NTNU
Norwegian University of Science and Technology

# A comparison study for bearing remaining useful life prediction by using standard stochastic approach and digital twin

PhD Candidate: **Jie Liu**

Supervisor: *prof.* **Shen Yin**

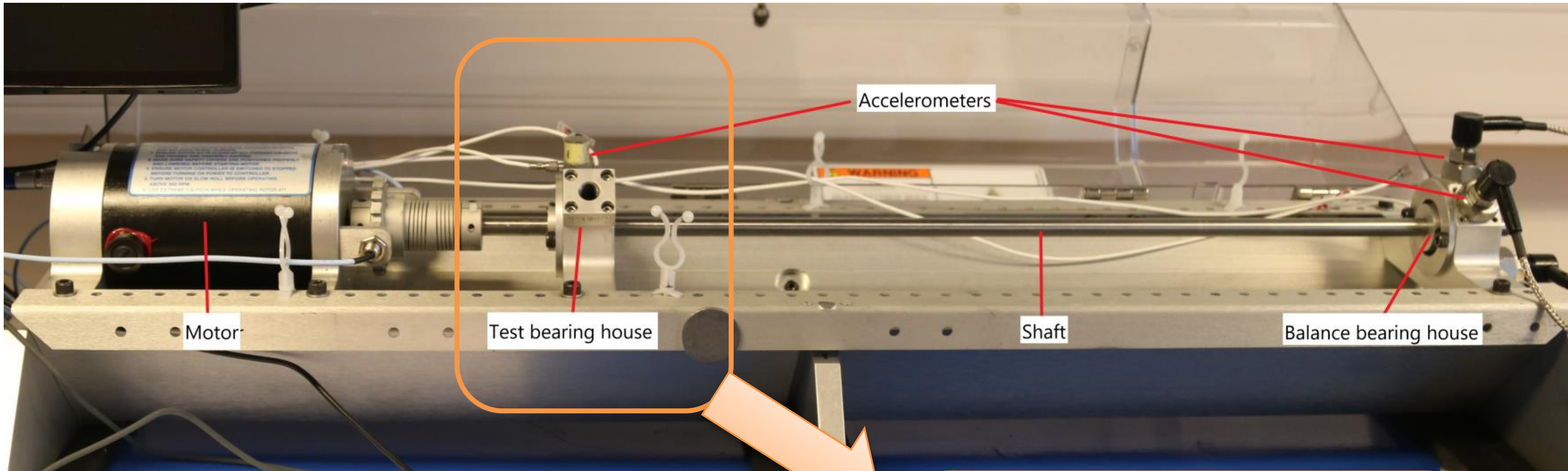
Co-supervisor: *prof.* **Jørn Vatn**

*Prepared for RAMS seminer*
*Date: 10/03/22*

# Comparison study for bearing remaining useful life prediction

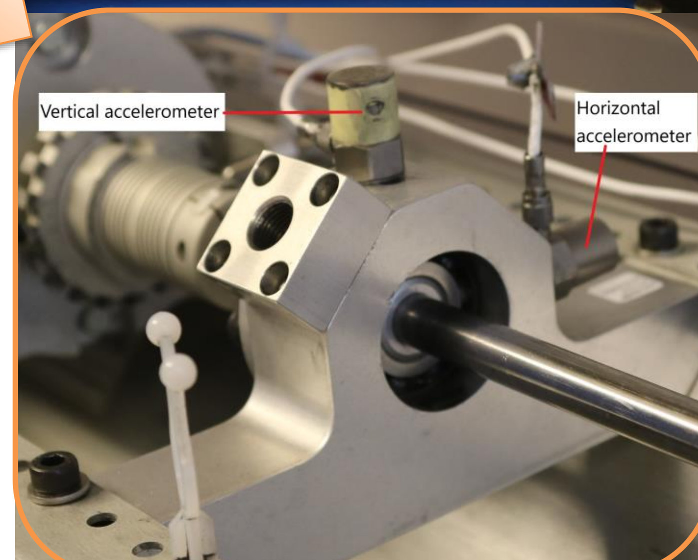
- Authors: Jie Liu, Jørn Vatn, Viggo Gabriel Borg Pedersen, Shen Yin
- Journal： International Journal of Reliability and Safety
- Bearing experiment
- RUL prediction
  - Stochastic approach - Wiener process
  - Stochastic approach - GBM
  - Matlab digital twin

