



HAL
open science

Adaptative data-driven approach for the remaining useful life estimation when few historical degradation sequences are available

Koceila Abid, Moamar Sayed-Mouchaweh, Laurence Cornez

► To cite this version:

Koceila Abid, Moamar Sayed-Mouchaweh, Laurence Cornez. Adaptative data-driven approach for the remaining useful life estimation when few historical degradation sequences are available. ICMLA 2020 - 19th IEEE International Conference on Machine Learning and Applications, Dec 2020, Miami, United States. pp.1145-1152, 10.1109/ICMLA51294.2020.00184 . cea-04559664

HAL Id: cea-04559664

<https://cea.hal.science/cea-04559664>

Submitted on 25 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptive Data-driven Approach for the Remaining Useful Life Estimation when Few Historical Degradation Sequences are Available

Koceila Abid
CEA Tech Hauts de France
Lille, France
koceila.abid@cea.fr

Moamar Sayed-Mouchaweh
IMT Lille Douai
Lille, France
moamar.sayed-mouchaweh@imt-lille-douai.fr

Laurence Cornez
CEA, LIST
Paris, France
laurence.cornez@cea.fr

Abstract—Fault prognostics is the estimation of the Remaining Useful Life (RUL) of a component until failure. It is a main part of the predictive maintenance strategy that can help to enhance the reliability and availability of industrial systems while reducing unscheduled downtime and maintenance cost. Applying fault prognostics to industrial systems is difficult because very few sequences about degradation are available. Hence, this paper proposes a new approach able to perform the fault prognostics when few historical degradation data are available. In offline, a library of Health Indicators (HIs) and library of adaptive models are defined. Then, the best HI-model pairs according to the accuracy of the a priori sequences are selected. In online, several RULs are predicted for the new sequence using the selected pairs. After that, the final RUL is computed by merging the RULs using a weighted mean. The approach is validated using a degradation scenario selected from an aircraft engine degradation dataset (C-MAPSS dataset). The obtained results are promising compared to some well-known similar state-of-the-art approaches.

Index Terms—Health Indicator (HI), Adaptive model, Remaining Useful Life (RUL)

I. INTRODUCTION

Major economic loss in industrial companies is due to the downtime of machinery and shutdown of production that result from systems failure. The economical loss for an unexpected one day stoppage in industry may cost until 100,000 € for power plant (coal) [1], 200,000 € for Pulp and Paper [1], and can reach up to 300,000 € for nuclear plant [2]. Traditional corrective maintenance of critical systems can cause a long period of downtime. Hence, applying Prognostics and Health Management (PHM) is very efficient. PHM is a maintenance strategy that can anticipate failures and avoid systems downtime. The failures can be anticipated by applying fault prognostics, which is the estimation of the current health state of a system and predicting its Remaining Useful Life (RUL), where the RUL is the time between the present time and the failure time.

RUL could be estimated based on two main approaches: experience based approaches (reliability and similarity) and degradation modelling based approaches (Model based and data-driven based modelling) [3]. Experience based approaches require a huge amount of run-to-failure sequences for RUL estimation, which is very costly and often not available

(rarely available) in industrial systems. Degradation modelling approaches are based on two methods (model based and data-driven based approaches). Model based approaches require a physical or mathematical model about the system operation. These methods are prohibitively costly to develop in real complex systems, due to the complex interactions between the different components of the system and the variations in the environmental conditions. Data-driven approaches require several historical degradation data for training the model to predict the RUL. Regardless of the interpretation difficulty of data-driven methods, they are the most suitable for prognostics of complex systems in terms of accuracy, applicability, and cost [3].

