A Wiener Process Model with Dynamic Covariate for Degradation Modeling and Remaining Useful Life Prediction

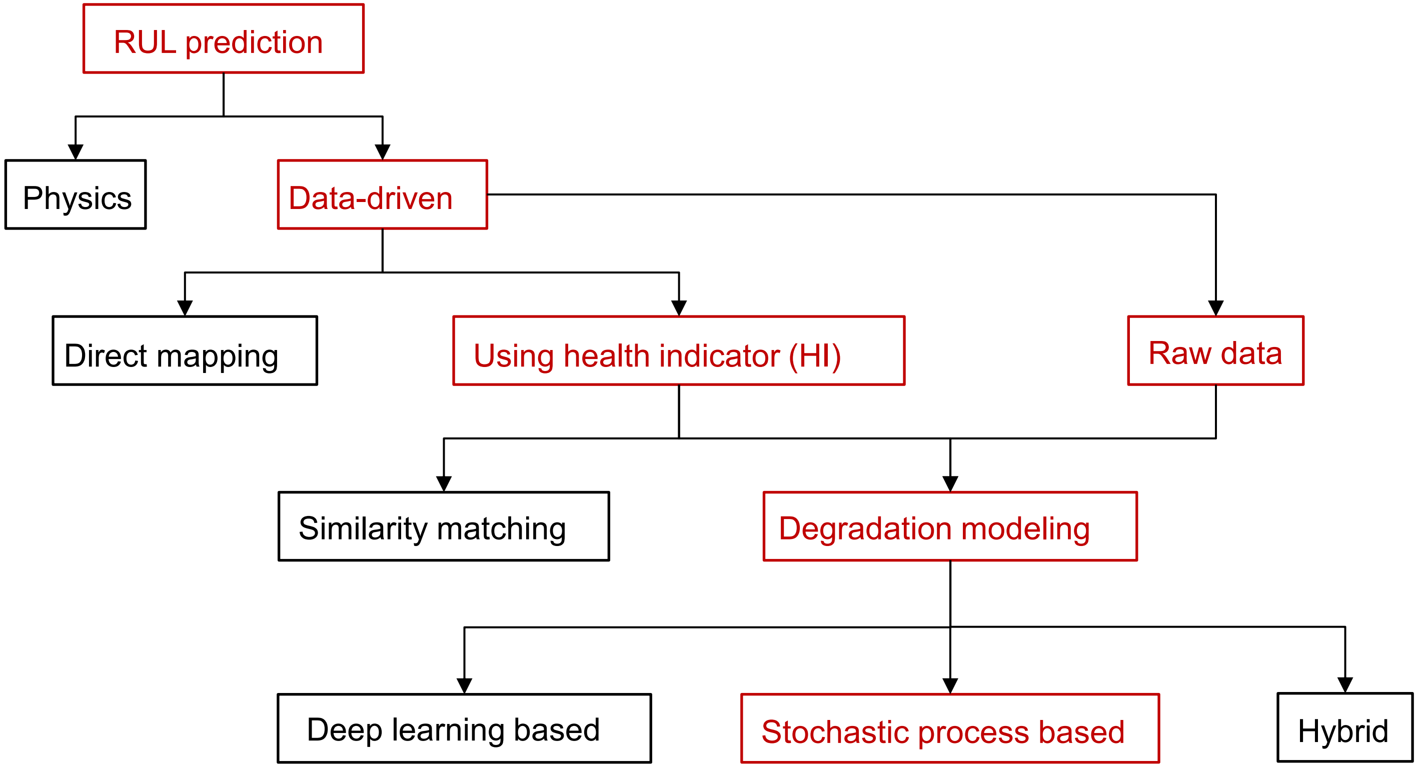
IEEE Transactions on Reliability, 2023
Shuyi Zhang, Qingqing Zhai, Xin Shi, and Xuejuan Liu (citation: 117)

Park Han Byeol
Department of Industrial Engineering, Pusan National University
parkhanbyeol104@gmail.com

Contents

- Mathematical background: wiener process (stochastic) for degradation modeling
  - Introduction to a general framework for probabilistic remaining useful life (RUL) prediction using stochastic process-based degradation models (stochastic process: wiener process)
- Paper review
  - A Wiener Process Model With Dynamic Covariate for Degradation Modeling and Remaining Useful Life Prediction
- Discussion
  - Strength
  - Weakness
  - Insight

Preliminary



Preliminary

- What is the first passing time prediction problem?
  - Problem of predicting the time at which the value of a function first reaches a specific threshold

- Application domains of first passing time prediction
  - Finantial engineering
    - Stock prices reach a critical level
  - System engineering
    - Network load exceeds a threshold
  - Mechanical and manufacturing engineering
    - RUL
  - Biotechnology
    - Particle reach to the specific target
      - Target is reached, it may trigger either a beneficial or a harmful process.

Preliminary

- Typical procedure for solving a first hitting time problem
  1. Establish a degradation model

  2. Transform the stochastic process for the degradation model X_t into the probability density function (pdf) of the htting time τ_h

  3. Derive the likelihood function
     - Under the assumption of independence between increments, the likelihood function of the degradation model (i.e., the stochastic process) can be expressed as the product of the likelihood at individual time points

  4. Perform statistical inference
     - If possible, use maximu likelihood estimation (MLE)
     - If not feasible, alternative include the following (not covered in this presentation)
       - Monte carlo simulation
       - Expectation maximization (EM) algorithm
       - Numerical analysis
       - Optimization methods

Preliminary

- Explain the full procedure for degradation modeling and RUL distribution estimation basedon a Wiener process with a linear drift term, in the following order:
    1. Brownian motion

    2. Wiener process (degradatiom model X_t)

    3. Reformulate the degradation model pdf X_t as a first hitting time problem for reaching the threshold h, that is, finding c_h
        - Here, the parameters of τ_h are same as those of X_t

    4. Statistical inference
        - Derive the likelihood function of X_t, perform statistical inference using MLE, and then use the resulting parameter estimates in the pdf of τ_h to estimate the RUL distribution

Preliminary-definition

DEFINITION 1.6.– $B = (B_t)_{t \geq 0}$ *is a Brownian motion if, and only if,*

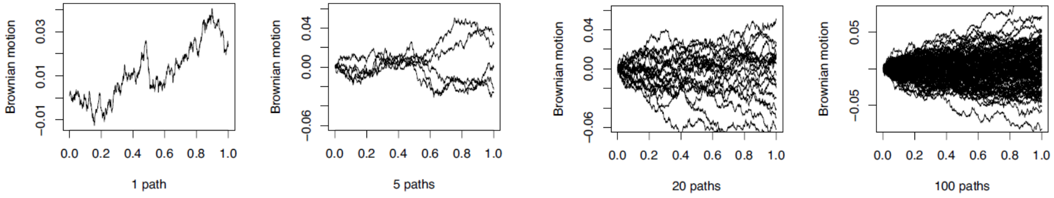
1) $B_0 = 0$ *a.s.;*

2) B *has stationary increments: for any* $t, s, h > 0$, $B_t - B_s \stackrel{d}{=} B_{t+h} - B_{s+h}$;

3) B *has independent increments: for any* $n \geq 1$ *and for any* $t_1 < \cdots < t_n$, *the random variables* $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ *are independent;*

4) *for any* $t > 0$, $B_t \sim \mathcal{N}(0, t)$.

