

Remaining Useful Life Estimation on Sensor Data

National Research University Higher School of Economics

Faculty of Computer Science

Programme ‘Master of Data Science’

Mikhail Zaitsev

Moscow, 2023

Abstract

Abstract goes here

Contents

1	Overview of Remaining Useful Life Models	3
1.1	RUL prediction approaches classification	3
1.2	Physics-based approaches	3
1.3	Hybrid approaches	5
1.4	Data driven approaches	5
	References	8

Chapter 1

Overview of Remaining Useful Life Models

1.1 RUL prediction approaches classification

Based on the methods and strategies employed to create a prediction model, remaining useful life (RUL) prediction approaches can be divided into different categories. Analytical-based, model-based, knowledge-based, and hybrid-based techniques could all be distinguished as four separate strategies [1]. Further classification could be done by the methodology of the approach taking into account the technique being used.

So, the class of models based on physical failure modeling technique is represented by analytical-based RUL prediction approaches. The technique of calculating the rough physical model of the understudied process or the model of deterioration propagation is referred to as physical failure modeling.

Model-based RUL Prediction refers to statistical and computational methods. In order to estimate the parameters of the underlying mathematical models of the deterioration process, such methods frequently utilise historical run-to-failure data.

Computational techniques are combined with domain knowledge to create knowledge-based methods. where the model logic is determined by expert knowledge of the process behavior and heuristic rulesets. Hybrid techniques combine any of the aforementioned techniques.

As research in this area has grown, new techniques and combinations of techniques have begun to be used, making it difficult to categorize them into separate categories. As a result, the previous classification can now be further broken down into three main approaches: physics-based approaches, data-driven approaches, and hybrid approaches [2].

Table 1.1 lists the benefits and drawbacks of the aforementioned strategies.

1.2 Physics-based approaches

Published application of physics-based approaches for RUL prediction is limited to degradation models of relatively simple components [3, 4, 5, 6]. Even for such relatively simple components, underlying physical-based models are quite complicated. Thus, despite such advantages of physical-based models as high accuracy,

Prognostic Approaches	Advantages	Drawbacks
Physics-based approaches	<ul style="list-style-type: none"> • High Accuracy • High Interpretability • Estimation at variable time increments steps • Extrapolation outside training data 	<ul style="list-style-type: none"> • Too domain specific, limited transferability • Vast domain expertise needed to build a model • Require assessment of simplifying assumption • Becomes too stochastic if the modeled process is complicated • May require empirical estimation of parameters
Data-driven approaches	<ul style="list-style-type: none"> • Do not require vast domain knowledge • Fast implementation • Can be transferred and applied for other domains • Do not require empirical estimation of parameters 	<ul style="list-style-type: none"> • Require a large amount of data • Usually difficult to interpret • Models do not incorporate knowledge of the physics of the system • Very dependent on training strategies due to overfitting and overgeneralization • Less accurate on longer prediction horizons
Hybrid approaches	<ul style="list-style-type: none"> • Eliminates some drawbacks of one method by applying different approach to the problem • Reduce amount of data needed to train models 	<ul style="list-style-type: none"> • Physics-based model is still required • Inherits drawbacks of both methods to some extent

Table 1.1: Advantages and drawbacks of prognostic approaches.

interpretability and moderate amount of data requirement their application to complex systems is limited.

There is published application of physical-based approach to model turbofan engine degradation via Wiener process [7], however the dataset used in this paper is unique, which makes it hard to compare the results with the results obtained on

C-MAPSS dataset [8].

1.3 Hybrid approaches

Despite that both physics-based and data-driven approaches have their drawbacks when applied independently, the combined use of both approaches might potentially lead to performance gains by utilizing advantages of both approaches. Specifically, physics-based approaches are limited by underlying model incompleteness and difficulties of estimating parameters of the models if the modeled process is relatively complicated. On the other hand, physics-based models do not require large amount of data, have high interpretability and can be used to extrapolate training or generate synthetic data. Data-driven approaches, in contrast, require representative training data to train but can discover complex patterns if training data is representative and large. Data-driven approaches are usually difficult to interpret. Thus, data-driven approaches can potentially be used in combination with physics-based approaches to enhance the latter.

