversity among the particles, called particle impoverishment. This arises due to the fact that in resampling stage, samples are drawn from a discrete distribution rather than a continuous one. The particles with high weight are duplicated too many times while the ones with small weight are discarded. After iterations, the presentation of the a posteriori density is approximated by only a few high weight particles. If this problem is not addressed properly, it may lead to "particle collapse", which is a severe case of sample impoverishment that all particles occupy the same point in the state space, giving a poor representation of the posterior density.

For the problem of degeneracy, the solution of resampling is a mature technique and has been widely accepted and used by many researchers. However, the solution of solving the impoverishment problem is still in the stage of exploration. Some publications that talk about impoverishment problem including [57, 89-94]. According to these publications the possible ways to address impoverishment include Markov Chain Monte Carlo (MCMC) move method and Regularized particle filter (RPF), in which the particles are samples from a continuous distribution instead of a discrete distribution in the resampling step. Here one of the most common MCMC methods, the Metropolis-Hasting (M-H) algorithm is adopted to mitigate the particle impoverishment.

The M-H algorithm employs a conditional density, known as proposal distribution. The standard Metropolis-Hasting algorithm is summarized as follows [91]:

- Choose a starting point x_0, and set i=0;
- Given current state x_t, draw x^* from $T(x^*| x_t)$ and draw a random number u from U(0,1);
- Accept $x_t = x^*$ if

$$u \le \min\{1, \frac{f(x^*)T(x_t \mid x^*)}{f(x_t)T(x^* \mid x_t)}\}$$

  Otherwise, $x_{t+1}=x_t$;
- Set $i=i+1$, go back to the first step.

The target distribution and the proposal distribution are denoted by *f* and *T* respectively. Usually, the proposal distribution is chosen to be symmetric. Then an acceptance-rejection rule is applied to generate the chain. The proposal distribution can be the prior, while in this study, a special case of M-H algorithm, namely, the Random-Walk Metropolis (RWM) is implemented in SIR particle filter, which is straightforward and efficient.

The theory behind RWM is to perturb the current state of the chain by adding some "noise", *i.e.*, $x' = x_t + \varepsilon_t$, where $\varepsilon_t \sim g_\sigma(\bullet)$ is i.i.d. for different time *t*. The new candidate still stays in the neighborhood of the state x_t. The mission is to determine if the new value is of interest or not by applying the rejection rules. Usually, $\varepsilon_t \sim g_\sigma(\bullet)$ is chosen to be symmetric and a common choice is Gaussian or uniform distribution. The RWM is summarized below:

- Choose a starting point x_0;
- Given the current state x_t, generate $\varepsilon_t \sim g_\sigma(\bullet)$, then evaluate $x' = x_t + \varepsilon_t$;
- Draw a random number *u* from U(0,1), and accept $x_{t+1}=x'$ if

$$u \le \frac{f(x')}{f(x_t)}$$

  Otherwise, $x_{t+1}=x_t$;
- Set $i=i+1$, go back to the first step.

3.3 Damage propagation modeling

Fatigue damage is one of the major failure modes of aircraft structures. Especially, repeated pressurization/depressurization cycles during take-off and landing cause many loading and unloading cycles which lead to fatigue cracks in the fuselage panels. Prediction of fatigue crack propagation and estimation of remaining useful life (RUL) can be implemented to improve maintenance strategies. This is particularly crucial for predictive maintenance, which tries to reduce unnecessary scheduled maintenance stops.

Cracks or damages in this paper refer to existing flaws on the fuselage panel of an aircraft and are modeled as through-the-thickness center straight cracks in an infinite plate. This assumption is well verified if the crack size is small compared to the distance between fuselage stiffeners. For larger crack sizes the model can be adjusted by considering corrective terms in the calculations of

the stress intensity factors to account for boundary conditions effect of stiffeners. Crack propagation can be modeled in myriad ways depending on different phenomena to which the critical crack site is subject [104-106]. Based on airframe fatigue tests on various military aircrafts, Molent et al. concluded that a simple crack growth model adequately represented a typical crack growth [107]. In this work, the celebrated Paris' law is selected to describe the crack growth behavior since it is commonly used for fatigue analysis due to its simplicity. The Paris' law is given by [108]:

$$\frac{da}{dk} = C(\Delta K)^m \tag{3-70}$$

where a is the half-crack size in meter, k is the number of load cycles. da/dk is the crack growth rate in meter/cycle. C and m are the Paris' law parameters which are associated with material properties. ΔK is the range of stress intensity factor in $\mathrm{MPa}\sqrt{m}$, which is given in Eq. (3-71) as a function of the pressure differential p, fuselage radius r and panel thickness t. The coefficient is a corrector factor intends to compensate for modeling the fuselage as a hollow cylinder.

$$\Delta K = A\frac{pr}{t}\sqrt{\pi a} \tag{3-71}$$

Although the crack propagation is a continuous accumulation process, the length of crack size is measured every flight cycle in practice, which can be modeled by a discrete process. Furthermore, the EKF algorithm that we seek to apply usually needs to be implemented numerically and there might not be adequate computational power to integrate the system dynamics as necessary in a continuous-time EKF. Hence, system dynamics is discretized such that a discrete-time EKF can be used [109]. Euler method is employed to discretize Eq.(3-70). By using Euler method, Eq.(3-70) can be rewritten as $a_k = a_{k-\Delta k} + C(\Delta K)^m \Delta k$. The precision of discretization depends on the discrete step. To reduce the discrete error, Δk is set to be one, *i.e.*, the discretization step of the Paris' law is taken at every flight cycle k, which is the minimal possible value from physical and practical point of view. Then the discrete Paris' law is written in a recursive form, given in Eq.(3-72).

$$\begin{aligned} a_k &= a_{k-1} + C\left(A\frac{p_{k-1}r}{t}\sqrt{\pi a_{k-1}}\right)^m \\ &= g(a_{k-1}, p_{k-1}) \end{aligned} \tag{3-72}$$

Note that C and m are the parameters associated with material properties and remain constant once

determined, while the pressure differential p can vary at every flight cycle. Then at each cycle, the pressure p_k is a random variable, which is expressed as

$$p_k = \overline{p} + \Delta p_k \tag{3-73}$$

The disturbance Δp_k around the given average pressure $\overline{p}$ is modeled as a centered normal distribution with variance σ^2. Note that the corresponding variations in p_k are intended to model variations in cruise altitudes of the aircraft for different flights. It would also be possible to model different cruise altitudes by variations in $\overline{p}$ directly but in this work we considered to model this through a perturbation term, given that the various possible cruise altitudes are relatively close to each other for a given aircraft. Then Eq.(3-72) becomes

$$a_k = g(a_{k-1}, \overline{p} + \Delta p_{k-1}) \tag{3-74}$$

Since uncertainty on pressure is generally small, a Mean-Value First Order Second Moment (MVFOSM) approach [110] based on first order Taylor expression is considered in this paper. This gives:

$$a_k = g(a_{k-1}, \overline{p}) + \frac{\partial g(a_{k-1}, \overline{p})}{\partial p} \Delta p_{k-1} \tag{3-75}$$

where $\frac{\partial g(a_{k-1}, \overline{p})}{\partial p}$ is the first order partial derivative of g with respect to the second variable p at the point $(a_{k-1}, \overline{p})$ and can be obtained analytically.

Taking $\frac{\partial g(a_{k-1}, \overline{p})}{\partial p} \Delta p_{k-1}$ as the addictive process noise and considering that $\overline{p}$ is a given constant, Eq.(3-75) could be written as:

$$a_k = f(a_{k-1}) + w_{k-1} \tag{3-76}$$

where

$f(a_{k-1}) = g(a_{k-1}, \overline{p})$ and

$$w_{k-1} = \frac{\partial g(a_{k-1}, \overline{p})}{\partial p} \Delta p_{k-1} \tag{3-77}$$

According to Eq.(3-76) the additive process noise w_k follows a normal distribution with mean zero and variance Q_k:

$$Q_k = \left(\frac{\partial g(a_{k-1}, \overline{p})}{\partial p} \sigma_p \right)^2 = \left(Cm(Ar/t)^m (\overline{p})^{m-1} (\pi a_{k-\Delta k})^{m/2} \sigma_p \right)^2 \tag{3-78}$$

Since the crack size is measured by sensors, the measured crack size always contains noise due to the harsh working environment and sensor limitations. The measurement data is modeled as

$$z_k = a_k + v_k \tag{3-79}$$

in which v_k is the measurement noise such that $v_k \sim N(0,R)$.

3.4 State-parameter identification using filtering methods – an application to fatigue damage model

In this section, EKF, UKF and PF are employed to implement the state-parameter estimation problem in the context of fatigue damage model. Specifically, the state refers the crack size and the parameters are the two Paris' law parameters m and C. The EKF/UKF are grouped and use the same estimation framework while the PF is hereafter presented separately.

The two Paris' law constants $\{m, C\}$ are treated as parameters to be estimated. This makes sense as for an aircraft containing hundreds of fuselage panels, it is normal that the variability are present in the material property parameter among the panel population. Although the nominal values of the Paris' law constants can be known since the material used to construct large commercial aircraft can be well characterized, fatigue is very dependent on minor microstructure irregularities which are difficult to control during the manufacturing process. This leads to various production batches of the material (Al-alloy) having somewhat different fatigue properties. In a large fleet of aircraft it thus cannot be assumed that all panels of all aircraft have the same fatigue properties as the nominal specification of the material manufacturer. Currently to account for this variability between different panels of an aircraft and between different aircraft, large safety factors are used when calculating fatigue life: a safety factor of 3 is typically assumed by the major manufacturers [111, 112]. The corresponding variability is thus significant. Variability in $\{m, C\}$ present through the panel population in a fleet affects the crack propagation rate of each panel, resulting in different time-to-fail for each panel. This fact is important for the operation and management team of an airline to plan the optimal maintenance strategy. It is with this motivation that the Paris' law constants are treated as parameters that need to be identified for each individual panel. This is beneficial for crack growth

prognostics and further for optimizing aircraft maintenance decision-making. In addition, we also consider the uncertainty present during crack propagation processes, which is characterized by assuming the pressure differential is uncertain and varies in each flight cycle.

3.4.1 Crack size Paris' law parameter estimation by EKF/UKF

3.4.1.1 Estimation framework

Estimation of parameters by a filtering approach can be classified into two categories, joint filtering and dual filtering [113-115]. Both of them use a similar filter to estimate state and parameters simultaneously. Joint filtering defines the parameter vector of interest as an additional state variable and artificially appends it onto the true state vector. The appended proportion of augmented state vector does not change beyond the effects of process noise during the time-update process while the augmented error covariance matrix is propagated as a whole (*i.e.* The parameters inherently do not depend on time evolution and keep constant).

In the aforementioned crack growth model, m and C are the unknown parameters that need to be estimated. Therefore, a two-dimensional parameter vector is defined as

$$\Theta=[m,C]^T \tag{3-80}$$

Appending Θ to the state variable, that is crack length a, the augmented state vector is defined as

$$\mathbf{x}_{au}=[a\ m\ C]^T \tag{3-81}$$

Using subscript "*au*" to denote all the augmented variables, the extended system equation is represented as Eq.(3-82) and expanded in the form of matrix for the sake of clarity, given by Eq.(3-83). Note that in Eq.(3-83), the function f refers to the one in Eq.(3-76).

$$\mathbf{x}_{au,k}=f_{au}(\mathbf{x}_{au,k-1})+\mathbf{w}_{au,k-1} \tag{3-82}$$

$$\begin{bmatrix} a_k \\ m_k \\ C_k \end{bmatrix}=\begin{bmatrix} f(a_{k-1}) \\ m_{k-1} \\ C_{k-1} \end{bmatrix}+\begin{bmatrix} w_{k-1} \\ 0 \\ 0 \end{bmatrix} \tag{3-83}$$

The augmented process noise covariance matrix, denoted by $Q_{au,k}$ is written in Eq.(3-84), in which Q_k is the variance of process noise on crack length given in Eq.(3-76).

$$Q_{au,k} = \begin{bmatrix} Q_k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = diag(Q_k, 0, 0) \tag{3-84}$$

Reminder that the measurement is modeled as Eq.(3-10), namely, $z_k=a_k+v_k$, where v_k is Gaussian with noise with variance R.

3.4.1.2 Estimation by EKF

For the augmented system of Eq.(3-82), the EKF consists of initialization, extrapolation and update.

Initialization:

Four variables, process noise covariance matrix Q_{au}, measurement noise variance R, estimated initial state value $\hat{\mathbf{x}}_{au,0}$ and initial state error covariance matrix P_0 should be initialized. Q_{au} is a 3-by-3 matrix including only one nonzero element Q_k, which is given in Eq.(3-76). R can be obtained from the sensor specifications. $\hat{\mathbf{x}}_{au,0}$ is a 3-by-1 vector containing 3 elements, $\hat{a}_0$, $\hat{m}_0$ and $\hat{C}_0$, which are the initial guess for crack length and parameters, respectively. In this paper, $\hat{a}_0, \hat{m}_0$ and $\hat{C}_0$ are generated randomly from uniform distributions. Note that the uniform distributions are used here only for making sure the initial guesses are random and biased from the true value. Once the guesses have been drawn, the uniform distributions will not be used in EKF process and hence have no impact on estimation of EKF on $\hat{a}_k$, $\hat{m}_k$ and $\hat{C}_k$. P_0 represents the confidence in the initial estimate of the state vector. In the absence of any a priori knowledge of the initial state values it is common to assume high values for the P_0 matrix. In contrast if one is confident of the initial estimate of the state, P_0 is generally set small values.

Extrapolation:

According to EKF recurrence formulas, the system propagates as follow.

$$\hat{\mathbf{x}}_{au,k|k-1} = f_{au}(\hat{\mathbf{x}}_{au,k-1|k-1}) \tag{3-85}$$

Rewriting Eq.(3-85) in matrix form for sake of clarity, the above equation is equivalent to:

$$\begin{bmatrix} \hat{a}_{k|k-1} \\ \hat{m}_{k|k-1} \\ \hat{C}_{k|k-1} \end{bmatrix} = \begin{bmatrix} f(\hat{a}_{k-1|k-1}) \\ \hat{m}_{k-1|k-1} \\ \hat{C}_{k-1|k-1} \end{bmatrix} \tag{3-86}$$

The error covariance P propagates as follow:

$$\mathbf{P}_{k|k-1} = \Phi_{k-1}\mathbf{P}_{k-1|k-1}\Phi_{k-1}^{T} + \mathbf{Q}_{au,k-1} \tag{3-87}$$

where Φ is the Jacobian matrix of the augmented system equation f_{au} at point $\hat{x}_{au,k-1|k-1} = [\hat{a}_{k-1|k-1}, \hat{m}_{k-1|k-1}, \hat{C}_{k-1|k-1}]$. It is computed as follows:

$$\Phi_{k-1} = \frac{\partial f_{au}(\hat{\mathbf{x}}_{au,k-1|k-1})}{\partial \mathbf{x}_{\mathbf{au}}} = \begin{bmatrix} 1 + C(m/2)\pi^{m/2}(Apr/t)^{m} a^{(m/2-1)} & C(Apr/t\sqrt{\pi a})^{m}\ln(Apr/t\sqrt{\pi a}) & (Apr/t\sqrt{\pi a})^{m} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\Bigg|_{\substack{a=\hat{a}_{k-1|k-1},\\ m=\hat{m}_{k-1|k-1},\\ C=\hat{C}_{k-1|k-1}}}$$

Update:

Calculate the Kalman gain.

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^{T}[\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{T} + R]^{-1} \tag{3-88}$$

where $\boldsymbol{H}_k$ is the Jacobin matrix of the measurement equation at point $\hat{\mathbf{x}}_{au,k|k-1} = [\hat{a}_{k|k-1}, \hat{m}_{k|k-1}, \hat{C}_{k|k-1}]$.

In this case, $\boldsymbol{H}_k$=[1 0 0].

The estimated measurement can be computed as

$$\hat{z}_k = \hat{a}_{k|k-1} \tag{3-89}$$

The a posteriori estimate of state is obtained from Eq.(3-89). It is expanded in matrix form for clarity, as given in Eq.(3-90)

$$\hat{\mathbf{x}}_{au,k|k} = \hat{\mathbf{x}}_{au,k|k-1} + \mathbf{K}_k(\mathbf{z}_{au,k} - \hat{\mathbf{z}}_{au,k}) \tag{3-90}$$

$$\begin{bmatrix} \hat{a}_{k|k} \\ \hat{m}_{k|k} \\ \hat{C}_{k|k} \end{bmatrix} = \begin{bmatrix} \hat{a}_{k|k-1} \\ \hat{m}_{k|k-1} \\ \hat{C}_{k|k-1} \end{bmatrix} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_k) \tag{3-91}$$

Finally, the error covariance matrix is updated as follow, in which **I** is unit matrix.

$$\mathbf{P}_{k|k} = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\mathbf{P}_{k|k-1} \tag{3-92}$$

3.4.1.3 Estimation by UKF

The detailed algorithm of UKF has been elaborated in 2.4. Here only the process of generating σ and the corresponding sigma points are detailed. In the case of Paris' law parameter estimation, the dimension of the augmented state vector is 3. So 7 sigma points ($\chi_i, i=0,1...6$) are generated. Supposing at time step *k-1*, the posterior mean and error covariance $\hat{\mathbf{x}}_{k-1|k-1}$ are and $\mathbf{P}_{k-1|k-1}$. Let us denote the matrix $\sqrt{(n+\kappa)\mathbf{P}_{k-1|k-1}}$ as below:

$$\sqrt{(n+\kappa)P_{k-1|k-1}} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

Then σ_i is obtained as follows:

$$\begin{matrix} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 & \sigma_6 \end{matrix} \\ \begin{bmatrix} P_{11} & P_{12} & P_{13} & -P_{11} & -P_{12} & -P_{13} \\ P_{21} & P_{22} & P_{23} & -P_{21} & -P_{22} & -P_{23} \\ P_{31} & P_{33} & P_{33} & -P_{31} & -P_{33} & -P_{33} \end{bmatrix} \tag{3-93}$$

Each sigma point $\chi_{i,k-1}$ is a 3-by-1 vector and the 7 sigma points constitute a 3-by-7 matrix which is expressed below:

$$\begin{matrix} \chi_{0,k-1} & \chi_{1,k-1} & \chi_{2,k-1} & \chi_{3,k-1} & \chi_{4,k-1} & \chi_{5,k-1} & \chi_{6,k-1} \end{matrix} \\ \chi_{k-1} = \begin{bmatrix} \hat{a}_{k-1|k-1} & \hat{a}_{k-1|k-1}+P_{11} & \hat{a}_{k-1|k-1}+P_{12} & \hat{a}_{k-1|k-1}+P_{13} & \hat{a}_{k-1|k-1}-P_{11} & \hat{a}_{k-1|k-1}-P_{12} & \hat{a}_{k-1|k-1}-P_{13} \\ \hat{m}_{k-1|k-1} & \hat{m}_{k-1|k-1}+P_{21} & \hat{m}_{k-1|k-1}+P_{22} & \hat{m}_{k-1|k-1}+P_{23} & \hat{m}_{k-1|k-1}-P_{21} & \hat{m}_{k-1|k-1}-P_{22} & \hat{m}_{k-1|k-1}-P_{23} \\ \hat{C}_{k-1|k-1} & \hat{C}_{k-1|k-1}+P_{31} & \hat{C}_{k-1|k-1}+P_{32} & \hat{C}_{k-1|k-1}+P_{33} & \hat{C}_{k-1|k-1}-P_{31} & \hat{C}_{k-1|k-1}-P_{32} & \hat{C}_{k-1|k-1}-P_{33} \end{bmatrix}$$

For the augmented system in Eq.(3-85), the procedure of propagating the system state from time step *k*-1 to time step *k* through UKF is summarized in Table 3-7.

3.4.2 Crack size-Paris' law parameter estimation by PF

The sequential importance resampling (SIR) particle filter will be used. When using the PF framework to the specific crack size-Paris' law parameter estimation problem, the following issues are customized.

(1) The parameter vector could be either a scalar of *m* or *C*, or a vector contains both *m* and *C*, depending on the parameter of interests.

(2) The initial particles are drawn from uniform distribution such as $a_0^{(i)} \sim U(a_{low,} a_{up})$, $\Theta_0^{(i)} \sim U(\Theta_{low}, \Theta_{up}), i=1,2,...N$. Note that a_{low}, a_{up}, Θ_{low}, Θ_{up} are the lower and upper bound of

Table 3-7 Summary of UKF algorithm

1. Initialize $\hat{\mathbf{x}}_{au,0}$ $\boldsymbol{P_0}$

For each time step: $k = 1,2...end$

2. Generate the sigma point matrix χ_{k-1}

3. Compute weight of each point as:

$$W_0 = \frac{\kappa}{n+\kappa}$$

$$W_i = \frac{1}{2(n+\kappa)} \qquad i = 1,2,...2n$$

4. Use the augmented system equation f_{au} ***to transform each sigma point*** $\chi_{i,k-1}$ ***into*** $\chi_{i,k}$***:***

$$\chi_{i,k} = f_{au}(\chi_{i,k-1})$$

5. Predict the state mean and error covariance as:

$$\hat{\mathbf{x}}_{au,k|k-1} = \sum_{i=0}^{2n} W_i \chi_{i,k}$$

$$P_{k|k-1} = \sum_{i=0}^{2n} W_i (\chi_{i.k} - \hat{\mathbf{x}}_{au,k|k-1})(\chi_{i.k} - \hat{\mathbf{x}}_{au,k|k-1})^T$$

6. Use the augmented measurement equation to transform each sigma point $\chi_{i,k}$ ***into*** $\gamma_{i,k}$ ***:***

$$\gamma_{i,k} = \chi_{i,k}$$

7. Compute the mean and covariance of predicted measurement as:

$$\hat{z}_{au,k} = \sum_{i=0}^{2n} W_i \gamma_{i,k}$$

$$P_{\hat{z}\hat{z}} = \sum_{i=0}^{2n} W_i (\gamma_{i,k} - \hat{z}_{au,k})(\gamma_{i,k} - \hat{z}_{au,k})^T$$

8. Compute cross-correlation covariance of state-measurement

$$P_{x\hat{z}} = \sum_{i=0}^{2n} W_i (\chi_{i,k} - \hat{x}_{au,k|k-1})(\gamma_{i,k} - \hat{z}_{au,k})^T$$

9. Compute Kalman gain

$$K_k = P_{x\hat{z}} P_{\hat{z}\hat{z}}^{-1}$$

10. Update the state mean and error covariance

$$\hat{\mathbf{x}}_{au,k|k} = \hat{\mathbf{x}}_{au,k|k-1} + K_k (z_{au,k} - \hat{z}_{au,k})$$

$$P_{k|k} = P_{k|k-1} + K_k P_{\hat{z}\hat{z}} K_k^T$$

uniform distribution for a and Θ, respectively. Then the initial particle is formed as $\mathbf{x}_{au,0}^{(i)} = [a_0^{(i)}, \Theta_0^{(i)}], i = 1,2,...N$

(3) The initial weight of each particle is assigned *1/N*, *i.e.*, $we_0^{(i)} = 1/N$. Note here use the denotation "*we*" instead of "*w*" to denote the particle weight in order to distinguish with the system noise "*w*".

(4) The importance sampling function is chosen as the state function f_{au}. That is to say, the "sampling from the prior distribution" in the general PF framework is implemented by propagating each particle through the state function, as shown in Eq.(3-94). Eq.(3-94) is expended as the vector

form for clarity, shown in Eq.(3-95).

$$\mathbf{x}_{au,k}^{(i)} = f_{au}(\mathbf{x}_{au,k-1}^{(i)}) + \mathbf{w}_{au,k-1}^{(i)} \tag{3-94}$$

$$\begin{bmatrix} a_k^{(i)} \\ \Theta_k^{(i)} \end{bmatrix} = \begin{bmatrix} f(a_{k-1}^{(i)}) \\ \Theta_{k-1}^{(i)} \end{bmatrix} + \begin{bmatrix} w_{k-1}^{(i)} \\ 0 \end{bmatrix} \tag{3-95}$$

(5) The weight of each particle has the same form as the PDF of the measurement noise but with the mean value equaling to the current arrived measurement z_k. Reminder the measurement noise v_k is Gaussian white noise with variance R. Therefore, the weight $we_k^{(i)}$ is calculated as

$$we_k^{(i)} = we_{k-1}^{(i)} \times \frac{1}{\sqrt{2\pi R}} \exp\left(-\frac{(a_k^{(i)} - z_k)^2}{2R} \right) \tag{3-96}$$

(6) The optimal estimates of the system state is calculated by Eq.(3-97), where $\widetilde{we}_k^{(i)}$ is the normalized weight.

$$\hat{\mathbf{x}}_{au,k} = \sum_{i=1}^{N} \widetilde{we}_k^{(i)} \mathbf{x}_{au,k}^{(i)} \tag{3-97}$$

In summary, the procedure of using SIR particle filter to the specific fatigue damage model is summarized in Table 3-8.

Table 3-8 Procedure of applying SIR particle filter to fatigue damage model

(1)For k=0, draw N particles from the initial state distribution

$a_0^{(i)} \sim U(a_{low}, a_{up}), i = 1,2,..N$

$\Theta_0^{(i)} \sim U(\Theta_{low}, \Theta_{up}), i = 1,2,..N$

$\mathbf{x}_{au,0}^{(i)} = [a_0^{(i)}, \Theta_0^{(i)}], i = 1,2,..N$

Assign each particle $\mathbf{x}_{au,0}^{(i)}$ the initial weight 1/N.

(2)For k=1,2,…

(a) Propagate each particle $\mathbf{x}_{au,0}^{(i)}$ through Eq.(3-94)

(b) Calculate the weight of each particle using Eq.(3-96).

(c) Normalize the weight and then the particles and their weight are represented as $\{ \mathbf{x}_{au,0}^{(i)}, \widetilde{we}_k^{(i)} \}$.

(d) Calculate $\hat{N}_{eff}$ according to Eq.(3-69). If $\hat{N}_{eff}$ is smaller than some threshold, do resample, obtained the resampled particles and their weight $\{ \mathbf{x}_{au,k}^{(i)*}, 1/N \}$

(e) Calculate the estimated state according to Eq.(3-97)

End for

3.5 Numerical study

3.5.1 Performance of EKF/UKF

This part investigates the application of the EKF and UKF algorithms to the estimation of the parameters m and C and true crack size a based on noisy measurements of the crack size in an aircraft fuselage panel. The noisy measurements are not real measurements but simulated ones based on the crack size evolution law to which we add simulated noise. The values in Table 3-9 are used for the numerical simulations. The values defining the aircraft geometry are characteristic of a short range commercial aircraft. The total number of simulation steps is set to be 60000 cycles based on the typical lifetime of such short-range commercial aircrafts. The parameter vector Θ, composed of the two material property parameters $\{m, C\}$, has been defined in Eq.(3-80).

Table 3-9 Numerical values for case study

Parameter	Denotation	Type	Value	Unit
Fuselage radius	r	Deterministic	1.95	m
Panel thickness	t	Deterministic	2e-3	m
Correct factor for stress intensity factor	A	Deterministic	1.25	-
Pressure differential	p	Normally distributed	N(0.06,0.003)	MPa
True initial crack length	a_0	Deterministic	0.5e-3	m
True Paris' law parameter	m	Deterministic	3.8	-
True Paris' law parameter	C	Deterministic	1.5e-10	-
Estimated initial crack length	$\hat{a}_0$	Uniformly distributed	U(a_0-0.25a_0, a_0+0.25a_0)	m
Estimated initial Paris' law parameter	$\hat{m}_0$	Uniformly distributed	U(m_0-0.25m_0, m_0+0.25m_0)	-
Estimated initial Paris' law parameter	$\hat{C}_0$	Uniformly distributed	U(C_0-0.25C_0, C_0+0.25C_0)	-
Initial error covariance matrix	P_0	Deterministic	$diag(1\times10^{-4},1,1\times10^{-10})$	-
Measurement noise variance	R	Deterministic	$(10\%a_0)^2$	-

According to the numerical experiments, the performance of EKF and UKF based on only one run is difficult to evaluate. Sometime UKF outperforms EKF while sometime on the contrary. In order to compare the average performance of EKF against UKF, A series of 100 numerical experiments under the same initial conditions have been implemented. In each experiment, we pre-generate simulated measurement data and these data are used in both EKF and UKF. The parameter convergence paths of these 100 experiments given by EKF and UKF are illustrated in Figure 3-4 and Figure 3-5.

To characterize the convergence behavior over these 100 samples, 4 indicators, $\hat{\Theta}_k$, Θ_{error}, $\hat{MSE}_k$ and MSE_k are established. $\hat{\Theta}_k$ is the average value of these 100 samples at the k-th flight cycle, as shown in Eq.(3-98), in which n_s is the number of repetitions of the simulations, here, n_s=100. Θ_{error} is the absolute value of relative error on $\hat{\Theta}_k$, calculated through Eq.(3-99), in which Θ denotes the "true" value of parameter.