For an extensive description of Brownian motion and its properties, the reader could refer to [MÖR 10], for instance.



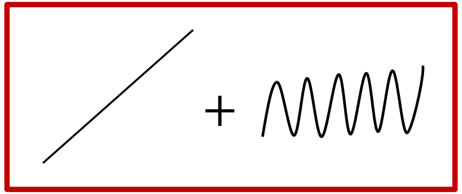
부산대학교 PUSAN NATIONAL UNIVERSITY BAE La Bigdata Analytics & Enginee...

Preliminary-definition

DEFINITION 1.11.– *The stochastic process* $W = (W_t)_{t \geq 0}$ *is called a* Wiener process *if*

$$\forall t \geq 0, \quad W_t = \mu t + \sigma B_t, \qquad [1.2]$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ *and* $(B_t)_{t \geq 0}$ *is a Brownian motion:* μ *is a drift parameter (that we expect to be positive in our case) and* σ^2 *is a variance parameter (or volatility parameter).*



$$W_t \sim N(\mu t, \sigma^2 t)$$

Preliminary-definition

Increment

$$W_t = \mu t + \sigma B_t$$

$$W_t - W_s \sim N\big(\mu(t-s), \sigma^2(t-s)\big)$$

$$B_{t-t_0} \sim N\big(0, \sigma^2(t-t_0)\big)$$

$$\mathbb{E}(W_t - W_s) = \mu(t-s) \quad \text{and} \quad \mathrm{Var}(W_t - W_s) = \sigma^2(t-s).$$

$$\forall t \geq t_0, \quad X_t = x_0 + W_{t-t_0} = x_0 + \mu(t-t_0) + \sigma B_{t-t_0}$$

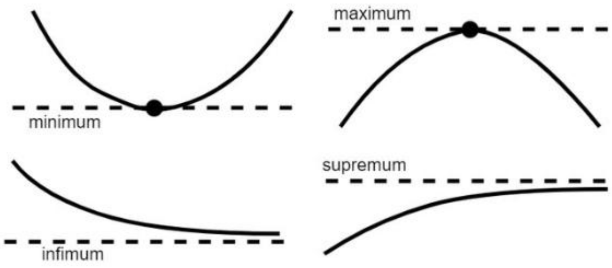
$$X_t \sim N\big(\mu(t-t_0), \sigma^2(t-t_0)\big)$$

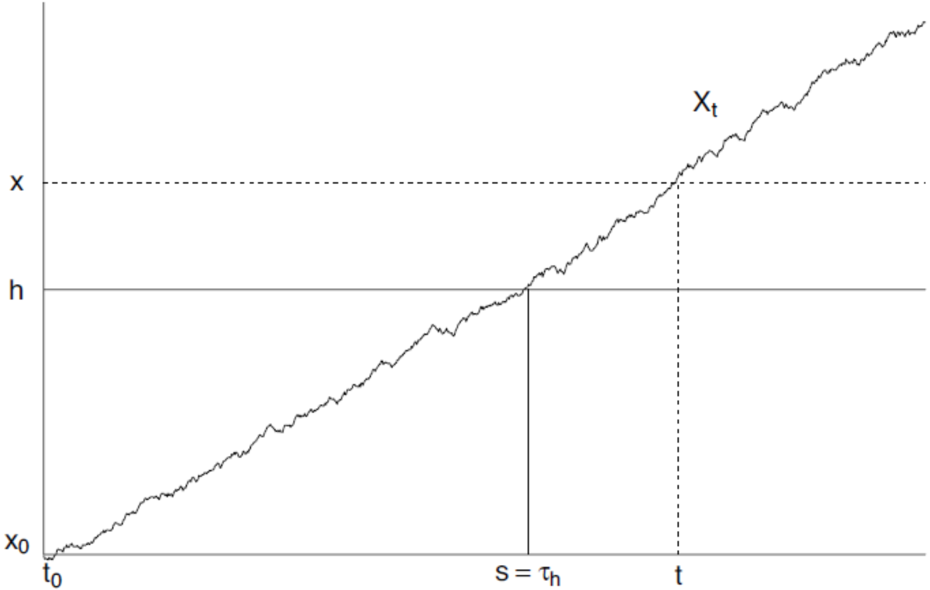
Degradation model

Preliminary-definition of first hitting time τ_h

First hitting time $\tau_h = \inf\{t \geq t_0; X_t \geq h\}$.

$$\forall t \geq t_0, \quad X_t = x_0 + W_{t-t_0} = x_0 + \mu(t-t_0) + \sigma B_{t-t_0},$$





Preliminary-derivation of the basic ingredients required for the proof

Transformation of a drift-free stochastic-process-based degradation model into a first hitting time problem

Let $x > h$. *The event* $X_t \geq x$ *implies that* $\tau_h < t$ (see figure).

$$P(X_t > x, \tau_h < t) = P(X_t > x|\tau_h < t)P(\tau_h < t) = \int_{t_0}^{t} P(X_t > x|\tau_h = s)P(\tau_h = s)ds$$

$$\mathbb{P}(X_t > x) = \mathbb{P}(X_t > x; \tau_h < t) = \int_{t_0}^{t} \mathbb{P}(X_t > x|\tau_h = s)\,\mathrm{d}F_{\tau_h}(s).$$

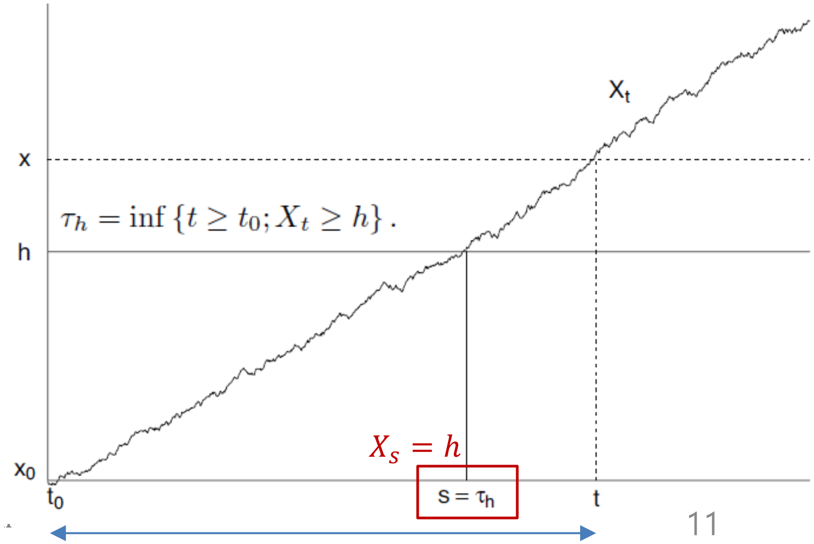
$$\mathbb{P}(X_t > x) = \int_{t_0}^{t} \mathbb{P}(X_t > x|X_s = h)\,\mathrm{d}F_{\tau_h}(s)$$