# Experiment setup

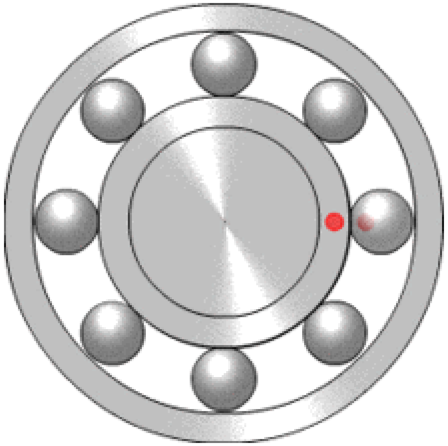




**Amplifier: coupler**



# Bearings



| | |
|---|---|
| Type of Bearing | Open roller/ball bearings |
| Number of balls | 10 Balls |
| Pitch diameter (B) | 70 mm |
| Ball diameter | 4.7 mm |
| Inner diameter(d) | 15.9 mm |
| Outer diameter(D) | 34.9 mm |

# Experiment output

Experiment form

| | | | |
|---|---|---|---|
| Bearing No. | 5 | Date | 2020-05-22 |

| Activity | Start time | End time | Reason |
|---|---|---|---|
| 1 | 8:48 | 10:08 | Compute required to restart. |
| 2 | 10:13 | 18:06 | No stress between data 16 and 16-1, saved for comparing between stops. |
| 3 | | | |
| 4 | | | |
| Data time interval | | | 5mins |

Stress level
(Target stress shall be marked)

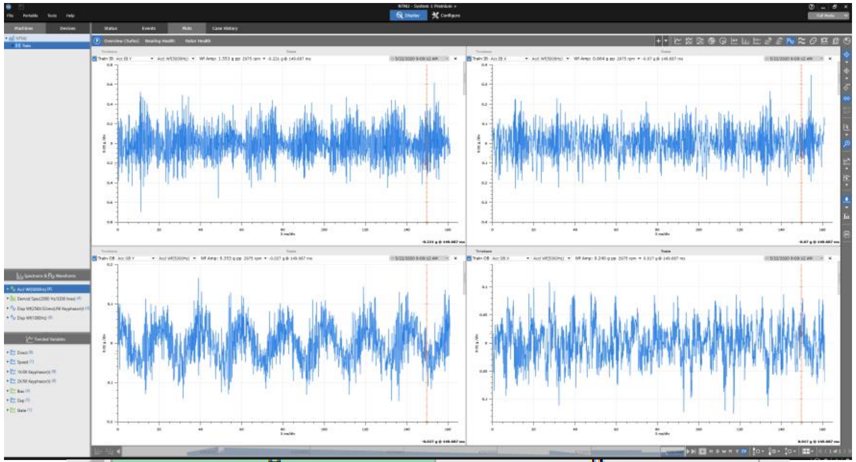
| Speed | Size of silicon | Density of silicon |
|---|---|---|
| 3005rpm(input), 2975rpm(system) | S M L | 1 pour /time, 0.72 grams |

Stress change log

| | Time | | Time | | Time | | Time | | Time | | Time |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 8:53 | 7 | 11:28 | 13 | 13:58 | 19 | 16:28 | | | | |
| 2 | 9:18 | 8 | 11:53 | 14 | 14:23 | 20 | 16:53 | | | | |
| 3 | 9:43 | 9 | 12:18 | 15 | 14:48 | 21 | 17:18 | | | | |
| 4 | 10:13 | 10 | 12:43 | 16 | 15:13 | 22 | 17:43 | | | | |
| 5 | 10:38 | 11 | 13:08 | 17 | 15:38 | 23 | | | | | |
| 6 | 11:03 | 12 | 13:33 | 18 | 16:03 | 24 | | | | | |