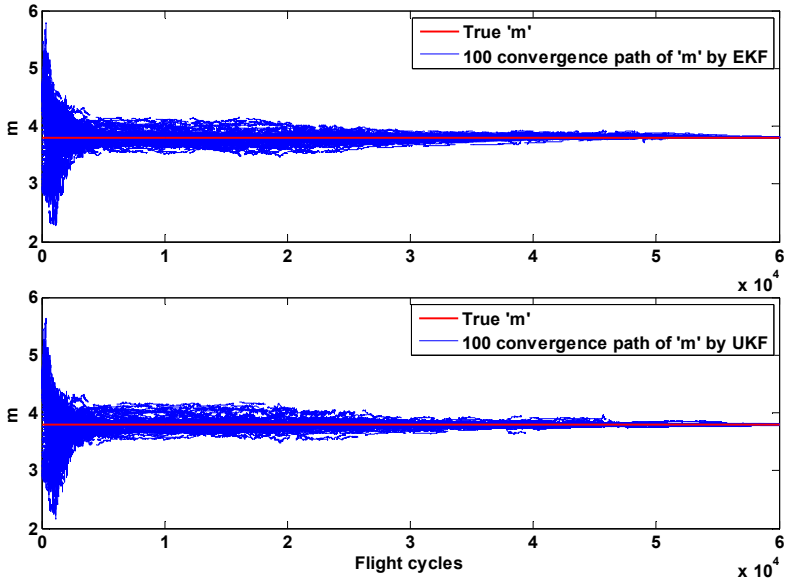

Figure 3-4 100 convergence path of '*m*' by EKF and UKF

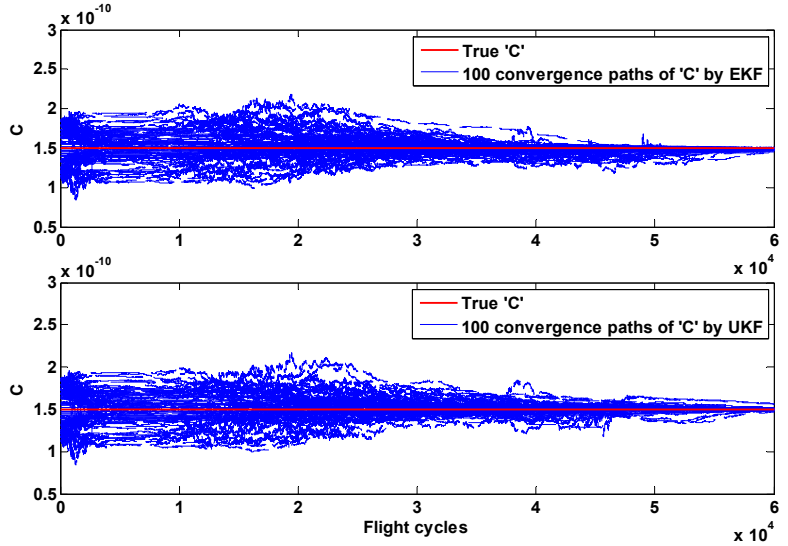

Figure 3-5 100 convergence path of '*C*' by EKF and UKF

$$\hat{\Theta}_k = \frac{1}{n_s}\sum_{j=1}^{n_s}\hat{\Theta}_{k,j} \tag{3-98}$$

$$\Theta_{error} = \frac{|\hat{\Theta}_k - \Theta|}{\Theta} \times 100\% \tag{3-99}$$

To measure the spread of the EKF estimation, the estimated variance at the *k*-th flight cycle in presence of the true values of the parameters, MSE_k, is computed through Eq.(3-100). Moreover, the estimated variance at *k*-th flight cycle in absence of the true values of the parameters, $\hat{MSE}_k$, is also given by Eq.(3-101). This indicator can be useful for determining the confidence interval when

the true values are unknown, which is typically the case in prognostics of residual life.

$$MSE_k = \frac{1}{n_s - 1}\sum_{j=1}^{n_s}(\hat{\Theta}_{k,j} - \Theta)^2 \tag{3-100}$$

$$\hat{MSE}_k = \frac{1}{n_s - 1}\sum_{j=1}^{n_s}(\hat{\Theta}_{k,j} - \hat{\Theta})^2 \tag{3-101}$$

Note the difference between *MSE* and $\hat{MSE}_k$. *MSE* is used to indicate how far on average the collection of estimates are from the true values of parameters being estimated, which in practice, are generally unknown, while $\hat{MSE}_k$ indicates how far, on average, the collection of estimates are from the average value of estimates. Table 3-10 lists the four indicators of the EKF and UKF estimator for *m* and *C* after 100, 1000, 10000, 30000 and 60000 cycles.

Table 3-10 Comparison $\hat{\Theta}_k / \Theta_{error} / \hat{MSE}_k$ /MSE of *m* and *C* over 100 simulations

No. of cycles	Para.	Filter	$\hat{\Theta}$	Θ_{error} (%)	$\hat{MSE}$	***MSE***
100 cycles	m	EKF	3.8420	1.1046	0.3215	0.3233
		UKF	3.8410	1.0791	0.3169	0.3185
	C	EKF	1.5067E-10	0.4462	5.0522E-22	5.0567E-22
		UKF	1.5066E-10	0.4433	5.0006E-22	5.0050E-22
1000 cycles	m	EKF	3.7122	2.3105	0.1830	0.1907
		UKF	3.6460	4.0518	0.1715	0.1952
	C	EKF	1.4988E-10	0.0783	5.2087E-22	5.2089E-22
		UKF	1.4942E-10	0.3878	5.1842E-22	5.1876E-22
10000 cycles	m	EKF	3.8017	0.0453	0.0142	0.0142
		UKF	3.8410	0.1664	0.0155	0.0155
	C	EKF	1.5222E-10	1.4797	3.7710E-22	3.8208E-22
		UKF	1.5141E-10	0.9423	3.7221E-22	3.7420E-22
30000 cycles	m	EKF	3.7917	0.2190	0.0023	0.0024
		UKF	3.7894	0.2793	0.0019	0.0020
	C	EKF	1.5156E-10	1.0393	8.7134E-23	8.9589E-23
		UKF	1.5255E-10	1.7015	7.4442E-23	8.0956E-23
60000 cycles	m	EKF	3.8025	0.0651	0.0000	0.0000
		UKF	3.7995	0.0122	0.0000	0.0000
	C	EKF	1.4920E-10	0.5360	1.2643E-24	1.9173E-24
		UKF	1.5014E-10	0.0952	1.5426E-24	1.5630E-24

The performance comparisons between EKF and UKF for estimating the parameter *m* is given in Figure 3-6. The left panel presents the comparison between *m* and $\hat{m}_k$. Note that $\hat{m}_k$ is the average value over 100 simulations. The right panel in Figure 3-6 gives the absolute value of relative error, *i.e.*$|\hat{m}_k - m|/m$. Similar, The performance comparison between EKF and UKF for estimating the parameter *C* is given in Figure 3-7. The left panel illustrates the comparison of *C* and $\hat{C}_k$. $\hat{C}_k$ is the average value over 100 simulations. The right panel in Figure 3-7 gives the absolute value of relative error, *i.e.*$|\hat{C}_k - C|/C$.

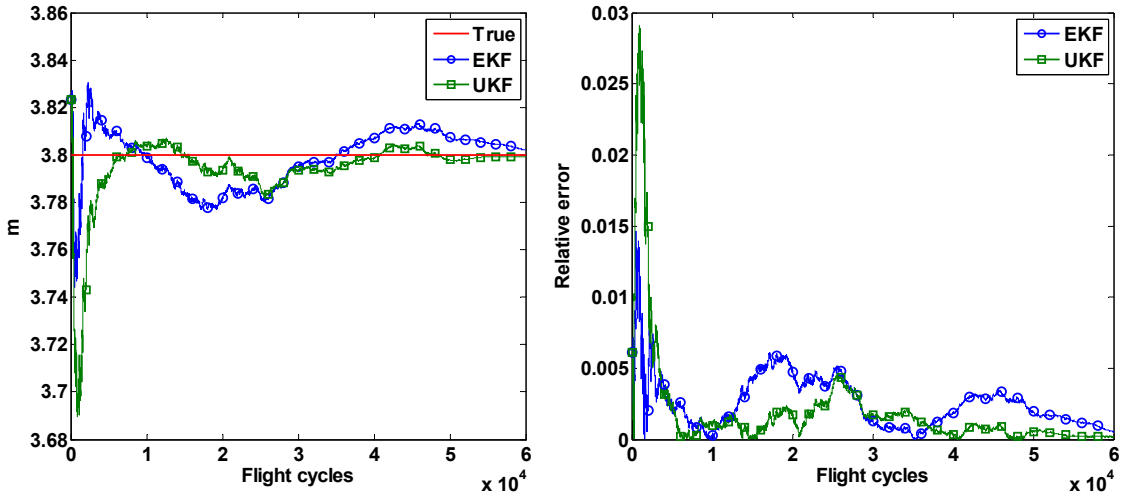

Figure 3-6 Performance comparisons of EKF and UKF for estimating the parameter *"m"*

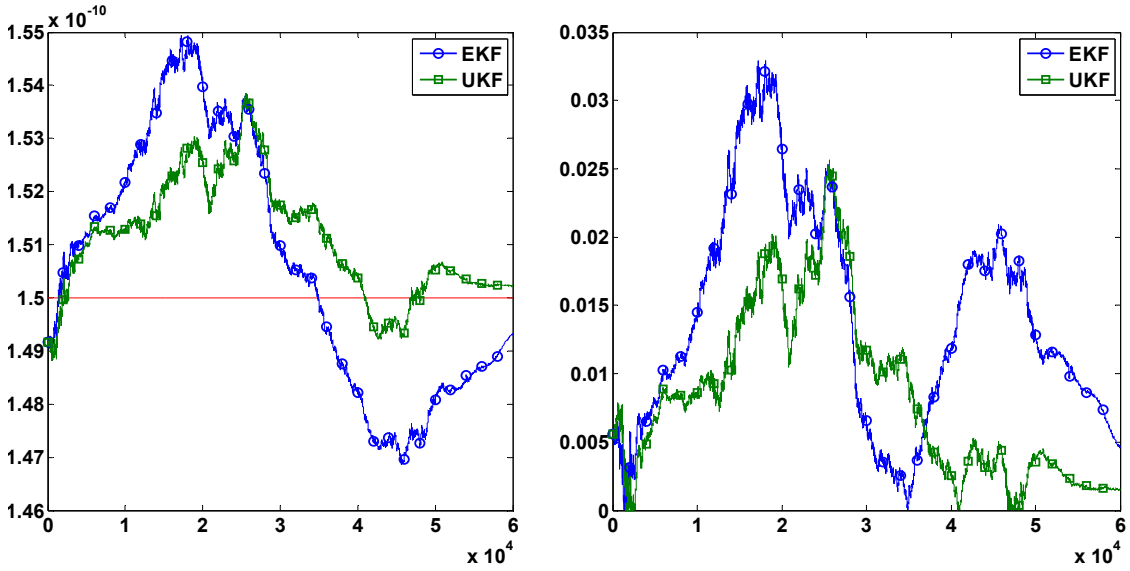

Figure 3-7 Performance comparisons of EKF and UKF for estimating the parameter *"m"*

Crack size evolution based on one simulation is illustrated in Figure 3-8. The left panel is the comparison between true crack size a_k and estimated crack size given respectively by EKF ($\hat{a}_k^{(E)}$) and UKF ($\hat{a}_k^{(U)}$). Each point represents a noisy measurement. Note that only one point every 100 cycles is represented in order not to overload the figure. It can be seen that the estimated crack size by both EKF and UKF fits well the true one (solid line) even with polluted measurements (solid points). The EKF outperforms slightly the UKF during the first several hundred flight cycles while finally the UKF gives better estimate to crack size that is closer to the true value. The right subplot gives the relative error between a_k and $\hat{a}_k$, *i.e.* $\left|\hat{a}_k - a\right| / a$. Both of the relative error of EKF and UKF begin with a high value of 0.4% due to a large biased initial estimate for crack size and decreases rapidly to nearly 0.1% after 1000 flight cycles. Then they fluctuate to a small extend but keep very

low (less than 2%) and overall decline progressively. However, note that UKF is much more time-consuming than that of EKF (12.38s vs 3.48s, on an Intel(R) i3-4130, 3.40GHz CPU) since a transformation of 7 sigma points is required at each time step compared to only one point performed in EKF.

According to the numerical experiments, it can be seen that when providing the same measurements and running the simulation once, the performance of UKF over EKF shows randomness and is difficult to evaluate. But when runs the simulation many times (*e.g.*, 100 times) and comparing the avergae value of the estimates, the UKF performs better than EKF in terms of the relative error, as shown in Figure 3-6 and Figure 3-7. However, UKF has its shortcomings. Firstly, UKF is sensitive to the scaling parameter κ and has a strict constraint for the choice of the initial state error covariance matrix $\boldsymbol{P}_0$. Improper selection of κ will cause estimator divergence while an inappropriate value of $\boldsymbol{P}_0$ may lead to a non-positive semi definite state error covariance matrix after several iterations, which makes Cholesky factorization impossible. Secondly, UKF is much more time-comsuing than EKF since a transformation of 7 simga points are required compared with EKF which needs to process only one point. Accordingly, the author concludes that for fatigue crack propagation problem EKF is an accurate and efficient recursive state/parameter estimation approach.

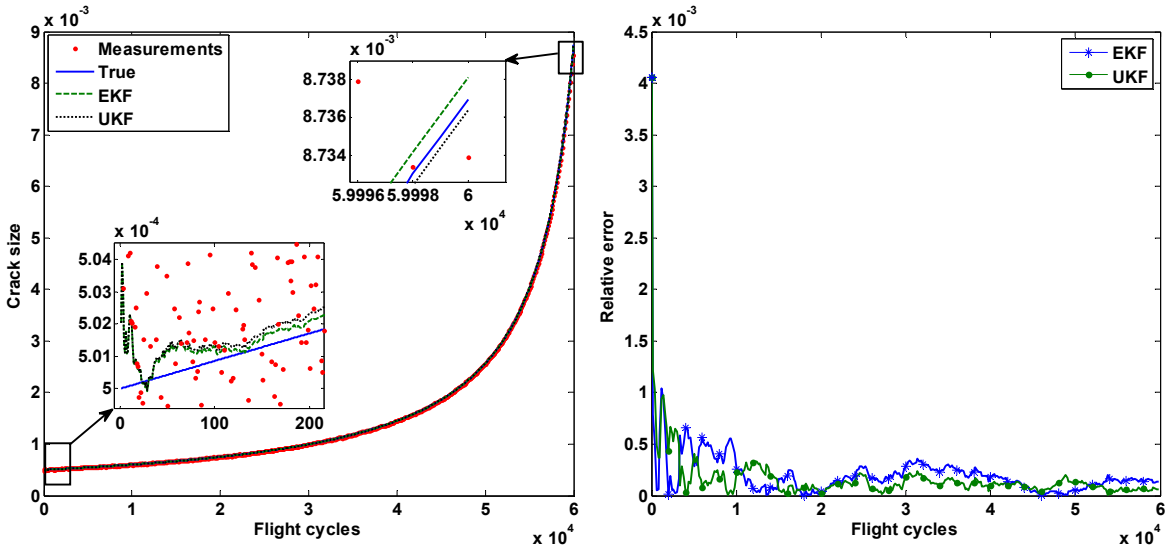

Figure 3-8 Crack size evolution based on one simulation

3.5.2 Performance of particle filter

This part investigates the application of the particle filter algorithms to the estimation of the

parameters m and C and crack size a based on noisy measurements. The numerical values used in particle filter experiments are reported in Table 3-11. Note that the initial distribution in the last three lines refers to the distribution from which the initial particles, $a_0^{(i)}$, $m_0^{(i)}$, $C_0^{(i)}$, are drawn. The initial distribution is assumed as a uniform distribution with the range of 50% around the true value. N=200 particles are used. In order to mitigate the particle impoverishment phenomena, a shorter time-scale is preferred to reduce the recurrence number. To this end, the crack is set to propagate from a relative larger initial value a_0=10mm, compared with the 0.5mm used in EKF/UKF. A threshold a_{th}=40mm is set to prevent the crack from growing too large to some impractical value. The numerical experiments are separated in three parts. In the first two parts, the parameter m and C are estimated separately one after another. Then in the third part m and C are estimated simultaneously.

Table 3-11 Numercial values used in particle filter experiments

Parameter	Denotation	Type	Value	Unit
Fuselage radius	r	Deterministic	1.95	m
Panel thickness	t	Deterministic	2e-3	m
Correct factor for stress intensity factor	A	Deterministic	1.25	-
Pressure differential	p	Normal	N(0.06,0.003)	MPa
True initial crack length	a_0	Deterministic	10e-3	m
True Paris' law parameter	m	Deterministic	3.8	-
True Paris' law parameter	C	Deterministic	1.5e-10	-
Measurement noise variance	R	Deterministic	$(10\%a_0)^2$	-
Initial distribution of a_0	-	Uniform	U[7.5e-3,12.5e-3]	m
Initial distribution of m		Uniform	U[3,4]	-
Initial distribution of C		Uniform	U[0.5e-10, 3e-10]	-

Firstly, the parameter C is fixed as the true value 1.5e-10 and m is treated as unknown. In this case, the augmented state vector has the form such as $\mathbf{x}_{au}=[a,m]$. The initial particles are drawn such that $a_0^{(i)}$~U[7.5e-3,12.5e-3], $m_0^{(i)}$~U[3,4], i=1,2,…N. The estimation result of m is shown in Figure 3-9(a). It indicates that m converge to its true value with time evolving. The comparison between the true crack size and the PF estimate is illustrated in Figure 3-9 (b), which shows that the estimated crack size well tracks the true one. Note that the measurements are collected every flight cycle but plotted every 100 cycles in order to not overdensify the figure.

Then the parameter m is fixed as the true value 3.8 and C is treated as unknown. In this case, the augmented state vector has the form such as $\mathbf{x}_{au}=[a,C]$. The initial particles are drawn such that $a_0^{(i)}$~U[7.5e-3,12.5e-3], $C_0^{(i)}$~U[0.5e-10, 3e-10], i=1,2,…N. The convergence process of C is shown in Figure 3-10(a). The crack size estimation by PF is shown in Figure 3-10(b), in which the measurements are plotted every 100 cycles. Similar to the previous case, the parameter C

converges to its true value with time evolving and the estimated crack size fits well the true one.

Finally, both the m and C are assumed as unknown and estimated simultaneously. In this case, the augmented state vector has the form such as $\mathbf{x}_{au}=[a,m,C]$. The initial particles are drawn such that $a_0^{(i)}$~U[7.5e-3,12.5e-3], $m_0^{(i)}$~U[3,4], $C_0^{(i)}$~U[0.5e-10, 3e-10], i=1,2,…N. The convergence process of m and C is shown in Figure 3-11. The estimated crack size is illustrated in Figure 3-12, in which the measurements are plotted every 100 cycles.

From the above results, it can be seen that particle filter is well performance in the case of dealing one-parameter estimation. When the augmented state is high dimensional including several unknown parameters, the estimation quality can be unsatisfactory. In addition, when dealing with the long-term iteration, particle filter suffers serious particle impoverish problem.

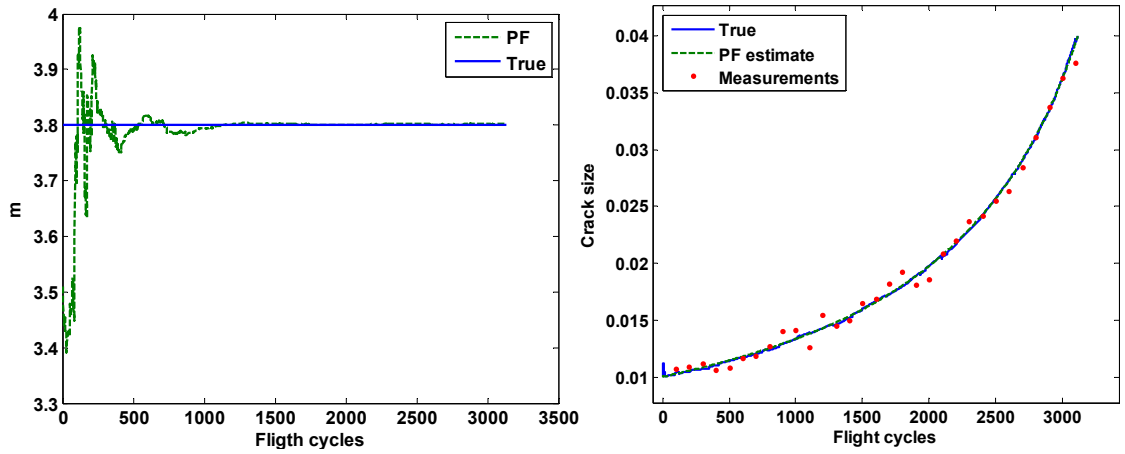

Figure 3-9 (a) The convergence process of *"m"* by PF (b) Crack size estimation by PF, in the case of only m is estimated

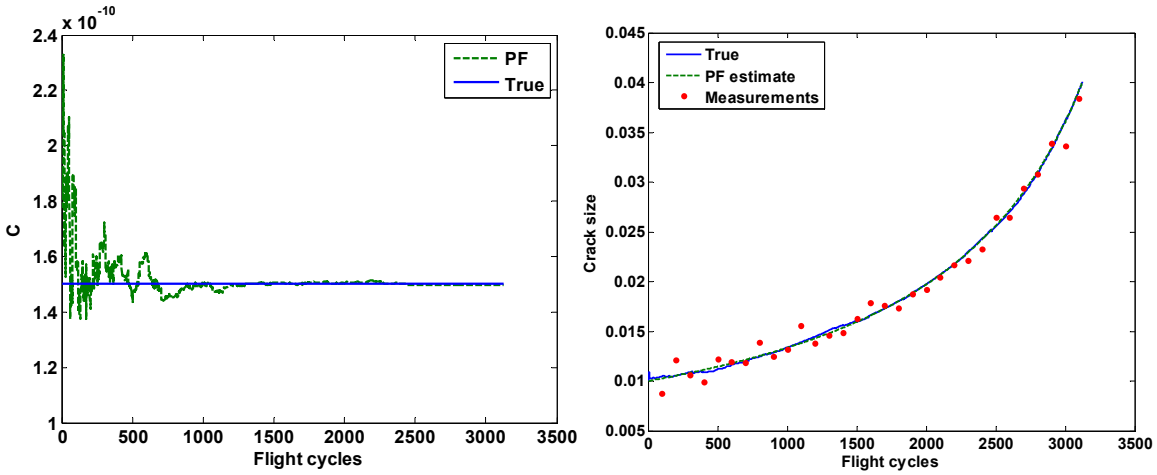

Figure 3-10 (a) The convergence process of *"C"* by PF (b) Crack size estimation by PF, in the case of only C is estimated

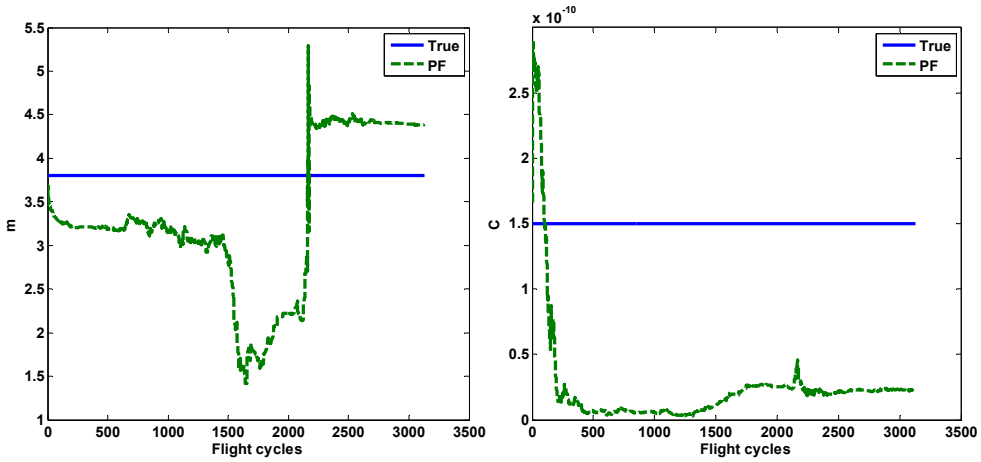

Figure 3-11 The convergence process of *"m"* and *"C"* by PF when estimating *"m"* and *"C"* simultaneously

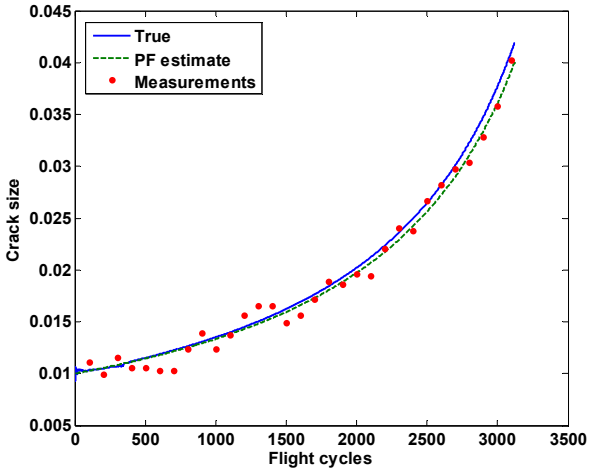

Figure 3-12 Crack size estimation by PF, in the case of only *C* is estimated in the case of estimating *"m"* and *"C"* simultaneously

3.6 Summary

This chapter introduced the Bayesian filter theory and its three numerical implementation methods, EKF, UKF and PF. These three numerical appraoches are applied to a specific state-paramter estimation problem in the context of fatigue damage model presented by Paris' law. Considering the roubustness, good performance in idenfitiying the model parameters as well as cheap computational cost of the EKF algorithom, EKF is chosen as the state-parameter estimation solution in the following model-based prognostics method.

CHAPTER 4 A MODEL-BASED PROGNOSTICS METHOD BASED ON COUPLING EXTENDED KALMAN FILTER AND FIRST-ORDER PERTURBATION METHOD FOR FATIGUE DAMAGE PROGNOSTICS ON FUSELAGE PANELS

4.1 Introduction

Fatigue damage is one of the major failure modes of aircraft structures. Especially, repeated pressurization/depressurization during take-off and landing cause many loading and unloading cycles which can lead to fatigue crack in the fuselage panels. Prognostics, focusing on the estimation of the future damage evolution and potential risk associated to a fuselage panel is crucial to aircraft safety. For the problem discussed at hand, the model-based prognostics method is adopted since the fatigue damage model for metal materials have been well researched and easy to obtain. Here the well-known Paris' law is employed. The two material property parameters in Paris' law and the initial crack size are assumed unknown. Moreover, both the crack size evolution and the loading condition at each flight cycle are affected by uncertainties. The model-based prognostics method applied in this work is tackled with two sequential phases: **(1) estimation of fuselage fatigue crack size and identification of unknown model parameters, and (2) prediction of the crack size distribution in future evolution**.

The first step of the model-based prognostics is implemented by the Extended Kalman filter that has been elaborated in CHAPTER 3. For the second phase, there are two main prediction types. The most obvious and widely used one is to predict remaining useful life (RUL), which is defined as the operation time of a component left before a damage indicator exceeds a threshold, given the component's current condition and past operation profile. Alternatively, it can also consider the evolution of crack size distribution for a given time interval. In some situations, especially when a fault or failure is catastrophic, (*e.g.* fatigue damage of the aircraft fuselage panel), inspection and maintenance are implemented regularly to avoid such failures by replacing or repairing the components that are in danger. In this case, it would be more desirable to predict the probability

that a component operates normally before some future time (*e.g.* next maintenance interval). This can be useful in order to help define the maintenance planning based on a prescribed probability of unscheduled maintenance, which imply extra cost and delay.

Both of the above two prediction types are kinds of predicting reliability, *i.e.*, the ability of a system or component to perform its required functions under stated conditions for a specified period of time [116]. Many different types of uncertainty sources that present in the dynamic system make it computationally expensive to evaluate the reliability of the system when considering the traditional Monte Carlo (MC) method or Latin Hypercube Sampling (LHS) [117] methods. Several alternative methods have been proposed to overcome this difficulty such as first-order reliability method (FORM) [118] and second-order reliability method (SORM) [119]. One of the limitations of these methods is that a system dynamic model that describes the system degradation mechanism is generally required. For the complex systems whose degradation model is difficult to obtain, the above methods are not able to be used. However, once the system dynamic model is obtained, the strength of such methods is manifested, namely, they can provide a relative long-term prediction for system reliability.

With the progress of sensor technology and the data acquisition and processing techniques, real time monitoring of the system health state is gradually considered in many engineering domains. Real time reliability prediction technique used in condition monitoring and maintenance decision is highly concerned by engineers. The abundant real time data make it possible to infer the system degradation path and then to predict the reliability. The relative methods for real time reliability prediction can be roughly classified into two categories: time series modeling and regression analysis. Both these two methods learn the degradation path from the historical on-line measurements and do not require much knowledge of degradation mechanism. The shortcoming of time series methods is that they cannot give a long prediction of degradation propagation and is only suitable in some short-term prediction cases.

The techniques belong to time series analysis include autoregressive moving average (ARMA) model, neural networks (NNs), support vector machines (SVMs) and other machine learning methods. Dang et al. [120] proposed a reliability prediction method based on degradation measure distribution and wavelet neural network. Wu et al. [121] and Karunanithi et al. [122] employed neural network to implement the software reliability prediction. Choi et al. [123] examined the

performance of two types of classifiers, namely, probabilistic neural network (PNN) and SVM in order to find the classification method with the highest accuracy for the estimation of system reliability, and compared them with the traditional MC method. Qin [124] constructed a software reliability prediction model based on PSO and SVM in which the Particle Swarm Optimization (PSO) was introduced to automatically optimize the parameters of SVM. Wu et al. [125] employed the least square support vector machine (LSSVM) combined with nonlinear filters, specifically, the Extended Kalman filter and the Unscented Kalman filter, to update the reliability data. The proposed method was compared with the existing NNs and SVMs and was proved to have better performance. Xu et al. [126] studied a dynamic system suffering a hidden degradation process. The dynamic system was characterized by state-space model and the particle filtering method was used to identify the model parameters. Based on the on-line estimation of the unknown parameters, a real-time reliability prediction framework was proposed. Amin [127] proposed a time series ARIMA (autoregressive integrated moving average) model to address the software reliability prediction issue. Lu et al [126] proposed a time series model that was computationally recursive and provided short-term, real time forecasts linked directly to conditional reliability estimates. Wang et al. [128] presented a reliability prediction for systems with multiple degradation measures to characterize the reality that a system may consist of multiple components or a component may have multiple degradation paths. These multiple measures may be correlated.