$$= \int_{t_0}^{t} \mathbb{P}(X_t - X_s > x - h)\,\mathrm{d}F_{\tau_h}(s)$$

$$= \int_{t_0}^{t} \mathbb{P}(X_{t-s} > x - h)\,\mathrm{d}F_{\tau_h}(s).$$

Partial derivative of x

$$f_{X_t}(x) = \int_{t_0}^{t} f_{\tau_h}(s) f_{X_{t-s}}(x-h)\,\mathrm{d}s,$$



Preliminary-linking the distribution of the degradation model to the first hitting time distribution

Easy example (wiener process with no drift)

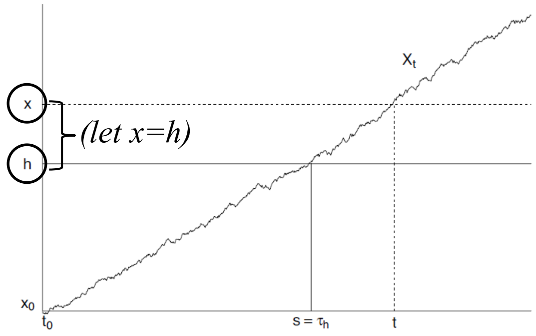
PROOF.– For these parameters, we have $X_t = \sigma B_t$ and, since $\frac{B_t}{\sqrt{t}}$ is standard Gaussian distributed,

(Standardization)

$$\mathbb{P}(X_t > x) = \mathbb{P}(\sigma B_t > x) = \mathbb{P}\left(\frac{B_t}{\sqrt{t}} > \frac{x}{\sigma\sqrt{t}}\right) = 1 - \Phi\left(\frac{x}{\sigma\sqrt{t}}\right),$$

$B_t \sim N(0,t)$
$\rightarrow E[B_t] = 0$
$\rightarrow Var[B_t] = t$

$\rightarrow \frac{B_t}{\sqrt{t}} \sim N(0,1)$



where Φ is the cumulative distribution function of the standard Gaussian distribution. Furthermore, for $x = h$, we have

$$\mathbb{P}(X_{t-s} > x - h) = F_{X_{t-s}}(0) = \frac{1}{2},$$

because X_{t-s} is Gaussian distributed with zero expectation. Now, we get, from equation [1.7],

$$\mathbb{P}(X_t > h) = \int_0^t \frac{1}{2} \mathrm{d}F_{\tau_h}(s) = \frac{1}{2}\int_0^t P(\tau_h = s)ds = P(\tau_t < t) = \frac{1}{2}F_{\tau_h}(t)$$

PPT Page 11 proof

$$\begin{aligned}\mathbb{P}(X_t > x) &= \int_{t_0}^{t} \mathbb{P}(X_t > x | X_s = h)\, \mathrm{d}F_{\tau_h}(s) \\ &= \int_{t_0}^{t} \mathbb{P}(X_t - X_s > x - h)\, \mathrm{d}F_{\tau_h}(s) \\ &= \int_{t_0}^{t} \mathbb{P}(X_{t-s} > x - h)\, \mathrm{d}F_{\tau_h}(s).\end{aligned}$$

and the c.d.f. of the first passage time is

First hitting time CDF

$$F_{\tau_h}(t) = 2\,\mathbb{P}(X_t > h) = 2\left(1 - \Phi\left(\frac{h}{\sigma\sqrt{t}}\right)\right).$$

Degradation CDF

Preliminary

$X_t = \sigma B_t$

$$F_{\tau_h}(t) = 2\,\mathbb{P}(X_t > h) = 2\left(1 - \Phi\left(\frac{h}{\sigma\sqrt{t}}\right)\right).$$

Partial derivative of t

$$\forall t \geq 0,\ f_{\tau_h}(t) = \frac{\mathrm{d}F_{\tau_h}(t)}{\mathrm{d}t} = \frac{h}{\sigma\sqrt{t^3}}\Phi'\left(\frac{h}{\sigma\sqrt{t}}\right) = \frac{h}{\sqrt{2\pi\sigma^2 t^3}}\exp\left(-\frac{h^2}{2\sigma^2 t}\right).$$

This proof is restricted to the case $X_t = \sigma B_t \sim N(0, \sigma^2 t)$.

PROPOSITION 1.12.– For $\mu = 0$, $x_0 = 0$ and $t_0 = 0$,

$$\forall t \geq 0, \quad F_{\tau_h}(t) = 2\left(1 - \Phi\left(\frac{h}{\sigma\sqrt{t}}\right)\right)$$

is the unique solution of equation [1.7]. The p.d.f. is then given by:

$$\forall t \geq 0, \quad f_{\tau_h}(t) = \frac{h}{\sqrt{2\pi\sigma^2 t^3}}\exp\left(-\frac{h^2}{2\sigma^2 t}\right).$$

Parameter of Degradation model X_t

Preliminary

What about a stochastic process with drift? (proof omitted)

$Let\ X_t = x_0 + \mu(t - t_0) + \sigma B_{t-t_0} \sim N\left(x_0 + \mu(t - t_0), \sigma^2(t - t_0)\right)$

With linear drift μ
With start point x_0
With start time t_0

PROPOSITION 1.12.– For $\mu = 0$, $x_0 = 0$ and $t_0 = 0$,

$$\forall t \geq 0, \quad F_{\tau_h}(t) = 2\left(1 - \Phi\left(\frac{h}{\sigma\sqrt{t}}\right)\right)$$

is the unique solution of equation [1.7]. The p.d.f. is then given by:

$$\forall t \geq 0, \quad f_{\tau_h}(t) = \frac{h}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{h^2}{2\sigma^2 t}\right).$$

THEOREM 1.14.– Remembering that the distribution of X_t is given by equation [1.5], then

$$\forall t \geq t_0, \quad f_{\tau_h}(t) = \frac{h - x_0}{\sqrt{2\pi\sigma^2 (t - t_0)^3}} \exp\left(-\frac{(h - x_0 - \mu(t - t_0))^2}{2\sigma^2 (t - t_0)}\right). \quad [1.11]$$

linking the distribution of the degradation model to the first hitting time distribution

Next: definition of likelihood function (degradation model)
Satistical inference
-> estimators of x_0, t_0, μ, σ^2 ??

Preliminary-establishing the likelihood function

Suppose that there are i=1,…,n individual run-to-failure degradation trajectories.

