Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5183-4_S20-02-184-cd

Evaluation of remain useful life prediction models from a resilience perspective

Jie Liu

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: jie.liu@ntnu.no

Yiliu Liu

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: yiliu.liu@ntnu.no

Shen Yin

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: shen.yin@ntnu.no

Jørn Vatn

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Norway. E-mail: jorn.vatn@ntnu.no

The remaining useful life of components is crucial in maintenance planning. However, it is difficult to determine whether or not a model is reliable and trustworthy. There is no clear instruction about how to evaluate a remaining useful life prediction model since the output results are not displayed in the physical system. Furthermore, the models may generate an inaccurate prediction due to missing or outlier data. The idea of resilience is presented to evaluate the performance of the remaining useful life prediction model in order to increase model dependability and robustness. This study offers a definition for the resiliency of remaining useful life prediction models. The study then provides methods for evaluating the models used for the remaining useful life prediction. The work could serve as a springboard for further research into topics such as resilience and life prediction models.

Keywords: Resilience, maintenance, remaining useful life prediction.

1. Introduction

The remaining useful life (RUL) of a component is the amount of time it can be used before it needs to be maintained or replaced, Vaidya and Rausand (2011). The prediction result is crucial for plans and activities related to maintenance. According to search results from the ”Web of Science”, there are about 13,000 publications with the topic of ”remaining useful life” by the end of 2021. Furthermore, the number of articles is rising at a constant rate from 1989 to 2021. The results indicate the significance of RUL and the research interest. Physical model-based methods, statistical methods, and condition monitor data-driven methods are all utilized to predict RUL, Liu et al. (2018). In practice, these algorithms may fail due to missing or outlier input data caused by sensor's failure, communication failure, or hacking. Therefore, methods must be developed to assure the reliability or robustness of these models when there are disruptions caused by input data. The goal is to ensure that the models work properly not just in normal conditions but also when external interruptions occur. Resilience becomes a better word to characterize the situation and evaluate the performance of models than reliability and robustness. Resilience is initially defined as a person's ability to bounce back from adversity, Cyrulnik (2009). Some studies use the term to characterize the ability of a system to bounce back, robustness, graceful extensibility, and architectures for long-term adaptability, Woods (2015). The study of resilience could then help systems to avoid the influence of disruptions and quickly return to a steady-state after interruptions, Gu et al. (2015). Only a few studies have used resilience to evaluate statistical or data-driven models. Vrabic et al. propose an intelligent agent-based architecture to enhance the resilience of digital twins in manufacturing which uses a learning agent to compare the accuracy of digital twins, Vrabič et al. (2021). Clark et al. present an assessment metric for the resilience of the cyber-physical system to due with cyber-attack, Clark and Zonouz (2017). According to the "Web of Science" research results, more than 139,000 publications have been published on resilience by the end of 2021, but only a few of them focus on resilience models. Furthermore, the outcome of RUL prediction is typically a lifetime distribution, which is not always straightforward to compare to the genuine lifespan distribution. The objective of the paper is to bridge the gap between real-world needs and the existing methods. Our research presents a method for evaluating the resilience of RUL prediction models and improving forecast reliability. The proposed method uses the accuracy of prediction as the indicator to evaluate the resilience of RUL prediction models.

The remaining parts of the paper are organized as follows: Section 2 presents variance definitions of resilience in different systems, and frequently used RUL prediction models according to our literature review. Section 3 presents the concept of resilience for RUL prediction. Then it proposes the method to evaluate the models used for remaining useful life prediction. Section 4 gives an example based on proposed methods. Section 5 concludes the paper.

2. Context and background

There are a variety of models that can be used to forecast RUL for various systems. Some of the models are linked to an online system to improve prediction accuracy. The online data, on the other hand, cause external interruptions to the RUL prediction models. Some questions must be answered before definitions for resilience RUL prediction models can be given.

- What is the definition of resilience for other systems?
- What types of models are used to predict RUL?
- Why is resilience introduced to evaluate the models?