In data-driven approaches, the RUL estimation procedure starts by constructing a Health Indicator (HI). The computed HI can show the state of the system (performance or condition) and its evolution over time. It can be based on raw measurements collected by sensors, or extracted features from the measurements (i.e., time domain, frequency domain, or time-frequency domain features), or it can also be a merge of several features (e.g., computing euclidean distance between normal and evolving class in the feature space [4]). Choosing a suitable HI can enhance the system health monitoring and RUL estimation. A good HI for prognostics can be chosen based on domain experience or it can be selected automatically using some criteria. J. Coble [5] proposed two metrics that are frequently used for HI selection called monotonicity and trendability. Monotonicity evaluates the increasing or decreasing trend of the HIs, while the trendability measures the correlation between HIs and the operating time. In [6], spectral kurtosis is selected as the best HI by using monotonicity and trendability metrics. In [7], a HI selection score is computed by using the two previous metrics while the HIs with a score higher than a threshold are merged using neural networks. The same metrics are used in [8] for HI selection among classical and other trigonometric features (e.g., asinh and atan). The selection of the best HI (or HIs) requires some a priori run-to-failure sequences (historical degradation sequences) in order to select the best HI. When the HI is selected and computed for the new measurement, it is extrapolated until failure threshold

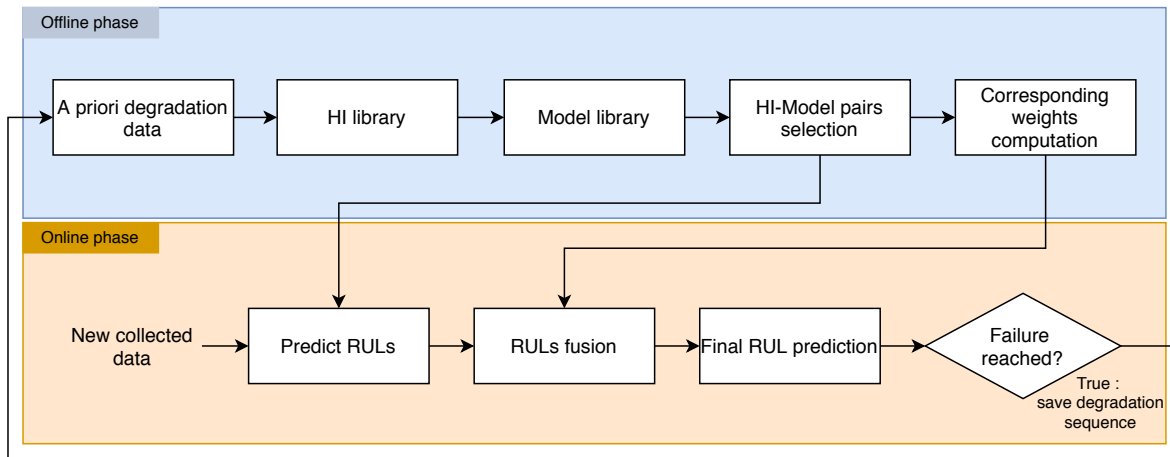


Fig. 1: Flowchart of the proposed prognostic approach

in order to compute the failure time, then, the RUL is deduced by the time between failure time and present time. The HI extrapolation can be achieved using statistical or machine learning techniques. Statistical methods are widely used to extrapolate the health indicator over time. In [9], Auto Regressive Moving Average model (ARMA) is used to extrapolate the HI of an elevator door motion system. For health monitoring of bearings, many works fit models following an exponential trend to predict the RUL, because degradation evolution of bearings presents an increasing exponential trend. In [10], the HI is fitted by an exponential model for RUL estimation of bearing. The double exponential model is the summation of two exponential models, which is used to extrapolate the computed HI in [7] [11]. Generalized Linear Model with a log link function is used in [12] for RUL estimation of a wind turbine shaft bearing. In [13], quadratic regression is applied to fit the HI for RUL estimation of bearings. Machine learning methods attempt to learn the degradation evolution of the system by using the available a priori sequences about the degradation. Support Vector Regression (SVR) is a machine learning technique that is widely used for RUL estimation. SVR is used in [14], [15] to predict a health indicator evolution in order to estimate bearing's RUL. In [6], SVR is used to extrapolate the selected HI to estimate the RUL of the high speed shaft bearing of a wind turbine. Artificial Neural Network (ANN) are also used for RUL estimation, a life percentage (life percentage is the inverse of RUL) estimation method for bearing is developed in [16] based on ANN. An ANN based method is developed in [17] for RUL estimation of equipment using as input the age and conditions monitoring values at the present and past measurement. The developed method is validated using real world vibration monitoring data collected from pump bearings. The trend evolution of the computed HI is predicted in [18] using Recurrent Neural Network (RNN) for the RUL estimation of rolling bearing. LSTM is a recurrent neural network designed to learn the long term dependencies and can remember information for long

periods of time. They have been applied for RUL estimation of lithium-Ion batteries [19], aircraft turbofan engines [20], and bearings [21]. Although Machine learning techniques have proven their effectiveness for the RUL estimation, they require several historical degradation sequences for efficient estimation.

