Stochastic Modeling for Two-Stage and Multivariate Degradation Processes

TANG Yincai (汤银才)

School of Statistics

East China Normal University

Joint work with：徐安察, 王平平, 庄亮亮

SRSE2024, Hangzhou, October 11-14, 2024

Outline

Traditional types of failure data

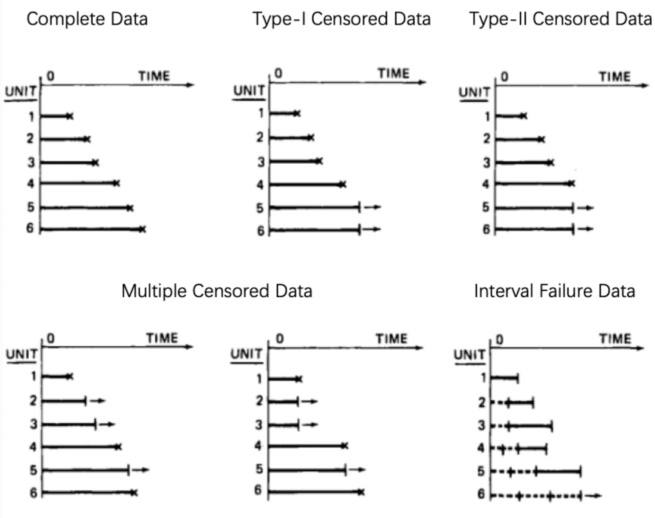


Figure 1.1: Types of failure data.

Accelerated life tests

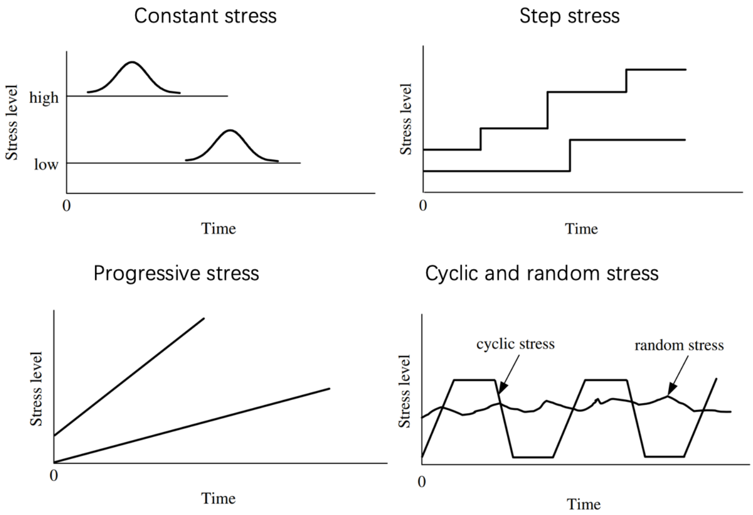


Figure 1.2: Types of accelerated life tests.

Degradation data from DT or ADT

- From hard failure to soft failure: with high technology, many products are designed with high reliability, and failure data are hard to collected for these products, even using accelerated life test.
- Degradation data provide a useful resource for obtaining reliability information for highly reliable products. Examples:
  - Loss of light output from an LED array
  - Power output decrease of photovoltaic arrays
  - Corrosion in a pipeline
  - Vibration from a worn bearing in a wind turbine
  - Loss of gloss and colour of an automobile finish
- Accelerated degradation tests can help more quickly reveal lifetime-related information for high-reliability products.

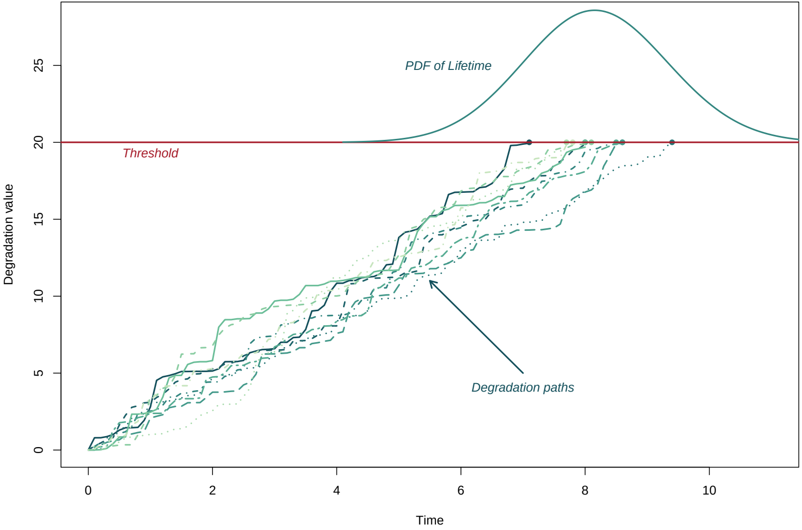


Figure 1.3: Degradation test and life time distribution

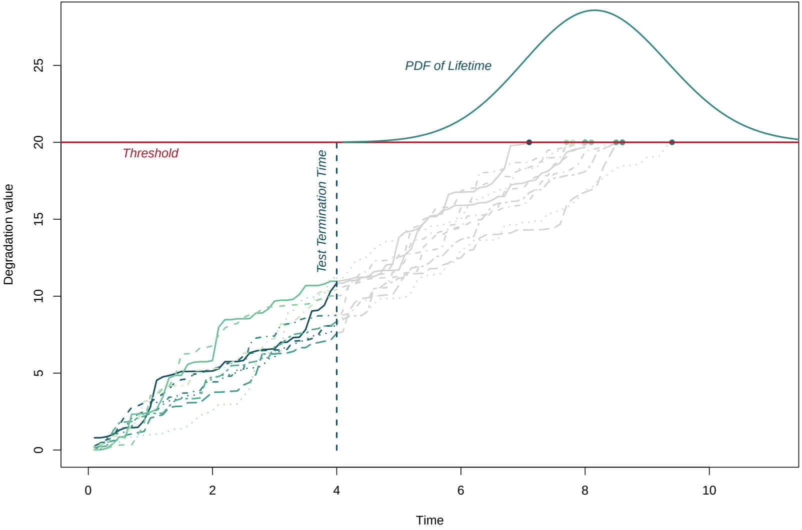


Figure 1.4: Time censoring degradation test

- Let $Y(t)$ be the degradation process of the performance characteristic (PC), and ω be the failure threshold level.
- Define that the lifetime of product $T = \inf\{t : Y(t) \geq \omega\}$.

Degradation models

- General degradation path models

$$Y(t) = D(t|\boldsymbol{\beta}, \boldsymbol{b}) + \epsilon.$$

- Stochastic degradation models, i.e., Wiener process (Liao and Tseng, 2006), gamma process (Park and Padgett, 2005), inverse Gaussian process (Wang and Xu, 2010), exponential dispersion process (Zhou and Xu, 2019), variance gamma, Ornstein–Uhlenbeck, etc.
- Two review papers: Ye and Xie (2015), Zhang et al. (2018).

Outline

Outline

Motivated example: OLED degradation data

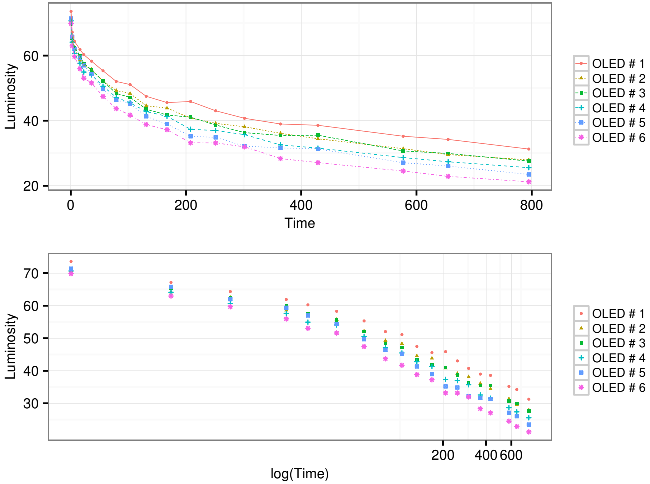


Figure 2.1: Degradation paths of OLEDs: luminosity against time (top) and luminosity against (log) transformed inspection time (bottom).

Related Literature for two-phase degradation models (I)

- Tseng et al. (1995) to analyze the two-phase degradation data tend to delete early degradation measurements.
- Bae and Kvam (2006) introduces a change-point regression model to fit degradation paths.
- A bi-exponential model with random-coefficients is proposed in Bae et al. (2008) and compared with a exponential model.
- Bae et al. (2015) adopt a Bayesian approach to model the two-phase degradation by using a change-point regression model under the continuity constraint.
- With the prior information taken into account, the bi-exponential model is reestablished in Yuan et al. (2016) under the Bayesian framework.

Related Literature for two-phase degradation models (II)

Two-phase degradation modeling

1. Wiener process: Wang et al. (2018a, 2018b), Zhang et al. (2019), Lin et al. (2021), Ma et al. (2023), etc.
2. Gamma process: Ling et al. (2019), Lin et al. (2021).
3. IG process: Duan and Wang (2017). Limitations：
   (i) Constraints on locations of change points;
   (ii) Insufficient considerations for deriving the lifetime distribution;
   (iii) Neglecting the uncertainty in estimation.

Wiener process with measurement error

- From the physical point of view, for many products, the degradation increment in an infinitesimal time interval can be viewed as an additive superposition of a large number of small external effects.
- Wang (2010) studies Wiener process with random effects for degradation data.
- The objective Bayesian method is developed for the accelerated degradation test based on Wiener process in Guan et al. (2016).
- Ye et al. (2013) incorporate the measurement error in the Wiener process on account of the imperfect inspection.

Contributions

Contributions

- We propose a change-point Wiener process with measurement error (CPWPME) through specifying the drift of the Wiener process as a two-phase linear function of time.
- Besides, the variability of the degradation paths for different OLEDs drives us to consider the unit-specific coefficients and change-points in the drift function.

Wiener process

Definition of Wiener process

- $W(t)$: the observed degradation character at t
- $Y(t) = W(0) - W(t)$: degradation value at t
- A well-adopted form of the Wiener process is written as

$$X(t) = m(t) + \sigma \mathcal{B}(t), \tag{2.1}$$

where $m(t)$ is the drift, σ is the diffusion coefficient, and $\mathcal{B}(t)$ is the standard Brownian motion with properties: i) $\mathcal{B}(0) = 0$; ii) $\mathcal{B}(t), t > 0$, has stationary independent Gaussian increments, i.e. $\Delta\mathcal{B} = \mathcal{B}(t + \Delta t) - \mathcal{B}(t)$ follows a normal distribution $\mathcal{N}(0, \Delta t)$.

Two-phase Wiener degradation process

Drift function of ith unit

The drift function of ith unit, $i = 1, \dots, n$, where n is the number of units, $m_i(t; \beta_i^H, \beta_i^L, \tau_i)$ is formulated as

$$m_i(t; \beta_i^H, \beta_i^L, \tau_i) = \begin{cases} \beta_i^H t, & \text{if } t \le \tau_i \\ \beta_i^L (t - \tau_i) + \beta_i^H \tau_i, & \text{if } t > \tau_i, \end{cases} \tag{2.2}$$

where β_i^H is the higher degradation rate at the early stage, β_i^L is the lower degradation rate at the stable stage, and τ_i is the change-point for the ith individual unit.