Different approaches have been proposed to combine physics-based and data-driven approaches. One of the approaches, described in [9], uses data-driven model, namely Support Vector Machine (SVM) and Support Vector Regression (SVR) to classify cutting tool wear stage and predict observed value of tool wear from input sensor signals, then the classification results of the tool wear stage are used to determine the prior distribution of the wear model, and the observed values of the tool wear are applied to update the physics-based model parameters. Another approach [10] uses Wiener process physical-based model to generate additional training data which is then utilized together with initial sensors data as input of data-driven Long Short-Time Memory (LSTM) bidirectional neural-network model to enhance RUL prediction. In this work [11] on aero gas turbine engine fault classification, the outputs of physical-based model, namely, health parameters are passed to data-driven model together with initial sensor data as features for Decision Tree Classifier which performs fault classification. Here [12], the authors are proposing the following framework: one Deep Neural Network (DNN) to approximate the physics-based model and the second prognostics DNN used for predicting RUL, based on combined input. The inputs for second prognostics DNN are sensor inputs and the outputs and parameters of approximated physics-based system model.

1.4 Data driven approaches

In recent years, data-driven approaches have been widely applied to RUL prediction. The increased availability of training data, no prior knowledge of modeled physical systems requirements and fast implementation made them a popular choice in recent works. Among them are classical machine learning methods, such as SVR, Decision Trees and Random Forest [13, 14], which showed good results on predicting RUL.

With the growing development of machine learning techniques and availability of computational resources, neural networks became more and more popular due to their ability to automatically construct features. Among which, Convolutional Neural Networks (CNN) were the first architectures to make a notable advancements

due to the fact that CNN can learn local, spatial structure within data. In this work [15] CNN architecture consisting of two layers of 1D convolutions following 1D pooling layer and fully connected final layer was applied to segmented multi-variate time series input data. The results obtained in this research proved that CNN approach can outperform classical machine learning methods. Later, [16] introduced slightly different CNN architecture, consisting of five 2D convolutional layers, each except the last with ten 1D kernels, which got even higher results. Such architecture was aimed to capture longer temporal structure due to increased depth of CNN and thus bigger receptive field.

However, CNNs are not designed to capture long temporal dependencies in data, to obtain a relatively large receptive field CNN should be deep. The second drawback of CNN in relation to RUL prognostics is that temporal dependencies learned by CNN are not causal, meaning that both previous and following time steps are available for CNN kernel at any particular time step. Due to this fact, Recurrent Neural Networks (RNN) showed better performance than CNNs. RNNs were designed to capture long temporal dependencies by the ability of remembering and updating information about previous time steps in the hidden state. This work [17] proposed LSTM model for RUL estimation which outperforms both classical machine learning methods as well as CNNs. As further development this paper suggests using CNN as feature extractor before LSTM to reduce noise and frequency of input features. Another work [18] enhances LSTM approach by proposing bidirectional LSTM to extract more information from sequential sensor data.

Recent developments in Graph Neural Networks (GNN) have been reflected in their application to RUL prediction tasks. In contrary to sequential and hierarchical structure of common neural network architectures, GNN allows more flexible and complex relationships between NN structures. In this study [19] Gated Graph Neural Network (GGNN) [20] was proposed in which subsets of features were represented as graph nodes, and each edge models causal relations between those subsets. In this way the model was capturing the complex structure of the equipment.

Most deep learning architectures treat each feature vector with equal importance, not considering the extent by which this vector contribute to predicted values. In order to focus on the features which contribute the most to the prediction the Attention mechanism was introduced and widely applied to different domains, such as natural language translation, computer vision many other fields of study.

The Attention mechanism was widely adopted and combined with many existing RUL prediction architectures as well as used as stand alone approach. In [21] the attention mechanism was introduced which weights LSTM outputs of each time step and combines them with final step output. The LSTM model with introduced Attention mechanism outperforms pure LSTM model. In the following work [22], on contrary, the Attention mechanism was applied to the input features to weight each feature according learned importance to enhance representativeness of input data.