In industrial systems, a large amount of historical data about the normal (healthy) operating systems are often available, contrasting with a lack of historical data about degraded or faulty conditions. Those data are often unavailable because of their high cost, or for safety reasons, or in the case of newly installed machines. Generating degradation data in the laboratory is very costly, and the degradation behavior is not similar to reality due to the variation of environmental and operating conditions.

This paper proposes a new adaptive data-driven approach that can overcome the issue of RUL estimation when there is a lack of historical degradation sequence. The proposed approach can start to predict the RUL when having only one a priori degradation sequence. The RUL prediction accuracy is improved over time when as new incoming measurements become available. Also, the RUL prediction accuracy increase when more a priori sequences are collected. The proposed approach is based on an ensemble of adaptive models, where each model is fitted with its specific HI. Hence, several HI-model pairs are employed to predict multiple RULs, while the final RUL is obtained by merging the multiple RULs using a weighted average fusion method. The proposed approach is validated using a degradation scenario with different degradation speeds selected from a well-known aircraft engine degradation dataset (C-MAPSS dataset).

The paper is organized as follows. Section II presents the proposed approach for RUL estimation. The experimental data and the selected scenarios are detailed in Section III. The obtained results are discussed in Section IV. Finally, Section V ends the paper with concluding remarks and future work.

II. PROPOSED APPROACH

The proposed approach is illustrated in Fig. 1, it is splitted into two phases: offline and online. In the offline phase, the HI library and the model library are defined. Then, the best HI-model pairs (combinations) that provide the best RUL estimation accuracies on the a priori degradation sequences are selected, and their corresponding weights are computed. In the online phase, the data are collected then the selected HI-model pairs are used to estimate multiple RULs. The multiple RULs are merged using the weighted mean in order to predict the final RUL. When the failure is reached for the new sequence, this sequence is used for updating the selected HI-model pairs in order to improve the RUL prediction accuracy for the next new sequence.

A. HI and model libraries for RUL estimation

In offline, a library of HIs (features) sensitive to the degradation evolution is defined. The HIs in the library are chosen according to the application. For example when collecting vibration measurement, the chosen HI can be root mean square [13], kurtosis [6], peak values [18], or fast Fourier transform [22].

In addition, a library of extrapolation based models is defined, while the defined models should be able to predict the degradation evolution over time. For example, when the degradation evolution has an exponential trend (e.g., bearings degradation), different models can be selected such as Generalized Linear Model (GLM) [12], quadratic regression [13], exponential model [10], and double exponential model [11].

The model should be fitted with the first samples of the HI from the beginning of the degradation sequence (the degradation sequence starts from the degradation detection and ends with the failure occurrence). The model is then used to extrapolate the HI until the failure threshold. The time when the selected HI exceeds a predefined failure threshold is named the time of failure, or also the time of end of life t_{EOL} . The RUL is the time difference between the present time and the time of end of life as illustrated in Fig. 2.

The chosen models are used in an adaptive manner which means that for each time cycle, the model is re-fitted with the new collected points in order to update the model parameters. This will improve the RUL accuracy when more data points are collected (more accurate when we approach failure).

B. HI-model pairs selection and weights computation

In this work, an ensemble of HI-model pairs (combinations) are used to predict the RUL. Hence, the first step is to select the best HI-model pairs. The latter are selected according to the RUL accuracy of the stored sequence offline. The stored run-to-failure sequences are used to estimate the RUL using all the combinations of of HIs and models in the corresponding libraries. Then, an accuracy a_{ij} (Eq. (3)) is calculated for each HI-model pair (combination), where i represents the index of a HI in the HI library, and j the index the model in the model library.