Assume that degradation model is

$X_t = x_0 + \mu(t - t_0) + \sigma B_{t-t_0} \sim N\left(x_0 + \mu(t - t_0), \sigma^2(t - t_0)\right)$

t_0: *initial time point (discrete time index=0)*

$t_{i,1}$: *time point (discrete time index=1)*

$s_{i,j} = t_{i,j} - t_{i,j-1}$

$X_{i,1}$: degradation model for the **increment from the** $t_{i,1}$

$Y_{i,j}$: **increment over** $s_{i,j}$

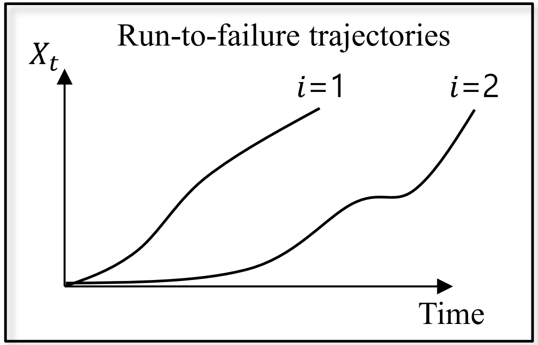
For any $t \geq t_0$, the p.d.f. of X_t is then

$$\forall x \in \mathbb{R}, \quad f_{X_t}(x) = \varphi\left(\frac{x - x_0 - \mu(t - t_0)}{\sigma(t - t_0)}\right),$$

where φ is the p.d.f. of the standard Gaussian distribution.

$$X_{i,1} \sim \mathcal{N}(x_0 + \mu(t_{i,1} - t_0), \sigma^2(t_{i,1} - t_0)), \quad \forall i \in \{1, \ldots, n\}$$
$$Y_{i,j} \sim \mathcal{N}(\mu s_{i,j}, \sigma^2 s_{i,j}), \quad \forall j \in \{1, \ldots, m_i\}, \forall i \in \{1, \ldots, n\}$$

$$L(x_0, t_0, \mu, \sigma^2 | \mathbf{D}_{obs}) = \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2(t_{i,1} - t_0)}} \varphi\left(\frac{x_{i,1} - x_0 - \mu(t_{i,1} - t_0)}{\sqrt{\sigma^2(t_{i,1} - t_0)}}\right) \times \prod_{i=1}^{n} \prod_{j=2}^{m_i} \frac{1}{\sqrt{\sigma^2 s_{i,j}}} \varphi\left(\frac{y_{i,j} - \mu s_{i,j}}{\sqrt{\sigma^2 s_{i,j}}}\right), \quad [1.14]$$



THEOREM 1.14.– Remembering that the distribution of X_t is given by equation [1.5], then

$$\forall t \geq t_0, \quad f_{\tau_h}(t) = \frac{h - x_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - x_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right). \quad [1.11]$$

Preliminary

- Statistical inference

Markovian property

$$\{X_t\}_{t=1}^{T} = X_1, \dots, X_T = \prod_{t=1}^{T} P(X_t|X_1, \dots, X_{t-1}) = \prod_{t=1}^{T} P(X_t|X_{t-1})$$

=Increments

$$X_{i,1} \sim \mathcal{N}(x_0 + \mu(t_{i,1} - t_0), \sigma^2(t_{i,1} - t_0)), \quad \forall i \in \{1, \dots, n\}$$

$$Y_{i,j} \sim \mathcal{N}(\mu s_{i,j}, \sigma^2 s_{i,j}), \quad \forall j \in \{1, \dots, m_i\}, \forall i \in \{1, \dots, n\}$$

$$L(x_0, t_0, \mu, \sigma^2 | \mathrm{D}_{obs}) = \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2(t_{i,1} - t_0)}} \varphi\left(\frac{x_{i,1} - x_0 - \mu(t_{i,1} - t_0)}{\sqrt{\sigma^2(t_{i,1} - t_0)}}\right)$$
$$\times \prod_{i=1}^{n} \prod_{j=2}^{m_i} \frac{1}{\sqrt{\sigma^2 s_{i,j}}} \varphi\left(\frac{y_{i,j} - \mu s_{i,j}}{\sqrt{\sigma^2 s_{i,j}}}\right), \qquad [1.14]$$

MLE

THEOREM 1.14.– Remembering that the distribution of X_t is given by equation [1.5], then

$$\forall t \geq t_0, \quad f_{\tau_h}(t) = \frac{h - x_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - x_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right) \qquad [1.11]$$

Preliminary-other stochastic process-based degradation modeling papers

- Efforts to nonlinear drift
  - Time scaling

    t: Linear time $\qquad\qquad$ $\Lambda(t;\alpha) = t^{\alpha}$: Non linear time

    $$X_t = \mu t + \sigma B_t \Longrightarrow X_t = \mu\Lambda(t;\alpha) + \sigma B_{\Lambda(t;\alpha)}$$

  - Nonlinear modeling of the drift function itself
    - Limitation: obtaining a closed-form solution is difficult or, in some cases, impossible
      - » This problem is same as solving the Fokker-Planck-Kolmogorov (FPK) equation with boundary conditions...??

    μ: linear drift $\qquad\qquad$ $\int \mu(t)dt$: non linear drift

    $$X_t = \mu t + \sigma B_t \Longrightarrow X_t = \int \mu(t)dt + \sigma B_t$$

- Other research
  - Two-stage drift function (drift)
    - Paper: Remaining useful life prediction for two-phase degradation model based on reparameterized inverse gaussian process
  - Fractional Brownian motion based diffusion (diffusion)
    - Paper: Predicting remaining useful life based on a generalized degradation with fractional Brownian motion

Preliminary-Summary

- Framework of Stochastic process-based RUL prediction
  - Define degradation model
    - Using stochastic differential equation
  - Construction of the likelihood function
  - Statistical inference
  - Wald (inverse gaussian) distribution

$$\left(or\ start\ dX_t = \mu dt + \sigma B_{dt} \sim N(\mu dt, \sigma^2 dt)\right)$$

$$Let\ X_t = x_0 + \mu(t - t_0) + \sigma B_{t-t_0} \sim N\left(x_0 + \mu(t - t_0), \sigma^2(t - t_0)\right)$$

$$L(x_0, t_0, \mu, \sigma^2 | \mathbf{D}_{obs}) = \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2(t_{i,1} - t_0)}} \varphi\left(\frac{x_{i,1} - x_0 - \mu(t_{i,1} - t_0)}{\sqrt{\sigma^2(t_{i,1} - t_0)}}\right) \times \prod_{i=1}^{n} \prod_{j=2}^{m_i} \frac{1}{\sqrt{\sigma^2 s_{i,j}}} \varphi\left(\frac{y_{i,j} - \mu s_{i,j}}{\sqrt{\sigma^2 s_{i,j}}}\right), \quad [1.14]$$

Statistical inference

$\hat{x}_0, \hat{t}_0, \mu, \sigma$

$$f_{\tau_h}(t) = \frac{h - x_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - x_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right).$$

Predict RUL

Paper review

- Paper review (recent stochastic process-based degradation modeling paper)
  - A Wiener Process Model With Dynamic Covariate for Degradation Modeling and Remaining Useful Life Prediction
  - Research question
    - What if the drift of degradation is dependent on other factors?

