Modélisation de la détérioration basée sur les données de surveillance conditionnelle et estimation de la durée de vie résiduelle

T. T. Le , C. Bérenguer, F. Chatelain

Univ. Grenoble Alpes, GIPSA-lab, F-38000 Grenoble, France

CNRS, GIPSA-lab, F-38000 Grenoble, France

Journées de l'Automatique GdR MACS

Context

FP7 European project: SUPREME

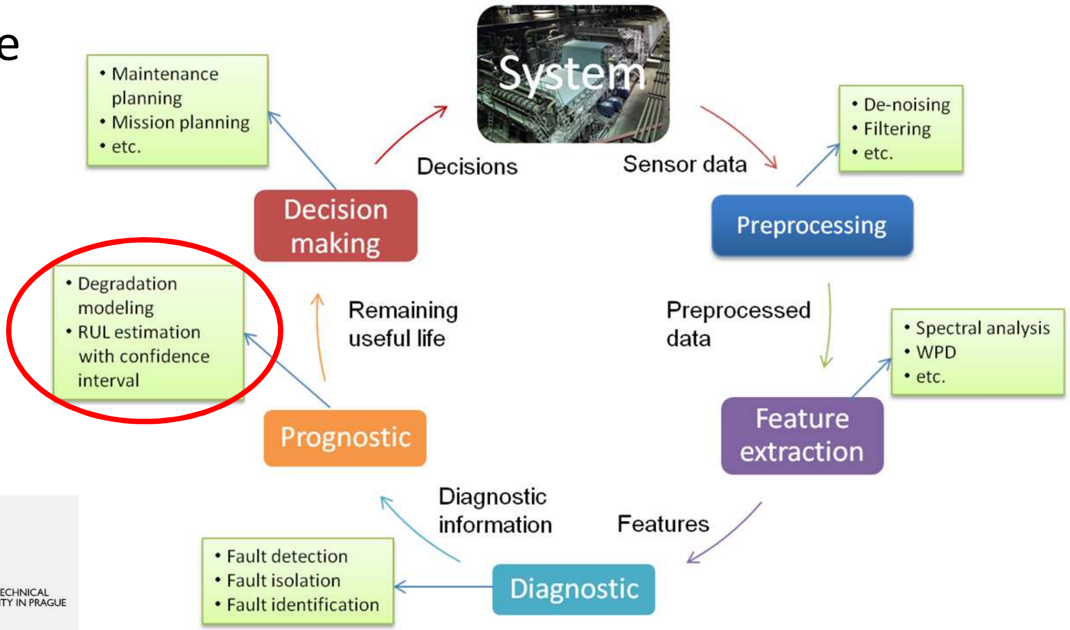
- “SUstainable PREdictive Maintenance for manufacturing Equipment”
- Purpose: Development of new tools for predictive maintenance to improve productivity, reduce machine downtimes and increase energy efficiency.

⇒ Application case: Paper machine

⇒ **Main objectives:**

✓ **Deterioration modeling**

✓ **Remaining Useful Life (RUL) estimation**



Predictive maintenance

This project has received funding from the European Community's Seventh Framework Programme (FP7-2012-NMP-ICT-FoF) under grant agreement no 314311

SUPREME partners

Problem statement

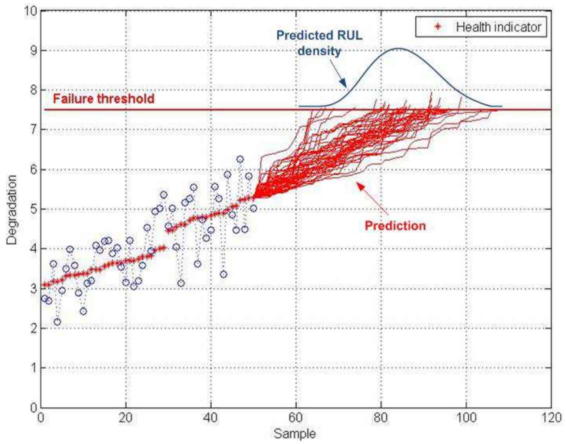
Deterioration models

- Represent temporal evolution of defects, i.e. of health indicators.
- Health indicators: unobservable
- Observations: noisy condition monitoring data

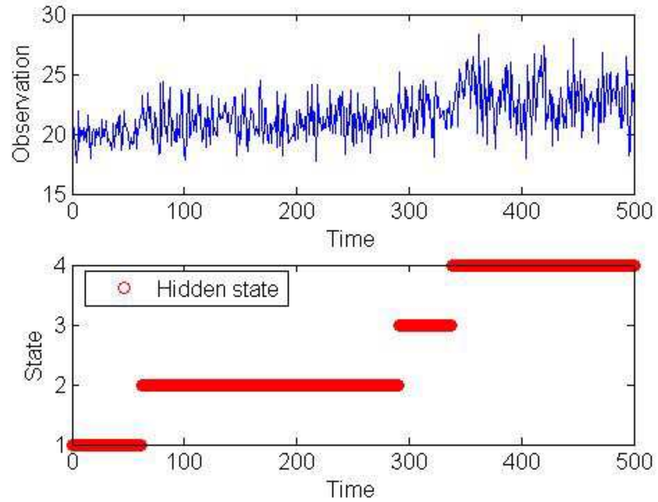
=> State-space representation

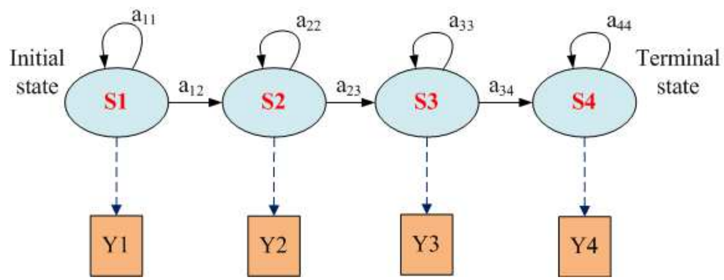
$$\begin{cases} x_t = f(x_{t-1}, \omega_t) & \text{: Hidden states} \\ y_t = g(x_t, v_t) & \text{: Observations} \end{cases}$$

- Two type of states
  - Continuous
  - Discrete



Continuous-state modeling





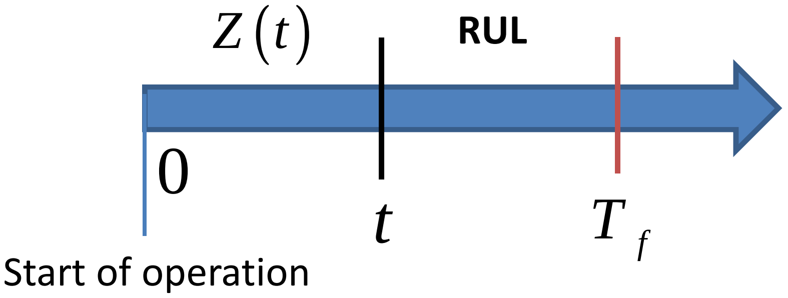
Discrete-state modeling

Problem statement

Remaining Useful Life (RUL)

- Conditional random variable:

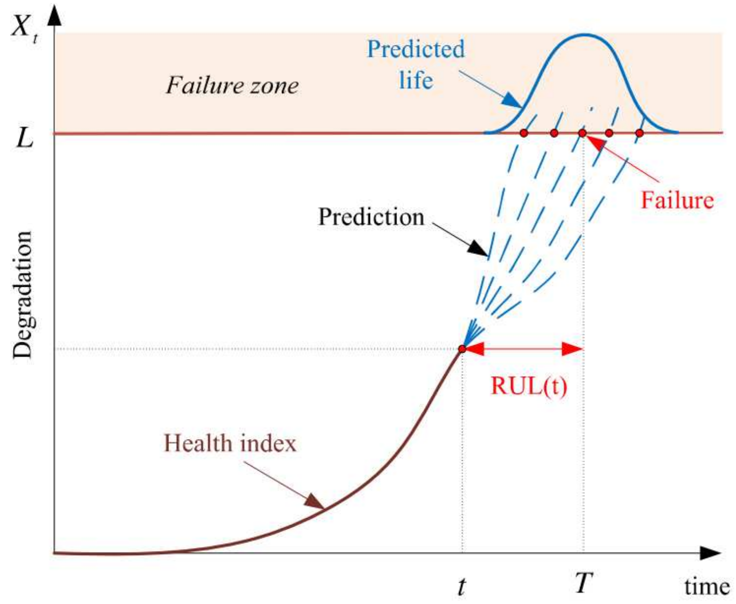
$$T_f - t \,\middle|\, T_f > t, Z(t)$$



Z(t): information up to time *t ;* T_f *:* time to failure

- Uncertainties assessment: characterize the RUL by a probabilistic distribution

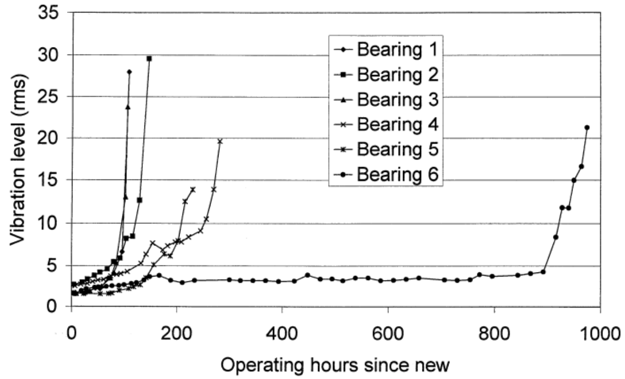
RUL prediction example



Problem statement

Problems

- Co-existence of multiple deterioration modes in competition



Proposed solution: Multi-branch modeling

- **Discrete health state:**
  – Markov based => *Multi-branch Hidden Markov model (MB-HMM)*
  – Semi-Markov based => *Multi branch Hidden semi-Markov model (MB-HsMM)*
- **Continuous health state :**
  – *Jump Markov linear systems*