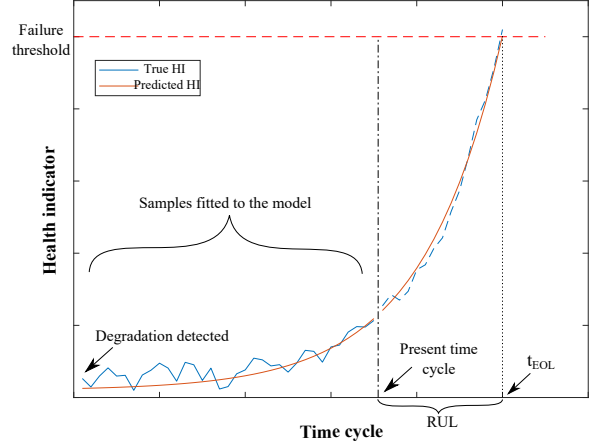


Fig. 2: Illustration of the RUL estimation.

The Cumulative Relative Accuracy (CRA) is used to evaluate the RUL prediction according to the actual RUL [23]. CRA is the weighted average of the Relative Accuracy (RA), where the RA is a measure of the error in RUL prediction (rul) relative to the actual RUL ($true_rul$) at a specific time index t . The time index for RUL prediction t starts when the degradation sequence starts and stops when the HI exceeds the failure threshold at t_{EOL} . The accuracy a_{ij} is computed as follows:

$$a_{ij}^s = CRA_{ij}^s \quad (1)$$

$$CRA = \frac{1}{n_t} \sum_{t=1}^{n_t} w(t) RA(t) \quad (2)$$

$$RA(t) = 1 - \frac{|rul(t) - true_rul(t)|}{true_rul(t)} \quad (3)$$

$$w(t) = \frac{\frac{1}{true_rul(t)}}{\sum_{t=1}^{n_t} \frac{1}{true_rul(t)}} \quad (4)$$

Where s is the sequence number, n_t is the number of time index (time cycle) from RUL estimation start until failure (EOL), and $w(t)$ is a weight factor defined as a function of the actual RUL [24]. It aims to improve the RUL evaluation by penalizing more the errors closer to t_{EOL} , since the RUL should be more accurate near the failure.

Table I presents an example of 4 HIs in the HI library and 3 models in the model library, where the accuracy a_{ij} is computed for each HI-model pair.

When there are two or more run-to-failure sequences available offline, the combinations accuracies a_{ij}^s (as in Table I) are computed for each sequence. Then, the mean is computed between the combinations accuracies of all the run-to-failure sequences, as shown in Eq. (5).

TABLE I: An example of accuracy computation for each combination HI-model

	Model ₁	Model ₂	Model ₃
HI ₁	a_{11}^s	a_{12}^s	a_{13}^s
HI ₂	a_{21}^s	a_{22}^s	a_{23}^s
HI ₃	a_{31}^s	a_{32}^s	a_{33}^s
HI ₄	a_{41}^s	a_{42}^s	a_{43}^s

$$a_{ij} = \frac{1}{n_s} \sum_{s=1}^{n_s} a_{ij}^s \quad (5)$$

Where a_{ij}^s is the combinations accuracy for each run-to-failure sequence s , and n_s is the number of available sequences. More the degradation sequences are available better the estimation accuracy of RUL for a new degradation sequence is. Therefore, at each time, a new degradation sequence is available, the selected HI-model pairs and their corresponding weights are updated. This is done to represent or cover better the variation of different degradation sequences dynamics.

When the combinations accuracies are calculated, the HI-model pairs are ranked according to the computed accuracies a_{ij} . Thereafter, the best combinations are selected when the accuracy is greater than a predefined threshold. A weight W_k is computed for each selected combination as a function of its accuracy. The weights can be defined as the normalization of the corresponding accuracy with a power factor γ as follows:

$$W_k = \frac{(a_{ij})_k^\gamma}{\sum_{k=1}^{n_k} (a_{ij})_k^\gamma} \quad (6)$$

Where n_k is the number of selected HI-model pairs, $(a_{ij})_k$ is the ranked combination accuracy where $(a_{ij})_1$ is the combination with the highest accuracy and $(a_{ij})_{n_k}$ is the lowest accuracy greater than the predefined threshold. γ is a power factor used to give more weight for combination with the highest accuracy. Table II is an example of 4 combinations selected (accuracy greater than a threshold). They are ranked according to the RUL estimation accuracy (the highest accuracy has the first rank), and the according weights are computed using Eq. (6).

