Thesis for the Degree of Master of Engineering

Model-based and Data-driven Approach
for Machine Prognostics
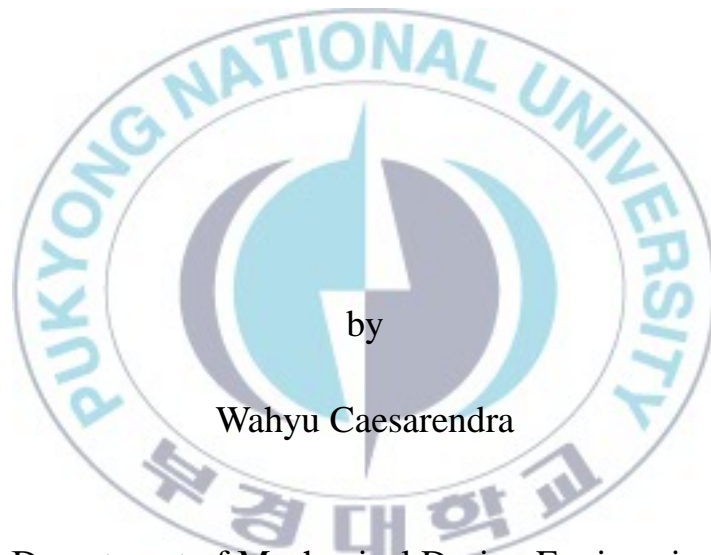
by

Wahyu Caesarendra

Department of Mechanical Design Engineering
The Graduate School

Pukyong National University

February 2010

Model-based and Data-driven Approach
for Machine Prognostics

기계 예지를 위한 모델 기반 및 데이터
주도 방법

Advisor: Prof. Bo-Suk Yang

by

Wahyu Caesarendra

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Engineering

in the Department of Mechanical Design Engineering
The Graduate School
Pukyong National University

February 2010

Model-based and Data-driven Approach
for Machine Prognostics

A thesis

by

Wahyu Caesarendra

Approved as to style and content by:

_____

(Chairman)   Prof. Seon-Jin Kim

_____   _____

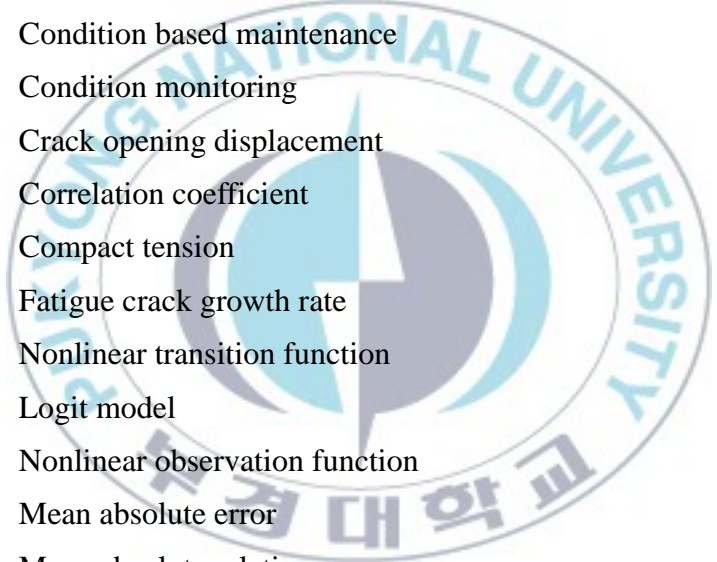(Member)   Prof. Byung-Tak Kim   (Member) Prof. Bo-Suk Yang

February 2010

Contents

List of Symbols

a	Crack length (mm)
BPFI	Ball pass frequency of the inner race (Hz)
BPFO	Ball pass frequency of the outer race (Hz)
BSF	Ball spin frequency (Hz)
C	Crack growth rate coefficient
CBM	Condition based maintenance
CM	Condition monitoring
COD	Crack opening displacement
CR	Correlation coefficient
CT	Compact tension
FCGR	Fatigue crack growth rate
$f(\cdot)$	Nonlinear transition function
$g(\vec{\mathbf{x}})$	Logit model
$h(\cdot)$	Nonlinear observation function
MAE	Mean absolute error
MARE	Mean absolute relative error
MLE	Maximum likelihood estimation
N	Number of cycles
n	Crack growth rate exponent
LR	Logistic regression
PDF	Probability density function
PF	Particle filter
$p(y\,|\,x)$	Data distribution
$p(x)$	Prior distribution

$p(y)$	Marginal distribution
$p(x\,|\,y)$	Posterior distribution
$P(\vec{x})$	Probability of some output event
RMSE	Root mean square error
RUL	Remaining useful life
RVM	Relevance vector machine
SIS	Sequential importance sampling
SIRs	Sequential importance sampling and resampling
VAE	Variance absolute error
VRE	Variance relative error
$da\,/\,dN$	Crack growth rate or crack propagation
ΔK	Stress intensity factor
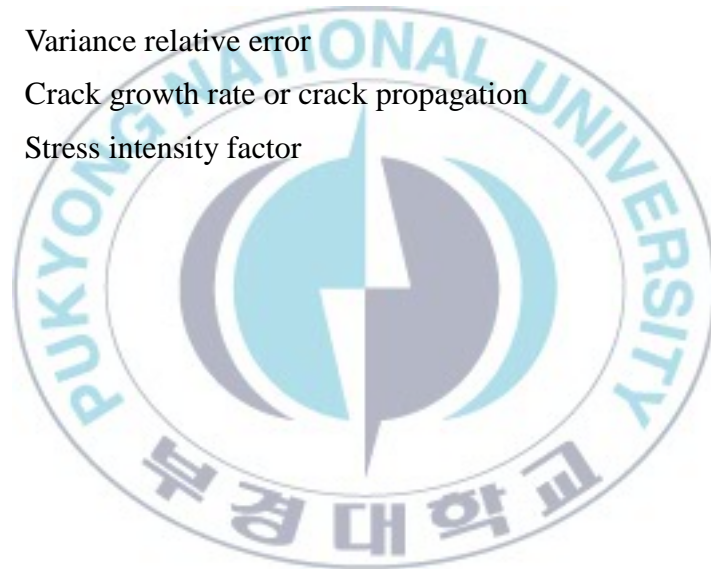
Model-based and Data-driven Approach
for Machine Prognostics

Wahyu Caesarendra

Department of Mechanical Engineering, The Graduate School,
Pukyong National University

Abstract

The severe competition in the market forced the industrial field to become cost-effective and reduce catastrophic failure. The maintenance activities including condition monitoring, fault diagnostic and prognostic are expected to overcome these issue. In maintenance work, the determination and prediction of final failure is the main objective. Therefore, prognostic part is become useful besides condition monitoring and fault diagnosis. Moreover, studies of prognostic approach are needed to be applied in industrial field. This paper proposes the application of particle filter (PF) method toward model-based prognosis approach and relevance vector machine (RVM) and logistic regression (LR) regarding data-driven prognosis approach. PF based sequential important sampling and resampling algorithm is employed to predict the trending data of low methane compressor and calculate the residual life of crack growth data of SPV50 steel. In terms of data-driven prognosis approach, RVM is combining with LR for failure degradation assessment. LR is used for calculate the failure degradation model of run-to-failure bearing data then RVM is employed to predict the final failure of individual bearing for case of simulated data and experimental data.

I. Introduction

1. Background of Maintenance Strategies

Today's maintenance and physical asset managers face great challenges to increase output, reduce equipment downtime, and lower costs. Therefore manage maintenance effectively is a key issue. High maintenance cost and unwanted downtime in industries is often caused by machine fault problems. Appropriate maintenance techniques are necessarily to overcome this problem. The prime objective of a maintenance technique is to keep machinery and plant equipments in good operating condition to prevent failure and production loss [1]. The evolutionary stages of maintenance technique start from corrective maintenance or breakdown maintenance. Corrective maintenance is a strategy where maintenance, in the form of repair work or replacement is only performed when machinery has failed. The next stage is schedule maintenance (also called planned maintenance or preventive maintenance), which applied to obtain a periodic interval maintenance based on the health conditions of a component or system. The third maintenance stage is condition based maintenance (CBM). CBM also known as predictive and proactive maintenance is a technique to assess the actual condition of machinery. CBM usually used to optimally schedule maintenance. CBM consist of continuously evaluating the conditions of monitored machines and identifying faults before catastrophic breakdown occurs. By organizing CBM, the cost benefit can be achieved. There are substantial factors to support CBM, machine fault diagnosis and machine prognosis. Machine fault diagnosis is

defined as a method for detecting, isolating and identifying a failure condition of a part or system, while the affected components are still operating even though they are in a degradation mode. Machine prognosis will be discussed in section 2. The strength and weaknesses of different maintenance strategies are shown in Fig. 1.1.

Philosophy: "Fix it when it breaks"

Corrective Maintenance

Weaknesses:
- High production downtime
- High costs of spare parts
- Overtime labor

Strength:
- Machines are not "over maintained"
- No condition monitoring related costs

Philosophy: "Fix it before it breaks"

Preventive Maintenance

Weaknesses:
- Machines are repaired when there are no faults
- Premature work

Strength:
- Reduced catastrophic breakdowns
- Greater control over stored and costs

Philosophy: "If it isn't broke, don't fix it"

Predictive Maintenance

Weaknesses:
- High investment costs
- Additional skill required
- Requires effective use of information

Strength:
- Unexpected breakdown is reduced
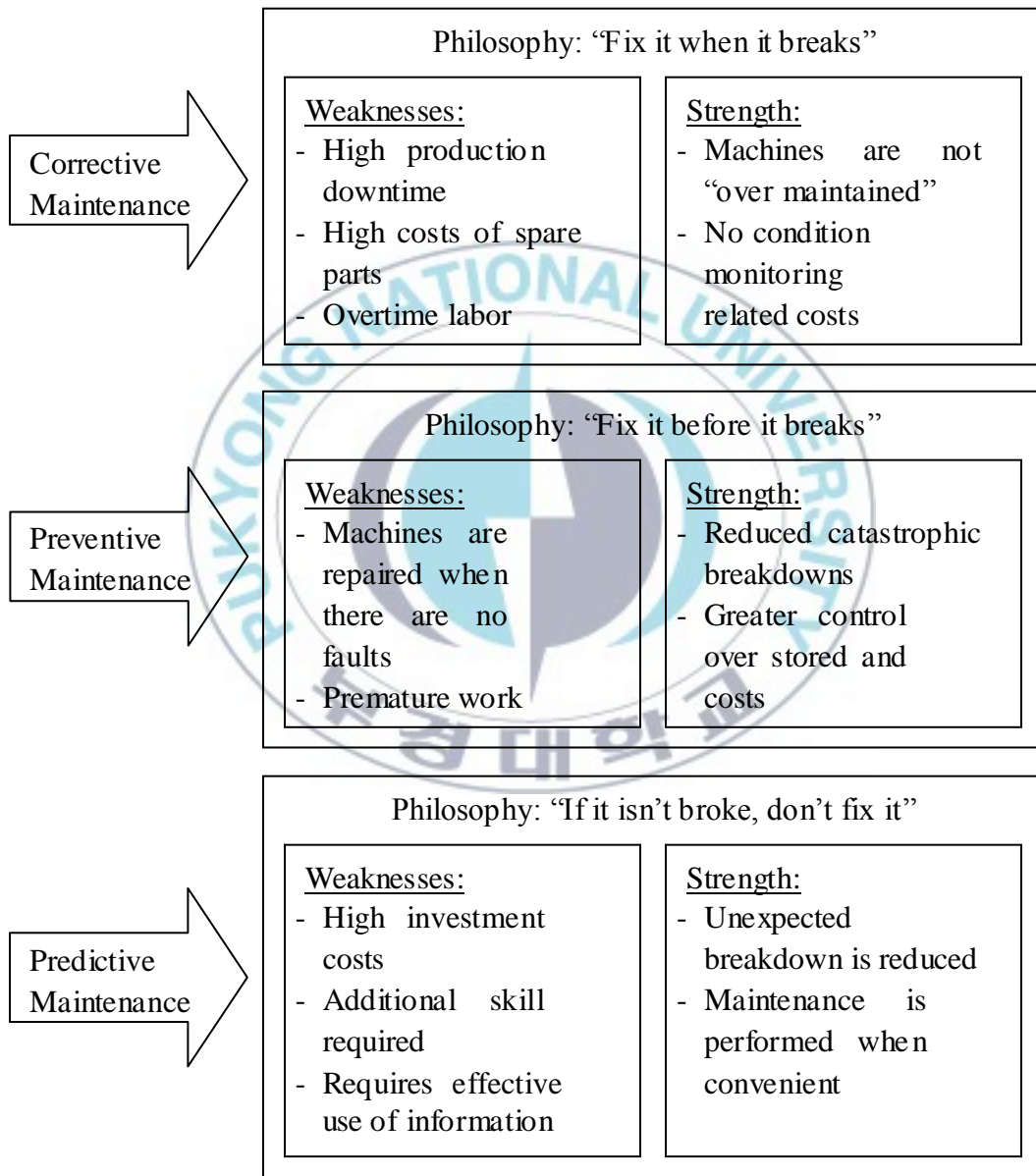- Maintenance is performed when convenient

Fig. 1.1 Strength and weaknesses of different maintenance strategies

2. Prognosis Definition and Classification

Prognosis is a composite word consisting of the Greek words *pro* and *gnosis*, literally translating into the ability to acquire knowledge (*gnosis*) about events before (*pro*) they actually occur. Prognosis (also called Prognostics) can be defined the ability to assess the current condition of part or system, observe the future condition of the part or system and predict the time left before catastrophic failure occurs. Generally, the lifetime of equipment can be divided into two stages. The first stage referred to a normal zone where no significant deviation from the normal operating state observed. The second stage is abnormal zone; this stage is initiated by potential failure and progressively develops into actual failure. In maintenance planning purposes, the prediction of the initiation point of the second stage is important [2]. The advanced prognosis is focused on performance degradation monitoring and assessment, so that failures can be predicted and prevented.
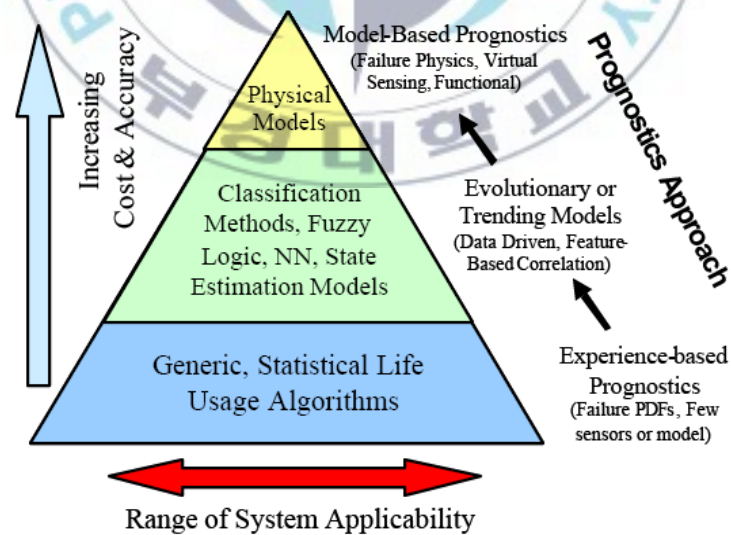
Fig. 1.2 Prognosis classification

To fulfill the goal of prognosis, two crucial steps are needed. (1) Failure prediction. (2) Remaining useful life (RUL) estimation. Failure prediction means, how to detect the initial failure occur and the failure progression. RUL refers to the time left before final failure occurs. The approaches of prognosis fall into three main categories; *statistical approaches*, *artificial intelligent approaches* and *model-based approaches* [3]. Fig. 1.2 re-summarizes the range of possible prognostics approaches as a function of the applicability to various systems and their relative implementation. Fig. 1.3 shows the application of prognostics.

Fig. 1.3 Forecasting application [4]

2.1 Statistical or Probability Approach

There are two main prediction types in machine prognostics. The most obvious and widely used prognostics is to predict how much tim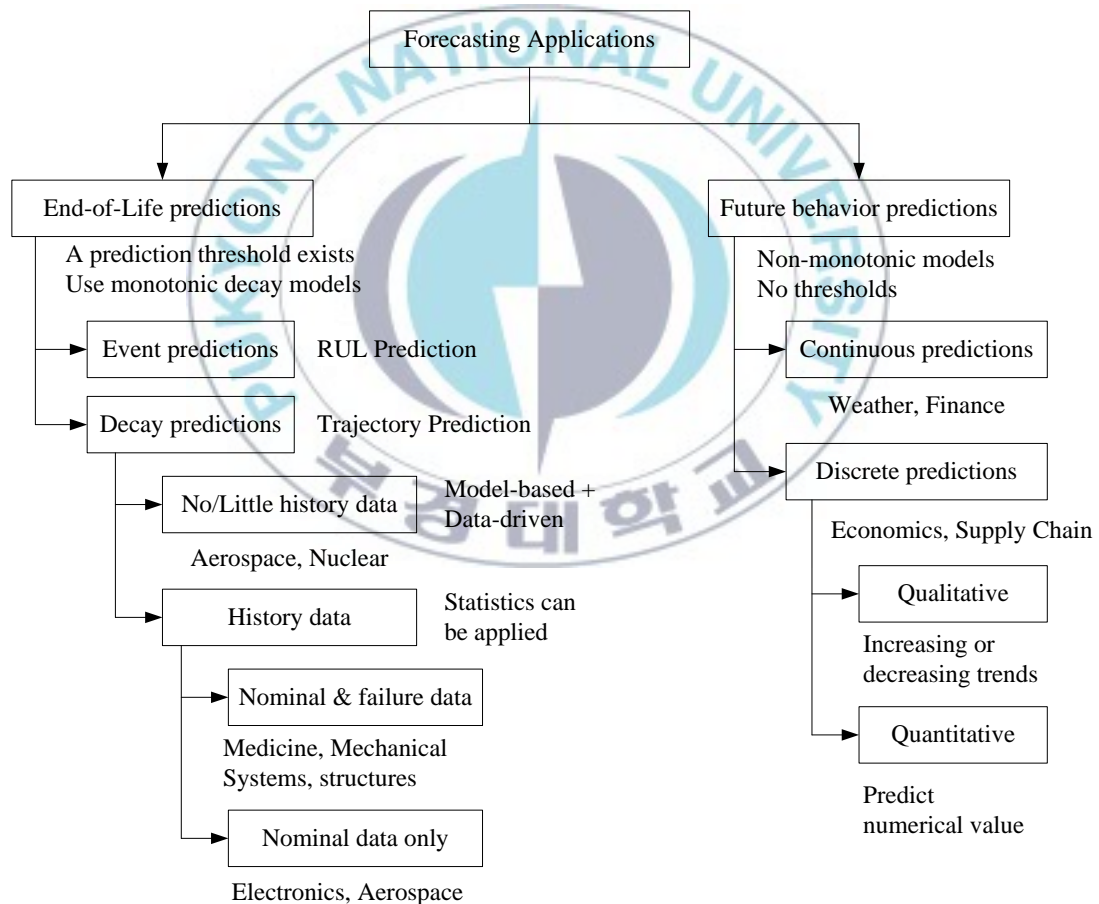e left before a failure occurs given the current machine condition and past operation profile which commonly called remaining useful life (RUL). In some situation, especially when a fault is catastrophic (e.g., nuclear power plant), it would be desirable to predict the chance that a machine operates without a fault or a failure up to some future time given the current machine condition and past operation profile [3]. This issue can addressed by statistical or probability prognosis approach. For example, ref. [5] separates two intervals of whole machine life: the I-P (Installation-Potential failure) interval in which the machine is running normally and the P-F (Potential failure-Functional failure) in which the machine condition has a problem. Based on two Weibull distributions assumed for the I-P and P-F intervals, failure prediction was derived in the two intervals and the RUL was estimated. Another example, ref. [6] proposed a model to predict the distribution of the residual life of water-pump using expert judgment based on the assumption that the hazard rate was a stochastic variable following a Gamma distribution. Ref. [7] applied proportional hazard model (PHM) to discuss the estimation of the residual life distribution.