Outline

- Multi-branch discrete-state model
  - Diagnostics and Prognostics framework
  - Numerical studies
- Jump Markov linear systems
  - Parameters learning
  - Health assessment and RUL estimation
  - Numerical study
- Conclusion & Perspectives

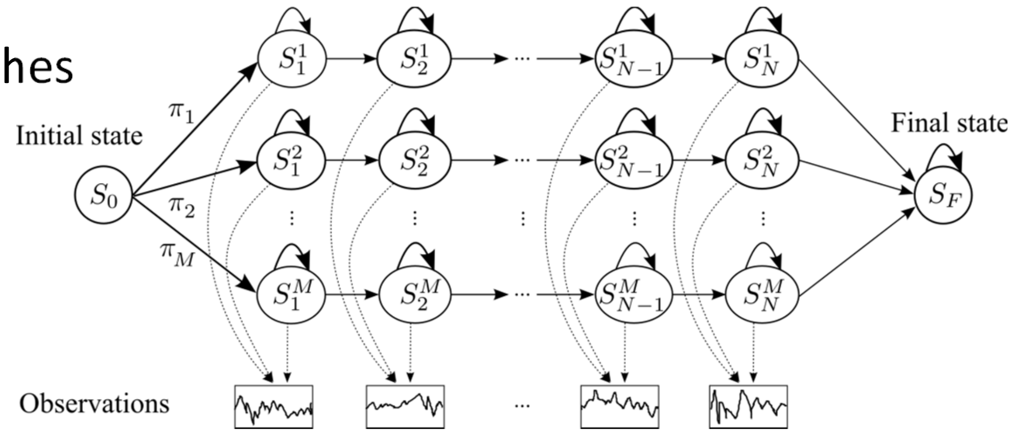
Multi-branch discrete-state models

Multi-branch discrete-state modeling

Multi-branch

- M deterioration modes ⇔ M branches
- Mode probability π_k : $\sum_{k=1}^{M} \pi_k = 1$
- Observations: continuous
- Final state: system failure
- State transition probabilities are different between the branches

=> Different deterioration rates



Assumptions

- Monotonic deterioration: Left-right topology
- Fault detection is perfect: Initial and failure states are non-emitting states
- Deterioration modes are exclusive once initiated => No branches switching

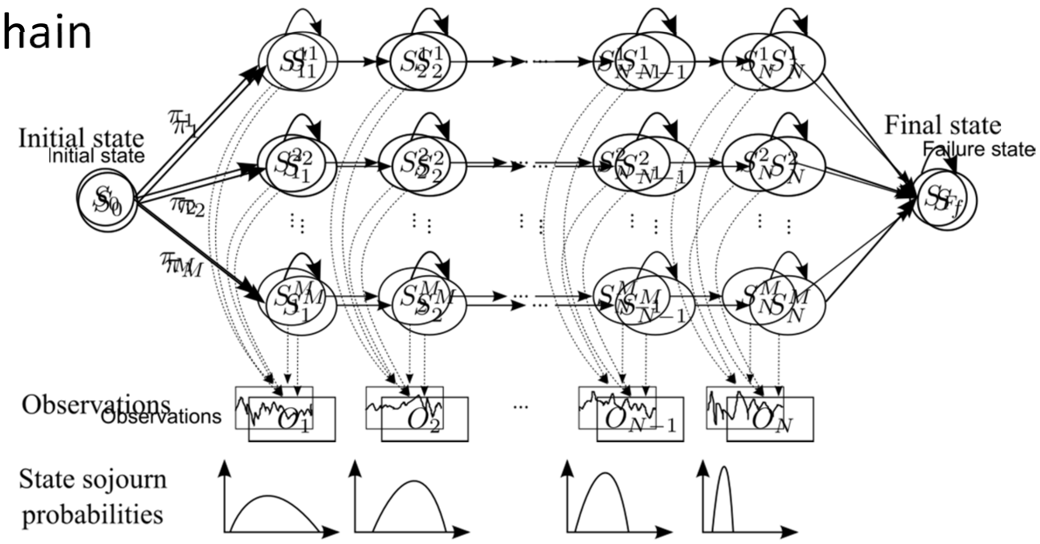
Multi-branch discrete-state modeling

Multi-branch Hidden Markov Model

- Each branch ~ left-right Markov chain
- Markovian property

$\Rightarrow$ State sojourn time: Exponential or geometrical distributed

$\Rightarrow$ May not be true in practice



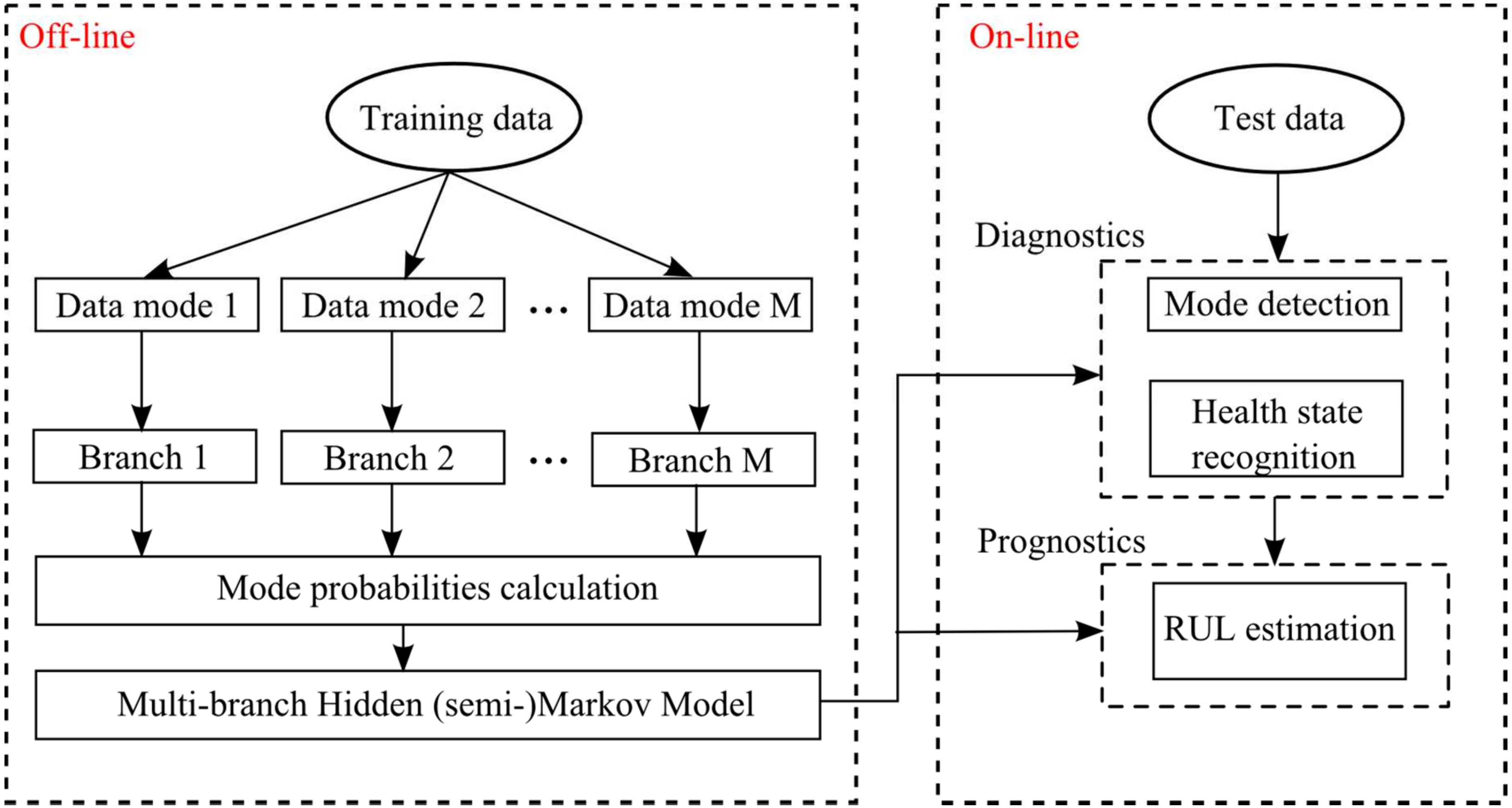
Multi-branch Hidden semi-Markov Model

- Each branch ~ left-right semi-Markov chain
- Semi-Markov property: relax the Markovian assumption