Notation

- $\boldsymbol{t}_i \equiv (t_{i,1}, \ldots, t_{i,n_i})$: the ordered inspection time points for the ith unit.
- $\boldsymbol{y}_i \equiv (y_{i,1}, \ldots, y_{i,n_i})$: the corresponding observed degradations of $\boldsymbol{Y}_i \equiv (Y_{i,1}, \ldots, Y_{i,n_i})$.
- n_i: the number of inspection time points.
- n: number of unit.
- $X_{i,j} = X(t_{i,j})$.
- $\Delta y_{i,j} \equiv (y_{i,j+1} - y_{i,j})$: the observed degradation increment of $\Delta Y_{i,j} \equiv (Y_{i,j+1} - Y_{i,j})$ on the time interval $(t_{i,j}, t_{i,j})$.
- $\Delta t_{i,j} = t_{i,j+1} - t_{i,j}$.

Proposed model: CPWPME

$$Y_{i,j} = X_{i,j} + \epsilon_{i,j}, \tag{2.3}$$

where $\epsilon_{i,j}$ is the measurement error and follows $\mathcal{N}(0, \gamma^2)$.

Statistical properties of $\Delta Y_{i,j}$

Expectation

$$\Delta m_{i,j} = \begin{cases} \beta_i^H \Delta t_{i,j}, & \text{if } \tau_i \geq t_{i,j+1}, \\ \beta_i^H(\tau_i - t_{i,j}) + \beta_i^L(t_{i,j+1} - \tau_i), & \text{if } t_{i,j} \leq \tau_i < t_{i,j+1}, \\ \beta_i^L \Delta t_{i,j}, & \text{if } \tau_i < t_{i,j}, \end{cases}$$

Covariance between $\Delta Y_{i,g}$ and $\Delta Y_{i,k}$

$$\text{cov}(\Delta Y_{i,g}, \Delta Y_{i,k}) = \begin{cases} \sigma^2 \Delta t_{i,1} + \gamma^2, & \text{if } k = g = 1, \\ \sigma^2 \Delta t_{i,k} + 2\gamma^2, & \text{if } k = g > 1, \\ -\gamma^2, & \text{if } k = g+1 \text{ or } g = k+1, \\ 0, & \text{otherwise}, \end{cases}$$

where $k, g = 1, \ldots, n_i - 1$.

Joint probability density function (PDF) of $\Delta \boldsymbol{Y}_i$

- $\Delta \boldsymbol{m}_i \equiv (\Delta m_{i,1}, \ldots, \Delta m_{i,n_i-1})$: the mean vector.
- Σ_i: the covariance matrix with the (k, g)th element given by $\mathrm{cov}(\Delta Y_{i,g}, \Delta Y_{i,k})$ for the ith degradation increment vector.
- $\Delta \boldsymbol{Y}_i \equiv (\Delta Y_{i,1} \ldots, \Delta Y_{i,n_i-1})$.

Joint PDF of $\Delta \boldsymbol{Y}_i$

$$f_{\Delta \boldsymbol{Y}_i}(\Delta \boldsymbol{y}_i) = (2\pi)^{-\frac{n_i-1}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left[-\frac{(\Delta \boldsymbol{y}_i - \Delta \boldsymbol{m}_i)^\intercal \Sigma_i^{-1} (\Delta \boldsymbol{y}_i - \Delta \boldsymbol{m}_i)}{2}\right],$$

where $\Delta \boldsymbol{y}_i \equiv \{\Delta y_{i,1} \ldots, \Delta y_{i,n_i-1}\}$ is the ith observed degradation increment vector.

Likelihood function

- $\boldsymbol{\beta}^H \equiv (\beta_1^H, \ldots, \beta_n^H)$: the higher degradation rate parameter vector.
- $\boldsymbol{\beta}^L \equiv (\beta_1^L, \ldots, \beta_n^L)$: the lower degradation rate parameter vector.
- $\boldsymbol{\tau} \equiv (\tau_1, \ldots, \tau_n)$: the change-point parameter vector.
- $\boldsymbol{\theta} = (\boldsymbol{\beta}^H, \boldsymbol{\beta}^L, \boldsymbol{\tau}, \sigma^2, \gamma^2)$: all the parameters in the CPWPME model.

Likelihood function of $(\boldsymbol{\beta}^H, \boldsymbol{\beta}^L, \boldsymbol{\tau}, \sigma^2, \gamma^2)$

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} (2\pi)^{-\frac{n_i-1}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left[-\frac{(\Delta\boldsymbol{y}_i - \Delta\boldsymbol{m}_i)^\intercal \Sigma_i^{-1} (\Delta\boldsymbol{y}_i - \Delta\boldsymbol{m}_i)}{2}\right]. \tag{2.4}$$

Prior specification

1. A truncated trivariate normal distribution is assigned for $\boldsymbol{\eta}_i$, for $i = 1, \ldots, n$, i.e. $\boldsymbol{\eta}_i \equiv (\beta_i^H, \beta_i^L, \tau_i) \sim \mathcal{MVN}(\boldsymbol{\omega}, \Omega)\mathcal{I}_{\{\beta_i^H>0,\ \beta_i^L>0,\ \tau_i>0,\}}$, where $\boldsymbol{\omega}$ is the mean vector and Ω is the covariance matrix, and $\mathcal{I}_{\{\beta_i^H<0,\ \beta_i^L<0,\ \tau_i<0,\}}$ is the indicator function.
2. The conjugate prior for $\boldsymbol{\omega}$ is also a trivariate normal distribution $\mathcal{MVN}(\boldsymbol{\kappa}, \Psi)$. Let the mean vector $\boldsymbol{\kappa} = \mathbf{0}_3$ and the covariance matrix $\Psi = 10^{-6}\boldsymbol{I}_3$, where $\mathbf{0}_3$ is a three dimensional zero vector and $\boldsymbol{I}_3$ is a 3×3 identity matrix.
3. Decompose the Ω as $\Omega = \Theta Q \Theta$, where $\Theta = \text{diag}\{\theta_1, \theta_2, \theta_3\}$. Assign the inverse-Wishart distribution $\mathcal{IW}(\rho, \boldsymbol{S})$ for Q. Specify the Gamma distribution $\mathcal{G}(a_\theta, b_\theta)$ as the prior distribution of θ_k for $k = 1, 2, 3$. Let $\rho = 4$, $S = \boldsymbol{I}_3$, and $a_\theta = 0.0001,\ b_\theta = 0.0001.$
4. The inverse Gamma distributions $\mathcal{IG}(a_\sigma, b_\sigma)$ and $\mathcal{IG}(a_\gamma, b_\gamma)$ are assigned for σ^2 and γ^2 respectively. Let $a_\sigma = b_\sigma = a_\gamma = b_\gamma = 0.001$.

Posterior inference

- $\boldsymbol{\theta} \equiv (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n, \sigma^2, \gamma^2, Q, \theta_1, \theta_2, \theta_3)$.

Joint posterior distribution of $\boldsymbol{\theta}$

$$\begin{aligned}\pi(\boldsymbol{\theta}|y) \propto & \mathcal{L}(\boldsymbol{\beta}^H, \boldsymbol{\beta}^L, \boldsymbol{\tau}, \sigma^2, \gamma^2)\left[\prod_{i=1}^{n} \pi(\boldsymbol{\eta}_i|\boldsymbol{\omega}, \Omega)\right] \pi(\boldsymbol{\omega}|\boldsymbol{\kappa}, \Psi)\pi(Q|\rho, S) \\ & \times \left[\prod_{k=1}^{3} \pi(\theta_k|a_k, b_k)\right] \pi(\sigma^2|a_\sigma, b_\sigma)\pi(\gamma^2|a_\gamma, b_\gamma)\end{aligned} \tag{2.5}$$

Posterior output: failure-time distribution

- The OLED devices are regarded to have failed if their luminosity fall below 50% of their initial luminosity.
- Define the 50% of six OLEDs' initial luminosity as a vector $(\mathcal{F}_1, \ldots, \mathcal{F}_6)$.
- Failure-time of the ith testing unit is defined as $T_i = \inf\{t | Y(t) \leq \mathcal{F}_i\}$, where $\mathcal{F}_i$ is the failure threshold of ith device.

Cumulative distribution function (CDF) of the failure-time

$$F_{T_i}(t) = \begin{cases} F_{IG}\left(t; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2}\right), & \text{if } t \leq \tau_i, \\ F_{IG}\left(t; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i)^2}{\sigma^2}\right), & \text{if } t > \tau_i, \end{cases} \quad (2.6)$$

for $i = 1, \ldots, N$. Here, $F_{IG}(x; \mu, \lambda)$ denotes an inverse Gaussian (IG) distribution with mean vector μ and shape parameter λ.

Posterior output: mean time to failure (MTTF)

MTTF of each OLED device

$$\begin{aligned}\mathcal{E}[T_i] =& \frac{\mathcal{F}_i}{\beta_i^H}\left[1 - F_{IG}\left(\frac{\mathcal{F}_i^2}{\tau_i {\beta_i^H}^2}; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2}\right)\right] + \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L} \\ &\times F_{IG}\left(\frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\tau_i {\beta_i^L}^2}; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\sigma^2}\right),\end{aligned}$$

for $i = 1, \ldots, N$.

Simulation Study

1. The CPWPME data are randomly generated under the following three different setup for the number of units and the number of inspection time points, i.e

   Scenario I: $n = 5$, $n_i = 16$;
   Scenario II: $n = 5$, $n_i = 21$;
   Scenario III: $n = 10$, $n_i = 21$.

2. The inspection time points are chosen from 0 to 18 with identical time intervals under each scenario.

Table 2.1: Parameter estimation results for scenario I.

Stat.	β_1^H	β_2^H	β_3^H	β_4^H	β_5^H	β_1^L	β_2^L	β_3^L	β_4^L	β_5^L
True	6.720	7.082	6.626	7.713	7.147	1.741	2.154	2.233	2.182	1.903
Bias	0.142	0.017	0.177	-0.259	-0.030	0.223	-0.061	-0.096	-0.122	0.128
SE	0.318	0.312	0.341	0.398	0.325	0.465	0.430	0.416	0.422	0.480
RMSE	0.348	0.312	0.384	0.474	0.326	0.516	0.434	0.427	0.439	0.496
CP	0.928	0.990	0.910	0.876	0.960	0.960	0.984	0.968	0.960	0.966
Stat.	τ_1	τ_2	τ_3	τ_4	τ_5	$\omega[1]$	$\omega[2]$	$\omega[3]$	σ^2	γ^2
True	12.828	12.214	11.660	10.787	12.616	7.000	2.000	12.000	2.000	1.000
Bias	-0.284	-0.085	0.074	0.291	-0.209	0.067	0.057	-0.021	0.048	0.185
SE	0.490	0.413	0.523	0.405	0.463	0.202	0.320	0.257	0.704	0.487
RMSE	0.566	0.422	0.527	0.498	0.508	0.213	0.324	0.257	0.705	0.520
CP	0.924	0.966	0.972	0.940	0.940	0.994	0.996	1.000	0.974	0.970

Table 2.2: Parameter estimation results for scenario II.

Stat.	β_1^H	β_2^H	β_3^H	β_4^H	β_5^H	β_1^L	β_2^L	β_3^L	β_4^L	β_5^L
True	6.720	7.082	6.626	7.713	7.147	1.741	2.154	2.233	2.182	1.903
Bias	0.149	-0.007	0.193	-0.297	-0.024	0.239	-0.034	-0.120	-0.136	0.119
SE	0.320	0.294	0.325	0.373	0.276	0.463	0.431	0.410	0.417	0.412
RMSE	0.353	0.294	0.378	0.477	0.276	0.521	0.432	0.427	0.438	0.429
CP	0.934	0.972	0.924	0.884	0.986	0.946	0.974	0.972	0.974	0.986
Stat.	τ_1	τ_2	τ_3	τ_4	τ_5	$\omega[1]$	$\omega[2]$	$\omega[3]$	σ^2	γ^2
True	12.828	12.214	11.660	10.787	12.616	7.000	2.000	12.000	2.000	1.000
Bias	-0.292	-0.052	0.069	0.331	-0.155	0.060	0.057	0.002	0.159	0.085
SE	0.432	0.403	0.448	0.450	0.347	0.191	0.303	0.221	0.717	0.393
RMSE	0.521	0.406	0.453	0.559	0.379	0.200	0.308	0.221	0.733	0.402
CP	0.930	0.974	0.970	0.918	0.970	1.000	1.000	0.998	0.944	0.960

Table 2.3: Parameter estimation results for scenario III.