TABLE II: An example of HI-model combination selection and ranking

Rank	Selected HI	Selected Model	Accuracies	Weights
1	HI ₂	Model ₁	$(a_{21})_1$	W_1
2	HI ₄	Model ₃	$(a_{43})_2$	W_2
3	HI ₁	Model ₁	$(a_{11})_3$	W_3
4	HI ₃	Model ₂	$(a_{32})_4$	W_4

It is worth mentioning that the RUL can be predicted using all the HI-model combinations, and weights near to zero will be given to the combinations with low accuracy, since the weights are computed according to the CRA accuracy in

an exponential manner (see Eq. (6)). However, using all the combinations for RUL estimation online can lead to a high computation time. For this reason, setting a threshold in order to select only the best HI-model pairs is necessary to reduce the computation time.

C. RUL estimation for a new degradation sequence

When the HI-model combinations are selected and their corresponding weights are computed offline, the final RUL is calculated using a weighted average of the computed RULs with the selected pairs (as shown in Fig. 3) estimated by the selected HI-model pairs as follows:

$$rul(t) = \sum_{k=1}^{n_k} W_k \cdot rul(t)_k \quad (7)$$

Where $rul(t)$ is the final RUL estimated at each time cycle t , $rul(t)_k$ is the RUL estimated by the HI-model pair k at each time cycle, and W_k is the corresponding weight to this pair.

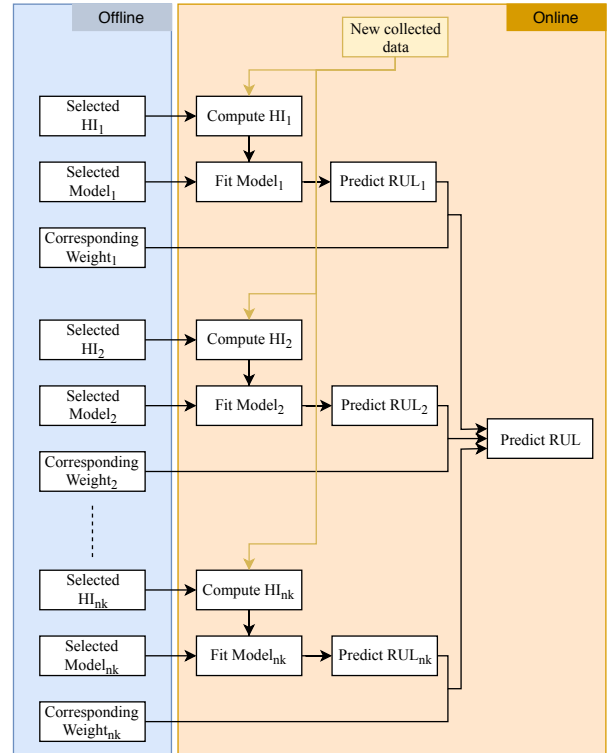


Fig. 3: Scheme of the HI-model fusion for RUL prediction.

III. EXPERIMENTAL DATA

The performance of the proposed approach is validated using a benchmark dataset named Commercial Modular Aero-P propulsion System Simulation (C-MAPSS) dataset. The dataset represents the damage propagation of the aircraft gas turbine engines. C-MAPSS dataset is generated by the NASA [25], this data has been widely used to compare RUL prediction methods in the literature [20]. The subdataset FD001

TABLE III: Degradation scenario

Sequence #	Degradation speed	Time cycle length
1	Fast	128
2		135
3		137
4	Medium	198
5		199
6		199
7	Slow	336
8		341
9		362

is considered for this validation, which includes one type of fault and one mode of operating conditions (flight conditions). This subdataset contains 100 run-to-failure sequences in the training set, and 100 sequences for testing that stop at some time before failure. For each sequence, there are 21 sensors measurements about the system conditions (e.g., temperature, pressure, and rotational speed).