2.2 Artificial Intelligence (AI) or Data-driven Approach

The data driven approaches are derived directly from routine condition monitoring (CM) data of monitored system (e.g. temperature, vibration, oil debris, currents, etc). The conventional data-driven methods include simple projection models, such as exponential smoothing and autoregressive model [8]. Data-driven techniques rely on comparative assessment of the status of a system under testing

with other known occurrences. They are based on the selected features that correlate with the failure progression and produce the desired output prediction of the time-to-failure (TTF) based on training process. In addition, for as long as the behavior of the system under testing remains similar to that of a previously known, healthy configuration, the former is deemed to be healthy. When the measured behavior deviates from this reference, a fault is detected, and a comparison with the conditions previously observed in analogous faulted systems can take place. Under the appropriate conditions, this new comparison has the potential to isolate and identify the fault efficiently. Thus, the ability of data-driven techniques to perform the task of diagnosis is given by the *training* of classification algorithm. A main drawback of data-driven approaches is their dependency on the quality of the operational data, and there is no physical understanding is usually provided. In data-driven approaches, the proper selection of a trending parameter is the key for successful prognosis of the remaining life. The selection criteria for such parameter should include the diagnosis ability, sensitivity, consistency and the amount of calculation required [9].

Ref. [10] has shown that a range of vibration parameters can be trended, such as: overall levels, specific frequency components, cepstrum amplitude and integrated frequency band amplitude. A number of time domain parameters were computed and trended throughout various tests including the peak value, RMS value, crest factor, and kurtosis.

2.3 Model-based or Physic-based Approach

The model-based methods assume that an accurate mathematical or physical model is available. Model based provide a technically comprehensive approach that has been used traditionally to understand component failure mode progression. One of the main advantages of the model-based approach is that it incorporates

physical understanding of the system being monitored. Physical model that often used in model-based approach is Paris law [11]. Fatigue crack growth in such typical machinery components as bearings, gears, shafts, and aircraft wings is affected by a variety of factors, including stress states, material properties, temperature, lubrication, and other environmental effects. Variations of available empirical and deterministic fatigue crack propagation models are based on Paris' formula:

$$\frac{da}{dN} = C\left(\Delta K\right)^{n} \tag{1.1}$$

where a is the instantaneous length of the dominant crack and N represents running cycles. The parameters C and n are regarded as material-dependent constants and related to factors such as material properties, environment, etc. The term ΔK represents the range of stress intensity over the loading cycle.

Table 1.1 Strength and weaknesses of three prognostic approaches

	Statistical or probability approach	Data-driven or artificial intelligent approach	Model-based or physic-based approach
Strength	Calculate the probability of machines health up to some future time	Have ability to transform high-dimensional data into lower dimensional information	Have ability to incorporate physical understanding of the system
Weaknesses	Requires a plenty quantity of previous data and information of historical data	Dependency on the quality of the operational data and there is no physical understanding of monitored system	Inflexibility; can only be applied to specific types of components

3. Prognosis and Health Management (PHM) Overview

Nowadays, prognosis research have been developed and create own research field namely Prognostic and Health Management (PHM). According to PHM, prognosis is the process of prediction the future reliability of a product by assessing the extent of deviation or degradation of a product from its expected normal operating conditions. While health monitoring is a process of measuring and recording the extent of deviation and degradation from a normal operating condition. Most products and systems contain significant content to provide needed functionality and performance. If one can assess the extent of deviation or degradation from an expected normal operating condition, this information can be used to meet several powerful goals, which include (1) providing advanced warning of failure; (2) minimizing unscheduled maintenance, extending maintenance cycles, and maintaining effectiveness through timely repair cost; (3) reducing the life-cycle cost of equipment by decreasing inspection costs, downtime, and inventory; and (4) improving qualification and assisting in the design and logical support of fielded and future system [12]. PHM includes diagnostic and health management solutions that identify problems before they become expensive to remedy, enabling guaranteed level and reducing life-cycle costs. In addition PHM is a method that permits the assessment of the reliability of product (or system) under its actual application conditions. When combined with physics-of-failure (PoF) models, it is thus possible to make continuously updated predictions based on the actual environmental and operational conditions. PHM techniques combine sensing, recording, interpretation of environmental, operational, and performance-related parameters to indicate a system's health. PHM can be implemented through the use of various techniques to sense and interpret the parameters indicative of:

- Performance degradation, such as deviation of operating parameters from

their expected values

- Physical degradation, such as material cracking, corrosion, interfacial delamination, or increase in electrical resistance or threshold voltage
- Changes in a life-cycle profile, such as usage duration and frequency, ambient temperature and humidity, vibration, and shock

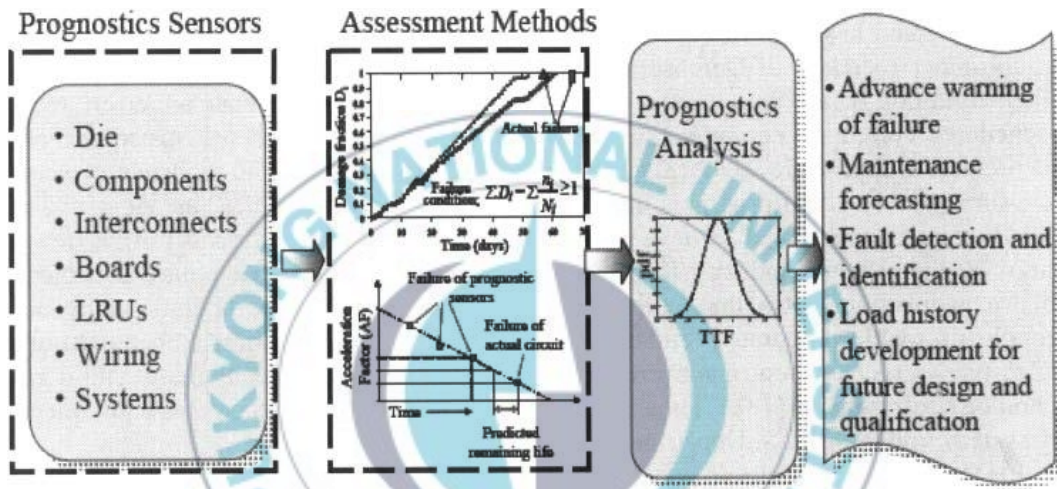The framework for prognostics and health management is shown in Fig. 1.4.

Fig. 1.4 Framework for prognostics and health management [13]

4. Objective of the Study

In this thesis, a study of model-based and data-driven approach was conducted to gather understanding and find appropriate methods for machine prognostics. Particles filter (PF) which also known sequential Monte Carlo method is proposed in model-based approach, while relevance vector machine (RVM) and logistic regression (LR) are employed for data-driven approach.

PF methods are still under significant development and the utilization of PF in real system is still infancy. PF offer the great advantage of not being subject to assumption of linearity, Gaussianity and stationary. Many researchers have been

introducing PF for time series prediction (14-18). Unfortunately, most of them utilized PF for simulation. It means the input or actual data are generated by artificial time series models. This problem motivated this study for using PF to predict the real trending data of low methane compressor. At the same time, the utilization of logistic regression in mechanical domain is still rare and the utilization of RVM for prognostic is still infancy. LR is frequently used in statistic science and biomedical research, e.g. [19, 20]. In this study, LR was used to build degradation model based randomly failure data then predict the final failure using RVM. RVM is considered as data-driven approach due to training step was first introduced by [21] is known has a good performance in classification. Therefore, this paper brings RVM into a prognostic area.

The goal of this thesis is to study the appropriate method implementing in model-based and data-driven approach. In the case of model-based approach, low methane compressor data acquired from condition monitoring (CM) routine was utilized for PF application in term of real data. Another application of PF is employed for fatigue crack growth rate data of SPV50 steel to estimate the residual life before final crack occurred. In the case of data-driven approach, LR is used to calculate failure degradation model based on failure probability data obtained from simulated bearing fault data and actual data from NASA data repository then predict the final failure using RVM. The result shows the potential method in machine prognostics.

5. Thesis Structure

The contents of this thesis are divided into six chapters. Except for the introduction chapter, the left chapters can be summarized as follows:

Chapter II introduces the necessary background knowledge of this research. Section 2.1 presents the Bayesian estimation which is fundamental knowledge for

particle filter method. Section 2.2 summarizes particle filter (PF) method based on sequential important sampling and resampling algorithm. Section 2.3 describes the relevance vector machine (RVM) algorithm for prognostics. Section 2.4 presents the logistic regression (LR).

Chapter III presents the application of particle filter method to predict the trending data (vibration data) of low methane compressor.

Chapter IV presents the application of particle filter method for fatigue crack growth data of SPV50 pressure vessel steel for determining the residual life before final crack occurred.

Chapter V describes the application of RVM and LR in terms of data-driven approach using simulated bearing fault data and experimental run-to-failure bearing data provided by NASA data repository site to predict the final failure of individual bearing.

Chapter VI contains two parts, section 6.1 summarizes the thesis and section 6.2 discusses future work of prognostics system.

6. Scientific Contribution

This thesis proposes model-based prognosis approach using PF based on sequential important sampling and resampling algorithm which is applied in real system. A propose method is employed to predict envelop and acceleration data of low methane compressor. PF method also employed for determining the residual time of crack propagation data of SPV50 pressure vessel steel before final failure occurred which is combine with finite element analysis. The second contribution is the combination of probability case and data-driven approach to predict the final failure of bearing fault. The application of LR is frequently used in medical science. In other hand, the utilization of RVM is still developing in prognosis field. LR is used initially for calculate the failure degradation model which contain of

bearing failure data and censored data. Then, RVM is utilized to predict the final failure of bearing fault based training process. The proposed method is evaluated using bearing simulated data and experimental data.

References

[1]   P. Jayaswal, A.K. Wadhwani, K.B. Mulchandani, Machine fault signature analysis, International Journal of Rotating Machinery (2008) 1-10.

[2]   W. Wang, A two-stage prognosis model in condition based maintenance, European Journal of Operational Research 182 (2007) 1177-1187.

[3]   A.K.S. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical System and Signal Processing 20 (2006) 1483-1510.

[4]   http://ti.arc.nasa.gov/m/event/industry-day/IDay_DataDrivenProg_poster.pdf

[5]   K.B. Goode, J. Moore, B.J. Roylance, Plant machinery working life prediction method utilizing reliability and condition-monitoring data, Proceeding of the Institution of Mechanical Engineering Part E – Journal of Process Mechanical Engineering 214 (2000) 109-122.

[6]   W. Wang, P.A. Scarf, M.A.J. Smith, On the application of a model of condition-based maintenance, Journal of the Operational Research Society 51 (2000) 1218-1227.

[7]   D. Banjevic, A.K.S. Jardine, Calculation of reliability function and remaining useful life for a Markov failure time process, IMA Journal of Management Mathematics 17(2) (2006) 115-130.

[8]   A. Heng, S. Zhang, A.C.C. Tan, J. Mathew, Rotating machinery prognostics: State of the art, challenges and opportunities, Mechanical Systems and Signal Processing, 23 (2009) 724-739.

[9] Y. Shao, K. Nezu, Prognosis of remaining bearing life using neural networks, Proceeding of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 214(3) (2000) 217-230.

[10] R.B. Randall, Computer aided vibration spectrum trend analysis for condition monitoring, Maintenance Management International 5 (1985) 161-167.

[11] P.C. Paris, F. Erdogan, A critical analysis of crack propagation laws, ASME Journal of Basic Engineering 85 (1963) 528-534.

[12] N. Vichare, M. Pecht, Prognostics and health management of electronics, IEEE Transactions on Components and Packaging Technologies 29(1) (2006) 222-229.

[13] M. Pecht, Prognostics and health management of electronics, John Wiley and Sons, 2008.

[14] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo in practice, Springer-Verlag, (2000a).

[15] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, Statistics and Computing 10 (2000b) 197-208.

[16] D. Guo, X. Wang, R. Chen, New sequential Monte Carlo methods for nonlinear dynamics systems, Statistics and Computing 15 (2005) 135-147.

[17] G. Kitagawa, Non-Gaussian state-space modeling of nonstationary time series, Journal of the American Statistical Association 82 (1987) 1032-1063.

[18] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, Journal of Computational and Graphical Statistics 5 (1996) 1-25.

[19] X.J. Xie, J. Pendergast, W. Clarke, Increasing the power: A practical approach to goodness-of-fit test for logistic regression models with continuous predictors, Computational Statistics & Data Analysis 52 (2008)

2703-2713.

[20] P.C. Austin, Absolute risk reductions, relative risks, relative risk reductions, and numbers needed to treat can be obtained from a logistic regression model, Journal of Clinical Epidemiology (in press).

[21] M.E. Tipping, The relevance vector machine, In S. Solla, T. Leen, K.R. Muller: Advances in Neural Information Processing System, 12, MIT Press, Cambridge, MA, 2000, pp. 287-289.

II. Background Knowledge

1. Bayesian Estimation

Due to the estimation of the future state based on sequential time seems to offer interesting potential for successful prognosis application. Here, Bayesian methods become the key point. Bayesian methods applied to combine a prior distribution of unknown state, based on available prior knowledge, and the likelihood of the current state. In particular, the observed data have to give for obtaining the conditional distribution of the unknown state. Regarding this purpose, the formality can accomplish by the application of Bayes' theorem [1]. For illustration, let X denote unknown state and Y our data. The full probability can be factored into components:

$$p(x, y) = p(y \mid x) p(x) = p(x \mid y) p(y)$$

Applying Bayes' rule, we obtain

$$p(x \mid y) = \frac{p(y \mid x) p(x)}{p(y)},$$

A few words used in order of notation. Random variables will be denoted by capital letters, and observed states will be denoted by lower case letters. In addition, we will use $p(\cdot)$ refer to a probability density function (PDF).

Data distribution, $p(y \mid x)$: Statisticians often refer to this as a sampling

15

distribution or measurement model. It is simply the distribution of the data, given the unobservable.

Prior distribution, $p(x)$: This distribution quantifies our a priori understanding of the unobservable quantities of interest.

Marginal distribution, $p(y) = \int p(y \mid x) p(x) \, dx$: We assume continuous *X* but note that there are analogous forms (sums) for discrete *X*. This distribution is also known as the prior predictive distribution. For the observations *Y*, $p(y)$ can be thought as the normalizing constant in *Bayes' rule*.

Posterior distribution, $p(x \mid y)$: This distribution of the unobservable given the data is our primary interest to be applied in the PF method. It is proportional to the product of the data model and the prior probability. The posterior is the update of our prior knowledge about *X* as summarized in $p(x)$ given the actual observations in *y*. In this sense, The Bayesian approach is inherently scientific in that it is analogous to the scientific method: one has prior belief (information), collects data, and then updates that belief given the new data (information).

2. Particle Filter

2.1 Background

Particle filtering (PF) is an emerging and powerful methodology for sequential signal processing with a wide range of applications in science and technology. It has captured the attention of many researchers in various communities, including those in signal processing, statistics, and econometrics.

Founded on the concept of sequential importance sampling (SIS) and the use of Bayesian theory, particle filtering is very suitable in the case when the system is

16

nonlinear or in the presence of non-Gaussian process/observation noise as in engines, gas turbine, gearboxes and the like, where the nonlinear nature and ambiguity of the rotating machinery world is significant when operating under fault conditions. Furthermore, particle filter allows information from multiple measurement sources to be fused in a principled manner, which is an attribute of decisive significance for fault detection/diagnostic purposes.

Many problems of nonlinear dynamic system can formulate in state-space model. Eqs. (2.1) and (2.2) show the state equation and the observation equation, respectively.

$$x_t = f(x_{t-1}) + u_t \quad \leftrightarrow \quad \underbrace{p(x_t \mid x_{t-1})}_{Transition\,Density} \tag{2.1}$$

$$y_t = h(x_t) + v_t \quad \leftrightarrow \quad \underbrace{q(y_t \mid x_t)}_{Observation\,Density} \tag{2.2}$$

where $f(\cdot)$ refer to the nonlinear state transition function, and $h(\cdot)$ is a nonlinear observation function. The state noise $u_t \sim N(0, \sigma_u^2)$ and observation noise $v_t \sim N(0, \sigma_v^2)$ is a white noise.

The state x_t depends only on the state value at the preceding time and noise. The observation y_t depends only on the current state x_t of the system and noise value. A transition equation describes the prior distribution of a hidden Markov process $\{x_t; t \in \mathbb{N}\}$, the so-called hidden state process, and an observation equation describes the likelihood of the observation $\{y_t; t \in \mathbb{N}\}$, with t being a discrete time index. Within a Bayesian framework, all relevant information about $\{x_0, x_1, ..., x_t\}$ given observation up to and including time t can be obtained from the posterior distribution $p(x_0, x_1, ..., x_t \mid y_0, y_1, ..., y_k)$. The condition is natural when the process x_t generated from the model in the increasing time order. Then

17

x_t is a homogenous Markov chain. The transition distribution $p(x_t | x_{t-1})$ referred to the conditional probability density of x_t given the posterior state $x_{0:t-1} = (x_0,...,x_{t-1})$ depends only on x_{t-1}. Hence, the conditional probability $q(y_t | x_t)$ is derived from the conditional probability density of y_t given the state x_{0-t} and the past observations $y_{1:t-1}$ depends only on x_t.
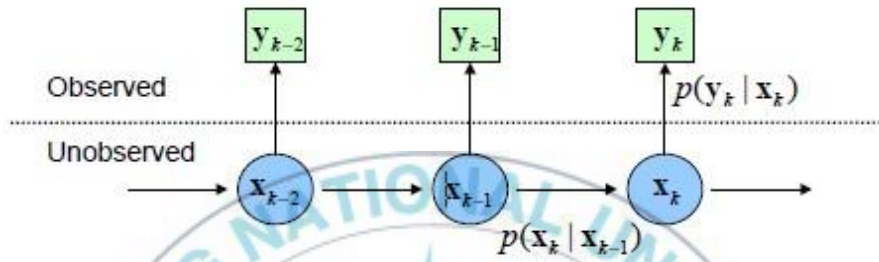

Fig. 2.1 Dynamic state space model

Monte Carlo methods for nonlinear filtering are based on sequential versions of the importance sampling paradigm, a technique that amounts to simulating samples under an instrumental distribution and then approximate the target distributions by weighting these samples using appropriately defined importance weight. In the nonlinear filtering context, importance sampling algorithm can be implemented sequentially. This algorithm is called sequential importance sampling, often abbreviated to SIS. Sequential importance sampling helps to reduce the number of samples required to approximate distributions with appropriate precision, being faster and more computationally efficient than classical Monte Carlo methods. Although the SIS algorithm has been known since the early 1970s, its use in nonlinear filtering problem was rather limited at the time [2]. As the number of iteration increases, the importance weights tend to degenerate, a phenomenon known as *weight degeneracy*. Basically, in the long run most of the samples have very small normalized importance weights and thus do

not significantly to the approximation of the target distribution. The solution is to allow rejuvenation of the set of samples by replicating the samples with high importance weights and removing samples with low weights. This method is called sequential importance sampling and resampling (SIRs) algorithm [3].

