Chapter 7. Health Prognosis

Prognostics and Health Management (PHM)

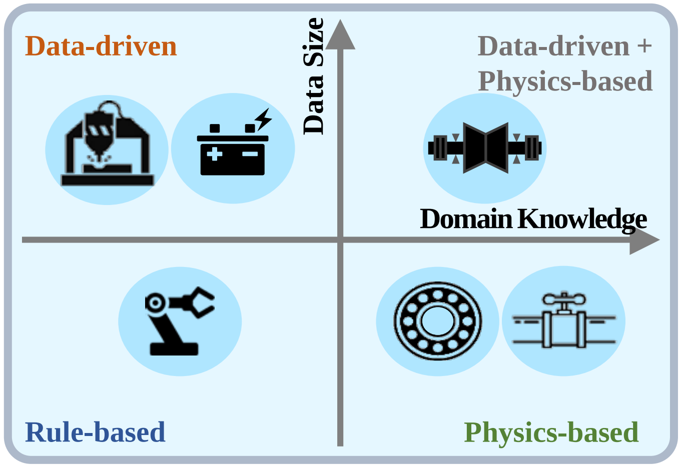
Byeng D. Youn

System Health & Risk Management Laboratory
Department of Mechanical & Aerospace Engineering
Seoul National University

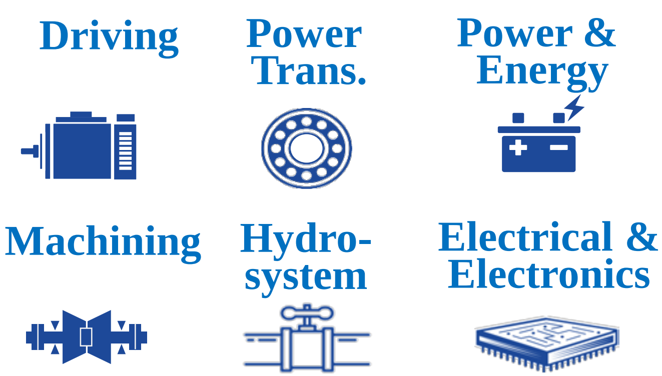
CONTENTS

Introduction

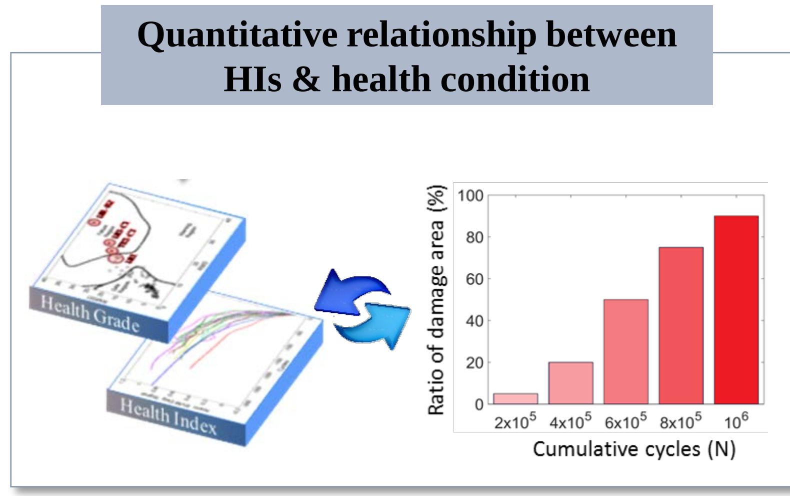
- **Prognostics Approaches**

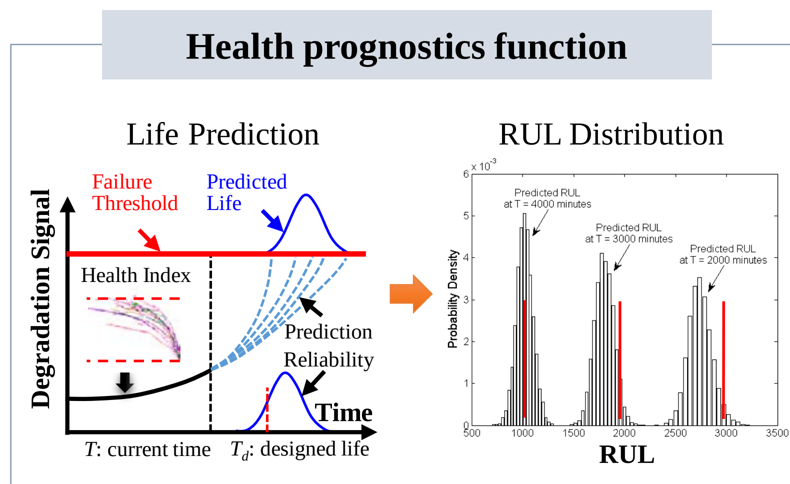


- **6 Core Assets in Manufacturing**



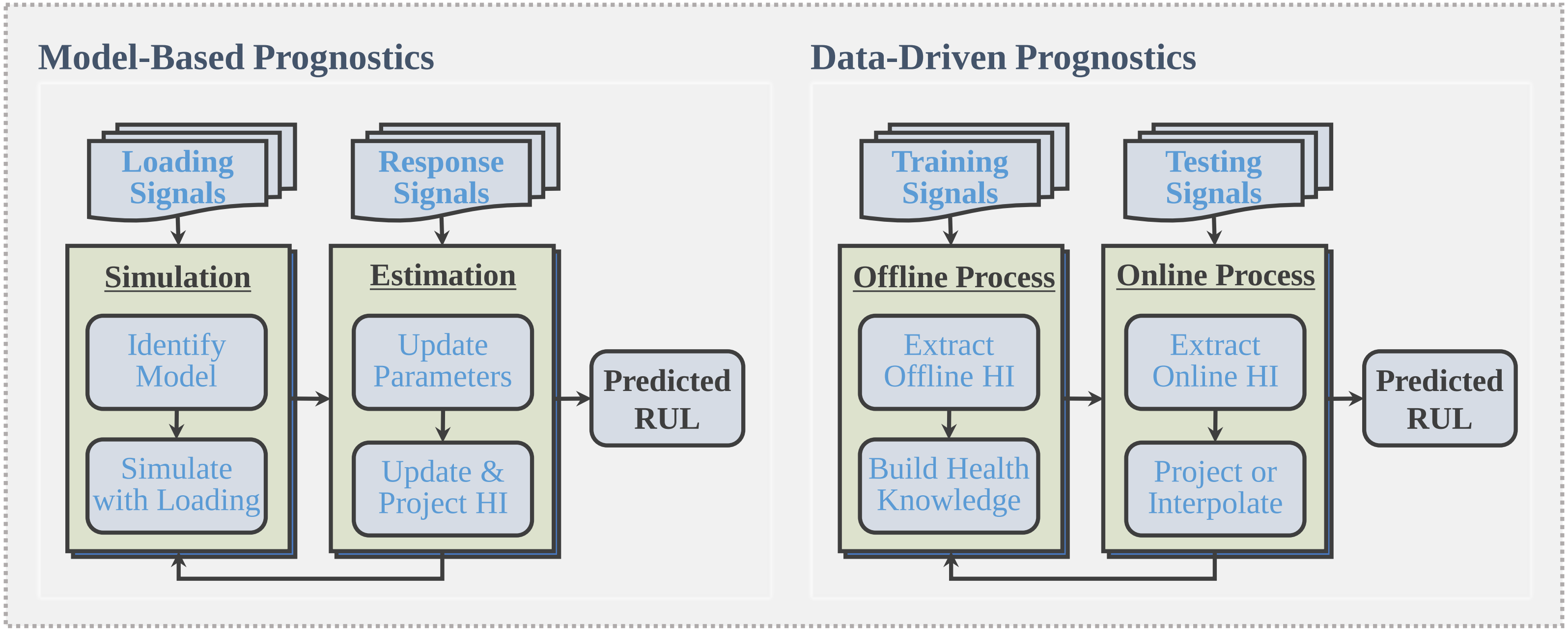
- **Prognostics Procedure**





Introduction

To **predict future health condition and remaining useful lives (RULs)** in real-time



- **Pros:** Assess RUL in early stages
  Require less failure data
- **Cons:** Require physics model
  Applicable for component level
- **Examples:** *PoF*-based models*
  Bayesian updating approaches
  Kalman/Particle Filter

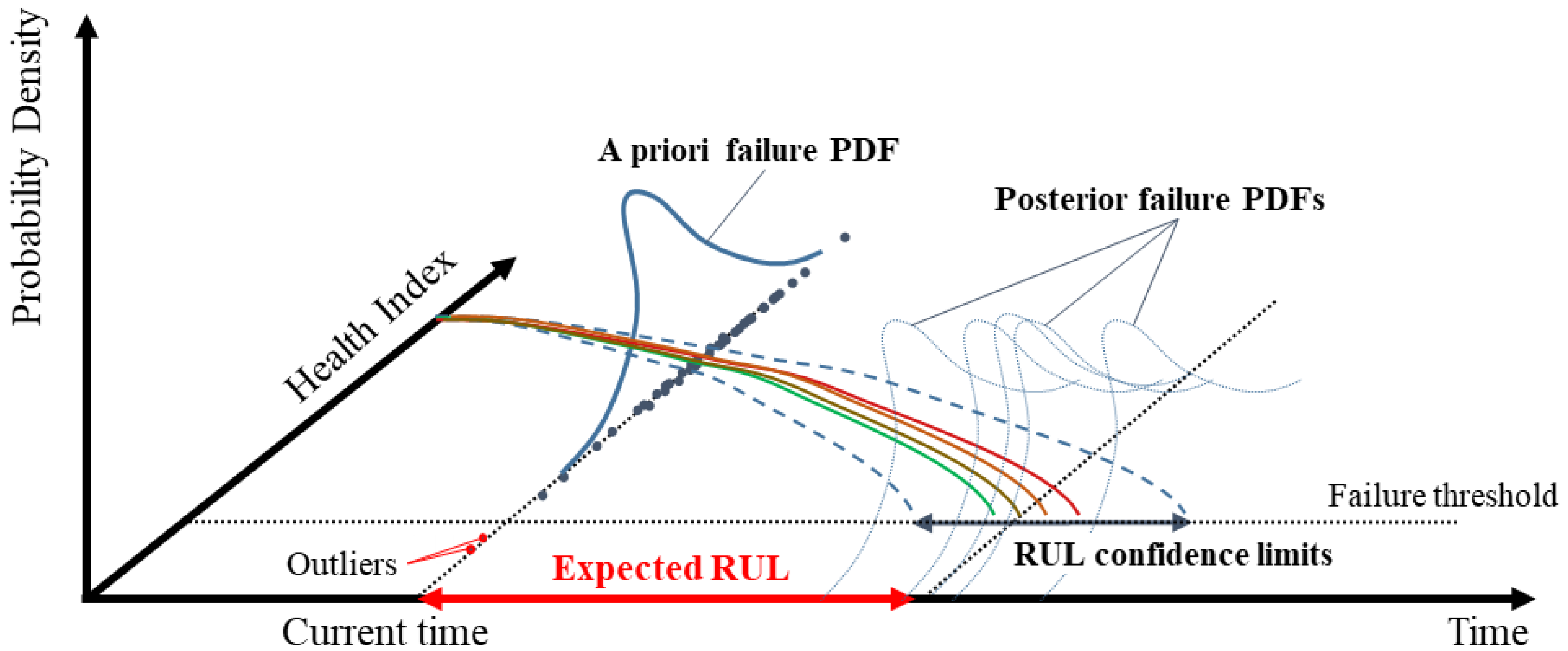
- **Pros:** Don't require assumptions about model
  Applicable to complex systems
- **Cons:** Require large amount of data
  Require heavy computational load
- **Examples:** *Interpolation-based approach*
  Extrapolation-based approach
  Machine learning-based approach

*Physics-of-Failure

Introduction

Remaining Useful Life Probability Density Function

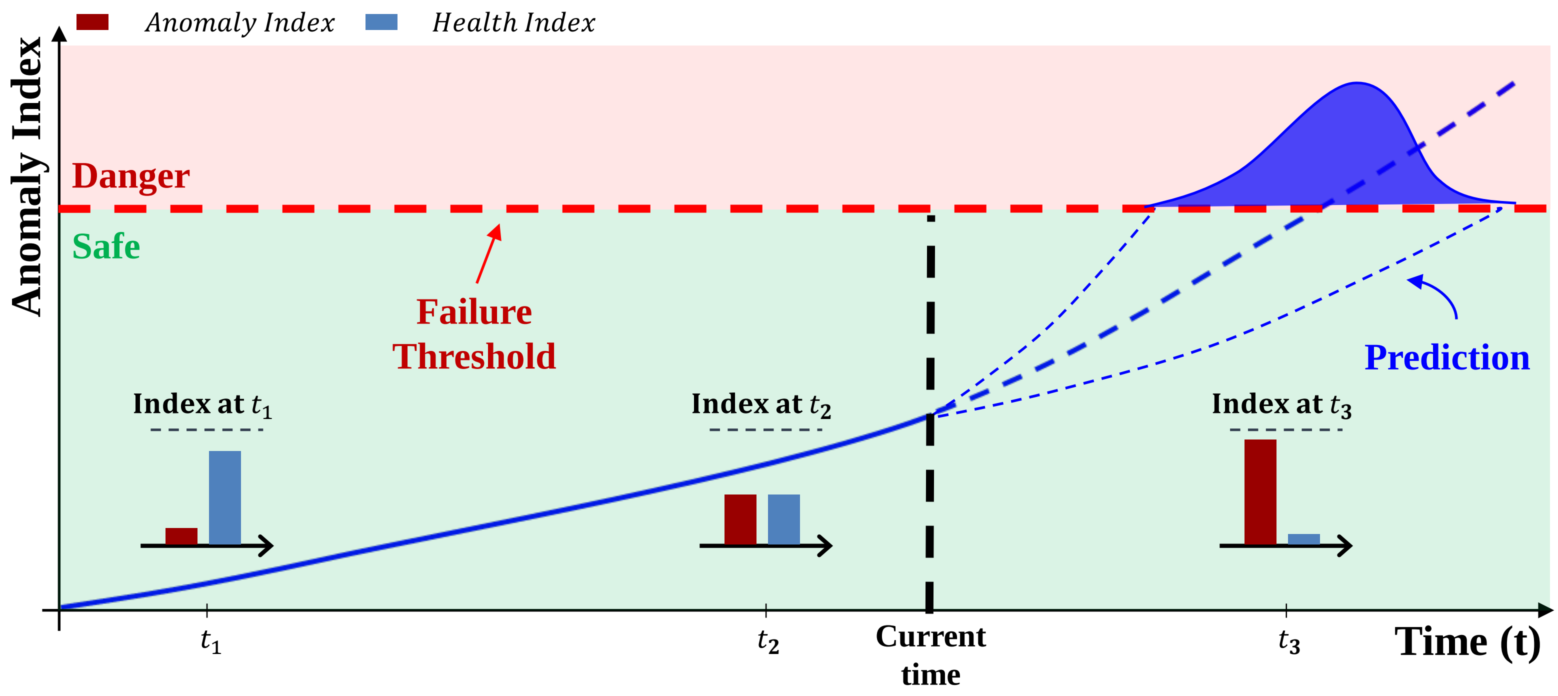
- Health prognostics function predicts the time remaining before the fault progresses to an unacceptable level, in other words, the **remaining useful life (RUL)**
- Depending on how uncertainty is handled in the prediction process, machine health can be regarded as probability distribution, degradation can be regarded as evolution of distribution



