DATA CHALLENGE PHME 2026 PROPOSAL

IUT Besançon-Vesoul | Marie & Louis pasteur University (www.iut-bv.univ-fcomte.fr)

Institut FEMTO-ST/AS2M/DATA-PHM (https://www.femto-st.fr)

E-mail: moncef.soualhi@umlp.fr

Outlines

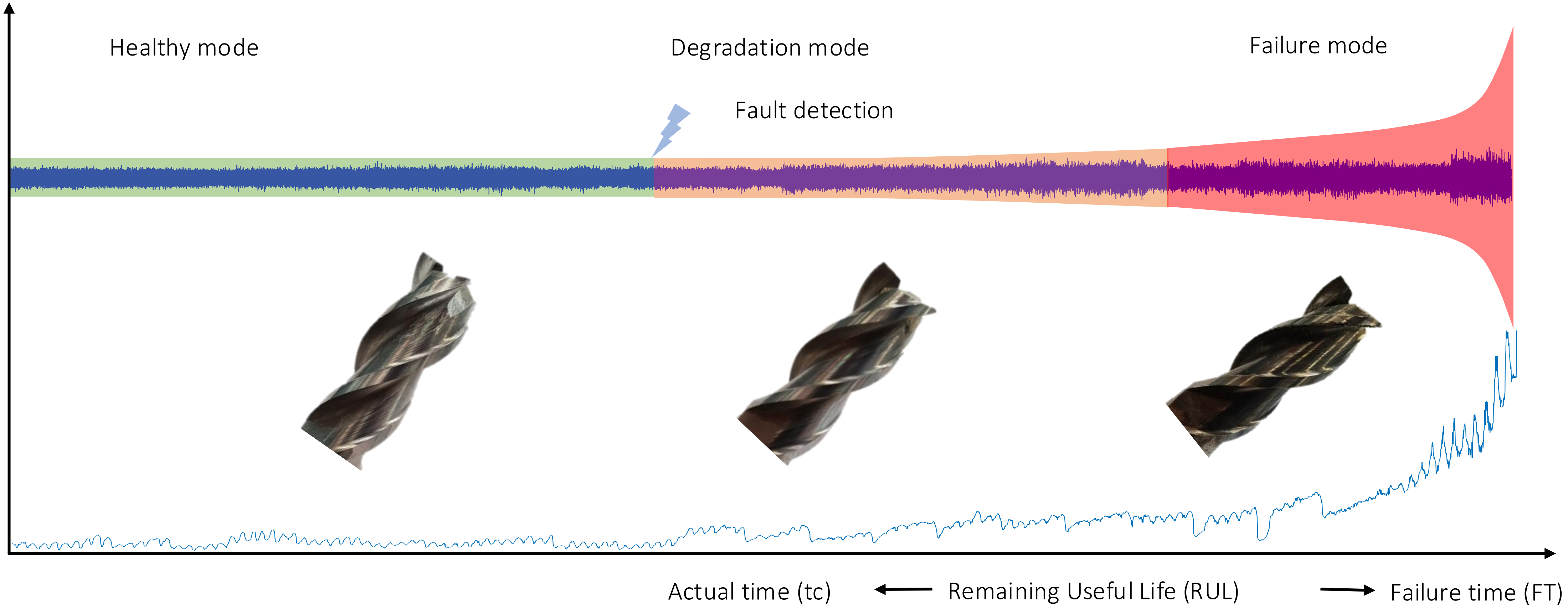
- Introduction to failure prognostics and challenges
- Case study description and failure scenario
- Collected data and visualization
- Evaluation metrics and scoring function

Remaining Useful Life techniques

Remaining useful life estimation for failure prognostics

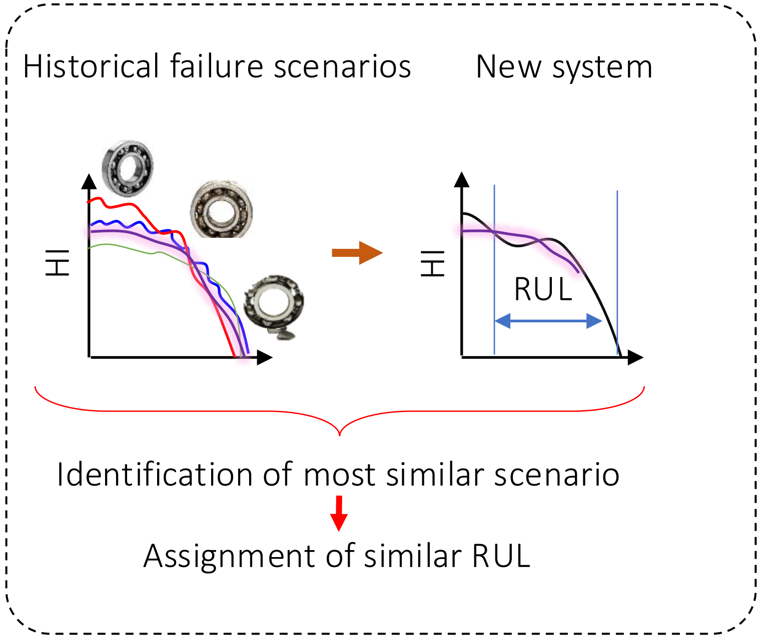
Failure prognostic is the analysis of the symptoms of faults to predict future condition and residual life within design parameters

[ISO 13372, 2012 1.5]