Paper

- Framework of Stochastic process-based RUL prediction
  - Define degradation model
    - Using stochastic differential equation
  - Construction of the likelihood function
  - Statistical inference
  - Wald (inverse gaussian) distribution

Paper

- Define degradation model ⟹ $X(t) = X(0) + v\Lambda(t;\alpha) + \sigma B(\Lambda(t;\alpha))$
  - Using stochastic differential equation

Ornstein-uhlenbeck (OU) process (stochastic differential equation)

$$dH(t) = \eta(\mu_H - H(t))\,dt + \sigma_H dW(t),$$

$$H(t) = (1 - e^{-\eta t})\mu_H + H(0)e^{-\eta t} + \frac{\sigma_H}{\sqrt{2\eta}} e^{-\eta t} W(e^{2\eta t} - 1)$$

$$E(H(t)) = H(0)e^{-\eta t} + (1 - e^{-\eta t})\mu_H$$

$$Var(H(t)) = \frac{\sigma_H^2}{2\eta}(1 - e^{-2\eta t}).$$

A key characteristic of the Ornstein-uhlenbeck (OU) process lies in its drift function
μ_H: *long-term average*

If $\mu_H > H(t) \rightarrow (\mu_H - H(t)) > 0$
This means that the long-term average is greater than the current observed value.
As a result, $dH(t)$ *tends to be positive.*

In other words, when the current value is below the long-term average,
a positive increment derives the process back toward that long-term average.

The parameter η *determines how strongly the process reverts toward the long-term average.*

For this reason, the OU process is also referred to as a mean-reverting process.

Paper

- Define degradation model $X(t) = X(0) + v\Lambda(t;\alpha) + \sigma B(\Lambda(t;\alpha))$

$$v(t) = \mu e^{\beta H(t)}$$

⇩

$$v(t) = \mu e^{\beta H(t)}$$

$$X(t) = \int_0^t v(\tau)\, d\Lambda(\tau;\alpha) + \sigma B(\Lambda(t;\alpha)).$$

Ornstein-uhlenbeck (OU) process

$$dH(t) = \eta(\mu_H - H(t))\,dt + \sigma_H dW(t),$$

$$H(t) = (1 - e^{-\eta t})\mu_H + H(0)e^{-\eta t} + \frac{\sigma_H}{\sqrt{2\eta}} e^{-\eta t} W(e^{2\eta t} - 1)$$

$$E(H(t)) = H(0)e^{-\eta t} + (1 - e^{-\eta t})\mu_H$$

$$Var(H(t)) = \frac{\sigma_H^2}{2\eta}(1 - e^{-2\eta t}).$$

Wiener process parameter

OU process parameter

Time scaling parameter

The model parameters are $\Theta = \{\mu_H, \eta, \sigma_H^2, \mu, \beta, \sigma^2, \alpha\}$.

Summary

- Statistical inference

A. Parameter Estimation in the Covariate Process

Considering the Markov property of the OU process, the joint distribution of the covariate $\mathbf{H}$ can be factorized as

$$f(\mathbf{H}) = \prod_{i=1}^{m} f(H_i|H_{i-1}) \tag{11}$$

where H_i follows a normal distribution conditional on H_{i-1}

$$(H_i|H_{i-1}) \sim N\left(\left(1-e^{-\eta\Delta t_i}\right)\mu_H + H_{i-1}e^{-\eta\Delta t_i}, \frac{\sigma_H^2}{2\eta}\left(1-e^{-2\eta\Delta t_i}\right)\right). \tag{12}$$

Therefore, the log-likelihood function can be obtained as

$$\begin{aligned}\ell\left(\eta,\mu_H,\sigma_H^2|\mathbf{H}\right) &= \sum_{i=1}^{m} \ln f(H_i|H_{i-1}) \\ &= -\frac{m}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{m} \ln\left(\frac{\sigma_H^2}{2\eta}\left(1-e^{-2\eta\Delta t_i}\right)\right) - \frac{\eta}{\sigma_H^2}\sum_{i=1}^{m} \frac{\left(H_i-\left(\mu_H+\left(H_{i-1}-\mu_H\right)e^{-\eta\Delta t_i}\right)\right)^2}{1-e^{-2\eta\Delta t_i}}.\end{aligned} \tag{13}$$

Taking partial derivative of the log-likelihood function with respect to μ_H, the MLE of $\widehat{\mu_H}$ is obtained as

$$\widehat{\mu_H} = \sum_{i=1}^{m} \frac{H_i - H_{i-1}e^{-\eta\Delta t_i}}{1+e^{-\eta\Delta t_i}}\left(\sum_{i=1}^{m}\frac{1-e^{-\eta\Delta t_i}}{1+e^{-\eta\Delta t_i}}\right)^{-1}. \tag{14}$$

Similarly, the MLE of $\widehat{\sigma_H^2}$ can be obtained as

$$\widehat{\sigma_H^2} = \frac{2\eta}{m}\sum_{i=1}^{m}\frac{\left(H_i-\left(\widehat{\mu_H}+\left(H_{i-1}-\widehat{\mu_H}\right)e^{-\eta\Delta t_i}\right)\right)^2}{1-e^{-2\eta\Delta t_i}}. \tag{15}$$

Substitute $\widehat{\mu_H} = \widehat{\mu_H}(\eta)$ and $\widehat{\sigma_H^2} = \widehat{\sigma_H^2}(\eta)$ into the log-likelihood function $\ell(\eta,\mu_H,\sigma_H^2|\mathbf{H})$, we can obtain the profiled likelihood as

$$\ell_P(\eta|\mathbf{H}) = \ell(\eta,\widehat{\mu_H}(\eta),\widehat{\sigma_H^2}(\eta)|\mathbf{H}). \tag{16}$$

Then, the estimate $\hat{\eta}$ can be obtained by maximizing the profiled likelihood with respect to η, and the estimates $\widehat{\mu_H}$ and $\widehat{\sigma_H^2}$ can be obtained by substituting $\hat{\eta}$ back to (14) and (15), respectively.

The model parameters are $\Theta = \{\mu_H, \eta, \sigma_H^2, \mu, \beta, \sigma^2, \alpha\}$.