=> allow arbitrary distributions for sojourn time: Gaussian, Weibull, …

Diagnostics and prognostics framework

Two-phase implementation: offline & online



Off-line phase

Model training

- Training data: High-level features extracted from condition monitoring data
- Topology selection: BIC criterion
- Data classification => M groups
- Each group is used to train a constituent branch:
  - MB-HMM: Adaption of the Baum-Welch algorithm
  - MB-HsMM: Adaption of the Forward-Backward procedure [Yu 2006]
- *A priori* mode probabilities

$$\pi_k = P(\lambda_k) = K_k / K \ , \qquad k = 1 \ldots M$$

K_k is the number of training sequences corresponding to the mode k

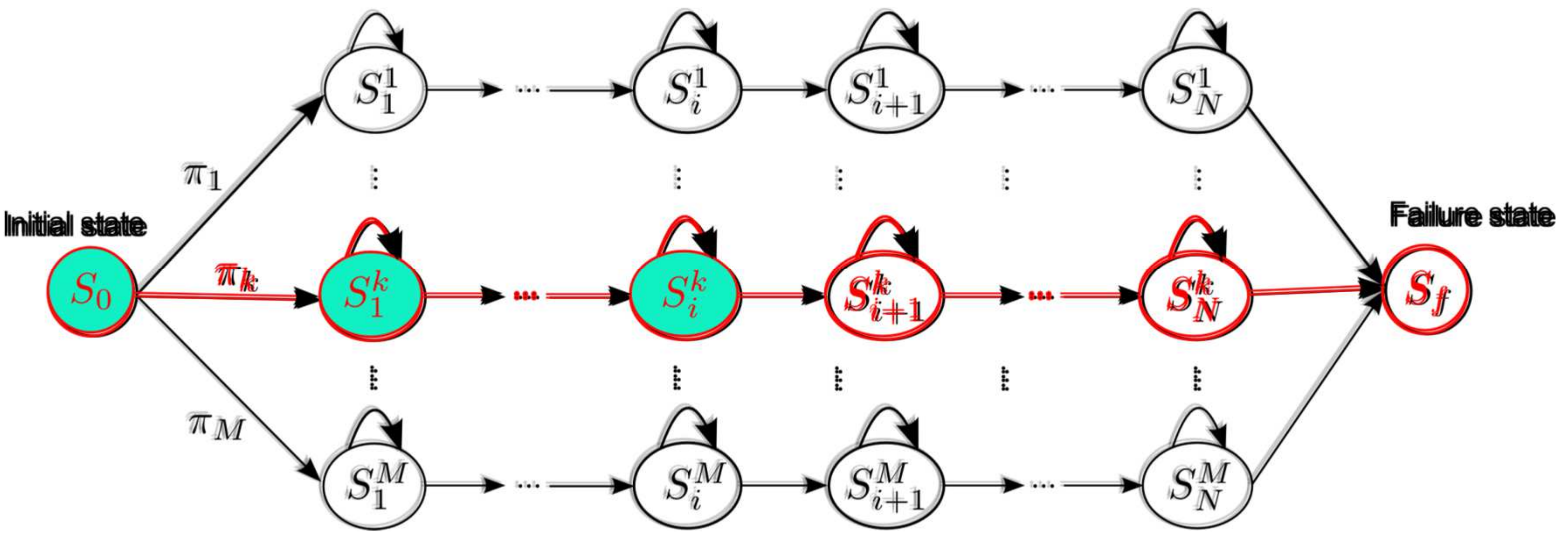
*[Yu06]: Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE Transactions on Signal Processing*, *54*(5), 1947-1951.

On-line phase

Diagnosis

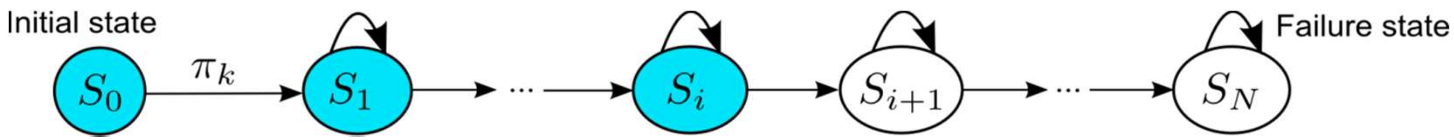
- Mode detection: $\hat{k} = \arg\max_{k} P\left(\lambda_k \mid \mathbf{O}\right)$

- Health-state assessment: Viterbi algorithm
  - Determine of the “best” state sequence: $Q^* = \arg\max_{Q_{\hat{k}}} P\left(\mathbf{O}, Q_{\hat{k}} \mid \lambda_{\hat{k}}\right)$
  - Consider the last state as the actual state



RUL estimation

One branch (HMM case)



- Suppose that the system is following the mode *k*

- RUL : discrete time assumption => number of transition steps to reach ***for the 1st time*** the failure state:

$$RUL_i^{(l)} = P\left(RUL = l \mid q_t = S_i\right) = P\left(q_{t+l} = S_N, q_{t+l-1} \neq S_N, \ldots, q_{t+1} \neq S_N \mid q_t = S_i\right)$$

- Left-right HMM: Given the current state, the system can either stay in the same state or jump to the next one

$\Rightarrow$ ***Recursive computation***

RUL estimation *(cont.)*

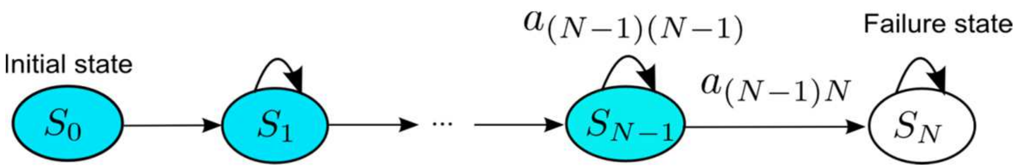
One branch (HMM case)

- At state S_{N-1} :

$$RUL_{N-1}^{(1)} = a_{(N-1)N}$$

$$\cdots$$

$$RUL_{N-1}^{(l)} = a_{(N-1)(N-1)} RUL_{N-1}^{(l-1)}$$

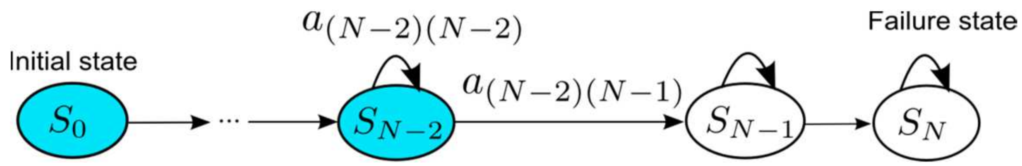


- At state S_{N-2} :

$$RUL_{N-2}^{(1)} = a_{(N-2)N} = 0$$

$$\cdots$$

$$RUL_{N-2}^{(l)} = a_{(N-2)(N-2)} RUL_{N-2}^{(l-1)} + a_{(N-2)(N-1)} RUL_{N-1}^{(l-1)}$$



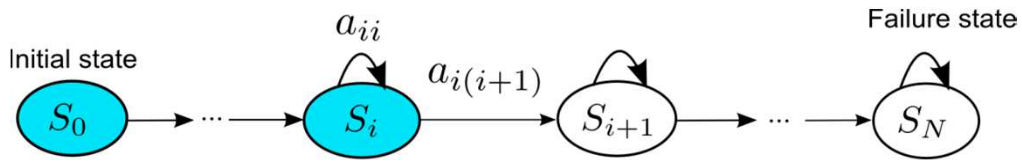
$\cdots \;\; \cdots$

- At state S_i :

$$RUL_i^{(1)} = a_{iN}$$

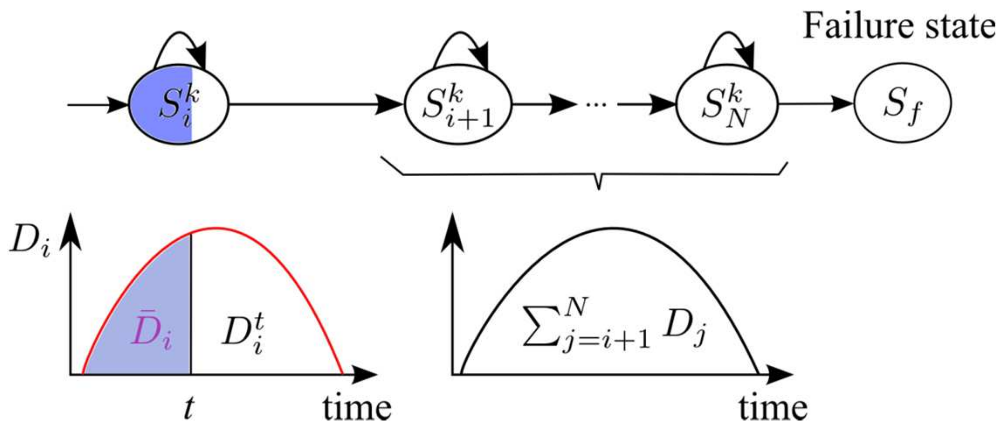
$$\cdots$$

$$RUL_i^{(l)} = a_{ii} RUL_i^{(l-1)} + a_{i(i+1)} RUL_{i+1}^{(l-1)}$$



RUL estimation *(cont.)*

One branch (HsMM case)



- Strictly left-right model:

$$\mathrm{RUL}_i^t = D_i^t + \sum_{j=i+1}^{N} D_j$$

  - D_j : Sojourn time in states *j*
  - $D_i^t = D_i - \bar{D}_i \mid D_i > \bar{D}_i$ ~ truncated Normal distribution
  - $\sum_{j=i+1}^{N} D_j$ ~ Normal distribution