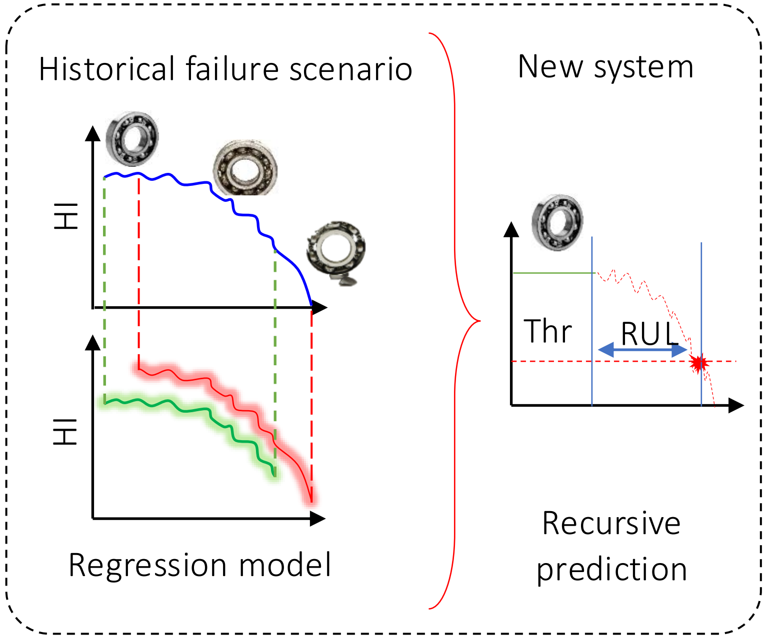


Remaining Useful Life techniques

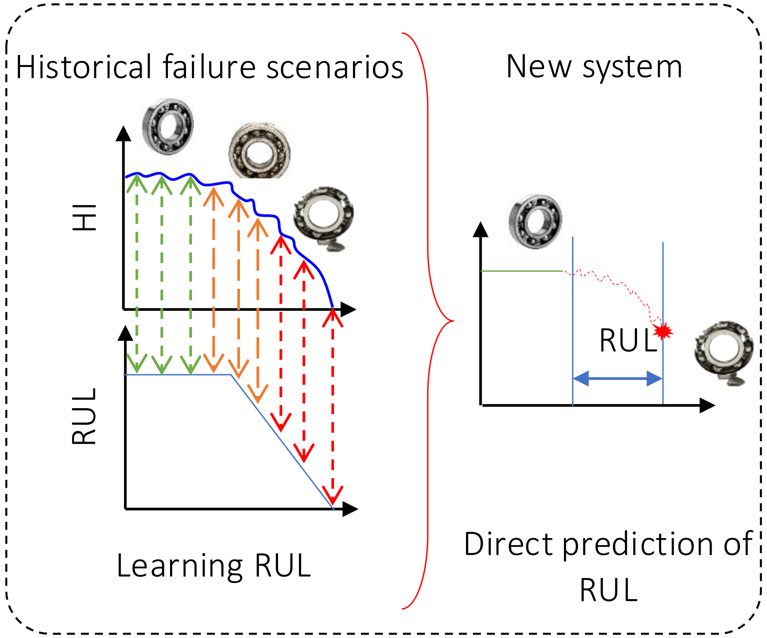
Similarity-based RUL estimation



Recursive RUL estimation



Direct RUL estimation



Similarity-based RUL estimation

- ✓ Easy implementation and fast computing time
- ✓ Perform good predictions
- ✗ Sensitive to condition variations
- ✗ Require large amount of similar data

Recursive RUL estimation

- ✓ Pre-determination of failure threshold
- ✓ Appropriate for incomplete data
- ✗ Difficult in non-stationary conditions
- ✗ Reduced modeling quality due to low data

Direct RUL estimation

- ✓ Perform high accuracy prediction
- ✓ Adaptive to the system variation
- ✗ Require large amount of similar data
- ✗ Take a lot of computation time

Address challenges in the proposed data

1

- **Addressing the variability of end-of-life on predictors:** failures at different times with high variance level exist.