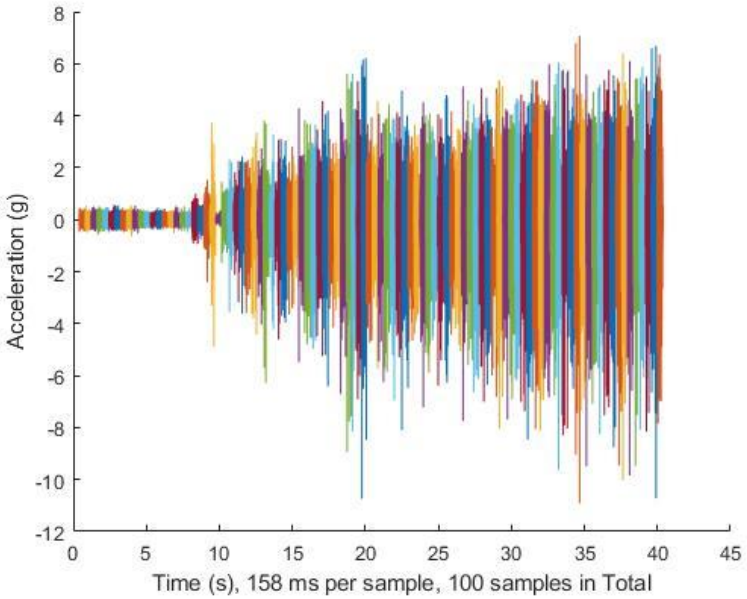
Summary

| | | | |
|---|---|---|---|
| Highest accelerate value | 9.711 g (110b) | Lowest accelerate value | -10.162 g (110b) |
| Note | Train IB ACC IBY ia.csv<br>Tested bearing, Vertical data<br>Train OB ACC OBY ic.csv<br>Balance bearing, Vertical data<br>No mixture left. To be changed next time. | Train IB ACC IBX ib.csv<br>Tested bearing, Horizontal data<br>Train OB ACC OBX id.csv<br>Balance bearing, Horizontal data | |
| Number of data file | 110*4 | Prepared by | Jie Liu |



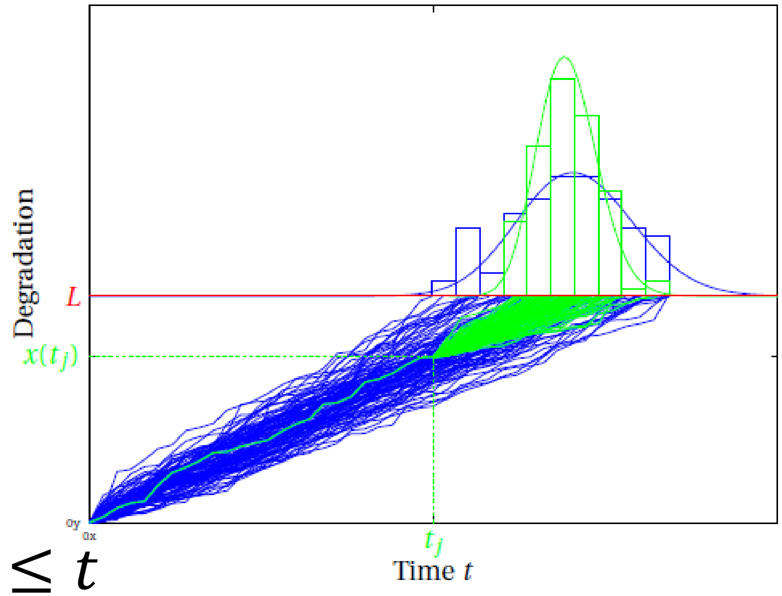
| | A | B |
|---|---|---|
| 1 | Machine Name | Train |
| 2 | Point Name | Acc IB Y |
| 3 | Wf Amp | 0.006 |
| 4 | Number of Revs | 21.07337761 |
| 5 | X-Axis Unit | ms |
| 6 | Y-Axis Unit | g |
| 7 | Sample Speed | 1975 rpm |
| 8 | Sample Status | Valid |
| 9 | Timestamp | 5/15/2020 10:06 |
| 10 | Variable | Accl Wf(5000Hz) |
| 11 | X-Axis Value | Y-Axis Value |
| 12 | 0 | -0.001 |
| 13 | 0.078 | 0.001 |
| 14 | 0.156 | 0.001 |