2.1. *Definition for resilience of system*

Resilience has been applied to a variety of technological systems, including civil infrastructure systems, Gay and Sinha (2013), energy systems, Wang et al. (2015), and transportation systems, Wan et al. (2018). For different systems, there are multiple definitions. According to Erik Hollnagel, resilience refers to the ability of a system to retain a healthy state over time while being exposed to negative and damaging events, as well as its ability to stay within a safe state during adverse conditions, Hollnagel et al. (2006). The concept of resilience in this paper is based on Erik Hollnagel's work, but it also considers the ability of a system or model to recover from shock or disturbances. The approaches of resilience evaluation include resilience metrics, Cai et al. (2018), quantitative approaches, Hosseini et al. (2019), qualitative approaches, and the resilience framework which combined technical and social factors, Jain et al. (2018). The resilience factor is present throughout the life cycle of a system, which includes the normal or pre-hazard phase, the hazard period, and the recovery period, Cheng et al. (2021). As a result, to improve system resilience, it must first determine how to assess system resilience. The methods could be quantitative, qualitative, or a combination of both approaches. Second, it should look at the complete life cycle of a system to see where it might increase its resilience. The improved points could be the system's ability to survive a disruptive external event. It could also have a high repair capability, allowing it to return to its original state quickly and at a lower cost. The failure mechanism for RUL prediction models differs from other models. Furthermore, evaluating the models is challenging because most of them are based on probability theories, and it is tough to define a standard. As a result, the definition of RUL prediction model's robustness must be given with caution.

2.2. *Models for RUL prediction and sensor errors*

There are several types of RUL prediction models. These are experience-based models, data-driven models, and physics-based models. Hybrid prognostics models, which mix two or more types of models, are also available, Liao and Köttig (2014). All these models are created based on the existing information of the system. The quality of data is critical for data-driven related models to produce high-quality output. With improved sensors and communication technology, it is feasible to improve the prediction results during operation since there is more knowledge about the system condition based on real-time sensor data, Si et al. (2011). However, sensor errors may result in incorrect predictions, which is much worse than offline prediction. There are many researchers working on error detection in sensors. The most common errors for sensors are outliers and missing data. The three most popular error detection methods are principal component analysis, artificial neural network, and Ensemble Classifiers, Teh et al. (2020). To ensure the quality of RUL prediction models, strategies to reduce the negative effects of outliers and missing data must be developed. This condition is the same as the definition of resilience, hence resilience is used to evaluate models and assure the accuracy of predictions.

3. Resilience of RUL prediction

3.1. *Definition for resilience of RUL prediction models*

Resilience is the ability of a system or piece of equipment to rebound from failures. The most likely failure of RUL prediction models is considered as incorrect outcomes from models. It could lead to an overestimation or underestimation of the RUL in a system. To avoid additional system loss, the failure must be discovered and addressed as quickly as possible.

The following definitions of RUL prediction model's resilience are provided based on the prior discussion of resilience in Sections 1 and Section 2. The resilience of a RUL prediction model is defined as the capacity of a model to retain its planned quality on its own. There are a few points that distinguish the definition of resilience from other concepts, as follows:

- **Remark 1** Resilience is not the same as reliability. The probability of a system's success at any time is known as reliability. While resilience refers to the ability of a system to continue its function in the face of a shock or interruption.
- **Remark 2** The terms ”resilience” and ”robustness” are not interchangeable. The ability of a system to withstand disruptions during operation is known as robustness, Wieland and Wallenburg (2012). While resilience refers to the ability of a model to continue its intended function, which includes the ability to withstand disturbance and recover from the shock.
- **Remark 3** It is a continuous process to evaluate the resilience of RUL prediction models.

3.2. *Evaluation of resilience of models*

The output of a RUL prediction model is the distribution of expected failure time. Since the estimated RUL cannot be found in the real system, the method proposed by Vrabic et al. (Vrabič et al. (2021)) cannot be used directly. Therefore, the historical lifetime of a system is used to compare with the prediction result. The statistical distance between these two distributions could be used as the indicator of the performance of a model. If these two distributions are in perfect coincidence, then the prediction result of the model is fulfilled with knowledge of the system which means the performance of the model is good. Otherwise, the accuracy of the model is low. There are many methods to evaluate the distance between two distributions, such as Euclidean distance, Manhattan distance, and standardized Euclidean distance.