2.2 Sampling and Resampling: SIRs Particle Filter

A SIRs PF essentially consists of the following three computational steps:

Step 1 (Sampling): In this step, present a set of random drawn samples (particles) of unknown state are generated based on the given sampling function which provide an estimate of the current state of the system and also propagate the particles from the previous time step to the current time.

Step 2 (Weight calculation): Based on the observation, the importance weight are estimate from those particles.

Step 3 (Resampling): This step involves redrawing particles from the same probability density based on some function of the particle weights such as the weights of the particles are approximately equal. The SIRs algorithm is summarized in Fig. 2.2.

Importance sampling

for $i = 1, ..., N$ **do** ▷ Initialization

Sample $\tilde{x}_0^{(i)} \sim q_0(x_0 \mid y_0)$.

Assign initial importance weights

$$\tilde{w}_0^{(i)} = \frac{q(y_0 \mid \tilde{x}_0^{(i)}) \, \pi_0(\tilde{x}_0^{(i)})}{q_0(\tilde{x}_0^{(i)} \mid y_0)}.$$

end for

---

for $t = 1, ..., T$ **do**

    for $i = 1, ..., N$ **do**

        Propagate particles

$$\tilde{x}_t^{(i)} \sim q_t(\tilde{x}_t^{(i)} \mid \tilde{x}_{t-1}^{(i)}, y_t).$$

        Compute importance weight

$$\tilde{w}_t^{(i)} = w_{t-1}^{(i)} \frac{g(y_t \mid \tilde{x}_t^{(i)}) f(\tilde{x}_t^{(i)} \mid \tilde{x}_{t-1}^{(i)})}{q_t(\tilde{x}_t^{(i)} \mid \tilde{x}_{t-1}^{(i)}, y_t)}$$

    end for

    Normalize weights

$$w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^{N} \tilde{w}_t^{(j)}}, \qquad i = 1, ..., N$$

    Resampling

    If Resampling **then**

    Select N particle index $j_i \in \{1, ..., N\}$ according to weights

$$\{w_{t-1}^{(j)}\}_{1 \le j \le N}$$

    Set $x_{t-1}^{(i)} = \tilde{x}_{t-1}^{(j_i)}$, and $w_{t-1}^{(i)} = 1/N, \ i = 1, ..., N.$

    else

        Set $x_{t-1}^{(i)} = \tilde{x}_{t-1}^{(i)}, \ i = 1, ..., N.$

    end if

end for

Fig. 2.2 Algorithm for sequential importance sampling and resampling (SIRs)

The N trajectories $\tilde{x}_{0:t}^{(1)}, ..., \tilde{x}_{0:t}^{(N)}$ are independent and identically distributed. Following the terminology in use in the nonlinear filtering community, we shall refer to the sample at the time index t, $\tilde{x}_t^{(1)}, ..., \tilde{x}_t^{(N)}$, as the population (or system) of *particles* and to $\tilde{x}_{0:t}^{(i)}$ for a specific value of the particle index i as the history (or trajectory) of the ith particle.

One of the main difficulties that must be addressed in the implementation of SIS particle filters is the degeneracy problem in the particle population. The degeneracy phenomenon consists in the fact that, as the algorithm evolves in time, the weight variance increase [4] and the importance weight distribution becomes progressively more skewed, at the point where (after a few iterations) all but one particle will have a negligible weight [4], [5], [6]. Since the degeneracy in the particle population is directly related to the variance of the importance weights, several authors have proposed measuring it by using an estimate $\hat{N}_{eff}$ of the effective sample size N_{eff} [7],[8],[9].

$$N_{eff} = \frac{N}{1 + \mathrm{var}_{\pi(\cdot|y_{0:t})}(w_{0:t})}, \qquad \hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N}(w_t^{(i)})^2}$$

The basic resampling method is calculating the best prediction result from the current population of particles using the normalized weights as probabilities of selection. The trajectories with small importance weights are eliminated, whereas those with large importance weights are replicated. After resampling, the normalized importance weights are reset to 1/N. The side effect of a resampling technique is that, theoretically, the simulated paths are no longer statistically independent. Although this condition may suggest that convergence results for the SIS algorithm are no longer valid, [6] and [10] have already established a solid

theoretical foundation that proves convergence even in the SIR case. In addition to the algorithm presented in Fig. 2.2, there are other versions intended to perform particle filter resampling, including multinomial resampling [11], residual resampling [12, 13], and minimum variance resampling [8].

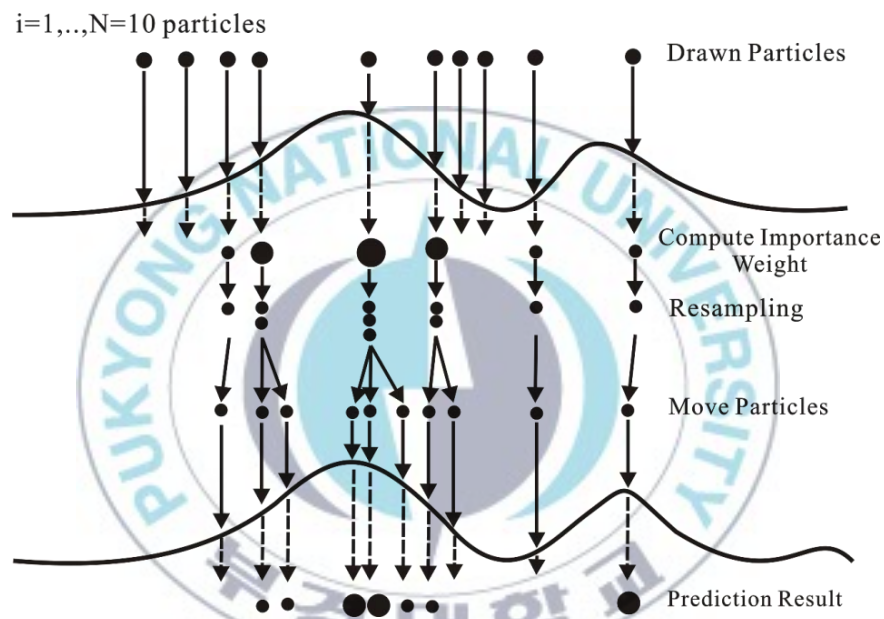Fig. 2.3 shows the visualization of SIRs algorithm.

Fig. 2.3 Visualization of sequential importance sampling and resampling (SIRs).

3. Relevance Vector Machine (RVM)

RVM is a Bayesian form representing a generalized linear model of identical functional form of support vector machine (SVM). It differs with SVM in the case of solution which provides probabilistic interpretation of its outputs [14]. RVM evades the complexity by producing models which have both a structure and a

parameterization process that, together, are appropriate to the information content of the data. As a supervised learning, RVM starts with a set of data input $\{\mathbf{x}\}_{n=1}^{N}$ and their corresponding target vector $\{\mathbf{t}\}_{n=1}^{N}$. The aim is to learn a model of the dependency of the target vectors on the inputs in order to make accurate prediction of $\mathbf{t}$ for unseen value of $\mathbf{x}$. Typically, the predictions are based on a function $y(\mathbf{x})$ defined over the input space, and learning the process of inferring the parameter of this function. In the context of SVM, this function takes form

$$y(\mathbf{x}) = \sum_{1}^{N} \mathbf{w}_i K(\mathbf{x}, \mathbf{x}_i) + w_0 \tag{2.3}$$

where $\mathbf{w} = \{w_1, w_2, ..., w_N\}$ is weight vectors, w_0 is bias and $K(\mathbf{x}, \mathbf{x}_i)$ is a kernel function.

RVM seeks to predict target $\mathbf{t}$ for any query of $\mathbf{x}$ according to

$$\mathbf{t} = y(\mathbf{x}) + \varepsilon_n \tag{2.4}$$

where ε_n are independents samples from noise process with mean 0 and variance σ^2.

The likelihood of data set can be written as

$$p(\mathbf{t} \mid \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{N-2} \exp\left\{-\frac{1}{2\sigma^2} \| \mathbf{t} - \mathbf{\Phi}\mathbf{w} \|^2 \right\} \tag{2.5}$$

where $\mathbf{\Phi}$ is the $N \times (N+1)$ design matrix with $\mathbf{\Phi}_{nm} = \{1,\ \mathrm{K}(\mathbf{x}_i, \mathbf{x}_1),\ \mathrm{K}(\mathbf{x}_i, \mathbf{x}_2), ...,\ \mathrm{K}(\mathbf{x}_i, \mathbf{x}_N)\}^{\mathrm{T}}$

Maximum likelihood estimation of $\mathbf{w}$ and σ^2 in Eq. (2.5) often results overfitting. Therefore, [15] recommended imposition of some prior constrains on the parameters $\mathbf{w}$ by adding a complexity to the likelihood or error function. This a priori information controls the generalization ability of the learning process. Typically, new higher-level parameters are used to constrain an explicit zero-mean Gaussian prior probability distribution over the weights

$$p\left(\mathbf{w} \mid \boldsymbol{\alpha}\right) = \prod_{i=0}^{N} N\left(w_i \mid 0, \alpha_i^{-1}\right) \qquad (2.6)$$

where $\boldsymbol{\alpha}$ is a vector of $\left(N \times 1\right)$ hyperparameters that controls how far from zero each weights is allowed to deviate [16].

Using Bayes' rule, the posterior overall unknowns could be computed, given the defined non-informative prior-distributions

$$p\left(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 \mid \mathbf{t}\right) = \frac{p\left(\mathbf{t} \mid \mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right) p\left(\mathbf{w}, \boldsymbol{\alpha}, \sigma\right)}{\int p\left(\mathbf{t} \mid \mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right) p\left(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right) d\mathbf{w} d\boldsymbol{\alpha} d\sigma^2} \qquad (2.7)$$

However, we cannot compute the solution of the posterior in Eq. (2.7) directly since we cannot perform the normalizing integral $p\left(\mathbf{t}\right) = \int p\left(\mathbf{t} \mid \mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right) p\left(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right) d\mathbf{w} \, d\boldsymbol{\alpha} \, d\sigma^2$. Instead, we decompose the posterior as

$$p\left(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 \mid \mathbf{t}\right) = p\left(\mathbf{w} \mid \mathbf{t}, \boldsymbol{\alpha}, \sigma^2\right) p\left(\boldsymbol{\alpha}, \sigma^2 \mid \mathbf{t}\right) \qquad (2.8)$$

to facilitate the solution. The posterior distribution of weights is given by

$$p\left(\mathbf{w}\,|\,\mathbf{t},\alpha,\sigma^2\right)=\frac{p\left(\mathbf{t}\,|\,\mathbf{w},\sigma^2\right)p\left(\mathbf{w},\alpha\right)}{p\left(\mathbf{t}\,|\,\alpha,\sigma^2\right)} \tag{2.9}$$

Eq. (2.9) has an analytical solution where the posterior covariance and mean are

$$\Sigma=\left(\mathbf{\Phi}^T\mathbf{B}\mathbf{\Phi}+\mathbf{A}\right)^{-1} \tag{2.10}$$

$$\mu=\Sigma\mathbf{\Phi}^T\mathbf{B}\mathbf{t} \tag{2.11}$$

with $\mathbf{A}=diag\left(\alpha_1,...,\alpha_{N+1}\right)$ and $\mathbf{B}=\sigma^{-2}\mathbf{I}$

Note that σ^2 is also treated as a hyperparameter, which may be estimated from the data. Therefore, machine learning becomes a search for the hyperparameter posterior most probable. Predictions for a new data are then made according to integration out the weights to obtain the marginal likelihood for the hyperparameter

$$p\left(\mathbf{t}\,|\,\alpha,\sigma^2\right)=\int p\left(\mathbf{t}\,|\,\mathbf{w},\sigma^2\right)p\left(\mathbf{w}\,|\,\alpha\right)\,d\mathbf{w}$$

$$\tag{2.12}$$

$$=\left(2\pi\right)^{N/2}\,|\,\mathbf{B}^{-1}+\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T\,|^{-1/2}\,\exp\left\{-\frac{1}{2}\mathbf{t}^T\left(\mathbf{B}^{-1}+\mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T\right)^{-1}\mathbf{t}\right\}$$

4. Logistic Regression (LR)

Logistic regression is a variation of ordinary regression method which is used when the dependent variable is a dichotomous variable (which is usually represented the occurrence or non-occurrence of some output event, usually coded as 0 and 1). The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of the dependent variable and a set of independent variables [17]. Here, logistic regression is used to obtain the failure probability between incipient failure and final failure. Then the probability result is regards as a target vector input of RVM.

Logistic regression can also be used to predict a dependent variable on the basis of continuous and/or categorical independents and to determine the percent of variance in the dependent variable explained by the independents; to rank the relative importance of independents; to asses interaction effects; and to understand the impact of covariate control variable. The impact of predictor variables is usually explained in terms of odds ratios. In the logistic regression method, the dependent variable is the probability that an event will occur; hence the output value has a discrete number of responses which is constrained between 0 (normal) and 1 (failure). The logistic function is:

$$\text{Prob(event)} = P(\vec{x}) = \frac{1}{1 + e^{-g(\vec{x})}} = \frac{e^{g(\vec{x})}}{1 + e^{g(\vec{x})}} \qquad (2.13)$$

where $P(\vec{x})$ the probability of some output event, $\vec{x}(x_1, x_2, ..., x_k)$ is an input vector, corresponding to the independent variables (predictors), and $g(\vec{x})$ is the

26

logit model. The logit model of multiple logistic regressions can be written as:

$$g\left(\vec{\mathbf{x}}\right) = \log\left(\frac{P\left(\vec{\mathbf{x}}\right)}{1 - P\left(\vec{\mathbf{x}}\right)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \tag{2.14}$$

where $g\left(\vec{\mathbf{x}}\right)$ is a linear combination of the independent variables $x_1, x_2, ..., x_k$ and $\alpha, \beta_1, \beta_2, ..., \beta_k$ are known as the regression coefficient. Logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable in order to calculate the parameter $\alpha, \beta_1, \beta_2, ..., \beta_k$.

References

[1] C.K. Wikle, L.M. Berliner, A Bayesian tutorial for data assimilation, *Physica* D 230 (2007) 1-16.

[2] O. Cappe, S.J. Godsill, E. Moulines, An overview of existing methods and recent advances in sequential Monte Carlo, In *Proceeding of The IEEE* 95 (2007) 899-924.

[3] D.B. Rubin, A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when the fraction of missing information is modest: The SIR algorithm (discussion of Tanner and Wong), Journal of the American Statistical Association 82 (1987) 543-546.

[4] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filering, Statistics and Computing 10 (2000) 197-208.

[5] M.S. Arumpalan, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, IEEE Transactions on Signal Processing 50 (2002) 174-188.

[6] G. Casella, C.P. Roberts, Rao-blackwellisation of sampling schemes, Biometrika 83 (1) (1996) 81-94.

[7] V. Verman, G. Gordon, R. Simmons, S. Thrun, Particle filters for rover fault diagnosis, IEEE Robotics & Automation Magazine (2004) 56-64.

[8] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, Journal of Computational and Graphical Statistics 5 (1) (1996) 1-25.

[9] S. Godsill, T. Clapp, Improvement strategies for Monte Carlo particle filters, in Sequential Monte Carlo Methods in Practice, A. Doucet, N. de Freitas, N. Gordon, Eds. New York: Springer-Verlag, 2001.

[10] D. Crisan, Particle filters – A theoretical perspectives, in Sequential Monte Carlo Methods in Practice, A. Doucet, N. de Freitas, N. Gordon, Eds. New York: Springer-Verlag, 2001.

[11] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceeding-F* 140 (2) (1993) 107-113.

[12] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamic systems, Journal of the American Statistical Association 93 (1998) 1032-1044.

[13] J.S. Liu, R. Chen, W.H. Wong, Rejection control and sequential importance sampling, Journal of the American Statistical Association 93 (1998) 1022-1031.

[14] M.E. Tipping, The relevance vector machine, In S. Solla, T. Leen, K.R. Muller: Advances in Neural Information Processing System, 12, MIT Press, Cambridge, MA, 2000, pp. 287-289.

[15] M.E. Tipping, Spare Bayesian learning and the relevance vector machine, Journal of Machine Learning Research 1 (2001) 211-244.

[16] B. Scholkopf, A.J. Smola, Learning with kernels: Support vector machines, Regularization, Optimization, and beyond, Cambridge, MA, MIT Press, 2002.

[17] J. Yan, M. Koc, J. Lee, A prognostic algorithm for machine performance assessment and its application, Production Planning and Control 15(8) (2004) 796-801.