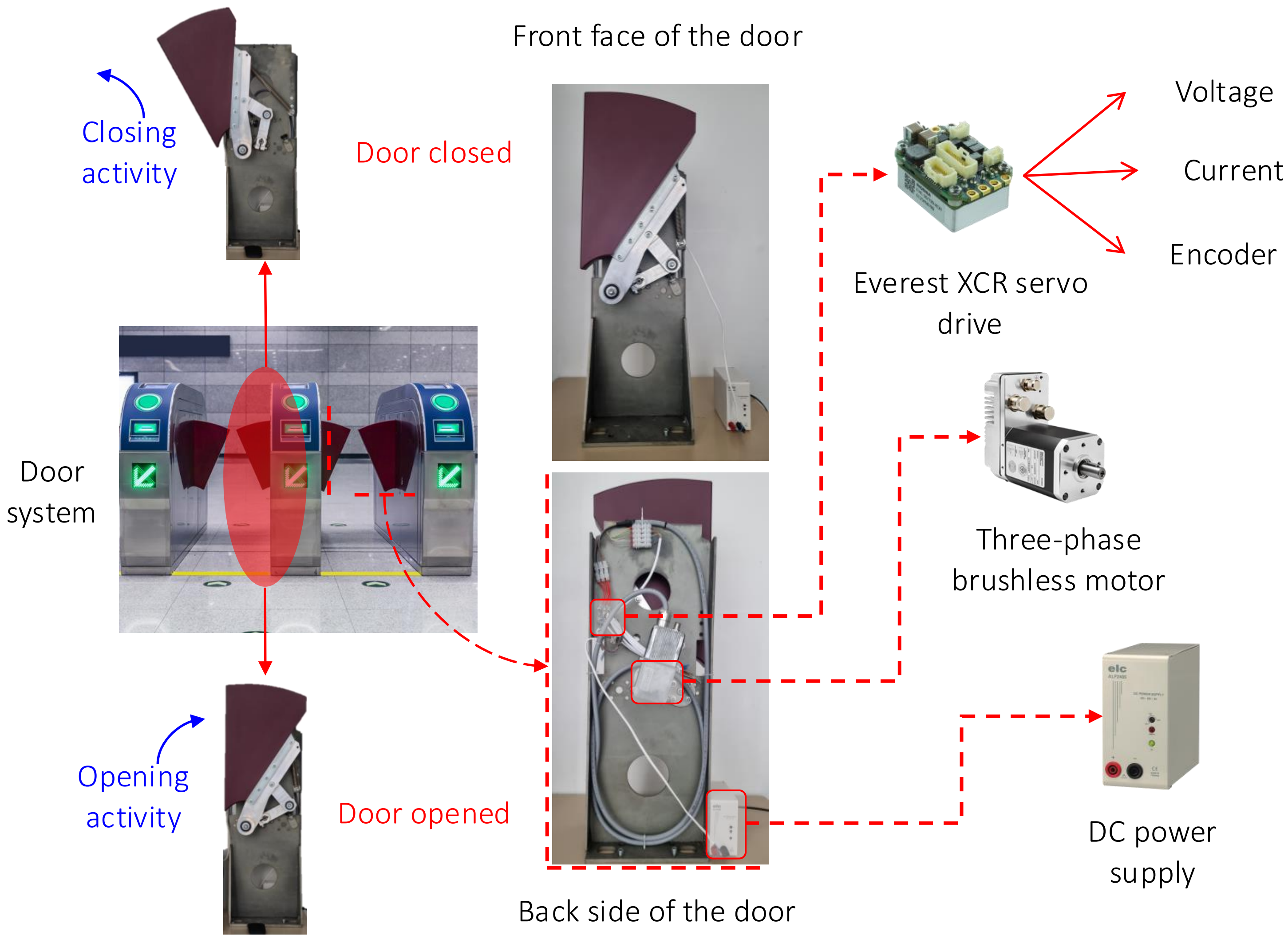
2

- **Addressing the uncertainty of predictions of end-of-life amplitudes**: the prognostic indicator failure time does not end at the same time.

3

- **Addressing the variability of degradation speed**: we can find a high life cycle with a rapid degradation trend and sometimes with a slow degradation trend, and reversely.

Case study – subway door



Case study – subway door

In this case study, to generate the failure scenario, the degradation time as well as the negative impact of the random shocks are simulated the Everest XCR controller.

During the closing activity of the door, a degradation on maximal position degree is triggered randomly over time.

$$X_{\max}(t_{s_n}) = [1 - p(t_{s_n})] \cdot X_{\max}(t_{s_{n-1}})$$

$X_{\max}(t_{s_n})$ and $X_{\max}(t_{s_{n-1}})$ are the **maximal position** degrees of the subway at the nth and the $(n-1)$-th times, respectively.

$p(t_{s_n})$ is the **degradation percentage**, and it is chosen randomly from predefined ranges to reflects the severity ' of degradation.

t_{s_n} is the **occurrence time** of the shocks and it reflects the different degradation rates (degradation speed) of the door.

The variation of these two parameters allows to generate multiple **possibilities** of failure, representing a dynamic behavior of the system degradation.

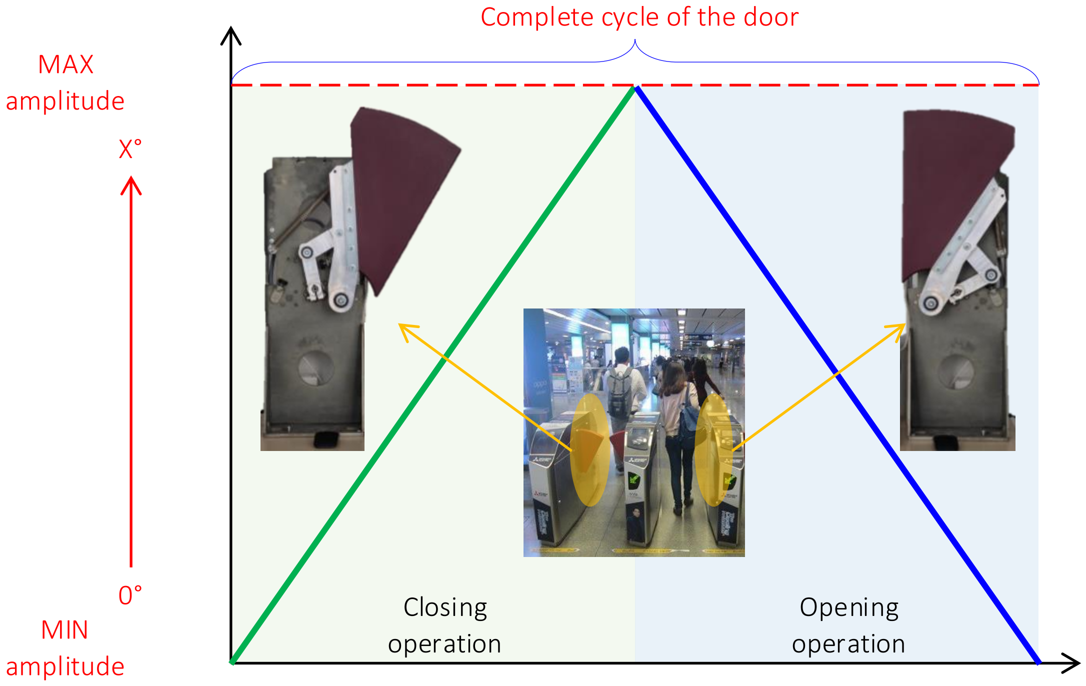


Illustration of degradation generation

$$X_{max}\left(t_{s_n}\right) = \left[1 - p\left(t_{s_n}\right)\right] \cdot X_{max}\left(t_{s_{n-1}}\right)$$

$P_1[2:5]\%$. $t_{s1}[3600:10000]$s → p_1 = 3% & t_{s1} = 7200s → $X_{max}(t_{s1})$ = (1-0.03)x90° = 87,3°

$P_2[4:9]\%$ $t_{s2}[3600:10000]$s → p_2 = 8% & t_{s2} = 4000s → $X_{max}(t_{s2})$ = (1-0.08)x87,3° = 80,31°

$P_3[20:30]\%$ $t_{s3}[500:1500]$s → p_3 = 25% & t_{s2} = 1000s → $X_{max}(t_{S3})$ = (1-0.25)x 80,31° = 20,23°