Moreover, the result of RUL prediction is used to make maintenance decisions. As a result, a meaningful result is a prediction of the most likely failure time of the subject that is as close to the true value as possible. Because the sum of probability is 1 for both distributions, the RUL of the highest probability parts will be compared. Two factors must be examined in this scenario. The first is the RUL values that are most likely to

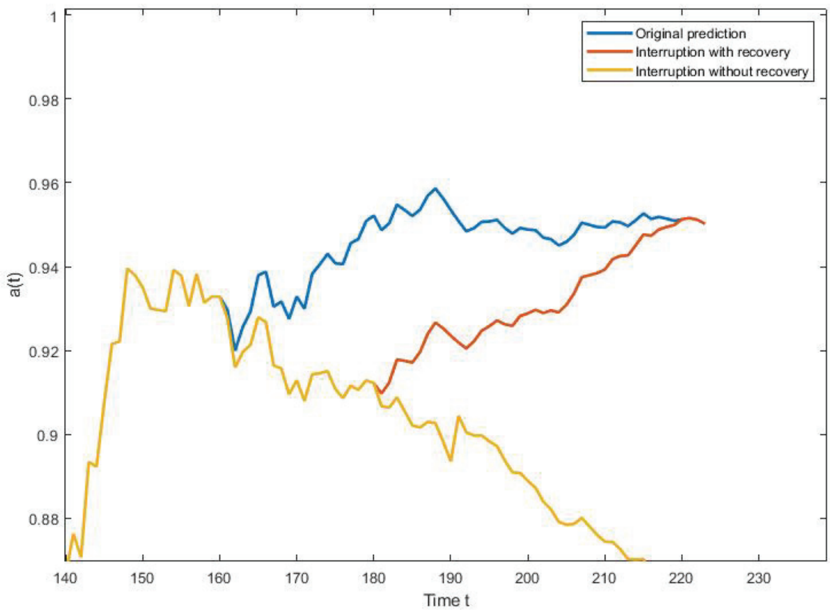

Fig. 1. Resilience of RUL prediction model

occur. The second is the associated probability of the value. The bounds of the most likely values must be explicitly established based on experience. Equation (1) can be used to compute the average RUL.

$$RUL_{avg} = \frac{\sum_{i=1}^{n} Pr_i * RUL_i}{\sum_{i=1}^{n} Pr_i} \qquad (1)$$

where Pr_i presents the probability of considering RUL_i.

Then the accuracy of prediction could be calculated with Equation (2).

$$a(t) = 1 - \left| \frac{RUL_{pre.}(t) - RUL_{hist.}(t)}{RUL_{hist.}(t)} \right| \qquad (2)$$

where $RUL_{pre.}(t)$ is the average RUL calculated according to the prediction of models at time t, $RUL_{hist.}(t)$ is the average RUL calculated according to historical data at time t. The maximum value of the accuracy is 1 and it is possible to be less than 0.

If the prediction is close to the true value, the result of accuracy is close to 100% which means the performance fulfills the historical data. If the prediction is far from the historical data, it means the system is not fitted with the experience. Then there is a risk to use the result from the model without further analysis. For a system with sufficient data and well-developed RUL prediction models, the performance of the model must be stable and keep a high level. Hence, if there is a shock to the system, the performance will decrease. The shock might be caused by the working load changing in a real system or by external disruption. However, if the performance could not return to normal states automatically, it means the accuracy of the model is decreased. The resilience of the model could be calculated according to the loss of accuracy which is the area between 1 and the loss of accuracy of the model.

Figure 1 is schematic diagram of resilience evaluation for RUL prediction. The blue line represents the accuracy of the original prediction, the yellow line represents the accuracy of the prediction, which is disrupted without recovering, and the red line represents the accuracy of the prediction that recovers from a disturbance. As shown in the graph, the original prediction accuracy is growing over time. This is because there is more data regarding the degradation entering the model as the condition monitoring continues. The prediction results may be more accurate than they were at the start. When the system is disrupted, the model is misled by the new data that arrives. Thus, the performance of prediction is decreased. If no more steps are done, the accuracy of the model will not recover or will recover extremely slowly to the normal state. Hence, inaccurate results may lead to incorrect maintenance decisions, resulting in increased maintenance costs. When considering the resilience of RUL prediction, it is easy to see when the prediction results are from knowledge. Then steps could be taken to improve the accuracy of prediction to a standard level, preventing further losses because of poor decisions.

In comparison to traditional methods, the above method considers more information of the prediction results and past deterioration knowledge. It could conduct a comprehensive review of model performance and then encourage consideration of disruptions in the design phase of the RUL prediction models. For RUL prediction models, the technique could reduce the impact of missing and outlier data.

4. Case study

We consider a system with continuous system monitoring and the predicted RUL is used for maintenance decisions. The indicator which presents the degradation level has been selected according to experience. The system starts from a perfect state and the performance is reduced during operation with increased indicators. Assuming the degradation process follows an exponential distribution. In addition, there is noise following a normal distribution with parameters μ and σ. The degradation process is present as Equation (3)

$$Y(t) = e^{\lambda t} + \varepsilon(t) \tag{3}$$

where λ is the parameter of exponential distribution and noise $\varepsilon(t)$ follows normal distribution $\mathbb{N}(\mu, \sigma^2)$.