Introduction

Remaining Useful Life Probability Density Function

- Health prognostics function predicts the time remaining before the fault progresses to an unacceptable level, in other words, the **remaining useful life (RUL)**
- Depending on how uncertainty is handled in the prediction process, machine health can be regarded as probability distribution, degradation can be regarded as evolution of distribution



Model-Based Prognostics

PoF-based models

- PoF-based models have been investigated to capture various degradation phenomena in engineered systems
- Defect (e.g., cracks and anomalies) initiation and propagation can be derived by using principles of physics

Updating PoF-based models by Bayesian approach

Sensor data contain rich information about system degradation behavior, and model-based prognostics incorporates new sensor information to update PoF-based models. Among the various approaches available to incorporate these evolving sensor data, Bayesian updating is the most widely used

- **Iterative Bayesian updating approaches**
  - Commonly used simulation approaches include iterative Markov Chain Monte Carlo (MCMC) methods (e.g., Metropolis-Hastings and Gibbs Sampling)
- **Non-iterative Bayesian updating approached**
  - Bayesian updating with analytical methods (e.g, importance sampling and rejection sampling)

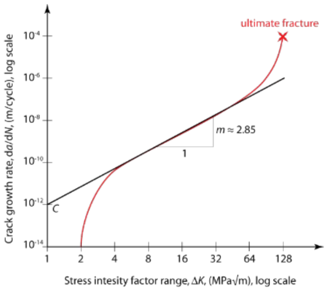
Model-Based : PoF-Based Models

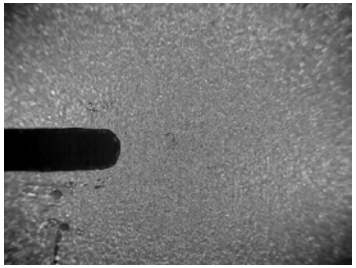
Physics of Failure Models for Prognostics

- PoF-based models, clearly of interest from the prognosis viewpoint, are separated into two major categories; 1) deterministic models 2) stochastic models
- Variations of available deterministic damage propagation models are based on physical law (ex. Paris' law, Fick's law)
  - Paris' formula (fatigue crack propagation model)

$$\frac{d\alpha}{dN} = C_0(\Delta K)^n$$

where α = instantaneous length of dominant crack
N = running cycles
C_0, n = material dependent constants
ΔK = range of stress-intensity factor over one loading cycle





- Stochastic degradation models consider all parameters as random quantities thus the resulting degradation equation is a stochastic differential equation.
  - Cumulative damage model

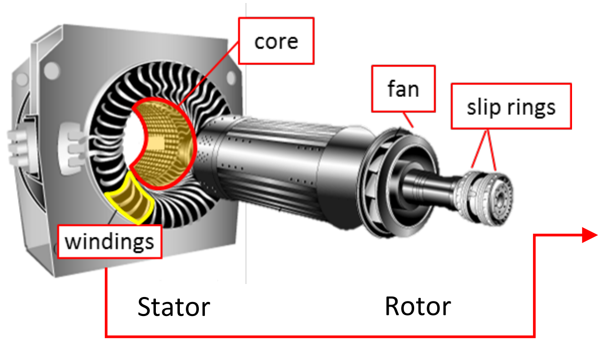
$$c(X_{n+1}) = c(X_n) + D_n h(X_n) \longrightarrow \int_0^t \frac{1}{h(X_u)} \mathrm{d}c(X_u) = \int_0^t dD_u = D_t - D_0$$

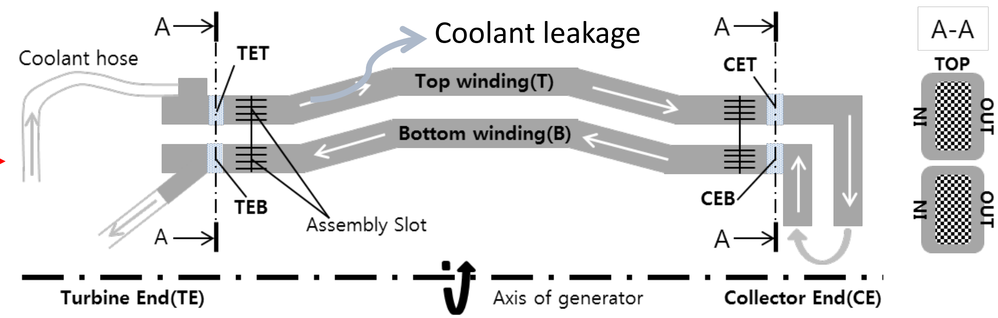
where $c(\cdot)$ = damage accumulation function
X_n = cumulative damage after n
D_n = damage incurred at the (n+1)st increment
$h(\cdot)$ = damage model function

Model-Based : PoF-Based Models

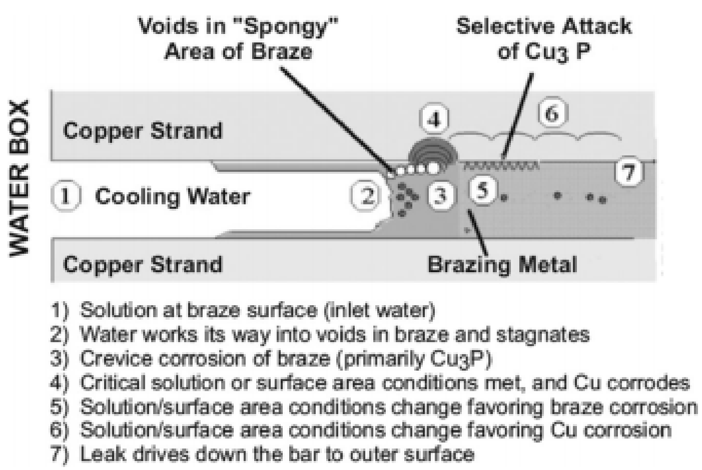
Example - Health prognosis for power generator stator windings (1/2)

- Overview of stator windings





- Crevice corrosion mechanism of stator winding



Capacitance measurement

$$\boxed{C} = \varepsilon_r \varepsilon_0 \frac{A}{t}$$

Capacitance [pF]

ε_r: Relative static permittivity

ε_{water} : 80.4

ε_{mica} : 5.6 - 6.0

ε_0: Electric constant (~ 8.854x10-12 F/m)

A: Area of tester

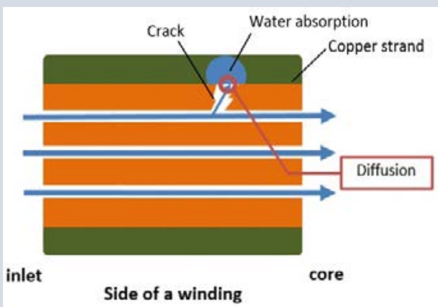
t : Thickness of insulation

Model-Based : PoF-Based Models

Example - Health prognosis for power generator stator windings (2/2)

- Fick's Law of diffusion describes the time course of the transfer of a solute between two compartments
- Calculation of concentration change of water impregnated into insulator over time using Fick's second law

Failure analysis



Fick's second law

$$\frac{\partial m}{\partial t} = D\,\frac{\partial^2 m}{\partial x^2}$$

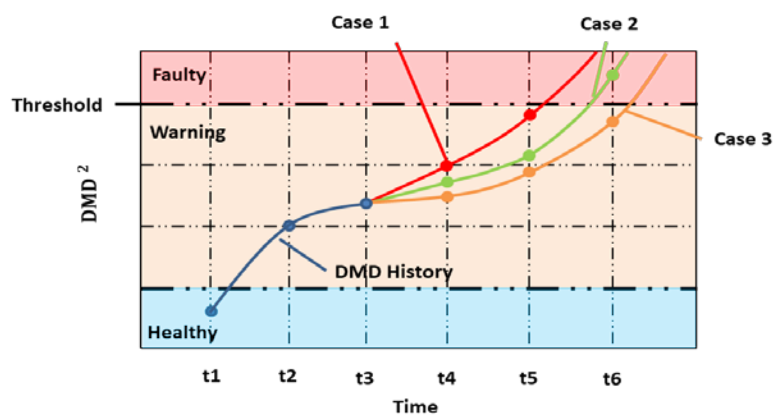
m = concentration of water in insulator
D = diffusion coefficient
x = position in sample

Analytical model

$$m = \left(1 - \exp\left[-7.3\left(\frac{D(t-t_i)^{0.75}}{h^2}\right)\right]\right) m_\infty$$

m_∞ = Concentration of water at steady state
t_i = Time when water absorption happened
h = Thickness of insulation

RUL prediction



DMD: Directional Mahalanobis Distance

Model-Based : Iterative Bayesian Updating Approaches

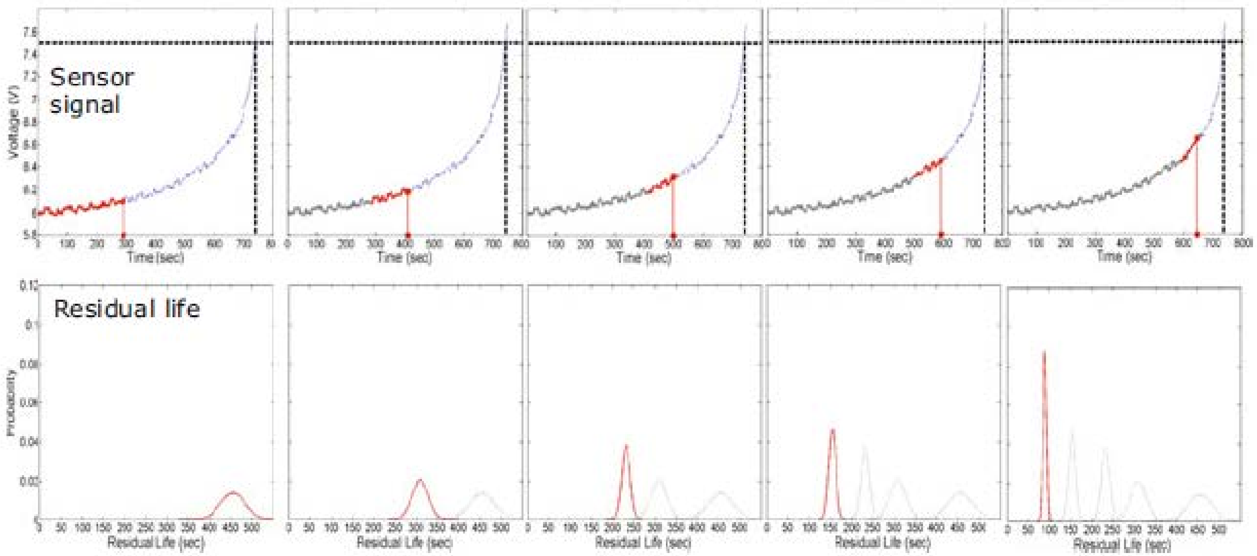
- Bayesian updating with simulation approach
- Commonly used simulation approaches include iterative Markov Chain Monte Carlo (MCMC)
  - e.g, Metropolis-Hastings, Gibbs Sampling

$$S(t_i) = S_0 + \delta \cdot \exp\left(\alpha t_i^2 + \beta t_i + \varepsilon(t_i) - \frac{\sigma^2}{2}\right)$$

where $S(t_i)$ = degradation signal at time t_i

δ, α, β = stochastic model parameters representing the uncertainty of generator operating conditions

ε = random error term modeling possible sensor noise that follows a zero-mean Gaussian distribution with std. deviation σ



Updating of a degradation model and RUL distribution

Model-Based : Non-Iterative Bayesian Updating Approaches

- Bayesian updating with analytical approaches
  Ex) Rejection sampling, Importance sampling→**Particle Filter**
- Consider a dynamic time sequential system with **probabilistic state space model**

$$\text{Transition : } \mathbf{x}_i = f(\mathbf{x}_{i-1}) + \mathbf{u}_{i-1}$$
$$\text{Measurement : } \mathbf{y}_{i+1} = g(\mathbf{x}_i) + \mathbf{v}_i$$

Here $\mathbf{x}_i$ is the vector of (hidden) system states at time $t_i = i \cdot \Delta t$, Δt is a fixed time step between two adjacent measurement points, and *i* is the index of the measurement time step, respectively; $\mathbf{y}_i$ is the vector of system observations (or measurements); and $\mathbf{u}_i$ is the vector of process noise for the states; $\mathbf{v}_i$ is the vector of measurement noise; and $f(\cdot)$ and $g(\cdot)$ are the state transition and measurement functions. It is aim to infer the system states **x** from noisy observations **y**

	Kalman filter	Particle filter
State space model	Linear Model $\mathrm{x_i} = Fx_{i-1} + \mathrm{u_{i-1}}$ $\mathrm{y_{i+1}} = Gx_i + \mathrm{v_i}$	Non-Linear Model $\mathrm{x_i} = f(x_{i-1}) + u_{i-1}$ $\mathrm{y_i} = g(x_i) + v_i$
Noise type	Gaussian, Unimodal	Any distribution, Multimodal
Solution	Solving exact solution (linear-Gaussian model)	Approximate solution (Importance Sampling)
Computational speed	Fast	Slow

Model-Based : Kalman Filter

Example – Battery Health Prognosis (1/2)

- Overview of Battery Health Management

Legend

Voltmeter

Amperemeter

Thermocouple

100%

0%

Charge level

Sensing

- Voltage
- Current
- Temperature

Reasoning/Prognostics

- State of charge (SOC)
- State of health (SOH)
- State of life (SOL)