…

$P_n[5:15]\%$ $t_{sn}[60:1000]$s → p_n = 14% & t_{sn} = 700s → $X_{max}(t_{Sn})$ = (1-0.14)x20,23 = 17,39°

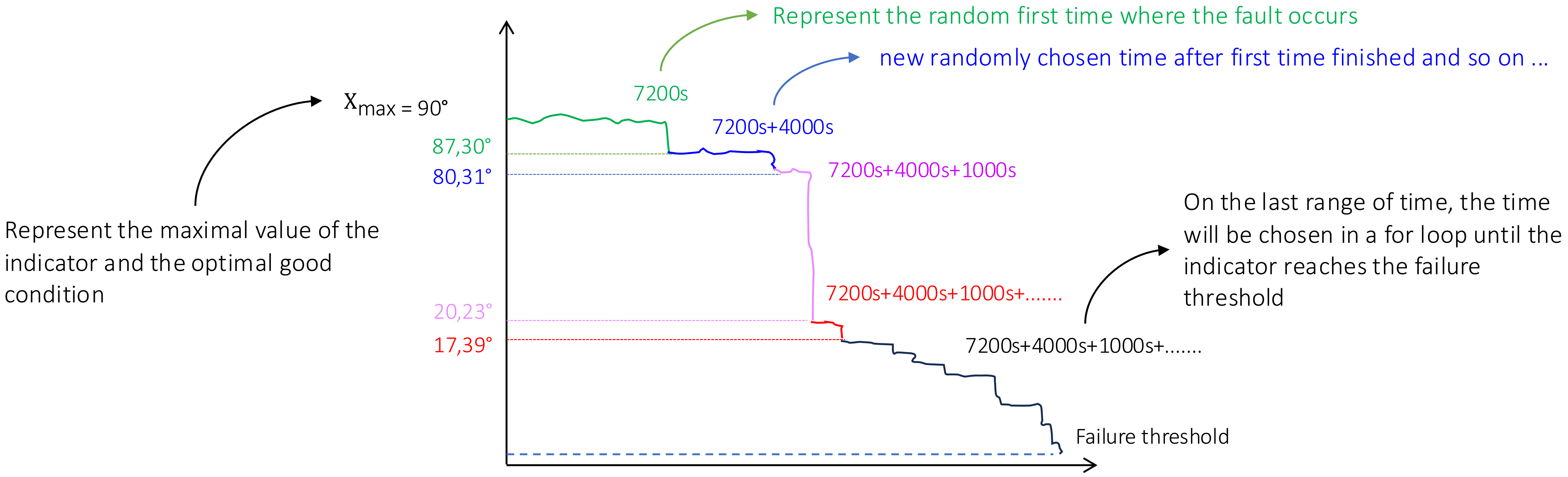
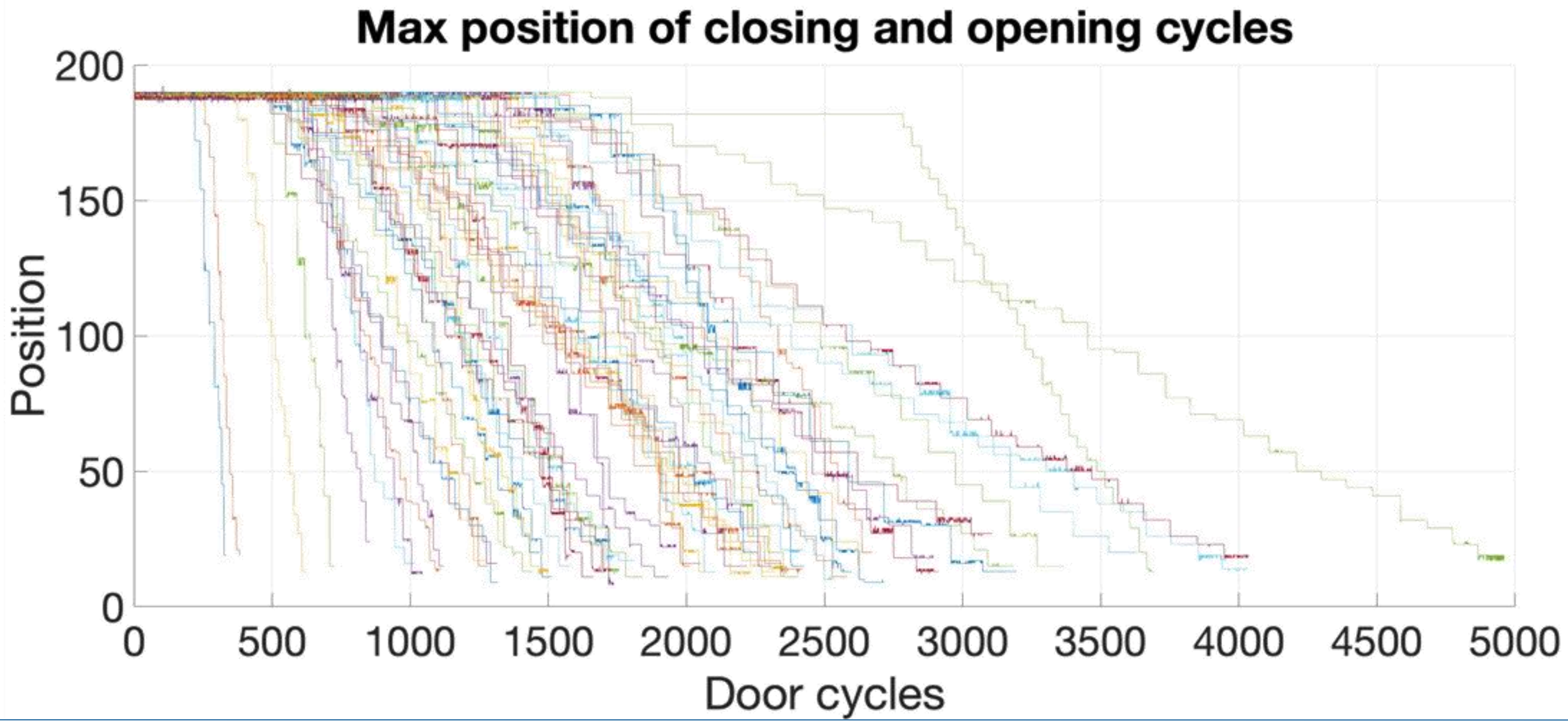


Illustration of degradation generation

Example of a health indicator using the MAX value as feature of position cycles



Address challenges in the proposed data

1

- **Addressing the variability of end-of-life on predictors:** failures at different times with high variance level exist.