Assume parameters $\lambda = 0.01$, $\mu = 0.07$, $\sigma = 0.6$, and threshold $L = 10$. Then the probability density function (pdf) of system lifetime could be obtained by simulating the degradation process. Repeating simulation by 10,000 time, the pdf of system failure time is presented as Figure 2.

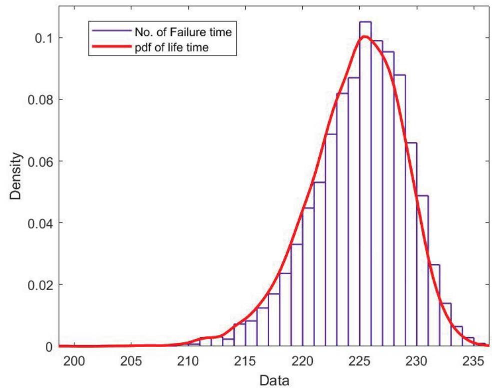

Fig. 2. Simulated historical of lifetime

On the other hand, a RUL prediction model is tested with previously generated data according to the evaluation method mentioned in Section 3.2. The model is provided by MATLAB - Predictive Maintenance Toolbox release R2021b The Mathworks, Inc. (2021). One group of simulated data is picked and used in the MATLAB toolbox. The prediction result of the last data set is shown as Figure 3. The upper part of the figure shows the real health indicators from the beginning until the current state compared with the current degradation model. The lower part of the figure shows the distribution of the last day's prediction with the confidence interval.

To simplify the calculation, only the highest probability of RUL is used to calculate the accuracy of prediction. According to Equation (1) and (2), the plot of accuracy is calculated and shown as the blue line in Figure 4. Data set before 100 is used as training data so they are not considered here. From the figure, the accuracy of the model fluctuates after the training process and becomes stable after data set 150. It means the accuracy of prediction is stable when there is enough data inputted in the model. In addition, the model improves the accuracy when there is more data and

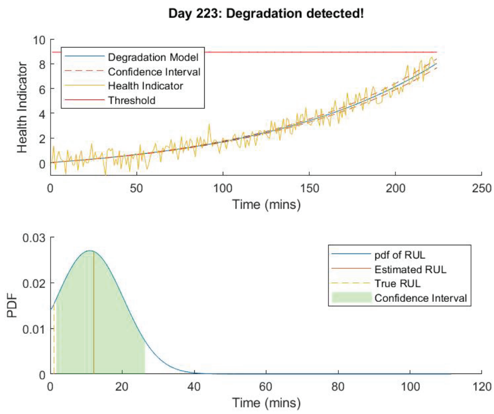

Fig. 3. Prediction result of the last data set

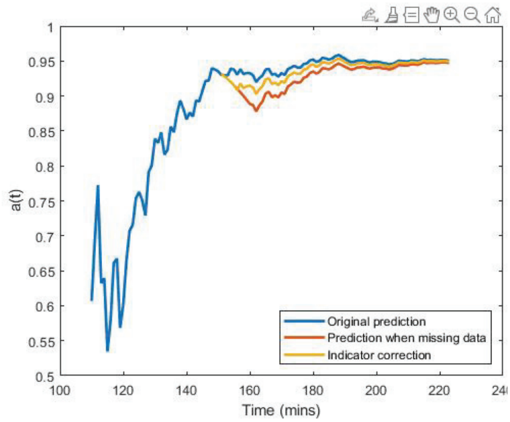

Fig. 4. Resilience of system during shock

keeps the accuracy level around 95% when it is stable.

Now, assume there is a shock to the system which makes the system missing data sets from 151 to 160. Thus, the system repeats the previous data which is 150 to make predictions until the real data enters. The accuracy of the model decreased. If there is no action to the model, the accuracy of the model could recover by itself slowly as shown with the red line of Figure 4. However, if the abnormal data is noticed, then the indicator could be corrected on time and help the accuracy back to a normal state earlier. The correction process is shown as the yellow line in Figure 4. The resilience of the model could be obtained by calculating the area between the accuracy curve and 100% accuracy. From the figure, the comparison resilience result of these three situations is the resilience of the original situation is larger than the resilience of indicator correction after shock, and larger than resilience with the shock of indicators if the resilience is sorted by size.