# Original data vs feature selection




| Feature | Expression |
|---|---|
| Mean | $\frac{1}{N}\sum_{i=1}^{N} x_i$ |
| Standard deviation (Std) | $\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - mean)\right)^{\frac{1}{2}}$ |
| Skewness | $\frac{1}{N}\cdot\sum_{i=1}^{N}\frac{(x_i-\bar{x})^3}{\rho^3}$ |
| Kurtosis | $\frac{\frac{1}{N}\cdot\sum_{i=1}^{N}(x_i-\bar{x})^4}{\left(\frac{1}{N}\cdot\sum_{i=1}^{N}(x_i-\bar{x})^2\right)^2}$ |
| Peak to Peak | $x_{max} - x_{min}$ |
| RMS (Root Mean Square) | $\left(\frac{1}{N}\sum_{i=1}^{N} x_i{}^2\right)^{\frac{1}{2}}$ |
| CrestFactor | $\frac{x_{max}}{RMS}$ |
| ShapeFactor | $\frac{RMS}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$ |
| ImpulseFactor | $\frac{x_{max}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$ |
| MarginFactor | $\frac{x_{max}}{\left(\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert\right)^2}$ |
| Energy | $\sum_{i=1}^{N} x_i^2$ |
| Absolute Maximum | $\lvert x_{max}\rvert$ |

# Stochastic approach I - Wiener process




- Degradation process
- $Y(t) = y_0 + \mu t + \sigma W(t)$
- $\Delta Y \sim N\left(\mu(t-s), \sigma^2(t-s)\right), 0 \le s \le t$
- Inverse Gaussian distribution method
- The CDF of first passage time
- $F_T(t) = \Phi\left(\frac{\sqrt{\lambda}}{v}\sqrt{t} - \frac{\sqrt{\lambda}}{\sqrt{t}}\right) + \Phi\left(-\frac{\sqrt{\lambda}}{v}\sqrt{t} - \frac{\sqrt{\lambda}}{\sqrt{t}}\right) e^{\frac{2\lambda}{v}}$
- Where $v = L/\mu$ and $\lambda = L^2/\sigma^2$
- $E(T) = \frac{L}{\mu}$ $\qquad Var(T) = \frac{L\sigma^2}{\mu^3}$

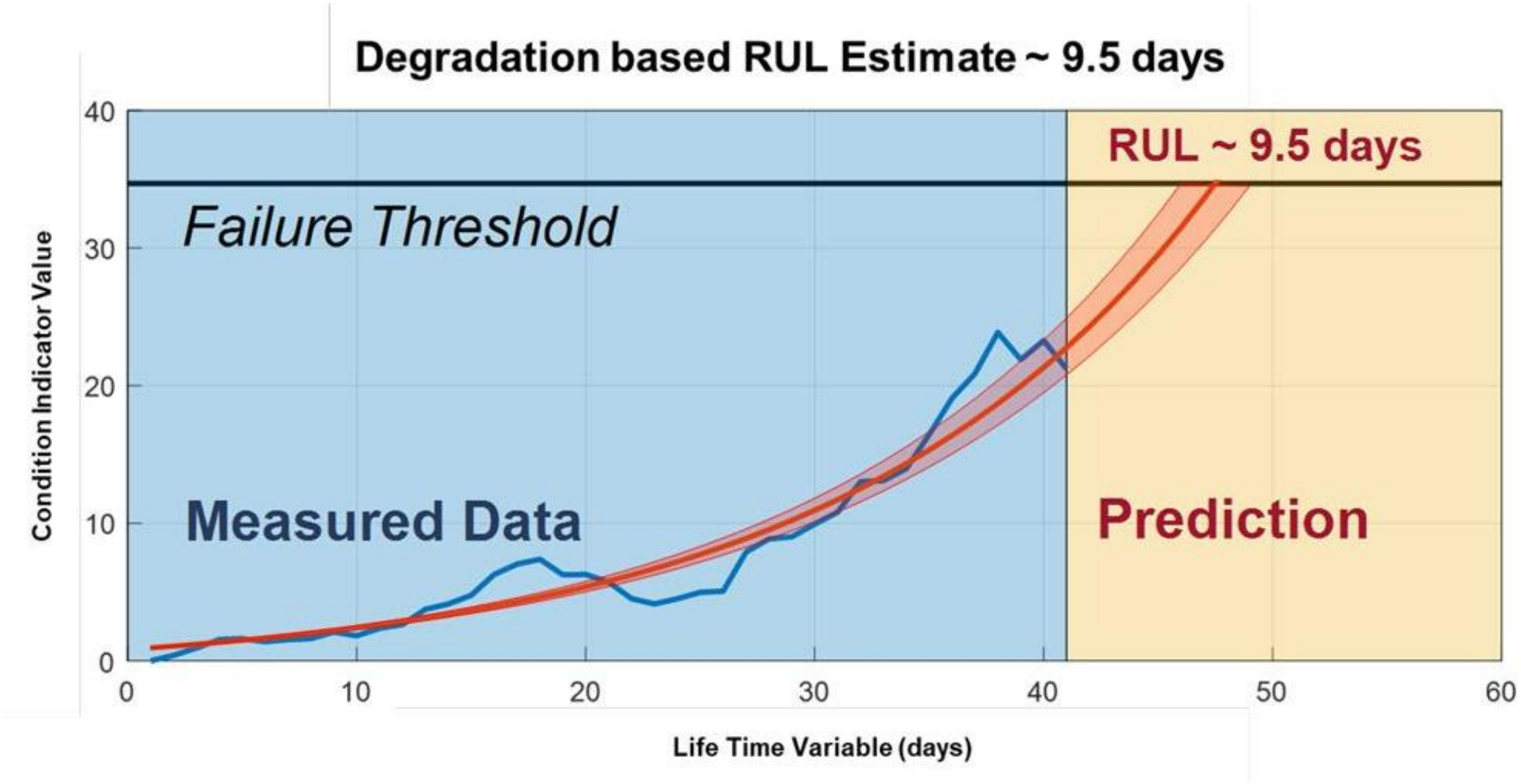
# Stochastic approach II
# - Geometric Brownian Motion