In order to show the performance of our proposed approach, a degradation scenario is considered for this dataset. The sequences in the training set are sorted according to their length, then three sequences are selected for the smallest length, the largest length, and the medium length. Hence, a degradation scenario of 9 sequences is obtained with different degradation speeds, as seen in Table III.

IV. RESULTS AND DISCUSSION

A. Data processing and HIs-models definition

First, some sensors are not selected from the 21 sensors measurement in the dataset. The sensors # 1, 5, 6, 10, 16, 18, and 19 are discarded because the values of the sensors remain unchanged during operation, and the sensors # 9 and 14 are discarded because there is a high variation in the values of these sensors. The sensors measurements # 2, 3, 4, 7, 8, 11, 12, 13, 15, 17, 20, and 21 are selected and presented in Fig. 4.

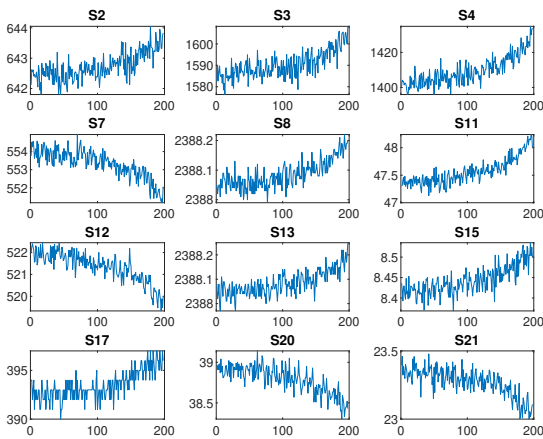


Fig. 4: Selected sensor values for the sequence #4

The HIs can be multiple features extracted from a single sensor measurements, or they can be based on multiple sensors measurements. Hence, a moving mean is applied in order to smooth the sensor values and to remove the noise. Then, the sensors that exhibit a negative exponential trend are flipped in order to obtain an exponential trend for all the sensor values. The constructed HI library is presented in Table IV.

TABLE IV: HI library

HI index	HI ₁	HI ₂	HI ₃	HI ₄	HI ₅	HI ₆
Sensor	S2	S3	S4	S7	S8	S11
HI index	HI ₇	HI ₈	HI ₉	HI ₁₀	HI ₁₁	HI ₁₂
Sensor	S12	S13	S15	S17	S20	S21

The model that can fit an exponential trend are selected in the model library: Model₁ : Generalized Linear Model (GLM) [12], Model₂ : quadratic regression [13], Model₃ : exponential model [10], and Model₄ : double exponential model [11]. The model library is shown in Table V.

TABLE V: Model library

Model index	Model ₁	Model ₂	Model ₃	Model ₄
Model	Generalized linear model	Quadratic regression	Exponential model	Double exponential model

1) *RUL estimation*: In offline, when some a priori sequences are available, the HI-model pairs that give a good CRA accuracy for the a priori sequences are selected. Since 12 HIs are in the HI library and 4 models in the model library, 48 HI-model pairs are obtained. Then, the HI-model pairs that show better accuracy than a threshold are selected. The CRA accuracies greater than a threshold of 0.6 are acceptable because the CRA accuracy is between 0 and 1. An example of the computed accuracies for each HI-model pair are shown in Table VI, where the accuracies greater than the threshold are shown in bold font. When the error of prediction is very high, the CRA accuracy value becomes negative. Hence, the accuracies with negative values are set to zero.

After that, the corresponding weight for each pair is computed using Eq. (6), where the exponent γ should be too high $\gamma = 10$ because the CRA accuracies are between 0 and 1. The selected HI-model pairs from Table VI are ranked, and

TABLE VI: Example of HI-model pairs accuracy when having sequence #1 as a priori

	Model ₁	Model ₂	Model ₃	Model ₄
HI ₁	0.529	0	0.492	0
HI ₂	0.295	0.714	0.314	0
HI ₃	0.392	0.780	0.494	0.292
HI ₄	0.606	0.119	0.766	0.634
HI ₅	0.291	0.676	0.407	0
HI ₆	0.619	0.508	0.707	0.005
HI ₇	0.616	0.552	0.647	0.415
HI ₈	0.756	0	0.684	0.150
HI ₉	0.398	0.547	0.545	0.182
HI ₁₀	0.033	0	0	0
HI ₁₁	0.473	0.291	0.337	0
HI ₁₂	0.630	0	0.674	0.403

the corresponding weights are computed and shown in Table VII.