- RUL = sum of Normal distribution + truncated Normal distribution

Bayesian Model Averaging

- Take into account model uncertainty: $P(\mathrm{RUL} \mid \mathbf{O}) = \sum_{k=1}^{M} P(\mathrm{RUL} \mid \lambda_k, \mathbf{O}) P(\lambda_k \mid \mathbf{O})$

Numerical examples

Simulated deterioration data

- **Fatigue Crack Growth (FCG)** model to represent the evolution of a crack depth

$$x_{t_i} = x_{t_{i-1}} + e^{w_{t_i}} C\left(\beta_b e^{\gamma_e}\sqrt{x_{t_{i-1}}}\right)^n \Delta t$$

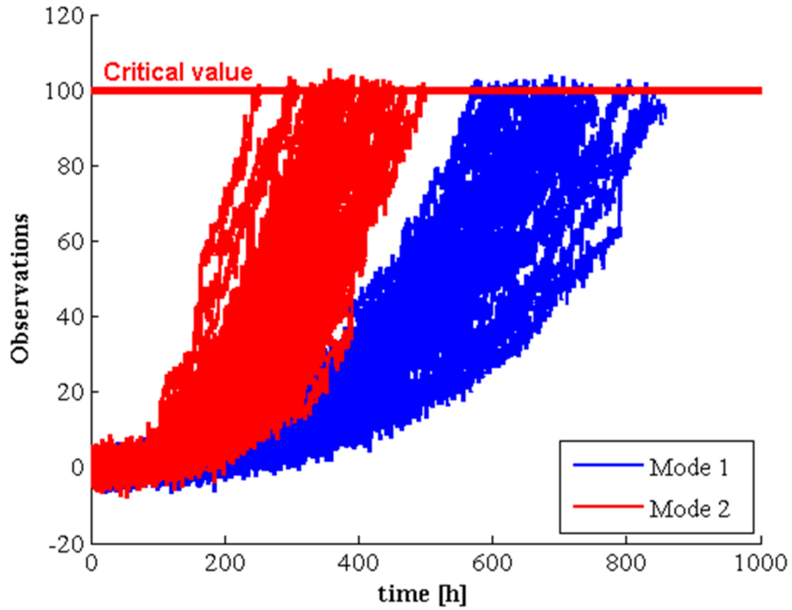
- Observation model: $y_{t_i} = x_{t_i} + \xi_{t_i}$

- Multi-mode: Two propagation rates
  - Crack depth is proportional with γ_e

$$\gamma_e = [0 \;\; 0.75]^T$$

$$C = 0.005, \quad n = 1.3, \quad \sigma_w = 1.7$$

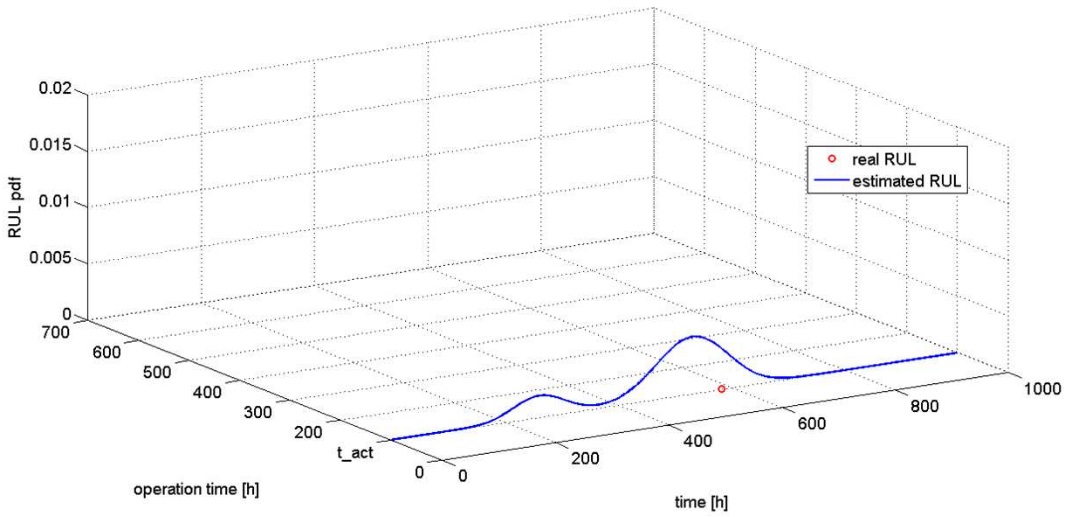
$$\sigma_\xi^2 = 10, \;\; L = 100, \;\; \pi_1 = \pi_2 = 0.5$$

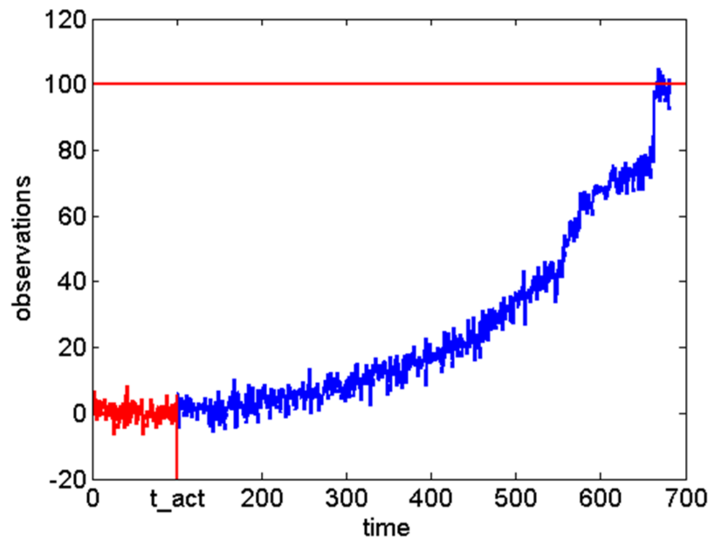


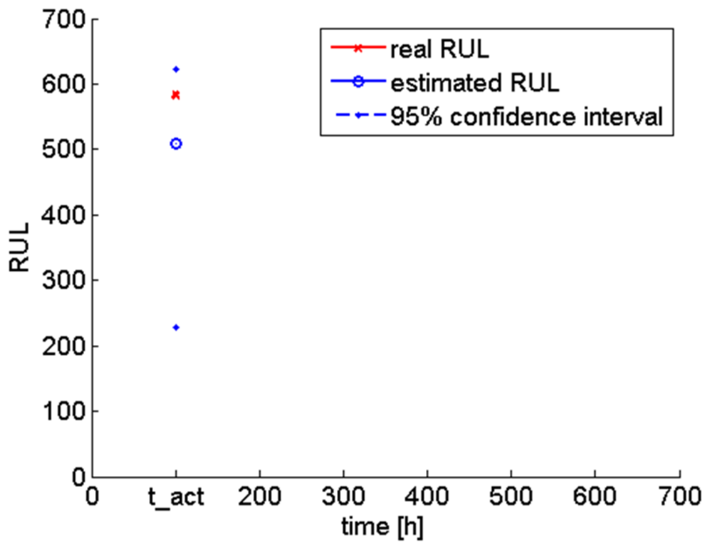
Two-mode training data

Numerical examples

Online RUL estimation (MB-HsMM model)



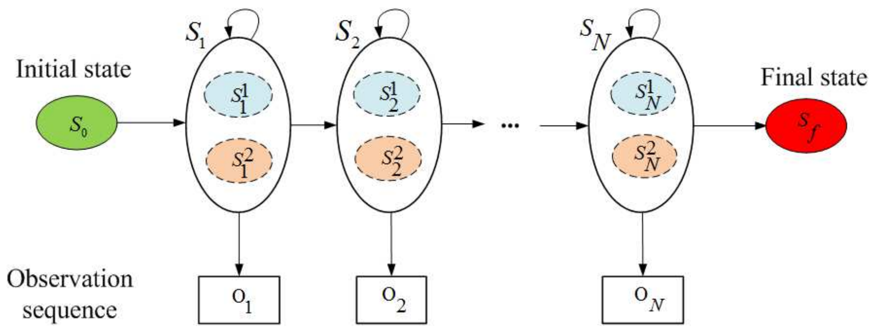




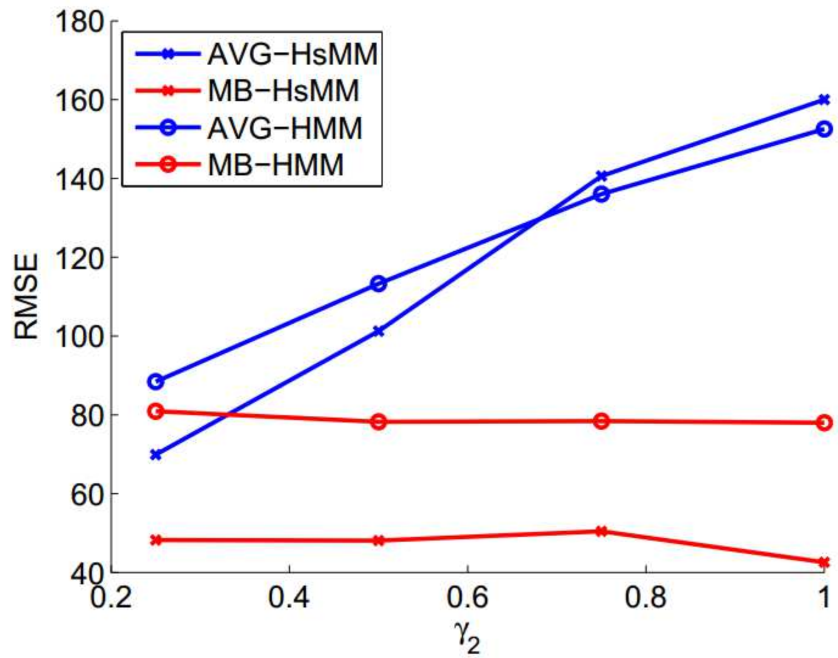
Numerical examples

Multi-branch model vs. Average model

- Aims: Investigate the advantages of the multi-branch models
- Method: Evaluate RUL estimation results at different mode "distances"
  - Mode "distance": proportional to γ_e
  - Criterion: Root mean squared error (RMSE)
- Result:



Average model



Case study (MB-HsMM model):

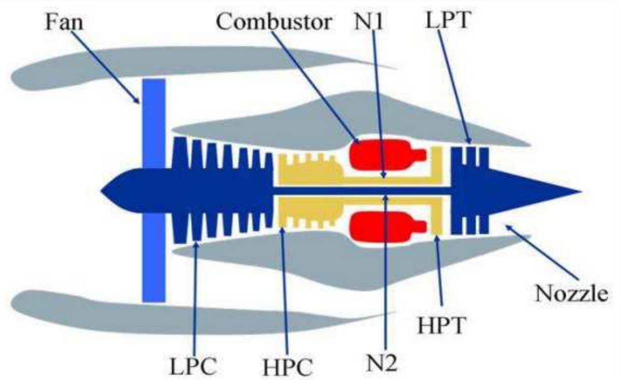
PHM08 competition

- C-MAPSS: Modeling a large realistic commercial turbofan engine
- 2 data set for training and test
- 218 identical and independent units
- Objective:
  - Construct a prognostic method basing on training data set
  - Use it to estimate the RUL of each unit in test data set
- Evaluation criterion:

$$S = \sum_{i=1}^{218} S_i$$

Where S_i is penalty score for unit *i*:

$$S_i = \begin{cases} e^{-d_i/13} - 1, & d_i \le 0 \\ e^{d_i/10} - 1, & d_i > 0 \end{cases}$$ where $d_i = RUL_{est}^{i} - RUL_{real}^{i}$



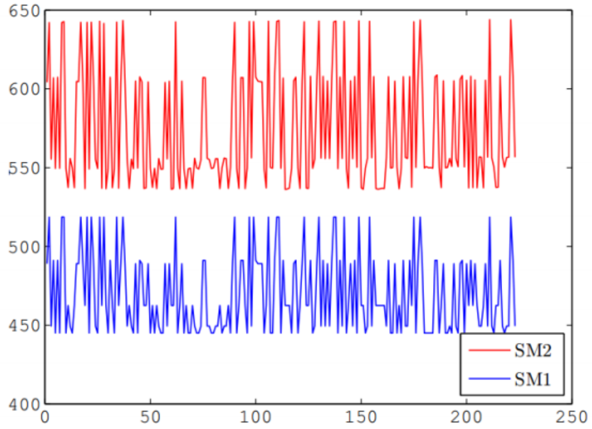
Simplified diagram of engine simulated in C-MAPSS

C-MAPSS: Commercial Modular Aero-Propulsion System Simulation

PHM2008 data

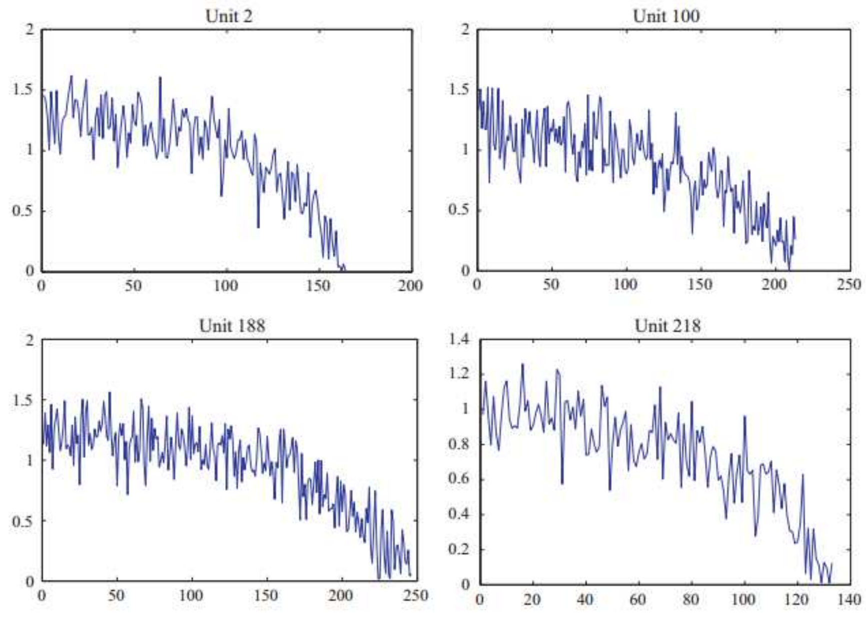
Health indicator construction

- From [Le Son *et al.*]*



Raw data

PCA-based



Constructed health indicators

- Clear tendency of indicator temporal evolution
- Better score than the winners of the competition:
  - S = 5520 with the Wiener process based method
  - S = 4170 with non-homogeneous Gamma based method

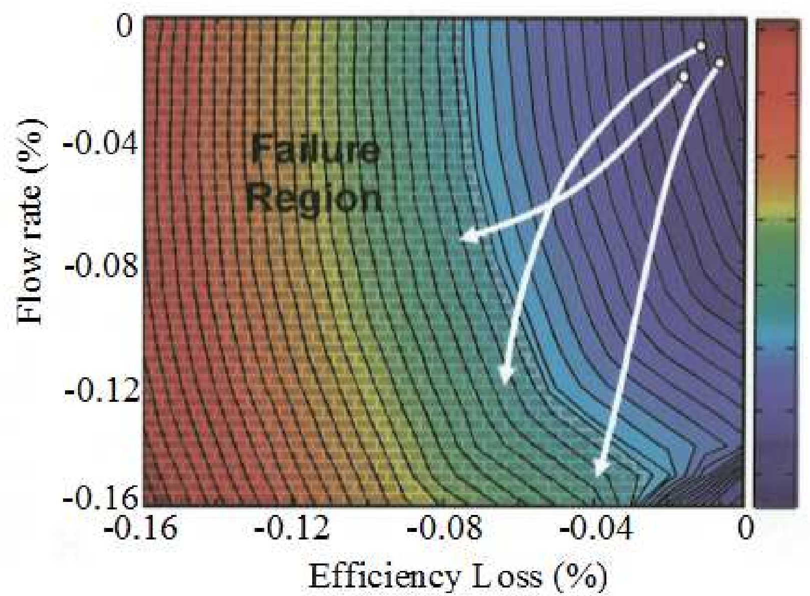
* [Le Son *et al.*] Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety 2012*

Application of the MB-HsMM model

Number of deterioration modes

- Different fault propagation trajectories depending on the decrease rates of the flow rate (f) and efficiency (e) parameters
- Consider 3 modes of deterioration:

	2 modes	3 modes	4 modes
Mode 1	f < e	f < e	f << e
Mode 2	f > e	f ≈ e	f < e
Mode 3		f > e	f > e
Mode 4			f >> e
Nb of branches	2	3	4



Fault propagation trajectories

Application of the MB-HsMM model

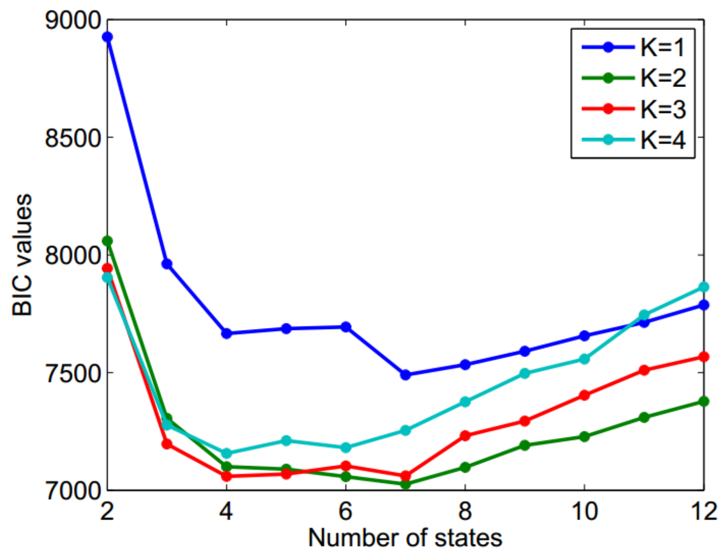
Observations model

- Mixture of Gaussian: $b_j(\mathbf{x}) = \sum_{k=1}^{K} c_{jk} \mathrm{N}\left(\mathbf{x}; \mu_{jk}, \Sigma_{jk}\right)$