The success of the attention-based models led to its further development and appearance of more advanced self-attention and multi-head attention mechanisms and architectures based mainly on them such as Transformers. This work [23] shows that sequentially applying Multi-Head Attention on features and then on sequences before LSTM layer can produce outstanding results by enhancing information retrieval from input features. Another study [24] obtained similar results by applying

modified Transformer architecture to multi-variate time series. In particular, the modified part was Encoder block of the Transformer, sensor encoder and time step encoder of the same architecture, processing inputs in parallel were introduced as part of this block to extract temporal and spatial features. After extracting, both features are merged together to represent information from both aspects.

The growing demand in NN architecture which can capture long temporal dependencies without drawbacks of RNNs, such as computational complexity due to sequential input processing and vanishing gradients results in invention of Temporal Convolutional Networks [25]. TCN consists of dilated, causal convolutions, meaning that no information leaks from past time steps to future. The same paper showed that TCNs outperforms RNNs across a broad variety of sequence modeling tasks. Application of TCN to RUL prediction is described in [26], a proposed consists of spatial and temporal attention blocs and TCN as feature extractors. Another work [27] uses TCN in parallel with bidirectional Gated Recurrent Unit (GRU) after attention block as feature extractor. As further development of the TCN ideas, Sample Convolution and Interaction Network (SCINet) was introduced to address some of TCN limitations. The application of SCINet together with LSTM feature extractors is described in [28].

Another approach is based on modified Variational Autoencoder (VAE), the idea behind this approach is to use the latent representation of encoded input sensor signals for RUL prediction. In this research [29] authors substitute the decoder part of VAE by Regressor, the loss function was modified accordingly. This forced the latent representation of input signal to be representation of the system rate of deterioration. The benefit of this approach is that not only the model is highly accurate but also the latent space obtained from input signals is interpretable.

Table 1.2 shows the results comparison of different methods on C-MAPSS dataset.

Methods	Year	FD001		FD002		FD003		FD004	
		RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
CNN [15]	2016	18.448	1286	30.294	13570	19.817	1596	29.156	7886
LSTM [17]	2017	16.14	338	24.49	4450	16.18	852	28.17	5550
DCNN [16]	2018	12.61	274	22.36	10412	12.64	284	23.31	12466
Attn-DLSTM [21]	2019	12.22	-	23.16	-	15.61	-	28.30	-
GNMR [19]	2020	12.14	212	20.85	3196	13.23	370	21.34	2795
AGCNN [22]	2021	12.42	226	19.43	1492	13.39	227	21.50	3393
DA-TCN [26]	2021	11.78	229	16.95	1842	11.56	257	18.23	2317
RVE [29]	2022	13.42	323	14.92	1379	12.51	256	16.37	1845
MDSA [23]	2022	11.43	209	13.32	1058	11.47	187	14.38	1618
DAST [24]	2022	11.43	203	15.25	924	11.32	155	18.36	1490
TCN-BiGRU [27]	2022	11.17	207	14.06	842	10.42	207	16.34	1498
Time-feature interaction model [28]	2023	11.59	208	-	-	10.9	187	-	-

Table 1.2: Results comparison of different methods on C-MAPSS dataset.