TABLE VII: Example of selected pairs and corresponding weights

Rank	HI index	Model index	Accuracy	Weight
1	3	2	0.780	0.209
2	4	3	0.766	0.176
3	8	1	0.756	0.153
4	2	2	0.714	0.087
5	6	3	0.707	0.079
6	8	3	0.684	0.057
7	5	2	0.676	0.050
8	12	3	0.674	0.049
9	7	3	0.647	0.032
10	4	4	0.634	0.026
11	12	1	0.630	0.025
12	6	1	0.619	0.021
13	7	1	0.616	0.020
14	4	1	0.606	0.017

In order to validate our proposed approach, the RUL is predicted for all the 9 sequences in the degradation scenario. When only one a priori sequence is available, the RUL is predicted for each sequence by using each sequence as a priori. Then, the accuracies are computed as presented in Table VIII. Table VIII presents the CRA accuracies for the RUL estimation of the sequences in the scenario using one a priori sequence. It can be observed that the diagonal CRA accuracies in the table are the highest because the RUL is estimated for each sequence using as a priori the same sequence. When three and five a priori sequences are available, randomly sequences are added to the a priori sequence as shown in Table IX and Table X, respectively. The improvement in accuracy when more a priori sequences are collected can be shown by computing the average of each row (predicted sequence). Table XI shows the average accuracies of each row and the overall accuracy for each number of a priori sequence. From Table XI, the accuracy improvement when using more a priori sequences can be seen in the overall accuracy. Moreover, the Standard Deviation (STD) is computed for each row in order to show the robustness of prediction. It can be seen that the overall STD is decreasing when collecting more a priori sequences. Hence, our proposed approach can improve the accuracy of prediction as well as the robustness when more a priori sequences are available.

TABLE VIII: RUL estimation accuracies for the degradation scenario with 1 a priori sequence

		A priori sequences (1 sequence)								
		S1	S2	S3	S4	S5	S6	S7	S8	S9
Predicted sequences	S1	0.863	0.781	0.555	0.719	0.576	0.752	0.738	0.776	0.736
	S2	0.796	0.818	0.654	0.706	0.608	0.78	0.738	0.715	0.659
	S3	0.587	0.78	0.895	0.767	0.797	0.807	0.746	0.728	0.675
	S4	0.63	0.741	0.69	0.853	0.838	0.863	0.862	0.726	0.775
	S5	0.505	0	0.194	0.868	0.89	0.407	0.731	0.865	0.412
	S6	0.821	0.88	0.565	0.714	0.646	0.807	0.708	0.67	0.718
	S7	0.645	0.617	0.571	0.86	0.783	0.797	0.917	0.789	0.859
	S8	0.601	0.671	0.366	0.504	0.677	0.592	0.7	0.86	0.503
	S9	0.648	0.542	0.488	0.87	0.809	0.735	0.853	0.842	0.886

TABLE IX: RUL estimation accuracies for the degradation scenario with 3 a priori sequences

		A priori sequences (3 sequences)								
		S2	S7	S4	S8	S1	S3	S5	S9	S6
Predicted sequences	S1	0.819	0.731	0.603	0.732	0.691	0.722	0.671	0.718	0.704
	S2	0.762	0.769	0.626	0.748	0.721	0.676	0.658	0.73	0.721
	S3	0.801	0.672	0.809	0.726	0.706	0.791	0.793	0.67	0.624
	S4	0.779	0.779	0.883	0.804	0.844	0.689	0.848	0.758	0.789
	S5	0.13	0.517	0.868	0.73	0.844	0.8	0.864	0.848	0.668
	S6	0.79	0.727	0.645	0.662	0.665	0.74	0.698	0.667	0.714
	S7	0.59	0.865	0.873	0.89	0.843	0.81	0.868	0.796	0.831
	S8	0.649	0.721	0.67	0.82	0.782	0.841	0.751	0.876	0.722
	S9	0.672	0.805	0.838	0.844	0.842	0.877	0.855	0.869	0.845