- Degradation process
- $dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$
- $S_t = S_0 \exp(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t))$

- The CDF of first passage time
- $F_T(t) = \Phi\left(\frac{\sqrt{\lambda}}{v}\sqrt{t} - \frac{\sqrt{\lambda}}{\sqrt{t}}\right) + \Phi\left(-\frac{\sqrt{\lambda}}{v}\sqrt{t} - \frac{\sqrt{\lambda}}{\sqrt{t}}\right) e^{\frac{2\lambda}{v}}$
- Where $v = \frac{lnL - lnS_0}{\mu - \frac{1}{2}\sigma^2}$ and $\lambda = \frac{(lnL - lnS_0)^2}{\sigma^2}$
- $E(T) = \frac{lnL - lnS_0}{\mu - \frac{1}{2}\sigma^2}$ $Var(T) = \frac{\sigma^2(lnL - lnS_0)}{(\mu - \frac{1}{2}\sigma^2)^3}$

# Matlab Digital twin

- $h(t) = \phi + \theta * \exp(\beta t + \epsilon - \frac{\sigma^2}{2})$



# Results

| Status | Standard Deviation | Peak to Peak | RMS | Energy |
|---|---|---|---|---|
| Initial status | 3.49E-01 | 2.53E+00 | 3.49E-01 | 1.71E+01 |
| Threshold | 2.15E+00 | 1.65E+01 | 2.15E+00 | 9.54E+03 |
| Lifetime | 500 minutes | | | |