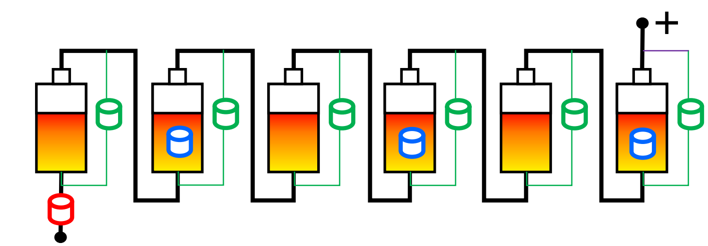
Restoration action

- Decision making
- Cell balancing
- Cell replacement

Passive capacity: C_{dis} = 20%, C_{char} = 10%

- ***Reduced charge & discharge capacity of pack***

Adaptive capacity: C_{dis} = 60%, C_{char} = 40%

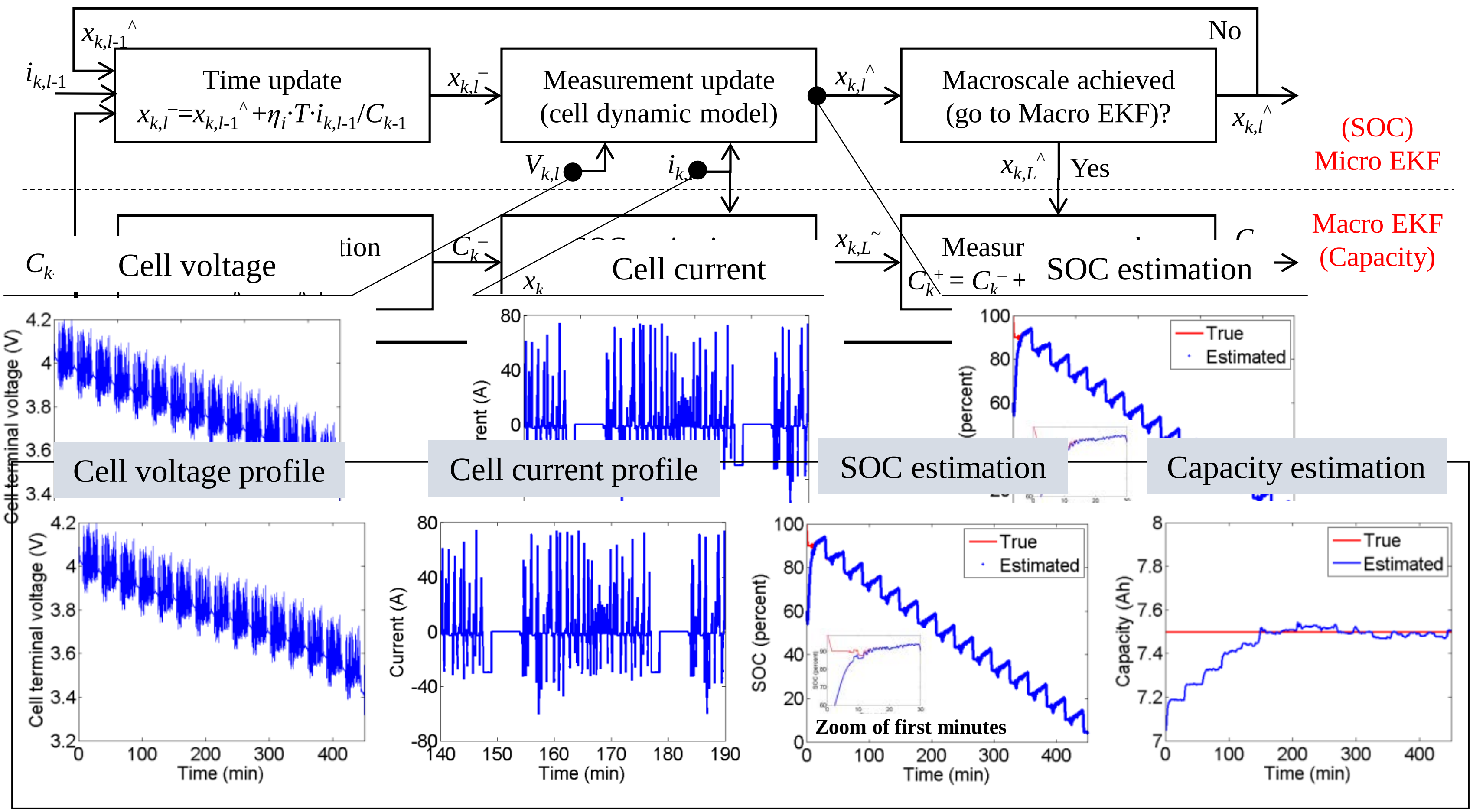


- ***Maximized charge & discharge capacity of pack***

Model-Based : Kalman Filter

Example – Battery Health Prognosis (2/2)

- Multiscale Extend Kalman Filter(EKF) for SOC & Capacity Estimation

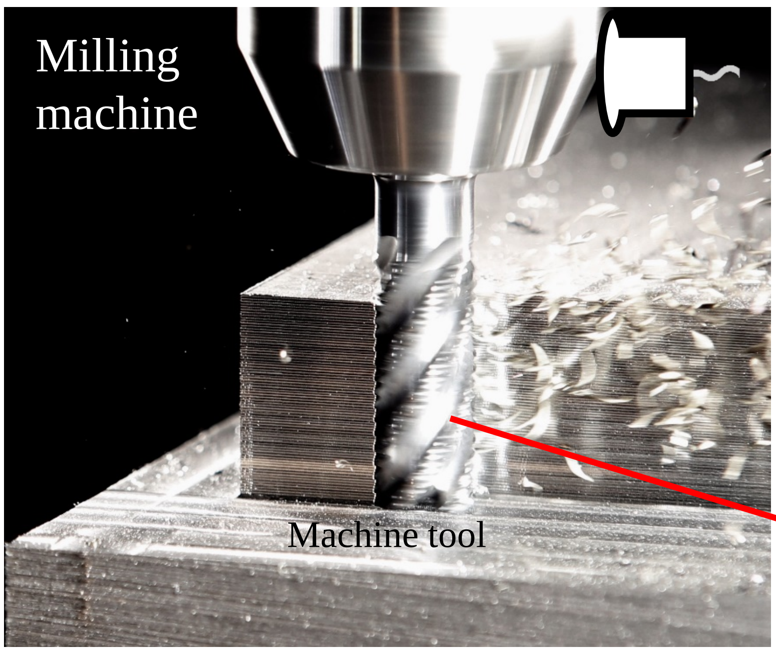


Model-Based : Particle Filter

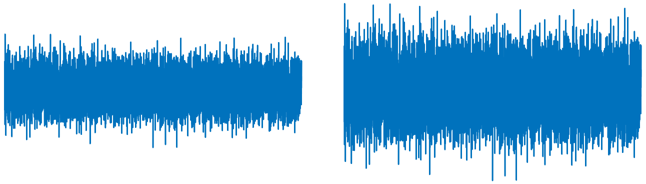
Example – Machine Tool Prognosis (1/2)

- 20% of machine downtime from the failure of machine tools
- Indirect measurement of wear from the sensors such as force, vibration, and so on
- Non-linear process of tool wear growth with non-Gaussian noise

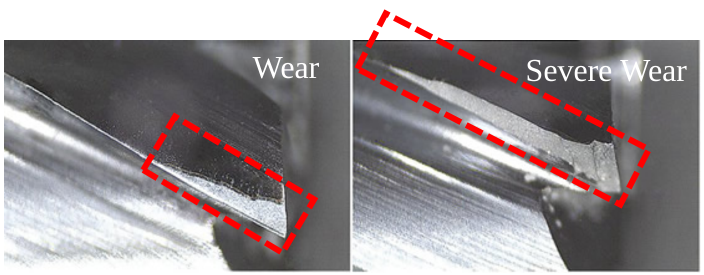
Accelerometer



Indirect measurement of wear with vibration



Direct measurement of wear



System Health & Risk Management

Model-Based : Particle Filter

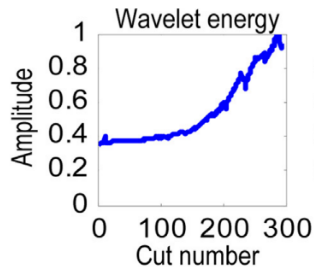
Example – Machine Tool Prognosis (2/2)

- RUL prediction with wear growth model and particle filter

Wear growth model

$$\frac{dx}{dt} = Cx^m$$

x: wear width
C, m: model constant



discretization

Measurement

$$z_k = x_k + v_k$$

v_k: measurement noise
z_k: feature from signals

System model

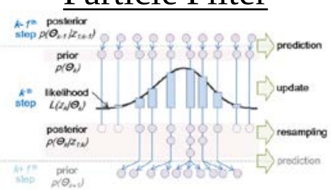
$$x_k = C_{k-1} x_{k-1}^{m_{k-1}} dt + x_{k-1}$$

x_k: wear state
C_k, m_k: uncertain constant

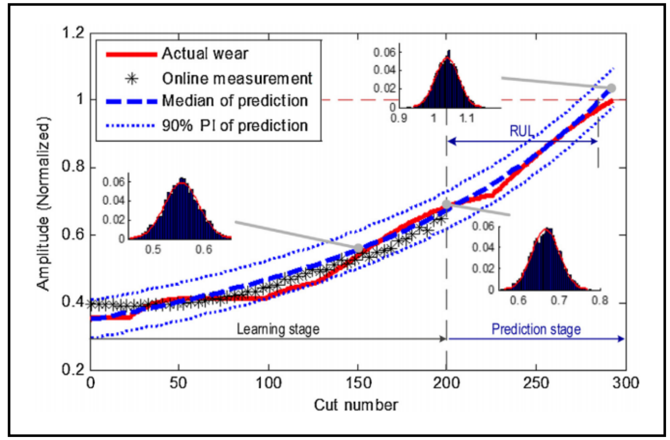
$p(x_k|z_k)$

$p(x_k|x_{k-1})$

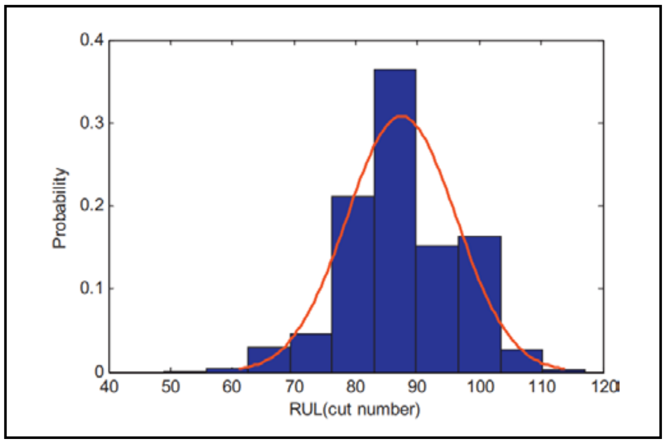
Particle Filter



Wear Tracking

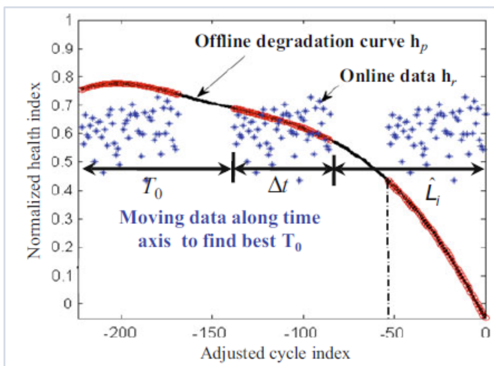


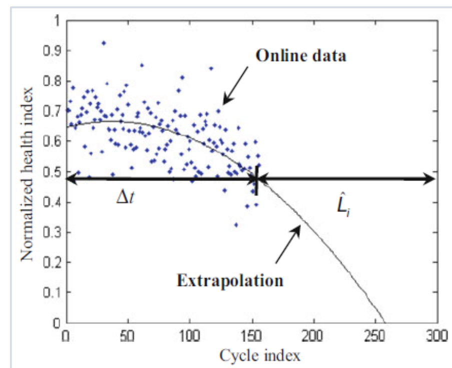
RUL Prediction

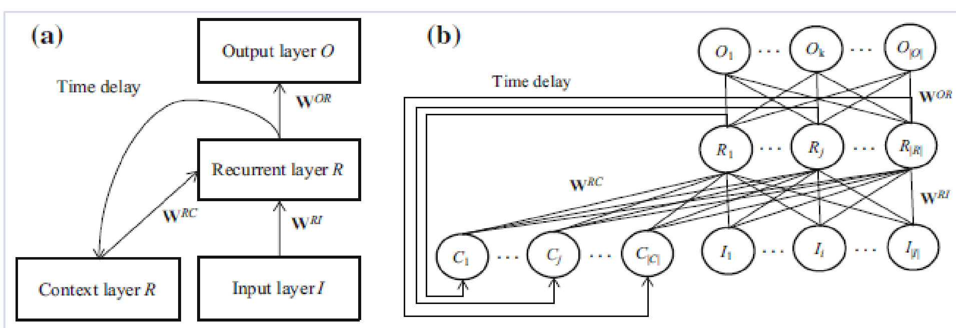


Data-Driven Prognostics

- Data-driven prognostic techniques utilize monitored operational data related to system health
- The major advantage of data-driven approaches is that they can be deployed more quickly and often at a lower cost, as compared to other approaches
- In addition, data-driven techniques can provide system-wide coverage
- Three approaches can be used for online RUL prediction in data-driven approaches; interpolation, extrapolation, and machine learning







Data-Driven : Interpolation-based approach

Similarity-Based Interpolation

- Background health knowledge model using relevance vector machine (RVM)

$$h(t)=\sum_{i=1}^{N}\omega_i\phi(t-t_i)+\varepsilon(t) \quad \text{or} \quad \mathbf{h}=\mathbf{\Phi}\cdot\boldsymbol{\omega}+\boldsymbol{\varepsilon}$$

where ϕ is a kernel density function

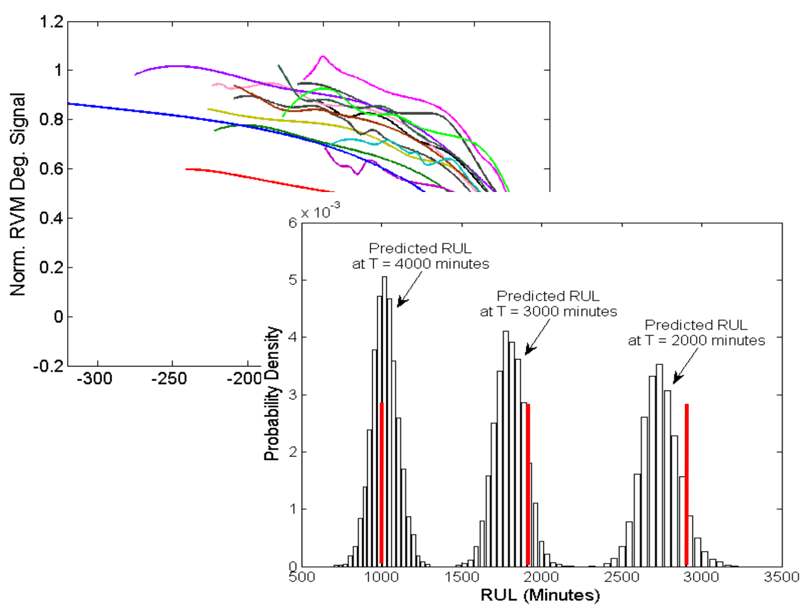
- Conditional remaining useful life (CRUL) from each $h_i(t)$
- Given: CRULs (CRUL_i) from background knowledges
- To find the similarity weights (W_i) for all RULs