2

- **Addressing the uncertainty of predictions of end-of-life amplitudes**: the prognostic indicator failure time does not end at the same time.

3

- **Addressing the variability of degradation speed**: we can find a high life cycle with a rapid degradation trend and sometimes with a slow degradation trend, and reversely.

Information on collected data

Operating conditions of the INGENIA platform				
Velocity	Acceleration	Deceleration	Experiments	Acquisition parameters
15	18	18	22 scenarios of degradation	Hardware: INGENIA servomotor
10	15	15	22 scenarios of degradation	File extension: .csv
15	18	15	23 scenarios of degradation	Time: 600 samples/file

Collected data (.csv files) Description							
Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Time	POS_REF	POS_FBK	VEL_FBK	VEL_FBK	FBK_DIGHALL	FBK_DIGENC1	DRV_PROT_VBUS
Duration of acquisition	Reference of desired position	Feedback of actual position	Reference of desired velocity	Feedback of actual velocity	Feedback of digital hall	Feedback of digital encoder	Voltage bus level of the driver
Column 9	Column 10	Column 11	Column 12	Column 13	Column 14	Column 15	Column 16
MOT_PROT_TEMP	FBK_CUR_A	FBK_CUR_B	FBK_CUR_C	DRV_PROT_TEMP	FBK_VOL_A	FBK_VOL_B	FBK_VOL_C
Temperature of the motor	Feedback current A	Feedback current B	Feedback current C	Temperature of the driver circuitry	Feedback voltage A	Feedback voltage A	Feedback voltage C

Structure of folders of data sets

An opening and a closing file acquisition constitute a cycle

Each file have 600 of points or 600 of points/position

There is 47 training failures : data are from begining of life to failure

A RUL file/failure is provided at the end of each folder

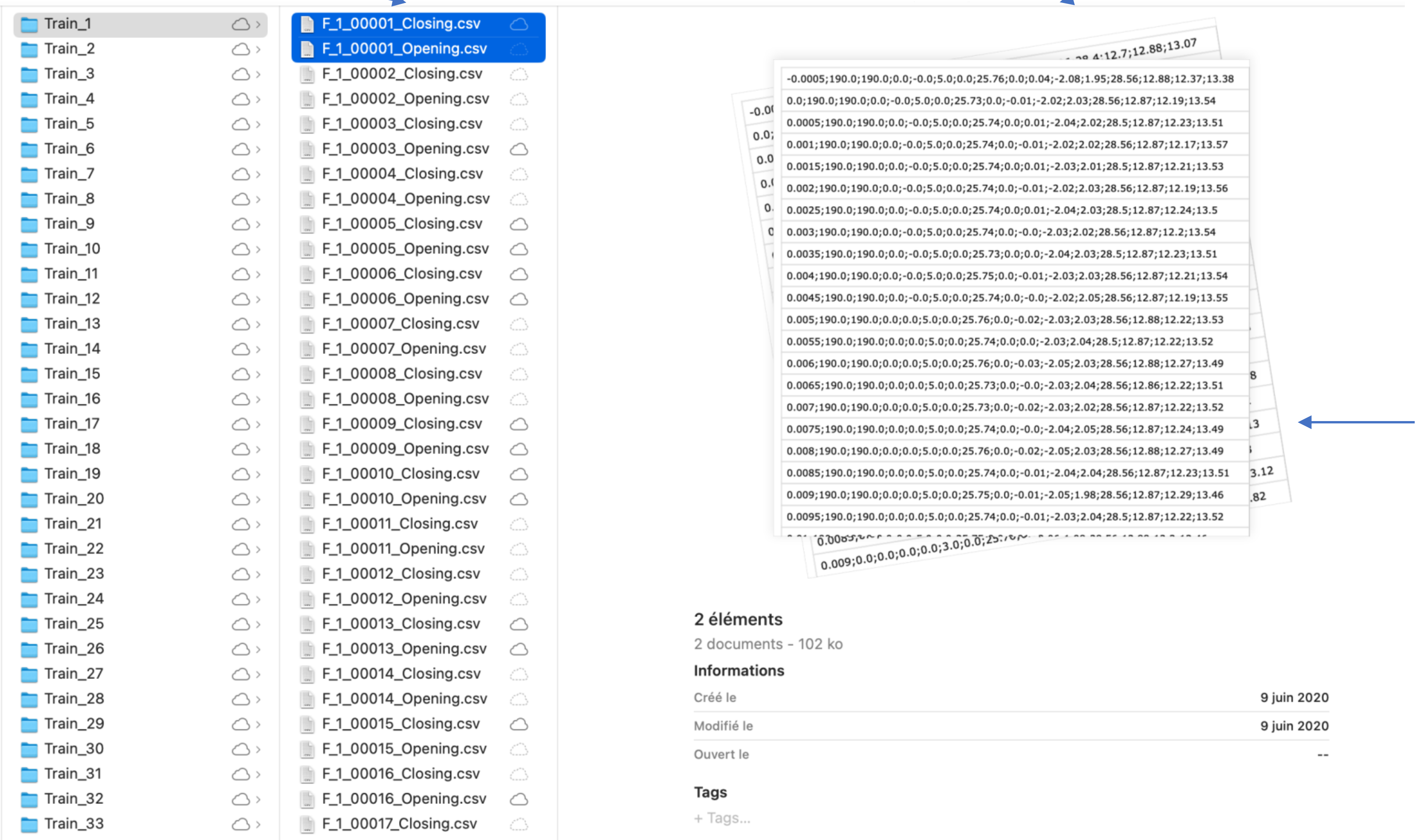
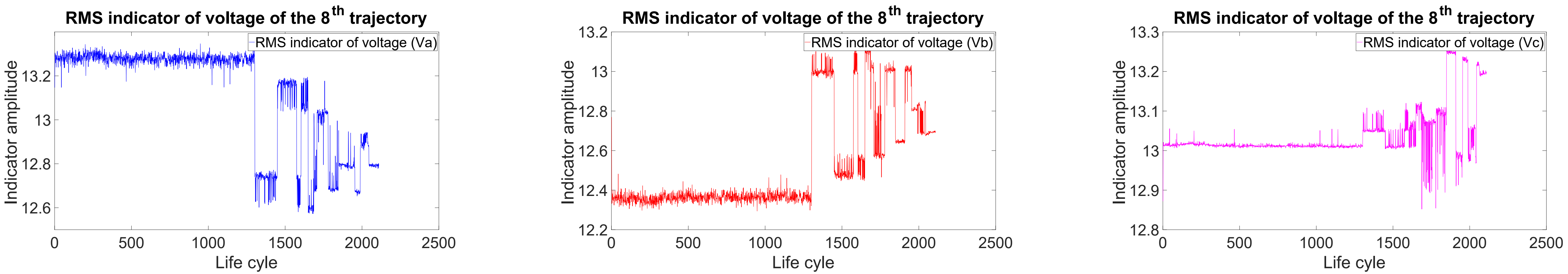