Summary

- Construction of the likelihood function

B. Parameter Estimation in the Degradation Process

Let $\Delta X_1 = X_1$ and $\Delta X_i = X_i - X_{i-1}$ for $i = 2, 3, \ldots, m$. From the state equations in (10), we can readily obtain the following log-likelihood function with respect to $\{\mu, \sigma^2, \beta\}$

$$\ell\left(\mu, \sigma^2, \beta | \mathbf{X}, \mathbf{H}\right) = -\frac{1}{2}\sum_{i=1}^{m} \ln\left(2\pi\sigma^2\lambda_i\right) - \sum_{i=1}^{m} \frac{\left(\Delta X_i - \mu e^{\beta H_{i-1}}\lambda_i\right)^2}{2\sigma^2\lambda_i}. \quad (17)$$

Taking partial derivatives with respect to $\{\mu, \beta, \sigma^2\}$, the estimators for the model parameters are obtained as follows:

$$\hat{\mu} = \frac{\sum_{i=1}^{m} \Delta X_i e^{\hat{\beta} H_{i-1}}}{\sum_{i=1}^{m} e^{2\hat{\beta} H_{i-1}}\lambda_i} \quad (18)$$

$$\sum_{i=1}^{m} e^{\hat{\beta} H_{i-1}} = \frac{1}{\hat{\mu}} \sum_{i=1}^{m} \frac{\Delta X_i}{\lambda_i} \quad (19)$$

$$\widehat{\sigma^2} = \frac{1}{m}\sum_{i=1}^{m} \frac{\left(\Delta X_i - \hat{\mu} e^{\hat{\beta} H_{i-1}}\lambda_i\right)^2}{\lambda_i}. \quad (20)$$

In the case that α is unknown, the estimators $\{\hat{\mu}, \widehat{\sigma^2}, \hat{\beta}\}$ are functions of α. Therefore, the estimate for α can be obtained by substituting $\{\hat{\mu}, \widehat{\sigma^2}, \hat{\beta}\}$ into the log-likelihood function $\ell(\mu, \sigma^2, \beta | \mathbf{X}, \mathbf{H})$ and maximizing the resulting profiled log-likelihood.

The model parameters are $\Theta = \{\mu_H, \eta, \sigma_H^2, \mu, \beta, \sigma^2, \alpha\}$.

Use wald (inverse gaussin) distribution -> predict RUL

Summary

- Summary
  - Advantages
    - Closed-form expressions
    - Interpretability
  - Disadvantages
    - Assumptions
      - Independent increments
      - Stationary increments
        - » Assumption of linear drift or brownian diffusion term
    - Can only be modeled using known functional forms

Assumtion: gaussian diffusion

$B_{t-t_0} \sim N(0, \sigma^2(t - t_0))$

$$X_t = x_0 + \mu(t - t_0) + \sigma B_{t-t_0},$$

Assumtion: known drift family

1. *How can it be identified?*
2. *Is it truly the optimal drift function?*

- Real industry
  - The increments may be correlated
    - Difficulty: constructing the likelihood function may become difficult
  - The increments may be non-stationary

Related works (hybrid degradation modeling)

- A long short-term memory neural network based Wiener process model for remaining useful life prediction (2022)
  - Reliability engineering and system safety

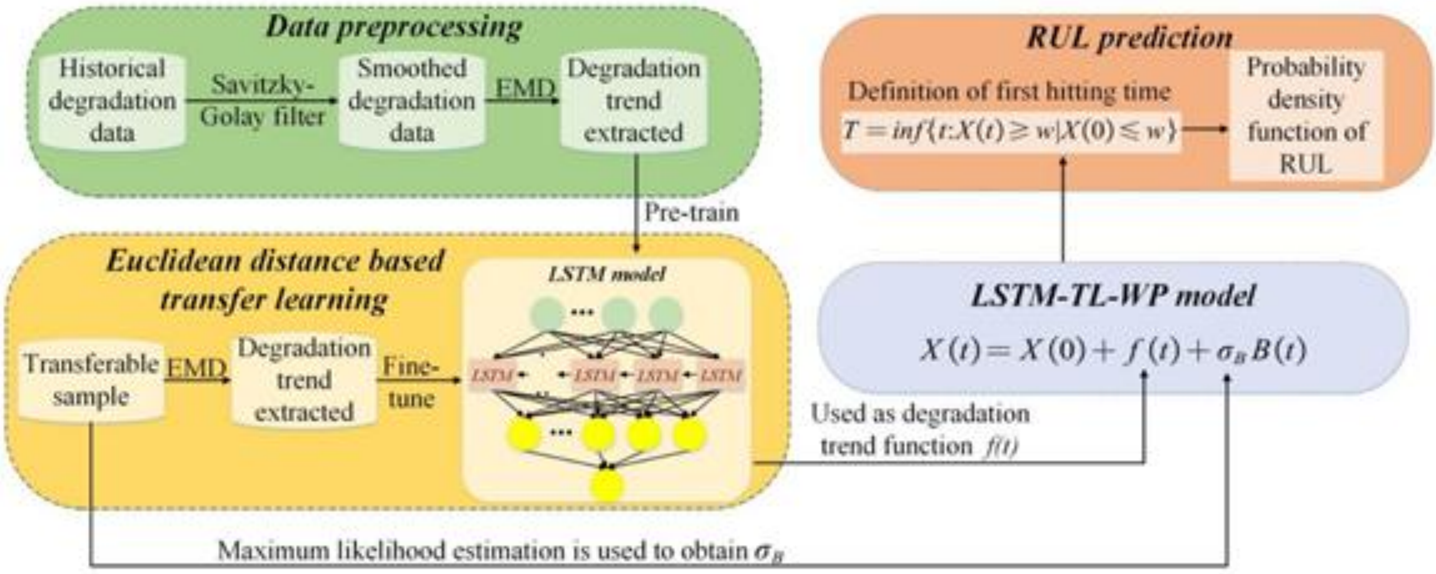
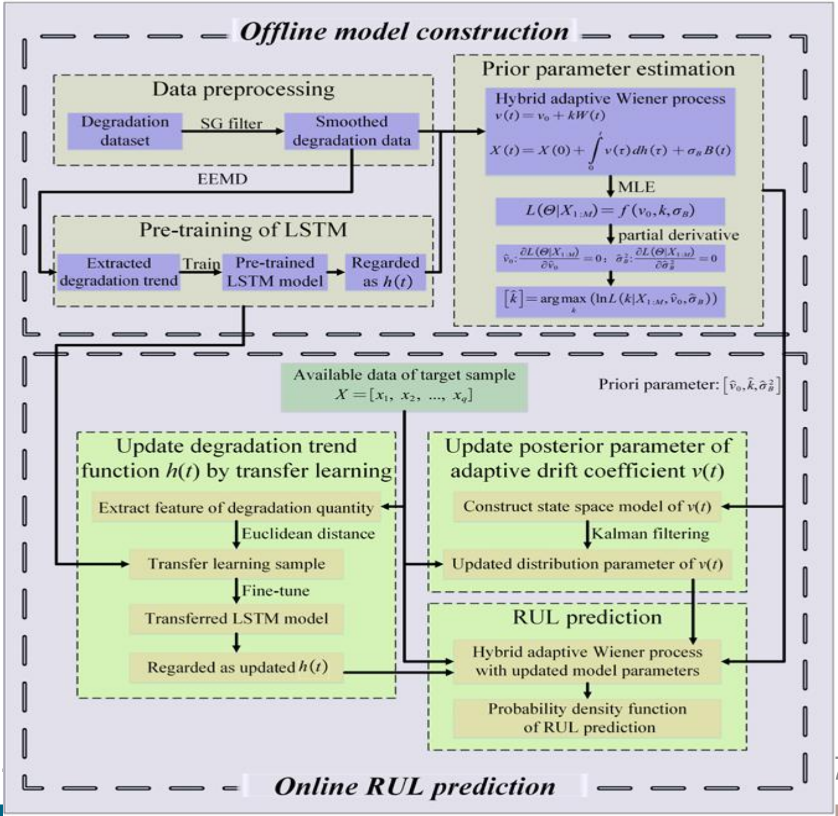