Similarity weights

$$W_i=\left[\sum_{j=1}^{N}\left(h_i(t_j)-h_i^p(t_j)\right)^2\right]^{-1}$$

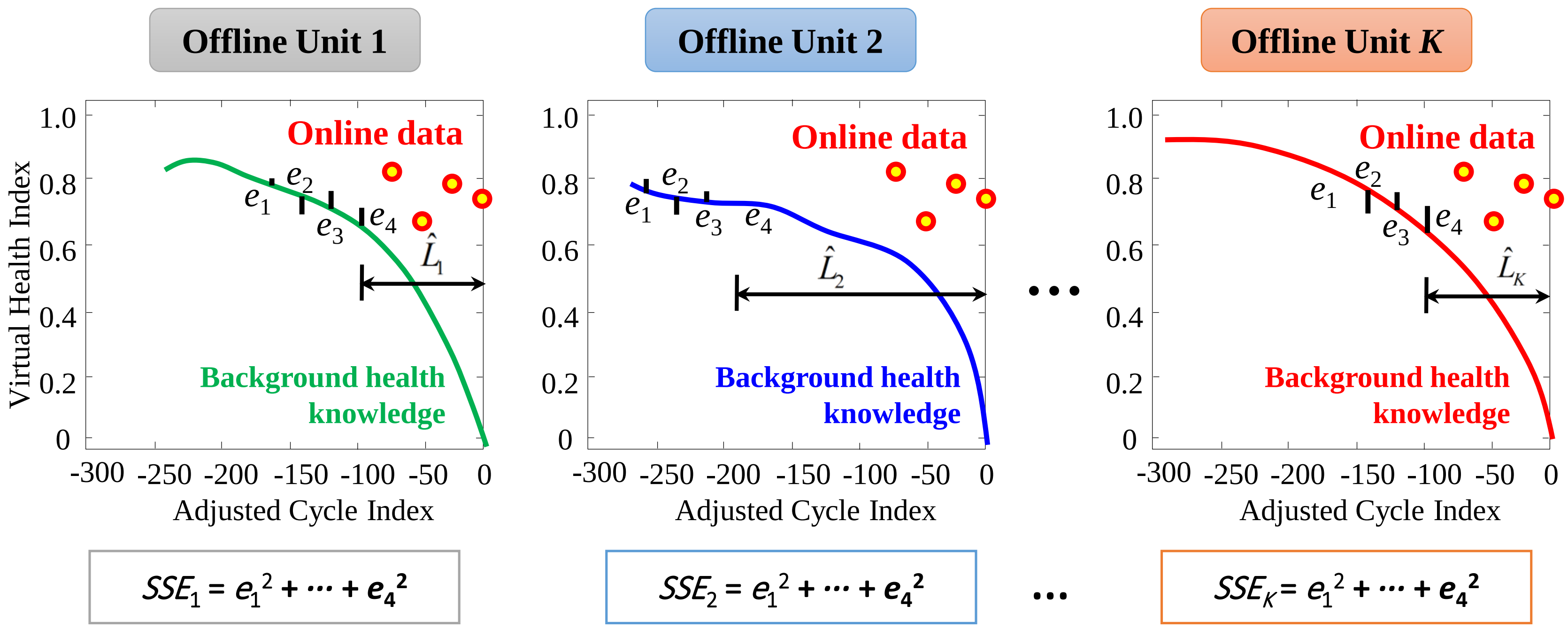
Remaining useful life (RUL) of an online unit

$$RUL=\frac{1}{W}\sum_{i=1}^{K}\left(W_i\cdot CRUL_i\right) \quad \text{where} \quad W=\sum_{i=1}^{K}W_i$$



Data-Driven : Interpolation-based approach

Similarity-Based Interpolation



Initial health state determination

Similarity-based interpolation

$$\hat{L} = \frac{1}{W}\left(W_1 \cdot \hat{L}_1 + W_2 \cdot \hat{L}_2 + \cdots + W_K \cdot \hat{L}_K\right), \quad W_i = SSE_i^{-1}$$

Data-Driven : Extrapolation-based approach

Bayesian linear regression

$$h(t) = b_1 t^2 + b_2 t + b_3$$

Prior for linear model parameters

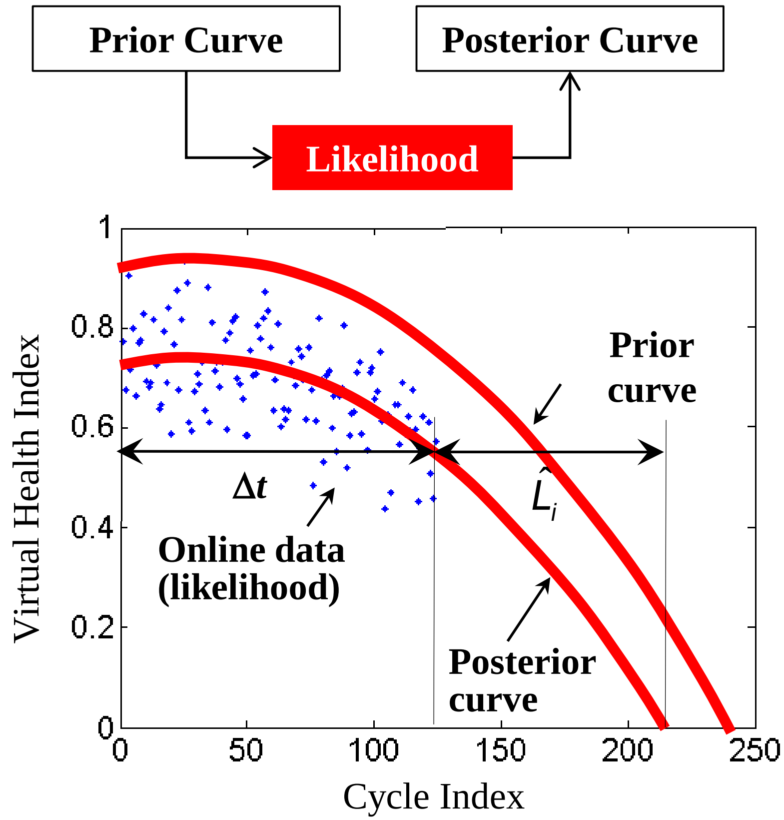
$$b_j \sim N(\mu_j, \sigma_j)$$

Offline fitting

Posterior for Bayes linear model parameters

$$\mathbf{b} = (\Phi^T \Sigma^{-1} \Phi)^{-1} \Phi^T \Sigma^{-1} \mathbf{h}$$

Φ: Online design matrix



Data-Driven : Extrapolation-based approach

Example –Power Transformer Prognosis (1/2)

- Classified health grades **without failure data** (unsupervised) in a **statistical** manner

$$\underbrace{h(x_1,x_2)}_{\text{Health condition}} = 1 - \underbrace{F_1(x_1)}_{\text{Marginal CDFs}} - \underbrace{F_2(x_2)}_{} + \underbrace{C\big(F_1(x_1),F_2(x_2)\big)}_{\text{Joint CDF (copula)}}$$

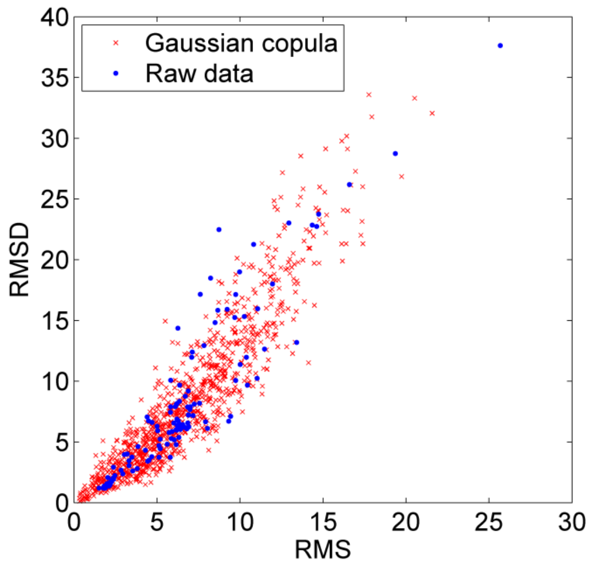
→ **Statistical health grades**

Grade A (healthy): $h > \Phi(-1.0\sigma)$

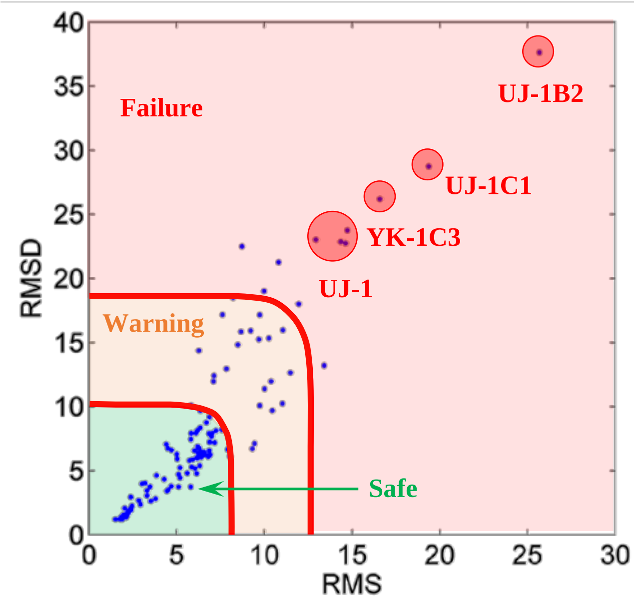
Grade B (warning): $\Phi(-2.0\sigma) \le h < \Phi(-1.0\sigma)$

Grade C (faulty): $h \le \Phi(-2.0\sigma)$

Statistical Modeling (Copula)



Statistical Health Grade



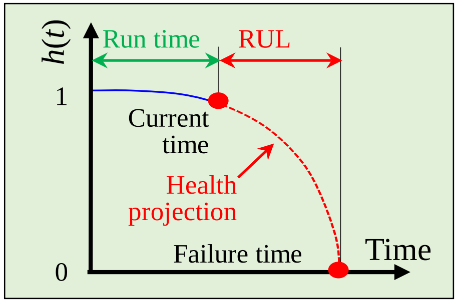
Health Reasoning : Statistical Health Grade System

Data-Driven : Extrapolation-based approach

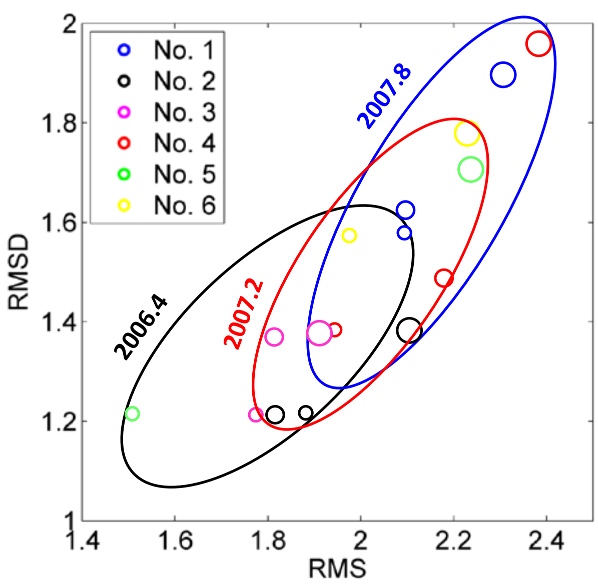
Example –Power Transformer Prognosis (2/2)

- Proved feasibility in health prognostics with limited data obtained in 2 years

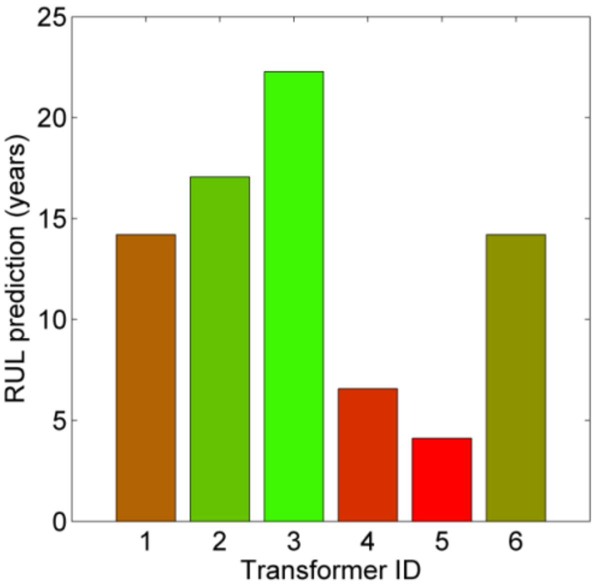
$$h(t) = \underbrace{1 - a}_{\text{Initial health condition}} \cdot \underbrace{\exp(bt)}_{\text{Health degradation}}$$



Health Degradation History



RUL Prediction Results



Data-Driven: Machine Learning-Based Approach

- In contrast to the interpolation- or extrapolation-based approaches, the machine learning-based approach does not involve any visible manipulation →black box model
- It requires the training of a prognostics model using the offline data
- It is capable of learning nonlinear dynamic temporal behavior, e.x, RNN (Chapter 6)

W^{RI} : weights of input and recurrent layer

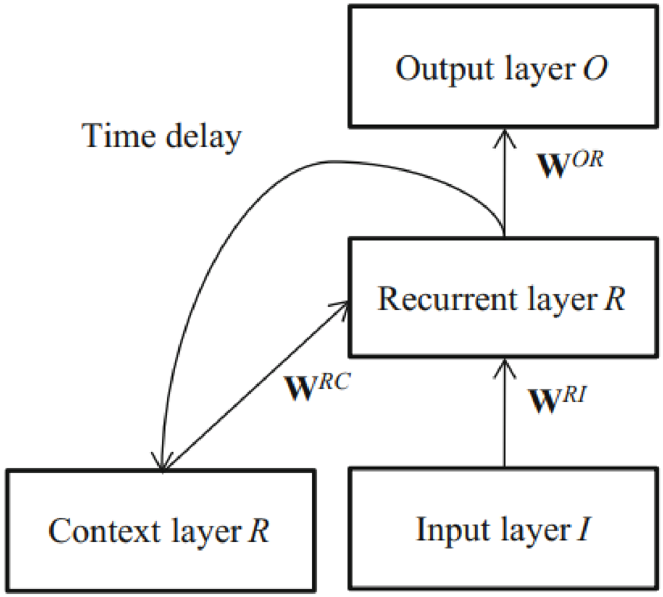
W^{OR} : weights of recurrent and output layer

W^{RC} : recurrent weights

$R^{(t)}$: recurrent units

$R^{(t-1)}$: previous recurrent units

f : activation function(sigmoid, tanh, ReLu)



$$\tilde{R}_i^{(t)} = \sum_j W_{ij}^{RI} I_j^{(t)} + \sum_j W_{ij}^{RC} R_j^{(t-1)}$$

$$R_i^{(t)} = f\left(\tilde{R}_i^{(t)}\right) = \left[1 + epx\left(-\tilde{R}_i^{(t)}\right)\right]^{-1}$$