Therefore, the resilience of the model could be evaluated with the accuracy of the prediction of RUL. The accuracy is increased after shock by itself. But if the reduced performance could be noticed earlier, it is possible to increase resilience and improve the performance of the system.

5. Conclusions

In this paper, the concept of resilience is introduced for RUL prediction models. It describes the current state of the art in resilience and RUL prediction models and proposes a definition of RUL prediction model resilience, based on the literature review. Methods for evaluating the resilience of RUL prediction models are proposed and investigated. Finally, the accuracy of prediction could be used to assess the resilience of RUL prediction models. The method requires sufficient knowledge about the remaining useful life and an online monitor of resilience. When operational resilience is reduced, an analysis must be conducted to identify if a system failure has occurred. If this is not the case, then data interpolation can be employed to assist RUL prediction models in maintaining their intended function.

6. Acknowledgement

This paper has been written under the Norwegian Center for Research based Innovation on Subsea Production and Processing (SUBPRO, project 9001 8912). The authors would greatly acknowledge for helpful suggestions from anonymous reviewers.

References

Cai, B., M. Xie, Y. Liu, Y. Liu, and Q. Feng (2018). Availability-based engineering resilience metric and its corresponding evaluation methodology. *Reliability Engineering & System Safety 172*, 216–224.

Cheng, Y., E. A. Elsayed, and Z. Huang (2021). Systems resilience assessments: a review, framework and metrics. *International Journal of Production Research*, 1–28.

Clark, A. and S. Zonouz (2017). Cyber-physical resilience: Definition and assessment metric. *IEEE Transactions on Smart Grid 10*(2), 1671–1684.

Cyrulnik, B. (2009). *Resilience: How your inner strength can set you free from the past*. Penguin UK.

Gay, L. F. and S. K. Sinha (2013). Resilience of civil infrastructure systems: literature review for improved asset management. *International Journal of Critical Infrastructures 9*(4), 330–350.

Gu, X., X. Jin, J. Ni, and Y. Koren (2015). Manufacturing system design for resilience. *Procedia Cirp 36*, 135–140.

Hollnagel, E., D. D. Woods, and N. Leveson (2006). *Resilience engineering: Concepts and precepts*. Ashgate Publishing, Ltd.

Hosseini, S., D. Ivanov, and A. Dolgui (2019). Review of quantitative methods for supply chain resilience analysis. *Transportation Research Part E: Logistics and Transportation Review 125*, 285–307.

Jain, P., H. J. Pasman, S. Waldram, E. Pistikopoulos, and M. S. Mannan (2018). Process resilience analysis framework (praf): A systems approach for improved risk and safety management. *Journal of Loss Prevention in the Process Industries 53*, 61–73.

Liao, L. and F. Köttig (2014). Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Transactions on Reliability 63*(1), 191–207.

Liu, H., Z. Mo, H. Zhang, X. Zeng, J. Wang, and Q. Miao (2018). Investigation on rolling bearing remaining useful life prediction: A review. In *2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, pp. 979–984. IEEE.

Si, X.-S., W. Wang, C.-H. Hu, and D.-H. Zhou (2011). Remaining useful life estimation–a review on the statistical data driven approaches. *European journal of operational research 213*(1), 1–14.

Teh, H. Y., A. W. Kempa-Liehr, I. Kevin, and K. Wang (2020). Sensor data quality: A systematic review. *Journal of Big Data 7*(1), 1–49.

The Mathworks, Inc. (2021). *MATLAB version (R2021b)*. Natick, Massachusetts: The Mathworks, Inc.

Vaidya, P. and M. Rausand (2011). Remaining useful life, technical health, and life extension. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 225*(2), 219–231.

Vrabič, R., J. Erkoyuncu, M. Farsi, and D. Ariansyah (2021). An intelligent agent-based architecture for resilient digital twins in manufacturing. *CIRP annals 70*(1), 349–352.

Wan, C., Z. Yang, D. Zhang, X. Yan, and S. Fan (2018). Resilience in transportation systems: a systematic review and future directions. *Transport reviews 38*(4), 479–498.

Wang, Y., C. Chen, J. Wang, and R. Baldick (2015). Research on resilience of power systems under natural disasters—a review. *IEEE Transactions on Power Systems 31*(2), 1604–1613.

Wieland, A. and C. M. Wallenburg (2012). Dealing with supply chain risks: Linking risk management practices and strategies to performance. *International journal of physical distribution & logistics management*.

Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety 141*, 5–9.