Ⅲ. PF Application for Vibration Data

1. Industry Data Acquisition

Methane compressors are momentous equipment used in petrochemical plants in which normal condition are required to maintain performance. Therefore, condition monitoring and prognosis are needed in the petrochemical industry. The vibration data was acquired from Samsung Heavy Industry Company. The particle filter method is applied in real system to predict the trending data of low methane compressors as shown in Fig. 3.1. This compressor is driven by a 440 kW, 6600 Volt motor with 2 poles operating at a speed of 3563 rpm. The related information of the system summarized in Table 3.1.
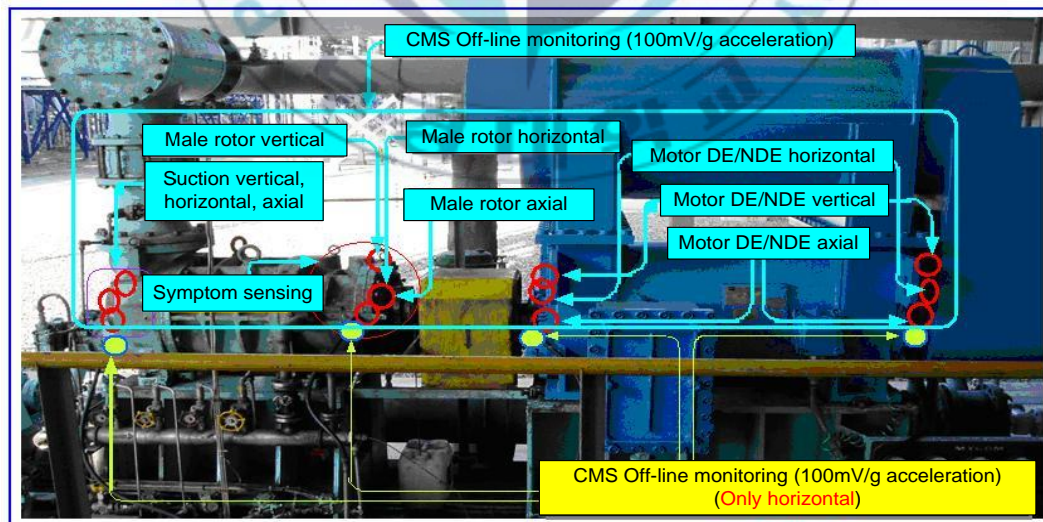
Fig. 3.1 Low methane compressor: wet screw type

Table 3.1 Description of system

Induction motor		Compressor	
Voltage	6600 V	Type	Wet screw (unload system)
Power	440 kW	Lobe	Male rotor (4 lobes)
Pole	2 Pole		Female rotor (6 lobes)
Bearing	NDE/6216, DE/6216	Bearing	Thrust: 7321 BDB
RPM	3565 rpm		Radial: Sleeve type

The system consists of two types of condition monitoring, namely off-line and on-line systems. In the off-line system, several vibration sensors installed on selected locations of the motor and compressor, such as drive-end (DE) motor, non drive-end (NDE) motor, the male rotor compressor and the suction part of the compressor. Each location consists of three directions of measurement: axial, vertical and horizontal. The circle shows the male component of the rotor compressor and the normal sensing location of the system. The on-line monitoring system consists of an acceleration sensor located horizontally at four locations: DE and NDE motor, male rotor compressor and the suction part of the compressor. Due to the importance of the methane compressor in petrochemical industry, it is imperative to carry out long term condition monitoring and to predict the future trend of performance degradation. The trending data of peak acceleration and envelope acceleration from compressor coupling sides has been acquired using data acquisition software installed in condition monitoring systems. Peak acceleration data are original trending data of peak-to-peak acceleration acquired from the machine; whereas envelope acceleration data are trending data of peak acceleration after the enveloping process. The trending data recorded from August 2005 to November 2005. The sampling period of the originally measured signals was 6 h with four time measurements a day. The total time-series data consisted of 400 measurements, which involved a process of performance degradation from normal to abnormal running conditions.

2. Proposed Method for Machine Condition Prognosis

The increase vibration levels in the associated machine can indicate the degradation condition of a machine. In order to predict the degradation trend based on the vibration indicator in rotating machinery, a PF algorithm is used. The flowchart of propose algorithm shown in Fig. 3.2.
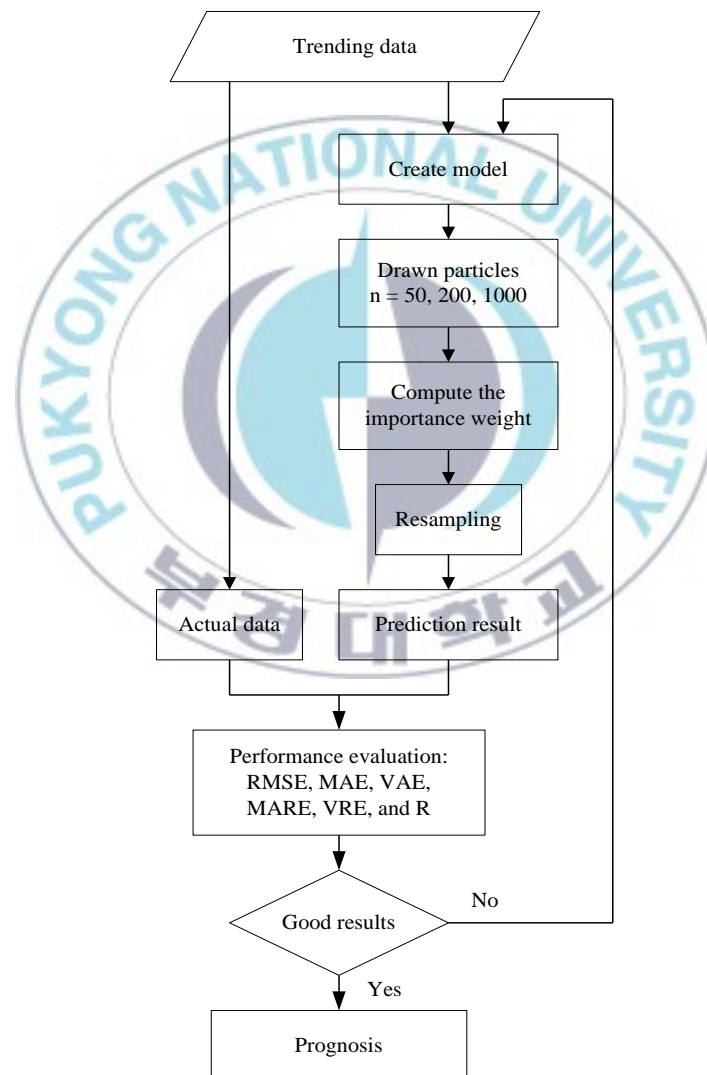

Fig. 3.2 Flowchart of particle filter algorithm for machine condition prognosis

The proposed algorithm for machine condition prognosis consists of the following steps:

Step 1 (Data acquisition): The used data is the trending information of system based on vibration signal, which contain histories of machine condition from the normal running process until faults occurred.

Step 2 (Prediction): The one-step ahead prediction of PF algorithm is used to predict the evolution by the steps of sampling method, computing importance weight and resampling.

Step 3 (Evaluation): To validate the proposed method, several performance evaluations are considered.

3. Performance Evaluation

In order to evaluate the predicting performance, absolute error and relative error criteria are used [1]. Absolute error is the amount of physical error in a prediction, while relative error gives an indication of how good a prediction is relative to the size of the parameter being predicted. Root mean square error (RMSE), mean absolute error (MAE), and variance absolute error (VAE) is utilized for absolute error. For relative error, mean absolute relative error (MARE) and variance relative error (VRE) are employed. In addition, correlation coefficient (CR) was also considered. The above statistic formulas list as follows:

Root mean square error:
$$RSME = \frac{1}{n}\sqrt{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad (3.1)$$

Mean absolute error:
$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \qquad (3.2)$$

Variance absolute error: $\qquad VAE = \dfrac{1}{n}\sum_{t=1}^{n}(\mid y_t - \hat{y}_t \mid - MAE)^2 \qquad$ (3.3)

Mean absolute relative error: $\quad MARE = \dfrac{1}{n}\sum_{t=1}^{n}\left|\dfrac{y_t - \hat{y}_t}{y_t}\right| \qquad$ (3.4)

Variance relative error: $\qquad VRE = \dfrac{1}{n}\sum_{t=1}^{n}\left(\left|\dfrac{y_t - \hat{y}_t}{y_t}\right| - E_{MAR}\right)^2 \qquad$ (3.5)

Correlation coefficient: $\qquad CR = \dfrac{\sum_{t=1}^{n}(y_t - \bar{y}_t)(\hat{y}_t - \bar{\hat{y}}_t)}{\sqrt{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2}\sqrt{\sum_{t=1}^{n}(\hat{y}_t - \bar{\hat{y}}_t)^2}} \qquad$ (3.6)

The smaller value of each error criteria indicates a higher accuracy of prediction. Meanwhile, a high correlation coefficient indicates a good accordance between the predicted values and actual values. The actual value of a time-series signal is expressed as y_t, while the predicted value is denoted as $\hat{y}_t$. In addition, $\bar{y}_t$ and $\bar{\hat{y}}_t$ means the average results of actual time series and predicted time series respectively. The symbol of n denotes the number of predictable data.

4. Results and Discussion

The trending data of peak acceleration plot in root mean square (RMS) values shown in Fig. 3.3. RMS is a common feature and the indicator used for machine condition monitoring. Fig. 3.4 shows the trending data of envelope acceleration. When detecting narrow pulse signals or displaying modulating signals, the envelope mode is useful. As mentioned above, the acquired data consists of 400 points of measurement that represents the machine condition (vibration amplitude)

over the time sequence.
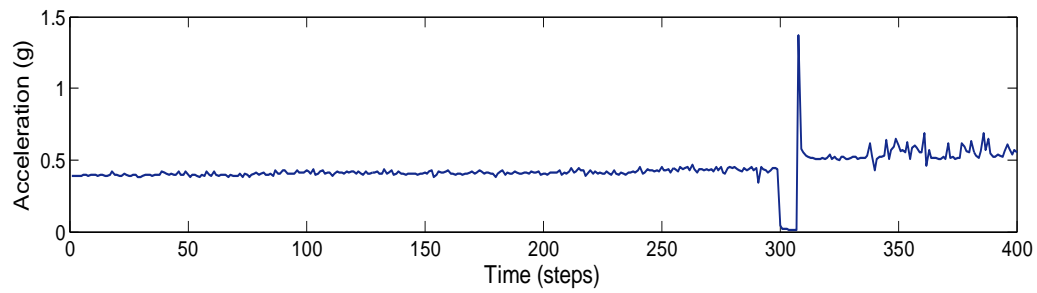
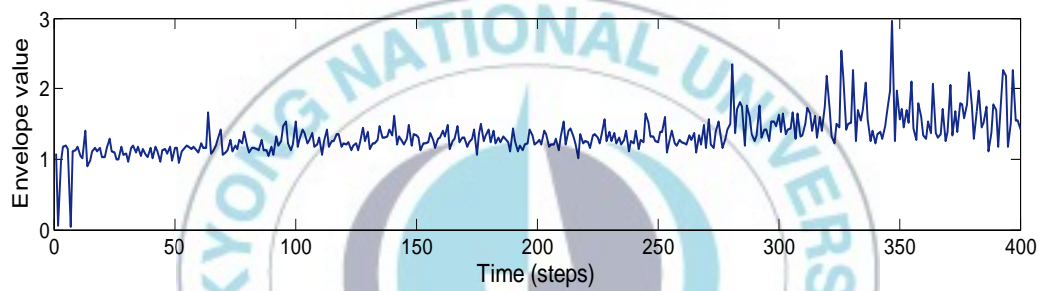Fig. 3.3 Peak acceleration data of low methane compressor plot in RMS value

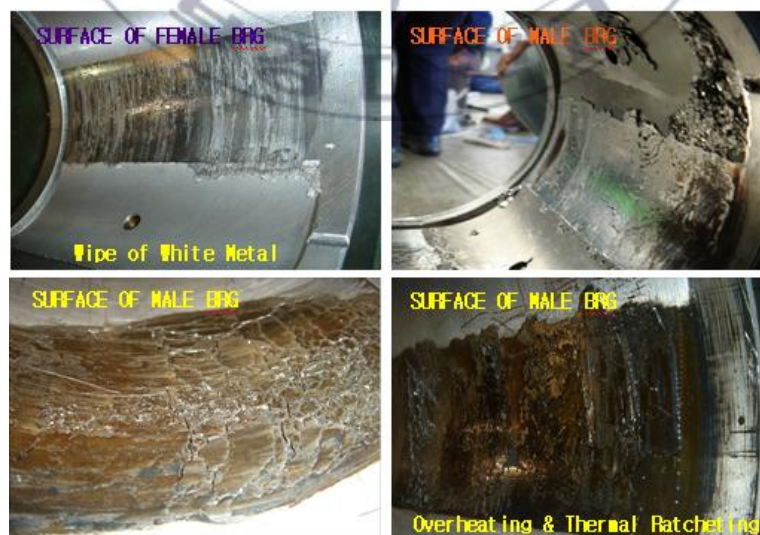Fig. 3.4 Trending data of envelope acceleration

Fig. 3.5 The faults of main bearings of compressor

35

The condition of the machine is shown in Fig. 3.3. It can be seen that at the beginning during the first 300 points of measurement, machine condition are normal, where the peak acceleration tends to be steady. Over points 300, amplitude of machine vibration increased drastically. It indicates that there are some faults occurring in this machine. These faults were the damages of main bearings of compressor (notation Thrust: 7321 BDB) due to the insufficient in Fig. 3.5.

Due to the nonlinearity and nonstationary of the real data, the propose method directly incorporate the actual data as the prior state. Hence, the posterior is defined as the following equation:

$$y(t) = x(t) + 0.01\cos(0.2t) + v_t \tag{3.7}$$

Where t is time sequence, $x(t)$ is actual data (prior state) and $v(t)$ is a white noise.

Regarding the definition of PF, number of particles sets is given to represent the actual data of $x(t)$. Thus, Eq. (3.7) becomes:

$$y(t) = xparticle(t,n) + 0.01\cos(0.2t) + v_t \tag{3.8}$$

Where n is a number of particles sets and $v_t \sim N(0,10^2)$

The proposed method used for predicting the future conditions of vibration amplitude (posterior state) based on actual data (prior state). In the present study, we use one-step-ahead prediction operations to generate number of particle over the time sequence and calculate the importance weight of the number of particles in discrete time. To obtain normalize constant, the weight normalization is calculated as follows:

$$w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^{N} \tilde{w}_t^{(j)}}, \qquad i = 1,...,N \tag{3.9}$$

Doucet et al. [2] introduce different types of resampling algorithm: residual

resampling, deterministic resampling and multinomial resampling. Each resampling algorithm has different degree of result as shown in Figs. 3.6 and 3.7. The resampling algorithm which has minimum RMSE value is utilized in this work.
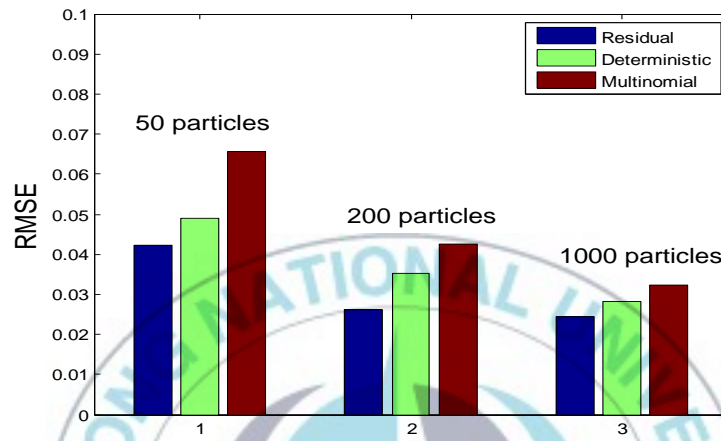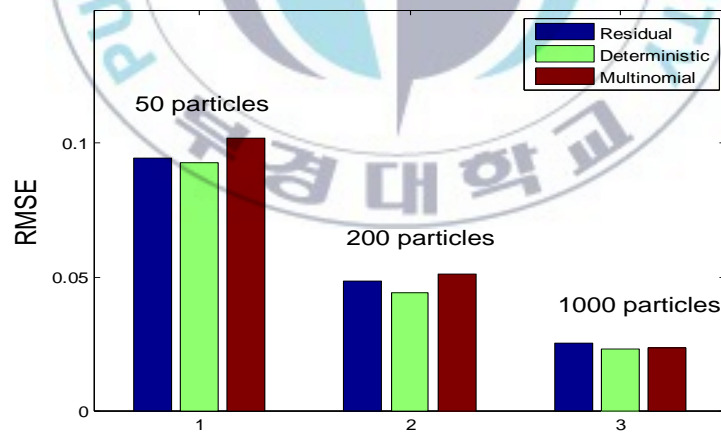