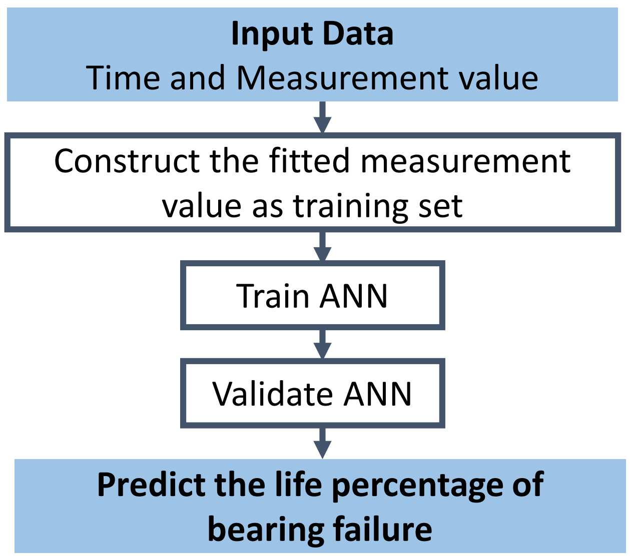
$$\tilde{O}_i^{(t)} = \sum_j W_{ij}^{OR} R_j^{(t)}$$

$$O_i^{(t)} = f\left(\tilde{O}_i^{(t)}\right) = \left[1 + epx\left(-\tilde{O}_i^{(t)}\right)\right]^{-1}$$

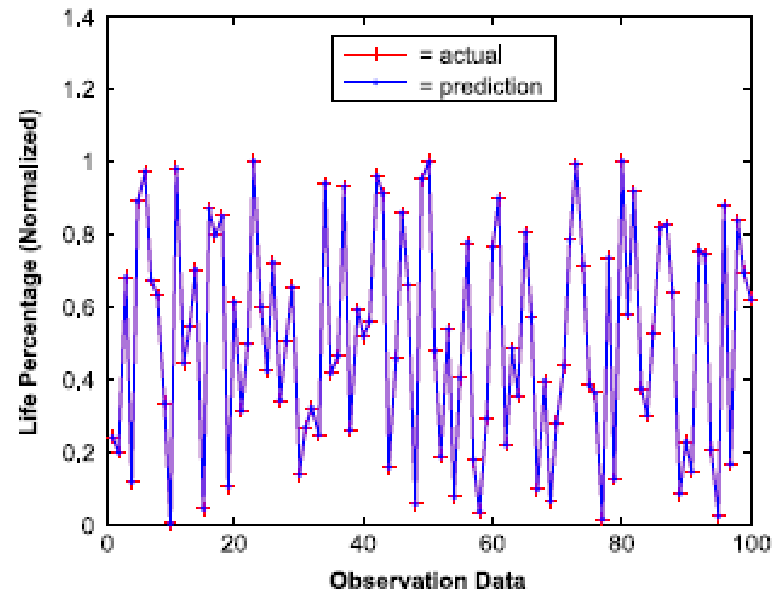
Data-Driven: Machine Learning-Based Approach

Example – (Mahamad et al, 2010)

- Using IMS bearing test bed signal (RMS, kurtosis)
- Input data : 6 nodes
  - Current and previous time
  - RMS at current and previous time (fitted with Weibull hazard rate function)
  - Kurtosis at current and previous time (fitted by Weibull hazard rate function)
- Output : life percentage (normalized)



Flow chart



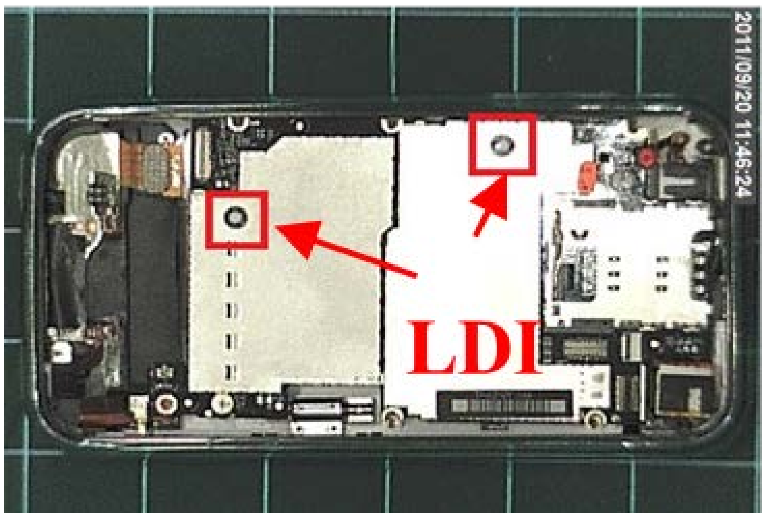
The output performance

Case Study: LDI (Model-Based)

Overview of Liquid Damage Indicator (LDI) Prognostics (Oh, et al. 2015)

- Liquid Damage Indicator / Liquid Contact Indicator
  - A small indicator that turns from white into another color, typically red, after contact with water
  - Numerous complaints are reported regarding the performance of LDIs
  - Several law suits were filed for denying warranty service to customers based on inaccurate LDIs (ex. iPhone 3G)

- Objective
  - To devise an deficient scheme for evaluating the performance of LDIs
  - To develop a performance degradation model

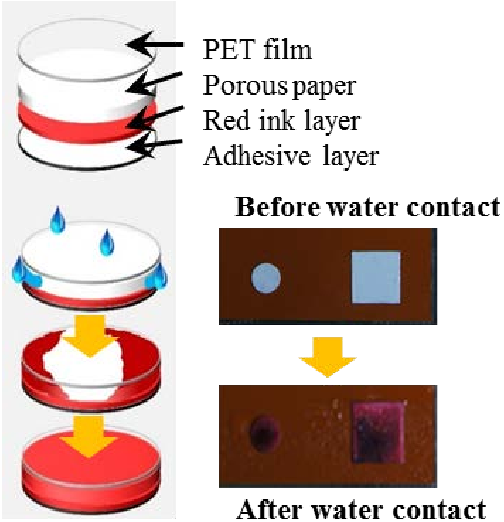


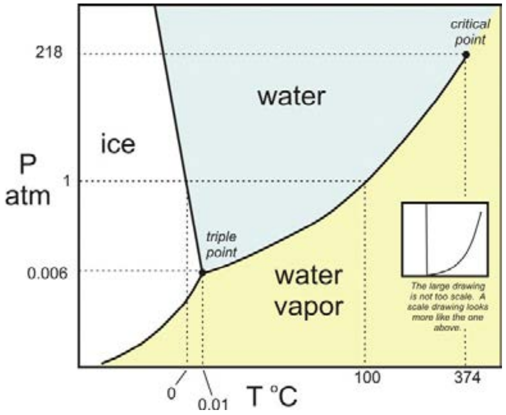
Case Study: LDI (Model-Based)

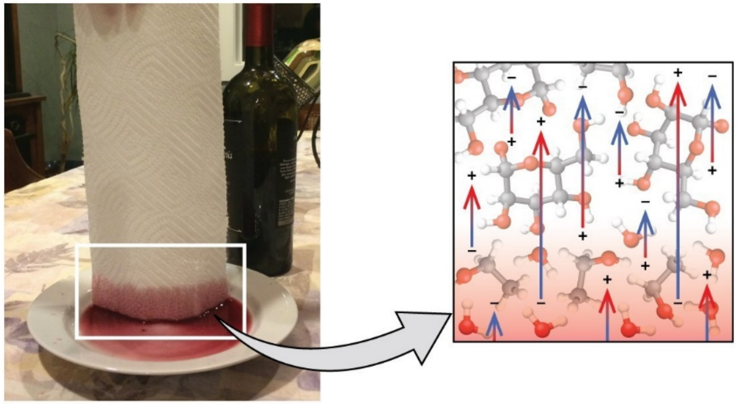
Two Main Physics Mechanisms

- Phase change from vapor to water
  - $V \propto (\Delta T)^m$
  - V is the volume of condensed water after phase change
  - ΔT is the temperature difference
  - m is the model constant

- Water transport in the paper (capillary action)
  - $x \propto n^k$
  - x is the distance of penetration
  - n is the number of thermal cycles
  - k is the model constant







Case Study: LDI (Model-Based)

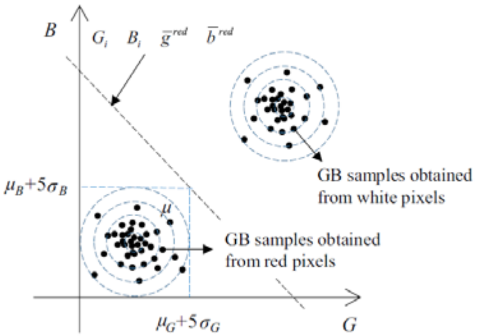
Accelerated Life Test for LDIs

- Step 1: Determination of test conditions
  - According to technical data, the operational condition is -40°C to 60°C and highly humid condition (95% RH at 55°C)
  - The use conditions for portable electronics with a use temperature range of -20°C to 45°C, and a relative humidity of 5-95%

Test no.	Chamber 1 (°C)	Chamber 2
1	-30	25°C, 95% RH
2	-25	
3	-20	1) LDIs on the FPCB substrate 2) LDIs on the glass substrate
4	-15	

- Step 2: Steps for life tests
  - I. Execute a 30-min test for a sample in chamber 1
  - II. Execute a 5-min test in chamber 2 as soon as the sample is taken out of chamber 2
  - III. Take a picture of the sample after II under a predefined light and angle condition
  - IV. Repeat I-III until the sample experiences 50 cycles or the LDI turns entirely red.

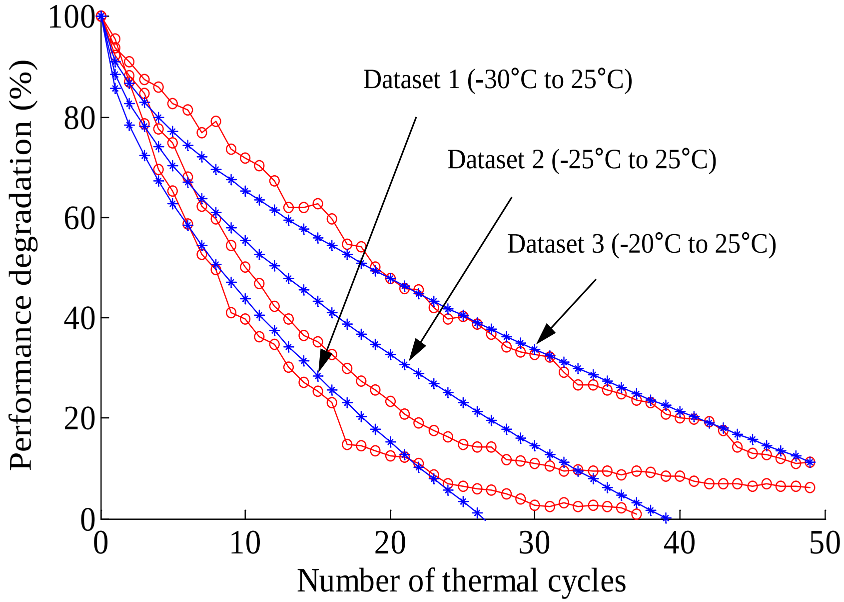
- Step 3: Quantification of performance degradation
  - A pixel = $\begin{cases} \text{white pixel,} & if\ G_i + B_i > \bar{g}^{red} + \bar{b}^{red} \\ \text{red pixel,} & otherwise \end{cases}$
  - $\bar{g}^{red} + \bar{b}^{red} = \mu_{G^{red}} + \mu_{B^{red}} + 5\{\sigma_{G^{red}} + \sigma_{B^{red}}\}$
  - G_i and B_i denote the G and B values of the ith pixel
  - $\bar{g}^{red}$ and $\bar{b}^{red}$ indicate the G and B value margins in red

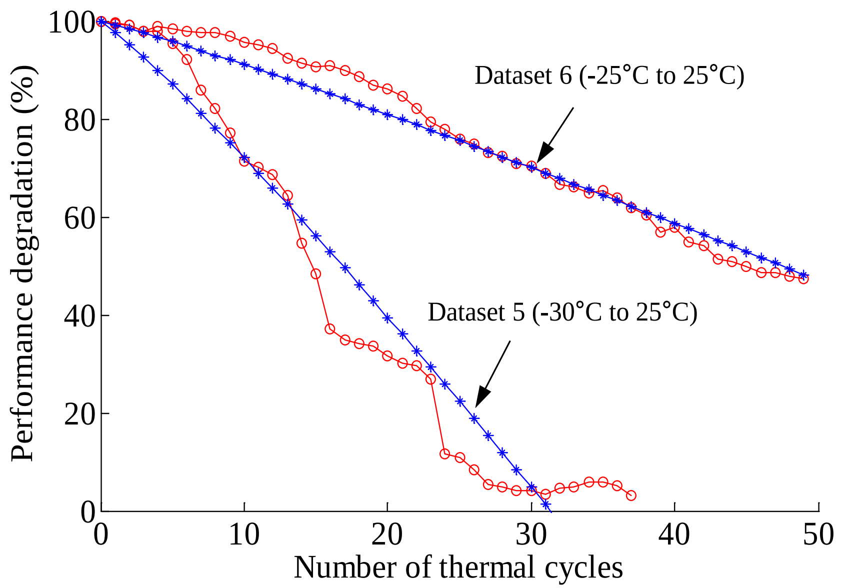


Case Study: LDI (Model-Based)

Degradation Model for LDIs

- A Novel Performance Degradation Model for LDIs
  - $D(n;\Delta T) = 100 - a(\Delta T)^b n^c$
  - D is the index that represents the performance degradation for LDIs (%)
  - a, b and c are the model constant





Case Study: LDI (Model-Based)

Validation of Proposed Model

- The cycles to failure was calculated using the proposed model.
- Two actual iPhone 3G were tested between -15 °C and 25 °C with 95% relative humidity.
- The **amount of error** in the prediction is **reasonable** considering inherent randomness in the specimen and measurement error.

$$D\left(n;\Delta T\right)=100-a\left(\Delta T\right)^{b}n^{c}$$

a = 0.005312
(lower bound, upper bound) =
(0.0008927, 0.009731)
b = 2.015
(lower bound, upper bound) =
(1.81, 2.22)
c = 0.5199
(lower bound, upper bound) =
(0.4886, 0.5512)

Predicted cycles to failure (CTF): 21 cycles

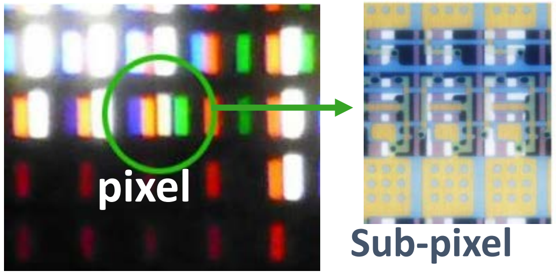
Actual CTF of the two iPhones: 4 and 16 cycles

Case Study: OLED (Model-Based)

Overview of Organic Lighting-Emitting Diode (OLED) Prognostics (Kim, et al. 2017)

- Issues of OLED Prognosis
  - The organic light-emitting diode (OLED) technology are more visual compelling and power efficient than liquid-crystal displays (LCDs)
  - OLED TV with layered structure and materials is subject to a great deal of manufacturing and operational uncertainties
  - Light-emitting layer and TFT are the major contributors to the degradation of OLED TV, which are correlated in a complicated manner

- Objective
  - To propose a reliable lifetime model of large OLED panels that incorporates manufacturing & operational uncertainty under various usage condition
  - To develop an effective scheme that predicts OLED TV reliability accurately and efficiently at an early product development stage.