Support vector regression (SVR) is used as a regression method for reliability prediction. Zhao et al. [129] employed the SVR approach to address the system reliability prediction problem given a series of failure time data. For the very important SVR parameter selection problem, they proposed a combination method of an analytic selection of prior selection followed by a genetic algorithm for intelligent optimization to avoid the problems such as divergence, slow convergence, local optima, etc. Chen [130] also employed SVR combined with genetic algorithm to implement the reliability prediction. Zhao et al. [131] proposed a particle filter-support vector regression for reliability prediction in which the particle filter was used to dynamically update the SVR parameter when new observations were available. Besides the machine learning methods discussed above, other reliability prediction approaches include fuzzy logic method [132-134], Bayesian network method [135, 136], stochastic process [137, 138], etc.

In this work, the Paris' law is used as the system dynamic model describing the system

degradation mechanism, *i.e.*, the crack size propagation. As mentioned above, model-based prognostics method in this work is tackled with two steps, as shown in Figure 4-1. The first step, *i.e.*, **estimating the crack size and the unknown Paris' law parameters up to the present time (denoted by *S*) using a sequence of measurement data collected during the aircraft service life until *S***, is addressed by the Extended Kalman filter (EKF). The theory of the EKF and its implementation for estimating crack size and Paris' law parameters have been elaborated in chapter 2. This chapter mainly focuses on the second step, *i.e.*, **predicting the statistical distribution evolution of the crack size in the future *I* flight cycles based on the estimated states at time *S***. Once the crack size distribution after any future cycles is obtained, the probability that the crack size exceeds a threshold after specific cycles can be predicted.

A direct Monte Carlo method is first used then a first-order perturbation method is proposed for the same objective of predicting reliability while leading to significant computational cost saving. A series of numerical experiments are implemented for verifying the performance of the new proposed first-order perturbation method and the traditional Monte Carlo method under the same condition.

Reminder that at the end of the first phase (time *S*), the following information is considered available from EKF and will be used as initial conditions of the second phase for both the Monte Carlo method and the first-order perturbation method:

- expected value of the augmented state vector, $\hat{\mathbf{x}}_{au,S} = [\hat{a}_S \ \hat{m}_S \ \hat{C}_S]^{\mathrm{T}}$
- covariance of augmented state vector P_S.

According to the nomenclature in CHAPTER 3 in which the EKF is detailed, the posterior estimation for crack size and the material property parameters at time k should be denoted as $\hat{a}_{k|k}$, $\hat{m}_{k|k}$ and $\hat{C}_{k|k}$. However, they are denoted as $\hat{a}_k$, $\hat{m}_k$ and $\hat{C}_k$ for simplicity reason in CHAPTER 4 since it is not necessary to distinguish the prior and the posterior estimate in this chapter. All the estimates given by the EKF refer to the a posteriori estimates.

According to the philosophy of the EKF algorithm, the state vector $\mathbf{x}_{au,S} = [a_s \ m \ C]^{\mathrm{T}}$ follows a multivariate normal distributed with mean $\hat{\mathbf{x}}_{au,S}$ and covariance $\boldsymbol{P}_S$, presented as $\mathbf{x}_{au,S} \sim N(\hat{x}_{au,S}, P_S)$.

4.2 Monte Carlo method

Based on the available information obtained at time *S*, one straightforward approach for predicting the crack size at a future flight cycle is the Monte Carlo (MC) method. By using the MC method, *NC* samples of the state vector are drawn from their distribution at *k*=*S*, *i.e.* sample $\mathbf{x}_{au,S}^{i} \sim N(\hat{\mathbf{x}}_{au,S}, P_S)$ *(i=1,2,...NC)*. Propagate forward each sample through the *I* next flight cycles using the discretized Paris' law in Eq.(3-76) and at any cycle *k*, *NC* samples of state vector $\mathbf{x}_{au,k}^{j}$ (*i=1,2,...NC*) can be obtained. It should be noted that *k* is initialized with *S+1* and propagates until *k=S+I. i.e.*, *k=S+1, S+2, ... S+I.* A histogram of crack size can be built from these *NC* samples and some density estimation methods can be used to generate the approximation of the probability density function. Considering typical values that are used for maintenance and using goodness of fit tests like Kolmogorov–Smirnov, we found that the distribution of the crack size is always very close to a normal distribution. Accordingly, only the mean and standard deviation of crack size *a* at any cycle *k* are estimated, given in Eq.(4-1) and Eq.(4-2).

$$\mu_{ak}^{MC} = \frac{1}{NC}\sum_{i=1}^{NC} a_k^i \tag{4-1}$$

$$\sigma_{ak}^{MC} = \frac{1}{NC-1}\sum_{i=1}^{NC}\left(a_k^i - \mu_{ak}^{MC}\right)^2 \tag{4-2}$$

The symbol "*MC*" in the upper right corner represents "Monte Carlo" method in order to distinguish the "First-order perturbation" method that will be discussed in Section 4.3 The schematic diagram of predicting crack size using MC method that discussed above is illustrated in Figure 4-1.

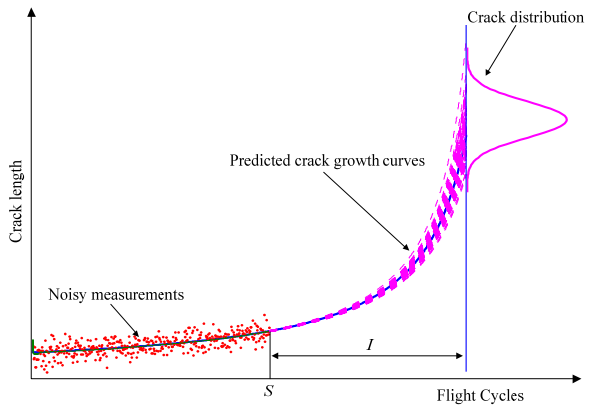

Figure 4-1 Schematic diagram of predicting crack size using Monte Carlo method

4.3 First-order perturbation method

The objective of the first-order perturbation method is the same as the MC method but it aims at predicting more efficiently the evolution of the crack size distribution from *S+1* to *S+I*. Let us define:

$$f_L(a,m,C,p) = C(A\frac{pr}{t}\sqrt{\pi a})^m \tag{4-3}$$

The Paris' law then becomes

$$a_k = a_{k-1} + f_L(a_{k-1},m,C,p_{k-1}) \tag{4-4}$$

Note that the index *k* starts from *S+1* and propagates until *S+I*, *i.e.*, *k=S+1, S+2,...S+I.* Before elaborating the first-order perturbation method, we first introduce the concept of "*Expected Trajectory*". A trajectory is a particular solution for a stochastic system, that is, with a particular instantiation for each random variable involved [102]. The term "*Expected Trajectory*" in this case refers to the trajectory that is obtained when the random variables assume their expected values. We use the hat symbol "–" to denote the expected value of a random variable, *e.g.*, $\overline{a}_k$ represents the expected value of a_k. For the problem discussed at hand, the "*expected trajectory*" of the crack size is the sequence $\{\overline{a}_k \,|\, k = S+1, S+2, ...S+I\}$ obtained as a solution of the following equation with zero process noise and with the expected value $\overline{a}_S$, $\overline{m}$, $\overline{C}$ and $\overline{p}$ as the initial conditions.

$$\overline{a}_k = \overline{a}_{k-1} + f_L(\overline{a}_{k-1},\overline{m},\overline{C},\overline{p}) \tag{4-5}$$

Due to the presence of random noise and uncertainties, a_k, m, C and p_k are considered random. Let the symbol "Δ" denote the perturbation from the expected values, then the real a_k, m, C and p_k can be modeled as

$$a_k = \overline{a}_k + \Delta a_k \tag{4-6}$$

$$m = \overline{m} + \Delta m \tag{4-7}$$

$$C = \overline{C} + \Delta C \tag{4-8}$$

$$p_k = \overline{p} + \Delta p_k \tag{4-9}$$

Δp_k is an uncertainty related to the cabin pressure differential, which varies from one flight cycle to another. On the other hand, Δm and ΔC are uncertainties related to the material properties of each panel and thus do not vary when time evolves. Recall the available information given by

EKF at k=S, which will be the initial condition in the following derivation.

$$\begin{bmatrix}\overline{a}_S & \overline{m} & \overline{C}\end{bmatrix}^T = \begin{bmatrix}\hat{a}_S & \hat{m}_S & \hat{C}_S\end{bmatrix}^T \tag{4-10}$$

$$\begin{bmatrix}\Delta a_S & \Delta m & \Delta C\end{bmatrix}^T \sim N(\mathbf{0}_{3\times 1}, P_S) \tag{4-11}$$

Subtracting Eq.(4-5) from Eq.(4-4), the perturbation of a_k is represented as

$$\Delta a_k = \Delta a_{k-1} + f_L(a_{k-1}, m, C, p_{k-1}) - f_L(\overline{a}_{k-1}, \overline{m}, \overline{C}, \overline{p}) \tag{4-12}$$

Since f_L is differentiable and the perturbation are considered to be small enough, we use a first order approximation to derive Eq.(4-12). Let $\boldsymbol{\lambda}_{k-1} = [\overline{a}_{k-1}, \overline{m}, \overline{C}, \overline{p}]$, which is a known vector, then Eq.(4-12) becomes

$$\Delta a_k = \Delta a_{k-1} + \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial a}\Delta a_{k-1} + \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial m}\Delta m + \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial C}\Delta C + \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial p}\Delta p_{k-1} \tag{4-13}$$

To make the Eq.(4-13) simpler we make the following substitution:

$$L_{k-1} = 1 + \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial a} \tag{4-14}$$

$$M_{k-1} = \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial m} \tag{4-15}$$

$$N_{k-1} = \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial C} \tag{4-16}$$

$$w_{k-1}^L = \frac{\partial f_L(\boldsymbol{\lambda}_{k-1})}{\partial p}\Delta p_{k-1} \tag{4-17}$$

in which $w^L_{k\text{-}1}$ is the process noise, a normal variable with mean zeros and standard deviation $\sigma_{k\text{-}1}$, calculated by Eq.(4-18).

$$\sigma_{k-1} = \frac{\partial f(\boldsymbol{\lambda}_{k-1})}{\partial p}\sigma_p \tag{4-18}$$

Then Eq.(4-13) becomes

$$\Delta a_k = L_{k-1}\Delta a_{k-1} + M_{k-1}\Delta m + N_{k-1}\Delta C + w_{k-1}^L, \quad S+1 \le k \le S+I \tag{4-19}$$

Eq.(4-19) provides a way to calculate the perturbation of crack size at any cycle. Recalling Eq.(4-11) that [Δa_S Δm ΔC] is multivariate normally distributed with zeros mean and known covariance P_S^+, then the distribution of Δa_k can be analytically calculated as the function of the distribution of [Δa_S Δm ΔC]. The following equations give the 3 steps forward derivation as an example and after k times iteration the analytical formula of calculating Δa_k is given in Eq.(4-20).

For simplicity, we use A_k, B_k and D_k to represent the coefficient of Δa_s, Δm and ΔC respectively while E_k denotes the noise term.

$$\Delta a_{S+1} = L_S \Delta a_S + M_S \Delta m + N_S \Delta C + w_S^L$$

$$\Delta a_{S+2} = L_S L_{S+1} \Delta a_S + (M_{S+1} + L_{S+1} M_S)\Delta m + (N_{S+1} + L_{S+1} N_S)\Delta C + (L_{S+1} w_S^L + w_{S+1}^L)$$

$$\Delta a_{S+3} = L_{S+2} L_S L_{S+1} \Delta a_S + (L_{S+2}(M_{S+1} + L_{S+1} M_S) + M_{S+2})\Delta m + (L_{S+2}(N_{S+1} + L_{S+1} N_S) + N_{S+2})\Delta C + (L_{S+2}(L_{S+1} w_S^L + w_{S+1}^L) + w_{S+2}^L)$$

$$\vdots$$

$$\Delta a_k = A_{k-1} \Delta a_S + B_{k-1} \Delta m + D_{k-1} \Delta C + E_{k-1} \tag{4-20}$$

Note that in Eq.(4-20), Δa_S , Δm and ΔC are stationary random variables whose statistical distributions are time invariant. A_k, B_k and D_k are deterministic and evolve with time, which are calculated recursively with their initial values A_S=L_S, B_S=M_S, D_S=N_S, as shown in Eq.(4-21) to Eq.(4-23). E_k is the only random variable which has a distribution varying from cycle to cycle and is derived recursively by Eq.(4-24). Since E_k is a linear combination of w^L_k, k=S, S+1, S+2,…, it is a normal variable such that $E_k \sim N(0, F_k)$, in which F_k represents the variance of E_k. Using the recurrence of Eq.(4-24), F_k can be obtained recursively with its initial value F_s=σ_s, given by Eq.(4-25). Note that w^L_k and σ_k in Eq.(4-24) and Eq.(4-25) refer to Eq.(4-17) and Eq.(4-18) respectively.

$$A_k = L_k A_{k-1} \tag{4-21}$$

$$B_k = L_k B_{k-1} + M_k \tag{4-22}$$

$$D_k = L_k D_{k-1} + N_k \tag{4-23}$$

$$E_k = L_k E_{k-1} + w_k^L \tag{4-24}$$

$$F_k = L_k^2 F_{k-1} + \sigma_k^2 \tag{4-25}$$

Since A_k B_k D_k are deterministic values and Δa_s, Δm, ΔC and E_k are random variables, Eq.(4-20) is rewritten as matrix form as $\Delta a_k = \boldsymbol{Y}_k \boldsymbol{\beta}_k$, where $\boldsymbol{Y}_k$=[A_k B_k D_k 1] and $\boldsymbol{\beta}_k$ =$[\Delta a_S \ \Delta m \ \Delta C \ E_k]^T$. Given that $[\Delta a_S \ \Delta m \ \Delta C]^T \sim N\left(\mathbf{0}_{3\times1}, P_{ks}^+\right)$ and $E_k \sim N(0, F_k)$, $\boldsymbol{\beta}_k$ is multivariate normal vector such that $\boldsymbol{\beta}_k \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, in which $\boldsymbol{\mu} = \left[\mathbf{0}_{4\times1}\right]$, $\boldsymbol{\Sigma} = diag(P_{ks}^+, F_k)$. Therefore, Δa_k is normally distributed with

mean $\boldsymbol{Y}_k\boldsymbol{\mu}$ and variance $\boldsymbol{Y}_k\boldsymbol{\Sigma}\boldsymbol{Y}_k^T$, which are calculated analytically,

$$\boldsymbol{Y}_k\boldsymbol{\mu} = 0 \tag{4-26}$$

$$\boldsymbol{Y}_k\boldsymbol{\Sigma}\boldsymbol{Y}_k^T = [A_k\ B_k\ D_k]P_s^+[A_k\ B_k\ D_k]^T + F_k \tag{4-27}$$

Given that $a_k = \overline{a}_k + \Delta a_k$ and $\overline{a}_k$ is deterministic, a_k is a normal variable that $a_k \sim N(\mu_{ak}^L, \sigma_{ak}^L)$, in which

$$\mu_{ak}^L = \overline{a}_k \tag{4-28}$$

$$\sigma_{ak}^L = \sqrt{\boldsymbol{Y}_k\boldsymbol{\Sigma}\boldsymbol{Y}_k^T} \tag{4-29}$$

The above formulas allow computing analytically the evolution of the crack size distribution from cycle *S+1* to cycle *S+I*. As summary, the process of calculating the mean and standard deviation of a_k is reported in Table 4-1.

Table 4-1 First-order perturbation procedure summary

1. Initialization

$A_S = L_S = 1 + \partial f_L(a_S, \overline{m}, \overline{C}, \overline{p})/\partial a$

$B_S = M_S = \partial f_L(a_S, \overline{m}, \overline{C}, \overline{p})/\partial m$

$D_S = N_S = \partial f_L(a_S, \overline{m}, \overline{C}, \overline{p})/\partial C$

$F_S = \sigma_S = \partial f_L(a_S, \overline{m}, \overline{C}, \overline{p})/\partial p$

2. for k=S+1,…S+I

2.1 Propagate the crack size:

$\overline{a}_k = \overline{a}_{k-1} + f_L(\overline{a}_{k-1}, \overline{m}, \overline{C}, \overline{p})$

2.2 Calculate L_k, M_k, N_k and σ_k :

$L_k = 1 + \partial f_L(\overline{a}_k, \overline{m}, \overline{C}, \overline{p})/\partial a$

$M_k = \partial f_L(\overline{a}_k, \overline{m}, \overline{C}, \overline{p})/\partial m$

$N_k = \partial f_L(\overline{a}_k, \overline{m}, \overline{C}, \overline{p})/\partial C$

$\sigma_k = \partial f_L(\overline{a}_k, \overline{m}, \overline{C}, \overline{p})/\partial p$

2.3 Calculate A_k, B_k, D_k and F_k:

$A_k = L_k A_{k-1}$

$B_k = L_k B_{k-1} + N_{k-1}$

$D_k = L_k D_{k-1} + M_{k-1}$

$F_k = L_k^2 F_{k-1} + \sigma_{k-1}^2$

2.4 Calculate the mean and standard deviation of a_k:

$\mu_{ak}^L = \overline{a}_k$

$\sigma_{ak}^L = \sqrt{[A_k\ B_k\ D_k]P_s^+[A_k\ B_k\ D_k]^T + F_k}$

End for

Now the mean and standard deviation of crack size at any future flight cycle from up to which the measurement data are collected are calculated analytically. The schematic diagram of the proposed EKF coupled first-order perturbation method is given in Figure 4-2. The dashed line in

the second phase represents the trajectory of the mean of the crack size estimated by the first-order perturbation method, *i.e.*, $\{\mu_k | k = S+1, S+2, \ldots\}$. For illustrative purpose, the crack size distribution at two arbitrary flight cycles *k1* (based on estimation μ_{k1} and σ_{k1}) and *k2* (based on μ_{k2} and σ_{k2}) are illustrated in Figure 4-2.

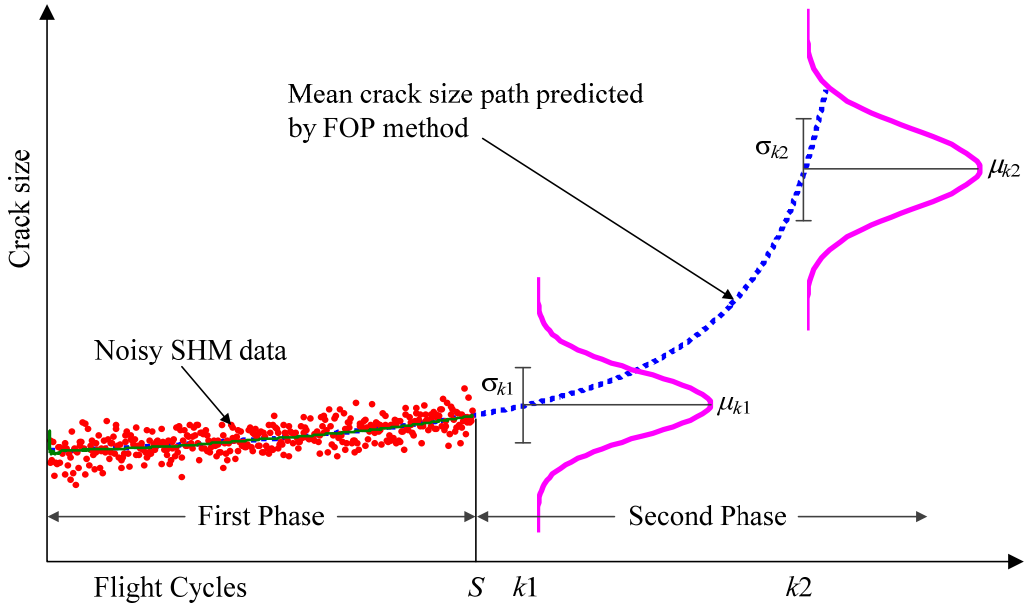

Figure 4-2 Schematic diagram of model-based prognostics

4.4 Verification of first-order perturbation method through numerical study

This section investigates the performance of the proposed first-order perturbation method. Firstly, the test case involving fatigue growth on a fleet of aircraft is described. Then the prediction comparisons between the first-order perturbation method and the traditional Monte Carlo method are provided.

4.4.1 Numerical experiment design

The parameters involved in the numerical application are classified into three categories: (1) parameters related to panel-to-panel variability; (2) parameters related to EKF algorithm; and (3) aircraft geometry parameters.

A fuselage is composed of several hundred panels. If all the manufactured panels are exactly the same and these panels work under exactly the same conditions and environment, then the panels will degrade totally identically. However, in practice, due to intrinsic variability in fatigue

propagation, the difference of material quality and the various working environment etc, uncertainties are considered in the degradation process. For example, each panel can have its own initial crack size, material properties and loading condition. We refer to such uncertainties related to variation between individual panels as panel-to-panel uncertainty. In this study, the panel-to-panel uncertainty is modeled by assuming that the initial crack size a_0, the material properties parameters m and C follow a predefined statistical distribution. Specifically, a_0 is assumed log normally distributed while m and $log_{10}C$ are assumed to follow a multivariate normal distribution with a correlation coefficient of -0.8, based on the research indicating that m and $log_{10}C$ are negatively linearly correlated [139-141].

For each individual panel, its a_0, m and C are drawn from the above-mentioned distributions representing the panel-to-panel uncertainty discussed in the above paragraph. These drawn values of a_0, m and C are regarded as the unobserved "true" values that need to be estimated by the EKF from measurements. The noisy measurements are simulated by adding Gaussian white noise to the "true" crack size value. This data is used as the actual measurements for the estimation process. During the EKF estimation process, only these noisy measurements data are available, from which the crack size and the material properties parameters should be estimated. Besides the measurement uncertainty, we also take into account the load uncertainty, *i.e.*, the pressure differential, which varies at every flight cycles and is modeled as a normal variable. The EKF algorithm needs to be initialized. In particular, the starting point of the estimation algorithm has to be provided. In our case, the starting point is drawn randomly from uniform distribution centered at the "true" value of a_0, m and C respectively and having a range of 50% around these true values. Once this initial guess has been provided, EKF makes estimation of $\hat{a}_k$ $\hat{m}_k$ and $\hat{C}_k$ at each cycle k. Note that to avoid a systematic bias, the starting point of the EKF algorithm is different (drawn randomly) for each panel in the aircraft. As for the initial error covariance matrix $\boldsymbol{P}_0$, it is chosen depending on how much confidence one has to the initial estimates.

Our potential application object is a typical short rang commercial aircraft such as A320 or B737 with a general lifetime of 60000 flight cycles. The values of the time invariant geometry parameters defining the fuselage (*e.g.* fuselage radius, panel thickness) used in the numerical experiment are related to such aircrafts. For recall, we define a correction factor A, which intends

to account for the fact that the fuselage is modeled as a hollow cylinder (without stringers and stiffness). A numerical value has been chosen from the studies by [142]. All the values of the above-mentioned three types of parameters are reported in Table 4-2.

Table 4-2 Parameters used in numerical study

Description	Notation	Type	Value
Parameter of panel-to-panel uncertainty			
Initial crack size in meter	a_0	Lognormal	LnN (0.2e-3, 0.07e-3)
Paris' law parameter	m C	Multivariate normal	$N(\mu_m, \sigma_m, \mu_C, \sigma_C, \rho)$
Mean value of m	μ_m	-	3.65
Standard deviation of m	σ_m	-	0.22
Mean value of C	μ_C	-	Log_{10}(2.75e-10)
Standard deviation of C	σ_C	-	0.25
Correlation coefficient of m and C	ρ	-	-0.8
Parameter related to EKF			
Estimated initial crack length in meter	$\hat{a}_0$	Uniform	U[a_0-25% a_0, a_0+25% a_0]
Estimated initial m	$\hat{m}_0$	Uniform	U[m_0-25% m_0, m_0+25% m_0]
Estimated initial C	$\hat{C}_0$	Uniform	U[C_0-25% C_0, C_0+25% C_0]
Initial error covariance matrix	P_0	-	diag ((1e-4)^2,(1e-2)^2,(1e-10)^2)
Measurement noise variance	Ra	-	(30% a_0)^2
Pressure differential in MPa	p	Normal	N(0.06, 0.003)
Aircraft geometry parameter			
Fuselage radius in meter	r	-	1.95
Panel thickness in meter	t	-	2e-3
Correction factor	A	-	1.25

To take into account the panel-to-panel uncertainty, we assume $N=100$ fuselage panels. The "true" a_0, m and C are simulated by drawing $a_0^{(i)}$, $m^{(i)}$, $C^{(i)}$ $(i=1,2,...N)$ from their respective distribution given in Table 4-2 and assigned to each panel. $[a_0^i\ m^i\ C^i]^T$ forms the initial condition of ith panel. For each panel, the numerical experiment is implemented as following. (1) Compute the true crack size $a_k^{(i)}$ at each cycle k for each panel i until a threshold a_{th}=39.5mm is reached and then determine for each panel its service life L^i (*i.e.* $L^i=k$ such that $a_k^{(i)}=a_{th}$). a_{th} represents the crack size threshold which should not be exceeded in order to maintain the panel integrity. Any panel with a crack size reaching a_{th} is regarded as threatening to the aircraft safety and need to be repaired or replaced. Among all the panels, we are only concerned about the "dangerous panels" whose service life L^i is shorter than 60000 flight cycles since 60000 cycles is the lifetime of a typical short-range commercial aircraft. (2) For each "dangerous panel", apply the EKF on the noisy measurements to carry out the state-parameter estimation from cycle $k=1$ until $k=L^i-J$. Here J is set to be 4000 flight cycles since this is a typical scheduled maintenance interval for the short-range commercial aircraft fuselage panels. (3) For each "dangerous panel", based on the estimated state vector and the a posteriori error covariance matrix given by EKF at $k=L^i-J$, the prediction of the statistical

distribution evolution of the crack size in the last J flight cycles is implemented. This prediction is computed by both first-order perturbation and MC methods. For MC method, NC=500 samples are used. More precisely, for a given threshold, we select the panels that reach this level before the end of the aircraft's life (*i.e.* 60000 cycles). For each of the dangerous panel, we predict the evolution of crack size distribution by using MC and first-order perturbation method in the last J cycles before the end of the panel's service life. This allows validating the proposed first-order perturbation method because we deal with the most non-linear part of the Paris' law. The steps of the numerical implementation are summarized below.

- *Step1*: For each panel, draw the samples of true initial crack size and the parameters respectively: $a_0^{(i)} \sim LnN(0.2e\text{-}3,\ 0.07e\text{-}3)$, $[m^{(i)}, C^{(i)}] \sim N(\mu_m, \sigma_m, \mu_C, \sigma_C, \rho)$, $i=1,2,\ldots N$.
- *Step 2:* Compute the true crack size $a_k^{(i)}$ at each cycle k for each panel i. Then determine for each panel its service life L^i (*i.e.* the number of cycle when the crack size reaches the threshold a_{th}: $L^i=k$ such that $a_k^{(i)}=a_{th}$). Consider only dangerous panels (*i.e.*, $L^i<60000$).
- *Step 3*: For each dangerous panel, compute the simulated measurements data from $k=1$ to $k=L^i\text{-}J$. Apply EKF algorithm on these measurements and obtain a posteriori estimated state vector $\hat{\mathbf{x}}_{au,k}^{(i)}$ and the a posteriori covariance matrix $P_k^{(i)}$, $k=1,2,\ldots L^i\text{-}J$
- *Step 4*: Predict the last J cycles (from $k=L^i\text{-}J+1$ to L^i) based on the previously calculated $\hat{\mathbf{x}}_{au,k}^{(i)}$ and $\boldsymbol{P}_k^{(i)}$, at $k=L^i\text{-}J$ using the MC method (see Section 4.2)
- *Step 5*: Predict the last J cycles (from $k=L^i\text{-}J+1$ to L^i) based on the previously calculated $\hat{\mathbf{x}}_{au,k}^{(i)}$ and $\boldsymbol{P}_k^{(i)}$ at $k=L^i\text{-}J$ using the first-order perturbation method (see Section 4.3)

4.4.2 Results and discussion

There are 8 "dangerous panels" over the population of 100, which means that the crack in these 8 panels will grow exceeding the threshold before the end of the aircraft's life. The initial condition of the each dangerous panel and its lifetime are listed in Table 4-3. Figure 4-3 illustrates all the 100 pairs of $\{m,C\}$ with plotting the dangerous one in asterisk.

Table 4-3 List of dangerous panels

No.	[a_0 m C]			Corresponding service life
1	[1.78e-4	3.31	8.45e-10]	50723
2	[1.93e-4	3.75	4.12e-10]	58255
3	[1.99e-4	3.31	7.45e-10]	53332
4	[2.27e-4	3.43	6.11e-10]	51270
5	[2.48e-4	3.31	8.20e-10]	42142
6	[2.50e-4	3.33	5.93e-10]	55795
7	[2.84e-4	3.40	8.17e-10]	33617
8	[2.87e-4	3.79	3.18e-10]	50628

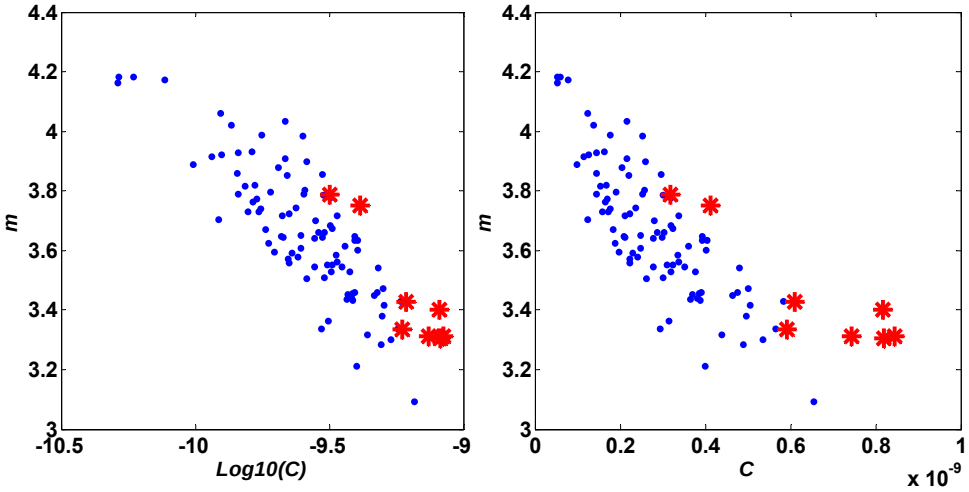

Figure 4-3 Illustration of 100 pairs of {m, C}

4.4.2.1 State-parameter estimation

This part presents and analyzes the state-parameter estimation results given by the EKF through reporting the results of panel No.7 as an example since it is not necessary to present all eight dangerous panels. Panel No.7 is representative since it has the shortest service life (33617 cycles) and is thus the most critical among the 8 panels. The crack size estimation of panel No.7 is shown in Figure 4-4. Note that the last 4000 cycles prior to the end of the service life is preserved to verify the performance of first-order perturbation method and will not be plotted. Therefore, the time range in Figure 4-4 is from one to 29618 (which equals to 33617 minus 4000). The left subplot is the comparison between true crack size a_k and the EKF estimate $\hat{a}_k$. Each point represents a noisy measurement used in the EKF. Note that only one point every 100 cycles is represented in order not to overload the figure. It can be seen that the estimated crack length (dashed line) fits very well the true one (solid line) even with polluted measurements (solid points). The right subplot gives the relative error between a and $\hat{a}_k$, *i.e.* $\left|\hat{a}_k - a\right| / a$, which begins with a high value of 6% due to a large biased initial estimate for crack size and decreases rapidly to nearly 1% after 100 flight cycles.