# Comparison of stochastic approaches

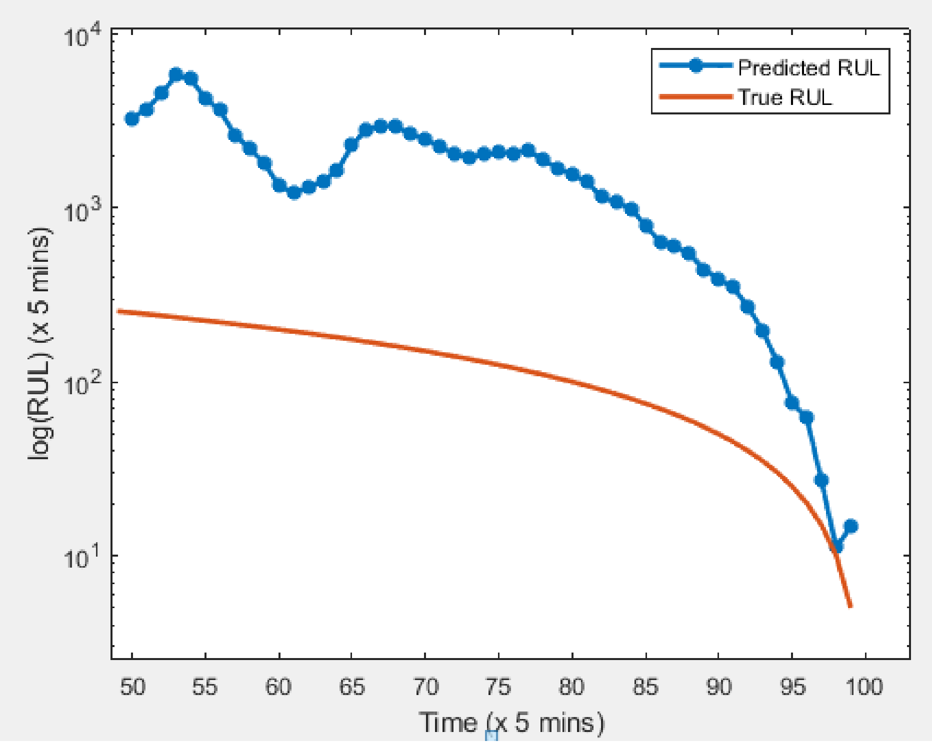
| Stochastic Process | Parameter | Standard Deviation | Peak to Peak | RMS | Energy |
|---|---|---|---|---|---|
| Wiener process | $\hat{\mu}_{50}$ | 1.57E-03 | 1.63E-02 | 1.57E-03 | 2.70E+00 |
| | $\hat{\sigma}_{50}$ | 8.44E-03 | 2.63E-01 | 8.43E-03 | 1.46E+01 |
| | $\hat{\mu}_{99}$ | 3.84E-03 | 3.00E-02 | 3.84E-03 | 1.87E+01 |
| | $\hat{\sigma}_{99}$ | 1.65E-02 | 3.10E-01 | 1.65E-02 | 8.74E+01 |
| GBM | $\hat{\mu}_{50}$ | 3.88E-03 | 4.48E-03 | 3.88E-03 | 8.56E-03 |
| | $\hat{\sigma}_{50}$ | 5.28E-02 | 4.41E-01 | 5.28E-02 | 2.89E+00 |
| | $\hat{\mu}_{99}$ | 5.23E-03 | 4.45E-03 | 5.23E-03 | 1.21E-02 |
| | $\hat{\sigma}_{99}$ | 5.37E-02 | 4.07E-01 | 5.37E-02 | 4.11E+00 |

| Stochastic Process | Data sets No. | Method | Parameter | Standard Deviation | Peak to Peak | RMS | Energy |
|---|---|---|---|---|---|---|---|
| WP | First 50 | Inverse Gaussian distribution | E(T) | 1370 | 1011 | 1370 | 3530 |
| | | | Var(T) | 3.97E+04 | 2.62E+05 | 3.97E+04 | 1.03E+05 |
| | | | Prediction error | 174% | 102% | 174% | 606% |
| | | Monte Carlo simulation | E(T) | 1371 | 1017 | 1371 | 3531 |
| | | | Var(T) | 3.96E+04 | 2.65E+05 | 3.93E+04 | 1.02E+05 |
| | | | Prediction error | 174% | 103% | 174% | 606% |
| | First 99 | Inverse Gaussian distribution | E(T) | 559 | 550 | 559 | 510 |
| | | | Var(T) | 1.03E+04 | 5.86E+04 | 1.03E+04 | 1.11E+04 |
| | | | Prediction error | 12% | 10% | 12% | 2% |
| | | Monte Carlo simulation | E(T) | 561 | 552 | 560 | 511 |
| | | | Var(T) | 1.04E+04 | 5.92E+04 | 1.04E+04 | 1.12E+04 |
| | | | Prediction error | 12% | 10% | 12% | 2% |
| GBM | First 50 | Inverse Gaussian distribution | E(T) | 730 | NA | 730 | NA |
| | | | Var(T) | 3.29E+05 | NA | 3.29E+05 | NA |
| | | | Prediction error | 46% | NA | 46% | NA |
| | | Monte Carlo simulation | E(T) | 729 | NA | 734 | NA |
| | | | Var(T) | 3.09E+05 | NA | 3.11E+05 | NA |
| | | | Prediction error | 46% | NA | 47% | NA |
| | First 99 | Inverse Gaussian distribution | E(T) | 479 | NA | 479 | NA |
| | | | Var(T) | 9.58E+04 | NA | 9.58E+04 | NA |
| | | | Prediction error | -4% | NA | -4% | NA |
| | | Monte Carlo simulation | E(T) | 484 | NA | 481 | NA |
| | | | Var(T) | 9.68E+04 | NA | 9.42E+04 | NA |
| | | | Prediction error | -3% | NA | -4% | NA |

There are some limitation to use GBM. GBM follows log-normally distribution. The mean value is $lnS_0+(\mu-\frac{\sigma^2}{2})*t$ and variance is $\sigma^2 t$. If $\mu-\frac{\sigma^2}{2}\leq 0$, the process might go backwards and being absorbed at zero. Therefore, the drift parameter $\mu$ shall always be larger than $\frac{\sigma^2}{2}$, otherwise there is no result for prediction.

