Title: Learning Complex Degradation Signal Manifolds for Accurate RUL Prediction via Conditional Diffusion Models

Authors: Donghyun Ko*, Xiaolei Fang**
* Edward P. Fitts Department of Industrial and Systems Engineering, Ph.D. Student
** Edward P. Fitts Department of Industrial and Systems Engineering, Associate Professor

Abstract:
Conventional Remaining Useful Life (RUL) prediction methods rely heavily on heuristic assumptions about signal distributions, such as exponential growth with Brownian motion error models [1]. While computationally efficient, these approaches struggle to capture the complex, nonlinear, and multimodal characteristics present in real-world degradation signals. In this research, we propose a fundamentally different paradigm: leveraging Conditional Diffusion Models to directly learn the manifold of degradation signals from data without parametric assumptions. Our experiments compare traditional Bayesian updating methods and a typical machine learning (ML)-based benchmark (PCA + Lognormal regression) against diffusion-based approaches under varying data complexities and noise structures. The results are expected to demonstrate that Conditional Diffusion Models provide superior adaptability and accuracy, especially in scenarios with complex signal behaviors and multiple failure modes. This work contributes to the AI4SE and SE4AI community by showcasing how advanced generative models can enhance prognostic capabilities and support trustworthy AI-enabled system health management.

Keywords: Remaining Useful Life (RUL), Time-to-Failure(TTF), Bayesian Updating, PCA, Lognormal regression, Degradation Signal Modeling, Conditional Diffusion Models

1. Introduction

Predicting the Remaining Useful Life (RUL) of systems based on degradation signals is critical for enabling proactive maintenance and ensuring mission readiness. Conventional approaches often impose predefined distributions such as 'Exponential' or 'Exponential + Brownian Motion Error' for RUL distribution and subsequently validate these assumptions post-hoc by applying bayesian updating or M/L based estimation technics [1]. However, real-world degradation data are rarely so simple — they often exhibit nonlinear trends, non-Gaussian noise, and multimodal behaviors arising from complex failure mechanisms and varying operational conditions [2]. To address these limitations, we propose the use of Conditional Diffusion Models, which offer a data-driven framework for learning the true manifold of degradation signals. Unlike traditional models constrained by parametric forms, diffusion models can approximate arbitrary continuous distributions with high fidelity, thereby capturing the intricate dynamics of degradation processes. Moreover, diffusion models can serve directly as RUL prediction models. They can outperform conventional ML-based RUL predictors, even those considered state-of-the-art in data-driven methods, such as the 'PCA + Lognormal regression' model [3][4]. Their performance is particularly notable when applied to complex degradation data.

2. Methodology

Our study investigates two key research points such that:

a) Point 1: “From Heuristic Assumptions to Data-Driven Manifold Learning”

Conventional approaches like Bayesian update models (i.i.d model, BM model) do not learn the actual degradation distribution. Instead, they select the model that is closest to the real RUL from predefined forms such as 'Exponential + Brownian motion error' and compute RUL by updating the posterior of prior parameters using closed-form solutions. In contrast, Conditional Diffusion Models learn the actual degradation signal manifold without any parametric assumptions, enabling more faithful and adaptive RUL prediction. To validate this, we compare Bayesian update models (i.i.d model, BM model) against Conditional Diffusion Models as shown in Fig. 1. RMSE and MAPE are used as performance metrics for RUL predictions. Experiments involve synthetic degradation signals of two types:

- Type 1: Exponential growth + Brownian Motion error
- Type 2: Exponential growth + Sudden drop + Exponential growth + Brownian Motion error
  (Real-world systems like Lithium-Ion Batteries or Turbofan Engines under mode switching)

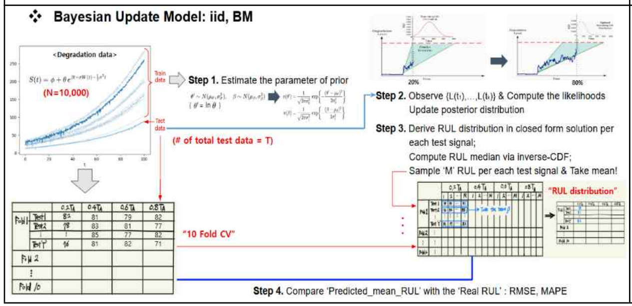
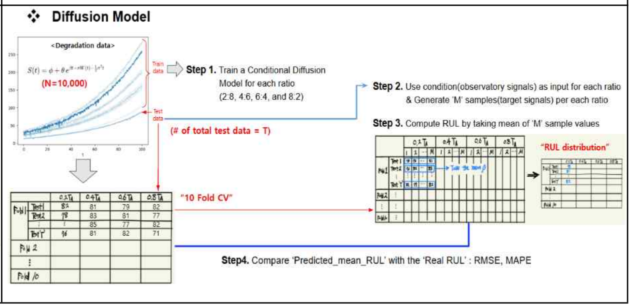
Bayesian Update Models	Diffusion Model
	

Fig. 1. Design of Experiment for Point 1

b) Point 2: Diffusion vs. Classical Machine Learning (PCA + Lognormal Regression)

Most ML-based RUL predictors use PCA to reduce the dimensionality of high-dimensional degradation data. However, since PCA projects the original high-dimensional data into a lower-dimensional linear subspace, many ML-based RUL predictors analyze RUL within this linear context. In contrast, diffusion models learn the full signal manifold without such assumptions, offering greater flexibility in modeling irregular degradation behaviors (e.g., heavy tails, multimodal trends) with sufficiently large generative deep learning networks, as supported by the Universal Approximation Theorem. To demonstrate this, we compare Conditional Diffusion Models against 'PCA + Lognormal regression' as illustrated in Fig. 2, which projects data into linear subspaces and fits RUL distributions under Gaussian assumptions. Evaluation metrics and data types used for this experiment are the same as in Point 1.