Then it fluctuates to a small extend but keep very low (less than 2%) and overall declines progressively.

The parameter estimation results are shown in Figure 4-5, which indicates that the m and C converge to their true values with time evolves. Figure 4-5 also gives the 95% confidence interval of the estimates, from which it can be observed that the spread decreases over time.

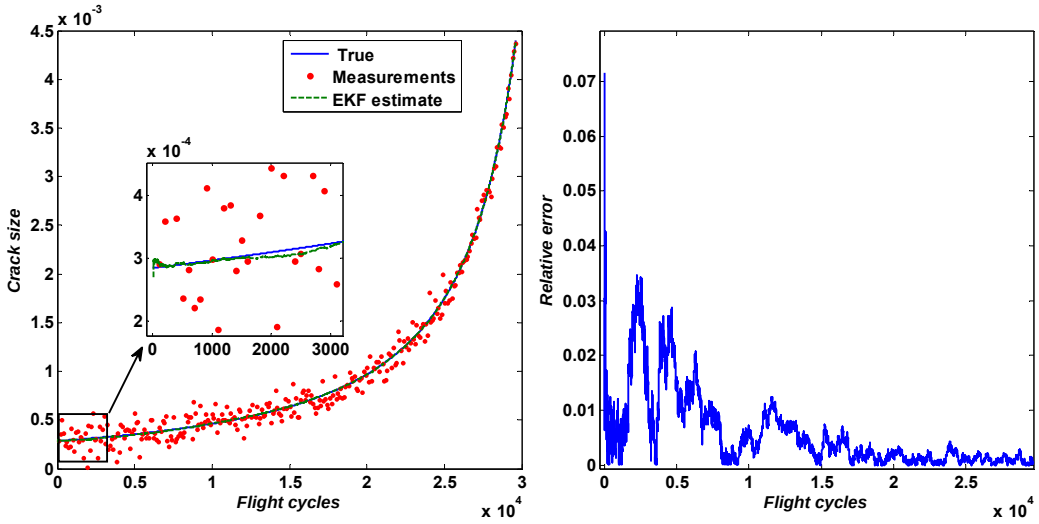

Figure 4-4 Crack size estimation by EKF-panel No.7

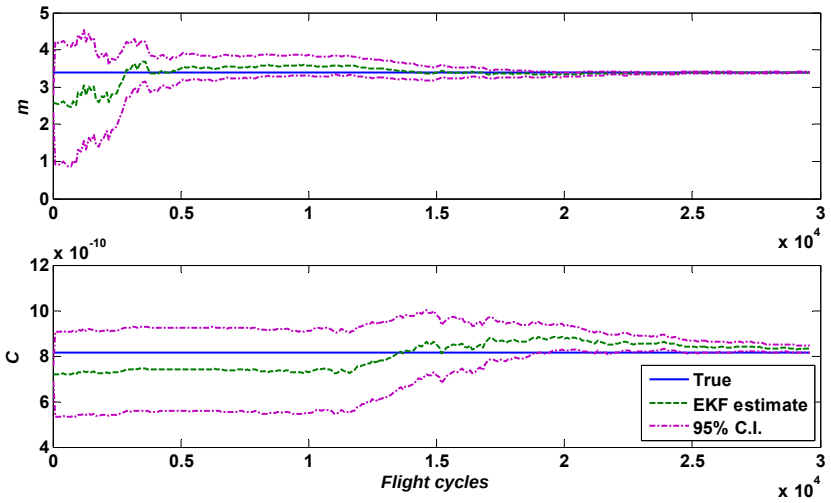

Figure 4-5 Convergence process of m and C - panel No.7

4.4.2.2 Crack size evolution prediction

This part shows and analyzes the prediction results of all the 8 dangerous panels. The aim is to compare the predicted distribution of crack size given by MC (5000 samples are used in MC) and the first-order perturbation methods in the last J=4000 cycles before the end of the panel's service life. According to the nature of the EKF, the crack size is normally distributed that is characterized

by two parameters, mean and standard deviation. Therefore, comparison of these two methods can be implemented through comparing the error between μ_{ak}^{L} and μ_{ak}^{MC}, σ_{ak}^{L} and σ_{ak}^{MC} *(k=L*^{i}*-J+1, L*^{i}*-J+2,… L*^{i}*)*, namely, the mean and standard deviation given by MC and the first-order perturbation methods, respectively. The trajectory comparison of μ_{ak}^{L} and μ_{ak}^{MC}, σ_{ak}^{L} and σ_{ak}^{MC} for all the 8 dangerous panels during the last 4000 flight cycles are given from Figure 4-6 to Figure 4-9. It can be seen that the trajectory of mean and standard deviation given by the first-order perturbation method fit well the ones given by the Monte Carlo method. This implies the proposed method gives satisfactory performance.

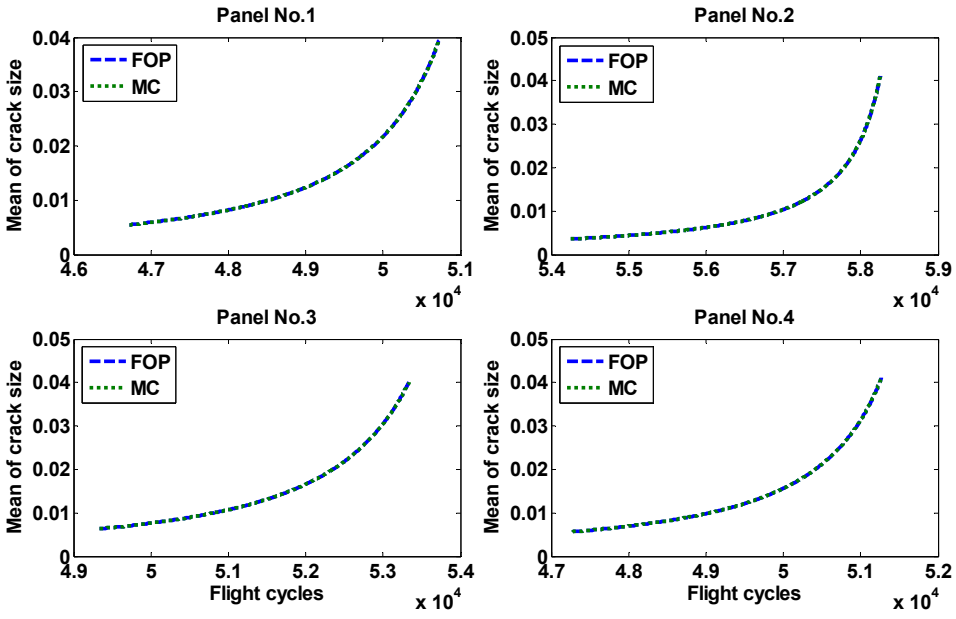

Figure 4-6 The trajectory comparison of μ_{ak}^{L} and μ_{ak}^{MC} for panel Nos.1-4

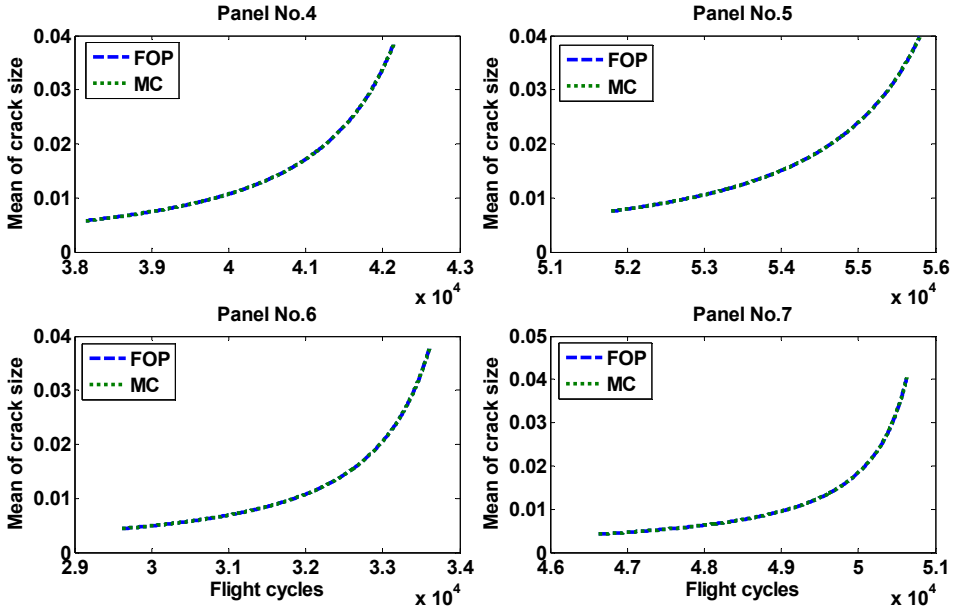

Figure 4-7 The trajectory comparison of μ_{ak}^{L} and μ_{ak}^{MC} for panel Nos.5-8

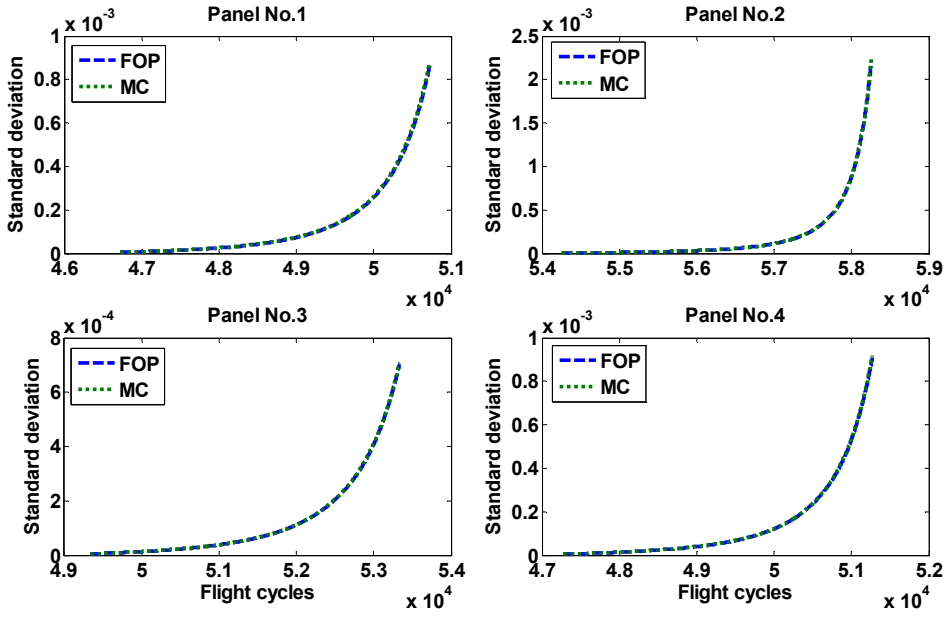

Figure 4-8 The trajectory comparison of σ_{ak}^{L} and σ_{ak}^{MC} for panels Nos.1-4

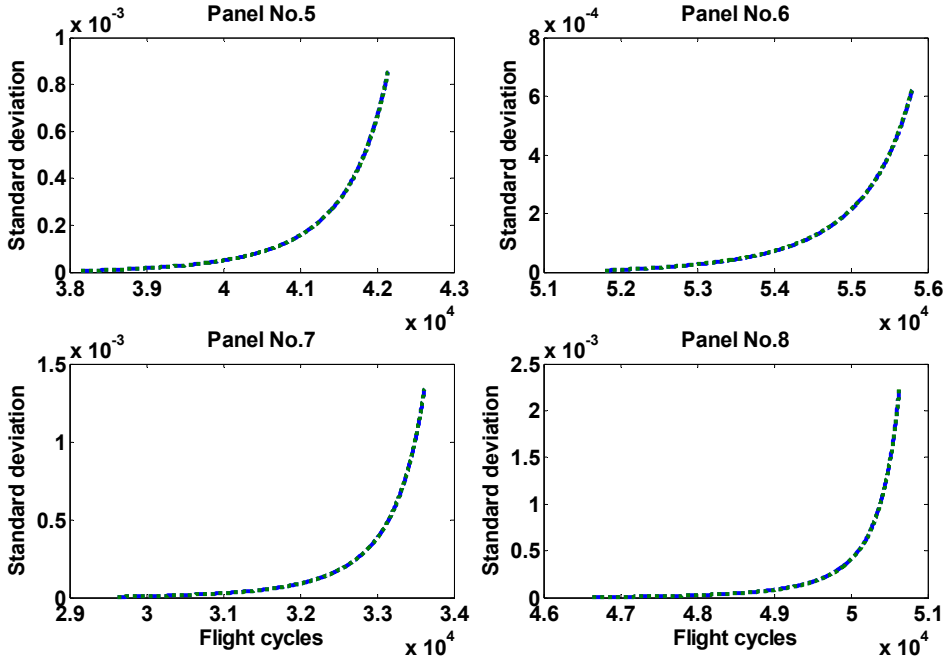

Figure 4-9 The trajectory comparison of σ_{ak}^{L} and σ_{ak}^{MC} for panels Nos.5-8

The results of the maximal error in the last J flight cycles are reported in Table 4-4. Note that the maximal error is obtained at the last cycle L^i. The first column is the index of the 8 dangerous panels whose initial condition and the corresponding service life have been presented in Table 4-3. According to the results, it can draw the following comments. (1) first-order perturbation method gives very close results to that of MC with maximal relative error 8.11e-2% for mean (No.4) and 1.87% for standard deviation (No.1). (2) For panel No. 2, 3, 4, 6, 8, the predicted means by both MC and first-order perturbation methods are greater than the true crack size, which equals to the

threshold a_{th}. This conservative prediction leads to a right decision of repairing or replacing the panels thereby ensuring the safety of the aircrafts. Note that even on panels 1, 5, and 7, for which the mean estimation is not conservative, when considering the 95% confidence interval, the prediction remains by a significant margin on the conservative side. The last column presented all the 95% confidence interval of the predicted mean. (3) The first-order perturbation method shows its great advantage in computational cost over MC (5000 samples are used) with 0.144s versus 889.254s, on a desktop with a processor Intel (R) Core(TM) i3-4130 CPU 3.4GHz. The computational savings are particularly important to process a large number of aircrafts.

Table 4-4 Comparison of the mean and standard deviation of the crack size of each "dangerous panel" given by MC and first-order perturbation method at the last flight cycle L^i

No.	μ_a^L (mm)	μ_a^{MC} (mm)	$\lvert \mu_a^L - \mu_a^{MC} \rvert / \mu_a^{MC}$ (%)	σ_a^L	σ_a^{MC}	$\lvert \sigma_a^L - \sigma_a^{MC} \rvert / \sigma_a^{MC}$ (%)	True crack size (mm)	95% C.I. based on μ_a^L and σ_a^L (mm)
1	39.4	39.5	6.01e-2	8.63e-4	8.79e-4	1.84	39.5	[37.7 41.1]
2	41.2	41.3	2.32e-1	2.21e-3	2.23e-3	0.95	39.5	[36.9 45.5]
3	40.0	40.0	2.17e-2	7.02e-4	7.06e-4	0.65	39.5	[38.6 41.4]
4	41.1	41.1	8.10e-2	9.09e-4	9.18e-4	1.01	39.5	[39.3 42.9]
5	38.5	38.5	6.17e-3	8.54e-4	8.51e-4	0.25	39.5	[36.8 40.2]
6	39.8	39.8	1.70e-2	6.19e-4	6.22e-4	0.58	39.5	[36.8 41.0]
7	38.2	38.2	5.72e-2	1.34e-3	1.34e-3	0.24	39.5	[35.6 40.8]
8	40.3	40.4	3.17e-1	2.22e-3	2.23e-3	0.37	39.5	[36.0 44.7]

4.5 Summary

In this chapter, an EKF coupled first-order perturbation method for model-based prognostics for fatigue damage propagation in fuselage panels is proposed. Both the material property parameters and the initial crack size are unknown and both the measurements of crack size and loading conditions are affected by uncertainties. The presented model-based prognostics method includes two sequential phases (1) state-parameter identification from noisy measurements data using the EKF algorithm and (2) future damage prediction using the proposed first-order perturbation method. The results show that the EKF performs well in state-parameter estimation process during which the crack size has been well estimated and the estimated mean values of the two unknown material property parameters converge to their true values over time. In the second phase, a new first-order perturbation method is employed to predict analytically the evolution of the crack size distribution in some future flight cycles from the present estimations given by the EKF. The results are compared to that of Monte Carlo method. The comparison shows that the first-order perturbation method gives similar prediction performance to MC with a great saving in

computational cost. Both Monte Carlo and first-order perturbation methods provide right decisions of repairing or replacing panels. This indicates that the first-order perturbation method is effective and could be used to replace Monte Carlo method in prognostics problem of fuselage panel fatigue crack growth. This is of particular interest dealing with large number of aircrafts, for example, planning maintenance strategies for a fleet of aircrafts.

CHAPTER 5 MAINTENANCE STRATEGIES BASED ON MODEL-BASED PROGNOSTICS METHOD

5.1 Introduction

Maintenance costs of commercial aircraft typically account for 10% to 20% of the aircraft operating costs. Developing efficient maintenance can be an important way for airlines to allow a new profit growth. An increasing number of researchers take into account the issues relating to safety, such as product reliability prediction and end-of-life prediction, during the early phase of product design [143, 144] to optimize the maintenance planning.

Aircraft maintenance can be classified into airframe maintenance and engine maintenance. The airframe maintenance that deals with non-structural items such as furniture and electronic systems is called non-structural airframe maintenance [145] while the one that concerns the fatigue damage in the structural section such as fuselage panels is called structural airframe maintenance. In this paper, maintenance refers to structural airframe maintenance while engine and non-structural airframe maintenance are out of scope of this study. Fatigue damage is one of the major failure modes of structures. Especially, repeated pressurization/depressurization cycles during take-off and landing cause many loading and unloading cycles, which can lead to fatigue cracks in the fuselage panels. The fuselage structure is designed to withstand small cracks. However, if left unattended, the cracks will grow progressively and finally cause panel failure. It is important to inspect the aircraft regularly so that all damage that have the risk of leading to panel fatigue failure can be repaired.

Traditionally, aircraft maintenance is performed corresponding to a fixed schedule. At the time of scheduled maintenance, the aircraft is sent to the maintenance hangar to undergo a series of maintenance activities including both engine and airframe maintenance. Structural airframe maintenance is a subset of these activities that focuses on detecting the large cracks that can possibly threaten the safety of the aircraft. Structural airframe maintenance is often implemented by the time-consuming techniques such as non-destructive inspection (NDI), general visual inspection, detail visual inspection (GVI), etc. Since the frequency of scheduled maintenance for commercial aircrafts is designed for a low probability of failure of 1e-7 [142], it is very likely that no large crack

exists during earlier life of an aircraft. Even so, the intrusive inspection by NDI or GVI for all panels needs to be performed to guarantee the absence of critical cracks that otherwise cause fatigue failure.

With the progress of the sensor technology, data acquisition and storage techniques as well as advanced data processing algorithms, structural health monitoring (SHM) system are gradually being introduced to aviation industry [2, 3, 8]. SHM employs a sensor network that are sealed inside the aircraft structures like fuselage, rotor blades, landing gears, bulkheads, etc., to monitor the damage state of these structures. The network may consist of a wide variety of sensor node types, strategically distributed in the structure to maximize data collection. The data is collected and processed using damaged detection algorithm and feature extraction methods to obtain the damage state of the structures. Once the damages in structures are able to be monitored automatically and continuously (or as frequently as needed in practice) by SHM system, more advanced condition-based maintenance (CBM) can be implemented instead of the scheduled maintenance, which could help save maintenance cost by planning maintenance policy based on the actual condition of the aircraft, rather than fixed schedules and inspection routines that might not be necessary, and thereby reduce aircraft's downtimes.

A lot of attention has been paid to CBM in literature and more recently to predictive maintenance (PdM). In some literature, PdM is referred to as CBM while in fact they have significant differences. Indeed, CBM and predictive maintenance share some characteristics in common that both of them rely on the damage-associated data collected by the SHM systems. The difference lies in that in CBM, the maintenance decision-making relies only on current damage level while the predictive maintenance makes use, in additional to current damage information, of a *prognostics index* to support the decision-making. The remaining useful life (RUL) is the most common prognostics index. By taking into account the RUL, the predictive maintenance strategy updates dynamically the maintenance schedule and predicts the time for next maintenance. Such predictive maintenance policies that adopt RUL as a prognostics index can be found in [23, 146, 147]. On the other hand, in some situations especially when a fault or failure is catastrophic, the periodic scheduled maintenance is necessary. Generally, the maintenance schedule is recommended by the manufacturer in concertation with safety authorities. Arbitrarily deciding the maintenance time only based on the estimated RUL without taking into account the periodic maintenance may be too disruptive to the original maintenance plan. The arbitrary trigged maintenance is unexpected and less optimal from

the economic point of view due to less notification in advance, *e.g.*, the absence of maintenance crew, the lack of spare parts, etc. In the above cases, instead of using the estimated RUL to dynamically decide the next maintenance time, it would be more desirable to predict the probability that a component operates normally before some future time (*e.g.*, before next scheduled maintenance). In other words, use the "system reliability in future time" as the prognostics index. This makes sense, since it helps make a decision whether or not the system needs to be repaired at current maintenance so that one can skip some unnecessary periodic maintenance. The predictive maintenance policy that incorporates the "future system reliability" information for supporting decision making can be found in [23, 24, 148].

For the application of structural airframe maintenance for an aircraft fleet in airlines, Pattabhiraman et.al proposed two type of CBM strategies aiming at reducing the number of traditional scheduled maintenance [142]. In his solution, the SHM system is proposed to monitoring the damage state of the aircraft, by which the damage state could be assessed as frequently as needed, not only at the time of scheduled maintenance. One strategy is purely CBM, *i.e.*, triggering maintenance anytime when needed only based on the current health state of the fuselage panels without considering the recommended maintenance schedule, while the other strategy does take into account the schedule. In Pattabhiraman's strategies, the decisions of asking a maintenance stop as well as replacing a panel depend on some fixed thresholds, thus called threshold-based maintenance strategies. These thresholds are determined for the entire fleet of aircraft to ensure a desirable level of reliability.

This chapter, building on Pattabhiraman's work, proposes two types of predictive maintenance strategies, which are based on the prognostics methods discussed in CHAPTER 4, *i.e.*, the predictive information is used in the maintenance planning. The main difference lies in the repair policy of how the predictive information is used. In general, in the first type of predictive maintenance policy, the decision of whether to repair a panel or not at a maintenance stop depends on the predicted crack size distribution in some future flight cycles. In the second type of predictive maintenance policy, the repair decision depends on the ratio that scheduled maintenance cost over unscheduled cost, hence is called cost driven predictive maintenance (CDPM).

This chapter is organized as follows. Section 5.2 presents the new maintenance idea in which the SHM systems are introduced. In Section 5.3, some prerequisite concepts and definitions needed

before introducing the predictive maintenance policy are given. Then the first type of predictive maintenance, which includes two variants PdM and PdM-skip, are detailed in Section 5.4 and the second predictive maintenance, called cost driven predictive maintenance, is given in Section 5.5. Numerical case study is implemented in Section 5.6. The traditional maintenance and the proposed predictive maintenance strategies are compared. Conclusions are drawn in Section 5.7.

5.2 Traditional scheduled maintenance works in tandem with unscheduled maintenance

Currently, aircraft maintenance is performed on a fixed schedule. Suppose that the aircraft undergoes the routine maintenance according to a schedule $T_n=T_1+(n-1)\Delta T$, where n=1,2,…, is the number of maintenance stop, T_n denotes the cumulative flight cycles at the n-th stop, T_1 is the number of flight cycles from the beginning of the aircraft's lifetime to the first scheduled maintenance stop. ΔT is the interval between two consecutive scheduled maintenance stops after T_1. Note that $T_1 > \Delta T$ because fatigue cracks propagate slowly during the earlier stage of the aircraft lifetime. With usage and ageing, the aircraft needs maintenance more frequently. The schedule is defined by aircraft manufacturers in concertation with certification authorities and aims at guaranteeing the safety using a conservative scenario [149]. For a given safety requirement, this schedule may not be optimal, in terms of minimizing maintenance cost. Indeed a specific aircraft may differ from the fleet's conservative properties used in calculating the maintenance schedule and possibly require fewer maintenance stops.

For a short range commercial aircraft with a typical lifetime of 60000 flight cycles, the maintenance schedule is designed such that the first maintenance is performed after 20000 flight cycles and the subsequence maintenance is every 4000 cycles until its end of life, thus adding up to 10 scheduled maintenance throughout its lifetime, as shown in Figure 5-1.

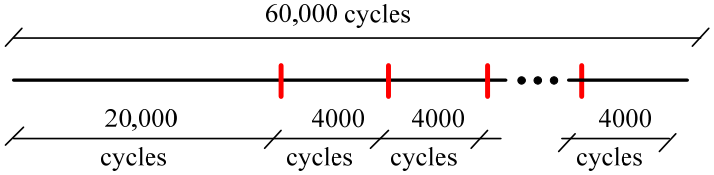

Figure 5-1 Schedule of the scheduled maintenance

By employing the SHM system, the damage state could be traced as frequently as needed and the maintenance can be asked at any time according to the aircraft health state rather than a fixed

schedule. This causes an unscheduled maintenance that could happen anytime throughout the aircraft's lifetime and generally occurs out of all scheduled maintenances time instant. Triggering a maintenance stop arbitrarily is distruptive to the current scheduled maintenance due to the factors of no advance notification, *e.g.*, less preparation of the maintenance team, unavailable tools, lack of spare parts etc. These factors lead unscheduled maintenance more expensive. Therefore, we attempt as much as possible to plan the structural airframe maintenance at the time of the scheduled maintenance and avoid the unscheduled maintenance in order to reduce the cost.

On the other hand, it makes sense to skip some scheduled maintenance stops. Since the frequency of scheduled maintenance for commercial aircrafts is designed for a low probability of failure (1e-7), it is very likely that no large crack exists during earlier life of the aircraft. Even so, according to the traditional scheduled maintenance, the intrusive inspection by NDI, GVI or DVI for all panels needs to be performed. Thanks to the on-board SHM system, the real time damage assessment could be done on site instead of exactly in the hangar, leading to the possibility of skipping some unnecessary scheduled maintenance if there are no life-threating cracks on the aircraft. If the crack missed at schedule maintenance grows too big to threaten the safety between two consecutive scheduled maintenances, the unscheduled maintenance is triggered at once. The frequent monitoring of the damage status would ensure the same level of reliability as scheduled maintenance. One thing that needs to be noted is that our objective is to plan the structural airframe maintenance by using predictive maintenance while the engine and non-structural airframe maintenance are always performed at the time of scheduled maintenance.

In summary, it is beneficial that in civil aviation industry the traditional scheduled maintenance works in tandem with the unscheduled maintenance.

5.3 Prerequisite concepts and definitions

In this section, some prerequisite concepts and definitions are given first. These concepts and definitions will be used in the following parts of introducing the two types of predictive maintenance policies.

5.3.1 Two types of uncertainties considered in the maintenance policy

Two different kinds of uncertainties are considered, which are called *panel-to-panel uncertainty* and *epistemic uncertainty*.

Panel-to-panel uncertainty considers the variability across the panel population. A fuselage composed of several hundreds of panels. If all the manufactured panels are exactly the same and these panels work under exactly the same conditions and environment, then the panels will degrade identically. However, in practice, due to the intrinsic variability in crack initiation and propagation, the difference of material quality and the various working environment, uncertainty exists in the degradation process. For example, each panel can have its own initial crack size and material properties. In this study, the panel-to-panel uncertainty is modeled by assuming that the initial crack size a_0, the material properties parameters m and C follow a predefined distribution. Specifically, the initial crack size a_0 is assumed lognormally distributed while m and $log_{10}C$ are assumed to follow a multivariate normal distribution with a negative correlation coefficient, based on the research indicating that m and $log_{10}C$ are negatively linearly correlated.