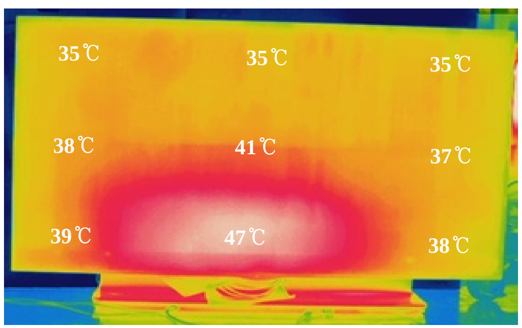


Case Study: OLED (Model-Based)

Failure Modes and Failure Mechanism

- OLED TV Panel
  - OLED TV degrades over time by a luminance changes and color shift
  - Light-emitting layer and TFT are the major contributors to the degradation of OLED TV, which are correlated in a complicated manner.

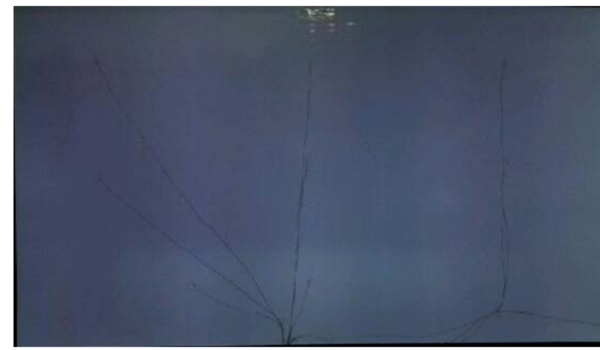
✓ Local heat source
✓ Natural convection



✓ Spatial deviation of temperature
✓ Degradation mechanism of two components – TFT & emissive layer

after long time usage

① Luminance change



② Color shift

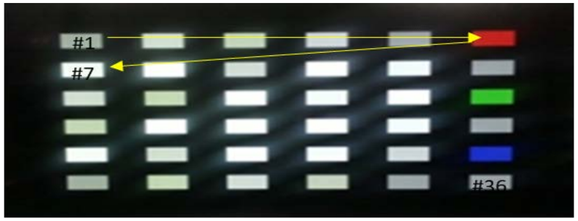
Case Study: OLED (Model-Based)

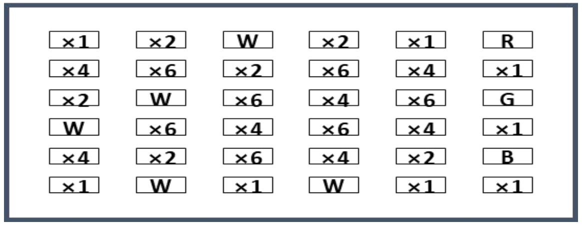
Life Tests for OLED

- Major acceleration factors
  - Ambient temperature and initial luminance intensity
  - Six sets of OLED panels for the accelerated degradation tests (ADTs)
  - Temperature: 3 sets in convection oven (25℃) and other 3 sets (40℃)
  - Luminance intensity: 4 levels of luminance intensity (×1, ×2, ×4, ×6)

- Measurement interval
  - Measurements were conducted at variable intervals between 24 - 180 hrs.
  - until an operating time reached 1,500 hours
  - Out of 36 patterns, R/G/B/W were excluded. As a result, 28 patterns were used for the measurement data

Panel	Temperature condition	Initial luminance intensity (The number of pattern)				Total number of patterns
		×1	×2	×4	×6	(168)
#1	25℃	7	7	6	8	28
#2	25℃	7	7	7	7	28
#3	25℃	7	7	7	7	28
#4	40℃	7	7	7	7	28
#5	40℃	7	7	7	7	28
#6	40℃	7	7	7	7	28

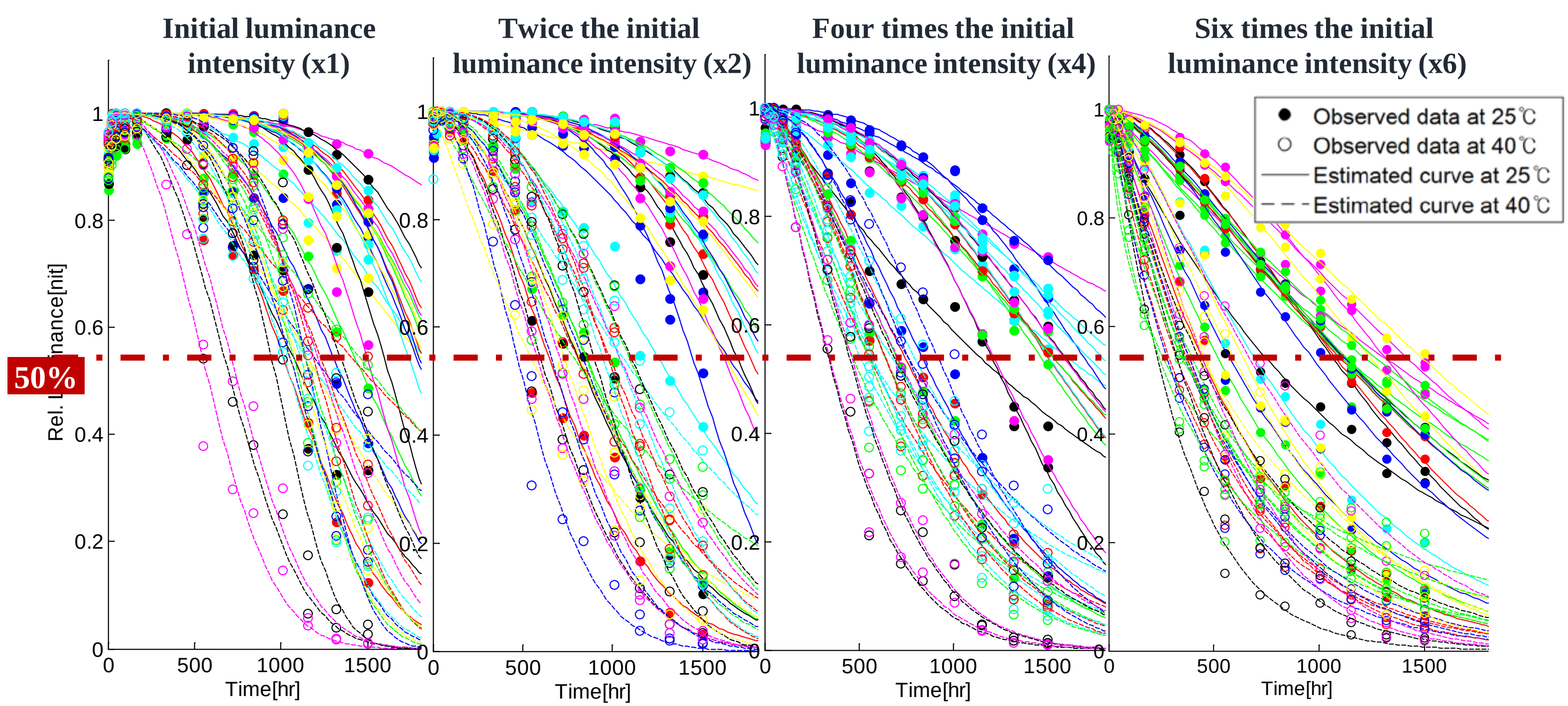




Case Study: OLED (Model-Based)

Degradation Model for OLED

- A Novel Bivariate Performance Degradation Model for OLEDs
  - $MTTF(T, I_{\text{lum}}) = \frac{A}{T} \cdot e^{\frac{B}{kT}} \cdot e^{I_{\text{lum}}(C+\frac{D}{kT})}$
  - By integrating the two lifetime models; Arrhenius equation (temperature), inverse power law (luminance)
  - k is the Boltzmann constant; T is the ambient temperature; I_{lum} is the initial luminance intensity; and A, B, C and $D$are the model parameters c

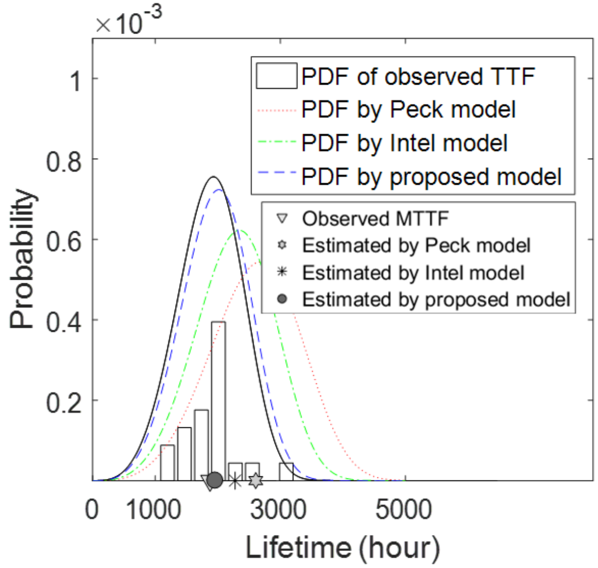


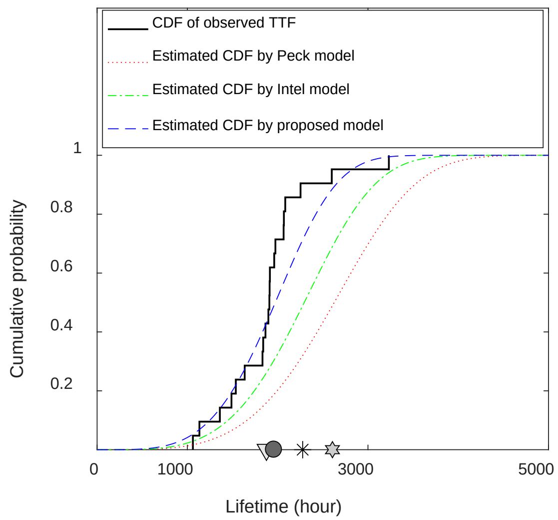
Case Study: OLED (Model-Based)

Validation of Proposed Model

- 21 failure data under use condition employed for validation.
- The MTTF of the 21 failure samples was 1,876 hours, whereas the MTTF estimated from the proposed model was 1,959 hours.

Model	Estimated lifetime		Chi-square GoF test		KS GoF test	
	$MTTF_{obs}{}^{*}$= 1875	Error*	Hypothesis	P-value	Hypothesis	P-value
Proposed	1959	4%	Accept	1.66×10^{-1}	Accept	6.38×10^{-2}
Peck's Model	2607	39%	Reject	8.09×10^{-5}	Reject	5.61×10^{-5}
Intel's Model	2277	21%	Reject	8.77×10^{-4}	Reject	4.72×10^{-5}



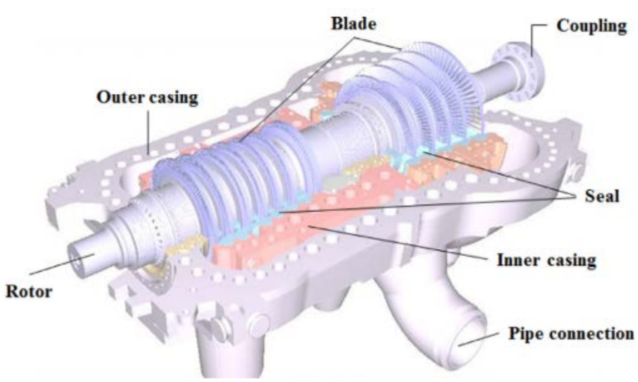


Case Study: Steam Turbine (Model-Based)

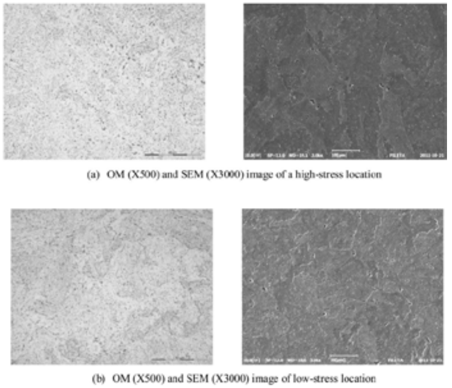
Overview of Steam Turbine Prognostics (Choi, et al. 2018)

- Issues of Steam Turbine Prognosis
  - The design life of steam turbine is typically 25 years of 200,000-250,000 h
  - Premature failure of the power plant machinery is the one of main interests for operators
  - The RUL of key elements could be predicted by metallurgical or theoretical analysis of as-received and degraded element but it is difficult to quantify the health conditions from results

- Objective
  - To quantify the result of the hardness measurement method that is most commonly and easily used in actual field
  - To propose a new damage growth model within the Bayesian statistical framework that can utilize sporadically measured and heterogeneous on-site data from stem turbines



Schematic of a steam turbine



Metallurgical changes of rotor steel after 146,708 h operation

Case Study: Steam Turbine (Model-Based)

FMEA for a Steam Turbine

- Among the components of the steam turbine, HIP rotor has the highest severity and risk
- Creep and low/high cycle fatigue are known to be the dominant mechanisms
- High temperatures and centrifugal force causes creep damage in high-stress regions
- Thermo-mechanical fatigue damage from the the thermal cyclic load causes cracking at the wheel corner

Comp.	Failure cause	Failure mechanism	Failure mode	Occurrence	Severity	Risk
HIP(High-Intermediate Pressure) steam turbine						
Rotor	Temp. cycling	Creep, LCF	Fracture	Not often	Very high	High
HP blade	Temp. cycling	LCF, HCF	Failure	Not often	High	Moderate
HP casing	Temp. cycling	CREEP, LCF	Crack	Not often	Moderate	Moderate
IP blade	Temp. cycling	LCF, HCF	Failure	Not often	High	Moderate
IP casing	Temp. cycling	CREEP, LCF	Crack	Not often	Moderate	Moderate
LP(Low Pressure) steam turbine						
Rotor	Wet. Cycling	Corrosion, LCF	Fracture	Not often	High	High
Blade	Wet. Cycling	LCF,HCF, Corrosion	Failure	Often	Moderate	Moderate
Bearing	Wear	Wear	Vibration	often	Low	Moderate

Case Study: Steam Turbine (Model-Based)