Illustration of current and voltage data

a) RMS health indicators constructed from the three-phase voltage signals



b) RMS health indicators constructed from the three-phase current signals

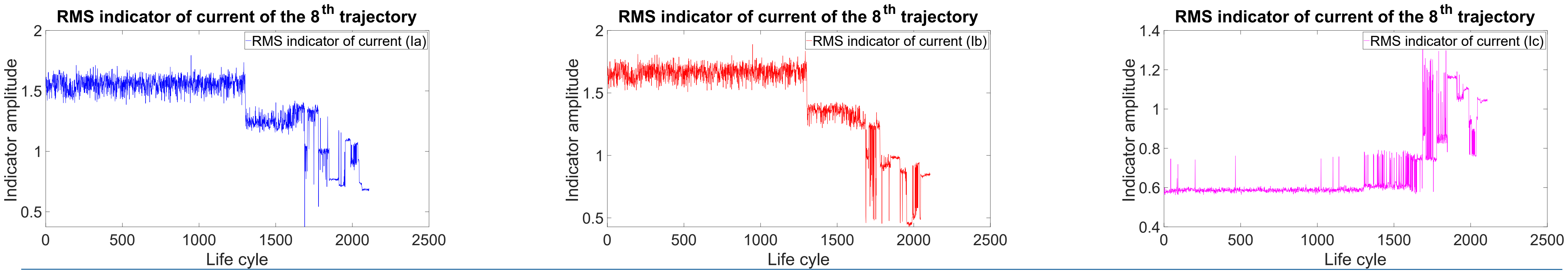
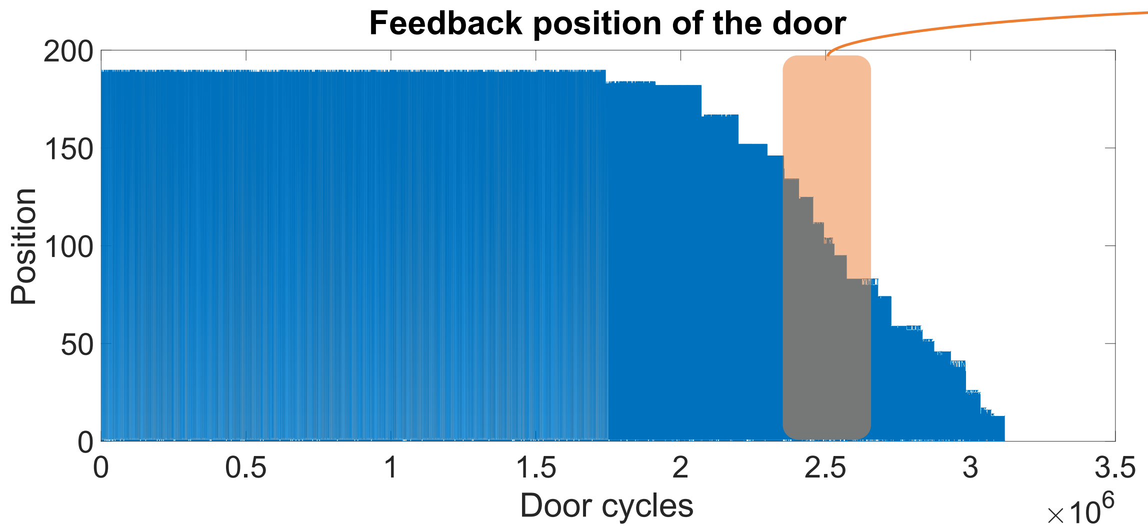
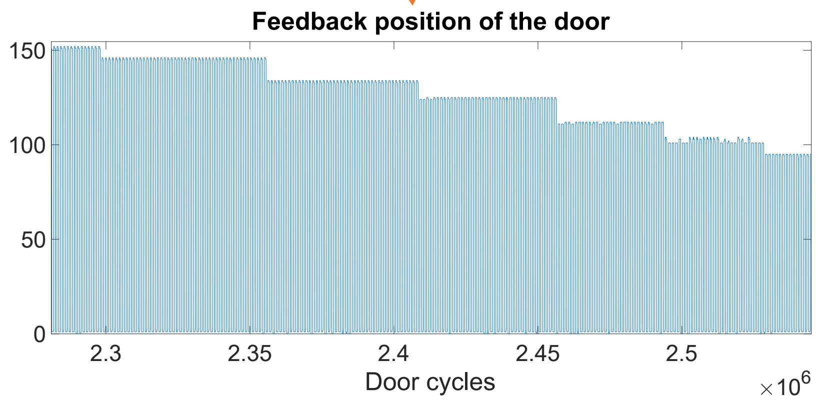
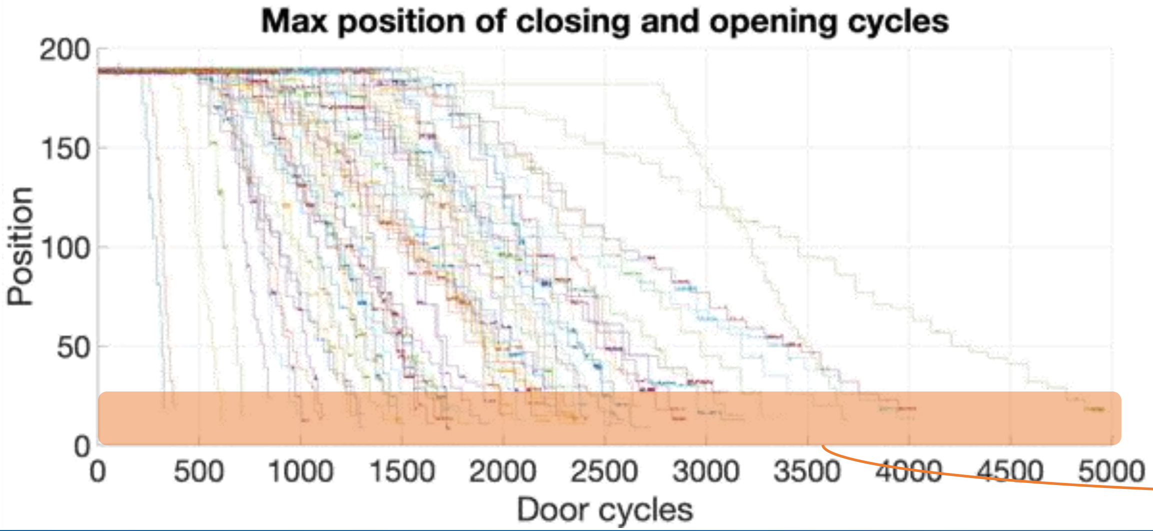
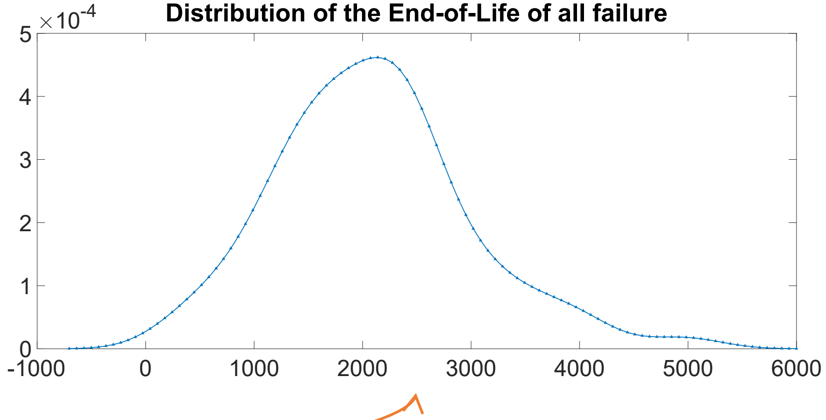