Process uncertainty, which is of epistemic nature, refers to the uncertainty present in each individual panel during its crack propagation process. This uncertainty represents the lack of knowledge on the information of each single panel, such as the uncertain initial crack size a_0 and the uncertain material property parameters $\{m, C\}$. This information should be inferred by EKF that is detailed in CHAPTER 3. During the EKF process, for the i-th panel, the initial guess for $a_0^{(i)}$, $m^{(i)}$ and $C^{(i)}$ are given and fed to EKF as the start point. As the noisy measurements arrives sequentially, at each time step k, EKF incorporates the measurements and gives the optimal posterior estimates to $a_k^{(i)}$, $m^{(i)}$ and $C^{(i)}$ along with the estimation variance, which reduces gradually as time evolves due to more measurements being are available. According to the nomenclature in CHAPTER 3 where the EKF is detailed, the posterior estimation for crack size and the material property parameters at time k should be denoted as $\hat{a}_{k|k}$, $\hat{m}_{k|k}$ and $\hat{C}_{k|k}$. However, they are denoted as $\hat{a}_k$, $\hat{m}_k$ and $\hat{C}_k$ for simplicity reasons in CHAPTER 5 since it is not necessary to distinguish a priori and a posteriori estimate in this chapter.

5.3.2 a_{maint} - The threshold for triggering an unscheduled maintenance

SHM system is assumed to monitor the damage state of the fuselage. The frequency of damage

status evaluation, henceforth called maintenance assessment, is assumed here to approximately coincide with A-checks of the aircraft (100 flights); *i.e.* a small maintenance task carried out overnight at the airline hub hangars. It would make sense to carry out the SHM-based maintenance at the A-check if only the sensors themselves are embedded in the aircraft and the monitoring system is ground-based to reduce flying weight and monitoring system cost. First, notation used during the introduction of the maintenance strategies in the following sub-sections is defined in Table 5-1.

Table 5-1 Notation used in maintenance strategies

Notation	Description
k	The current number of flight cycles of the aircraft
i	The panel index in the fuselage of the aircraft
a_{cr}	The critical half crack size beyond which the panel fails
a_{maint}	The predefined threshold for triggering an unscheduled maintenance

The critical half crack size that will cause a panel failure is defined in Eq.(5-1) in which K_{IC} is a conservative estimate of the fracture toughness in loading Mode I and p_{cr} is also a conservative estimate of the pressure p given its distribution.

$$a_{cr} = \left(\frac{K_{IC}}{A \frac{p_{cr} r}{t} \sqrt{\pi}} \right)^2 \tag{5-1}$$

Since the damage assessment is done every 100 cycles, if a crack size exceeding a_{cr} is present in a panel in between two damage assessments, it may cause panel failure at once. Therefore, a different safety threshold a_{maint}, which is smaller than a_{cr}, is determined to ensure safety. a_{maint} is calculated to maintain a 1e-7 probability of failure of the aircraft between two damage assessments. (*i.e.*, when a crack size below a_{maint} is present on the aircraft, its probability to exceed the critical crack size a_{cr} in future 100 cycles is less than 1e-7). Hence it ensures the safety of the aircraft until next damage assessment. a_{maint} is the safety threshold, the event that any crack size exceeding a_{maint} triggers an unscheduled maintenance.

5.4 The first type of predictive maintenance

5.4.1 Repair policy

The following repair policy is used in the two variants of the first type of predictive

maintenance, PdM and PdM-skip. According to CHAPTER 4, when measurement data are available up to time k, the EKF is used to estimate the crack size and to identify the Paris' law parameters at time k. For each panel, based on the estimated crack size and material parameters, *i.e.* $\{\hat{a}_k, \hat{m}_k, \hat{C}_k\}$, the first-order perturbation method is used to predict the evolution of the crack size distribution in the future I cycles. The mean and standard deviation at $k+I$, μ_{k+I} and σ_{k+I}, are estimated by the first-order perturbation method. Based on the predicted distribution information, the 0.95 quantile, denoted by a_q, is calculated.

$$a_q = \Phi^{-1}(0.95 \mid \mu_{k+I}, \sigma_{k+I}) \tag{5-2}$$

in which Φ^{-1} is the inverse cumulative distribution function of the normal distribution with mean and standard deviation μ_{k+I} and σ_{k+I}, respectively. If $a_q > a_{maint}$, the panel is considered in danger and should be repaired. Otherwise, this panel is left unattended. This decision is called repair policy, denoted by d, which could be considered as a binary value.

$$d = \begin{cases} 0 & if\ a_q \le a_{maint} \\ 1 & if\ a_q > a_{maint} \end{cases} \tag{5-3}$$

The underlying meaning behind the repair policy is that if a crack size is present on the panel, the probability that this crack size grows greater than the threshold a_{maint} at the next schedule maintenance is less than 5%.

5.4.2 PdM

The objective of PdM is to decide the maintenance according to the actual condition of the aircraft rather than based on a fixed maintenance schedule. Figure 5-2 illustrates the flowchart of PdM. In this strategy, the maintenance assessment is implemented every 100 cycles. At each maintenance assessment, if the largest crack size among all the aircraft's panels exceeds a_{maint}, an unscheduled maintenance is asked immediately and this aircraft is sent to the maintenance hangar. The panel with the largest crack size triggering the unscheduled maintenance is called critical panel. At the time of unscheduled maintenance, besides repairing the critical panel, other panels may be also repaired according to the repair policy to prevent frequent unscheduled maintenance. More specifically, the crack size distribution of each panel in the next $I=I_b$ cycles is predicted using the EKF-FOP method and the 0.95 quantile is a_q is calculated. The panels whose a_q are greater than

a_{maint} are repaired.

In this strategy, the selection of forward prediction step I_b can be determined in several ways. It can be designed to maintain a desired frequency of unscheduled maintenances or it can be just chosen based on the existing experience with scheduled maintenance interval recommended by certification authorities or aircraft manufacturers. The parameter I_b can be optimized taking into consideration of outcomes like frequency of unscheduled maintenance and maintenance cost. Here I_b is set to be 4000 except when otherwise stated, that is set to the scheduled maintenance interval.

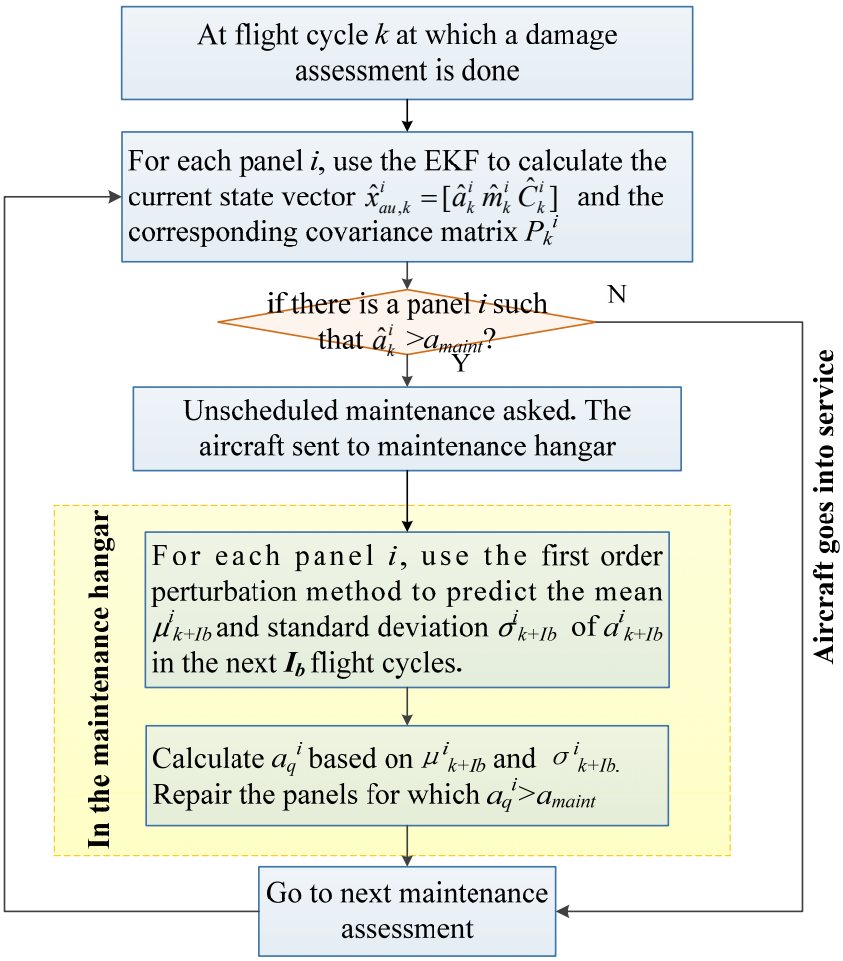

Figure 5-2 Flow chart of PdM strategy

5.4.3 PdM-skip

PdM-skip is a hybrid strategy that leverages the strength from both scheduled maintenance and PdM. The difference between PdM and PdM-skip is that PdM is designed completely independently without considering the time of scheduled maintenance, during which the engine and non-structural airframe maintenance are always performed. While PdM-skip ensures as much as possible that maintenance activities are carried out during the time of scheduled maintenance. The reason for this

is that having the structural airframe maintenance at the same time with engine and non-structural maintenance would tend to reduce cost.

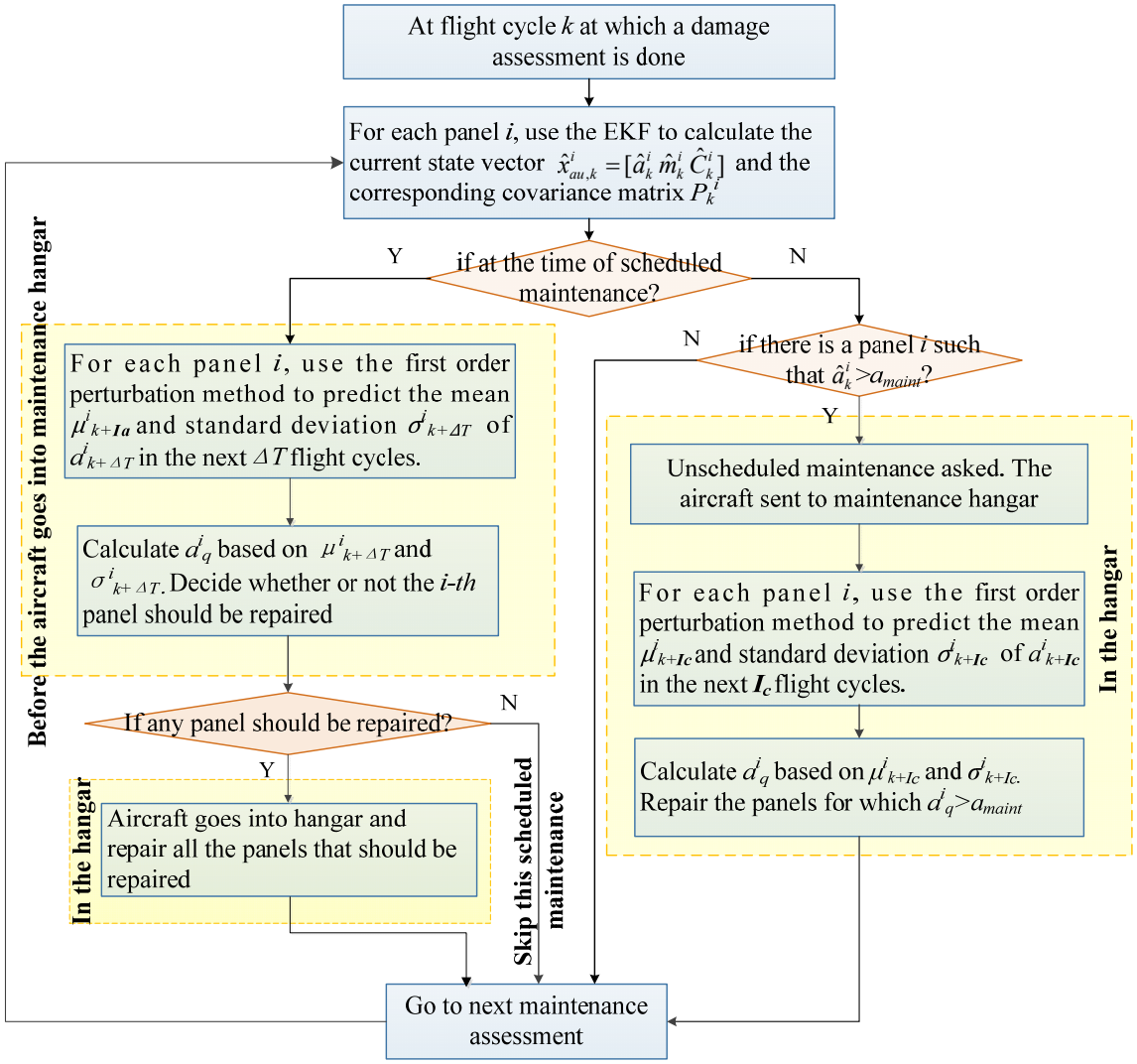

Figure 5-3 Flow chart of PdM-skip strategy

The PdM-skip strategy is described in Figure 5-3. The maintenance assessment is carried out every 100 cycles for unscheduled maintenance. At each scheduled maintenance stop, before the aircraft goes into the maintenance hangar, for each panel, the crack size distribution at next I_a flight cycles, *i.e.*, the distribution at the next scheduled maintenance, is predicted and the repair policy d is implemented to decide whether or not a panel needs to be repaired. If no panel needs to be repaired at this scheduled time, then PdM-skip recommends skipping this scheduled maintenance. If a crack that is missed at the time of scheduled maintenance exceeds a_{maint} between two consecutive scheduled maintenances, PdM-skip will recommend structural airframe maintenance to be performed immediately. This calls for unscheduled maintenance, which is costlier but guarantees safety. At an unscheduled maintenance stop, we predict the crack size distribution in the future I_c cycles for all

panels and the repair policy d is used to decide the ones that need to be repaired.

Here I_c is set to be the number of cycles from current to the scheduled maintenance after the next one. This is with a view to be able to skip the next scheduled maintenance and not have an unscheduled maintenance soon after. For example, if the scheduled maintenance is every 4,000 cycle and an unscheduled maintenance occurs at the 43,000th cycle, I_c will be set to 5,000 with a view to reach the 48,000 maintenance skipping the one at 44,000. Note that I_c is distinct from I_b that was used in the PdM strategy. In PdM strategy, I_b is constant while in PdM-skip, I_c is a variable depending on how many flight cycles are left from current to the after next scheduled maintenance time.

5.5 The second type of predictive maintenance

As elaborated in Section 5.2, unscheduled maintenance is more costly than scheduled maintenance due to less advance preparation. The objective of CDPM is to make a tradeoff between the cost ratio of scheduled cost over unscheduled cost and the probability of occurring unscheduled maintenance. In general, if the unscheduled cost is much more expensive than scheduled cost, then more panels are repaired at a scheduled maintenance to avoid a possible unscheduled maintenance. If the unscheduled cost is not much higher than the scheduled cost, then fewer panels are repaired. The decision of how many panels should be repaired depends on the relation between the cost ratio and the probability of triggering an unscheduled maintenance. The overall idea of CDPM policy is described below:

- The damage states of the fuselage panels are monitored continuously by the on-board SHM system and a damage assessment is performed every 100 flights (which approximately coincides with A-checks of the aircraft).
- At each assessment, as new arrived sensor data is available, the EKF is used to filter the measurement noise to provide the estimated crack size and parameters of crack growth model for each panel at current flight cycle.
- At the n-th scheduled maintenance stop, before the aircraft goes into the maintenance hangar, for each panel, the crack propagation trajectory from stop n to n+1 is predicted and the crack size distribution at next scheduled maintenance is obtained by using the first-order perturbation method. Taking into account this predicted information of each

panel, the cost optimal repair policy decides to skip or trigger the current *n*-th stop. If it is triggered, a group of specific panels is selected to be repaired based on the predicted information to minimize the expected maintenance cost. The algorithm of selecting a group of specific fuselage panels is called cost optimal repair policy and will be described in Section 5.5.3.

- During the interval of two consecutive scheduled maintenance stop, if there is a crack exceeding a safety threshold a_{maint} at damage assessment, an unscheduled maintenance is triggered immediately. The aircraft is sent to the hangar and this panel is repaired. The meaning and calculation of a_{maint} has been discussed in Section 5.3.2.

5.5.1 Reliability of individual panel

At the *n*-th scheduled maintenance stop (the cumulative cycles is T_n) the crack size distribution of each individual panel before the next scheduled stop is predicted. For the *i*-th panel, the probability of triggering an unscheduled maintenance before next scheduled maintenance stops is denoted by $P(us|a^i)$. It is approximated by Eq.(5-4), *i.e.*, the probability that the crack size of the *i*-th panel at next scheduled maintenance $a^i_{T_{n+1}}$ is greater than a_{maint}, given the information provided by EKF at current scheduled maintenance stop, more specifically, the estimated crack size and material property parameters, $[\hat{a}^i{}_{T_n}, \hat{m}^i{}_{T_n}, \hat{C}^i{}_{T_n}]$, and the covariance matrix $P^i{}_{T_n}$.

$$P(us \mid a^i) = \Pr(a^i_{T_{n+1}} > a_{maint} \mid [\hat{a}^i_{T_n}, \hat{m}^i_{T_n}, \hat{C}^i_{T_n}], P^i_{T_n}) \tag{5-4}$$

The evolution of the crack size distribution from T_n to T_{n+1} is predicted by the first-order perturbation method presented in CHAPTER 4. According to the first-order perturbation method, $a^i_{T_{n+1}}$ is normally distributed with parameters $\mu^{\,i}_{T_{n+1}}$ and $\sigma^{\,i}_{T_{n+1}}$, which are calculated analytically. Thus $P(us \mid a^i)$ is computed as

$$P(us \mid a^i) = \int_{a_{maint}}^{\infty} \Phi(a^i_{T_{n+1}} \mid \mu^i_{T_{n+1}}, \sigma^i_{T_{n+1}}) da^i_{T_{n+1}} \tag{5-5}$$

where Φ is the probability density function of normal distribution with the parameters $\mu^{\,i}_{T_{n+1}}$ and $\sigma^{\,i}_{T_{n+1}}$.

Note that the probability of triggering an unscheduled maintenance of a panel is not monotonic

with its current crack size, *i.e.*, it is not necessarily true that panel with a larger crack size is more likely to trigger an unscheduled maintenance. Due to the variability of crack growth rate among panels as well as the uncertainty presented in the crack propagation process, a larger crack size at *n*-th stop may have a lower probability of exceeding a_{maint} before next scheduled stop, compared with a smaller crack size.

5.5.2 Cost model

Some concepts as well as their notations are given firstly before the cost structure is introduced.

d_n^j - The repair decision for the *j*-th panel at the *n*-th scheduled maintenance stop. It is a binary value defined as

$$d_n^j = \begin{cases} 1 & \text{if panel j is repaired} \\ 0 & \text{if panel j is not repaired} \end{cases} \tag{5-6}$$

$\mathbf{d}_n$ - the decision vector such that $\mathbf{d}_n=[d_n^1, d_n^2, ...d_n^N]$.

c_0 - The scheduled set up cost, which is a fixed cost that occurs every time the scheduled maintenance is triggered. The set up cost is assigned only once even if more than one panel is replaced.

c_0^{un} - The unscheduled set up cost, which is a fixed cost that occurs when unscheduled maintenance is triggered. Due to less advance notification, $c_0^{un} > c_0$.

τ^d - A variable used to indicate the binary nature of scheduled maintenance. τ^d=1 means that the scheduled maintenance is triggered and the set up cost is incurred while τ^d =0 means this scheduled maintenance is skipped thus no set up cost.

c_s – The fixed cost of repairing one panel.

c_{us} - The repair cost at unscheduled maintenance, also called unscheduled repair cost, which is composed of two items, the unscheduled set up cost c^0_{un} plus the per panel repair cost c_s.

The expected maintenance cost at the *n*-th scheduled maintenance stop, denoted by $C(\mathbf{d}_n)$, is modeled as the function of the repair decision of each panel, as given in Eq.(5-7). Here we assume that the probability for a panel to have more than one unscheduled repair is negligible.

$$C(\mathbf{d}_n) = c_0\tau^d + c_s(\sum_{i=1}^{N} d_n^j) + c_{us}(\sum_{i=1}^{N}(1-d_n^j)P(us \mid a^j)) \tag{5-7}$$

The first two terms represent the scheduled repair cost. The last term represents the unscheduled repair cost.

5.5.3 Repair policy for the second type of predictive maintenance

The objective is to find the optimal grouping of several panels to be repaired to minimize the cost when the aircraft is at *n*-th scheduled maintenance stop. The algorithm is under the following assumptions:

- The probability for a panel to have more than one unscheduled repair is negligible.
- The probability to have more than one unscheduled repair at the same cycle is negligible.

The second assumption means that having more than one panel repaired during unscheduled maintenance do not reduce the average cost of each panel. Note that this is different from the previous strategies of section 5.4 where several panels were repaired at unscheduled maintenance in order to avoid too frequent unscheduled maintenance. A different assumption was made here in order to allow analytical tractability of the optimal policy.

At the *n*-th scheduled maintenance, for each panel, the probability of triggering an unscheduled maintenance between stop *n* and *n*+1 is calculated according to Section 5.3.2. Sort and arrange them in descending order such that

$$P(us \mid a^1) > P(us \mid a^2) > ...P(us \mid a^{i-1}) > P(us \mid a^i) > P(us \mid a^{i+1}) ... > P(us \mid a^N) \tag{5-8}$$

Eq.(5-8) implies that the panel that is most likely to trigger an unscheduled maintenance is arranged first. The motivation is that it is more concerned about the panels with higher probability of having unscheduled repair since unscheduled maintenance is more costly. In the following parts, the panel index refers to the order in Eq.(5-8). Two sets *I* and *J* are defined.

$$I = \{1 \le i \le N \mid c_s \le c_{us} P(us \mid a^i)\} \tag{5-9}$$

$$J = \{1 \le l \le N \mid c_0 + lc_s \le c_{us} \sum_{j=1}^{l} (P(us \mid a^j))\} \tag{5-10}$$

For zero set up cost (*i.e.*, c_0=0), the set *I* contains the element *i* such that repairing the *i*-th panel at current scheduled maintenance costs less than repairing it at an unscheduled maintenance stop. For any value of the set up cost, set *J* includes the element *j* such that repairing all these *j* panels at scheduled maintenance cost less than at unscheduled maintenance. B_I and b_J are defined as the

maximal value and the minimal value of set *I* and *J*, respectively. Note that B_I and b_J are scalars.

$$B_I = \max\{1 \le i \le N \mid c_s \le c_{us} P(us \mid a^i)\} \tag{5-11}$$

$$b_J = \min\{1 \le l \le N \mid c_0 + lc_s \le c_{us} \sum_{j=1}^{l} P(us \mid a^j)\} \tag{5-12}$$

A simple example is given below to illustrate the meaning of B_I and b_J intuitively. In the following example, B_I=4 and b_J=3.

$c_s \le c_{us} P(us \mid a^1)$
$c_s \le c_{us} P(us \mid a^2)$
$c_s \le c_{us} P(us \mid a^3)$
$c_s \le c_{us} P(us \mid a^4)$
$c_s > c_{us} P(us \mid a^5)$

and

$c_0 + c_s > c_{us}(P(us \mid a^1))$
$c_0 + 2c_s > c_{us}(P(us \mid a^1) + P(us \mid a^2))$
$c_0 + 3c_s \le c_{us}(P(us \mid a^1) + P(us \mid a^2)) + P(us \mid a^3))$

From Eq.(5-9) to Eq.(5-12), the following properties can be deduced.

$$1 \le b_J < B_I \le N \tag{5-13}$$

$$c_s \le c_{us} P(us \mid a^j), \text{ for } j = 1,2,...B_I \tag{5-14}$$

$$c_s > c_{us} P(us \mid a^j), \text{ for } j = B_{I+1}, B_{I+2}, ...N \tag{5-15}$$

$$c_0 + lc_s > c_{us} \sum_{j=1}^{l} P(us \mid a^j), \text{ for } j = 1,2,...b_{J-1} \tag{5-16}$$

$$c_0 + b_J c_s \le c_{us} \sum_{j=1}^{B_J} P(us \mid a^j) \tag{5-17}$$

The proof for Eq.(5-13) is given in Appendix B and the Eq.(5-14) to Eq.(5-16) can be easily derived from the definitions given in Eq.(5-9) to Eq.(5-12). Now we discuss the cost optimal policy at the *n*-th scheduled maintenance stop.

If set *I* is empty (*i.e.*, $I = \varnothing$) and the set up cost is zero (*i.e.*, c_0=0), for any panel the expected unscheduled repair cost is smaller than the scheduled one. In this case, the optimal repair policy is not to repair any panel at current scheduled maintenance stop, *i.e.*, $d_n^{j*}(a^j)$ =0, for *j*=1,2,…*N*. Note that d_n^{j*} denotes the optimal repair decision for the *j*-th panel at the *n*-th scheduled maintenance

stop.

If the set I is not empty (*i.e.*, $I \neq \varnothing$) and the set up cost is zero (*i.e.*, c_0=0), from Eq.(5-14) and Eq.(5-15), it can be known that for any panel j that $j<B_I$ the expected unscheduled repair cost is larger than the scheduled one, while for any panel j that $j>B_I$, the expected unscheduled repair cost is smaller than the scheduled one. In the case $I \neq \varnothing$, the set J could be either empty or non-empty.

If J is empty (*i.e.*, $J = \varnothing$), it means that repairing the N panels at scheduled maintenance stop costs more than at unscheduled maintenance. Then, for $J = \varnothing$, the optimal maintenance policy is not to repair any panel at current scheduled maintenance stop, *i.e.*, $d_n^{j*}(a^j)$ =0, for j=1,2,…N. Note that $I = \varnothing$ implies $J = \varnothing$ but we can have $J = \varnothing$ and $I \neq \varnothing$.

If J is not empty (*i.e.*, $J \neq \varnothing$), from Eq.(5-16) and Eq.(5-17), it can be known that for any panel j that $j<b_J$, repairing the j-first panels at scheduled maintenance stop costs more than at unscheduled maintenance, and for $j=b_J$, repairing the j-first panels at scheduled maintenance stop cost less than at unscheduled maintenance. As for $j>b_J$, repairing the j-first panels at scheduled maintenance stop can be either better or worse. For example, we can have:

$$c_0 + c_s > c_{us}(P(us \mid a^1))$$

$$c_0 + 2c_s < c_{us}(P(us \mid a^1) + P(us \mid a^2))$$

$$c_0 + 3c_s > c_{us}(P(us \mid a^1) + P(us \mid a^2) + P(us \mid a^3)) \text{ or}$$

$$c_0 + 3c_s < c_{us}(P(us \mid a^1) + P(us \mid a^2) + P(us \mid a^3))$$

From Eq.(5-13), it can be known that the range [1,N] are divided into three intervals by B_I and b_J, which are [1, b_J], [b_J+1, B_I] and [B_I+1, N]. To determine the optimal policy, it is clear that the b_J -first panels have to be repaired at the current scheduled maintenance (see Eq.(5-17)). In addition, since the expected unscheduled maintenance cost of panels in the interval [b_J+1, B_I] are larger than scheduled maintenance cost (see Eq.(5-14)), they should also be repaired at current scheduled maintenance stop. Finally, the optimal repair policy at n-th scheduled maintenance can be summarized as follows:

If $J = \varnothing$

$$d_n^{j^*} = 0,\ for\ j = 1,2,...N$$

Else

$$d_n^{j^*} = \begin{cases} 1 & for\ j = 1,2,...B_I \\ 0 & for\ j = B_I + 1,...N \end{cases} \tag{5-18}$$

The above decision implies that when J is empty (*i.e.*, $J = \varnothing$), the optimal decision is not to repair any panel at the n-th scheduled maintenance stop. The expected cost under this situation is

$$C(\mathbf{d}_n^*) = c_{us}(\sum_{j=1}^{N} P(us \mid a^j)) \tag{5-19}$$

When J is not empty (*i.e.*, $J \neq \varnothing$), the optimal decision is to repair the first B_I panels and leave unattended the remaining ones. Accordingly, the cost in this case is

$$C(\mathbf{d}_n^*) = c_0 + c_s B_I + c_{us}(\sum_{j=B_I+1}^{N} P(us \mid a^j)) \tag{5-20}$$

Then the optimized total maintenance cost during the aircraft lifetime, denoted as $C(\mathbf{d}^*)$ is the sum of the cost at each scheduled maintenance $C(\mathbf{d}_n^*)$:

$$C(\mathbf{d}^*) = \sum_n C(\mathbf{d}_n^*) \tag{5-21}$$

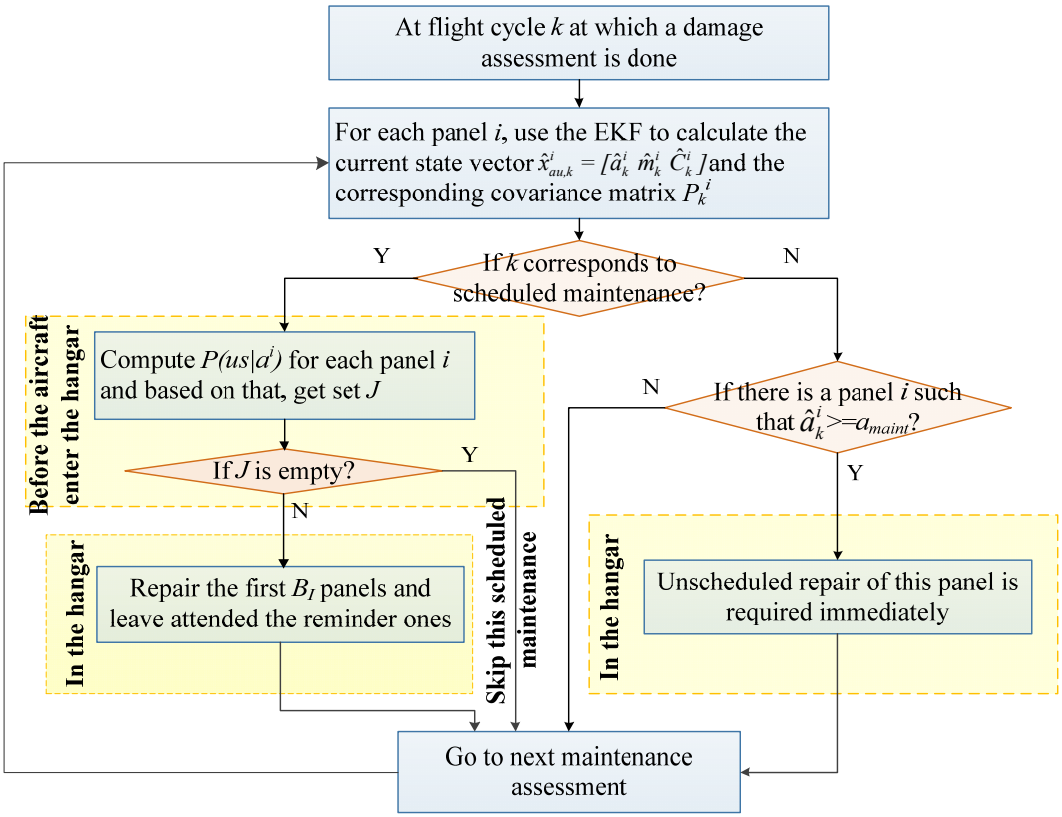

Figure 5-4 Flow chart of CDPM

The rigorous mathematical proof regarding $C(\mathbf{d}_n^*)<C(\mathbf{d}_n)$, *i.e.*, why $\mathbf{d}_n^*$ is the optimal decision is given in Appendix B. The cost optimal policy is integrated into the predictive policy, whose flowchart is illustrated in Figure 5-4. The above repair decision is made at each scheduled maintenance stop until the end of the aircraft's life. Then the total maintenance cost during aircraft lifetime $C(\mathbf{d}^*)$ can be calculated.