Characteristics of On-site Measurement Data

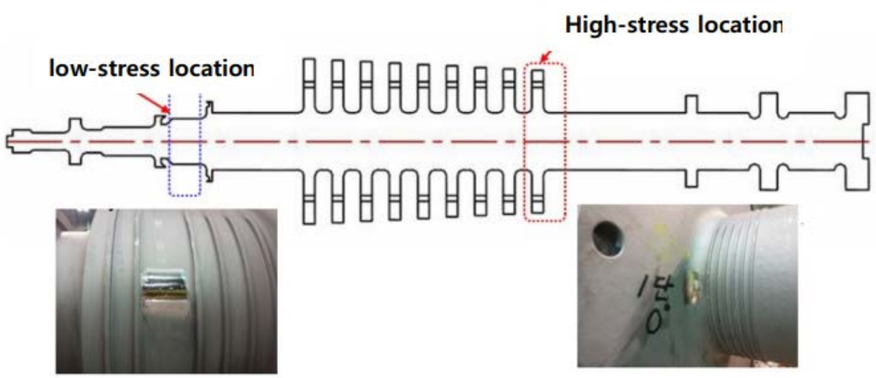
- It is extremely difficult to measure material degradation directly
- Material hardness data was achieved both low-stress and high-stress conditions
- The harness in a low-temperature region can be used as a reference hardness
- Ten sets of the hardness data set were sporadically measured at overhauls over 10 years

Quantitative Damage Index for the Hardness Data

- Hardness Damage index was introduced that takes into account both creep and fatigue damage

$$D = 1 - \frac{\widetilde{H}}{H} = 1 - \frac{H_a}{H_v}$$

where H_a and H_v are the hardness values measured at aged and virgin



System Health & Risk Management

Case Study: Steam Turbine (Model-Based)

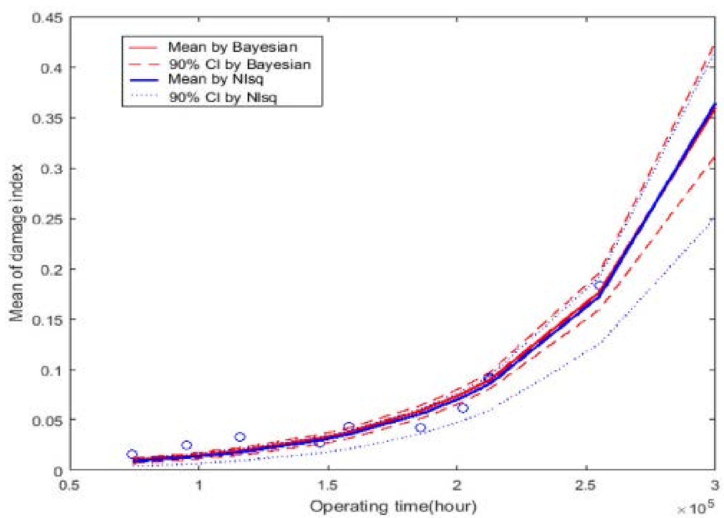
Damage Growth Model using Sporadically Measured and Heterogeneous On-site Data

- New damage growth model that utilizes the hardness damage indices
  - Bayesian inference and Markov Chain Monte-carlo(MCMC) techniques are used to update the parameters
  - The damage growth model can be defined in the form of a distribution as

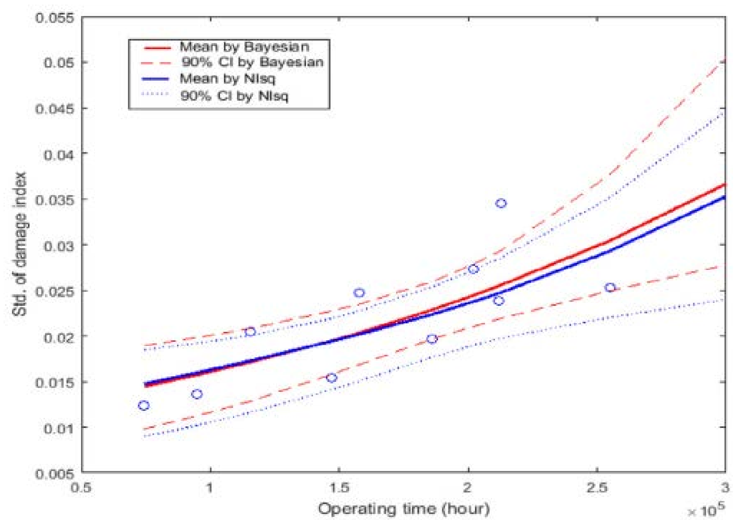
$$D(t) \sim \mathrm{N}\big(\mu_D(t), \sigma_D(t)\big)$$

where $\mu_D(t), \sigma_D(t)$ are the mean and standard deviation of the time-varying damage

  - Metropolis-Hasting algorithm to generate samples that MCMC simulation



(a) Mean



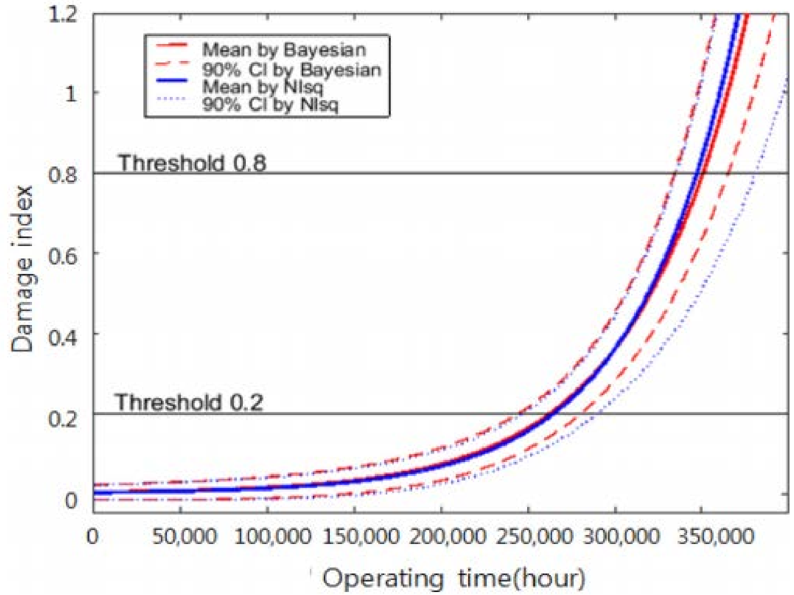
(b) Standard deviation

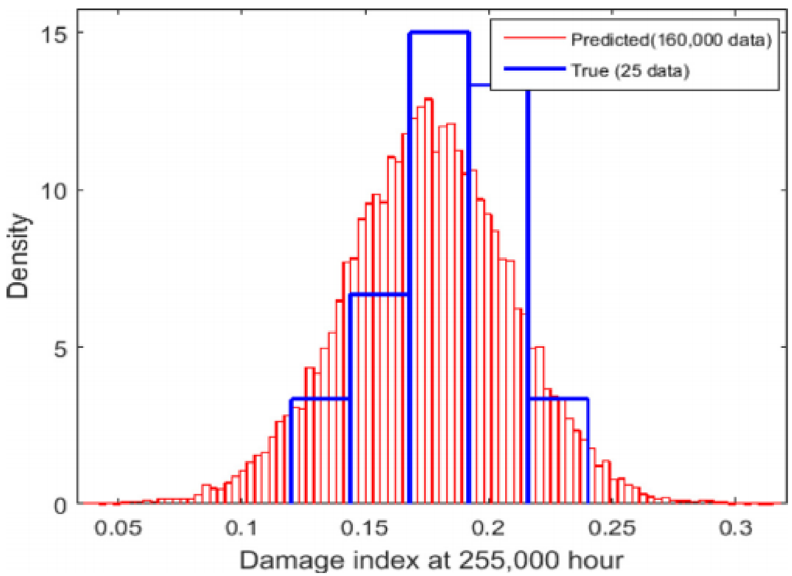
Mean and standard deviation results obtained by performing the Bayesian updating

Case Study: Steam Turbine (Model-Based)

Validation of Proposed Model

- RUL prediction at three operating times (0, 200,000h, and 250,000) to determine an appropriate failure criterion
- Failure criterion of the damage index 0.2 gives a reasonable RUL for steam turbine with the actual retirement history of steam turbines



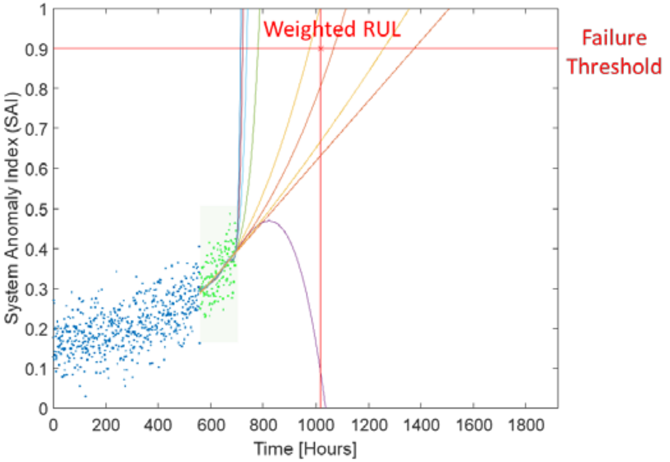


Case Study: Ensemble Method (Data-Driven)

Overview of Ensemble Method

- Limitation of single prognostic algorithm
  - Dependency of the algorithm's accuracy on training data set
    → Number and type of training data affects algorithm's accuracy
  - Weak for variable manufacturing, environmental and operational conditions
    → More robustness for various operating conditions is necessary
  - Difficulty reflecting various types of degradation trend.
    → Each algorithm can produce good results for only appropriate degradation trends

- Combination of multiple algorithms to form a hybrid algorithm
  - To improve the robustness
    → For type of algorithm
    → For operating and environmental conditions
    → For type or number of input data
  - To increase accuracy of algorithm
    → Sum of multiple algorithms' results

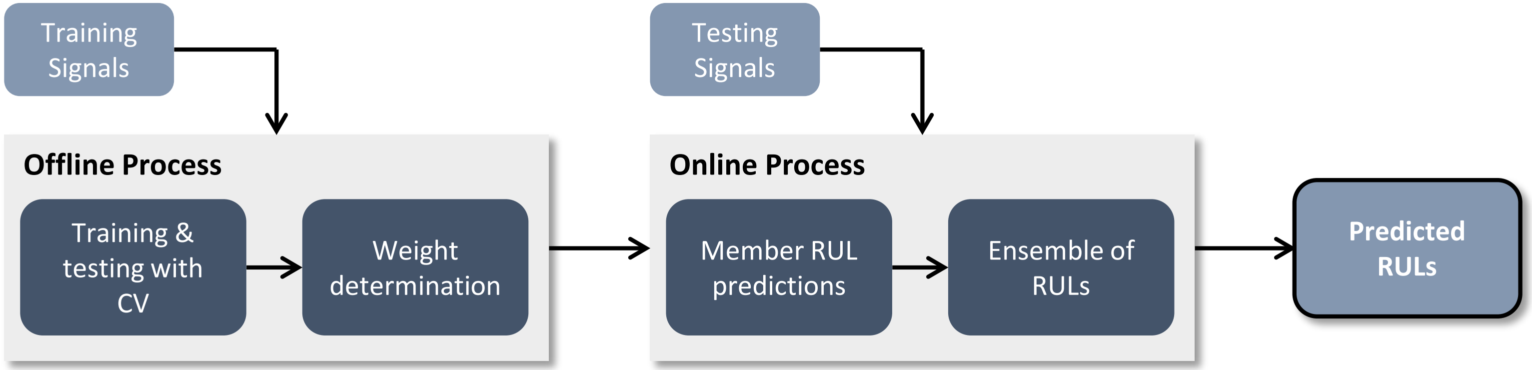


Example for Ensemble method

Case Study: Ensemble Method (Data-Driven)

Procedure of Ensemble approach

- **STEP 1 :** Acquire offline training sensory signals
- **STEP 2 :** Offline Process
  - **STEP 2a :** Perform the offline training and testing processes with k-fold cross validation(CV) with the training sensory signals to compute the CV error
  - **STEP 2b :** Determine the weights using 3 weighting schemes (accuracy-based, diversity-based, optimization-based)
- **STEP 3 :** Acquire online testing sensory signals
- **STEP 4 :** Online Process
  - **STEP 4a :** Predict RULs using the member algorithms through the online prediction process which employs health knowledge obtained from the offline training process
  - **STEP 4b :** Predict the ensemble RULs with the optimum weights obtained from STEP 2b



Flowchart of the ensemble method

SHRM System Health & Risk Management

Case Study: Ensemble Method (Data-Driven)

Details of Ensemble Method

- **Ensemble of 5 Prognostics Algorithm**
  - Similarity-based interpolation approach
    - RVM (Relevance Vector Machine), SVM (Support Vector Machine), Exponential fitting
  - Extrapolation-based approach
    - Bayesian linear regression
  - Recurrent neural network approach
- **Weighting schemes**
  - Accuracy-based weighting
    - To give larger weight to an algorithm with higher prediction accuracy
  - Diversity-based weighting
    - To give large weight to higher prediction diversity algorithm, contributing more to the ensemble RUL
  - Optimization-based weighting
    - To maximize the accuracy and robustness by synthesizing the accuracy & diversity
- **Weighted-sum of predicted RULs by 5 algorithms**

$$\hat{L} = \sum_{j=1}^{M} w_j \widehat{L_j}(\boldsymbol{y_t}, \boldsymbol{Y})$$

($\hat{L}$: ensemble predicted RUL, w_j : weight to the jth algorithm, $\widehat{L_j}$: RUL by jth algorithm)

Case Study: Ensemble Method (Data-Driven)

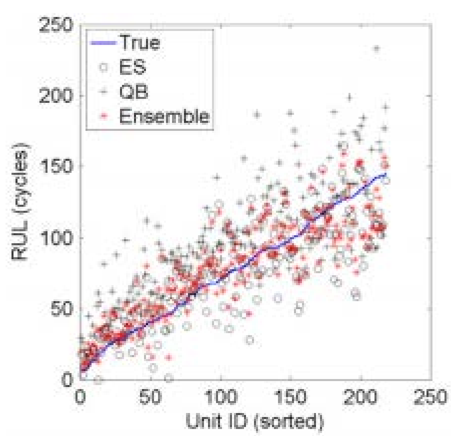
Case Studies for Validation of Proposed Model

- Apply Ensemble model to 3 sensory data
  - 2008 PHM challenge data, Power transformer, Electric cooling fan

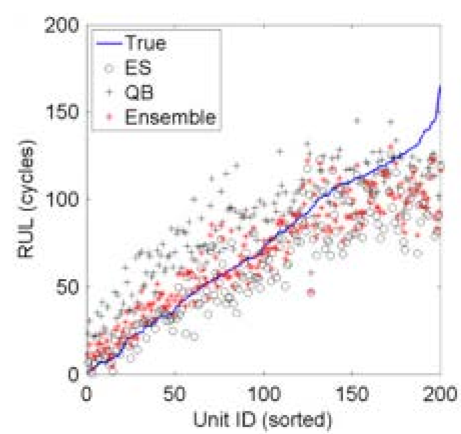
- Weighting results (Electric cooling fan)