Stat.	β_1^H	β_2^H	β_3^H	β_4^H	β_5^H	β_6^H	β_7^H	β_8^H	β_9^H	β_{10}^H
True	6.720	7.082	6.626	7.713	7.147	6.633	7.218	7.330	7.257	6.863
Bias	0.179	-0.022	0.259	-0.362	-0.064	0.249	-0.094	-0.152	-0.113	0.109
SE	0.255	0.227	0.248	0.308	0.219	0.245	0.233	0.251	0.246	0.226
RMSE	0.311	0.228	0.358	0.475	0.228	0.349	0.251	0.294	0.270	0.250
CP	0.908	0.978	0.894	0.814	0.980	0.890	0.978	0.946	0.966	0.976
Stat.	β_1^L	β_2^L	β_3^L	β_4^L	β_5^L	β_6^L	β_7^L	β_8^L	β_9^L	β_{10}^L
True	2.478	2.123	1.804	1.300	2.356	1.986	1.995	2.298	2.260	2.188
Bias	-0.174	0.031	0.195	0.347	-0.108	0.077	0.052	-0.162	-0.125	-0.038
SE	0.378	0.360	0.331	0.433	0.331	0.322	0.303	0.343	0.349	0.326
RMSE	0.416	0.361	0.384	0.555	0.348	0.331	0.307	0.379	0.371	0.328
CP	0.952	0.984	0.960	0.878	0.990	0.986	0.990	0.972	0.976	0.990
Stat.	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9	τ_{10}
True	12.503	12.428	12.041	10.910	12.339	11.969	11.915	11.194	11.738	12.229
Bias	-0.220	-0.199	-0.129	0.272	-0.128	-0.078	-0.012	0.329	0.093	-0.141
SE	0.369	0.311	0.331	0.349	0.331	0.327	0.292	0.358	0.286	0.345
RMSE	0.430	0.369	0.355	0.442	0.355	0.336	0.292	0.486	0.301	0.372
CP	0.930	0.946	0.970	0.894	0.962	0.968	0.992	0.858	0.980	0.960
Stat.	$\omega[1]$	$\omega[2]$	$\omega[3]$	σ^2	γ^2					
True	7.000	2.000	12.000	2.000	1.000					
Bias	0.058	0.088	-0.094	0.063	0.041					
SE	0.135	0.206	0.159	0.517	0.290					
RMSE	0.147	0.224	0.185	0.520	0.293					
CP	0.992	0.986	0.996	0.952	0.956					

Example: OLED data analysis

- The OLED degradation data was modeled using the CPWPME approach, with parameter estimation conducted via hierarchical methods.
- The Markov chains were initiated with a 20,000 iteration burn-in period, followed by an additional 30,000 iterations to obtain posterior samples for inference.
- Estimation results for the CPWPME model are summarized in Table 2.4. The estimated posterior means for ω and the covariance matrix Ω are given by:

$$\hat{\boldsymbol{\omega}} = (3.76, 9.74, 4.36)$$

$$\hat{\Omega} = \begin{pmatrix} 0.16350 & 0.00353 & -0.00368 \\ 0.00353 & 0.22370 & -0.00108 \\ -0.00368 & -0.00108 & 0.09092 \end{pmatrix}$$

Table 2.4: Parameter estimation based on the CPWPME model.

OLED	$\boldsymbol{\beta}^H$				$\boldsymbol{\beta}^L$			
	Est.	SE	2.5%	97.5%	Est.	SE	2.5%	97.5%
#1	3.665	0.224	3.204	4.100	9.800	0.302	9.217	10.440
#2	3.653	0.230	3.177	4.099	9.557	0.331	8.821	10.120
#3	3.697	0.222	3.243	4.143	9.819	0.294	9.271	10.460
#4	3.806	0.215	3.404	4.261	9.635	0.295	8.992	10.170
#5	3.775	0.230	3.323	4.250	9.808	0.290	9.251	10.420
#6	3.932	0.256	3.503	4.488	9.802	0.285	9.234	10.400

OLED	$\boldsymbol{\tau}$			
	Est.	SE	2.5%	97.5%
#1	4.475	0.131	4.210	4.749
#2	4.482	0.144	4.218	4.802
#3	4.364	0.124	4.101	4.586
#4	4.425	0.121	4.193	4.673
#5	4.165	0.152	3.872	4.447
#6	4.266	0.129	4.013	4.506

Models comparison

Benchmark models

- **CPWP**: The CPWP model is similar to our CPWPME model but omits measurement error.
- **TPLCP**: $y_{i,j} = \begin{cases} \zeta_i t_{i,j} - \kappa_i t_{i,j} + \epsilon_{i,j}, & j = 1, \ldots, \gamma_i \\ \zeta_i t_{i,j} - \kappa_i \varsigma_i + \epsilon_{i,j}, & j = \gamma_i + 1, \ldots, n_i \end{cases}$ for the ith item data, where $y_{i,j}$ is the jth observation measured at time $t_{i,j}$, and $\varsigma_i \in [t_{\gamma_i}, t_{\gamma_i+1})$ is the change-point of ith item. The error $\epsilon_{i,j}$ are assumed to be $i.i.d.$ $\mathcal{N}(0, \upsilon^2)$.
- **BE**: $y_{i,j} = \phi_i \exp(-(\gamma_i + \Delta\gamma_i)t_{ij}) + (1 - \phi_i)\exp(-\gamma_i t_{i,j}) + \epsilon_{i,j}$ where $i = 1, \ldots, I, j = 1, \ldots, n_i$, and error $\epsilon_{i,j}$ are assumed to be $i.i.d$ $\mathcal{N}(0, \omega^2)$. Denote $\boldsymbol{\phi} \equiv (\phi_1, \ldots, \phi_I)^\intercal$, $\boldsymbol{\gamma} \equiv (\gamma_1, \ldots, \gamma_I)^\intercal$, and $\Delta\boldsymbol{\gamma} \equiv (\Delta\gamma_1, \ldots, \Delta\gamma_I)^\intercal$.

Table 2.5: Parameter estimation of defferent benchmark models.

OLED	CPWP Model			TPLCP Model			BE Model		
	β^H	β^L	τ	ζ	κ	ς	ϕ	γ	$\Delta\gamma$
#1	3.799	9.269	4.373	-8.947	-5.440	4.506	0.647	4.741	-4.681
#2	3.797	9.300	4.369	-9.144	-5.640	4.486	0.647	4.741	-4.680
#3	3.815	9.305	4.310	-9.088	-5.560	4.256	0.634	4.742	-4.679
#4	3.849	9.334	4.364	-9.445	-5.555	4.439	0.622	4.743	-4.678
#5	3.850	9.429	4.210	-9.670	-5.936	4.025	0.609	4.745	-4.676
#6	3.903	9.407	4.287	-9.819	-5.654	4.215	0.596	4.745	-4.676

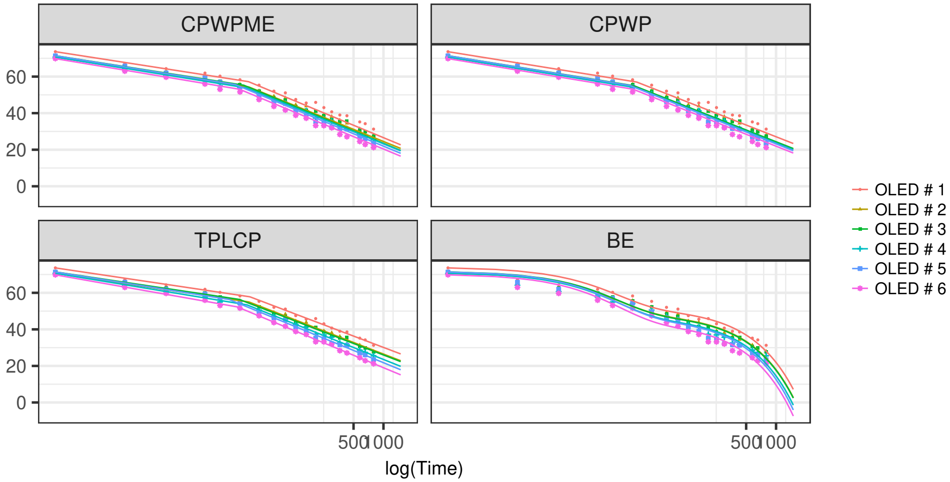


Figure 2.2: The posterior degradation path fits; luminosity vs. log(time) for each OLED data.

Mean squared prediction error (MSPE)

$$\mathrm{MSPE} = \sum_{j=1}^{n_i} \left(y_{i,j} - \hat{y}_{i,j}\right)^2, \tag{2.7}$$

where $\boldsymbol{y}_i = \{y_{i,1}, \ldots, y_{i,n_i}\}, i = 7$, is the degradation data of the 7th unit and $\hat{\boldsymbol{y}}_i = \{\hat{y}_{i,1}, \ldots, \hat{y}_{i,n_i}\}$ is the corresponding prediction value.

Table 2.6: MSPE for the 7th OLED degradation path.

Model	CPWPME	CPWP	TPLCP	BE
MSPE	360.04	406.04	436.29	678.53

- The CPWPME model's MSPE is much smaller than that of three other models, indicating its superiority.

MTTF

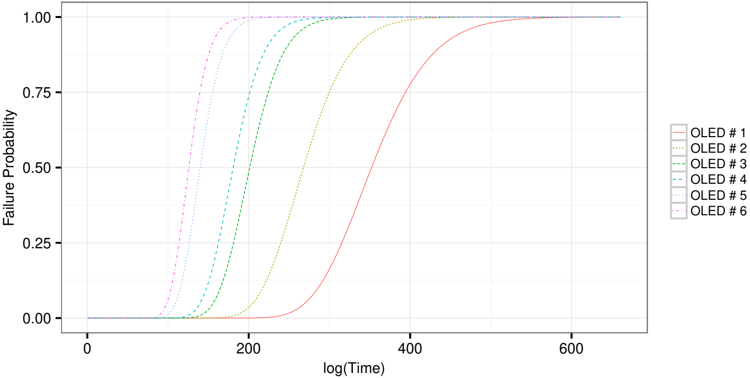


Figure 2.3: Posterior distribution of the failure-time for each OLED.

- The MTTF estimates for each unit are $(352.28, 267.84, 201.25, 180.67, 139.88, 125.77)$.

Outline

Motivated example: lithium batteries

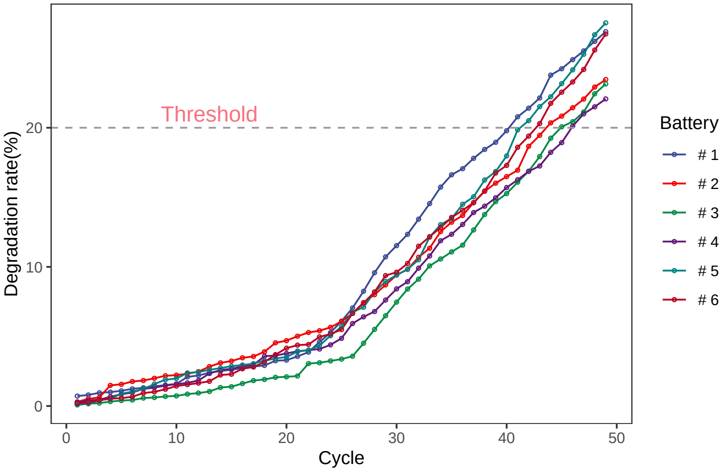


Figure 2.4: Capacity degradation data of lithium batteries.