5.6 Numerical study

5.6.1 The strategies involved

In this section, the performance of predictive maintenance is evaluated through comparison with threshold-based maintenance. The classification tree of maintenance strategy is summarized in Figure 5-5. All the specific strategies involved in this section are those in the third level of the tree.

The first type of threshold-based maintenance refers to traditional scheduled maintenance, in which the schedule is shown in Figure 5-1. At each scheduled maintenance, the aircraft is taken into a hangar and the inspection of all panels is done using techniques like NDI, GVI or DVI. Cracks detected with a size greater than a threshold a_{rep} are repaired. The threshold is determined to guarantee a desirable level of probability of failure between two scheduled maintenances.

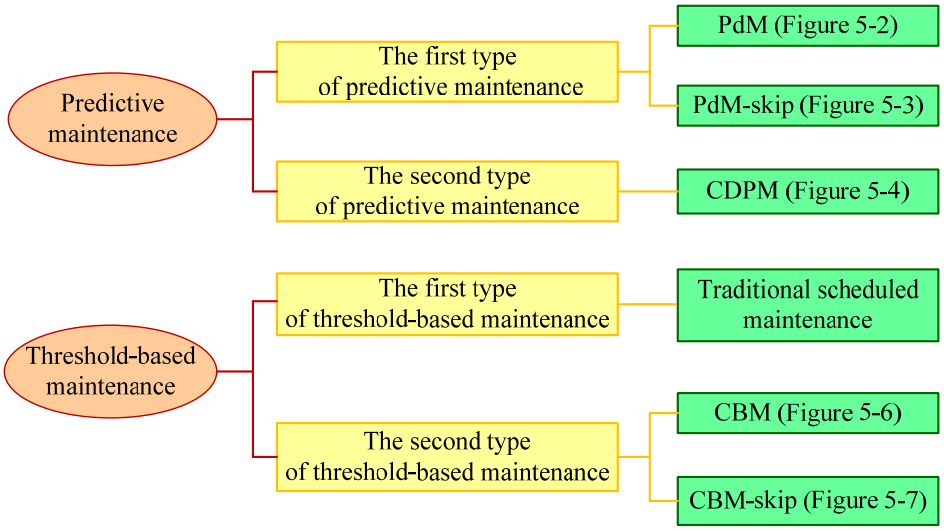

Figure 5-5 The maintenance strategies involved in the numerical study

The second type of threshold-based maintenance includes two variants, named CBM and CBM-skip. These maintenance strategies have been proposed by Pattabhiraman et al [142]. In CBM, the SHM is assumed to be used to track the damage and the damage assessment is performed

every 100 flight cycles (the same as in the predictive maintenance). After damage assessment via on-board sensors, maintenance is requested if the maximum damage size in an aircraft exceeds the threshold a_{maint}. Once maintenance is asked, all panels on the fuselage are inspected for damage in the hangar using the on-board SHM equipment, and panels with a damage size greater than a threshold a_{rep} are repaired. a_{rep} is set to prevent frequent maintenance for that aircraft. The flow chart of CBM is shown in Figure 5-6.

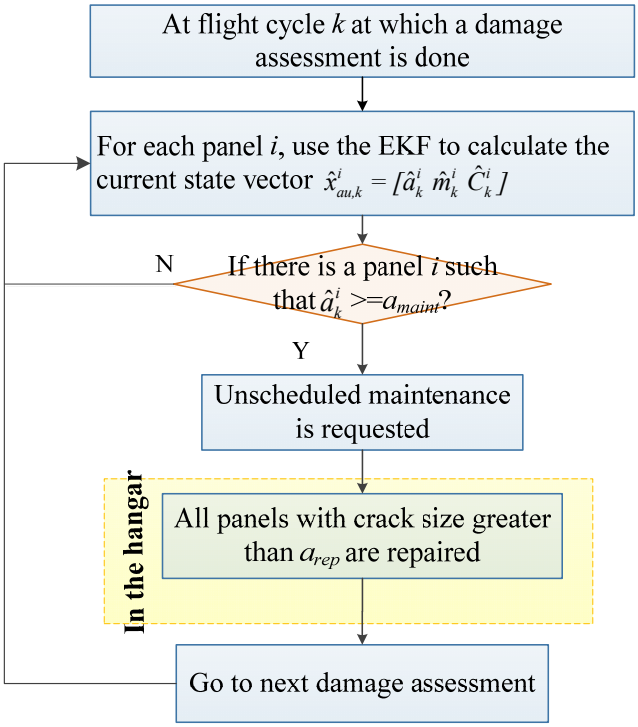

Figure 5-6 Flow chart of CBM

In CBM-skip, the damage assessment is performed every 100 flights. The aim is the same as PdM-skip to skip some unnecessary early scheduled maintenance while guarantee the safety by triggering unscheduled maintenance. Specifically, at each scheduled maintenance, if there is no crack size exceeding a threshold $a_{th\text{-}skip}$, then the current scheduled maintenance is skipped. Between two consecutive scheduled maintenance stops, if a crack grows beyond a_{maint}, the unscheduled maintenance is triggered and all panels whose crack size is greater than a_{rep} are repaired. The flowchart of threshold-based maintenance is given in Figure 5-7. Three design parameters characterize the CBM-skip. First a_{maint} ensures the safety. Second $a_{th\text{-}skip}$ is calculated such that the probability of one crack exceeding a_{maint} before next scheduled maintenance is less than 5%. Finally, the repair threshold a_{rep} is set the same value as in traditional maintenance. All the thresholds involved and their values are given in Table 5-4.

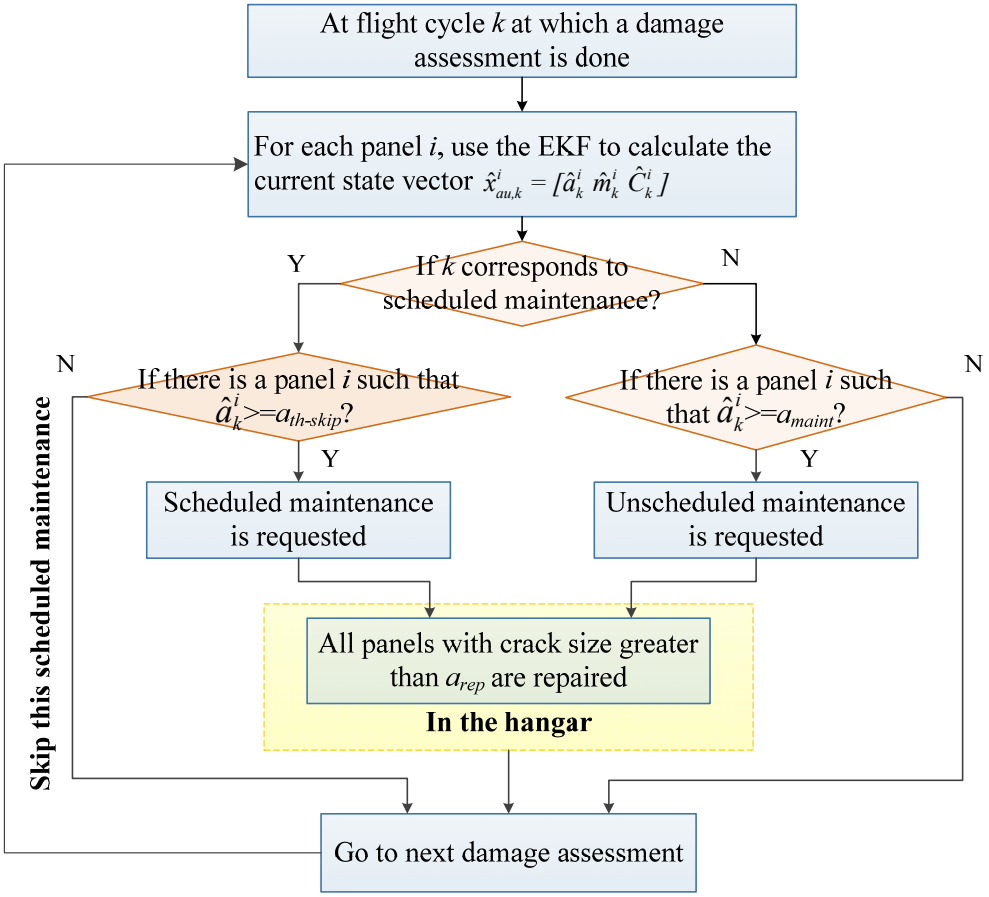

Figure 5-7 Flow chart of CBM-skip

5.6.2 Input data

5.6.2.1 The geometry values

The values of the geometry parameters defining the fuselage (*e.g.* fuselage radius, panel thickness) used here are related to short-range commercial aircrafts. These values are time-invariant. Recall that we define a correction factor for stress intensity factor A, which accounts for the fact that the fuselage is modeled as a hollow cylinder (without stringers and stiffeners). The numerical values for the geometry parameters have been chosen from [142] and are reported in Table 5-2.

Table 5-2 Aircraft geometry parameters

Description	Notation	Value
Fuselage radius	r	1.95 m
Panel thickness	t	2e-3 m
Correction factor	A	1.25

5.6.2.2 The panel-to-panel uncertainty values

To show the benefits of predictive maintenance over threshold-based maintenance, two degrees of spread on $\{m, C\}$ and the pressure p are considered by keeping the mean values constant while changing the coefficient of variation (COV). Since it is difficult to estimate the exact amount of variability on the material properties in reality, a small and a large uncertainty cases are considered. These distributions can be later updated with in service data to obtain a more accurate estimation of the variability. The assumed uncertainties on a_0, $\{m, C\}$ and p are reported in Table 5-3. 50000 samples of $\{m, C\}$ representing the material parameter for each panel of the 100 aircrafts are illustrated in Figure 5-8 according to the multivariate distribution given in Table 5-3.

Table 5-3 Panel-to-panel-uncertainties on a_0,$\{m, C\}$ and p

Description	**Notation**	**Type**	**Value**	
Initial crack size (meter)	a_0	Lognormal	LnN(0.3e-3,0.08e-3)	
Paris' law parameters	$\{m, C\}$	Multivariate	N (μ_m, σ_m , μ_C , σ_C , ρ)	
Mean of m	μ_m	-	3.6	
Mean of C	μ_C	-	Log_{10}(2e-10)	
C.C.[a] of m and C	ρ	-	-0.8	
			Small Uncertainty	Large uncertainty
Standard deviation of m	σ_m	-	0.5%COV	3%COV
Standard deviation of C	σ_C	-	0.5%COV	3%COV
Pressure	p	Normal	N(0.06, 0.5%COV)	N(0.06, 3%COV)

C.C[a]. means the correlation coefficient.

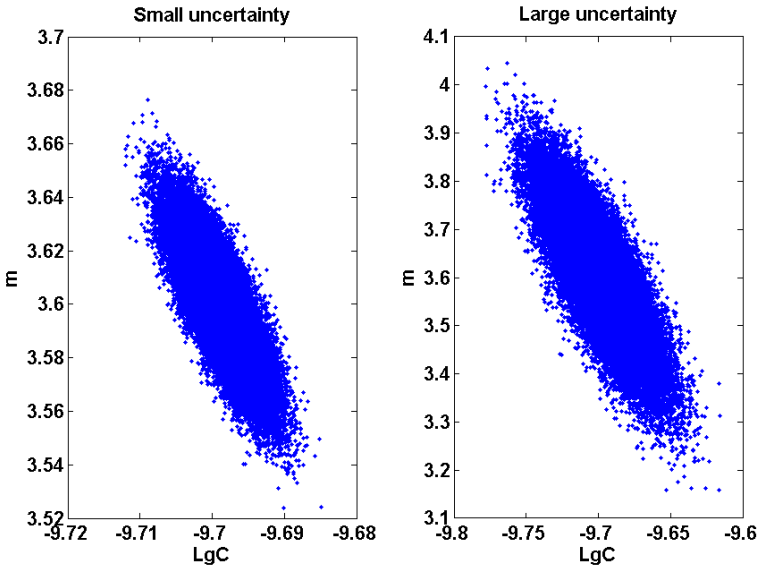

Figure 5-8 Illustration of the population of $\{m, C\}$ in large uncertainties case

5.6.2.3 Thresholds

The critical crack size a_{cr}, the thresholds a_{maint} , a_{rep} and $a_{th\text{-}skip}$ that are used in different

maintenance strategies are calculated based on the distribution of $\{m, C\}$ and p. Their values in small and large uncertainty case are given in Table 5-4.

Table 5-4 Numerical values of thresholds

Notation	Description	Values in small uncertainty case	Values in large uncertainty case
a_{cr}	The critical crack size cause panel fail (meter)	59.6e-3	75.7e-3
a_{maint}	The safety threshold for trigging unscheduled maintenance (meter)	47.4e-3	67.3e-3
a_{rep}	The repair threshold (meter)	4.3e-3	10.4e-3
$a_{th\text{-}skip}$	The skip threshold used in CBM-skip (meter)	4.3e-3	10.4e-3

5.6.2.4 The initial values used in EKF process

The panel-to-panel uncertainty discussed above represents the variability among panels population. This uncertainty is the intrinsic variability among panel's properties and thus irreducible. The initial crack size and the material property parameters of each panel, *i.e.*, $a_0^{(i)}$, $m^{(i)}$, $C^{(i)}$, $i=1,2,...N$, are sampled from the given distribution in Table 5-3. These values are regarded as the "true unknown draws" that need to be estimated by the EKF from noisy measurements.

For each individual panel, there is an EKF process. The EKF algorithm needs to be initialized. In particular, the starting point of the estimation algorithm has to be provided. In this case, the starting point of the i-th panel is drawn randomly from uniform distribution centered at the "true" value of $a_0^{(i)}$, $m^{(i)}$, $C^{(i)}$, $i=1,2,...N$ respectively and having a range of 50% around these true values. Once this initial guess has been provided, EKF gives an estimation $\hat{a}_k$, $\hat{m}_k$ and $\hat{C}_k$ at each cycle k. As for the initial error covariance matrix P_0, it is chosen depending on how much confidence one has to the initial estimates. The numerical values used in EKF process are given in Table 5-5.

Table 5-5 Numerical values used in the EKF

Description	Notation	Type	Value
Initial guess for crack size	$\hat{a}_0$	Uniform	$U[a_0\text{-}25\%\ a_0,\ a_0\text{+}25\%\ a_0]$
Initial guess for m	$\hat{m}_0$	Uniform	$U[m_0\text{-}25\%\ m_0,\ m_0\text{+}25\%\ m_0]$
Initial guess for C	$\hat{C}_0$	Uniform	$U[C_0\text{-}25\%\ C_0,\ C_0\text{+}25\%\ C_0]$
Initial error covariance matrix	P_0	-	$diag\ ((1e\text{-}4)^2,(1e\text{-}2)^2,(1e\text{-}10)^2)$
Measurement noise variance	Ra	-	10%

5.6.2.5 The cost-related quantities

The aircraft maintenance cost is composed of engine maintenance cost and airframe maintenance cost. The airframe maintenance cost is further divided into structural airframe maintenance and non-structural airframe maintenance. Some sophisticated aircraft cost models have been developed [150, 151]. Note that the engine and non-structural maintenance are always performed at the time of scheduled maintenance interval. A simple cost model regarding the engine maintenance cost is detailed in APPENDIX A, which is based on existing literature ([152]). Based on the empirical expressions and aircraft parameters in APPENDIX A (Table A-1), the engine maintenance cost is \$258/flight. The aircraft makes 60000 flights during its lifetime and undergoes 10 scheduled maintenances. Hence, the cost of one scheduled engine maintenance (EM) is \$1.55 million. The cost for non-structural airframe maintenance (NSM) is assumed to be \$5.04 million. The cost for the structural airframe part performed by traditional NDI, GVI or DVI technologies at the time of scheduled maintenance consists of two parts, the set up cost c_0^t and the repair cost, which is the cost of repairing one panel, denoted by c_s, multiplied by the number of repaired panel. c_0^t is assumed \$1.44 million and c_s is \$0.25 million.

For the SHM-based maintenance policies (PdM, PdM-skip, CDPM, CBM and CBM-skip), the scheduled set up cost c_0 is only a fraction of c_0^t due to the use of SHM system, leading to less labor intensive inspection compared to traditional inspection through DVI and NDI. This fraction is denoted as k_{SHM}. In contrast, the unscheduled set up cost c_0^{un} is higher than c_0^t due to less advance notice. A factor k_{un} is set to denote the higher set up cost incurred by unscheduled maintenance. Note that the per panel repair cost c_s is the same no matter of scheduled or unscheduled maintenance. It is the difference in set up cost that leads unscheduled maintenance to be costlier than scheduled maintenance.

At the n-th scheduled maintenance, the repair costs for different maintenance policies are given in the 8th and 9th lines of Table 5-6. The unscheduled repair cost is given in the 10th and 11th lines. The symbol "N_p" in the last column of lines 8-10 denotes the number of panels repaired at that corresponding maintenance stop. Note that the unscheduled repair cost of CDPM c_{us} is composed of the unscheduled set up cost and the cost of repairing one panel since there is only one panel repaired once unscheduled maintenance is triggered. For each maintenance policy, once the scheduled repair cost and the unscheduled repair cost are determined, the lifetime cost can be obtained by adding up these scheduled/unscheduled cost during the lifetime. The lifetime repair

cost for each maintenance policy is given in Table 5-7.

Note that for traditional maintenance, CBM, CBM-skip, PdM, PdM-skip, all cost-related quantities have no effect on the repair decision, while in CDPM, the repair decision depends on the cost ratio c_s/c_{us}, thus relating to $k_{un.}$ In the numerical experiments, c_0^t and c_s are constants and are set to be 1.44 and 0.25 (Million $) respectively. k_{SHM} does not affect the repair decision, so it is assumed to be a constant value of 0.9 for simplicity. Different scenarios under varying k_{un} are studied. A series of discrete value, 0.9, 3, 5, 10, are chosen for k_{un}. k_{un}=0.9 indicates the unscheduled set up cost is as cheap as scheduled CPDM set up cost. This is an extreme case.

Table 5-6 Cost-related quantities description

Notation	In which maintenance policy it involves?	Description	How to calculate?
c_0^t	Traditional scheduled maintenance	Set up cost	1.44 M$
k_{SHM}	PdM, PdM-skip, CDPM, CBM, CBM-skip	Coefficient	0.9
k_{un}	PdM, PdM-skip, CDPM, CBM, CBM-skip	Coefficient	0.9, 3, 5, 10
c_s	All maintenance policies	Per panel repair cost	0.25 M$
c_0	PdM-skip, CDPM, CBM-skip	Scheduled set up cost	$c_0 = k_{SHM} c_0^t$
c_0^{un}	PdM, PdM-skip, CDPM, CBM, CBM-skip	Unscheduled set up cost	$c_0^{un} = k_{un} c_0^t$
C_n^s	Traditional scheduled maintenance	Scheduled repair cost at *n*-th scheduled maintenance	$C_n^s = c_0^t + c_s N_p$
C_n^{thr}	PdM-skip, CBM-skip	Scheduled repair cost at *n*-th scheduled maintenance	$C_n^{thr} = k_{SHM} c_0^t + c_s N_p$
c_{us}^{thr}	PdM, PdM-skip, CBM, CBM-skip	Unscheduled repair cost	$c_{us}^{thr} = k_{un} c_0^t + c_s N_p$
c_{us}	CDPM	Unscheduled repair cost	$c_{us} = k_{un} c_0^t + c_s$

Table 5-7 Lifetime repair cost for each maintenance policy

Maintenance Policy	Lifetime repaire cost	Remarks
Traditional scheduled maintenance	$C_1 = N_1(c_0^t + c_s N_p)$	N_1 is the number of scheduled maintenance
PdM-skip, CBM-skip	$C_2 = N_2(k_{SHM} c_0^t + c_s N_p) + N_3(k_{un} c_0^t + c_s N_p)$	N_2 is the number of maintenance that occurs at scheduled maintenance, N_3 is the number of unscheduled maintenance
PdM, CBM	$C_3 = N_4(k_{un} c_0^t + c_s N_p)$	N_4 is the number of unscheduled maintenance

5.6.3 Results and discussion

5.6.3.1 Results of the first type of predictive maintenance

The first type of predictive maintenance strategies (PdM and PdM-skip) are compared with the threshold-based maintenance (the scheduled maintenance, CBM and CBM-skip). The process of no

maintenance intervention is also simulated, which enables to show that if no maintenance interferes, how many panels actually fail. Through comparing the "actually failed" panels in the *no maintenance intervention process* with the number of repaired panels in other maintenance strategies, it is able to know the "unnecessary repairs" of different strategies, thereby, to further quantify the conservativeness level of different maintenance strategies. The comparison results under large panel-to-panel uncertainty (large uncertainty) are reported in Table 5-8.

In Table 5-8, the 2nd row gives the number of total failures (in case of no maintenance intervention) or repaired panels over the entire fleet (*i.e.*, 100×500 panels). The 3rd row presents the number of "unnecessary repaired" panels, *i.e.*, panels that would not fail during the whole life but are nevertheless (unnecessarily) repaired according to the maintenance strategy. The 4th-6th row give the minimal, the maximal and the average number of maintenance stops among the 100 aircraft, respectively. The 7th row shows the average number of unscheduled maintenance stops. Note that for CBM and PdM, all maintenances are unscheduled. The 8th-10th rows give the minimal, the maximal and the average number of repaired panels among the 100 aircraft.

Table 5-8 Comparison of different processes

	No maintenance	Scheduled	CBM	CBM-skip	PdM	PdM-skip
Panels fail/repaired over the entire fleet	692 failures	1403 repaired	1454 repaired	1403 repaired	738 repaired	762 repaired
Unnecessary repairs	-	711	762	711	46	70
Minimal number of maintenance stop	-	10	1	2	1	1
Maximal number maintenance stop	-	10	3	6	4	5
Avg. number. of maintenance stop	-	10	1.9	3.9	2.4	3.1
Avg. number of unscheduled maintenance stop	-	0	1.9	0	2.4	0
Minimal number of repaired panels	-	5	2	5	2	2
Maximal number of repaired panels	-	21	29	21	16	16
Avg. number of repaired panels	-	14.0	14.5	14.0	7.3	7.6

It can be seen that if one lets cracks grow continuously without maintenance intervention, 692 panels over the whole fleet eventually fail. All these 692 panels are repaired in each maintenance strategy prior to their failure. In other words, all maintenance strategy can ensure safety. Each maintenance strategy has a different extent of "unnecessary repair". The number of unscheduled maintenance is 0 in both CBM-skip and PdM-skip, which indicates that all maintenances occur at the times of scheduled maintenance and no unscheduled maintenance is conducted. This does not

mean that there will never be any but it is a rare event that we do not capture with our fleet size.

Comparing the threshold-based maintenance strategies, it shows that CBM-skip and the scheduled maintenance has the same number of repaired panels because they share the same repair threshold a_{rep} and the synchronized maintenance schedule (see Figure 5-1) since no unscheduled maintenance is triggered in CBM-skip. CBM repairs more panels than CBM-skip even if they have the same repair threshold. The reason lies in the last 4000 cycles. During the last 4000 cycles, for CBM, the panel whose crack size is larger than a_{maint} triggers unscheduled maintenance and at the same time panels whose crack size are larger than a_{rep} are repaired. While in CBM-skip, there is a scheduled maintenance at 56000^{th} cycle and panels with crack size larger than a_{rep} are repaired. There is no chance (or very low probability) for panels to grow exceeding a_{maint} in the last 4000 cycles, thus no chance to trigger unscheduled maintenance. Therefore, in CBM-skip process, even if there are panels with crack size between a_{rep} and a_{maint} in the last 4000 cycles, they do not have the chance to be repaired.

One may notice that among threshold-based maintenance strategies, CBM and CBM-skip have the equal or larger number of repaired panels than the scheduled maintenance. This is due to the usage of a_{rep}, which is a conservative value calculated for scheduled maintenance to ensure 1e-7 reliability in 4000 cycles for the entire fleet, as the repair threshold for CBM and CBM-skip. There are two different contributions to the *conservativeness*, the inter-aircraft variability and intra-aircraft variability. The first one refers to that the worst aircraft in the fleet may have a large crack size much sooner than the average, while the second one is related to different crack sizes and crack growth rates in one aircraft. Indeed, the conservativeness could be further reduced by optimizing the repair threshold used in CBM and CBM-skip. However, this is not trivial due to various factors affecting the conservativeness level and it needs to be optimized in terms of a conservative value of the crack size and crack growth rate. Another option is to perform prognostics, which is the motivation of proposing PdM and PdM-skip. Both of these two strategies address the two contributions to the overall conservativeness, and thus decrease the number of maintenance stops and repaired panels.

Note also that the comparison of the unnecessary repairs allows comparing the conservativeness level of the various strategies. Scheduled maintenance is clearly the most conservative since it needs to cover a very conservative crack size and crack growth rate both over

the fleet and within and individual aircraft. In order to decrease the cost it makes sense to decrease the conservativeness level and the various maintenance strategies reduce the conservativeness to different extent. CBM and CBM-skip address the part that stems from inter-aircraft variability as well as the intra-aircraft variability related to different crack sizes. But it does not cover intra-aircraft variability related to different crack growth rates. Note that in order to quantify the conservativeness gains from CBM over the scheduled maintenance, we need to have a comparable number of maintenance stops, otherwise a higher number of maintenance stops would be traded off for a lower number of repaired panels. Accordingly, we set two stops for scheduled maintenance (closer to the number of stops in CBM 1.9) with a 20000 cycles interval, *i.e.*, the first maintenance stop is at 20000^{th} and the second is at 40000^{th} cycle. In this case, the repair threshold decreases to a very small value 0.8e-3mm to maintain a reliability of 1e-7 in 20000 cycles for the entire fleet and the number of repaired panels goes up to 8990. When comparing the conservativeness level of PdM and PdM-skip we can see that only very few panels are unnecessarily repaired (reduction by more than an order of magnitude over CBM and CBM-skip) which shows the interest of using prognostics in the maintenance strategy.

Figure 5-9 to Figure 5-11 illustrate the statistical characters of failure/repaired time of panels over the entire fleet, *i.e.*, 100×500 panels. The histogram of the failure time in the case of no maintenance intervention is given in Figure 5-9. The numbers in the x-axis is the center of the bin and the bin width is 2000 cycles. For example, the first bin means that the number of panels whose failure time is within the range of [36000, 38000] cycles is two. It can be seen that most failures occur at the second half of the lifetime and the number of failed panels gradually increases toward the end of life (EOL).

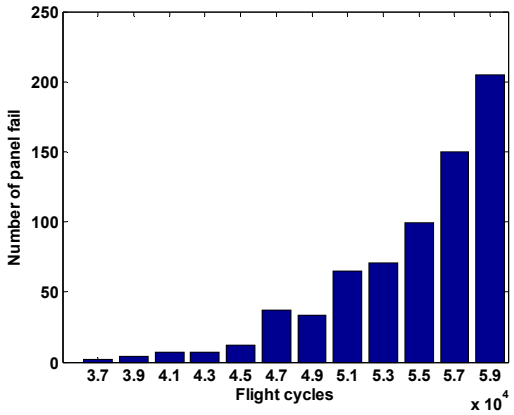

Figure 5-9 Number of panels fails within the time range of each bin in the case of no maintenance process

Figure 5-10 compares the scheduled maintenance, CBM-skip and PdM-skip strategies in terms of number of repaired panels at each scheduled stop (reminder there are 10 scheduled stops, see Figure 5-1). Scheduled maintenance and CBM-skip show the same pattern. The reason has been explained above. PdM-skip reduces by 90% the "unnecessary repair" over CBM-skip since PdM-skip decreases the *conservative level* by doing prognostics for each panel. The panels that are repaired in CBM-skip may not be necessary to repaired in PdM-skip due to their slow growth rates, thus not threatening the safety.

Figure 5-11 compares CBM and PdM in terms of the number of repaired panels within the time range of each bin. The numbers in x-axis is the center of the bin and the bin width is 2000 cycles. For example, the first bin means that there are two panels whose failure cycle is within the range of [36000, 38000] cycles for both CBM and PdM. It shows that PdM significantly reduces the number of repaired panels compared to CBM at the later period of the aircraft lifetime. That is because CBM repairs many cracks slightly larger than a_{rep} near the EOL but actually these panels do not affect safety. In contrast, PdM reduces these "unnecessary repair" by considering the reliability at EOL.