Fig. 1. Flowchart of LSTM-TL-WP model for RUL prediction.

부산대학교 PUSAN NATIONAL UNIVERSITY

Related works (hybrid degradation modeling)

- Remaining useful life prediction using a hybrid transfer learning-based adaptive wiener process model (2025)
  - Reliability engineering and system safety



Related works (hybrid degradation modeling)

- Remaining useful life prediction of rotating machine via long short-term memory neural network with uncertainty quantification (2026)
  - Engineering applications of artificial intelligence

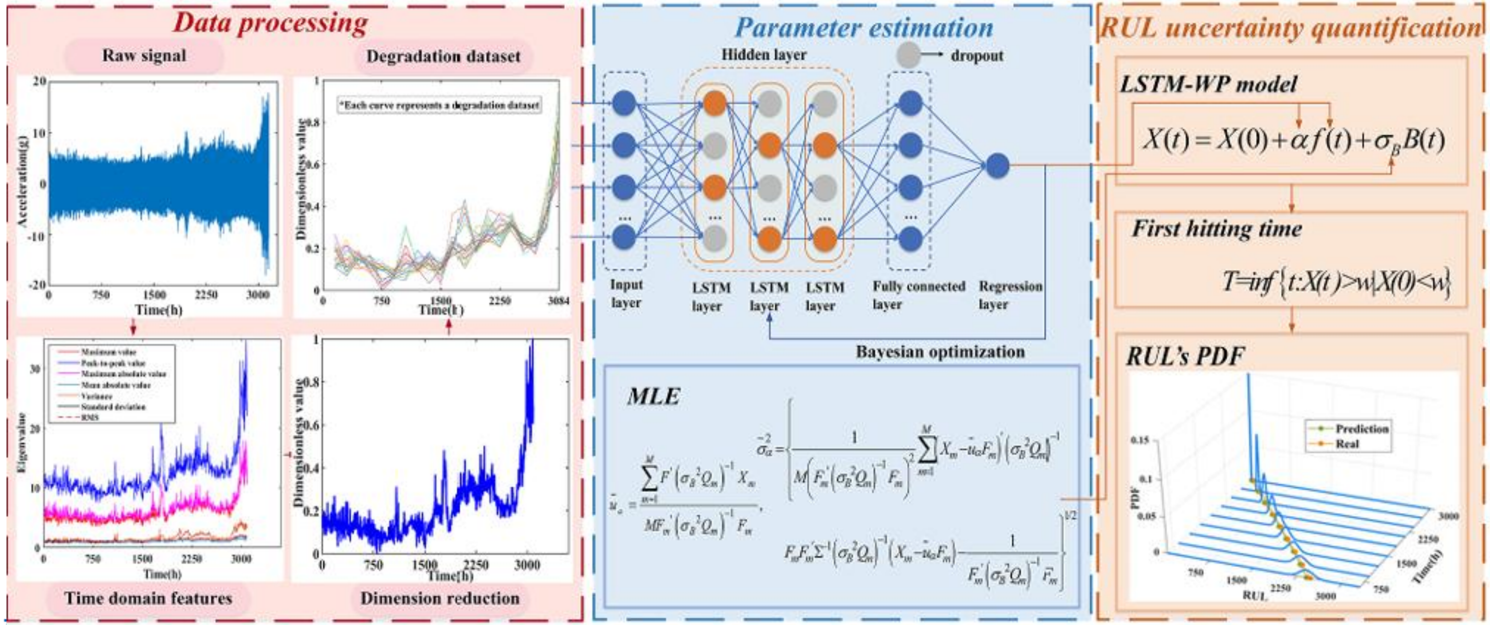
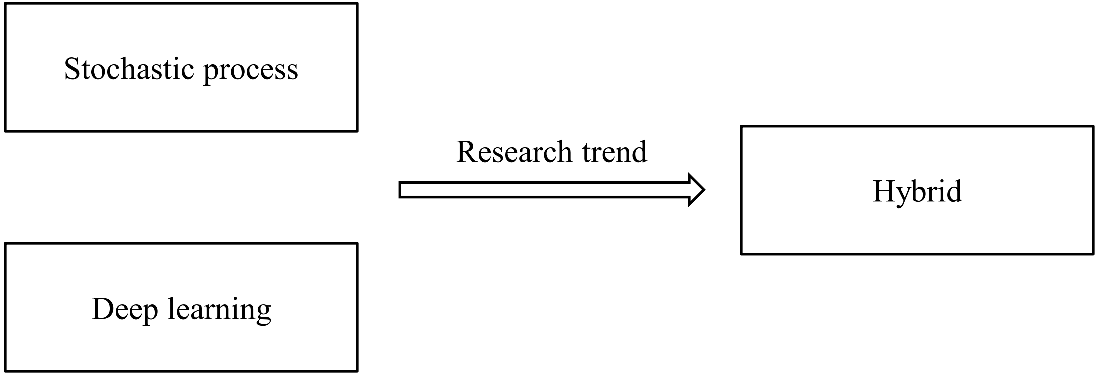
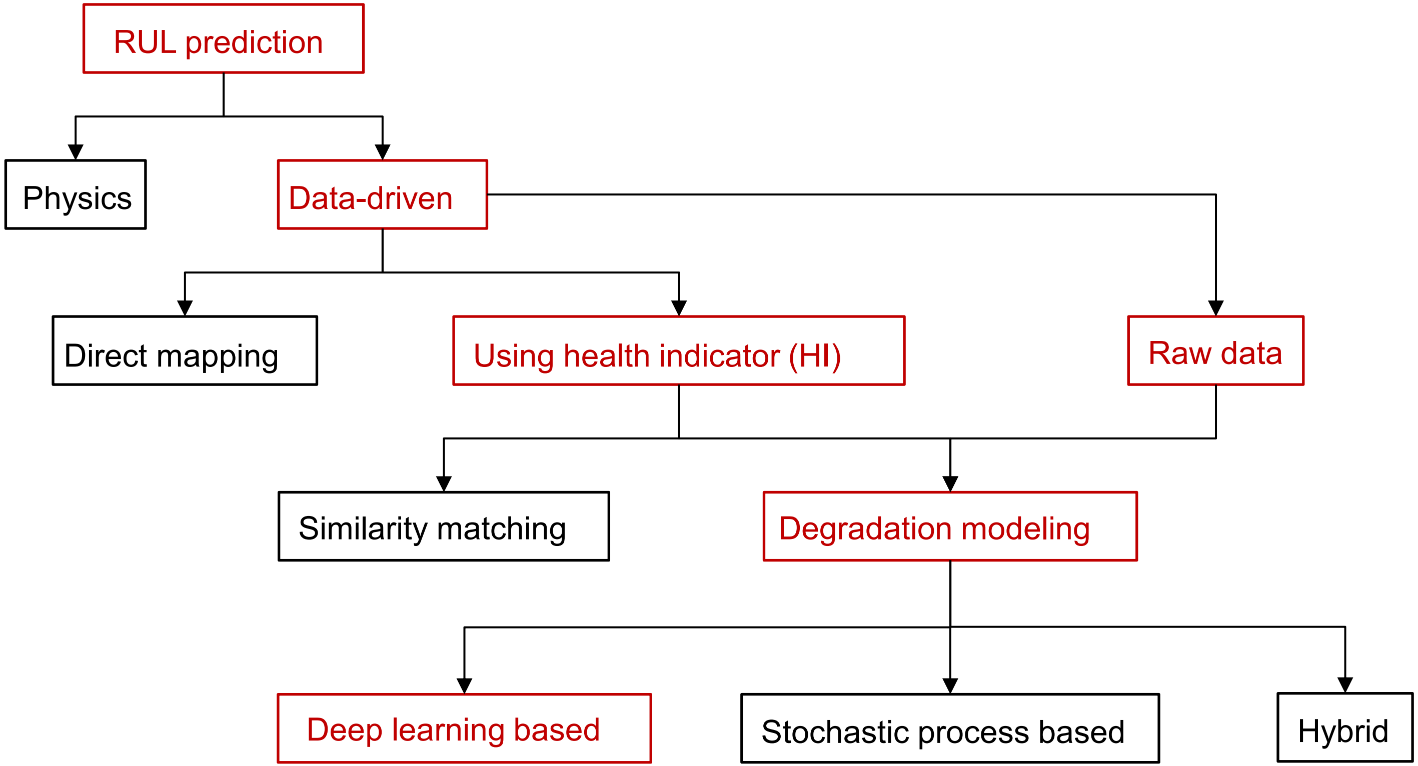