Contributions

(i) A novel two-phase reparameterized IG (rIG) degradation model with distinct change points and model parameters for each individual system;

(ii) Derive the distribution of failure time and remaining useful life (RUL), and propose an adaptive replacement policy;

(iii) Employ bootstrap and Bayesian approach to generate interval estimates for the parameters.

Reparameterized IG distribution

Connection to IG distribution

The rIG distribution $rIG(\delta, \gamma)$ relates to the traditional IG distribution $IG(a, b)$ as $a = \delta/\gamma$ and $b = \delta^2$.

Moment generating function (MGF)

$$M_Y(t) = E(e^{ty}) = e^{\delta\gamma\left(1-\sqrt{1-\frac{2t}{\gamma^2}}\right)}. \tag{2.8}$$

Additive property

If $Y_1 \sim rIG(\delta_1, \gamma), Y_2 \sim rIG(\delta_2, \gamma)$, then $Y_1 + Y_2 \sim rIG(\delta_1 + \delta_2, \gamma)$.

PDF

If a random variable Y follows rIG distribution, then its PDF is

$$f_{rIG}(y|\delta,\gamma) = \frac{\delta}{\sqrt{2\pi}} e^{\delta\gamma} y^{-3/2} e^{-\left(\delta^2 y^{-1} + \gamma^2 y\right)/2}, \ y > 0, \ \delta > 0, \ \gamma > 0. \quad (2.9)$$

CDF

$$F_{rIG}(y|\delta,\gamma) = \Phi\left[\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right] + e^{2\delta\gamma}\Phi\left[-\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right], \quad (2.10)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

rIG process

Definition of rIG process

rIG process $\{Z(t), t \geq 0\}$ satisfies the following properties:

(i) $Z(0) = 0$ with probability one;

(ii) $Z(t)$ has independent increments. Specifically, $Z(t_2) - Z(t_1)$ and $Z(s_2) - Z(s_1)$ are independent for all $t_2 > t_1 \geq s_2 > s_1 \geq 0$;

(iii) For all $t > s \geq 0$, $Z(t) - Z(s)$ follows the rIG distribution $rIG(\delta(\Lambda(t) - \Lambda(s)), \gamma)$, where $\Lambda(t)$ is a monotone increasing function with $\Lambda(0) = 0$, δ and γ are unknown parameters.

- Denoted as $r\mathcal{IG}(\delta\Lambda(t), \gamma)$.
- The mean and variance of $\{Z(t), t \geq 0\}$, which are $\delta\Lambda(t)/\gamma$ and $\delta\Lambda(t)/\gamma^3$, respectively.

Two-phase rIG degradation model

Two-phase rIG degradation model

Suppose a system's performance characteristic degrades in two distinct phases, separated by a single change point.

$$
\begin{aligned}
&Y(t)|\tau \sim r\mathcal{IG}\left(m(t;\delta_1,\delta_2,\tau),\gamma\right),\ \tau \sim N\left(\mu_\tau,\sigma_\tau^2\right), \\
&m(t;\delta_1,\delta_2,\tau) = \begin{cases} \delta_1 t, & t \le \tau, \\ \delta_2\left(t-\tau\right)+\delta_1\tau, & t > \tau, \end{cases}
\end{aligned} \tag{2.11}
$$

where δ_1 and δ_2 are the drift parameters for $t \le \tau$ and $t > \tau$, respectively.

Failure-time

Let $T = \inf\{t \mid Y(t) \geq \mathcal{D}\}$, and $Y(t) = \begin{cases} Y_1(t), & t \leq \tau, \\ Y_1(\tau) + Y_2(t-\tau), & t > \tau. \end{cases}$

Conditional reliability function of T

- $0 \leq t \leq \tau$

$$\bar{F}_1(t \mid \tau) = P(T > t \mid \tau \geq t) = P(Y_1(t) < \mathcal{D} \mid \tau \geq t) = F_{r\mathcal{IG}}(\mathcal{D}|\delta_1 t, \gamma). \qquad (2.12)$$

- $t > \tau$

$$\begin{aligned} \bar{F}_2(t \mid \tau) =& P\left(Y(t) < \mathcal{D} \mid \tau < t\right) = P\left(Y_1(\tau) + Y_2(t-\tau) < \mathcal{D} \mid \tau < t\right) \\ =& \int_0^{\mathcal{D}} F_{r\mathcal{IG}}(\mathcal{D} - y_\tau|\delta_2(t-\tau), \gamma) f_1(y_\tau \mid \tau) \mathrm{d}y_\tau, \end{aligned} \qquad (2.13)$$

where y_τ represents the degradation value at τ, and $f_1(y_\tau \mid \tau)$ is the PDF of y_τ.

Failure-time

Unconditional reliability function of T

$$\begin{aligned} R(t) &= P\left(Y(t) < \mathcal{D}, \tau \geq t\right) + P\left(Y(t) < \mathcal{D}, 0 < \tau < t\right) \\ &= \bar{F}_1\left(t \mid \tau\right) \bar{G}_\tau(t) + \int_0^t g_\tau(\tau|\mu_\tau, \sigma_\tau^2) \bar{F}_2\left(t \mid \tau\right) d\tau, \end{aligned} \tag{2.14}$$

where $\bar{G}_\tau(t)$ is the survival function of random variable τ.

MTTF

$$\text{MTTF} = E(T) = \int_0^\infty R(t)dt. \tag{2.15}$$

RUL

Let $S_t = \inf\{x; Y(t+x) \geq \mathcal{D} \mid Y(t) < \mathcal{D}\}$.

Conditional reliability function of S_t

(i) When $x + t \leq \tau$:

$$\bar{F}_{S_t,1}(x \mid \tau) = F_{r\mathcal{IG}}(\mathcal{D} - Y(t)|\delta_1 x, \gamma). \tag{2.16}$$

(ii) When $t < \tau < x + t$:

$$\begin{aligned}\bar{F}_{S_t,2}(x \mid \tau) &= P(Y(t+x) < \mathcal{D} \mid Y(t) \leq \mathcal{D}) \\ &= \int_0^{\mathcal{D}} F_{r\mathcal{IG}}(\mathcal{D} - y_\tau|\delta_2(t+x-\tau), \gamma) f_1(y_\tau \mid \tau)\mathrm{d}y_\tau.\end{aligned} \tag{2.17}$$

(iii) When $\tau \leq t$:

$$\bar{F}_{S_t,3}(x \mid \tau) = F_{r\mathcal{IG}}(\mathcal{D} - Y(t)|\delta_2 x, \gamma). \tag{2.18}$$

RUL

Unconditional reliability function of S_t

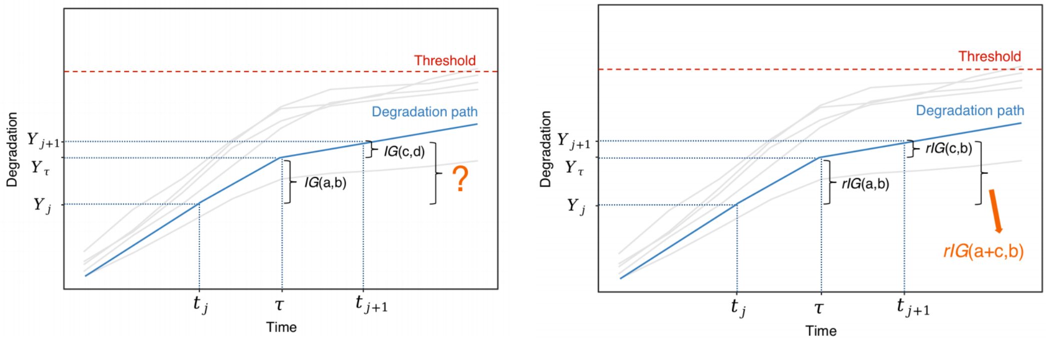
$$\begin{aligned} R_{S_t}(x) =& P(Y(t+x) < \mathcal{D}, t < x+t \leq \tau) \\ & + P(Y(t+x) < \mathcal{D}, t \leq \tau < x+t) + P(Y(t+x) < \mathcal{D}, t > \tau) \\ =& \bar{F}_{S_t,1}(x \mid \tau)\, \bar{G}_\tau(x+t) + \int_t^{x+t} g_\tau(\tau|\mu_\tau, \sigma_\tau^2) \bar{F}_{S_t,2}(x \mid \tau)\, d\tau \\ & + \int_0^t g_\tau(\tau) \bar{F}_{S_t,3}(x \mid \tau)\, d\tau. \end{aligned} \tag{2.19}$$

Mean of RUL at time t

$$\text{MRL} = E(S_t) = \int_0^\infty R_{S_t}(x) dx. \tag{2.20}$$

Data

- I systems under inspection in a degradation test.
- Deterioration pattern follows the two-phase rIG degradation model.
- $Y_{i,j}$ is the observed degradation value at the measurement time $t_{i,j}$, $i = 1 \ldots, I,\ j = 1, \ldots, n_i$, and $0 < t_{i,1} < \ldots < t_{i,n_i}$.
- Let $\Delta y_{i,j} = Y_{i,j} - Y_{i,j-1}$, $Y_{i,0} = 0$.
- Denote $\boldsymbol{\Delta Y}_i = (\Delta y_{i,1}, \ldots, \Delta y_{i,n_i})^\top$, $\boldsymbol{\Delta Y} = \left(\boldsymbol{\Delta Y}_1^\top, \cdots, \boldsymbol{\Delta Y}_I^\top\right)^\top$.



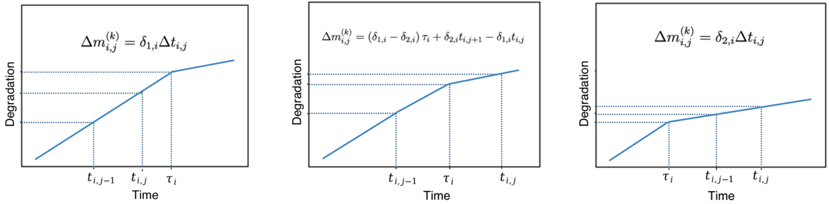
(a) IG process (b) Re-parameterized IG process

Conditional PDF of $\Delta y_{i,j}$

$$\Delta y_{i,j} \sim rIG\left(\Delta m_{i,j}^{(k)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right), \gamma\right),$$

$$\Delta m_{i,j}^{(k)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right) = \begin{cases} \delta_{1,i}\Delta t_{i,j} & k = 1, \\ \left(\delta_{1,i} - \delta_{2,i}\right)\tau_i + \delta_{2,i} t_{i,j} - \delta_{1,i} t_{i,j-1}, & k = 2, \\ \delta_{2,i}\Delta t_{i,j}, & k = 3, \end{cases}$$

$\Delta t_{i,j} = t_{i,j} - t_{i,j-1}$ and $t_{i,0} = 0$, $i = 1\ldots, I$, $j = 1, \ldots, n_i$.



(a) $\tau_i \geq t_{i,j}$ (b) $t_{i,j-1} \leq \tau_i < t_{i,j}$ (c) $\tau_i < t_{i,j-1}$

Figure 2.5: Three scenarios for change points and inspection time.