K: number of mixture components

Topology selection

- BIC criterion: N = 7; K = 2



RUL estimation result

$$RSE = \sqrt{\sum_{i=1}^{218} d_i^2}$$

$$MSE = \sum_{i=1}^{218} \frac{d_i^2}{218}$$

$$d_i = RUL_{est}^i - RUL_{real}^i$$

Method	Score	RSE	MSE
1-branch HsMM	12246	502	1157
2-branch HsMM	6456	451	936
3-branch HsMM	**5458**	**410**	**773**
4-branch HsMM	**3791**	**389**	**694**
Wiener-based method	5575	423	823
Gamma-based method	4107	434	864

Multi-branch continuous-state models

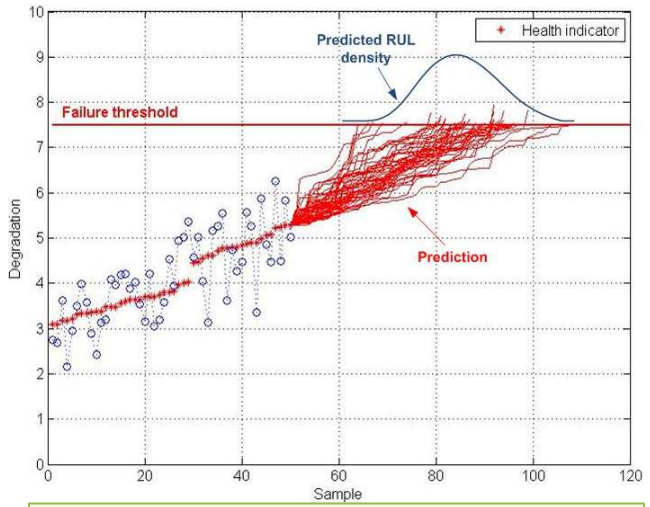
Motivations

Health states are continuous

- State-space representation
- Deterioration modes in competition

=> mode switching:

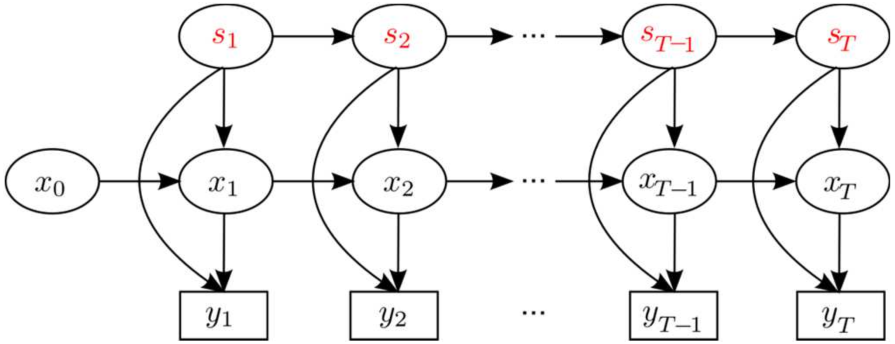
$$\begin{cases} x_t = f(x_{t-1}, \omega_t, s_t) & \text{: Hidden states} \\ y_t = g(x_t, v_t, s_t) & \text{: Observations} \end{cases}$$



Continuous-state modeling

- s_t : realization at time t of discrete variable S

  S ~ discrete-time Markov chain
- Graphical representation



Switching state-space model

Jump Markov Linear System

Assumption

- Deterioration dynamic can be approximated by linear model
- M deterioration modes: $s_t \in \{1, 2, \ldots, M\}$
- Model formulation

$$\begin{cases} x_t = A_{s_t} x_{t-1} + \omega_t \\ y_t = C_{s_t} x_t + v_t \end{cases} \quad \text{where} \quad \begin{cases} \omega_t \sim \mathcal{N}\left(0, Q_{s_t}\right) \\ v_t \sim \mathcal{N}\left(0, R_{s_t}\right) \\ x_0 \sim \mathcal{N}\left(\mu_0, \Sigma_0\right) \end{cases}$$

- Transition probability matrix

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} & \ldots & \pi_{1M} \\ \pi_{21} & \pi_{22} & \ldots & \pi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{M1} & \pi_{M2} & \ldots & \pi_{MM} \end{pmatrix} \quad \text{where} \quad \pi_{ij} = \mathrm{P}\left(s_{t+1} = i \middle| s_t = j\right)$$

- Initial state distribution: $\pi_1(i) = P(s_1 = i)$

Parameters learning problem

Model parameters

$$\Theta = \{A_i, C_i, Q_i, R_i, \mu_0, \Sigma_0, \Pi, \pi_1\}_{i=1,\ldots,M}$$

- Incomplete data => Expectation-Maximization algorithm:
  - E step: $$Q\left(\Theta \mid \Theta^{(k)}\right) = \mathbf{E}\left[\log P\left(X_T, S_T, Y_T \mid \Theta\right) \mid Y_T, \Theta^{(k)}\right]$$
  - M step: $$\Theta^{(k+1)} = \arg\max_{\Theta} Q\left(\Theta \mid \Theta^{(k)}\right)$$
- Problem: Presence of switching dynamic

$$Q\left(\Theta \mid \Theta^{(k)}\right) = \sum_{S_T}\left(P\left(S_T \mid Y_T, \Theta^{(k)}\right) \int p\left(X_T \mid S_T, Y_T, \Theta^{(k)}\right) \log P\left(X_T, S_T, Y_T \mid \Theta\right) dX_T\right)$$

- Computed over all possible sequences of discrete states S_T => Intractable

=> Approximated EM algorithm

Approximated EM algorithm

Pruning technique

- Approximation: Calculate the sum over the most “likely” state sequence
- Adaption of the Viterbi algorithm
- Do not guarantee the convergence, but still sufficient in several practical cases
- Most important: the algorithm is linear in number of time steps.

From the traditional Viterbi algorithm…

- Define the best “partial cost” at time t:

$$J_t(i) = \max_{\mathrm{S}_{t-1},\mathrm{X}_t} \log \mathrm{P}\left(\mathrm{X}_t, \mathrm{Y}_t, \mathrm{S}_{t-1}, s_t = i\right)$$

- Then, for each state transition j -> i, assign an “innovation cost”:

$$J_{t,t+1}^{j,i} = -\frac{1}{2}\left(y_t - C_i x_{t+1|t,j,i}\right)'\left(C_i \Sigma_{t+1|t,j,i} C_i' + R_i\right)^{-1}\left(y_t - C_i x_{t+1|t,j,i}\right)$$
$$-\frac{1}{2}\log\left|C_i \Sigma_{t+1|t,j,i} C_i' + R_i\right| + \log \Pi(j,i)$$

Recursion

- At time t+1: $J_{t+1}(i) = \max_{j} \left(J_{t,t+1}^{j,i} + J_t(j) \right)$

$$\psi_t(i) = \arg\max_{j} \left(J_{t,t+1}^{j,i} + J_t(j) \right)$$

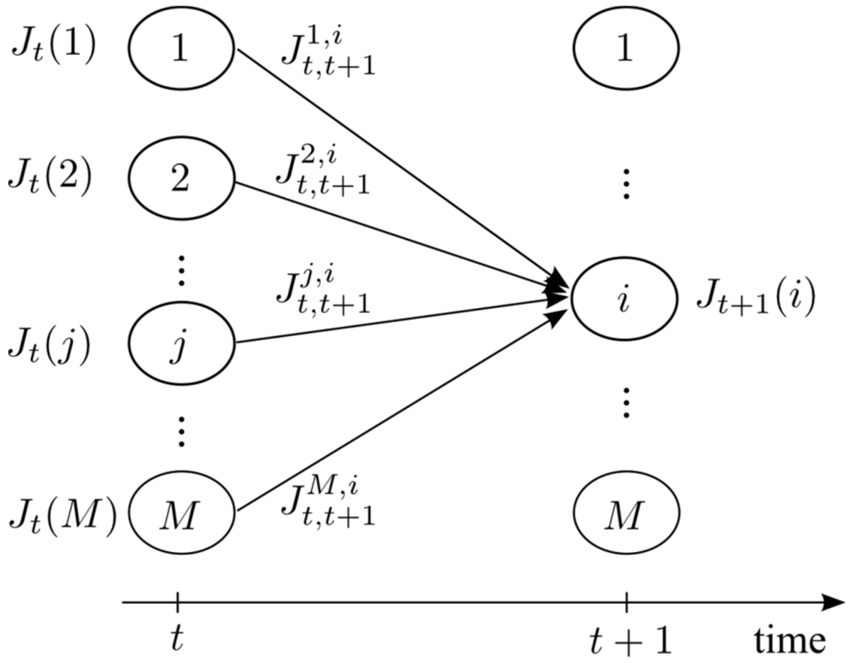


Illustration of partial cost calculation

Termination

- At time T: $J_T^* = \max_{i} J_{T,i}$

$$s_T^* = \arg\max_{i} J_{T,i}$$

Backtracking

For $t = T-1, T-2, \ldots, 1$: $s_t^* = \psi_{t+1}\left(s_{t+1}^* \right)$

Approximated Q function

Denote S_T^* the most likely state sequence:

$$Q\left(\Theta \mid \Theta^{(k)}\right) \approx \int p\left(X_T \mid S_T^*, Y_T, \Theta^{(k)}\right) \log P\left(X_T, S_T^*, Y_T \mid \Theta\right) dX_T$$