Fig. 3.6 Resampling of RMS data

Fig. 3.7 Resampling of envelope data

For RMS data, residual resampling has a minimum error shows in Fig. 3.6. While, deterministic resampling is appropriate for envelope data shows in Fig. 3.7.
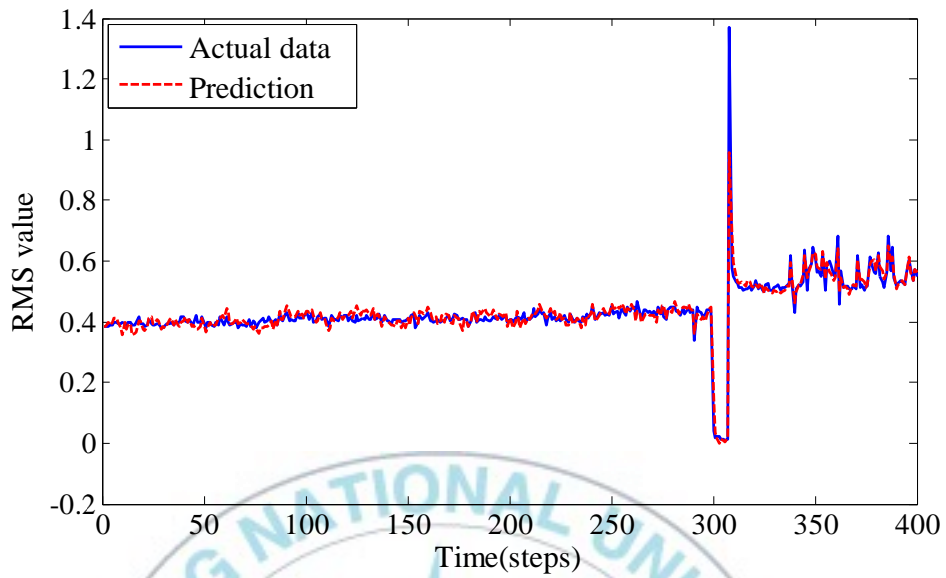
For the aim of prediction, the proposed method employed the peak acceleration data and envelope acceleration data as an actual data or prior state and the output is the prediction result or posterior state.
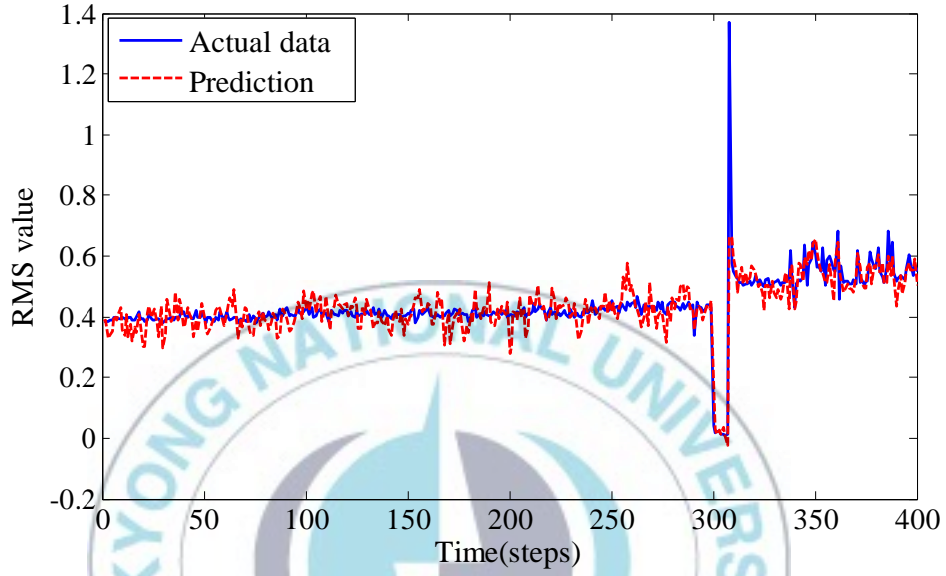
(a) 50 number of particles
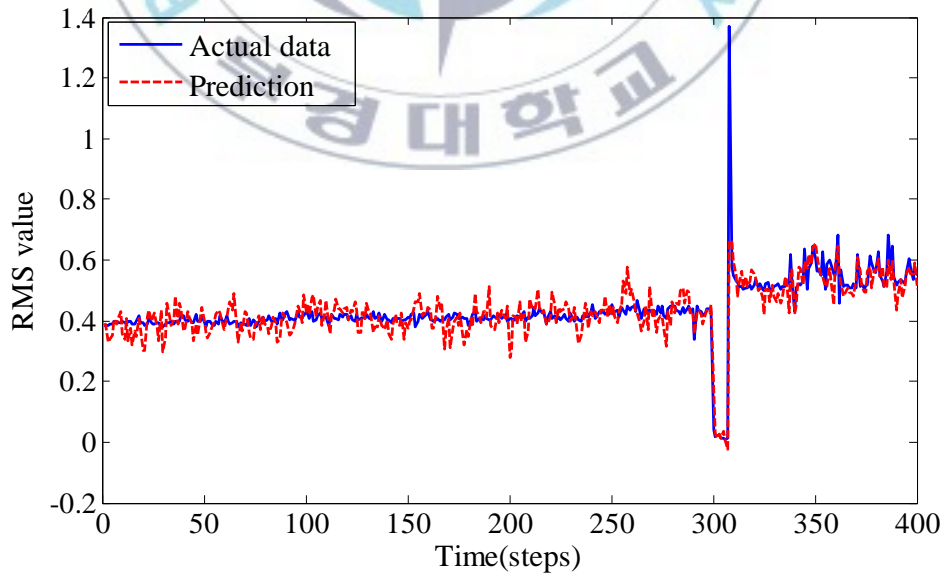
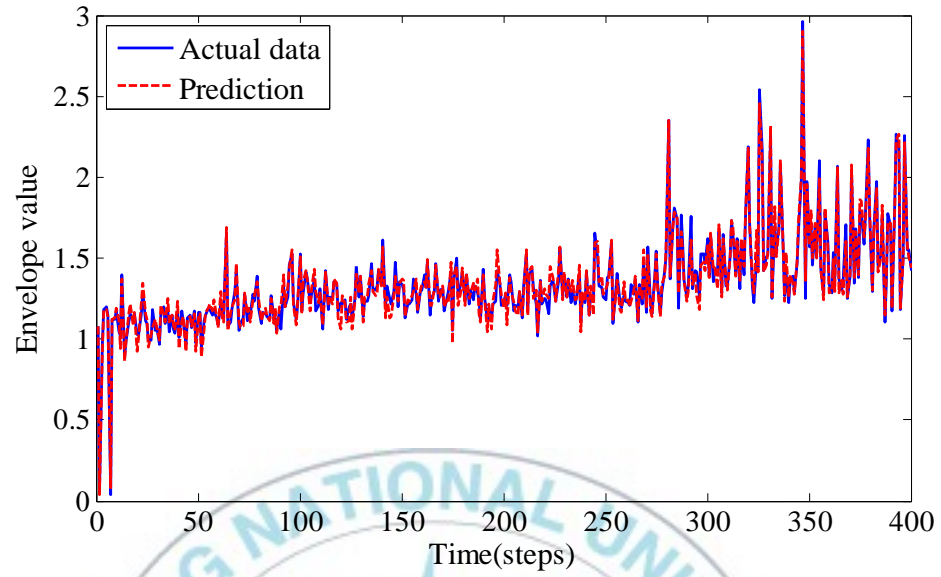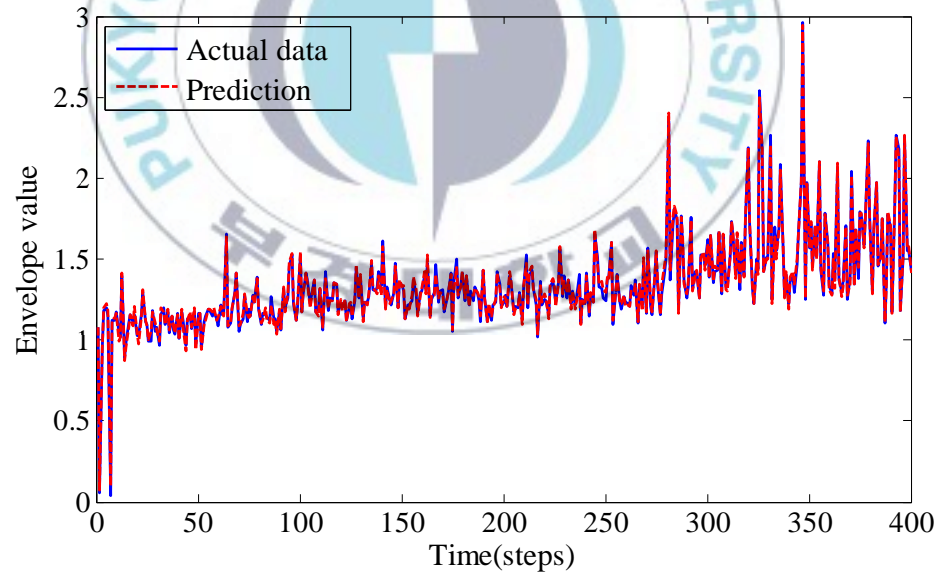(b) 200 number of particles

(c) 1000 number of particles

Fig. 3.8 Prediction result of RMS data
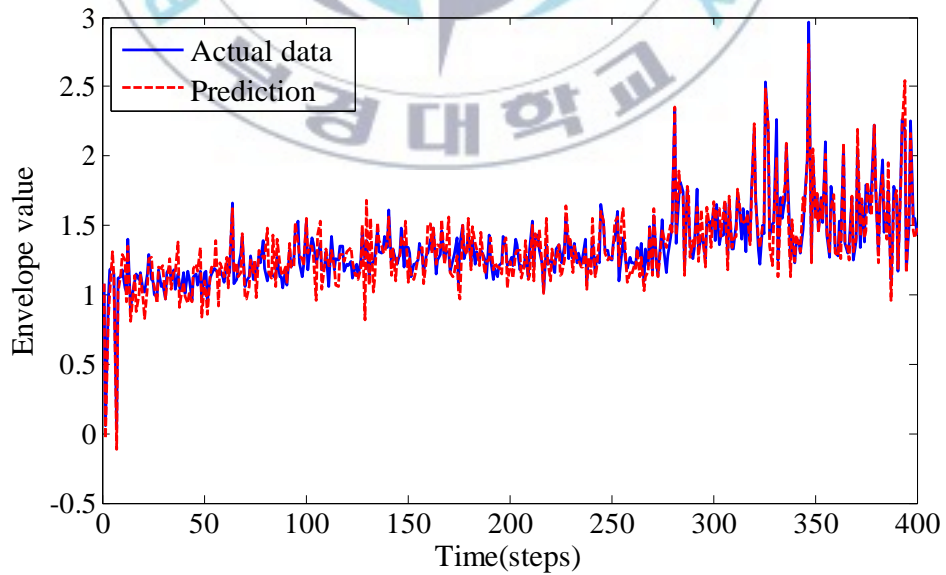
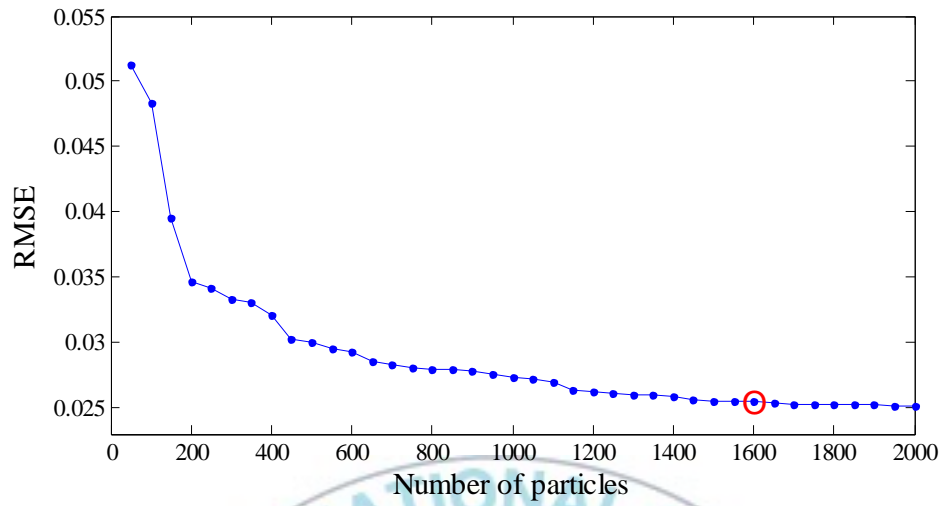(a) 50 number of particles

(b) 200 number of particles

(c) 1000 number of particles

Fig. 3.9 Prediction result of envelope data

The different number of particles sets shows the different results of prediction process. Figs. 3.8 and 3.9 show the prediction result of peak acceleration and envelope acceleration for a different number of particles. Both figure illustrated that the prediction becomes more accurate with the increased number of particles. The RMSE values for 1000 number of particles are 0.0243 and 0.0229 for RMS data and envelope data respectively. The prediction results are summarized in Table 3.2. By plotting the RMS error over the sequence number of particles, the optimal number of particles is obtained as shown in Fig. 3.10. The approximate number of particles is 1600 and 1100 for RMS data and envelope data respectively. Over the optimal number of particles, the RMSE value tends to be constant.

Table 3.2 Performance evaluation results of prediction

Data set	Error measure					
	RMSE	MAE	VAE	MARE	VRE	CR
RMS data						
$n = 50$	0.0422	0.0421	0.0067	0.1193	0.0466	0.7714
$n = 200$	0.0262	0.0259	0.0025	0.0718	0.0251	0.8994
$n = 1000$	0.0243	0.0175	0.0014	0.0653	0.1248	0.9448
Envelope data						
$n = 50$	0.0927	0.0804	0.0182	0.0656	0.0059	0.9378
$n = 200$	0.0441	0.0407	0.0043	0.0343	0.0034	0.9826
$n = 1000$	0.0229	0.0191	0.0009	0.0163	0.0008	0.9961

(a) RMS data
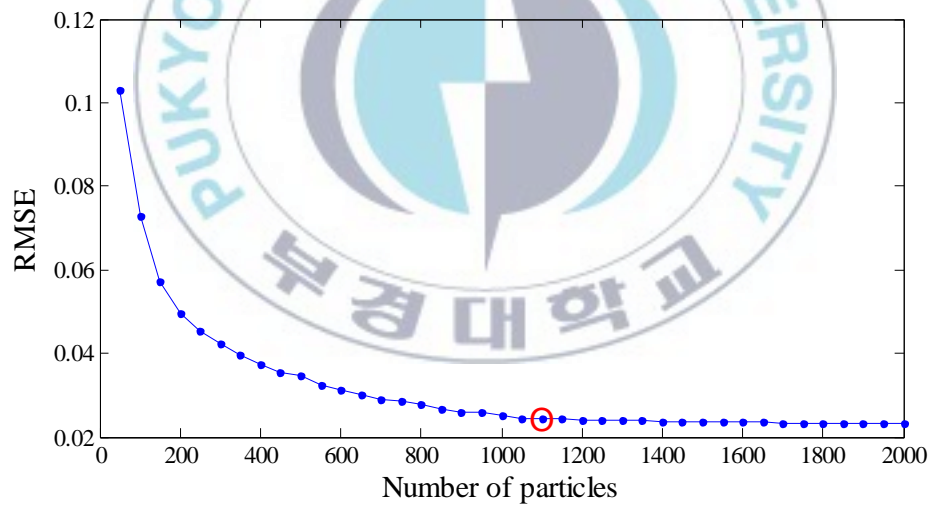
(b) Envelope data

Fig. 3.10 Error prediction degradation over number of particles

References

[1]   G. Niu, B.S. Yang, Dempster-Shafer regression for multi-step-ahead time series prediction towards data-driven machinery prognosis, Mechanical Systems and Signal Processing 23 (2009) 740-751.

[2]   A. Doucet, N. de Freitas, N. Gordon, Sequential Mont Carlo in practice, Springer-Verlag New York (2000).

IV. PF Application for Crack Propagation Data

1. Procedure of Experiment

SPV50 pressure vessel steel that used in the present study was available in the form of an original 12 mm-thick hot rolled plate with an identifiable rolling direction. The chemical composition and the mechanical properties of SPV pressure vessel steel are given in Table 4.1 and Table 4.2. Compact tension (CT) specimens with LT orientation were prepared according to ASTM E647 recommendation Fig. 4.1. Ten fatigue crack growth data have been acquired by fatigue crack growth test as shown in Fig. 4.2. All specimens were cut off with the same sheet and the specimen width, W = 50 mm and initially pre-cracked to produce a sharp crack front.
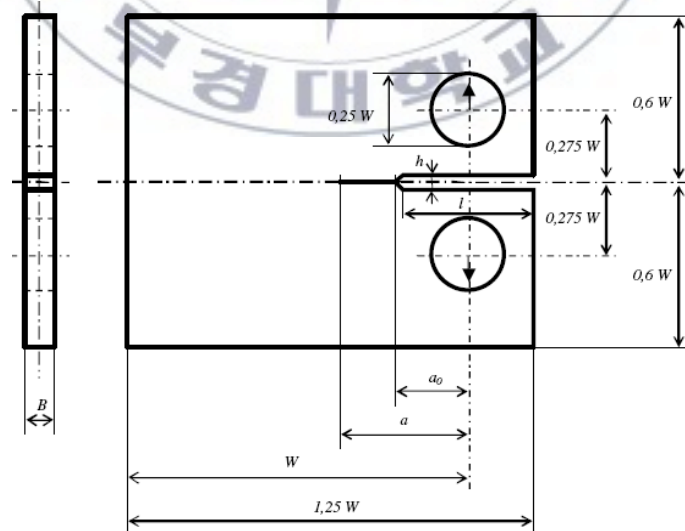

Fig. 4.1 CT specimen with highlight of applied load and symmetry planes

Fig. 4.2 Fatigue crack growth rate test

Table 4.1 Chemical composition of SPV50 pressure vessel steel (wt. %)

C	Si	Mn	P	S	Ni	Cr	Mo	Fe
0.13	0.3	1.27	0.016	0.004	0.12	0.01	0.05	Bal.

Table 4.2 Mechanical properties of SPV50 pressure vessel steel

Tensile strength, σ_t	Yield strength, σ_y	Elongation, ε	Modulus of elasticity, E	Poisson's ratio, v
656 MPa	572 MPa	35 %	2000 MPa	0.3

In order to align the specimens, circular washer were used to hold the specimens in the mid-plane of the clevis grips during fatigue crack growth testing. The experiment was taken in servo-hydraulic testing machine interfaced to a computer for control and data acquisition. All tests were conducted in air and at

45

room temperature and carried out in tension-tension and constants amplitude loading. The loading control mode for controlled loading range of 3.9 kN, the stress ratio of 0.1 and the frequency of 5 Hz. The ΔP level adjusted was found to be within $\pm 0.2\%$ of the range, and this had been considered satisfactory. The microcomputer generates a time series, which is used as the input to multiplying digital-to-analog converter (DAC). The crack opening displacement (COD) is measured by crack opening displacement gauge mounted on the face of the machine notch. This signal and the load cell output are simultaneously digitized and take into microcomputer. The crack lengths were mainly measured by the compliance method. In addition, one person measured the crack lengths on both specimen surfaces with the help of a travelling microscope (x100).

2. Physical Model of Residual Life Prediction

For predicting residual life, the rate at which a fatigue crack grows must be suitably described in terms of various cracks driving parameter [1]. The basic model for crack propagation rate was first proposed by Paris and Erdogan [2]. In the fracture mechanics approach to fatigue, the fatigue crack growth rate (FCGR) for metallic materials in related to the stress intensity factor (SIF) range in the form of Paris-Erdogan equation is given by

$$\frac{da}{dN} = C\left(\Delta K\right)^{m} \tag{4.1}$$

where da/dN is fatigue crack growth rate, C and m are the material constants, a is crack length, N is the number of cycles to load, and ΔK is SIF range. The material constant C and m, hereinafter called as crack growth rate coefficient and exponent, respectively.

3. Results

3.1 Experiment Result

The crack lengths, a, versus the corresponding number of cycles, N, under constant amplitude loading were measured through continuous monitoring during fatigue testing and plotted in Fig. 4.3. The crack propagation data is acquired from 26 mm to 37 mm of crack length, where the increment of crack length is 0.5 mm. As shown in Fig. 4.3, the crack propagation curve under the same loading condition for each specimen is different; if there is no experimental error, this refer to the effect of uncontrolled factors such as the presence of residual stresses or micro-structural heterogeneity in the material. This phenomenon called the randomness of fatigue crack growth rate [3].
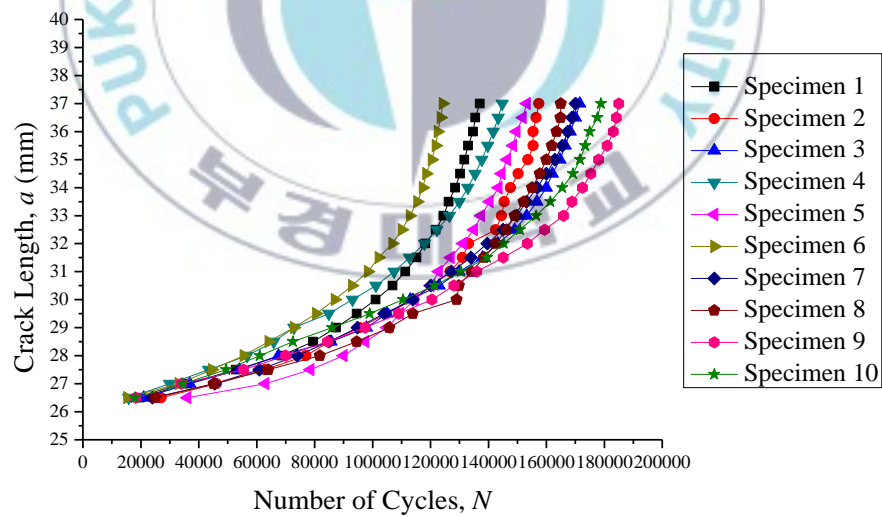

Fig. 4.3 Experimental crack length as a function of cycle's number

Table 4.3 shows the evident of randomness phenomena giving by the variation

of parameter C and m for each specimen. The crack growth rate, da/dN is obtained by taking the derivative of the above crack length, a, versus number of cycles, N, curve. By using these data, the value of da/dN can be calculated easily.