Conditional PDF of $\Delta y_{i,j}$

Let $\lambda_{i,j}^{(1)}=\mathcal{I}\left(\tau_i \geq t_{i,j}\right), \lambda_{i,j}^{(2)}=\mathcal{I}\left(t_{i,j-1} \leq \tau_i < t_{i,j}\right), \lambda_{i,j}^{(3)}=\mathcal{I}\left(\tau_i < t_{i,j-1}\right)$.

$$\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right) = \Delta m_{i,j}^{(1)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)^{\lambda_{i,j}^{(1)}} \times \Delta m_{i,j}^{(2)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)^{\lambda_{i,j}^{(2)}} \times \Delta m_{i,j}^{(3)}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)^{\lambda_{i,j}^{(3)}}.$$

$$\begin{aligned} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \tau_i, \gamma\right) = & \frac{\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)}{\sqrt{2\pi}} \exp\left\{\gamma \Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)\right\} \Delta y_{i,j}^{-3/2} \\ & \times \exp\left\{-\frac{\left[\Delta m_{i,j}\left(\delta_{1,i}, \delta_{2,i}, \tau_i\right)\right]^2 \Delta y_{i,j}^{-1} + \gamma^2 \Delta y_{i,j}}{2}\right\}. \end{aligned}$$

Likelihood function

- Let $\boldsymbol{\delta}_1 = (\delta_{1,1}, \ldots, \delta_{1,I})^\top$, $\boldsymbol{\delta}_2 = (\delta_{2,1}, \ldots, \delta_{2,I})^\top$ and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_I)^\top$.
- Denote $\boldsymbol{\eta} = \left(\boldsymbol{\delta}_1^\top, \boldsymbol{\delta}_2^\top, \gamma\right)^\top$, $\boldsymbol{\theta}_\tau = \left(\mu_\tau, \sigma_\tau^2\right)^\top$ and $\boldsymbol{\vartheta} = \left(\boldsymbol{\theta}_\tau^\top, \boldsymbol{\eta}^\top\right)^\top$.
- Given the observed data $\boldsymbol{\Delta Y}$, the likelihood function is

$$L_{obs}(\boldsymbol{\Delta Y}|\boldsymbol{\vartheta}) = \prod_{i=1}^{I} \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} f_{i,j}\left(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \tau_i, \gamma\right) g_\tau(\tau_i|\boldsymbol{\theta}_\tau)\mathrm{d}\tau_i. \tag{2.21}$$

Remark: Obtain a closed-form solution for the ML estimates of $\boldsymbol{\vartheta}$ is not feasible.

Bayesian analysis

$$
\begin{aligned}
&Y_i(t|\tau_i) \sim r\mathcal{IG}\left(m(t;\delta_{1,i},\delta_{2,i},\tau_i),\gamma\right),\ \tau_i \sim N\left(\mu_\tau,\sigma_\tau^2\right),\ i=1,\ldots,I,\\
&m(t;\delta_{1,i},\delta_{2,i},\tau_i) = \begin{cases} \delta_{1,i}t, & t \le \tau_i,\\ \delta_{2,i}\left(t-\tau_i\right)+\delta_{1,i}\tau_i, & t > \tau_i, \end{cases}\\
&\left(\mu_\tau,\sigma_\tau^2\right) \sim NIGa\left(\beta_\tau,\eta_\tau,\upsilon_\tau,\xi_\tau\right),\gamma \sim N(\omega,\kappa^2),\\
&\delta_{1,i} \sim N\left(\mu_1,\sigma_1^2\right),\delta_{2,i} \sim N\left(\mu_2,\sigma_2^2\right),\\
&\left(\mu_1,\sigma_1^2\right) \sim NIGa\left(\beta_1,\eta_1,\upsilon_1,\xi_1\right),\left(\mu_2,\sigma_2^2\right) \sim NIGa\left(\beta_2,\eta_2,\upsilon_2,\xi_2\right),
\end{aligned}
$$

where $NIGa(\cdot)$ denotes the normal-inverse gamma distribution.

Joint posterior distribution of $\boldsymbol{\theta}$

- Let $\boldsymbol{\theta} = \left(\boldsymbol{\vartheta}, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2\right)^\top$ be the parameter vector.
- According to Bayes' theorem, the joint posterior distribution of $\boldsymbol{\theta}$ can be derived as

$$\begin{aligned}\pi(\boldsymbol{\theta} \mid \boldsymbol{\Delta Y}) \propto \pi\left(\mu_\tau, \sigma_\tau^2\right) \pi\left(\mu_1, \sigma_1^2\right) \pi\left(\mu_2, \sigma_2^2\right) \pi\left(\gamma \mid \omega, \kappa\right) \pi\left(\tau \mid \mu_\tau, \sigma_\tau^2\right) \\ \times \pi\left(\boldsymbol{\delta}_1 \mid \mu_1, \sigma_1^2\right) \pi\left(\boldsymbol{\delta}_2 \mid \mu_1, \sigma_1^2\right) f_{\Delta Y}\left(\boldsymbol{\Delta Y} \mid \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\tau}, \gamma\right).\end{aligned} \tag{2.22}$$

- Employ the **Gibbs sampling algorithm** to generate posterior samples of the parameters, thereby facilitating Bayesian inference.

Adaptive replacement policy

- $0 = t_{i,0} < t_{i,1} < \cdots < t_{i,j}$ are discrete inspection times.
- $y_{i,j}$ represents the observed degradation value, $y_{i,1:j} = \{y_{i,1}, y_{i,2}, \ldots, y_{i,j}\}$.
- Iteratively update estimations of model parameters and RUL distributions, $f_{S_t}(x|y_{i,1:j})$.

Idea

1. Evaluate **candidate maintenance actions** at each inspection time point.
2. Determine **optimal preparation and maintenance actions** as data continues to be collected.

Policy assumption

- Maintenance is executed perfectly by replacing the system spare parts.
- Failure is detected only by inspections, and the cost of each inspection is c_i.
- An adequate supply of spare parts.
- Maintenance preparation time ϖ is usually required.

Two maintenance actions

At $t_{i,j}$, the decision maker has the option: replace the system or wait until the next inspection.

- **Corrective replacement**: implement if the system is found to have failed during the inspection, incurring a corrective replacement cost denoted as c_c.
- **Preventive replacement**: implement when it is expected that the system is nearing the failure state, incurring a preventive replacement cost denoted as c_p.

Candidate replacement time at $t_{i,j}$

$$\mathcal{T}_{i,j} = \inf_{T_{i,j}} \left\{ \int_0^{T_{i,j}-t_{i,j}} \frac{c_c + c_i\lfloor x + t_{i,j} \rfloor + c_b}{x + t_{i,j} + \varpi} f_{S_t}(x|y_{i,1:j})dx \right. \\ \left. + \int_{T_{i,j}-t_{i,j}}^{+\infty} f_{S_t}(x|y_{i,1:j}) \frac{c_p + c_i\lfloor T_{i,j} - \varpi \rfloor}{T_{i,j}} dx \right\},$$

where $\lfloor \psi \rfloor = \max\{h \in \mathbb{Z} \mid t_{i,h} \leq \psi\}$, and c_b is the downtime cost during the preparation time after system failure.

Optimal preparation and replacement time

As the values of $\mathcal{T}_{i,j}$ are successively updated,

$$\mathcal{T}_i' = \inf_{t_{i,j}}\{\mathcal{T}_{i,j} - t_{i,j} \leq \varpi\}, \quad \text{and} \quad \mathcal{T}_i^* = \mathcal{T}_i' + \varpi. \tag{2.23}$$

Performance evaluation

- Consider a set of I systems, each of which operates for a single cycle.
- Let $\mathbb{X}_i = \min\{\mathcal{T}_i^*, \mathcal{T}_i^{\mathsf{f}}\}$, where $\mathcal{T}_i^*$ represents predicted optimal maintenance time, and $\mathcal{T}_i^{\mathsf{f}}$ represents actual failure time.

Actual cost rate of the i-th system

$$CR_i = \begin{cases} \dfrac{c_p + c_i \lfloor \mathbb{X}_i - \varpi \rfloor}{\mathcal{T}_i^*}, & \mathbb{X}_i = \mathcal{T}_i^*, \\ \dfrac{c_c + c_i \lfloor \mathbb{X}_i \rfloor + c_b}{\mathcal{T}_i^{\mathsf{f}} + \varpi}, & \mathbb{X}_i = \mathcal{T}_i^{\mathsf{f}}, \end{cases} \tag{2.24}$$

Average cost rate for all systems

$$\overline{CR} = \frac{\sum_{i=1}^{I} CR_i}{I}. \tag{2.25}$$

Simulation study

Simulation settings

- (I) $I = 5$ and $n_i = 20$; (II) $I = 5$ and $n_i = 40$; (III) $I = 8$ and $n_i = 20$.
- Considering the heterogeneity, we generate $\delta_{1,1}, \ldots, \delta_{1,I}$ from $N(4, 1)$, $\delta_{2,1}, \ldots, \delta_{2,I}$ from $N(15, 1)$, and $\tau_1, \ldots, \tau_I$ from $N(10, 1)$.
- For each scenario, we generate 500 samples to reduce the effects of randomness on the results.

Simulation study

- **Bayesian method**:
  - Flat priors: $(\mu_\tau, \sigma_\tau) \sim NIGa(8, 100, 0.01, 0.01)$, $(\mu_1, \sigma_1) \sim NIGa(1, 100, 0.01, 0.01)$, $(\mu_2, \sigma_2) \sim NIGa(2, 100, 0.01, 0.01)$, and $\gamma \sim N(5, 100)$.
  - Initiate a burn-in period comprising $\mathcal{L} = 5000$ iterations, and an additional $\mathcal{S} - \mathcal{L} = 5000$ iterations are conducted to obtain posterior samples.
- **ML method**: the point estimates are calculated by the EM algorithm, corresponding interval estimates are calculated by parametric bootstrap method with $\mathcal{B} = 500$.
- Indexes of assessing different methods: **relative bias** (RB), **rooted mean squared error** (RMSE) and 95% **coverage probability** (CP).

Parameter estimation performance of two methods

Table 2.7: Parameter estimation from Bayes and ML methods for two scenarios.

Scen.	Meth.	Stat.	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$	$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$	γ
I	Bayes	RB	0.024	0.029	-0.007	0.015	0.012	-0.026	0.019	0.023	0.056	0.003	0.011
		RMSE	1.326	1.363	1.357	1.332	1.330	0.422	0.424	0.476	0.422	0.431	0.168
		CP	0.956	0.953	0.946	0.953	0.957	0.941	0.925	0.900	0.928	0.926	0.964
	MLE	RB	0.057	0.039	0.040	0.057	0.050	0.065	0.071	0.057	0.078	0.060	0.057
		RMSE	1.315	1.381	1.302	1.401	1.508	0.641	0.645	0.576	0.667	0.739	0.308
		CP	0.889	0.922	0.878	0.900	0.833	0.922	0.922	0.900	0.889	0.867	0.811
Scen.	Meth.	Stat.	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$	$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$	γ
II	Bayes	RB	-0.005	0.007	0.023	0.011	-0.005	-0.019	0.000	0.016	0.000	0.012	0.001
		RMSE	1.068	1.011	1.065	1.015	1.044	0.349	0.283	0.275	0.355	0.332	0.124
		CP	0.930	0.945	0.950	0.944	0.927	0.902	0.925	0.947	0.885	0.902	0.914
	MLE	RB	0.036	0.035	0.017	0.032	0.039	0.029	0.041	0.036	0.025	0.042	0.039
		RMSE	0.944	1.010	0.880	0.900	0.985	0.331	0.358	0.323	0.328	0.346	0.150
		CP	0.905	0.890	0.905	0.920	0.900	0.895	0.890	0.930	0.930	0.920	0.865

Model comparison in reliability estimation

- Linear $\Lambda(t) = t$; Power $\Lambda(t;\alpha) = t^{\alpha}$; Exponential $\Lambda(t;\alpha) = \exp(\alpha t) - 1$.