The above three figures show the statistical behavior of the failure/repaired time of panels over the entire fleet. In order to give more insight into what happens in different processes, we take the simulation results of aircraft No.30 as an example and illustrate in Figure 5-12 which panels are repaired at which flight cycle in each process. The symbol "#" represents the panel index. The numbers in parentheses along x-axis in subplots 1, 5 and 6 are the cycles corresponding to the failure/repair in the process of no maintenance intervention, CBM and PdM, respectively. The red solid dots and the green solid squares along y-axis represent the "real failed" panels and the "unnecessary repaired" panels. It can be seen that all the "real failed" panels shown in the first subplot are repaired in all other maintenance process priori to their failure, that is to say all maintenance strategies ensure safety. Scheduled maintenance and CBM-skip show exactly the same pattern. CBM wastes many panels near the EOL while PdM-skip and PdM have the least unnecessary repair.

Now discuss the difference between PdM and PdM-skip. PdM is designed completely independent without considering the time of maintenance schedule, during which the engine and non-structural airframe maintenance are always performed. All the stops are unscheduled

maintenance that occurred out of the time of scheduled maintenance. This can be costly due to no advance preparation. In PdM-skip, the average number of maintenance stop is slight higher than that in PdM (1.5 vs 1.2 in small uncertainties and 3.1 vs 2.4 in large uncertainties), but all their structural airframe maintenances are implemented during one of the 10 scheduled maintenance stops, when the engine and non-structural airframe are performed. The results indicate that PdM-skip fits well the objective that it ensures as much as possible that maintenance activities are carried out during the time of scheduled maintenance, and thus is more recommended than PdM.

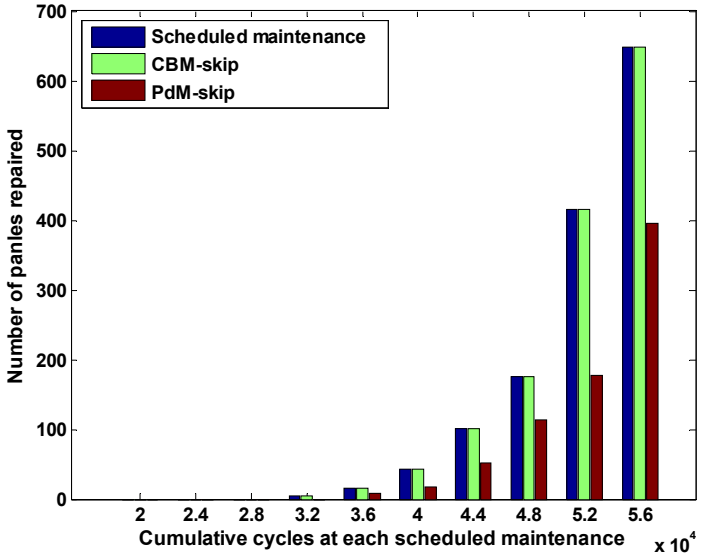

Figure 5-10 Comparison of maintenance strategies in terms of the number of repaired panels at each scheduled maintenance stop

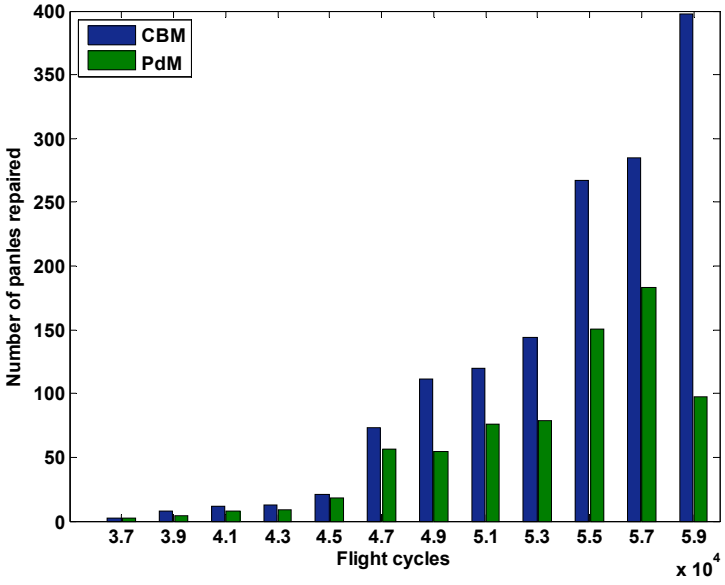

Figure 5-11 Comparison of CBM and PdM in terms of number of repaired panels within the time range of each bin

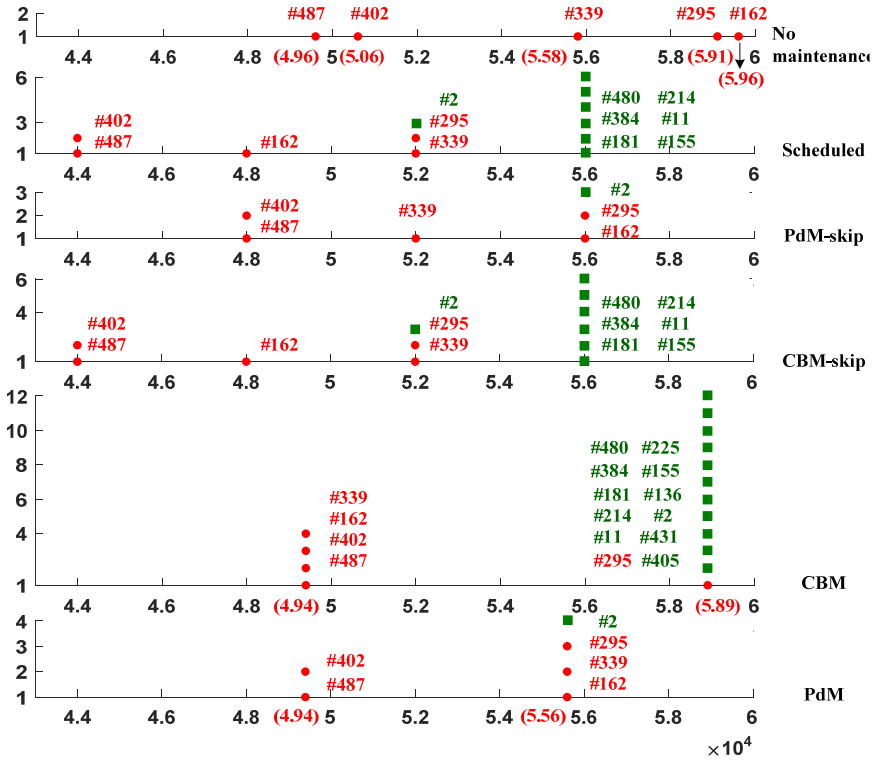

Figure 5-12 Different processes of aircraft #30

Table 5-9 Comparison of different strategies under small uncertainty case

	No maintenance	Scheduled	CBM	CBM-skip	PdM	PdM-skip
Panels fail/repaired over the entire fleet	223 failures	276 repairs	340 repairs	276 repairs	236 repairs	249 repairs
Unnecessary repaires	0	53	117	53	13	26
Minimal number of maintenance stops	-	10	1	0	0	0
Maximal number maintenance stops	-	10	4	4	4	4
Avg. number. of maintenance stops	-	10	1.2	1.6	1.3	1.5
Avg. number of unscheduled maintenance stops	-	0	1.2	0	1.3	0
Minimal number of repaired panels	-	0	0	0	0	0
Maximal number of repaired panels	-	8	11	8	8	8
Avg. number of repaired panels	-	2.7	3.4	2.7	2.3	2.4

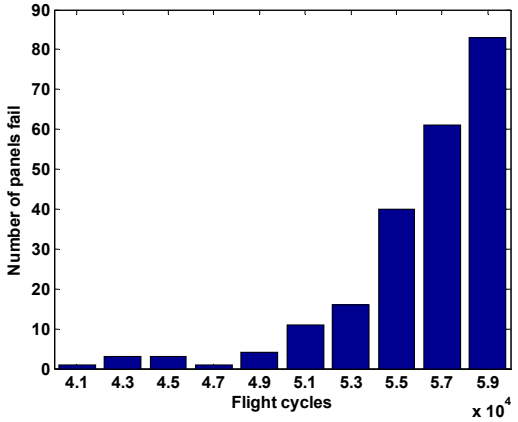

Figure 5-13 Number of panels fail within the time range of each bin in the case of no maintenance process intervention, in small uncertainty case

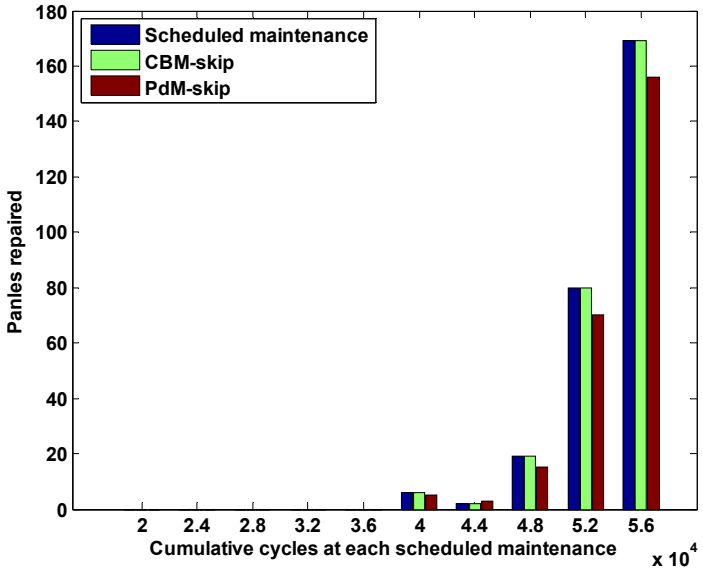

Figure 5-14 Comparison of maintenance strategies in terms of the number of repaired panels at each scheduled maintenance stop, in small uncertainty case

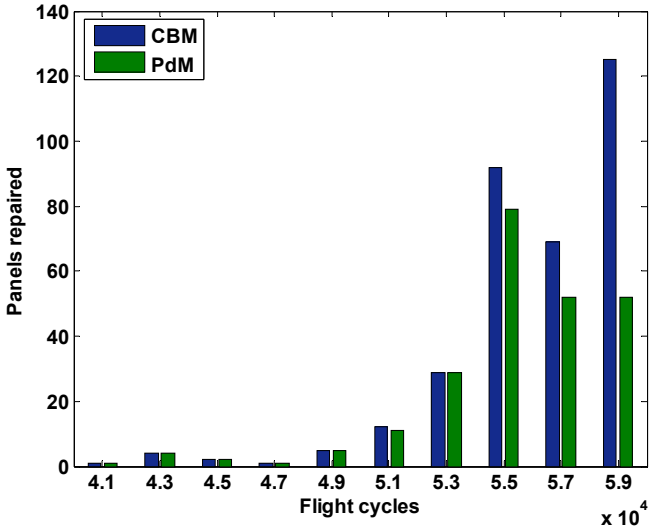

Figure 5-15 Comparison of CBM and PdM in terms of number of repaired panels within the time range of each bin, in small uncertainty case

In order to investigate the gains from prognostics-based maintenance over threshold-based maintenance, the six processes under small uncertainty case is simulated. The results are given in Table 5-9. Figure 5-13 to Figure 5-15 illustrate the statistical characters of failure/repaired time of panels over the entire fleet, *i.e.*, 100×500 panels in small uncertainty case. Through comparing the results of small uncertainty with the large uncertainty case, it sees that the prognostics-based strategies (PdM and PdM-skip) gain slightly over the threshold-based ones (CBM and CBM-skip) in terms of repaired panels in small uncertainty case, while significantly outperform the threshold-based strategies when large uncertainties are present. This is caused by the different philosophies of these two types of strategies. The prognostics based strategies repair a panel based

on its individual crack growth behavior while threshold based strategies has the same repair threshold for all panels. Specifically, when the uncertainties on material property parameters $\{m, C\}$ and on pressure p are small, both the panel-to-panel variability and the variability presenting in the crack propagation process are small, leading the cracks in the panels have similar propagation behavior. In this situation, the two types of strategies have similar performance. In contrast, when large uncertainties are present on $\{m, C\}$ and p, the cracks have big variability in propagation rate among the panel population. In the threshold-based strategies, due to the constant repair threshold, all panels with a crack size greater than the repair threshold are repaired, even if some of them have a very low growth rate and are not likely to fail until the aircraft's end of life. Prognostics-based strategies have an advantage in this situation since they treat the panels individually. Combined with the crack size and the material property parameters of each panel at current time, PdM/PdM-skip predict its crack growth trajectory in a future period and make the decision of whether or not replacing this panel based on this predicted behavior.

The 7th column gives the average maintenance costs per aircraft of different maintenance policies. According to the simulation results, no unscheduled maintenance is found in threshold-based maintenance. This does not mean that there will never be any but it is a very rare event which we do not capture with our fleet size. Therefore, the varying k_{un} has no effect to the cost of threshold-based maintenance. It can be seen that the CDPM leads to a significant cost savings compared with both traditional maintenance and threshold-based maintenance. The savings could be attributed to two aspects. Firstly, compared with the traditional scheduled maintenance, the CDPM skipped some unnecessary maintenance stops, thus reduced the set up cost. Secondly, unlike the scheduled maintenance and threshold-based maintenance, in which one repair criterion is applied for all panels (*i.e.*, panels with crack size greater than a_{rep} are repaired), by using the prognostics method, the CDPM treats each panel individually. Based on the predicted information of each panel, the number of repaired panels is optimized. This reduces the number of repaired panels at each maintenance stop.

Note that the difference in structural maintenance cost for different cost ratios is about 5%. This means that the optimal maintenance policy allows to squeeze out these last few percent in terms of cost gains based on the objective measure of the cost ratio, without having to tune any additional parameters. It is also important to note how the optimal cost driven policy is affected by the level

of uncertainties. We found that the cost optimal policy is most sensitive to the parameters of the maintenance decision (cost ratio) when the panel-to-panel variability is low compared to the prediction uncertainty. This can be explained as following: there are two items when predicting the crack size distribution at each scheduled maintenance, the first is predicting the mean and the second one is predicting the standard deviation after some additional cycles. If the panel-to-panel variability is large compared to the prediction uncertainty, then it is mainly the predicted mean value of crack size that matters and if the panel-to-panel variability is small compared to the prediction uncertainty then both the mean and standard deviation matter. The cost optimal policy is thus less sensitive in a large panel-to-panel variability case than in a low one even though the potential cost gains over traditional or threshold based maintenance would be larger with large panel to panel variability. On the other hand in a low panel-to-panel variability case, while the potential cost gains become smaller, the maintenance policy becomes much more sensitive to maintenance decision parameters (cost ratio) and using the cost optimal policy makes an increasingly significant difference. The cost optimal policy would be even more sensitive to the cost ratios in applications where the distribution of unscheduled events between two scheduled maintenances is more gradual. This would be for example the case when the variability in material properties would be smaller and the prediction uncertainty due to measurement noise would be larger. The optimality of the maintenance strategy also guarantees that the structural maintenance cost is minimal without having to tune any additional parameters in the maintenance strategy. In addition, it allows to avoid having to choose a quantile (for example 95%) of the predicted distribution after some additional cycles when determining which panels to replace.

Table 5-10 Comparison results of different maintenance policies

Scenario	**Cost ratio (c_s/c_{us})**	**Maintenance policy**	**Avg. No. of M.S.[a]/aircraft**	**Avg.No.of U.M.S.[b] /aircraft**	**Avg. No. of R.P.[c]/aircraft**	**Avg. M.C.[d]/aircraft**
-	-	Traditional Scheduled	10	-	14.2	17.9
-	-	CBM-skip	3.6	0	14.2	8.2
kun=0.9	0.16	CDPM	2.9	0.36	7.3	5.7
kun=2	0.08	CDPM	3.0	0.02	7.4	5.8
kun=3	0.05	CDPM	3.0	0.02	7.4	5.8
kun=5	0.03	CDPM	3.1	0	7.5	5.9
kun=10	0.01	CDPM	3.1	0	7.5	5.9

The cost difference between the CDPM and the traditional scheduled maintenance helps make the decision concerning the implementation of an SHM system on aircraft. More specifically, if the

cost incurred by installing and operating an SHM system is less than the cost saved by using SHM, then it is worth to install it on aircraft.

5.7 Summary

In this chapter, the on-board SHM system is proposed to monitor the damage state of the fuselage instead of the traditional NDI and DVI inspection, which is time-consuming since the aircraft needs to be partly disassembled in order to allow inspection. The information collected by SHM is used in the prognostics method elaborated in CHAPTER 4. Based on the prognostics method, two types of predictive maintenance strategies are developed.

The first type includes two variants, PdM and PdM-skip. They are designed in order to skip the unnecessary scheduled maintenance. The difference is that PdM is designed completely independent without considering the time of scheduled maintenance, during which the engine and non-structural airframe maintenance are always performed. While PdM-skip ensures as much as possible that maintenance activities are carried out during the time of scheduled maintenance. Having the structural airframe maintenance at the same time with engine and non-structural maintenance would tend to reduce cost. The proposed PdM /PdM-skip are compared with the traditional scheduled maintenance, and CBM/CBM-skip proposed in [142] on a fleet of short range commercial aircraft. A cost model is used to quantify and compare the cost effectiveness among different strategies. It is found that PdM/PdM-skip gain significantly over CBM/CBM-skip when large uncertainties are present because the remaining useful life is calculated individually for each panel. As for the comparison between the PdM-skip and PdM, PdM-skip leads to slightly more maintenance stops than PdM. However, all maintenance stops incurred by PdM occur as unscheduled maintenance, which is more expensive due to less advance notice, while almost all maintenance stops incurred by PdM-skip happen at scheduled maintenance.

The second type of predictive maintenance is CDPM, which leverages the benefit from both the scheduled and unscheduled maintenance. It skips some unnecessary scheduled maintenance stops but at the same time guarantees the aircraft safety by querying the health state of the fuselage frequently. CDPM makes a tradeoff between the cost ratio of scheduled cost over unscheduled cost and the probability of occurring unscheduled maintenance. It selects a group of panels which are to

be repaired at a scheduled maintenance stop so as to minimize the cost. The CDPM is applied to an example of a fleet of short range commercial aircraft. The simulation results are compared with the traditional scheduled maintenance and CBM-skip in terms of the average number of maintenance stops, the average number of repaired panels and the average cost per aircraft under same operational conditions. The results show a significant cost reduction achieved by employing the CDPM. Furthermore CDPM allows to assure the cost optimality of the maintenance policy without having to tune any additional parameters. The cost optimality then allows to squeeze out the last few percent of cost savings from prediction based maintenance.

By comparing the cost difference between SHM-based maintenance (PdM, PdM-skip, CBM, CBM-skip) and the scheduled maintenance, one can make the decision concerning the implementation of the SHM system on aircraft. More specifically, if the cost incurred by installing and operating an SHM system is lower than the cost saved by employing SHM, then it is worth to install the SHM system on the aircraft.

CHAPTER 6 CONCLUSIONS

Surrounding the issues posed at the beginning of this work, *i.e.*, how to use the noisy SHM data to estimate accurately the crack size, and how to establish a link between the prognostics and maintenance strategy, this work investigated multiple aspects of these issues and proposed multiple building blocks to address these issues extensively. Firstly, several filtering methods are explored and the most appropriate one, the EKF, is chosen to address the problem of state-parameter estimation from noisy data in the context of fatigue damage growth in an aeronautical context. Then a model-based prognostics method coupling the EKF and the first-order perturbation method is developed. Finally, the prognostics method is linked to maintenance planning. Two types of predictive maintenance are proposed.

Specifically, in CHAPTER 3, it investigates the Bayesian filter theory and its three numerical implementation methods, EKF, UKF and PF. These three numerical approaches are applied to a specific state-paramter estimation problem in the context of fatigue damage model presented by Paris' law. The performance of the three methods are compared. The results show that for the specific application of fatigue damage model, the EKF gives satisfactory performance and has the lowest computation cost. Thus the EKF is chosen to be used in the model-based prognostics method.

Then in CHAPTER 4, an EKF coupled first-order perturbation method for model-based prognostics for fatigue damage propagation in fuselage panels is proposed. The presented model-based prognostics method includes two sequential phases (1) state-parameter identification from noisy measurements data using the EKF algorithm and (2) future damage prediction using the proposed first-order perturbation method. The first-order perturbation method is able to predict analytically the evolution of the crack size distribution in some future flight cycles from the present estimations given by the EKF. The first-order perturbation method is compared with Monte Carlo method. The comparison shows both of them provide right decisions of repairing panels but the first-order perturbation method leads to great savings in computational cost. This indicates that the first-order perturbation method is effective and could be used to replace Monte Carlo method in prognostics problem of fuselage panel fatigue crack growth.

In CHAPTER 5, the developed model-based prognostics method is integrated in the maintenance planning for a fleet of aircraft in an airline. Each panel of each aircraft is banded with the model-based prognostics method, *i.e.*, for each panel, its crack size and material property parameters $\{m,C\}$ estimation as well as its crack propagation prediction are done by the model-based prognostics method individually. By using the predictive information in two different ways, two types of predictive maintenance are proposed. Specifically, in the first type of predictive maintenance policy, the decision of whether repairing a panel or not at a maintenance stop depends on the predicted crack size distribution in some future flight cycles. In the second type of predictive maintenance policy, the ratio of scheduled repair cost over unscheduled repair cost is taken as an input of repair policy, *i.e.*, the repair decision depends on the relation between the cost ratio and the probability of occurrence of the unscheduled maintenance. In this case, the repair policy acts as a branch controller. The unscheduled maintenance is avoided as much as possible if the cost ratio is low (which indicated that the unscheduled cost is much higher than that of scheduled cost). In the contrast, if the unscheduled cost is not much higher than the scheduled cost, then fewer panels are repaired at scheduled maintenance and the occurrence of unscheduled maintenance is relaxed.

For perspectives and future work, one can extend from the followings aspects. (1) More complicated and practical fatigue damage models, such as Huang' s model [153], Wheeler's model [154] can be considered. (2) Some assumptions can be released to make the simulation more practical. For example, multiple cracks can be assumed in one fuselage panel. (3) Due to the usage of some more sophisticate model, the number of variables to be identified can be large and running a model-based method is computational expensive. In such cases, some data-driven methods can be considered. Data-driven methods are more appropriate in such cases that the understanding of the principle of degradation process is not comprehensive, or the system is sufficiently complex such that develop an accurate degradation model is prohibitively expensive. (4) The reliability of SHM system can be investigated. The probability of crack size detection by SHM can be considered. In addition, the balance between the maintenance cost saved by SHM system and the cost incurred due to the installation of SHM can be studied so that one can know whether it is worth installing an SHM system (5) From the industry application point of view, an aircraft Prognostics and Health Management (PHM) system including hardware and software can be developed. The hardware can be composed of sensor network, equipment of data acquisition, transformation and storage while

the software parts involve various data processing algorithms and prognostics methods. The aircraft PHM system can enable the airlines or manufactures to know the health state of an aircraft in future time, therefore, to plan the maintenance just in time based on the predicted health state rather than a fixed schedule.

APPENDIX A COST MODEL

The engine maintenance cost is given in terms of empirical expression in [152]. The parameters that will be used in the cost model are given in Table A-1. The engine maintenance cost per flight is given as

$$N_{\mathrm{e}} \times (\text{engine labor cost} + \text{material cost} \times \frac{t+1.3}{t-0.25}) \tag{A-1}$$

The engine labor cost is given by Eq.(A-2)., where C_1=1.27-0.2*BPR^2, C_3=0.032*n_c+K

$$0.21 \times R \times C_1 \times C_3 \times (1+T)^{0.4} \tag{A-2}$$

Engine material cost is given in Eq.(A-3), where C_2=0.4*(OAPR)$^{1.3}$+0.4

$$2.56 \times (1+T)^{0.8} \times C_1 \times C_2 \times C_3 \tag{A-3}$$

Table A-1 Parameters used in cost model

Notation	Description	Value
Ne	Number of engines	2
T	Sea level static thrust (tons)	24.6
BPR	Engine bypass ratio	6
nc	Number of compressor stages	9
K	Parameters relative to shaft number of compressor	0.57
OAPR	Overall pressure ratio of the engine	31.3

APPENDIX B PROOF OF THE COST OPTIMAL POLICY

This Appendix gives a mathematical proof of the cost optimal policy presented in Section 5.5.3. Eq.(5-13) is firstly proved as the prerequisites for the proof. Recall that in Eq.(5-13), it gives $1 \le B_I < b_J \le N$. Suppose the contrary

$$1 \le B_I < b_J \le N \tag{B-1}$$

Then we have

$$c_{us} \sum_{j=1}^{b_J} P(us \mid a^j) = c_{us} \left(\sum_{j=1}^{B_I} P(us \mid a^j) + \sum_{j=B_I+1}^{b_J} P(us \mid a^j) \right) \tag{B-2}$$

Since $B_I < b_J$, according to Eq.(5-16):

$$c_{us} (\sum_{j=1}^{B_I} P(us \mid a^j)) < c_0 + B_I c_s \tag{B-3}$$

And according to Eq.(5-14):

$$c_{us} (\sum_{j=B_I+1}^{b_J} P(us \mid a^j)) < c_s (b_J - B_I) \tag{B-4}$$

Sum the inequality Eq.(B-3) and Eq.(B-4):

$$c_0 + B_J c_s > c_{us} \sum_{j=1}^{B_J} P(us \mid a^j) \tag{B-5}$$

which is impossible since Eq.(5-17) is not satisfied. So $1 < b_J < B_I < N$

Now, we prove the cost optimal repair policy. Reminder that the optimal policy $\mathbf{d}_n^*$ is

If $J = \varnothing$

$$d_n^{j^*} = 0, \; for \; j = 1,2,...N$$

Else

$$d_n^{j^*} = \begin{cases} 1 & for \; j = 1,2,...B_I \\ 0 & for \; j = B_I + 1,...N \end{cases}$$

The maintenance cost is a function of decision. Our objective is to prove $C(\mathbf{d}_n) > C(\mathbf{d}_n^*)$ for any maintenance policy $\mathbf{d}_n$. Let us define the following set:

$$A = \{1 \le j \le B_I \mid d_n^j = 1\} \tag{B-6}$$

$$\overline{A} = \{1 \le j \le B_I \mid d_n^j = 0\} \tag{B-7}$$

$$B = \{B_I + 1 \le j \le N \mid d_n^j = 1\} \tag{B-8}$$

$$\overline{B} = \{B_I + 1 < j \le N \mid d_n^j = 0\} \tag{B-9}$$

$|A|, |\overline{A}|, |B|$ and $|\overline{B}|$ are the cardinality of the set A, $\overline{A}$, B and $\overline{B}$, respectively. Obviously, we have the following: $|A| + |\overline{A}| = B_I$ and $|B| + |\overline{B}| = N - B_I$. The maintenance cost $C(\mathbf{d}_n)$ is then computed as

$$C(\mathbf{d}_n) = c_0 + c_s \mid A \mid + c_{us}(\sum_{j \in \overline{A}} P(us \mid a^j)) + c_s \mid B \mid + c_{us} \sum_{j \in \overline{B}} P(us \mid a^j) \tag{B-10}$$

Since $c_{us} P(us \mid a^j) \ge c_s$, for j=1,2,…B_I (see Eq.(5-14)). Then:

$\sum_{j \in \overline{A}} c_{us} P(us \mid a^j) \ge c_s \mid \overline{A} \mid$, hence

$$\begin{aligned} & c_0 + c_s \mid A \mid + c_{us}(\sum_{j \in \overline{A}} P(us \mid a^j)) \\ & \ge c_0 + c_s \mid A \mid + c_s \mid \overline{A} \mid \\ & = c_0 + B_I c_s \end{aligned} \tag{B-11}$$

Since $c_{us} P(us \mid a) < c_s$, for j=B_I+1,…N (see Eq.(5-14)).Then:

$$\begin{aligned} & c_s \mid B \mid + c_{us} \sum_{j \in \overline{B}} P(us \mid a^j) \\ & > c_{us} \sum_{j \in B} P(us \mid a^j) + c_{us} \sum_{j \in \overline{B}} P(us \mid a^j) \\ & = \sum_{j=B_I+1}^{N} c_{us} P(us \mid a^j) \end{aligned} \tag{B-12}$$

Sum the inequality Eq.(B-11) and Eq.(B-12), then:

$$c_0 + c_s \mid A \mid + c_{us}(\sum_{j \in \overline{A}} P(us \mid a^j)) + c_s \mid B \mid + c_{us} \sum_{j \in \overline{B}} P(us \mid a^j) \ge c_0 + B_I c_s + \sum_{j=B_I+1}^{N} c_{us} P(us \mid a^j) \tag{B-13}$$

The left term of the inequality is the maintenance cost $C(\mathbf{d}_n)$ while the right term of the inequality if the optimal cost $C(\mathbf{d}_n^*)$, so we have $C(\mathbf{d}_n) > C(\mathbf{d}_n^*)$. Up to now, the cost under any other decision $\mathbf{d}_n$ is greater than the cost under the optimal decision $\mathbf{d}_n^*$ has been proved.

REFERENCES

[1] Lynch, J.P., *A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring.* The Shock and Vibration Digest, 2006. **38**(2): p. 91-128.

[2] Diamanti, K. and Soutis, C., *Structural health monitoring techniques for aircraft composite structures.* Progress in Aerospace Sciences, 2010. **46**(8): p. 342-352.

[3] Xiaoliang, Z., et al., *Active health monitoring of an aircraft wing with embedded piezoelectric sensor/actuator network: I. Defect detection, localization and growth monitoring.* Smart Materials and Structures, 2007. **16**(4): p. 1208.

[4] Jeong-Beom, I. and Fu-Kuo, C., *Detection and monitoring of hidden fatigue crack growth using a built-in piezoelectric sensor/actuator network: I. Diagnostics.* Smart Materials and Structures, 2004. **13**(3): p. 609.

[5] Ihn, J.-B. and Chang, F.-K., *Detection and monitoring of hidden fatigue crack growth using a built-in piezoelectric sensor/actuator network: II. Validation using riveted joints and repair patches.* Smart Materials and Structures, 2004. **13**(3): p. 621.