Table 4.3 Samples of random variables C and m

Specimen	C	m	R^2
1	3.07e-14	3.20	0.93
2	2.01e-14	3.25	0.83
3	1.51e-13	2.93	0.96
4	1.36e-12	2.62	0.97
5	1.96e-13	2.93	0.80
6	1.08e-14	3.37	0.97
7	3.12e-13	2.82	0.97
8	2.80e-14	3.19	0.73
9	4.16e-14	3.10	0.94
10	4.21e-13	2.76	0.95

According to ASTM E647 [4], stress intensity factor (ΔK) can be calculated by following equation:

$$\Delta K = \frac{\Delta P}{B\sqrt{W}} \frac{(2+\alpha)}{(1-\alpha)^{3/2}} (0.886 + 4.64\alpha - 13.32\alpha^2 + 14.72\alpha^3 - 5.6\alpha^4) \qquad (4.2)$$

where ΔP is force range can be written as $\Delta P = P_{max} - P_{min}$, B is thickness of specimen, W is width of specimen and $\alpha = a/W$.

Fig. 4.4 shows the relation between FCGR, da/dN and number of cycles, N. It can be seen that most of the life of the specimen is spent while the crack propagation is relatively small. It means there is a specified region where the crack propagation is rising rapidly.
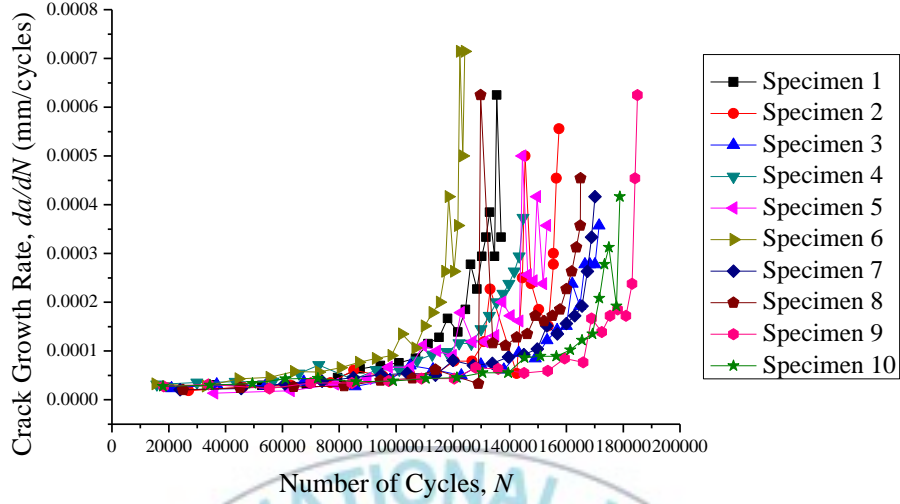
Fig. 4.4 Crack growth rate vs. number of cycles: experimental data

3.2 Determination of Parameter (C, m)

In order to obtain the parameter (C, m), the differential equation model (Paris law equation) given by Eq. (4.1) is fitted to the whole experimental data set by linear least square regression method and plotted in log-log scale [5]:

$$\log\left(\frac{da}{dN}\right) = \log(C) + m \log(\Delta K) \tag{4.3}$$

A plot of the logarithmic crack propagation versus the logarithmic SIF range is depicted in Fig. 4.5. The mean of crack growth rate coefficient and crack growth rate exponent for ten specimens are respectively equal to $C = 3.069 \times 10^{-14}$ and $m = 3.2$. The estimate of the crack growth rate exponent as 3.2 indicates that there is evidence that the crack increasing risk of failure as the number of cycles. The estimation of crack growth rate coefficient was very small, indicating that the mean lifetime of the specimens is extremely large as estimated from these data

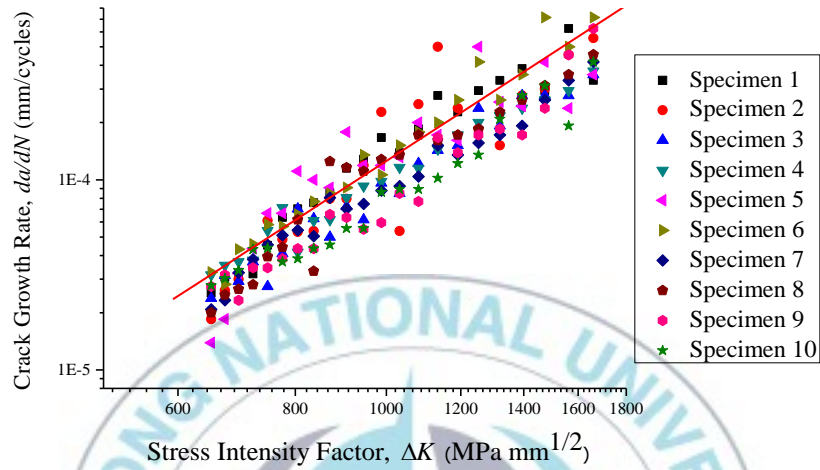(approximate 120,000 – 180,000 number of cycles).

Fig. 4.5 Crack growth rate vs. stress intensity factor: least square result

Due to the randomness phenomenon as shown in Table 4.3, it is very difficult to predict the crack propagation. Therefore, normal probability density function fit is employed in order to obtain the maximum density value of number of cycles for each crack length data point. The normal PDF formula can be written as follows

- Normal density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{4.4}$$

where, μ is the mean and σ is the standard deviation.

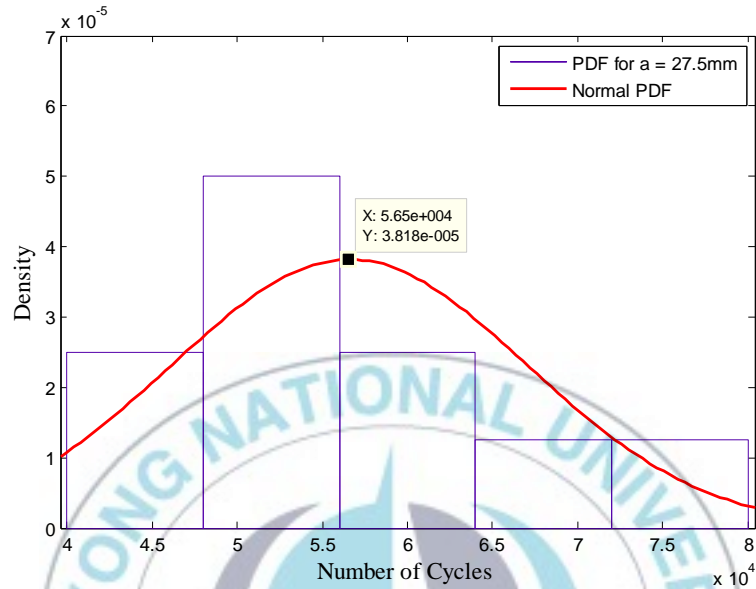Example result of normal PDF fit with number of bins equal to 5 are shown in Fig. 4.6.

Fig. 4.6 PDF fit using normal PDF for *a* = 27.5 mm

Figs. 4.7 and 4.8 illustrate the density plot for crack length vs. number of cycles and crack growth rate vs. number of cycles, respectively. Fig. 4.7 shows the comparison between normal PDF fit and mean of ten crack length data. It can be seen that there is no significant differences, indicated the two graph is coincided. According to Fig. 4.8, the crack growth rate seems more smoothly comparing with Fig. 4.4. Then we can use this data for prediction using particle filter method. The least square result of crack propagation vs. SIF for normal PDF fit is shown in Fig. 4.9. The mean of crack growth rate coefficient and crack growth rate exponent is equal to $C = 6.42 \times 10^{-14}$ and $m = 3.06$, respectively.
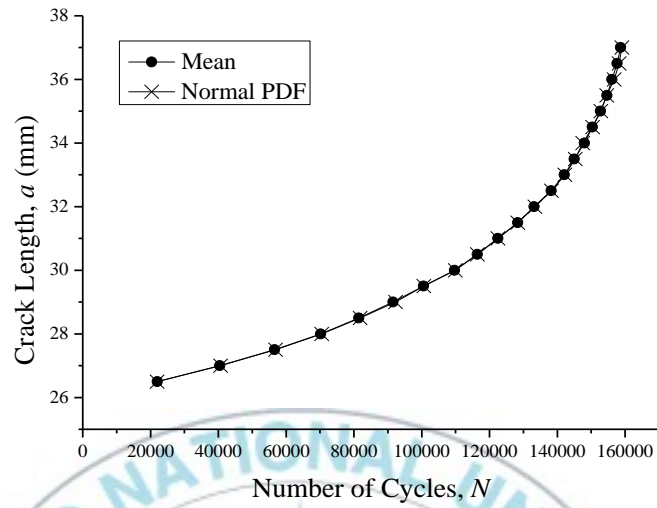
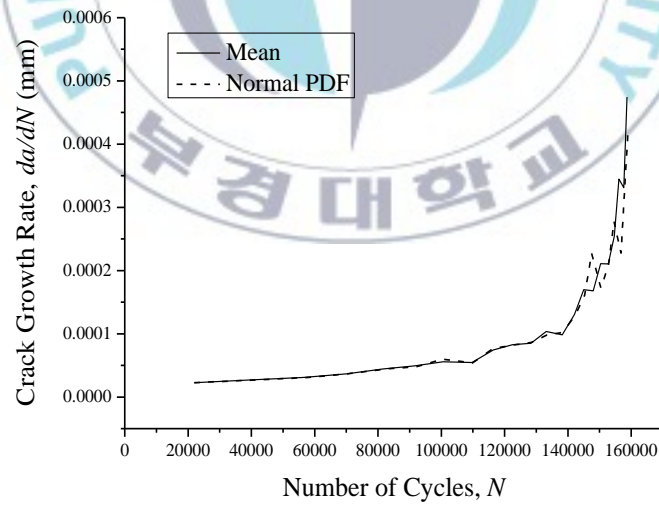Fig. 4.7 Crack length vs. number of cycles using normal PDF fit
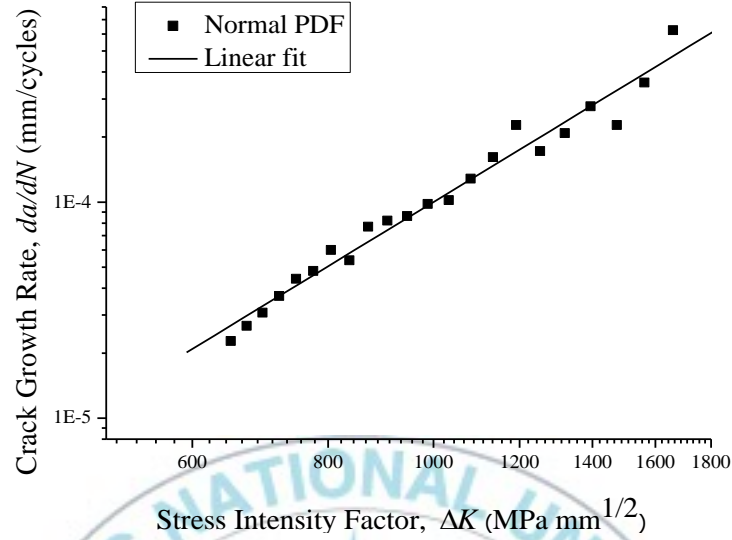

Fig. 4.8 Crack growth rate vs. number of cycles

Fig. 4.9 Least square result of crack propagation vs. SIF

3.3 Finite Element Analysis

The FE model is constructed using ANSYS Multiphysics 11, a commercial FE software package. In this study, the static analysis is employed to obtain SIF. The material behavior assumed as linear elastic material with a homogeneous, isotropic material near the crack region. Use particular command in ANSYS; it automatically generates singular elements around the specified key point.

Assume the SIF is equivalent along the thickness of 3D model, therefore 2D analysis is considered to determine the crack intensity factor Fig. 4.10. The uniform, homogeneous plate of CT specimen is symmetric about horizontal axes in both geometry and load. This means that the state of stress and deformation below horizontal centerline is a mirror image of above the centerline, and likewise for a vertical centerline. We can take advantage of the symmetry and, by applying the correct boundary conditions, use only a half of the plate for the FE model. Fig. 4.10 shows the 2D symmetric analysis with triangle meshing, constraints and load

direction, respectively. Plane 183 solid with triangle element shape and plane stress was used in static analysis. The Young's modulus is 2 GPa and a Poisson's ratio is 0.3.
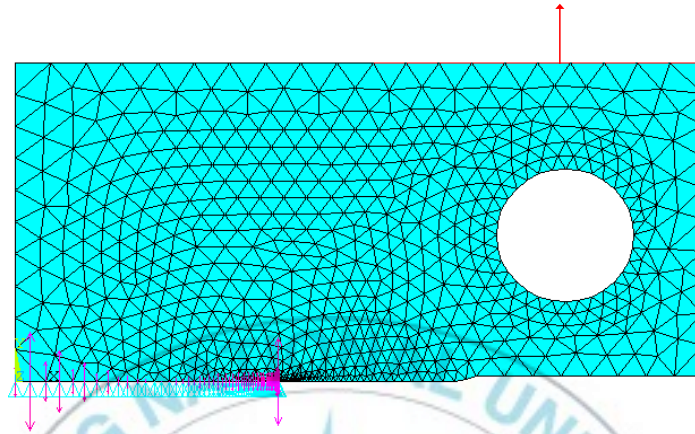

Fig. 4.10 2D FE meshing, constraint and load direction adopted for the 2D analysis (a = 26 mm)
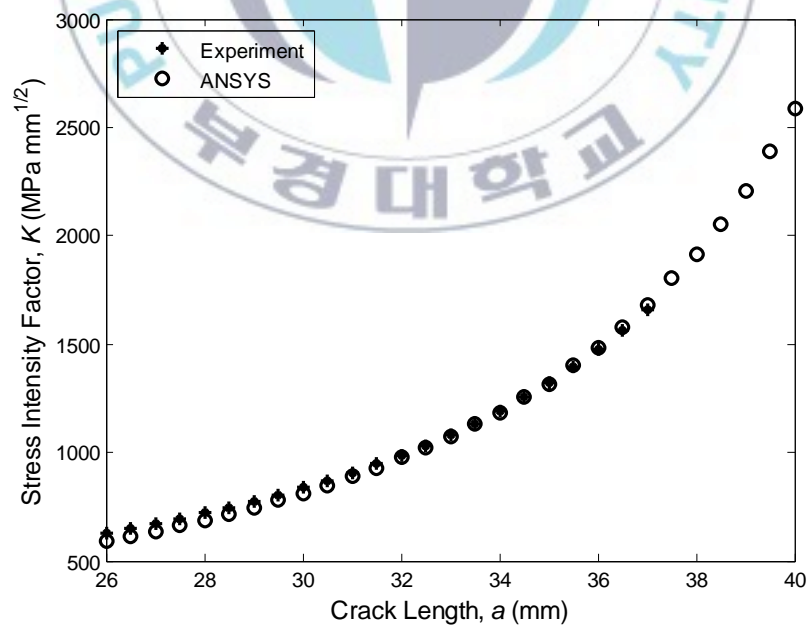

Fig. 4.11 Comparison of SIF

Furthermore, 2D stress intensity factor method using ANSYS is employed to obtain mode-I SIF from 26 mm crack length to 40 mm crack length. The comparison of stress intensity factor between experiment and ANSYS result was shown in Fig. 4.11. The mean error between experiment and ANSYS result is 2.42%.

3.4 Prediction Result

Fig. 4.12 shows the performance of PF due to the prediction of crack growth rate for crack length data of normal PDF fit. In order to evaluate the prediction performance, the root mean square error (RMSE) is employed as follows:

$$RSME = \frac{1}{n}\sqrt{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad (4.5)$$

where, y_t and $\hat{y}_t$ are the actual value of crack propagation data and the predicted value, respectively.

The PF prediction performance of 1000 number of particles from 26 mm crack length to 37 mm crack length was shown in Fig. 4.12.

In this case, experiment data are acquired up to 37 mm of crack length. Therefore, we set up 37 mm crack length as prediction start point or alarm setting. In addition, we assume 40 mm crack length is a final failure of crack. Moreover, the PF prediction from 37.5 mm to 40 mm of crack length was also shown in Fig. 4.12. Finally, the residual life from prediction start point (37 mm) to final failure (40 mm) is 3212 number of cycles as shown in Fig. 4.13.
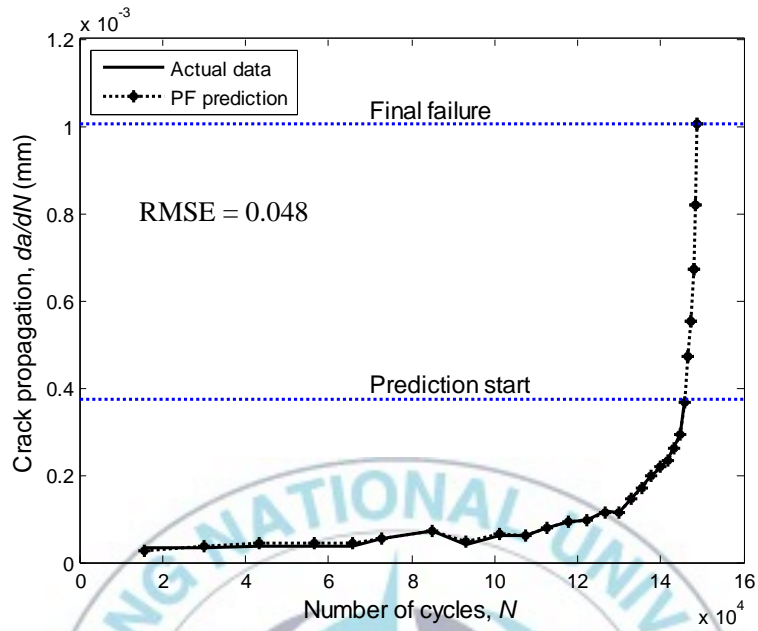
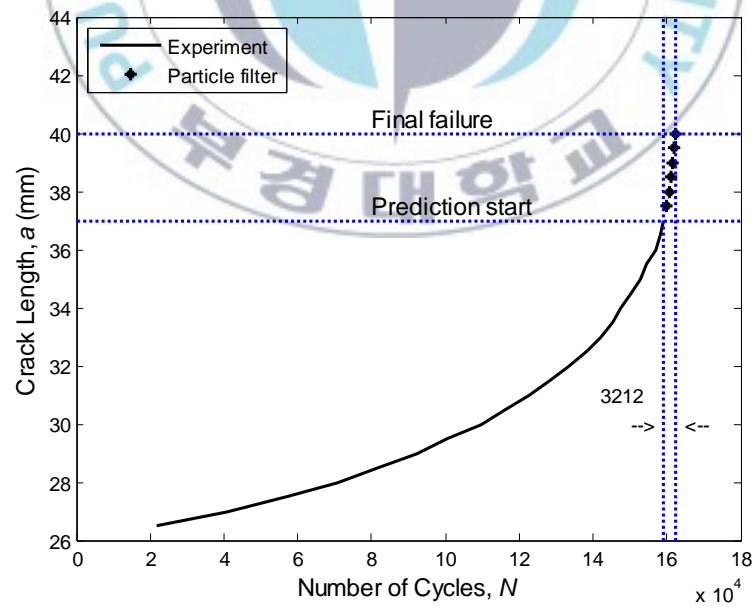Fig. 4.12 Crack growth rate prediction of Normal PDF data

Fig. 4.13 Residual life estimation of Normal PDF data

References

[1] E. Bechhoefer, A. Bernard, D. He, Use of Paris law for prediction of component remaining life, IEEE Aerospace Conference (2008) 1-9.