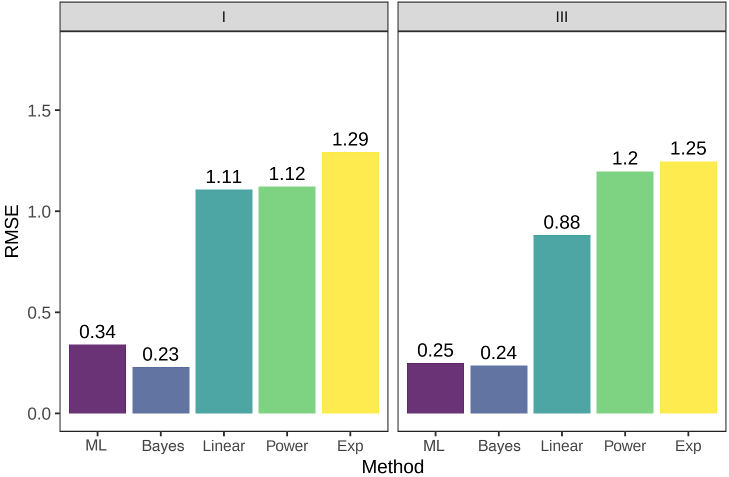


Figure 2.6: Average RMSE of MTTF estimators based on various models.

Parameter estimation with different models

Table 2.8: RMSE and RB results for different models.

Model	Training(30)		Prediciton (19)		Overall	
	RMSE	RB	RMSE	RB	RMSE	RB
Proposed	0.448	0.248	1.538	0.060	1.020	0.175
Linear	3.476	1.442	3.685	0.156	3.558	0.943
Power	2.057	0.568	2.475	0.113	2.229	0.391
Exp	0.908	0.313	1.611	0.065	1.230	0.217
Duan	0.434	0.239	1.976	0.075	1.276	0.175

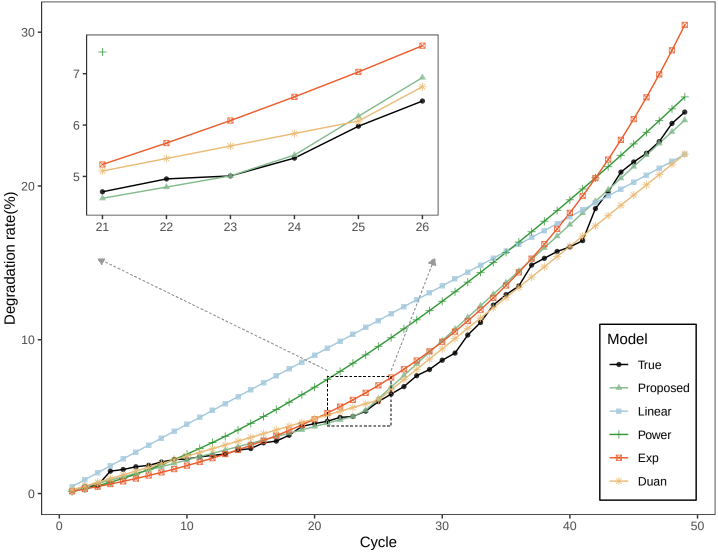
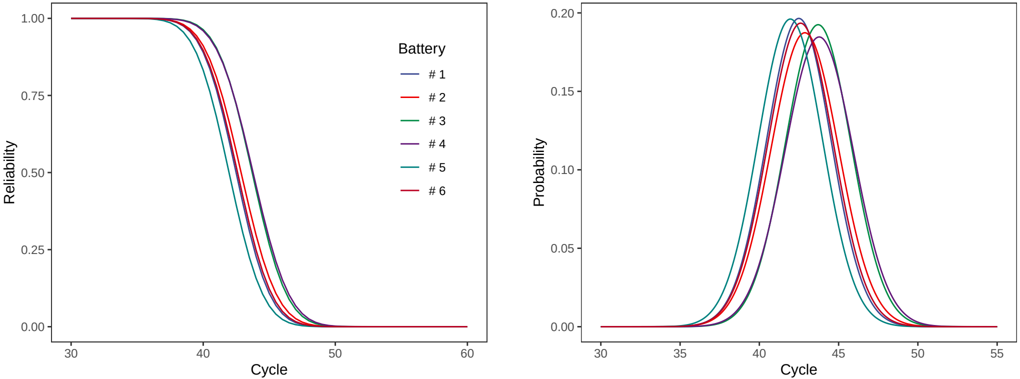
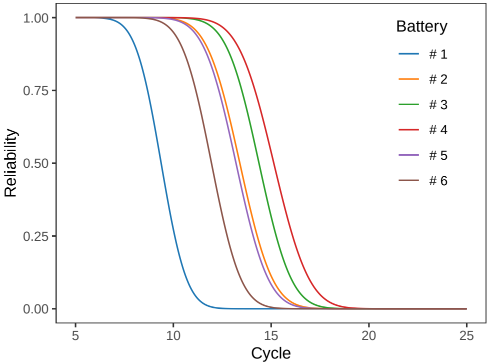


Figure 2.7: Degradation path training and prediction results for battery #2 using different methods, with a zoomed-in view of the potential change point locations.

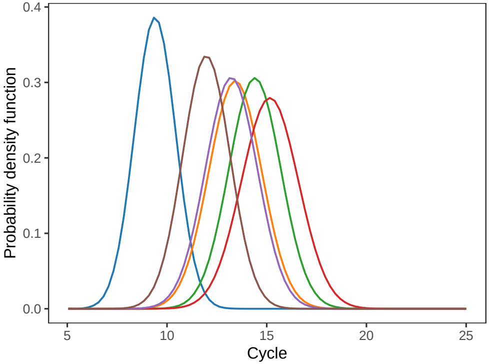


(a) (b)

Figure 2.8: Reliability and density functions of failure time based on HB method.



(a)



(b)

Figure 2.9: Reliability and density functions of RUL based on HB method.

RUL-based adaptive maintenance policy

1. Cycles 1-30 as historical data, continuously acquiring new data over time.
2. $c_i = 2, c_c = 600, c_p = 200$, and $c_b = 100$.
3. Maintenance preparation period is $\varpi = 1$.

Benchmark policies

i) **Classical replacement policy (CRP)**: preventive maintenance time is determined by the system's mean time to failure $\bar{\mathcal{T}}^F$.

ii) **Ideal replacement policy (IRP)**: the assumption of perfect predicted failure time $\mathcal{T}_i^P$.

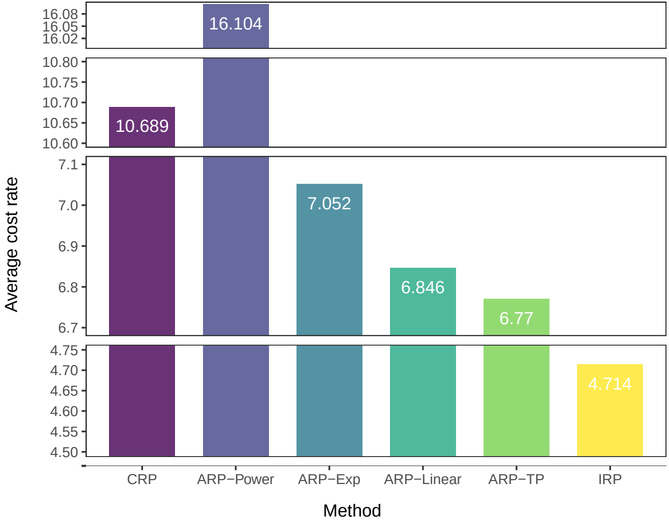


Figure 2.10: Average cost rate for each policy.

Outline

Outline

Motivated example: HMT degradation data

- To maintain the high availability and high efficiency of heavy machine tools, preventive maintenance and system health management are implemented.
- The heavy machine tools (HMT) have two important PCs: the positioning accuracy and the output power.
- HMT fails if the value of the positioning accuracy exceeds the threshold level $\omega_1 = 35$ or the value of the output power exceeds the threshold level $\omega_2 = 120$.

HMT with two PCs

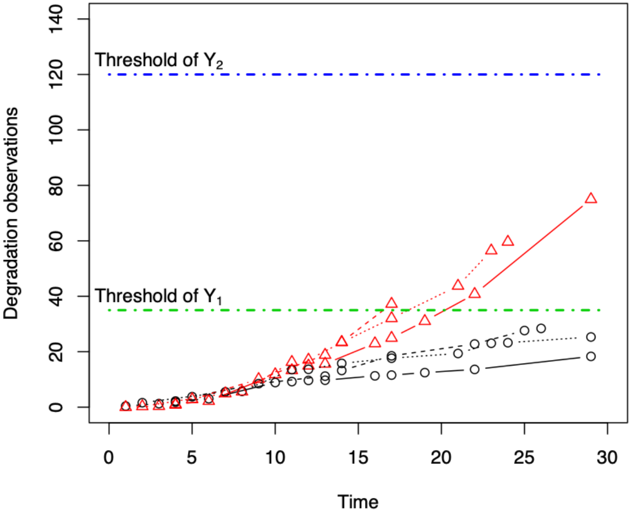


Figure 3.1: Degradation paths of the positioning accuracy and output power.

Objective

- The positioning accuracy is measured by programmed procedures, while measurements of the output power are recorded by the system operators, and may be missing at some time points.
- Historical information and experts' experience have indicated that these two performance indicators are correlated.

Objective

- How to build a model for bivariate degradation process?
- How to estimate the missing values of the output power?

Related Literature

- LED system consists of many LED lamps for different lighting purposes, and each LED lamp can be viewed as a PC in the LED system (Sari et al., 2009).
- A rubidium discharge lamp: The rubidium consumption and the light intensity (Sun and Balakrishnan, 2013).
- Modeling methods: using copula function (Sun et al. 2010,2012, Wang et al., 2014,2015, Peng et al., 2016, Duan and Wang, 2018).
    - Difficult to choose copula function.
    - Reliability function of product is not analytic.
    - No physical explanation.

Model

Bivariate Wiener degradation model

Assume two PCs in a system, degradation process of the s-th PC is:

$$Y_s(t) = \alpha\beta_s h_s(t,\gamma_s) + \sigma_s B_s(h_s(t,\gamma_s)), \ \ s = 1, 2, \tag{3.1}$$

- β_s and σ_s denote the drift parameter and the diffusion parameter.
- $h_s(t,\gamma_s)$ is a non-decreasing function of time with $h_s(0,\gamma_s) = 0$.
- $B_s(\cdot)$ is a standard Brownian motion, where $B_1(\cdot)$ and $B_2(\cdot)$ are independent.
- α is random, and follows normal distribution with mean 1 and variance δ^2.

Comments on α

- α could describe the unit-to-unit variation among the systems.
- With the same working environment for both PCs, α is a common factor affecting the degradation process.