Illustration of position data









Evaluation metrics

- **Precision Metric**

The precision metric quantifies the prediction error's dispersion around its mean. Given T be the total prediction, it can be calculated based on the following equation:

$$\sqrt{\frac{1}{\mathrm{T}}\sum_{\mathrm{t}=1}^{\mathrm{T}}(\epsilon(\mathrm{t}) - \bar{\epsilon})}$$

ϵ is the error at time t, defined as the difference between the real RUL and the estimated RUL at time t: $\mathrm{RUL}_{\mathbf{real(t)}} - \mathrm{RUL}_{\mathbf{test(t)}}$

$\bar{\epsilon}$ is the mean of all errors .

T is the total number of prediction points.

To implement this, lets assume that we have only 5 observations for a test failure scenario. Here, we calculate the metric for all predicted values of the gived observations and avoid to use only the last value.

- **Root Mean Square Error (RMSE) Metric Calculation**

The Root Mean Square Error (RMSE) is a standard way to measure the error of a model in predicting quantitative data. The RMSE is calculated using the following formula:

$$RMSE = \sqrt{\frac{1}{\mathrm{T}}\sum_{\mathrm{t}=1}^{\mathrm{T}}\epsilon(\mathrm{t})^2}$$

Evaluation metrics

- **Accuracy Metric**

The Accuracy (Acc) metric quantifies the precision of Remaining Useful Life (RUL) predictions over a series of prediction points, emphasizing the exponential penalty for deviations from the actual RUL values. This metric is calculated using the following formula:

$$Acc = \frac{1}{T}\sum_{t=1}^{T} e^{\frac{-|\epsilon(t)|}{RUL_{real(t)}}}$$

- **Prognostic Horizon**

The Prognostic Horizon (PH) measures the period before the actual end-of-life (EoL) during which the predicted Remaining Useful Life (RUL) values consistently stay within a specified tolerance of the true RUL. This metric is crucial for effective predictive maintenance planning. The Prognostic Horizon is defined by the equation:

$$PH = t_{EoL} - t_{\alpha}$$

t_{EoL} is the actual end-of-life time.
t_{α} is the earliest time at which the predicted RUL remains within the acceptable bounds consistently.

The acceptable bounds are defined by the formula:

$$[RUL_{real(t)} - \alpha \times t_{EoL,} RUL_{real(t)} + \alpha \times t_{EoL,}]$$

Where α is a percentage that defines the tolerance range around the true RUL.

Evaluation metrics

- **Normalization of Each Metric**

The idea is to use an exponential function that maps the metric value to a range between 0 and 1, with the function asymptoti cally approaching 1.