References

- [1] C. Okoh et al. “Overview of Remaining Useful Life Prediction Techniques in Through-life Engineering Services”. In: *Procedia CIRP* 16 (2014). Product Services Systems and Value Creation. Proceedings of the 6th CIRP Conference on Industrial Product-Service Systems, pp. 158–163. ISSN: 2212-8271. DOI: <https://doi.org/10.1016/j.procir.2014.02.006>. URL: <https://www.sciencedirect.com/science/article/pii/S2212827114001140>.
- [2] Youdao Wang, Yifan Zhao, and Sri Addepalli. “Remaining Useful Life Prediction using Deep Learning Approaches: A Review”. In: *Procedia Manufacturing* 49 (2020). Proceedings of the 8th International Conference on Through-Life Engineering Services – TESConf 2019, pp. 81–88. ISSN: 2351-9789. DOI: <https://doi.org/10.1016/j.promfg.2020.06.015>. URL: <https://www.sciencedirect.com/science/article/pii/S2351978920316528>.
- [3] Nathan Bolander et al. “Physics-based remaining useful life prediction for aircraft engine bearing prognosis”. In: *Annual Conference of the PHM Society* 1 (2009), p. 1. DOI: <https://doi.org/10.1007/s00170-022-09417-4>. URL: <https://papers.phmsociety.org/index.php/phmconf/article/view/1631>.
- [4] Yu Hui Lui et al. “Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction”. In: *Journal of Power Sources* 485 (2021), pp. 229–327. ISSN: 0378-7753. DOI: <https://doi.org/10.1016/j.jpowsour.2020.229327>. URL: <https://www.sciencedirect.com/science/article/pii/S0378775320316153>.
- [5] Matthew Daigle, Chetan Kulkarni, and George Gorospe. “Application of Model-based Prognostics to a Pneumatic Valves Testbed”. In: *IEEE Aerospace Conference Proceedings* (Mar. 2014). DOI: 10.1109/AERO.2014.6836319.
- [6] Matthew Daigle and Kai Goebel. “Model-Based Prognostics With Concurrent Damage Progression Processes”. In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43 (May 2013), pp. 535–546. DOI: 10.1109/TSMCA.2012.2207109.
- [7] Junqiang Liu and Jiwei Xie. “Remaining Useful Lifetime Estimation for Aero-Engines Based on Wiener-Process with Integrated Lifetime Data and Degradation Data”. In: (2017), pp. 293–298. DOI: 10.1109/SDPC.2017.63.
- [8] Abhinav Saxena et al. “Damage propagation modeling for aircraft engine run-to-failure simulation”. In: *2008 International Conference on Prognostics and Health Management*. 2008, pp. 1–9. DOI: 10.1109/PHM.2008.4711414.

- [9] Yifan Li et al. “A hybrid remaining useful life prediction method for cutting tool considering the wear state”. In: *The International Journal of Advanced Manufacturing Technology* 121 (2022), pp. 2583–3596. ISSN: 1433-3015. DOI: <https://doi.org/10.1007/s00170-022-09417-4>. URL: <https://doi.org/10.1007/s00170-022-09417-4>.
- [10] Xiaolei Liu et al. “A Hybrid Method of Remaining Useful Life Prediction for Aircraft Auxiliary Power Unit”. In: *IEEE Sensors Journal* 20.14 (2020), pp. 7848–7858. DOI: 10.1109/JSEN.2020.2979797.
- [11] Sangjo Kim et al. “Diagnostics using a physics-based engine model in aero gas turbine engine verification tests”. In: *Aerospace Science and Technology* 133 (2023), p. 108102. ISSN: 1270-9638. DOI: <https://doi.org/10.1016/j.ast.2022.108102>. URL: <https://www.sciencedirect.com/science/article/pii/S1270963822007763>.
- [12] Manuel Arias Chao et al. *Fusing Physics-based and Deep Learning Models for Prognostics*. 2020. DOI: 10.48550/arxiv.2003.00732. URL: <https://arxiv.org/abs/2003.00732>.
- [13] Xiao-lin Wang et al. “A SVR-Based Remaining Life Prediction for Rolling Element Bearings”. In: *Journal of Failure Analysis and Prevention* 15 (2015), pp. 548–554. ISSN: 1864-1245. DOI: 10.1007/s11668-015-9976-x. URL: <https://doi.org/10.1007/s11668-015-9976-x>.
- [14] Sangram Patil et al. “Remaining Useful Life (RUL) Prediction of Rolling Element Bearing Using Random Forest and Gradient Boosting Technique”. In: *ASME International Mechanical Engineering Congress and Exposition Volume 13: Design, Reliability, Safety, and Risk* (Nov. 2018). V013T05A019. DOI: 10.1115/IMECE2018-87623. eprint: <https://asmedigitalcollection.asme.org/IMECE/proceedings-pdf/IMECE2018/52187/V013T05A019/2504203/v013t05a019-imece2018-87623.pdf>. URL: <https://doi.org/10.1115/IMECE2018-87623>.
- [15] Giduthuri Sateesh Babu, Peilin Zhao, and Xiao-Li Li. “Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life”. In: *Database Systems for Advanced Applications*. Ed. by Shamkant B. Navathe et al. Cham: Springer International Publishing, 2016, pp. 214–228. ISBN: 978-3-319-32025-0.
- [16] Xiang Li, Qian Ding, and Jian-Qiao Sun. “Remaining useful life estimation in prognostics using deep convolution neural networks”. In: *Reliability Engineering and System Safety* 172 (2018), pp. 1–11. ISSN: 0951-8320. DOI: <https://doi.org/10.1016/j.ress.2017.11.021>. URL: <https://www.sciencedirect.com/science/article/pii/S0951832017307779>.
- [17] Shuai Zheng et al. “Long Short-Term Memory Network for Remaining Useful Life estimation”. In: *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*. 2017, pp. 88–95. DOI: 10.1109/ICPHM.2017.7998311.