Joint PDF of $Y_1(t)$ and $Y_2(t)$

$$\begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} \sim \boldsymbol{N}_2(\mu_H, \Sigma), \tag{3.2}$$

where $\mu_H = \binom{\beta_1 h_1(t,\gamma_1)}{\beta_2 h_2(t,\gamma_2)}$,

$$\Sigma = \begin{pmatrix} \sigma_1^2 h_1(t,\gamma_1) + \delta^2 \beta_1^2 h_1^2(t,\gamma_1) & \delta^2 \beta_1 \beta_2 h_1(t,\gamma_1) h_2(t,\gamma_2) \\ \delta^2 \beta_1 \beta_2 h_1(t,\gamma_1) h_2(t,\gamma_2) & \sigma_2^2 h_2(t,\gamma_2) + \delta^2 \beta_2^2 h_2^2(t,\gamma_2) \end{pmatrix}.$$

Failure-time distribution: joint CDF

- Denote that the threshold level of $Y_s(t)$ is ω_s, $s = 1, 2$.
- The lifetime of the s-th PC is defined as $T_s = \inf\{t : Y_s \geq \omega_s\}$.
- The joint CDF of T_1 and T_2 is

$$F(t_1, t_2) = A_1 + A_2 + A_3 + A_4,$$

where

$$A_1 = bvn\left(\frac{-\omega_1+\beta_1 h_1(t_1,\gamma_1)}{K_1}, \frac{-\omega_2+\beta_2 h_2(t_2,\gamma_2)}{K_2}, \frac{C_5}{K_1K_2}\right),$$

$$A_2 = \exp\left\{\frac{2\beta_2\omega_2}{\sigma_2^2} + \frac{2\beta_2^2\omega_2^2\delta^2}{\sigma_2^4}\right\} bvn\left(\frac{-\omega_1+\beta_1 h_1(t_1,\gamma_1)+C_1}{K_1}, \frac{-\omega_2-\beta_2 h_2(t_2,\gamma_2)-C_4}{K_2}, \frac{-C_5}{K_1K_2}\right),$$

$$A_3 = \exp\left\{\frac{2\beta_1\omega_1}{\sigma_1^2} + \frac{2\beta_1^2\omega_1^2\delta^2}{\sigma_1^4}\right\} bvn\left(\frac{-\omega_1-\beta_1 h_1(t_1,\gamma_1)-C_3}{K_1}, \frac{-\omega_2+\beta_2 h_2(t_2,\gamma_2)+C_2}{K_2}, \frac{-C_5}{K_1K_2}\right),$$

$$A_4 = \exp\left\{\frac{2\beta_1\omega_1}{\sigma_1^2} + \frac{2\beta_2\omega_2}{\sigma_2^2} + 2\delta^2\left(\frac{\beta_1\omega_1}{\sigma_1^2} + \frac{\beta_2\omega_2}{\sigma_2^2}\right)^2\right\}$$
$$\times bvn\left(\frac{-\omega_1-\beta_1 h_1(t_1,\gamma_1)-C_1-C_3}{K_1}, \frac{-\omega_2-\beta_2 h_2(t_2,\gamma_2)-C_2-C_4}{K_2}, \frac{C_5}{K_1K_2}\right).$$

Failure-time distribution: Reliability function

$$
\begin{aligned}
bvn(x_1, x_2, \theta) &= \frac{1}{2\pi\sqrt{1-\theta^2}} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \exp\left\{ -\frac{x^2 - 2\theta xy + y^2}{2(1-\theta^2)} \right\} \mathrm{d}x\mathrm{d}y, \\
K_1 &= \sqrt{\sigma_1^2 h_1(t_1, \gamma_1) + \beta_1^2 \delta^2 h_1^2(t_1, \gamma_1)}, \\
K_2 &= \sqrt{\sigma_2^2 h_2(t_2, \gamma_2) + \beta_2^2 \delta^2 h_2^2(t_2, \gamma_2)}, \\
C_1 &= 2\beta_1 h_1(t_1, \gamma_1)\beta_2 \omega_2 \delta^2 / \sigma_2^2, \\
C_2 &= 2\beta_1 \omega_1 \beta_2 h_2(t_2, \gamma_2)\delta^2 / \sigma_1^2, \\
C_3 &= 2\beta_1^2 h_1(t_1, \gamma_1)\omega_1 \delta^2 / \sigma_1^2, \\
C_4 &= 2\beta_2^2 h_2(t_2, \gamma_2)\omega_2 \delta^2 / \sigma_2^2, \\
C_5 &= \beta_1 h_1(t_1, \gamma_1)\beta_2 h_2(t_2, \gamma_2)\delta^2.
\end{aligned}
$$

Failure-time

- The lifetime of system is defined as $T = \min(T_1, T_2)$.

Reliability of system at time t

$$R(t) = F(t,t) + 1 - F_{T_1}(t) - F_{T_2}(t), \quad (3.3)$$

where $F_{T_s}(t)$ is the CDF of T_s:

$$F_{T_s}(t) = \Phi\left(\frac{\beta_s h_s(t,\gamma_s) - \omega_s}{\sqrt{\beta_s^2\delta^2(h_s(t,\gamma_s))^2 + \sigma_s^2 h_s(t,\gamma_s)}}\right)$$
$$+ \exp\left\{\frac{2\beta_s\omega_s}{\sigma_s^2} + \frac{2\beta_s^2\delta^2\omega_s^2}{\sigma_s^4}\right\}\Phi\left(-\frac{2\beta_s^2\delta^2\omega_s h_s(t,\gamma_s) + \sigma_s^2(\beta_s h_s(t,\gamma_s) + \omega_s)}{\sigma_s^2\sqrt{\beta_s^2\delta^2(h_s(t,\gamma_s))^2 + \sigma_s^2 h_s(t,\gamma_s)}}\right).$$

RUL

- The RUL of the s-th PC at time t_k:

$$L_{t_k}^{(s)} = \inf\{l : Y_s(l + t_k) \geq \omega_s | Y_s(t_j) < \omega_s, j = 1, 2, \cdots, k\},\ \ s = 1, 2,$$

where $t_1, \ \ldots, \ t_k$ are the measurement times.

- The RUL of the system:

$$L_{t_k} = \min(L_{t_k}^{(1)}, L_{t_k}^{(2)}).$$

- The reliability function of L_{t_k} at time l:

$$R_{L_{t_k}}(l) = F_{L_{t_k}}(l, l) + 1 - F_{L_{t_k}^{(1)}}(l) - F_{L_{t_k}^{(2)}}(l), \tag{3.4}$$

where $F_{L_{t_k}^{(s)}}(l)$ is the CDF of $L_k^{(s)}$, with analytical form

$$\begin{aligned}
F_{L_{t_k}^{(s)}}(l) =& \Phi\left(\frac{\tilde{\mu}\beta_s h_s(l,\gamma_s) - (\omega_s - Y_s(t_k))}{\sqrt{\beta_s^2\tilde{\delta}^2(h_s(l,\gamma_s))^2 + \sigma_s^2 h_s(l,\gamma_s)}}\right) \\
&+ \exp\left\{\frac{2\tilde{\mu}\beta_s(\omega_s - Y_s(t_k))}{\sigma_s^2} + \frac{2\beta_s^2\tilde{\delta}^2(\omega_s - Y_s(t_k))^2}{\sigma_s^4}\right\} \\
&\times \Phi\left(-\frac{2\beta_s^2\tilde{\delta}^2(\omega_s - Y_s(t_k))h_s(l,\gamma_s) + \sigma_s^2(\tilde{\mu}\beta_s h_s(l,\gamma_s) + (\omega_s - Y_s(t_k)))}{\sigma_s^2\sqrt{\beta_s^2\tilde{\delta}^2(h_s(l,\gamma_s))^2 + \sigma_s^2 h_s(l,\gamma_s)}}\right).
\end{aligned}$$

Data format

- Suppose that a total of n systems are tested in an experiment.
- For the i-th system, let y_{isj} be the j-th degradation observation of the s-th PC at the measurement time t_{isj}, $s = 1, 2, j = 1, 2, \ldots, m_{is}$.
- $y_{is0} = 0$. Let $z_{isj} = y_{isj} - y_{is(j-1)}$, and $\Lambda_{isj} = h_s(t_{isj}, \gamma_s) - h_s(t_{is(j-1)}, \gamma_s)$, $s = 1, 2, i = 1, 2, \ldots, n, j = 1, 2, \ldots, m_{is}$.
- Then for the i-th system, the model can be described as

$$z_{isj}|\alpha_i \sim \boldsymbol{N}(\alpha_i \beta_s \Lambda_{isj}, \sigma_s^2 \Lambda_{isj}), \text{and } \alpha_i \sim \boldsymbol{N}(1, \delta^2),$$

where $s = 1, 2,\ j = 1, 2, \ldots, m_{is}$.

Bayesian analysis

Prior

- $\beta_s \sim \boldsymbol{N}(1, 10^3); 1/\sigma_s^2 \sim \boldsymbol{IG}(0.01, 0.01); 1/\delta^2 \sim \boldsymbol{IG}(0.01, 0.01); \gamma_s \sim \boldsymbol{IG}(0.01, 0.01)$.

Gibbs sampling

- Full conditional posterior distribution of α_i is **normal distribution** with mean $\tilde{\mu}_i$ and variance $\tilde{\delta}_i^2$, where $\tilde{\delta}_i^2 = (\delta^{-2} + \sigma_1^{-2}\beta_1^2 h_1(t_{i1m_{i1}}, \gamma_1) + \sigma_2^{-2}\beta_2^2 h_2(t_{i2m_{i2}}, \gamma_2))^{-1}$, $\tilde{\mu} = \tilde{\delta}_i^2(\delta^{-2} + \sigma_1^{-2}\beta_1 y_{i1m_{i1}} + \sigma_2^{-2}\beta_2 y_{i2m_{i2}})$.
- Full conditional posterior distribution of β_s is **normal distribution** with mean $\tilde{\mu}_{\beta_s}$ and variance $\tilde{\sigma}_{\beta_s}^2$, where

$$\tilde{\sigma}_{\beta_s}^2 = (1/\sigma_{\beta_s}^2 + \sum_{i=1}^{n} \alpha_i^2 h_s(t_{ism_{is}}, \gamma_s)/\sigma_s^2)^{-1},$$

$$\tilde{\mu}_{\beta_s} = \tilde{\sigma}_{\beta_s}^2(\mu_{\beta_s}/\sigma_{\beta_s}^2 + \sum_{i=1}^{n} \alpha_i y_{ism_{is}}/\sigma_s^2), s = 1, 2.$$

Bayesian analysis

Gibbs sampling

- The full conditional posterior distribution of σ_s^2 is **inverse gamma distribution**
$$\boldsymbol{IG}\left(a_s+\sum_{i=1}^{n} m_{is},\ b_s+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m_{is}} {(z_{isj}-\alpha_i\beta_s\Lambda_{isj})^2}/{2\Lambda_{isj}}\right),\ s=1,2.$$
- The full conditional posterior density function of γ_s is proportional to
$$\prod_{i=1}^{n}\prod_{s=1}^{2}\prod_{j=1}^{m_{is}}\frac{1}{\sqrt{\Lambda_{isj}}}\exp\left\{-\frac{(z_{isj}-\alpha_{is}\Lambda_{isj})^2}{2\sigma_s^2\Lambda_{isj}}\right\}(\gamma_s)^{c_s-1}\exp\{-d_s\gamma_s\}.$$

Estimation of the missing values

- If we just observe the degradation value of $Y_1(t_k)$ at the time t_k, estimating the missing value $Y_2(t_k)$ is of our interest.
- Let $\Delta Y_s(t_k) = Y_s(t_k) - Y_s(t_{k-1})$, and $\Delta h_{sk} = h_s(t_k, \gamma_s) - h_s(t_{k-1}, \gamma_s),\ s = 1, 2.$
- We can obtain that

$$\begin{pmatrix} \Delta Y_1(t_k) \\ \Delta Y_2(t_k) \end{pmatrix} \sim \boldsymbol{N}_2(\Delta\mu_H, \Delta\Sigma), \tag{3.5}$$

where $\Delta\mu_H = \binom{\beta_1 \Delta h_{1k}}{\beta_2 \Delta h_{2k}}$,

$$\Delta\Sigma = \begin{pmatrix} \sigma_1^2 \Delta h_{1k} + \delta^2 \beta_1^2 (\Delta h_{1k})^2 & \delta^2 \beta_1 \beta_2 \Delta h_{1k} \Delta h_{2k} \\ \delta^2 \beta_1 \beta_2 \Delta h_{1k} \Delta h_{2k} & \sigma_2^2 \Delta h_{2k} + \delta^2 \beta_2^2 (\Delta h_{2k})^2 \end{pmatrix}.$$

- Given $\Delta Y_1(t_k)$, the conditional mean of $\Delta Y_2(t_k)$ is

$$\mathsf{E}(\Delta Y_2(t_k)) = \beta_2 \Delta h_{2k} + \frac{\delta^2 \beta_1 \beta_2 \Delta h_{2k}}{\sigma_1^2 + \delta^2 \beta_1^2 \Delta h_{1k}} (\Delta Y_1(t_k) - \beta_1 \Delta h_{1k}).$$

- The Bayesian estimation of $Y_2(t_k)$ can be obtained as

$$\tilde{Y}_2(t_k) = Y_2(t_{k-1}) + \int \left[\beta_2 \Delta h_{2k} + \frac{\delta^2 \beta_1 \beta_2 \Delta h_{2k}}{\sigma_1^2 + \delta^2 \beta_1^2 \Delta h_{1k}} (\Delta Y_1(t_k) - \beta_1 \Delta h_{1k}) \right] f(\Theta|z) \mathsf{d}\Theta,$$

where $f(\Theta|z)$ is the posterior PDF of Θ.