=> Calculated by Rauch-Tung-Streiber (RTS) smoother

M-step: Parameter re-estimation

Continuous part

$$\begin{cases} \hat{\mathrm{A}}_i = \mathrm{S}_{10,i}\ \mathrm{S}_{00,i}^{-1} \\ \hat{C}_i = \mathrm{S}_{20,i}\ \mathrm{S}_{11,i}^{-1} \\ \hat{Q}_i = \dfrac{1}{\sum_{k=1}^{K} T_i^{(k)}} \mathrm{S}_{11,i} - \hat{\mathrm{A}}_i \mathrm{S}_{10,i}^{'} \\ \hat{R}_i = \dfrac{1}{\sum_{k=1}^{K} T_i^{(k)}} \mathrm{S}_{22,i} - \hat{C}_i \mathrm{S}_{20,i}^{'} \\ \hat{\mu}_0 = \dfrac{1}{K} \sum_{k=1}^{K} \hat{\mu}_0^{(k)} \\ \hat{\Sigma}_0 = \dfrac{1}{K} \sum_{k=1}^{K} \left[\hat{P}_0^{(k)} + \hat{\mu}_0^{(k)} \left(\hat{\mu}_0^{(k)'} \right) \right] - \hat{\mu}_0 \left(\hat{\mu}_0 \right)^{'} \end{cases}$$

where

$$\begin{cases} \mathrm{S}_{11,i} = \sum_{k=1}^{K} \sum_{t \in \mathrm{F}_i(S_T^*)} \left(\hat{x}_{t|T}^{(k)} \left(\hat{x}_{t|T}^{(k)} \right)' + \hat{\Sigma}_{t|T}^{(k)} \right) \\ \mathrm{S}_{10,i} = \sum_{k=1}^{K} \sum_{t \in \mathrm{F}_i(S_T^*)} \left(\hat{x}_{t|T}^{(k)} \left(\hat{x}_{t-1|T}^{(k)} \right)' + \hat{\Sigma}_{t|T} \right) \\ \mathrm{S}_{00,i} = \sum_{k=1}^{K} \sum_{t \in \mathrm{F}_i(S_T^*)} \left(\hat{x}_{t-1|T}^{(k)} \left(\hat{x}_{t-1|T}^{(k)} \right)' + \hat{\Sigma}_{t-1|T} \right) \\ \mathrm{S}_{22,i} = \sum_{k=1}^{K} \sum_{t \in \mathrm{F}_i(S_T^*)} y_t^{(k)} \left(y_t^{(k)} \right)' \\ \mathrm{S}_{20,i} = \sum_{k=1}^{K} \sum_{t \in \mathrm{F}_i(S_T^*)} y_t^{(k)} \left(\hat{x}_{t|T}^{(k)} \right)' \end{cases}$$

M-step: Parameter re-estimation

Discrete part

- Similar to the Hidden Markov model:

$$\hat{\pi}_1(i) = \text{number of times in state } i \text{ at time } t = 1$$

$$\pi_{i,j} = \frac{\text{number of transitions from state } i \text{ to state } j}{\text{number of transitions from state } i}$$

- We obtain:

$$\hat{\pi}_1 = \frac{1}{K} \sum_{k=1}^{K} e_{S_T^{*(k)}(t=1)}$$

$$\hat{\Pi} = \frac{1}{K} \sum_{k=1}^{K} \sum_{t=1}^{T^{(k)}} \left(\xi_t^{(k)} \left(\xi_t^{(k)} \right)' \right) \operatorname{diag} \left(\sum_{t=1}^{T^{(k)}} \xi_t^{(k)} \right)^{-1}$$

$$e_i = \left[0 \;\mathrm{K}\; 0 \; \underset{\uparrow}{1} \; 0 \;\mathrm{K}\; 0 \right]$$

*i*th element

where $\xi_t^{(k)} = e_{S_T^{*(k)}(t)}$

JMLS based diagnostics

Mode probabilities

- Given test data $Y_t = \{y_1, y_2, \ldots, y_t\}$
- Denote $S^*_{t,i}$ the best state sequence at time t that ends in *i*

$$\mu_t(i) = P(s_t = i) = \frac{P(S^*_{t,i})}{\sum_{i=1}^{M} P(S^*_{t,i})} = \frac{1}{1 + \exp\left(\sum_{j \neq i} J_{t,j} - J_{t,i}\right)}$$

Health state assessment

$$\begin{cases} \hat{x}_t = \sum_{i=1}^{M} \mu_t(i) \hat{x}_{t,i} \\ \hat{\Sigma}_t = \sum_{i=1}^{M} \mu_i(t) \Sigma_{t,i} + \sum_{i=1}^{M} \mu_t(i) \left(\hat{x}_{t,i} - \hat{x}_t\right)\left(\hat{x}_{t,i} - \hat{x}_t\right)^T \end{cases}$$

where $\hat{x}_{t,i}$ and $\Sigma_{t,i}$ are the mean value and variance given $S^*_{t,i}$

RUL estimation

- Discrete time: $RUL = \left(\min k \geq 1 : x_{t+k} \geq L \mid x_t < L\right)$
- If future mode changes are deterministic => x_{t+k} can be estimated by multi-step ahead Kalman prediction
- In JMLS case: M fold increase in number of Gaussian distributions to consider

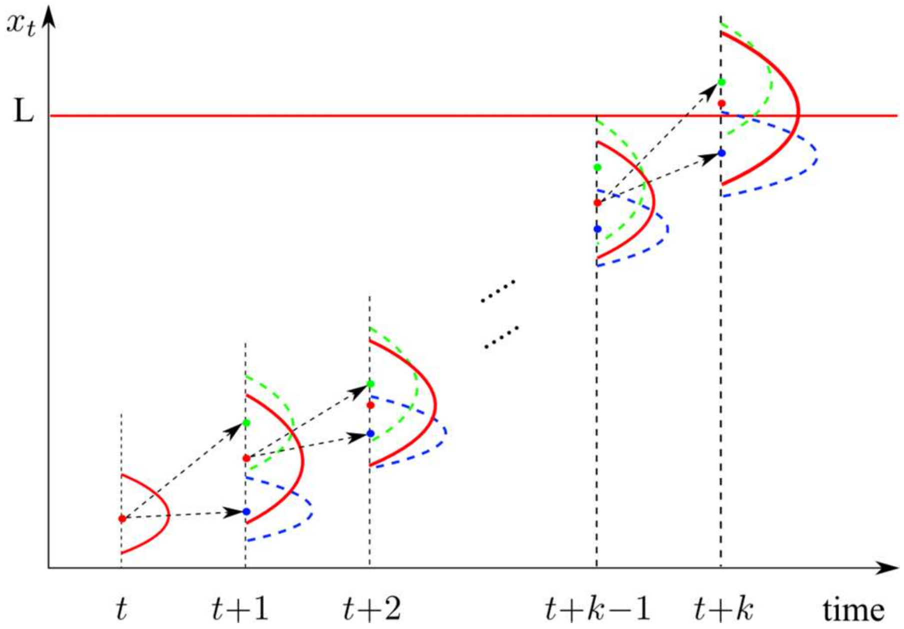
$\Rightarrow$ Intractable computation

$\Rightarrow$ Approximation: merge all one-step predicted Gaussian distributions into one Gaussian distribution

$$p\left(x_{t+1|t}\right) \approx \sum_{i=1}^{M} \mu_{t+1}(i) \mathrm{N}\left(x_{t+1|t,i}, \Sigma_{t+1|t,i}\right)$$

- Mode probability update

$$\mu_{t+1}(i) = \sum_{j=1}^{M} \pi_{j,i} \cdot \mu_t(j)$$



Numerical results

Mode training

- Two deterioration modes: quick-rate (mode 1) and normal-rate (mode 2)
- Real parameters

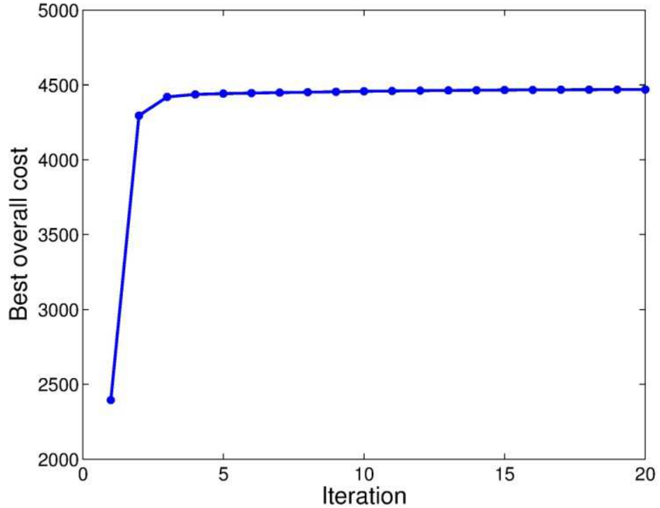
$$A_1 = 1.01 \quad C_1 = 1 \quad Q_1 = 0.015 \quad R_1 = 2.5$$
$$A_2 = 1.002 \quad C_2 = 1 \quad Q_2 = 0.005 \quad R_2 = 0.25$$
$$\mu_0 = 2 \quad \Sigma_0 = 0.2$$

$$\Pi = \begin{bmatrix} 0.99 & 0.01 \\ 0.02 & 0.98 \end{bmatrix} \quad \pi_1 = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$$