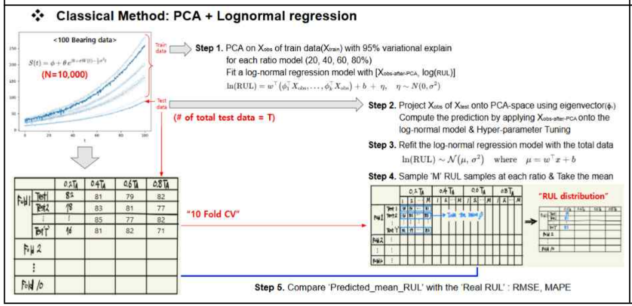
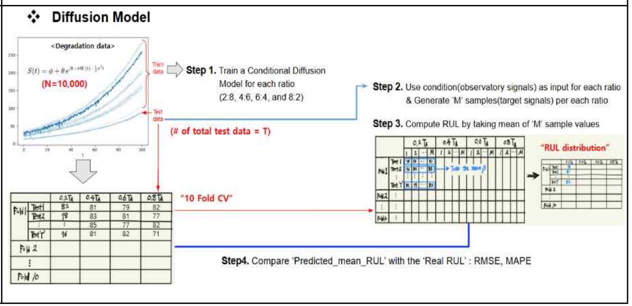
Bayesian Update Models	Diffusion Model
	

Fig. 2. Design of Experiment for Point 2

3. Expected Results

- **Point 1:** For Type 1 data, Bayesian models may perform comparably; However, as data complexity increase (e.g., Type 2), Conditional Diffusion is expected to outperform due to its capacity to model nonlinearity, nonrregularity and complex variations.

- **Point 2:** Diffusion models are expected to exceed ‘PCA + Lognormal regression’, particularly when the true degradation manifold deviates from linear structures and simple distributional forms.

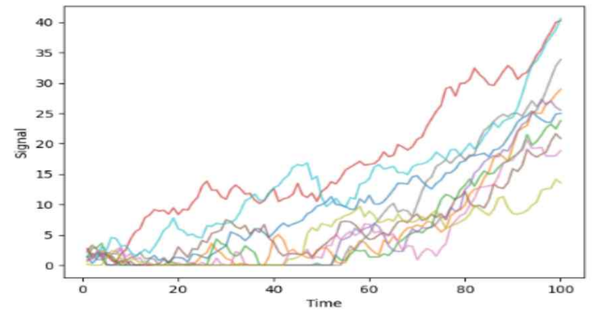
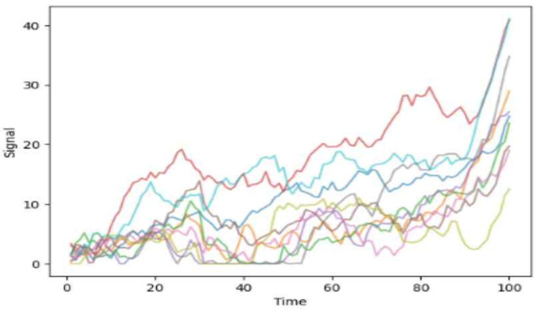
Type 1	Type 2
	

Fig. 3. Data Type

4. Significance and Contribution to AI4SE / SE4AI

This work demonstrates how advanced generative models—specifically Conditional Diffusion Models—can augment the capabilities of system health management frameworks within AI4SE and SE4AI contexts. By moving beyond rigid and parametric modeling, our approach supports the development of more robust, adaptive, and trustworthy prognostic tools, aligning with the workshop’s theme of enabling transformative capabilities in system engineering through AI.

References

[1] Gebraeel, N. Z., Lawley, M. A., Liu, R., & Parmeshwaran, V. (2005). "Residual Life Predictions from Vibration-Based Degradation Signals: A Neural Network Approach." IEEE Transactions on Industrial Electronics.

[2] Jeong, C., Yue, X., & Chung, S. (2025). "Fed-Joint: Joint Modeling of Nonlinear Degradation Signals and Failure Events for Remaining Useful Life Prediction using Federated Learning." arXiv preprint arXiv:2503.13404

[3] Gebraeel, N. Z. (2006). "Sensory-Updated Residual Life Distributions for Components with Exponential Degradation Patterns." IEEE Transactions on Automation Science and Engineering, 3(4), 382–393.

[4] Tang, L., & Che, A. (2019). "Data-Driven Remaining Useful Life Prediction with Degradation Data: A Review and Comparative Study." Applied Sciences, 9(23), 5112

[5] Ho, J., Jain, A., & Abbeel, P. (2020). "Denoising Diffusion Probabilistic Models." Advances in Neural Information Processing Systems.

[6] Tashiro, Y., Gritsenko, A. A., & Ermon, S. (2021). "CSDI: Conditional Score-Based Diffusion Models for Probabilistic Time Series Imputation." arXiv preprint arXiv:2107.03502.

[7] Ho, J., Jain, A., & Abbeel, P. (2022). "Classifier-Free Diffusion Guidance." arXiv preprint arXiv:2207.12598.

[8] Kazerouni, A., & Kochenderfer, M. J. (2023). "Diffusion Models for Time-Series Data Generation and Forecasting." arXiv preprint arXiv:2301.13016.

[9] Tashiro, T. et al. (2021). "Diffusion Models for Time-Series." arXiv preprint.

[10] Qiu, H. et al. (2024). "TS-Diffusion: Generating Highly Complex Time-Series." arXiv preprint.