Simulation study

- The mean degradation paths of the two PCs are $1.5t$ and $0.7t^2$. Thus, $(\beta_1, \beta_2) = (1.5, 0.7)$, and $(\gamma_1, \gamma_2) = (1, 2)$.
- The diffusion parameters $(\sigma_1^2, \sigma_2^2) = (0.4, 0.3)$, and $\delta^2 = 0.04$.
- A total number of n systems are put into test, and each systems are measured m times. We choose $n = 3, 4, 5$ and $m = 6, 10$.
- 10,000 independent datasets for each experimental setting are generated to compute the point estimates, the root mean square errors (RMSE) and the empirical coverage probabilities with nominal level 95%.
- We run the Gibbs sampling 80,000 times, and discard the first 20,000 times as the burn-in period. The length of the thinning interval is taken as 20.

Table 3.1: Bayesian estimates of the parameters based on 10,000 replications.

(n, m)	Estimates	β_1	β_2	σ_1^2	σ_2^2	δ^2	γ_1	γ_2
(3,6)	Mean	1.521	0.718	0.415	0.332	0.0376	1.113	2.215
	RMSE	0.212	0.116	0.175	0.117	0.0239	0.221	0.454
(3,10)	Mean	1.524	0.714	0.413	0.321	0.0382	1.112	2.224
	RMSE	0.205	0.110	0.138	0.101	0.0204	0.201	0.398
(4,6)	Mean	1.518	0.717	0.421	0.322	0.0377	1.095	2.150
	RMSE	0.194	0.105	0.168	0.107	0.0199	0.189	0.361
(4,10)	Mean	1.519	0.711	0.417	0.318	0.0382	1.091	2.121
	RMSE	0.185	0.099	0.128	0.091	0.0178	0.157	0.326
(5,6)	Mean	1.505	0.703	0.416	0.312	0.0386	1.043	2.103
	RMSE	0.167	0.092	0.152	0.089	0.0157	0.146	0.252
(5,10)	Mean	1.506	0.708	0.419	0.308	0.0386	1.051	2.107
	RMSE	0.150	0.085	0.113	0.085	0.0141	0.102	0.228

Table 3.2: Coverage probabilities of the interval estimates with nominal level 95%.

(n, m)	β_1	β_2	σ_1^2	σ_2^2	δ^2	γ_1	γ_2
(3,6)	0.912	0.913	0.974	0.979	0.934	0.969	0.982
(3,10)	0.926	0.924	0.976	0.977	0.938	0.965	0.975
(4,6)	0.928	0.922	0.974	0.980	0.940	0.968	0.977
(4,10)	0.931	0.938	0.969	0.976	0.936	0.963	0.968
(5,6)	0.934	0.934	0.969	0.976	0.948	0.964	0.962
(5,10)	0.944	0.942	0.968	0.961	0.948	0.959	0.963

Misspecification

- There might be a mis-specification of the distribution α. Another simulation is used to check the robustness of the normal assumption.
- We assume that α follows the normal, lognormal, Weibull and Gamma distributions.
- The proposed model is used to fit data generated under these distributions.
- The estimated 10% quantile of the failure time distribution is compared with the true quantile.
- The relative biases (RB) are computed using 10,000 Monte Carlo replications.

Table 3.3: RBs of the estimated 10% quantile under different distributions of α.

(n, m)	Normal	Lognormal	Weibull	Gamma
(3,6)	0.389	0.338	0.417	0.407
(3,10)	0.218	0.0248	0.144	0.096
(4,6)	0.0811	0.168	0.306	0.217
(4,10)	0.124	0.0421	0.0584	0.030
(5,6)	0.185	0.0960	0.178	0.0578
(5,10)	0.0400	0.0191	0.0758	0.0377

Case study

- Following Peng et al. (2016), we assume $h_1(t, \gamma_1) = t$ and $h_2(t, \gamma_2) = t^{\gamma_2}$.

Table 3.4: Bayesian estimation of model parameters using heavy machine tool data.

	Our model			Peng et al. (2016)		
Parameters	Mean	SD	95% CI	Mean	SD	95% CI
β_1	0.871	0.026	(0.826, 0.927)	0.875	0.132	(0.675, 1.172)
β_2	0.142	0.040	(0.079, 0.233)	0.162	0.051	(0.086, 0.281)
σ_1^2	0.951	0.164	(0.683, 1.322)	$\times$	$\times$	$\times$
σ_2^2	0.101	0.037	(0.050, 0.193)	$\times$	$\times$	$\times$
γ_2	1.915	0.091	(1.741, 2.092)	1.867	0.091	(1.690, 2.045)
δ^2	0.0084	0.0096	(0.0019, 0.030)	$\times$	$\times$	$\times$

Table 3.5: Prediction of the missing degradation observations.

	Our model			Peng et al. (2016)		
Parameters	Mean	SD	95% CI	Mean	SD	95% CI
$y_2(t_{1,11})$	76.09	3.38	(69.95, 83.32)	72.30	3.56	(65.64, 79.75)
$y_2(t_{2,9})$	57.05	1.58	(54.12, 60.33)	56.26	4.64	(48.71, 67.02)
$y_2(t_{2,10})$	71.52	1.84	(68.10, 75.38)	69.66	6.33	(59.03, 83.98)
$y_2(t_{2,11})$	76.73	2.08	(72.55, 80.76)	74.45	6.90	(62.82, 90.03)
$y_2(t_{3,11})$	85.52	2.49	(80.94, 90.79)	84.43	5.53	(75.19, 96.88)

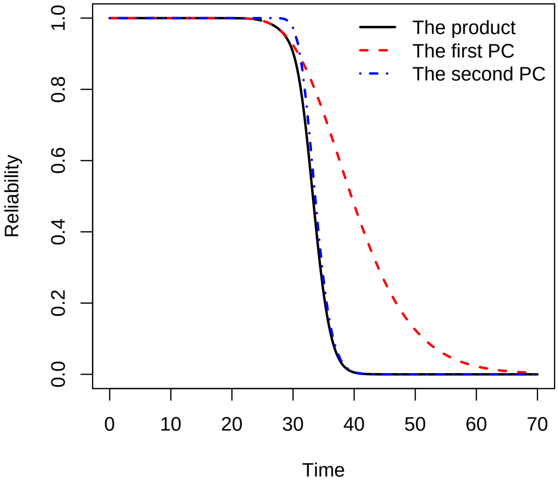


Figure 3.2: The reliability of the system and the two PCs.

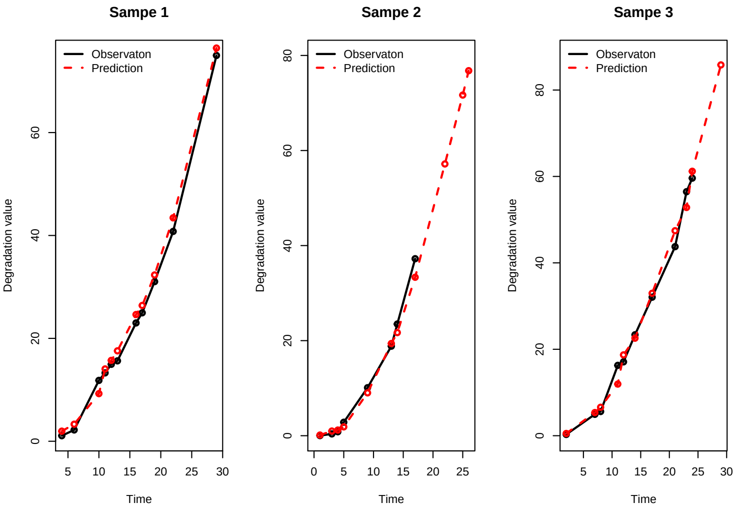


Figure 3.3: Estimation of degradation values of the second PC.

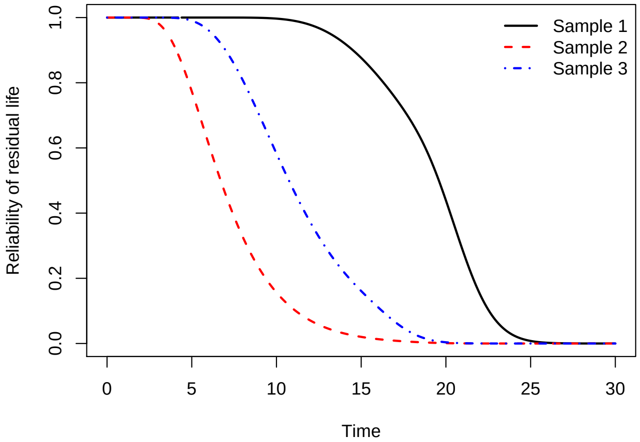
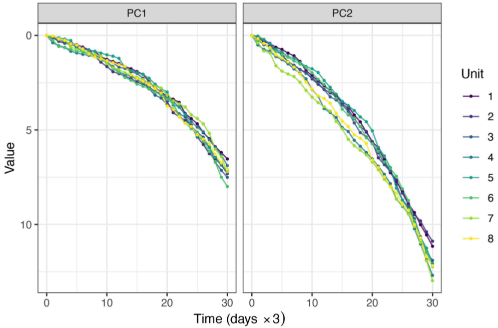


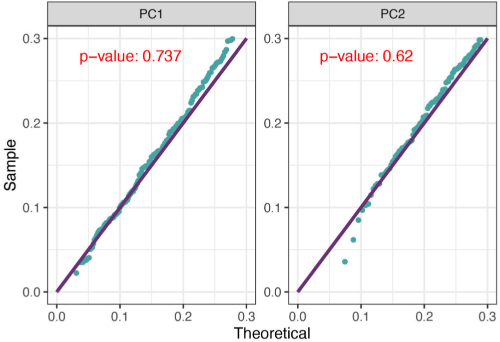
Figure 3.4: The reliability functions of the RUL for the three systems.

Outline

Motivated example: PMB degradation data



(a) Degradation paths



(b) Q-Q plots using IG distribution

Figure 3.5: Summary of Permanent magnet brake (PMB) data for two PCs: degradation paths and Q-Q plots.

PMB data with two PCs

Unit	1	2	3	4	5	6	7	8
Correlation	0.819	0.749	0.806	0.840	0.779	0.749	0.765	0.800

Figure 3.6: Pearson correlation coefficients of two PCs across various units.

- **Objective**: establish a multivariate IG process model incorporating common effects.

Outline

Conclusion

Bayesian statistical inference for

- Two-phase degradation models based on
  - Wiener model
  - Inverse Gaussian model
- Multivariate degradation models based on
  - Bivariate Wiener model
  - Multivariate inverse Gaussian model

Challenges for Bayesian statistical inference

- Fast approximation for MCMC-based Bayesian methods: ABC, INLA, VB
- ADT for stochastic process based degradation models
- More complex random effects models into the degradation model
- Deep Bayesian learning for PHM, with degradation data and sensoring data.

Thanks!