- Normalized PH

as its values increase, and inversely for RMSE and Pre, where the function should asymptotically approach 1 as the RMSE value decreases.
For Prognostic Horizon (PH):

$$PH_{norm} = \frac{t_{EoL} - t_{in}}{t_{EoL}}$$

- **Normalized Root Mean Square Error (RMSE) or Precision (Pre)**

Given that lower RMSE values are better and we want the normalized value to approach 1 as the RMSE value approaches 0, we sho uld consider a function that gently decays from 1 to 0 as m increases, while meeting the specific value requirements at :

$$Prec_{norm}, RMSE_{norm} = f(m) = \frac{a}{m^b + c} + d$$

Where a, b, c and d can be optimized using fsolver function with initial values (scipy.optimize package)

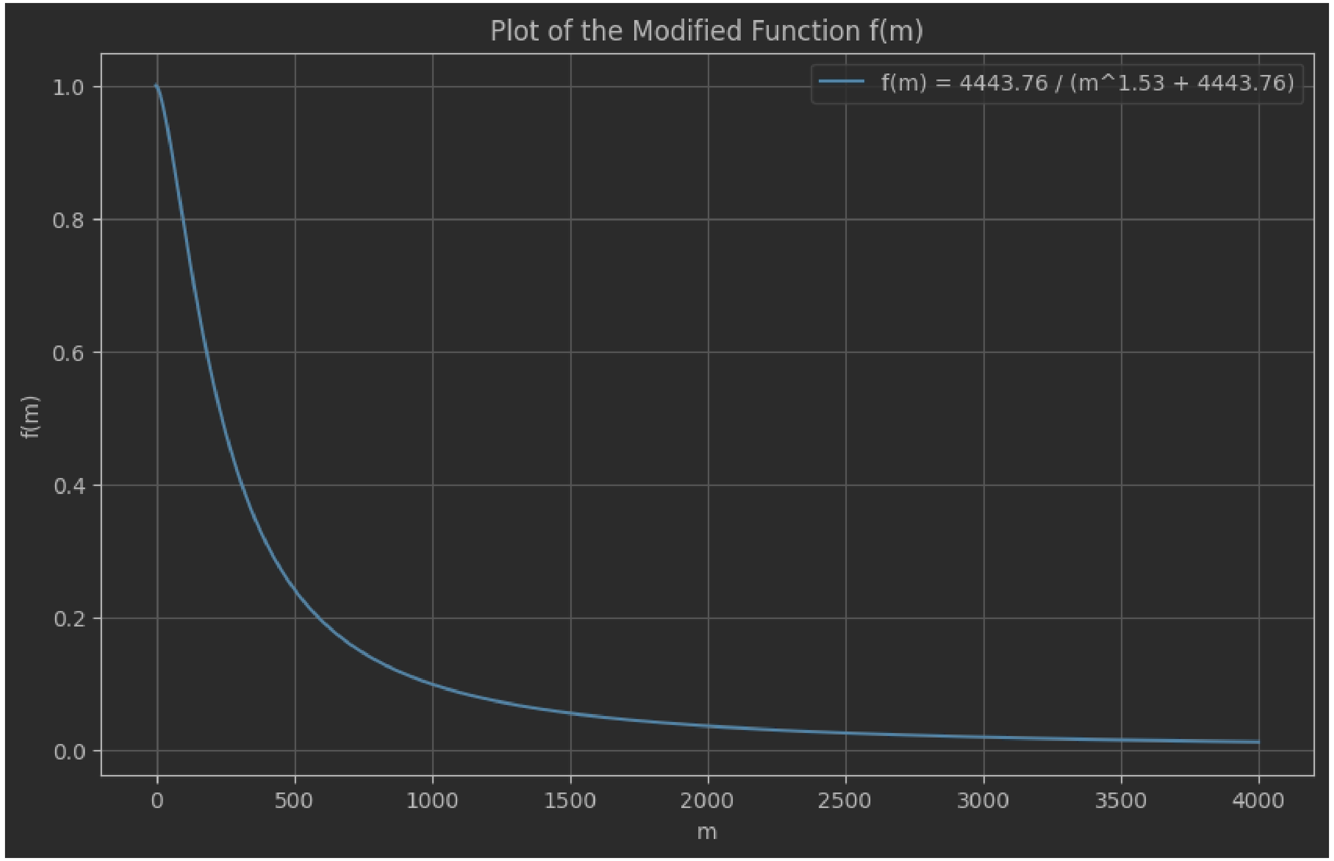
- Combine all metrics: Combinemetricsetric

$$Score = \frac{ACC + Prec_{norm} + RMSE_{norm} + \beta \times PH_{norm}}{3 + \beta}$$

Where β is applied for a balance between the (precision, accuracy, root mean square error), and the prognostic horizon

Soualhi, M., Nguyen, K. T., Medjaher, K., Nejjari, F., Puig, V., Blesa, J., ... & Marlasca, F. (2023). Dealing with prognostics uncertainties: Combination of direct and recursive remaining useful life estimations. *Computers in Industry*, *144*, 103766.

Evaluation metrics



$$\text{Score} = \frac{\text{ACC} + \text{Prec}_{\text{norm}} + \text{RMSE}_{\text{norm}} + \beta \times \text{PH}_{\text{norm}}}{3 + \beta}$$

Where β is applied for a balance between the (precision, accuracy, root mean square error), and the prognostic horizon.

Soualhi, M., Nguyen, K. T., Medjaher, K., Nejjari, F., Puig, V., Blesa, J., ... & Marlasca, F. (2023). Dealing with prognostics uncertainties: Combination of direct and recursive remaining useful life estimations. *Computers in Industry*, *144*, 103766.

Other information

- The overall and detailed description (with this presentation) will be online at the conference website
- Instructions for the submission of results are provided on the announcement page
- There will be a deadline to submit final results
- Prizes will be awarded to the first three winning groups
- Python/Matlab code will be provided to calculate results and metrics
- Data will be available on different servers

....

https://www.linkedin.com/in/moncef-soualhi-078240138/
E-mail: moncef.soualhi@umlp.fr

Thank you for your attention!