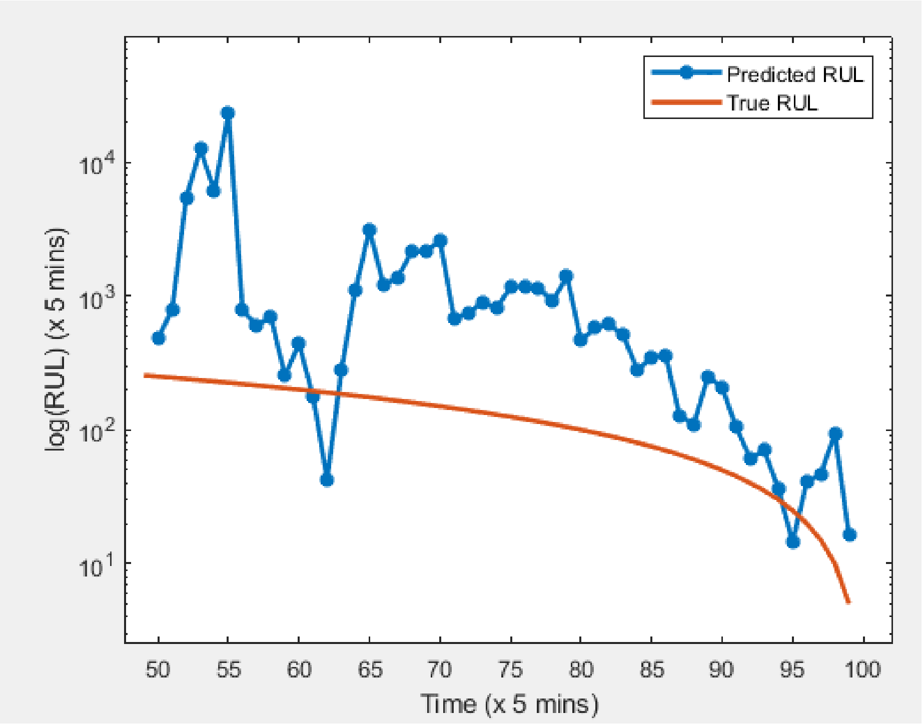
# Predicted RUL by Wiener process feature: Energy

| Samle No. | Current Time | Predicted Lifetime | True RUL | Predicate RUL | Variance | Predict error |
|---|---|---|---|---|---|---|
| 50 | 250 | 3530 | 250 | 3280 | 1.03E+05 | 606% |
| 55 | 275 | 4542 | 225 | 4267 | 3.68E+05 | 808% |
| 60 | 300 | 1654 | 200 | 1354 | 6.14E+04 | 231% |
| 65 | 325 | 2624 | 175 | 2299 | 3.84E+05 | 425% |
| 70 | 350 | 2807 | 150 | 2457 | 4.78E+05 | 461% |
| 75 | 375 | 2468 | 125 | 2093 | 3.17E+05 | 394% |
| 80 | 400 | 1954 | 100 | 1554 | 1.63E+05 | 291% |
| 85 | 425 | 1218 | 75 | 793 | 5.56E+04 | 144% |
| 90 | 450 | 840 | 50 | 390 | 2.54E+04 | 68% |
| 95 | 475 | 551 | 25 | 76 | 1.33E+04 | 10% |
| 99 | 495 | 510 | 5 | 15 | 1.11E+04 | 2% |