	RS	SS	ES	QB	RN	RS–SS–ES–QB–RN		
						AW	DW	OW
Weight by AW	0.3646	0.3767	0.2552	0.0008	0.0027	–	–	–
Weight by DW	0.1423	0.1427	0.1496	0.3285	0.2369	–	–	–
Weight by OW	0.1155	0.8845	0.0000	0.0000	0.0000	–	–	–
CV error	1.4770	1.4298	2.1100	717.8430	199.0067	1.5188	11.8520	1.4292
Validation error	0.7027	0.9223	0.7037	461.5064	84.3975	0.7185	11.0177	0.6984

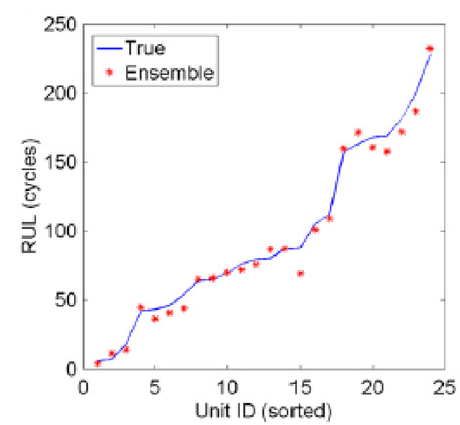
- RUL results plot



2008 PHM challenge data



Power transformer



Electric cooling fan

Case Study: Co-Training Method (Data-Driven)

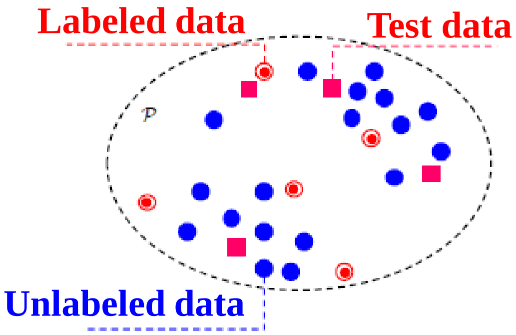
Co-training algorithm is a machine learning algorithm using small amounts of labeled data and large amounts of unlabeled data, it is often called as **Semi-Supervised Learning**

Overview of RUL prediction using Co-training method (Chao et. al., 2015)

- Issues of traditional data-driven prognostics
  - It requires some amount of failure data for achieving good prediction accuracy
  - **Failure data** are fairly expensive and time-consuming to obtain
  - **Suspension data*** are relatively easier to obtain than Failure data

- Objective
  - To improve the accuracy in RUL prediction using small amount of Failure data and large amount of suspension data

- Sample space
  - Failure data (Labeled, small amounts)
  - Suspension data (Unlabeled, large amounts)



Prognosis sample space

* Suspension data : condition monitoring data acquired from the very beginning of an engineered system's lifetime till planned inspection or maintenance when the system is taken out of service.

Case Study: Co-Training Method (Data-Driven)

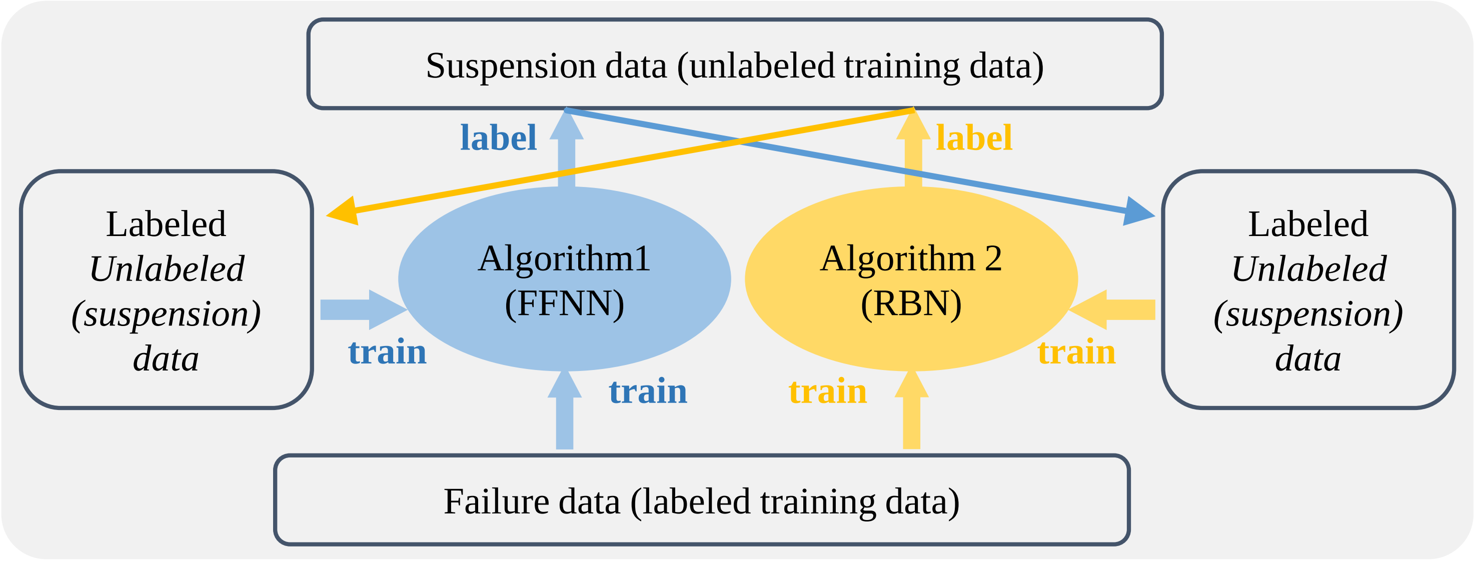
COPROG (CO-training PROGnostics) Process

- Algorithm1 : FFNN (Feed-forward neural network)
  - Prediction accuracy is quantified using the SSE performance function (or the validation error)

$$SSE = \sum e^2 = \sum (L^p - L^T)^2$$

L^p : Predicted Normalized RUL
L^T : True Normalized RUL

- Algorithm2 : RBN (Radial basis network)
  - It is ANN that uses radial basis functions as activation functions
  - The output layer weights is determined which the best approximate the training instances by a matrix pseudo-inverse technique



Flowchart of training process in COPROG

Case Study: Co-Training Method (Data-Driven)

COPROG (CO-training PROGnostics) Process

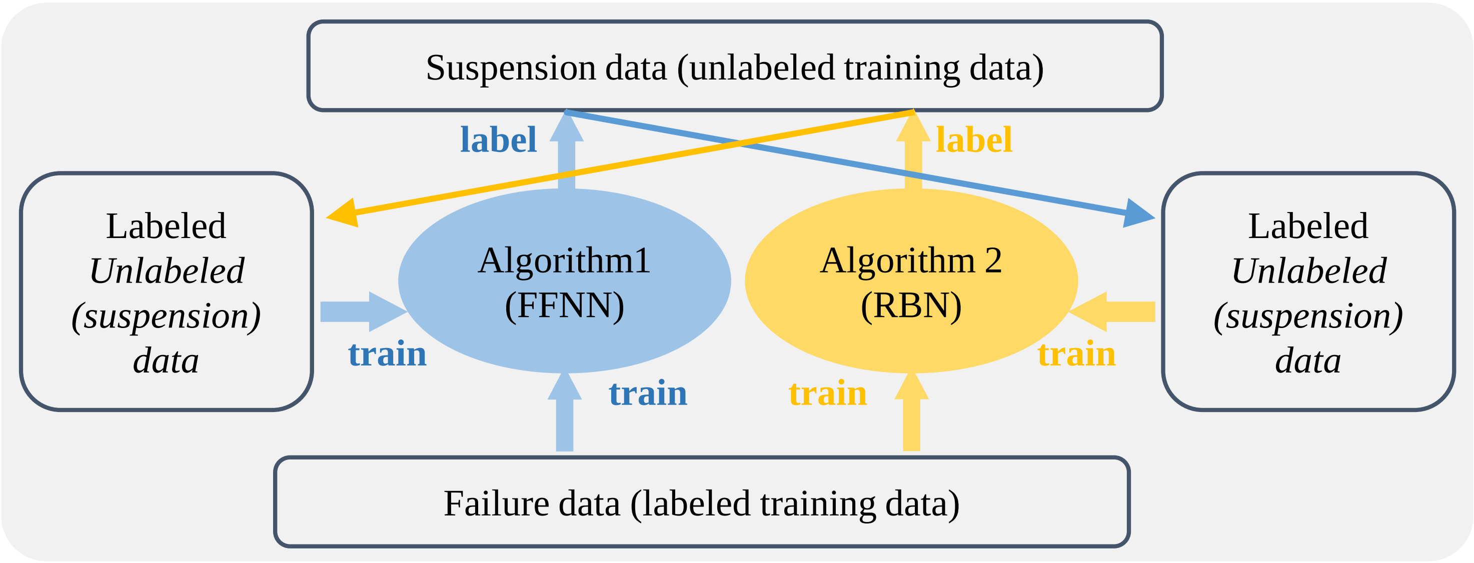
- Iterative Training process

  (1) training each algorithm using labeled data

  (2) predicting label of unlabeled data using trained algorithm

  (3) Training using labeled unlabeled data also

- Iterative training is repeated until no suspension unit can be found to be capable of reducing the prediction error of either algorithm on its training data set



Flowchart of training process in COPROG

Case Study: Co-Training Method (Data-Driven)

COPROG (CO-training PROGnostics) Process

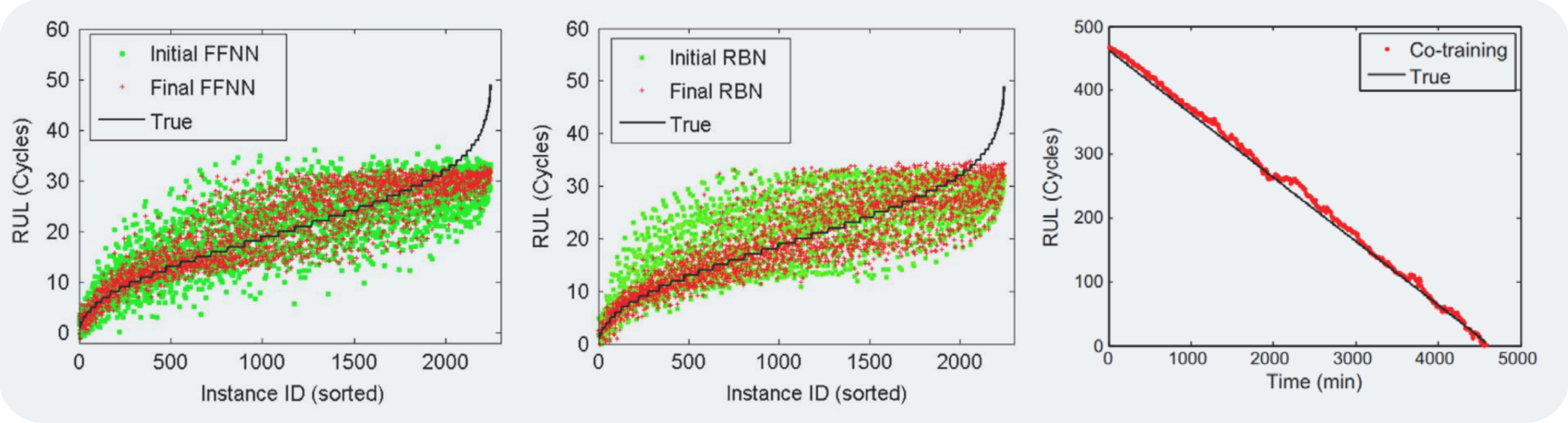
$$\mathit{Minimize} \quad SSE = \sum_{X_i \in \mathcal{L}} \left(L_i^T - \left(w_1 h_1(X_i) + w_2 h_2(X_i) \right) \right)^2$$

$$\mathit{Subject\ to} \quad w_1 + w_2 = 1, 0 \le w_1 \le 1, 0 \le w_2 \le 1$$

X_i : Training input instance
$\mathcal{L}$: labeled training data set
L_i^T : True Normalized RUL
h : Training Function (1=FFNN, 2=RBN)
w : weight

- Confidence Measure
  - It is need to identify the appropriate suspension unit and to minimize SSE in RUL prediction on the failure units

- Weight Optimization
  - the RUL predictions of these two algorithms are combined in a weighted-sum formulation as the final prediction



RUL predictions for a testing fan unit by co-training prognosis

THANK YOU
FOR LISTENING

Reference

[1] H. Chestnut, Systems Engineering Tools. New York: Wiley, 1965.

[2] R. Friedman and H. L. Eidenoff, ‘‘A Two-Stage Fatigue Life Prediction Method,’’ Grumman Advanced Development Report, 1990, Bethpage, NY.

[3] J. C. Newman, ‘‘FASTRAN-II: A Fatigue Crack Growth Structural Analysis Program,’’ NASA Technical Memorandum, 1992 Langley Research Center, Hampton, VA.

[4] L. Ljung, System Identification: Theory for the User, 2d ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[5] F. L. Lewis, Optimal Estimation: With an Introduction to Stochastic Control Theory. New York: Wiley, 1986.

[6] W. Weibull, ‘‘A Statistical Distribution Function of Wide Applicability,’’ Journal of Applied Mechanics 18:293, 1951.

[7] P. G. Groer, ‘‘Analysis of Time-to-Failure with a Weibull Model,’’ Proceedings of the Maintenance and Reliability Conference, MARCON 2000, Knoxville, TN.

[8] F. P. Lees, Loss Prevention in the Process Industries, 2d ed. Boston: Butterworth-Heinemann, 1996.

[9] G Vachtsevanos at al. Intelligent Fault Diagnosis and Prognosis for Engineering Systems, Wiley, 2006.

[10] Chao Hu, et al., “Ensemble of data-driven prognostics algorithms for robust prediction of remaining useful life,” Reliability Engineering and System Safety, 103, 2012, 120-135

[11] Vachtsevanos, G., et al., Intelligent Fault Diagnosis and Prognosis for Engineering Systems, First edition, Wiley, 2006, 314-333

[12] Nam-Ho, Kim, “Prognostics and Health Management of Engineering Systems”, First edition, Springer, 2016, 179-241

[13] Hu, Chao, et al. "A co-training-based approach for prediction of remaining useful life utilizing both failure and suspension data." Mechanical Systems and Signal Processing 62 (2015): 75-90.

[14] Ebden, Mark. "Gaussian processes for regression: A quick introduction." The Website of Robotics Research Group in Department on Engineering Science, University of Oxford, 2008

[15] Mahamad, et al, "Predicting remaining useful life of rotating machinery based artificial neural network." Computers & Mathematics with Applications 60.4, 2010, 1078-1087.

[16] Hyunseok Oh, et al., “An empirical model to describe performance degradation for warranty abuse detection in portable electronics,” *Reliability Engineering and System Safety*, 142, (2015), 92-99

[17] Kim, Dae Hwan, et al. "Bivariate Lifetime Model for Organic Light-Emitting Diodes." *IEEE Transactions on Industrial Electronics*, 64.3, (2017): 2325-2334.

[18] Choi, Woosung, et al. "A Bayesian approach for a damage growth model using sporadically measured and heterogeneous on-site data from a steam turbine." Reliability Engineering & System Safety (2018).e