- [18] Ahmed Elsheikh, Soumaya Yacout, and Mohamed-Salah Ouali. “Bidirectional handshaking LSTM for remaining useful life prediction”. In: *Neurocomputing* 323 (2019), pp. 148–156. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2018.09.076>. URL: <https://www.sciencedirect.com/science/article/pii/S0925231218311573>.
- [19] Jyoti Narwariya et al. *Graph Neural Networks for Leveraging Industrial Equipment Structure: An application to Remaining Useful Life Estimation*. 2020. DOI: 10.48550/arxiv.2006.16556. URL: <https://arxiv.org/abs/2006.16556>.
- [20] Yujia Li et al. *Gated Graph Sequence Neural Networks*. 2015. DOI: 10.48550/arxiv.1511.05493. URL: <https://arxiv.org/abs/1511.05493>.
- [21] Ankit Das et al. “Deep Recurrent Architecture with Attention for Remaining Useful Life Estimation”. In: *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*. 2019, pp. 2093–2098. DOI: 10.1109/TENCON.2019.8929267.
- [22] Hui Liu et al. “Remaining Useful Life Prediction Using a Novel Feature-Attention-Based End-to-End Approach”. In: *IEEE Transactions on Industrial Informatics* 17.2 (2021), pp. 1197–1207. DOI: 10.1109/TII.2020.2983760.
- [23] Zhi Lai et al. *Multi-Dimensional Self Attention based Approach for Remaining Useful Life Estimation*. 2022. DOI: 10.48550/ARXIV.2212.05772. URL: <https://arxiv.org/abs/2212.05772>.
- [24] Zhizheng Zhang, Wen Song, and Qiqiang Li. “Dual-Aspect Self-Attention Based on Transformer for Remaining Useful Life Prediction”. In: *IEEE Transactions on Instrumentation and Measurement* 71 (2022), pp. 1–11. DOI: 10.1109/tim.2022.3160561. URL: <https://doi.org/10.1109/tim.2022.3160561>.
- [25] Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. *An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling*. 2018. DOI: 10.48550/ARXIV.1803.01271. URL: <https://arxiv.org/abs/1803.01271>.
- [26] Yan Song et al. “Distributed Attention-Based Temporal Convolutional Network for Remaining Useful Life Prediction”. In: *IEEE Internet of Things Journal* 8.12 (2021), pp. 9594–9602. ISSN: 2327-4662. DOI: 10.1109/JIOT.2020.3004452.
- [27] Ran Gong, Jinxiao Li, and Chenlin Wang. “Remaining Useful Life Prediction Based on Multisensor Fusion and Attention TCN-BiGRU Model”. In: *IEEE Sensors Journal* 22.21 (2022), pp. 21101–21110. DOI: 10.1109/JSEN.2022.3208753.
- [28] Dmitry Zhevnenko, Mikhail Kazantsev, and Ilya Makarov. *Interaction models for remaining useful life estimation*. 2023. DOI: 10.48550/ARXIV.2301.05029. URL: <https://arxiv.org/abs/2301.05029>.
- [29] Nahuel Costa and Luciano Sánchez. “Variational encoding approach for interpretable assessment of remaining useful life estimation”. In: *Reliability Engineering and System Safety* 222 (2022), p. 108353. ISSN: 0951-8320. DOI: <https://doi.org/10.1016/j.res.2022.108353>. URL: <https://www.sciencedirect.com/science/article/pii/S0951832022000321>.