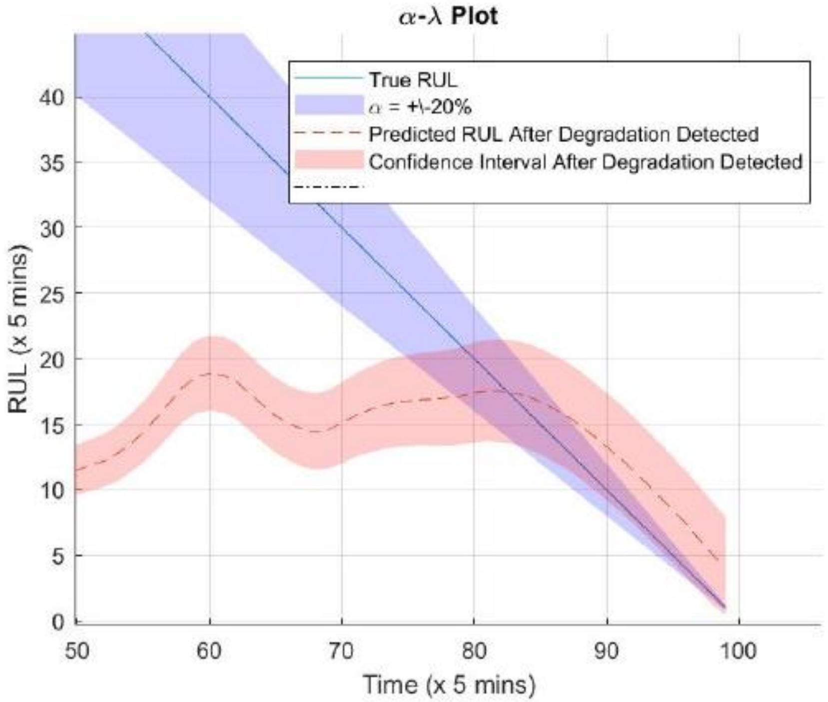
# Predicted RUL by GBM feature: standard deviation vs RMS

| Samle No. | Current Time | Predicted Lifetime | True RUL | Predicate RUL | Variance | Predict error |
|---|---|---|---|---|---|---|
| 50 | 250 | 730 | 250 | 480 | 3.29E+05 | 46% |
| 55 | 275 | -23227 | 225 | -23502 | -1.09E+10 | -4745% |
| 60 | 300 | 739 | 200 | 439 | 3.62E+05 | 48% |
| 65 | 325 | 3495 | 175 | 3170 | 4.39E+07 | 599% |
| 70 | 350 | 2965 | 150 | 2615 | 2.53E+07 | 493% |
| 75 | 375 | 1543 | 125 | 1168 | 3.48E+06 | 209% |
| 80 | 400 | 864 | 100 | 464 | 6.01E+05 | 73% |
| 85 | 425 | 773 | 75 | 348 | 4.13E+05 | 55% |
| 90 | 450 | 660 | 50 | 210 | 2.54E+05 | 32% |
| 95 | 475 | 489 | 25 | 14 | 9.95E+04 | -2% |
| 99 | 495 | 479 | 5 | -16 | 9.58E+04 | -4% |



# Predicted RUL by Matlab digital twin

| Sample No. | Current Time | Predicted RUL | Predicted lifetime | Variance | Prediction error |
|---|---|---|---|---|---|
| 50 | 250 | 57.45 | 307.45 | 2.60 | -39% |
| 55 | 275 | 71.69 | 346.59 | 3.51 | -31% |
| 60 | 300 | 94.34 | 394.34 | 5.25 | -21% |
| 65 | 325 | 78.22 | 403.22 | 5.13 | -19% |
| 70 | 350 | 75.34 | 425.34 | 6.14 | -15% |
| 75 | 375 | 83.84 | 458.84 | 6.11 | -8% |
| 80 | 400 | 86.98 | 486.98 | 6.41 | -3% |
| 85 | 425 | 83.61 | 508.61 | 5.76 | 2% |
| 90 | 450 | 66.39 | 516.39 | 12.37 | 3% |
| 95 | 475 | 42.32 | 517.32 | 12.17 | 3% |
| 99 | 495 | 19.83 | 514.83 | 11.92 | 3% |



# Conclusions

- All model's prediction results are improved when there are more information about the bearing's status. Number of required data sets: Matlab digital twin < GBM < Wiener process.
- When there is enough information, stochastic models have lower prediction error than the Matlab digital twin.
- After dataset 80, the prediction error of Matlab digital twin is around 3%.
- For stochastic approaches, single feature is used for prediction. While for Matlab digital twin, it used fusion features for prediction.
- Digital twin is that it takes longer time to run the model than stochastic approaches.


THANK YOU