- Estimated parameters

$$\hat{A}_1 = 1.01 \quad \hat{C}_1 = 1.01 \quad \hat{Q}_1 = 0.02 \quad \hat{R}_1 = 2.59$$
$$\hat{A}_2 = 1.002 \quad \hat{C}_2 = 1 \quad \hat{Q}_2 = 0.01 \quad \hat{R}_2 = 0.25$$
$$\hat{\mu}_0 = 2.19 \quad \hat{\Sigma}_0 = 0.285$$

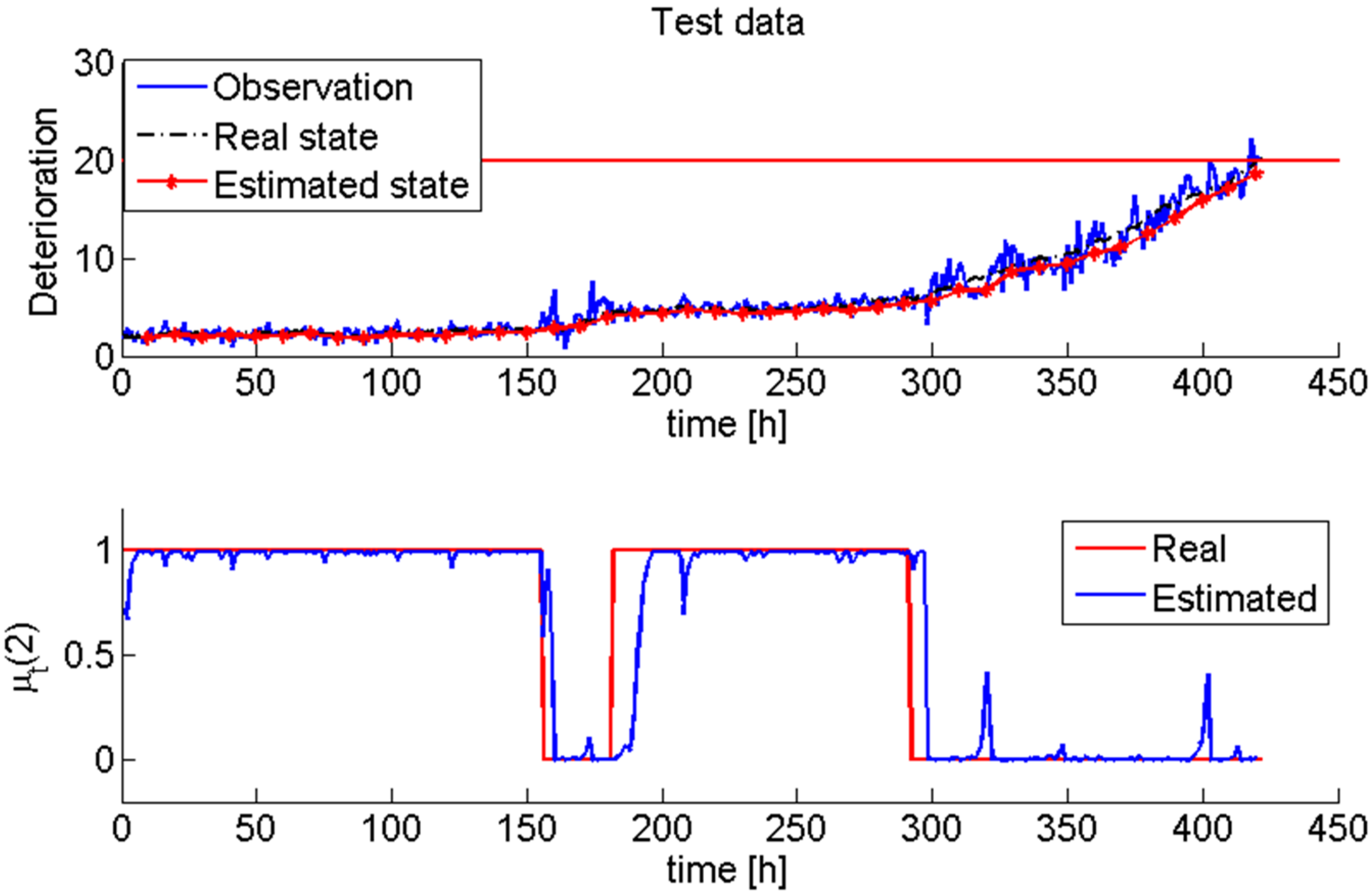
$$\hat{\Pi} = \begin{bmatrix} 0.994 & 0.006 \\ 0.011 & 0.989 \end{bmatrix} \quad \hat{\pi}_1 = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$$



Convergence of the EM algorithm

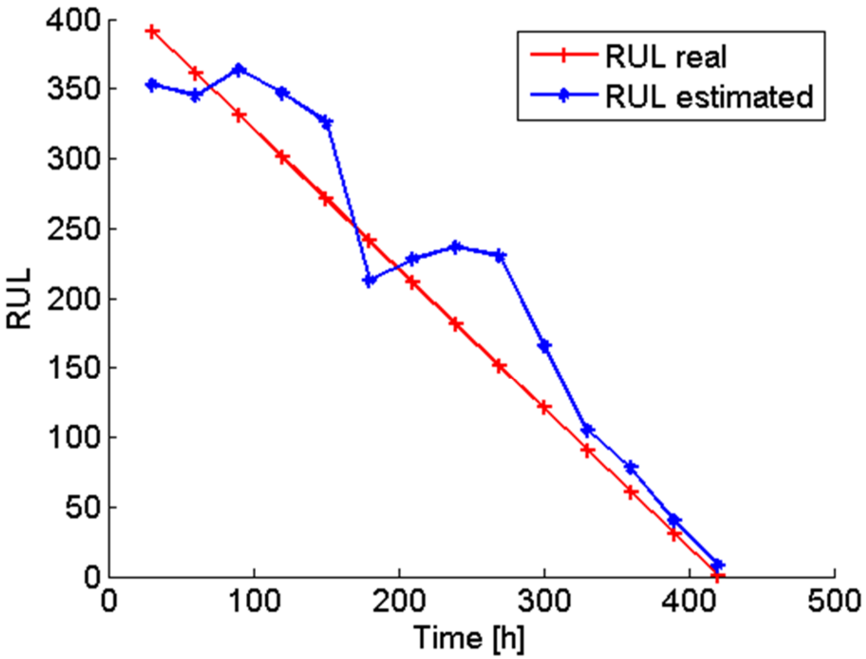
Diagnosis result

Mode detection and health assessment

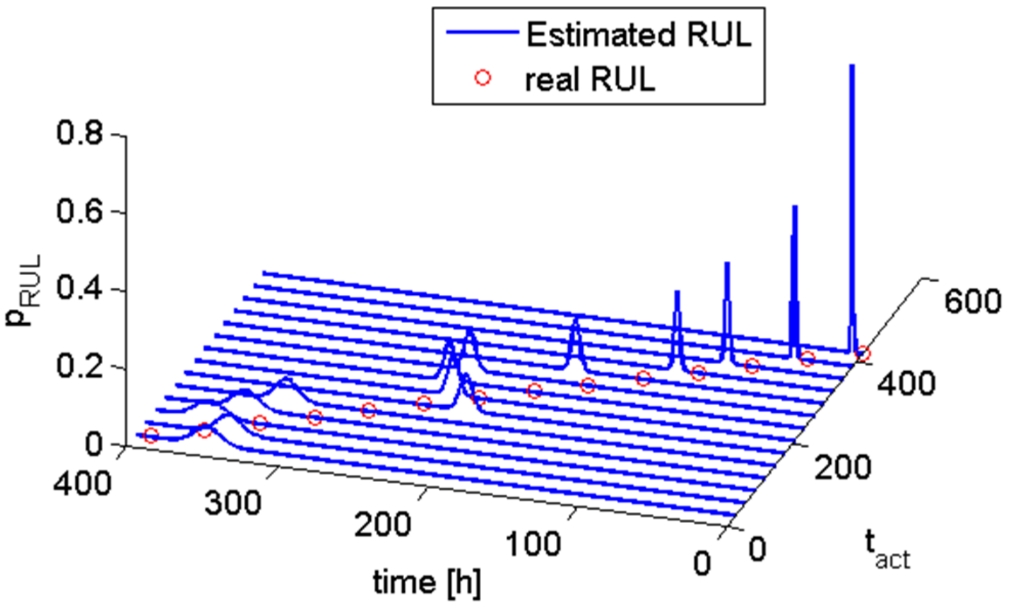


RUL estimation result

Time to failure: $T_f = 420[h]$



Mean estimated RUL at different times



Evolution of estimated RUL pdf

Conclusion & Perspectives

Conclusion

- Development of multi-branch models to deal with the co-existence problem of multiple deterioration modes
  - Discrete health states: MB-HMM & MB-HsMM
  - Continuous health states: JMLS
- Proposition of multi-branch models based framework for diagnostics & prognostics
  - Detection of actual deterioration mode
  - Assessment of current health status
  - Estimation of the RUL

Perspectives

- Extension of the JMLS model to the non-linear case
- Application with real-life data

Thank you for your attention!

LE Thanh Trung

PhD student

GIPSA-lab, Grenoble Insititut of Technology

11 rue des Mathématiques – 38402 Saint Martin d'Hères cedex

Phone: +33 4 76 82 71 59 - Fax: +33 4 76 82 63 88

Email: thanh-trung.le@gipsa-lab.grenoble-inp.fr