Fig. 2. Flowchart of LSTM-WP-Bo model for RUL prediction.

Summary



Preliminary



Summary

- Recursive forecasting (using deep learning)
  - Gated dual attention unit neural networks for remaining useful life prediction of rolling bearings
  - Dual-thread gated recurrent unit for gear remaining useful life prediction
  - Cocktail LSTM and its application into machine remaining useful life prediction
  - Gated transient fluctuation dual attention unit network for long-term remaining useful life prediction of rotating machinery using IIoT
  - Trend-fluctuation correlated attention unit for remaining useful life prediction
  - …

Related works (DL-based degradation modeling)

- Gated Transient Fluctuation Dual Attention Unit Network for Long-Term Remaining Useful Life Prediction of Rotating Machinery Using IIoT (2024)
  - IEEE Internet of Things Journal

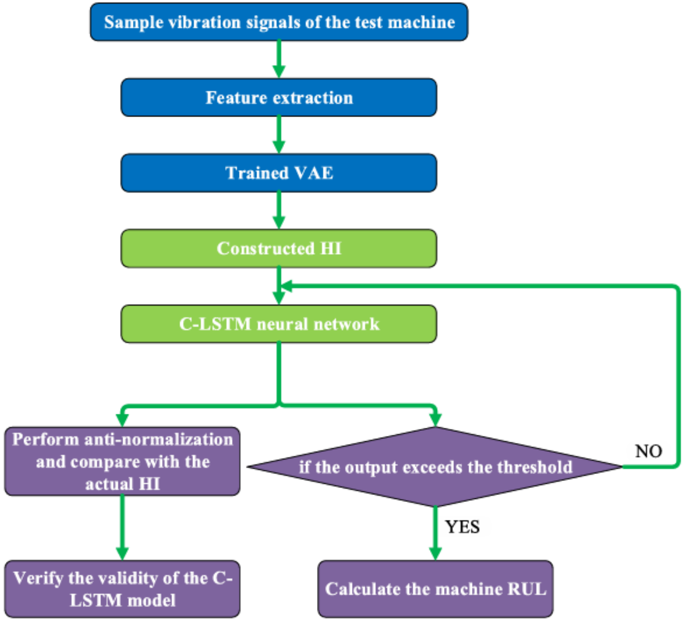
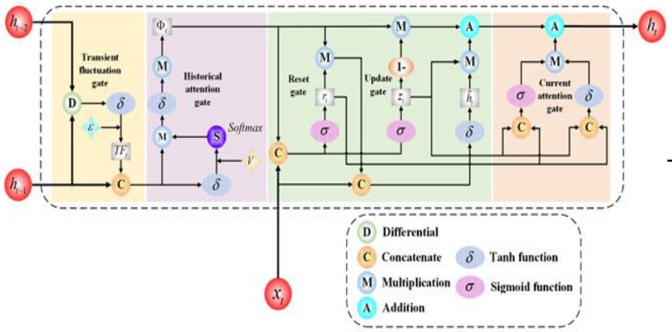
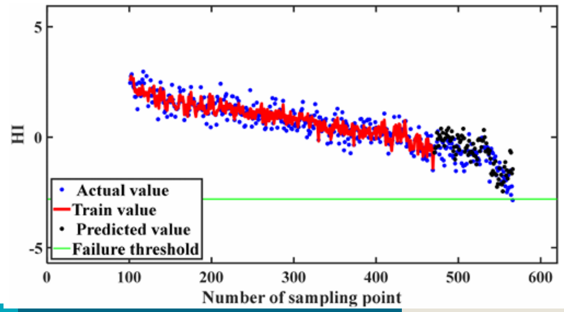


Fig. 3. Proposed machine RUL prediction flowchart.





If degradation start time: t, then

$$\hat{y}_{t+1} = f(y_{t-L}, \dots, y_t)$$

$$\hat{y}_{t+2} = f(y_{t-L+1}, \dots, \hat{y}_{t+1})$$

$$\hat{y}_{t+3} = f(y_{t-L+2}, \dots, \hat{y}_{t+1}, \hat{y}_{t+2})$$

$$\vdots$$

Until reach failure threshold