[6] Leong, W.H., et al., *Structural health monitoring using scanning laser vibrometry: III. Lamb waves for fatigue crack detection.* Smart Materials and Structures, 2005. **14**(6): p. 1387.

[7] Staszewski, W.J., et al., *Structural health monitoring using scanning laser vibrometry: I. Lamb wave sensing.* Smart Materials and Structures, 2004. **13**(2): p. 251.

[8] Ignatovich, S.R., et al., *Fatigue damage and sensor development for aircraft structural health monitoring.* Theoretical and Applied Fracture Mechanics, 2013. **65**: p. 23-27.

[9] McCall, J.J., *Maintenance Policies for Stochastically Failing Equipment: A Survey.* Management Science, 1965. **11**(5): p. 493-524.

[10] Sherif, Y.S. and Smith, M.L., *Optimal maintenance models for systems subject to failure–A Review.* Naval Research Logistics Quarterly, 1981. **28**(1): p. 47-74.

[11] Cho, D.I. and Parlar, M., *A survey of maintenance models for multi-unit systems.* European Journal of Operational Research, 1991. **51**(1): p. 1-23.

[12] Canfield, R.V., *Cost Optimization of Periodic Preventive Maintenance.* IEEE Transactions on Reliability, 1986. **35**(1): p. 78-81.

[13] Nakagawa, T., *Optimal policy of continuous and discrete replacement with minimal repair at failure.* Naval Research Logistics Quarterly, 1984. **31**(4): p. 543-550.

[14] Tam, A., Chan, W., and Price, J., *Optimal maintenance intervals for a multi-component system.* Production Planning and Control, 2006. **17**(8): p. 769-779.

[15] Baldin, A.E., *Condition based maintenance in chemical and chemical fiber industries.* Monitoring Diagnostics in the Industry Congress, Prague, 1975.

[16] Baidin, A.E., *The reliability of condition-based maintenance.* 3rd EFNMS Congress, Stockholm, 1976.

[17] Baldin, A.E., *Condition-based maintenance: A powerful tool for modern management.* Terotecnica I, Elsevier Sc Pub Co., 1979.

[18] Baidin, A.E., *Condition-based maintenance implementation and experiences.* VIAI Man Congress, Trieste, 1973.

[19] *Technical diagnostic-and condition-based maintenance for better plant availability.* Measurement, 1986. **4**(1): p. 7-22.

[20] Zhou, X., Xi, L., and Lee, J., *Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation.* Reliability Engineering & System Safety, 2007. **92**(4): p. 530-534.

[21] Van Horenbeek, A. and Pintelon, L., *A dynamic predictive maintenance policy for complex multi-component systems.* Reliability Engineering & System Safety, 2013. **120**: p. 39-50.

[22] Traore, M., Chammas, A., and Duviella, E., *Supervision and prognosis architecture based on dynamical classification method for the predictive maintenance of dynamical evolving systems.* Reliability Engineering & System Safety, 2015. **136**: p. 120-131.

[23] Nguyen, K.-A., Do, P., and Grall, A., *Multi-level predictive maintenance for multi-component systems.* Reliability Engineering & System Safety, 2015. **144**: p. 83-94.

[24] Curcurù, G., Galante, G., and Lombardo, A., *A predictive maintenance policy with imperfect monitoring.* Reliability Engineering & System Safety, 2010. **95**(9): p. 989-997.

[25] Li, H., et al., *Improving rail network velocity: A machine learning approach to predictive maintenance.* Transportation Research Part C: Emerging Technologies, 2014. **45**: p. 17-26.

[26] Bansal, D., Evans, D.J., and Jones, B., *BJEST: A reverse algorithm for the real-time predictive maintenance system.* International Journal of Machine Tools and Manufacture, 2006. **46**(10): p. 1068-1078.

[27] Ferreiro, S., et al., *Application of Bayesian networks in prognostics for a new Integrated Vehicle Health Management concept.* Expert Systems with Applications, 2012. **39**(7): p. 6402-6418.

[28] Ming Tan, C. and Raghavan, N., *A framework to practical predictive maintenance modeling for multi-state systems.* Reliability Engineering & System Safety, 2008. **93**(8): p. 1138-1150.

[29] Deloux, E., Castanier, B., and Bérenguer, C., *Predictive maintenance policy for a gradually deteriorating system subject to stress.* Reliability Engineering & System Safety, 2009. **94**(2): p. 418-431.

[30] Kaiser, K.A. and Gebraeel, N.Z., *Predictive Maintenance Management Using Sensor-Based Degradation Models.* IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2009. **39**(4): p. 840-849.

[31] Zhao, Z., et al., *Predictive maintenance policy based on process data.* Chemometrics and Intelligent Laboratory Systems, 2010. **103**(2): p. 137-143.

[32] Huda, A.S.N. and Taib, S., *Application of infrared thermography for predictive/preventive maintenance of thermal defect in electrical equipment.* Applied Thermal Engineering, 2013. **61**(2): p. 220-227.

[33] Orhan, S., Aktürk, N., and Çelik, V., *Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: Comprehensive case studies.* NDT & E International, 2006. **39**(4): p. 293-298.

[34] Hashemian, H.M., *Wireless sensors for predictive maintenance of rotating equipment in research reactors.* Annals of Nuclear Energy, 2011. **38**(2–3): p. 665-680.

[35] Carnero, M.C., *Selection of diagnostic techniques and instrumentation in a predictive maintenance program. A case study.* Decision Support Systems, 2005. **38**(4): p. 539-555.

[36] Efthymiou, K., et al., *45th CIRP Conference on Manufacturing Systems 2012On a Predictive Maintenance Platform for Production Systems.* Procedia CIRP, 2012. **3**: p. 221-226.

[37] Chen, D., Wang, X., and Zhao, J., *2012 International Workshop on Information and Electronics EngineeringAircraft Maintenance Decision System Based on Real-time Condition Monitoring.* Procedia Engineering, 2012. **29**: p. 765-769.

[38] Garcia, M.C., Sanz-Bobi, M.A., and del Pico, J., *SIMAP: Intelligent System for Predictive Maintenance: Application to the health condition monitoring of a windturbine gearbox.* Computers in Industry, 2006. **57**(6): p. 552-568.

[39] Bansal, D., Evans, D.J., and Jones, B., *A real-time predictive maintenance system for machine systems.* International Journal of Machine Tools and Manufacture, 2004. **44**(7–8): p. 759-766.

[40] Carnero, M., *An evaluation system of the setting up of predictive maintenance programmes.* Reliability Engineering & System Safety, 2006. **91**(8): p. 945-963.

[41] Lee, J., et al., *Industrial Big Data Analytics and Cyber-physical Systems for Future Maintenance & Service Innovation.* Procedia CIRP, 2015. **38**: p. 3-7.

[42] Sikorska, J.Z., Hodkiewicz, M., and Ma, L., *Prognostic modelling options for remaining useful life estimation by industry.* Mechanical Systems and Signal Processing, 2011. **25**(5): p. 1803-1836.

[43] An, D., Kim, N.H., and Choi, J.-H., *Practical options for selecting data-driven or physics-based prognostics algorithms with reviews.* Reliability Engineering & System Safety, 2015. **133**: p. 223-236.

[44] Kandukuri, S.T., et al., *A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management.* Renewable and Sustainable Energy Reviews, 2016. **53**: p. 697-708.

[45] Lee, J., et al., *Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications.* Mechanical Systems and Signal Processing, 2014. **42**(1–2): p. 314-334.

[46] 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, 2006. **20**(7): p. 1483-1510.

[47] Si, X.-S., et al., *Remaining useful life estimation – A review on the statistical data driven approaches.* European Journal of Operational Research, 2011. **213**(1): p. 1-14.

[48] Jouin, M., et al., *Particle filter-based prognostics: Review, discussion and perspectives.* Mechanical Systems and Signal Processing, 2016. **72–73**: p. 2-31.

[49] Hu, C., et al., *Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life.* Reliability Engineering & System Safety, 2012. **103**: p. 120-135.

[50] Mosallam, A., Medjaher, K., and Zerhouni, N., *Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction.* Journal of Intelligent Manufacturing, 2014: p. 1-12.

[51] Le Son, K., et al., *Remaining useful life estimation based on stochastic deterioration models: A comparative study.* Reliability Engineering & System Safety, 2013. **112**: p. 165-175.

[52] Gobbato, M., Kosmatka, J.B., and Conte, J.P., *A recursive Bayesian approach for fatigue damage prognosis: An experimental validation at the reliability component level.* Mechanical Systems and Signal Processing, 2014. **45**(2): p. 448-467.

[53] Ling, Y. and Mahadevan, S., *Integration of structural health monitoring and fatigue damage prognosis.* Mechanical Systems and Signal Processing, 2012. **28**: p. 89-104.

[54] Karandikar, J.M., Kim, N.H., and Schmitz, T.L., *Prediction of remaining useful life for fatigue-damaged structures using Bayesian inference.* Engineering Fracture Mechanics, 2012. **96**: p. 588-605.

[55] Gobbato, M., et al., *A reliability-based framework for fatigue damage prognosis of composite aircraft structures.* Probabilistic Engineering Mechanics, 2012. **29**: p. 176-188.

[56] Sun, J., et al., *Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model.* Mechanical Systems and Signal Processing, 2014. **45**(2): p. 396-407.

[57] Hu, Y., et al., *A particle filtering and kernel smoothing-based approach for new design component prognostics.* Reliability Engineering & System Safety, 2015. **134**: p. 19-31.

[58] Zio, E. and Peloni, G., *Particle filtering prognostic estimation of the remaining useful life of nonlinear components.* Reliability Engineering & System Safety, 2011. **96**(3): p. 403-409.

[59] Baraldi, P., et al., *Ensemble neural network-based particle filtering for prognostics.* Mechanical Systems and Signal Processing, 2013. **41**(1–2): p. 288-300.

[60] Chen, C., Vachtsevanos, G., and Orchard, M.E., *Machine remaining useful life prediction: An integrated adaptive neuro-fuzzy and high-order particle filtering approach.* Mechanical Systems and Signal Processing, 2012. **28**: p. 597-607.

[61] Si, X.S., *An Adaptive Prognostic Approach via Nonlinear Degradation Modeling: Application to Battery Data.* IEEE Transactions on Industrial Electronics, 2015. **62**(8): p. 5082-5096.

[62] Si, X.S., et al., *Estimating Remaining Useful Life With Three-Source Variability in Degradation Modeling.* IEEE Transactions on Reliability, 2014. **63**(1): p. 167-190.

[63] Wan, E.A. and Merwe, R.V.D. *The unscented Kalman filter for nonlinear estimation*. in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*. 2000.

[64] Julier, S.J. and Uhlmann, J.K., *A general method for approximating nonlinear transformations of probability distributions*. 1996, Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford.

[65] Garcia-Velo, J. and Walker, B., *Aerodynamic parameter estimation for high-performance aircraft using extended Kalman filtering*, in *20th Atmospheric Flight Mechanics Conference*. 1995, American Institute of Aeronautics and Astronautics.

[66] Chowdhary, G. and Jategaonkar, R., *Aerodynamic parameter estimation from flight data applying extended and unscented Kalman filter.* Aerospace Science and Technology, 2010. **14**(2): p. 106-117.

[67] Aksoy, S., Muhurcu, A., and Kizmaz, H. *State and parameter estimation in induction motor using the Extended Kalman Filtering algorithm*. in *Modern Electric Power Systems (MEPS), 2010 Proceedings of the International Symposium*. 2010.

[68] Bisht, S.S. and Singh, M.P., *An adaptive unscented Kalman filter for tracking sudden stiffness changes.* Mechanical Systems and Signal Processing, 2014. **49**(1–2): p. 181-195.

[69] D'Alfonso, L., et al., *Mobile robot localization via EKF and UKF: A comparison based on real data.* Robotics and Autonomous Systems, 2015. **74, Part A**: p. 122-127.

[70] Kihun, K., et al. *Estimation of hydrodynamic coefficients of a test-bed AUV-SNUUV I by motion test*. in *OCEANS '02 MTS/IEEE*. 2002.

[71] Bressel, M., et al., *Extended Kalman Filter for prognostic of Proton Exchange Membrane Fuel Cell.* Applied Energy, 2016. **164**: p. 220-227.

[72] Julier, S.J. and Uhlmann, J.K. *New extension of the Kalman filter to nonlinear systems*. in *AeroSense'97*. 1997. International Society for Optics and Photonics.

[73] VanDyke, M.C., Schwartz, J.L., and Hall, C.D., *Unscented Kalman filtering for spacecraft attitude state and parameter estimation.* Department of Aerospace & Ocean Engineering, Virginia Polytechnic Institute & State University, Blacksburg, Virginia, 2004.

[74] Crassidis, J.L. and Markley, F.L., *Unscented Filtering for Spacecraft Attitude Estimation.*

Journal of Guidance, Control, and Dynamics, 2003. **26**(4): p. 536-542.

[75] Qu, C.C. and Hahn, J., *Process monitoring and parameter estimation via unscented Kalman filtering.* Journal of Loss Prevention in the Process Industries, 2009. **22**(6): p. 703-709.

[76] Wendel, J., et al., *A Performance Comparison of Tightly Coupled GPS/INS Navigation Systems based on Extended and Sigma Point Kalman Filters.* Navigation, 2006. **53**(1): p. 21-31.

[77] Jan, W., et al., *Comparison of Extended and Sigma-Point Kalman Filters for Tightly Coupled GPS/INS Integration*, in *AIAA Guidance, Navigation, and Control Conference and Exhibit*. 2005, American Institute of Aeronautics and Astronautics.

[78] Tang, L., et al. *Filtering and prediction techniques for model-based prognosis and uncertainty management*. in *2010 Prognostics and System Health Management Conference*. 2010.

[79] An, D., Choi, J.-H., and Kim, N.H., *Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab.* Reliability Engineering & System Safety, 2013. **115**: p. 161-169.

[80] Tang, L., et al. *Novel metrics and methodologies for the verification and validation of prognostic algorithms*. in *Aerospace Conference, 2011 IEEE*. 2011.

[81] Dalal, M., Ma, J., and He, D., *Lithium-ion battery life prognostic health management system using particle filtering framework.* Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2011. **225**(1): p. 81-90.

[82] Jouin, M., et al., *Prognostics of PEM fuel cell in a particle filtering framework.* International Journal of Hydrogen Energy, 2014. **39**(1): p. 481-494.

[83] Jouin, M., et al. *Prognostics of Proton Exchange Membrane Fuel Cell stack in a particle filtering framework including characterization disturbances and voltage recovery*. in *Prognostics and Health Management (PHM), 2014 IEEE Conference on*. 2014.

[84] Kimotho, J.K., Meyer, T., and Sextro, W. *PEM fuel cell prognostics using particle filter with model parameter adaptation*. in *Prognostics and Health Management (PHM), 2014 IEEE Conference on*. 2014.

[85] Zhang, B., et al. *A multi-fault modeling approach for fault diagnosis and failure prognosis of engineering systems*. in *Annual conference of the prognostics and health management society, San Diego, CA*. 2009.

[86] Guo, L., et al. *Comparison of resampling algorithms for particle filter based remaining useful life estimation*. in *2014 International Conference on Prognostics and Health Management*. 2014.

[87] Dong, H., et al., *Lithium-ion battery state of health monitoring and remaining useful life prediction based on support vector regression-particle filter.* Journal of Power Sources, 2014. **271**: p. 114-123.

[88] Yan, H.C., Pang, C.K., and Zhou, J.H. *Precognitive maintenance and probabilistic assessment of tool wear using particle filters*. in *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*. 2013.

[89] Arulampalam, M.S., et al., *A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking.* IEEE Transactions on Signal Processing, 2002. **50**(2): p. 174-188.

[90] Yi, G., Shesheng, G., and Yu, G. *Improved regularized particle filter algorithm for SINS/SAR integrated navigation*. in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. 2010.

[91] Zhai, Y. and Yeary, M. *Implementing particle filters with Metropolis-Hastings algorithms*. in

Region 5 Conference: Annual Technical and Leadership Workshop, 2004. 2004.

[92] Li, T., Sattar, T.P., and Sun, S., *Deterministic resampling: Unbiased sampling to avoid sample impoverishment in particle filters.* Signal Processing, 2012. **92**(7): p. 1637-1645.

[93] Andrieu, C., Freitas, N.D., and Doucet, A. *Sequential MCMC for Bayesian model selection*. in *Higher-Order Statistics, 1999. Proceedings of the IEEE Signal Processing Workshop on*. 1999.

[94] Nordh, J. *Metropolis-hastings improved particle smoother and marginalized models*. in *Signal Processing Conference (EUSIPCO), 2015 23rd European*. 2015.

[95] An, D., Choi, J.-h., and Kim, N.H., *A Comparison Study of Methods for Parameter Estimation in the Physics-based Prognostics*, in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. 2012, American Institute of Aeronautics and Astronautics.

[96] Karandikar, J.M., Kim, N.H., and Schmitz, T.L., *Prediction of remaining useful life for fatigue-damaged structures using Bayesian inference.* Engineering Fracture Mechanics, 2012. **96**(0): p. 588-605.

[97] Andrieu, C., et al., *An Introduction to MCMC for Machine Learning.* Machine Learning, 2003. **50**(1): p. 5-43.

[98] Srinivasan, V. and Mason, C.H., *Technical Note—Nonlinear Least Squares Estimation of New Product Diffusion Models.* Marketing Science, 1986. **5**(2): p. 169-178.

[99] Gill, P. and Murray, W., *Algorithms for the Solution of the Nonlinear Least-Squares Problem.* SIAM Journal on Numerical Analysis, 1978. **15**(5): p. 977-992.

[100] Boukamp, B.A., *A Nonlinear Least Squares Fit procedure for analysis of immittance data of electrochemical systems.* Solid State Ionics, 1986. **20**(1): p. 31-44.

[101] Moré, J.J., *The Levenberg-Marquardt algorithm: implementation and theory*, in *Numerical analysis*. 1978, Springer. p. 105-116.

[102] Grewal, M.S. and Andrews, A.P., *Kalman filtering: Theory and Practice with MATLAB*. 2014: John Wiley & Sons.

[103] Choi, E.-J., et al., *Onboard orbit determination using GPS observations based on the unscented Kalman filter.* Advances in Space Research, 2010. **46**(11): p. 1440-1450.

[104] Harris, C.E., et al., *Analytical Methodology for Predicting Widespread Fatigue Damage Onset in Fuselage Structure.* Journal of Aircraft, 1998. **35**(2): p. 307-317.

[105] Mohanty, J.R., Verma, B.B., and Ray, P.K., *Prediction of fatigue crack growth and residual life using an exponential model: Part I (constant amplitude loading).* International Journal of Fatigue, 2009. **31**(3): p. 418-424.

[106] Nilsson, K.-F. *Elasto-plastic models for interaction between a major crack and multiple small cracks*. in *FAA-NASA Symposium on the Continued Airworthiness of Aircraft Structures, Atlanta, GA, Proceedings.* 1996.

[107] Molent, L. and Barter, S.A., *A comparison of crack growth behaviour in several full-scale airframe fatigue tests.* International Journal of Fatigue, 2007. **29**(6): p. 1090-1099.

[108] Paris, P. and Erdogan, F., *A Critical Analysis of Crack Propagation Laws.* Journal of Basic Engineering, 1963. **85**(4): p. 528-533.

[109] Sabet, M.T., Pouria, S., and Zarini, M., *Extended and Unscented Kalman filters for parameter estimation of an autonomous underwater vehicle.* Ocean Engineering, 2014. **91**: p. 329-339.

[110] Huang, B. and Du, X., *Probabilistic uncertainty analysis by mean-value first order Saddlepoint Approximation.* Reliability Engineering & System Safety, 2008. **93**(2): p.

325-336.

[111] Bristow, J.W. and Irving, P.E., *Safety factors in civil aircraft design requirements.* Engineering Failure Analysis, 2007. **14**(3): p. 459-470.

[112] Cavallini, G. and Lazzeri, R., *A probabilistic approach to fatigue risk assessment in aerospace components.* Engineering Fracture Mechanics, 2007. **74**(18): p. 2964-2970.

[113] Nelson, A.T. and Wan, E.A. *A two-observation Kalman framework for maximum-likelihood modeling of noisy time series*. in *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on*. 1998.

[114] Wan, E.A. and Nelson, A.T., *Dual Extended Kalman Filter Methods*, in *Kalman Filtering and Neural Networks*. 2002, John Wiley & Sons, Inc. p. 123-173.

[115] SABES, P.N. and JORDAN, M.I., *Advances in neural information processing systems*. 1995: In G. Tesauro & D. Touretzky & T. Leed.

[116] Jay, F. and Goetz, J., *IEEE Standard Dictionary of Electrical and Electronics Terms, ; The Institute of Electrical and Electronics Engineers.* Inc.: New York, 1984.

[117] McKay, M.D., Beckman, R.J., and Conover, W.J., *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code.* Technometrics, 1979. **21**(2): p. 239-245.

[118] Hasofer, A.M. and Lind, N.C., *Exact and invariant second-moment code format(for reliability analysis in multivariate problems).* American Society of Civil Engineers, Engineering Mechanics Division, Journal, 1974. **100**: p. 111-121.

[119] Hohenbichler, M., et al., *New light on first- and second-order reliability methods.* Structural Safety, 1987. **4**(4): p. 267-284.

[120] Xiangjun, D. and Jiang, T. *Reliability prediction based on degradation measure distribution and wavelet neural network*. in *Prognostics and System Health Management (PHM), 2012 IEEE Conference on*. 2012.

[121] Wu, H.L., Zhong, Y., and Yong-Hui, J. *An improved reliability prediction model based on BP network*. in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*. 2010.

[122] Karunanithi, N., Whitley, D., and Malaiya, Y.K., *Using neural networks in reliability prediction.* IEEE Software, 1992. **9**(4): p. 53-59.

[123] Gorguluarslan, R.M. and Choi, S.-K. *Predicting Reliability of Structural Systems Using Classification Method*. in *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2013. Portland, Oregon, USA.

[124] Qin, L.N. *Software reliability prediction model based on PSO and SVM*. in *Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on*. 2011.

[125] Wu, X., et al., *Failure and reliability prediction of engine systems using iterated nonlinear filters based state-space least square support vector machine method.* Optik - International Journal for Light and Electron Optics, 2016. **127**(3): p. 1491-1496.

[126] Xu, Z., Ji, Y., and Zhou, D., *Real-time Reliability Prediction for a Dynamic System Based on the Hidden Degradation Process Identification.* IEEE Transactions on Reliability, 2008. **57**(2): p. 230-242.

[127] Amin, A., Grunske, L., and Colman, A., *An approach to software reliability prediction based on time series modeling.* Journal of Systems and Software, 2013. **86**(7): p. 1923-1932.

[128] Peng, W. and Coit, D.W. *Reliability prediction based on degradation modeling for systems with multiple degradation measures*. in *Reliability and Maintainability, 2004 Annual Symposium - RAMS*. 2004.

[129] Zhao, W., Tao, T., and Zio, E., *System reliability prediction by support vector regression with analytic selection and genetic algorithm parameters selection.* Applied Soft Computing, 2015. **30**: p. 792-802.

[130] Chen, K.-Y., *Forecasting systems reliability based on support vector regression with genetic algorithms.* Reliability Engineering & System Safety, 2007. **92**(4): p. 423-432.

[131] Zhao, W., et al., *A dynamic particle filter-support vector regression method for reliability prediction.* Reliability Engineering & System Safety, 2013. **119**: p. 109-116.

[132] Bazu, M., et al. *Fuzzy-Logic Reliability Predictions in Microtechnologies*. in *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*. 2005.

[133] Haiyan, H., De-zi, Z., and Na, C., *Aeroengine Reliability Prediction Based on Fuzzy and Interval Number.* Procedia Engineering, 2015. **99**: p. 1284-1288.

[134] Chen, D., Yao, C., and Feng, Z., *Reliability Prediction Method of Hydraulic System by Fuzzy Theory.* IFAC Proceedings Volumes, 2013. **46**(5): p. 457-462.

[135] Yao, J.Y., et al. *Modeling system based on fuzzy dynamic Bayesian network for fault diagnosis and reliability prediction*. in *2015 Annual Reliability and Maintainability Symposium (RAMS)*. 2015.

[136] Yang, S., Lu, M., and Ge, L. *Bayesian Network Based Software Reliability Prediction by Dynamic Simulation*. in *Software Security and Reliability (SERE), 2013 IEEE 7th International Conference on*. 2013.

[137] Loon Ching, T. and Dong Shang, C., *Reliability prediction using nondestructive accelerated-degradation data: case study on power supplies.* IEEE Transactions on Reliability, 1995. **44**(4): p. 562-566.

[138] Wan, Y., et al., *Thermal reliability prediction and analysis for high-density electronic systems based on the Markov process.* Microelectronics Reliability, 2016. **56**: p. 182-188.

[139] Cortie, M.B. and Garrett, G.G., *On the correlation between the C and m in the paris equation for fatigue crack propagation.* Engineering Fracture Mechanics, 1988. **30**(1): p. 49-58.

[140] Benson, J.P. and Edmonds, D.V., *The relationship between the parameters C and m of Paris' law for fatigue crack growth in a low-alloy steel.* Scripta Metallurgica, 1978. **12**(7): p. 645-647.

[141] Bilir, Ö.G., *The relationship between the parameters C and n of Paris' law for fatigue crack growth in a SAE 1010 steel.* Engineering Fracture Mechanics, 1990. **36**(2): p. 361-364.

[142] Pattabhiraman, S., et al., *Skipping unnecessary structural airframe maintenance using an on-board structural health monitoring system.* Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2012. **226**(5): p. 549-560.

[143] Remery, M., Mascle, C., and Agard, B., *A new method for evaluating the best product end-of-life strategy during the early design phase.* Journal of Engineering Design, 2012. **23**(6): p. 419-441.

[144] Fajdiga, M., JurejevČIČ, T., and Kernc, J., *Reliability Prediction in Early Phases of Product Design.* Journal of Engineering Design, 1996. **7**(2): p. 107-128.

[145] Sriram, P., et al., *Synchronizing Condition-based Maintenance with Necessary Scheduled*

Maintenance, in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. 2012, American Institute of Aeronautics and Astronautics.

[146] Wang, Z.Q., et al., *A Prognostic-Information-Based Order-Replacement Policy for a Non-Repairable Critical System in Service.* IEEE Transactions on Reliability, 2015. **64**(2): p. 721-735.

[147] Langeron, Y., Grall, A., and Barros, A., *A modeling framework for deteriorating control system and predictive maintenance of actuators.* Reliability Engineering & System Safety, 2015. **140**: p. 22-36.

[148] Huynh, K.T., Barros, A., and Berenguer, C., *Multi-Level Decision-Making for The Predictive Maintenance of k-out-of-n: F deteriorating systems.* IEEE Transactions on Reliability, 2015. **64**(1): p. 94-117.

[149] Guan, X., Jha, R., and Liu, Y., *Model selection, updating, and averaging for probabilistic fatigue damage prognosis.* Structural Safety, 2011. **33**(3): p. 242-249.

[150] Scanlan, J., et al., *Cost modelling for aircraft design optimization.* Journal of Engineering Design, 2002. **13**(3): p. 261-269.

[151] Tammineni, S., et al., *A knowledge-based system for cost modelling of aircraft gas turbines.* Journal of Engineering Design, 2009. **20**(3): p. 289-305.

[152] Kundu, A.K., *Aircraft design*. Vol. 27. 2010: Cambridge University Press.

[153] Huang, X., Torgeir, M., and Cui, W., *An engineering model of fatigue crack growth under variable amplitude loading.* International Journal of Fatigue, 2008. **30**(1): p. 2-10.

[154] Wheeler, O.E., *Spectrum Loading and Crack Growth.* Journal of Basic Engineering, 1972. **94**(1): p. 181-186.

PUBLICATIONS

JOURNAL

[1] **Yiwei WANG**, Nicolas BINAUD, Christian GOGU, Christian BES. *Determination of Paris' law constants and crack length evolution via Extended and Unscented Kalman filter: An application to aircraft fuselage panels.* **Mechanical Systems and Signal Processing**, 2016. **80**: p. 262-281.

[2] **Yiwei WANG**, Nicolas BINAUD, Christian GOGU, Christian BES, Raphael T. Haftka, Nam H. KIM. *A cost driven predictive maintenance policy for structural airframe maintenance.* **Chinese journal of aeronautics, In press (International journal published by Elsevier)**.

[3] **Yiwei WANG**, Christian GOGU, Nicolas BINAUD, Christian BES. Coupling Extended Kalman Filter and a linearization-based approach for fatigue damage prognostic on fuselage panels, *Chinese journal of aeronautics* (under review)

[4] **Yiwei WANG**, Nam H. KIM, Raphael T. HAFTKA, Christian GOGU, Nicolas BINAUD, Christian BES. Noise-dependent ranking of prognostics algorithms based on discrepancy without true damage information. (submitted to Reliability Engineering and System Safety)

[5] **Yiwei WANG**, Christian GOGU, Nicolas BINAUD, Christian BES, Raphael T. HAFTKA, Nam H.Kim, Predictive airframe maintenance strategies using model-based prognostics. (submitted to Journal of Risk and Reliability)

CONFERENCE

[1] **Yiwei WANG**, Christian GOGU, Nicolas BINAUD, Christian BES. Predicting Remaining Useful Life by fusing SHM data based on Extended Kalman Filter 25th European Safety and Reliability Conference (ESREL2015), Zurich, Switzerland, 7-10th September 2015.

[2] **Yiwei WANG,** Christian GOGU, Nicolas BINAUD, Christian BES. Comparing structural airframe maintenance strategies based on probabilistic estimates of the remaining useful service life. *S12 Congrès de français mécanique mécaniques*, 2015.

[3] Lea D. COT, **Yiwei WANG,** Christian BES**,** Christian GOGU (2014). Scheduled and SHM Structural Airframe Maintenance Applications Using a New Probabilistic Model. In *EWSHM-7th European Workshop on Structural Health Monitoring*.