[2] P.C. Paris, F. Erdogan, A critical analysis of crack propagation laws, ASME Journal of Basic Engineering 85 (1963) 528-534.

[3] S.J. Kim, Y.S. Kong, Y.S. Kim, B.T. Kim, Y.J. Noh, H.K. Ku, H.C. Jeong, On probabilistic fatigue crack growth behavior under constant amplitude loads, Key Engineering Materials 353-358 (2007) 1219-1224.

[4] ASTM E 647-00. Standard test method for measurement of fatigue crack growth rates. Philadelphia: ASTM.

[5] F. Casciati, P. Colombi, L. Faravelli, Inherent variability of an experimental crack growth curve, Structural Safety 29 (2007) 66-76.
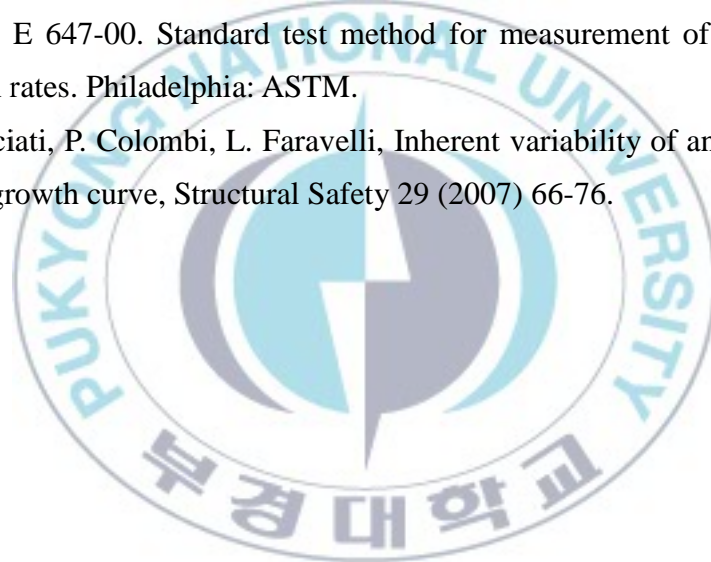
Ⅴ. RVM and LR Application

1. Methodology

The propose method as shown in Fig. 5.1 employs bearing failure simulation data and experiment nun-to-failure data. One dimensional feature, namely kurtosis is calculated initially. This features can be used to represents the information of bearing from normal to failure condition. Failure degradation is calculated using LR method for both case of simulation data and experimental data. The results are regarded as target vectors of failure probability. RVM is used for training the run-to-failure kurtosis data and target vectors of failure probability then predict of individual unit of machine component. Prediction model is obtained using training process which saved weight and bias in the model. To evaluate the prediction performance, root mean square error (*RMSE*) and correlation (*R*) are utilized. The *RMSE* and *R* formula given by the following equation:

$$RMSE = \left(\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \right)^{1/2} \tag{5.1}$$

$$R = \frac{Cov(y_t, \hat{y}_t)}{\sigma_{y_t} \sigma_{\hat{y}_t}}, \quad Cov(y_t, \hat{y}_t) = \frac{1}{N} \sum_{t=1}^{N} (y_t - \overline{y}_t)(\hat{y}_t - \overline{\hat{y}}_t) \tag{5.2}$$

where *R* is correlation and $Cov(y_t, \hat{y}_t)$ is covariance between actual and predicted values. y_t indicates actual value and $\hat{y}_t$ refers to predicted value. In addition, $\overline{y}_t$ and

$\overline{y_t}$ denotes the averages result of actual and predicted value, respectively. Moreover, σ_{y_t} is standard deviation of actual value and $\sigma_{\hat{y}_t}$ is standard deviation of predicted value. The symbol of *n* denotes the number of predictable data. The smaller value of *RMSE* indicates a higher accuracy of prediction and a high correlation value expressed a good prediction.
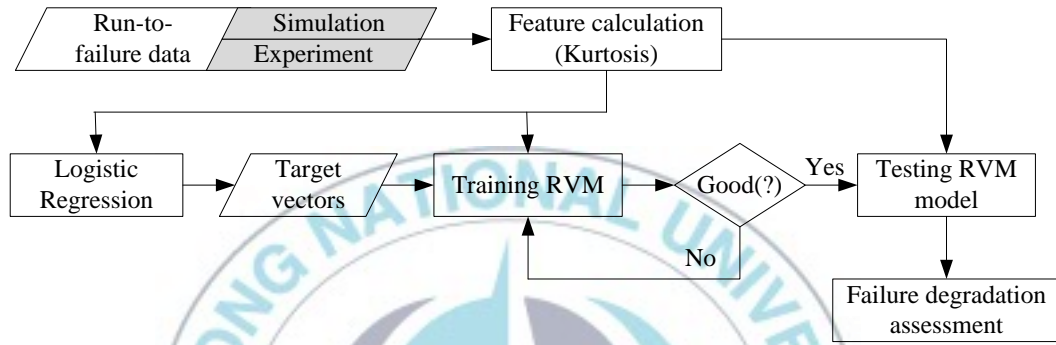

Fig. 5.1 Schematic diagram of failure degradation assessment model

2. Simulation and Experiment

2.1 Simulation Data

The proposed method is validated by using simulation data of bearing defect degradation and real data obtained from experimental work. In the simulation, we developed vibration CM data that represents defect propagation of rolling element bearing by MATLAB program. The properties of rolling element bearing in the simulation were as follows: pitch diameter of 23 mm, number of rolling elements of 9; roller diameter of 8 mm and contact angle of 0°. We conducted bearing outer-race, bearing inner-race and ball fault defect simulation under rotating speed 100 rpm and sampling frequency 5 kHz. Fig. 5.2 shows the accumulated signal of three bearing defect. Fig. 5.2(a) shows the simulated time domain. This signal was

converted to frequency domain using fast-Fourier transform (FFT) as shown in Fig. 5.2(b). This figure present that the spectrum was dominated by high-frequency resonant signals. To separate the bearing fault frequency signal from these dominant signals, the vibration signals were band-pass filtered and rectified. Fig. 5.2(c) depict the peaks were detected at 4.88 Hz (1xBPFO), 10.07 Hz (1xBPFI) and 8.39 Hz (2xBSF), respectively, which closely matched with the calculated outer-race fault, inner-race fault and ball fault defect frequency as indicated in the top of figure.
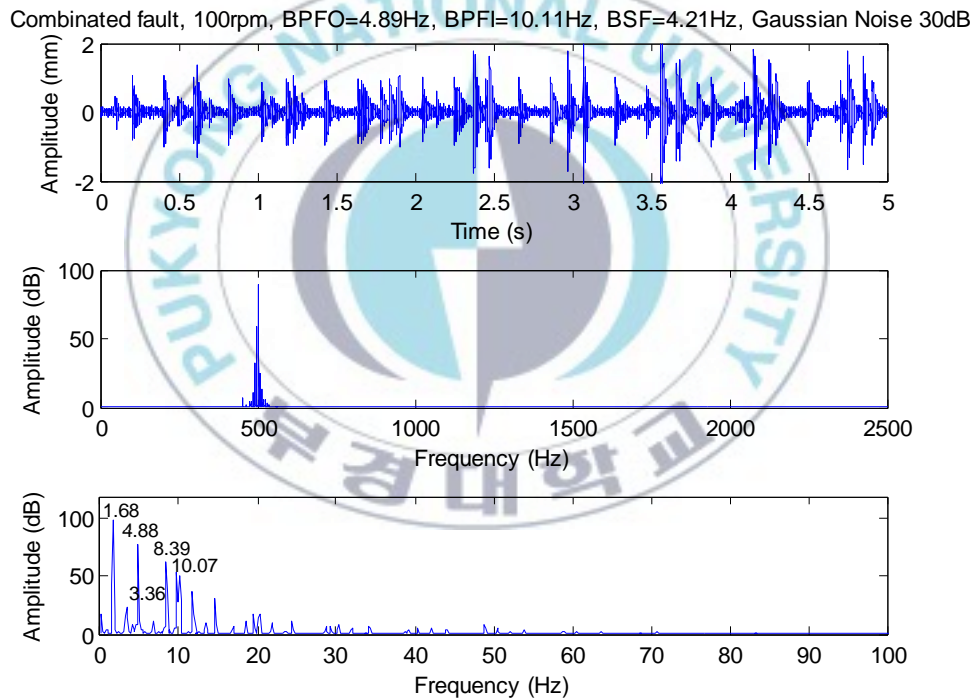

Fig. 5.2 Accumulated signal of three bearing defects: (a) Time domain plot of raw signal, (b) Frequency spectrum of raw signal, (c) Fault detection after demodulation

The simulated signals were repeatedly generated from the computer program

based on equations presented in [1, 2]. Every simulated signal has defect impulses that increase at different rates and time measurement. It has been observed in bearing life test that bearing degradation signals possess an inherent exponential growth [3, 4]. The example result of feature calculation using kurtosis is shown in Fig. 5.3. For further information of Fig. 5.3 will be discussed in section 3.
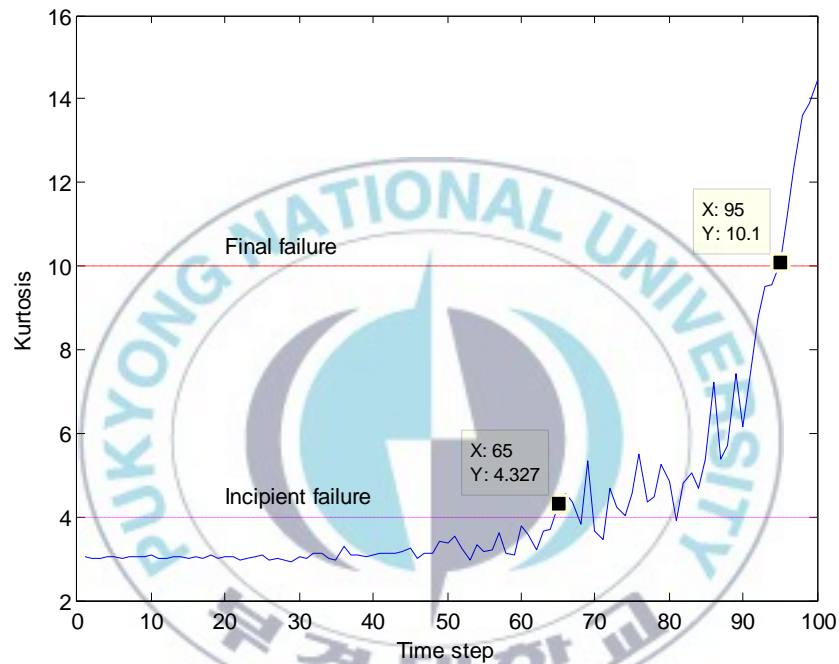

Fig. 5.3 Kurtosis of bearing sample No.15

2.2 Experimental Data

The experimental data was generated from bearing test rig that able to produce run-to-failure data. These data was downloaded from Prognostics Center of Excellence (PCoE) through prognostic data repository contributed by Intelligent Maintenance System (IMS), University of Cincinnati [5]. Bearing test rig consists of four bearings that were installed on one shaft as presented in Fig. 5.4. The

rotation speed of shaft was kept constantly at 2000 rpm and a radial load of 6000 lb was placed onto the shaft and bearing by a spring mechanism. All bearings are forced lubricated. The bearings used are Rexnord ZA-115 double row bearings that have 16 rollers at each row, a pitch diameter of 71.5 mm, roller diameter of 7.9 mm, and a tapered contact angle of 15.17°. The vibration signals were acquired by eight accelerometers from PCB 353B33 (a high sensitivity quartz ICP accelerometers) that were installed at vertical and horizontal directions. Four thermocouples were also installed to the outer-race of each bearing to record bearing temperature for monitoring lubrication purposes. Vibrations signals were collected every 20 minutes by NI-DAQ Card 6062E data acquisition card. The data sampling rate was 20 kHz and the data length was 20480 points.
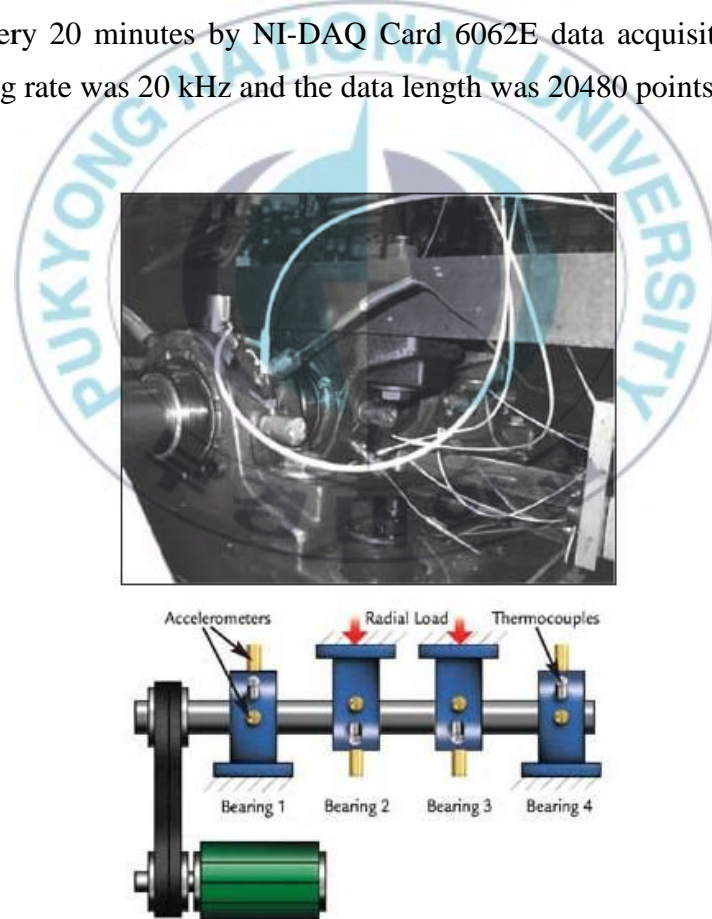

Fig. 5.4 Bearing test rig and sensor placement illustration [6]

The collected vibration data consist of 2 datasets which consist of 8 complete failure data with different failure time and 4 data regarded as normal condition. One dimensional feature namely kurtosis is calculated for validating the proposed method. The presentation of kurtosis of vibration data for dataset 1 and dataset 2 with threshold setting are shown in Fig. 5.5. Dataset 1 consists of 8 run-to-failure data and dataset 2 consists of 4 censored data (means, have not failure until experiment stopped at measurement points equal to 2156). Eleven data are used (consist of 7 failure data and 4 censored data) for training and remaining for testing. Fig. 5.5 shows the example kurtosis data of dataset 1 (bearing No. 1) and dataset 2 (bearing No. 1).
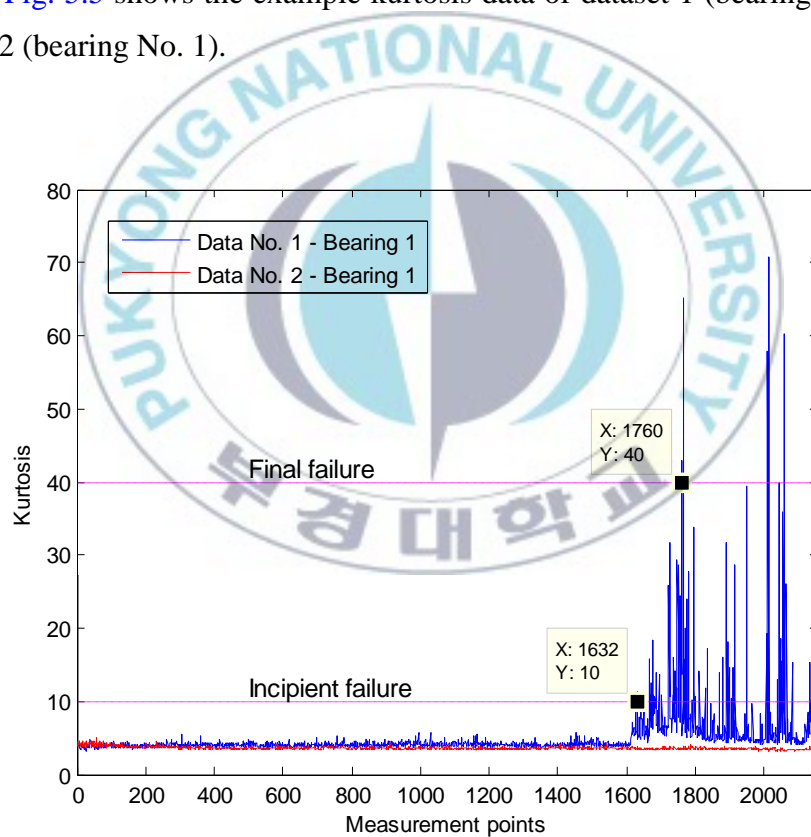

Fig. 5.5 Kurtosis of vibration data

3. Results and Discussion

3.1 Simulation

For the sake of illustration, we construct run-to-failure data set to which regression was applied. In the case of simulated data, 25 bearing samples are simulated in the range 0-100 of time step. After feature calculation of kurtosis, predefined threshold of incipient failure and final failure were set-up initially as shown in Fig. 5.3. When the kurtosis signal exceeds both thresholds, which means the incipient failure time and final failure time is noted. Then, these data are used for calculating the failure probability from incipient failure time until final failure using LR. Fig. 5.3 shows the example kurtosis signal of bearing No. 15 with the incipient failure time and final failure time are 65 and 95, respectively. For final failure, the results are 20 bearings have failed before time step equal to 100 and 5 are censored. In this works, single logistic regression is applied, means the independent variable is only one which estimated from incipient and final failure time of kurtosis data can be denoted as x_1. The parameter α and β_1 are calculated using maximum likelihood estimation (MLE). Therefore, the logit model corresponding to Eq. 2.14 is determined by

$$g(\vec{\mathbf{x}}) = \log\left(\frac{P(\vec{\mathbf{x}})}{1 - P(\vec{\mathbf{x}})}\right) = 4.3318 + 0.2172 * x_1$$

According to above equation $g(\vec{\mathbf{x}})$ can be calculated. Moreover, failure probability according to Eq. 2.13 can be estimated. The result of failure probability is shown in Fig. 5.6. This figure illustrated the failure probability of 24 bearings simulated data calculated from incipient failure (failure probability equal

to 0) until final failure occurred (failure probability equal to 1). Moreover, this result regarded as target vector and will be use for RVM training. The figure shows that the degradation from incipient failure until final failure occurred is 48-100 of time step. We can say, within the time step 0-47 the condition is normal.
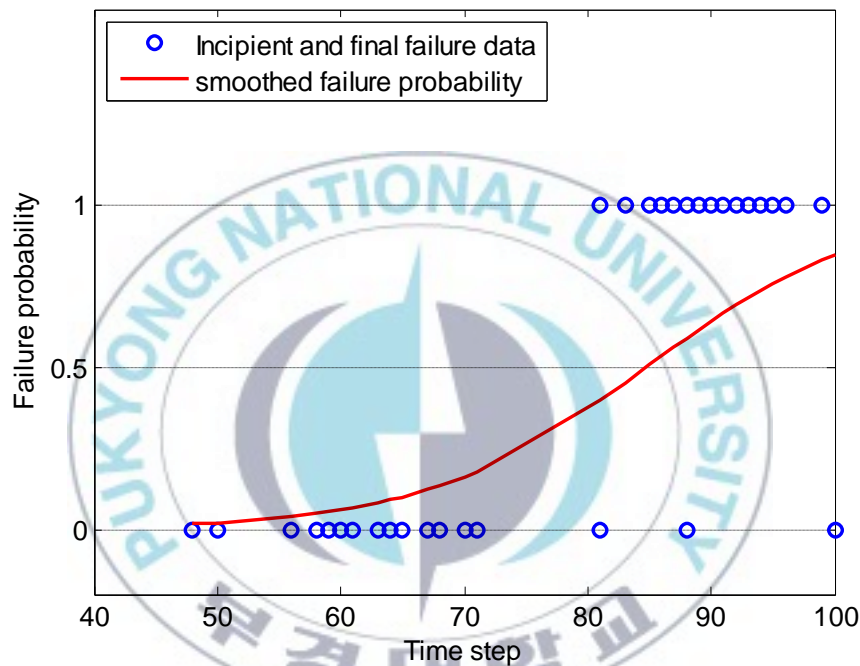

Fig. 5.6 Logistic regression result of simulated data

RVM then trained by input data obtained from kurtosis of bearing simulated data and target vector of failure probability estimated by LR. In RVM training, 24 kurtosis data are utilized and remaining 1 for testing. We employed Gaussian kernel and searched kernel-width in the range of $\{10^{-0.5}, 10^{-0.4}, 10^{-0.3}, ..., 10^{-6}\}$ to obtain optimized RVM training process. After training process which produced and saved weight and bias, RVM is employed to predict the failure time of bearing sample No. 15 which has final failure time at 95 as shown in Fig. 5.3.

The prediction result of bearing sample No. 15 is shown in Fig. 5.7. The effect of improper selection of kernel-width parameter is also presented in Fig. 5.7. It shows overfitting phenomena in prediction of failure probability of testing data using kernel-width 1.0×10^{-3}. In additions, Table 5.1 informs the performance of testing process after RVM training with respect to kernel-width. In this work, kernel-width values was studied in the range of $\{10^{-0.5}, 10^{-0.4}, 10^{-0.3}, ..., 10^{-6}\}$ while the kernel-width values of $\{10^{-1}, 10^{-2}\}$ produces serious overfitting. It can be seen in Table 5.1, the optimal value of kernel-width is $1.0 \times 10^{-0.4}$ and 1.0×10^{-5}, where the lower value than $1.0 \times 10^{-0.4}$ and the greater value than 1.0×10^{-5} give the same optimal value of *RMSE* and *R*. The range between 1.0×10^{-1} and 1.0×10^{-3} yields bad overfitting region shown in the higher and lower of *RMSE* and *R*, respectively. However, the optimal kernel-width value has been calculated in order to obtain the prediction result with final failure prediction is 96 as shown in Fig. 5.7.

Table 5.1 Performance of RVM testing w.r.t kernel-width using bearing simulation data

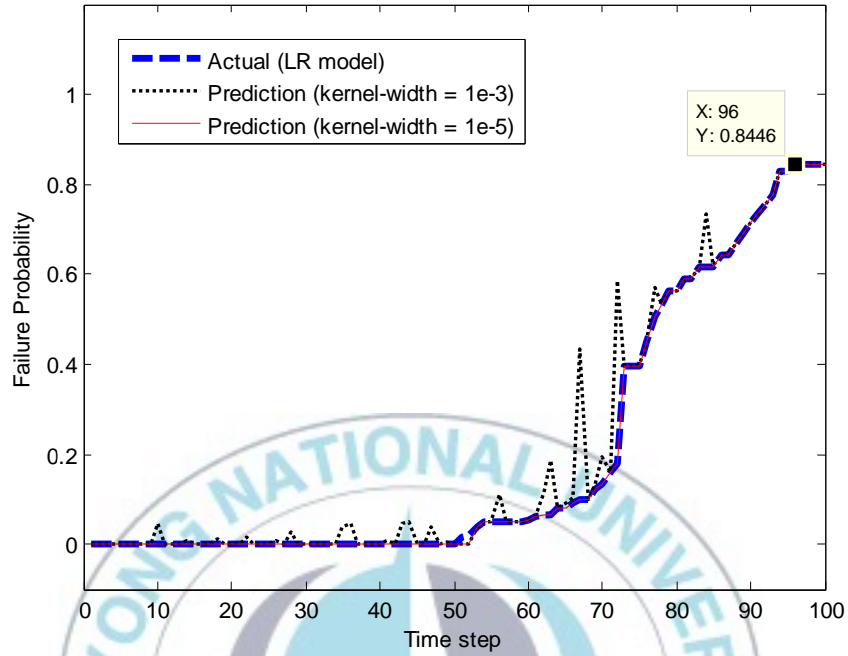| Kernel-width	*RMSE*	*R*
$1.0 \times 10^{-0.5}$	0.003	0.99
$1.0 \times 10^{-0.4}$	**0.003**	**0.99**
$1.0 \times 10^{-0.3}$	0.004	0.989
$1.0 \times 10^{-0.2}$	0.058	0.973
$1.0 \times 10^{-0.1}$	0.022	0.987
1.0×10^{-1}	0.201	0.884
1.0×10^{-2}	0.337	0.812
1.0×10^{-3}	0.058	0.973
1.0×10^{-4}	0.004	0.989
1.0×10^{-5}	**0.003**	**0.99**
1.0×10^{-6}	0.003	0.99

Fig. 5.7 Predicted failure degradation of simulated data

The performance of prediction can be simply calculated as

$$\text{Accuracy} = \left(1 - \frac{|t_a - t_p|}{t_a}\right) \times 100\% = \left(1 - \frac{|95 - 96|}{95}\right) \times 100\% = 98.95\%$$

where t_a refer to actual failure time and t_p is prediction failure time. This prediction result seems overestimate, but the accuracy of 98.95% is acceptable for building the prognostic model.

3.2 Experimental

It can be seen in Fig. 5.8 that the range of failure probability from incipient failure time until final failure time is 1621~1770. Based on this range, RVM was trained by 150 data points of 7 kurtosis of bearing failure data and 4 kurtosis censored data. The target vectors of failure probability in this training process were obtained from LR method (Fig. 5.8) which estimated parameter α and β equal to 7.4466 and 0.0324, respectively. Threshold value both incipient failure and final failure were set initially as shown in Fig. 5.5. Similar to simulation case, when the kurtosis signal exceeds both thresholds, that means the incipient failure time and final failure time is noted. It can be seen in Fig. 5.5, the incipient failure and final failure of dataset 1 – bearing No. 1 are 1632 and 1760, respectively.
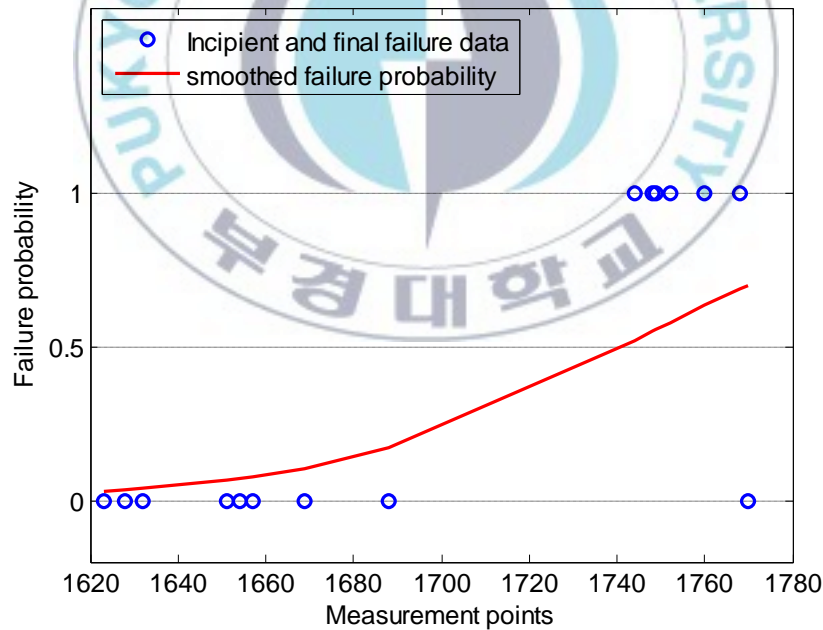

Fig. 5.8 Logistic regression result of experimental data

In the training process, we also used Gaussian kernel. We searched kernel value in

the range of {0.5, 0.1, …, 0.0001} to obtained optimized RVM training process. Table 5.2 illustrates the performance of testing process after RVM training with respect to kernel-width. The prediction result of dataset 1 − bearing No. 1 as a testing data is shown in Fig. 5.9. This figure also depicts the evident of the improper selection of kernel-width shown in the RVM prediction using kernel-width 0.001.

Table 5.2 Performance of RVM testing w.r.t kernel-width experiment bearing data

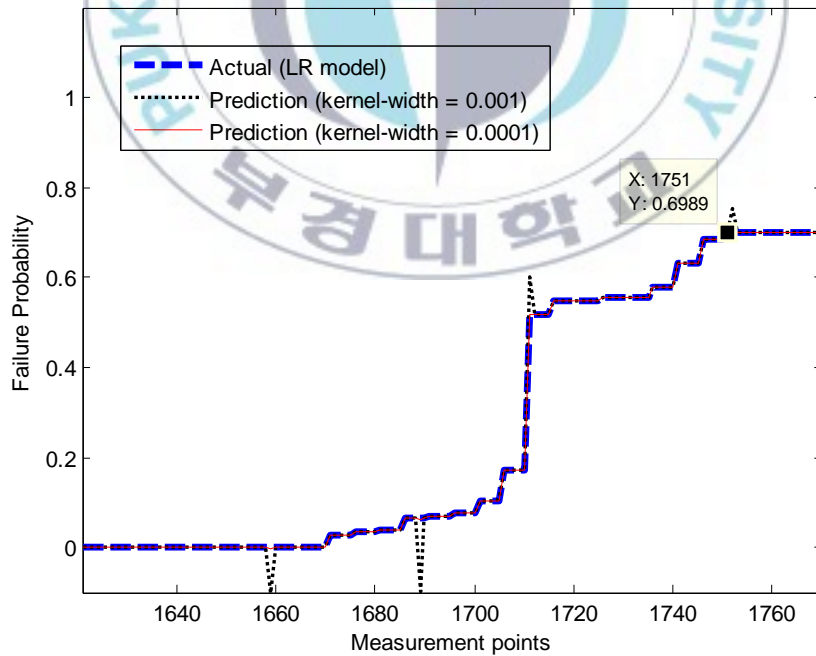| Kernel-width	*RMSE*	*R*
0.5	1.309	0.308
0.1	0.354	0.602
0.01	0.114	0.926
0.001	0.018	0.992
0.0001	**0.0002**	**0.99**

Fig. 5.9 Predicted failure degradation of experimental data

In this case, selection of relatively high kernel-width gave serious overfit prediction of failure probability or RVM testing. The best RVM testing performance was reached at kernel-width 0.0001 with *RMSE* and *R* are 0.0002 and 0.99, respectively. In addition, selection of kernel-width that lowers than 0.0001 affected high error and low correlation of failure probability. Finally, the prediction of failure time is 1751. The accuracy of prediction is calculated using the following formula

$$\text{Accuracy} = \left(1 - \frac{|t_a - t_p|}{t_a}\right) \times 100\% = \left(1 - \frac{|1760 - 1751|}{1760}\right) \times 100\% = 99.49\%$$

This prediction seems underestimation. However, in real application underestimation is more useful than overestimation for machine health prognostics.

References

[1] P.D. McFadden, J.D. Smith, Model for the vibration produced by a single point defect in rolling element bearing, Journal of Sound and Vibration 96 (1984) 69-82.

[2] Y.F. Wang, P.J. Kootsookos, Modeling of low shaft speed bearing faults for condition monitoring, Mechanical Systems and Signal Processing 12(3) (1998) 415-426.

[3] N.Z. Gebraeel, M.A. Lawley, R. Li, J.K. Ryan, Residual-life distribution from component degradation signals: A Bayesian approach, IEE Transaction 37(2005) 543-557.

[4] N. Gebraeel, Sensory-updated residual life distribution for components with exponential degradation pattern, IEEE Transactions on Automation Science

and Engineering 3(4) (2006) 382-393.

[5] J. Lee, H. Qiu, G. Yu, J. Lin, Rexnord Technical Services, 'Bearing Data Set', IMS, University of Cincinnati, NASA Ames Prognostics Data Repository, [http://ti.arc.nasa.gov/project/ prognostics-data-repository], NASA Ames, Moffett Field, CA, 2007.

[6] H. Qiu, J. Lee, J. Lin, Wavelet filter-based weak signature detection method and its application on roller bearing prognostics, Journal of Sound and Vibration 289 (2006) 1066-1090.

VI. Conclusions and Future Work

1. Conclusions

In this paper, the applications of particle filter (PF) for model-based prognosis approach and relevance vector machine (RVM) and logistic regression (LR) for data-driven prognosis approach has been explained. In case of PF application toward model-based prognosis approach which are shown in chapter 3 and chapter 4, mathematic model of monitoring system or part is very important in order to obtain a good prediction result. In chapter 3 we have difficulty to calculate the model of trending data of low methane compressor, thus the mathematic model are adopted from [1] which modified by trial and error. In chapter 4, the prediction model of fatigue crack growth data is Paris' Law equation. The parameters are obtained by least square method and finite element analysis. Finally, the residual life before final crack occurred can be determined. For both real cases, PF is employed to predict the future state or future value based on sequential important sampling and resampling algorithm where the optimal result obtained when the number of particle is increase.

In chapter 5 contain the application of LR and RVM for simulated bearing fault data and experimental data obtained from NASA data repository. To address the probability behavior in real system and then calculate the appropriate model that represents the failure degradation of part or system, LR is employed. LR works based on incipient failure (coded as 0) or final failure (coded as 1) decision given a number of bearing samples data. The result of failure degradation model calculated from LR is regarding as target vector of RVM. Kurtosis data and target

vector are utilized for RVM training. After training process, RVM is employed for predict the final failure time of individual bearing data known as testing process.

2. Future Work

After study about model-based and data-driven prognosis approach, the combination between probability and data-driven approach seems has potential result to address the real prognostics issues due to the model-based has main limitation which only can be applied in particular cases. Thus the future work will concentrate at fusion method between probability and data-driven prognosis approach. The future development method should be able to apply in real system not only in laboratory scale.

Reference

[1] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, Journal of Computational and Graphical Statistics 5 (1996) 1-25.

기계 예지를 위한 모델 기반 및 데이터 주도 방법

Wahyu Caesarendra

부경대학교 대학원 기계설계공학과

국문 요약

산업분야에서 비용대비 효율을 향상시키면서 심각한 결함의 발생 빈도를 낮추는 것은 치열한 시장 경쟁에 있어서 중요한 요소가 되고 있다. 상태 감시, 결함 진단 및 결함 예측을 포함한 정비 활동들을 통하여 이러한 문제점을 극복할 수 있다. 정비 작업의 주된 목적은 최종 고장에 대한 예측과 대응하는 정비활동에 관한 의사결정을 관리하는 것이다. 그러므로, 예지는 상태 감시 및 결함 진단과 더불어 중요한 분야가 되어가고 있다.

본 논문에서는 기계의 예지를 위한 모델 기반 예지 접근방법인 Particle Filter (PF) 방법과 데이터 주도 방법인 Relevance Vector Machine(RVM) 및 Logistic Regression (LR) 방법을 제안하였다. 제안된 방법의 유효성을 검증하기 위해, 순차적인 중요 샘플링과 재샘플링 알고리즘을 기반으로 하는 PF 방법이 저압 메탄 압축기에서 계측된 데이터의 경향 예측과 SPV50 강의 균열 성장에 대한 잔여 수명을 예측하는데 적용되었다. 또한 데이터 주도 예지 접근법에서는 고장으로 인한 성능 저하를 평가하기 위해 RVM 기법과 LR 기법을 결합하였다. LR 기법을 이용하여 고장 성능 저하 모델인 고장 베어링 데이터를 계산하고 RVM 기법을 통하여 시뮬레이션 데이터와 실험 데이터의 각 베어링에 대한 최종 고장을 예측하였다. 이러한 검증을 통하여 이 연구에서 제안된 방법이 신뢰할 수 있는 예지 방법으로서 이용될 수 있는 잠재력을 가지고 있음을 확인하였다.

ACKNOWLEDGMENTS

This is a major milestone in my life, for which I would like to say Thank you from the bottom of my heart to some greatest people.

My deepest appreciation and special thanks goes to my academic and research advisor, Prof. Bo-Suk Yang, for his valuable assistance, guidance, and encouragement in bringing this research work to a successful completion. He showed extraordinary patience and was always available for friendly advice. Thank you Prof. for your trust, effort and time and for all the opportunities you gave to me. And also I will not forget to say my thankful to Prof. Seon-Jin Kim who help me to provide data and gave me some valuable knowledge's, it's very honors to me. Thank you to Prof. Byung-Tak Kim for the valuable comments of my thesis.

I am also sincerely thankful to my senior laboratory: Dr. Achmad Widodo, Dr. Jong-Duk Son, Dr. Gang Niu, Dr. Van Tung Tran, Dr. Min-Chan Shim, and Dr. Jin-Dae Song for their advice and help.

I am grateful to my colleagues and friends in Intelligent Mechanics Laboratory, Pukyong National University, South Korea: Sung-Won Jo, Sung-Do Tae, Ae-Hee Song, Seung-Wook Yang, Jun-Seok Oh, Cho-Won Jung, Jung-Min Han, Md. Younus Ali, Thom Pham Hong, and Sun-Yong Jang. It has been a great honor for me to work with them and shared all great moments.

I would like to express my deepest gratitude to my parents, wife, brother and family in law for love, constant support and encouragements over the years which have been the most important factors in my achievements and my life.

There are many others who were indirectly supportive to my work and I would like to thank all of them from my heart.