TABLE X: RUL estimation accuracies for the degradation scenario with 5 a priori sequences

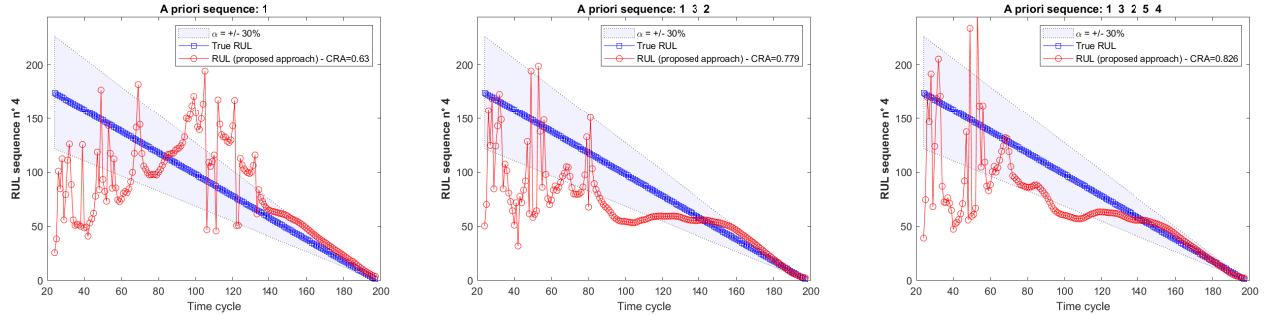
		A priori sequences (5 sequences)								
		S5	S9	S6	S3	S2	S1	S4	S7	S8
Predicted sequences	S1	0.686	0.717	0.618	0.728	0.756	0.736	0.676	0.769	0.744
	S2	0.714	0.727	0.698	0.754	0.661	0.703	0.665	0.729	0.722
	S3	0.711	0.665	0.788	0.74	0.635	0.694	0.807	0.622	0.61
	S4	0.826	0.794	0.867	0.796	0.829	0.745	0.865	0.795	0.769
	S5	0.841	0.801	0.831	0.796	0.83	0.84	0.874	0.823	0.823
	S6	0.669	0.702	0.682	0.696	0.674	0.689	0.682	0.664	0.671
	S7	0.828	0.832	0.865	0.866	0.839	0.795	0.883	0.841	0.801
	S8	0.798	0.848	0.721	0.831	0.771	0.849	0.716	0.862	0.87
	S9	0.852	0.861	0.834	0.848	0.838	0.876	0.855	0.855	0.836

The RUL estimation accuracy improvement is illustrated in Fig. 5. The figure shows the true RUL in blue, which is inversely proportional to the time cycle, the predicted RUL in red, and an interval of +/-30% on the true RUL plotted for better visualisation. The RUL is predicted for the sequence #4 by using as a priori one sequence (#9), three sequences (#9, #2, and #6), and five sequences (#9, #2, #6, #8, and #1). It can be observed that when collecting more a priori sequences the RUL estimation is improved, this is thanks to the proposed HI-model pairs selection that can update the selected pairs in offline for each new collected sequence.

The overall accuracies of the RUL estimation using the proposed approach are compared in Table XII with the overall accuracies using only the first best HI-model pair without merging several pairs. For example, in Table VII the best pair is HI₃ and Model₂. The proposed approach is also compared with SVR and LSTM. SVR and LSTM are machine learning techniques that have shown their effectiveness for the RUL estimation [14], [15], [19], [20]. From the table, it can be observed that the proposed approach outperforms the other techniques when few a priori sequences are available, this is thanks to the adaptability of our proposed approach and also thanks to the combination of several HI-model pairs. However, when several a priori sequences are available in offline (more than 6 sequences), the RUL estimation accuracy using our approach converge while the LSTM outperforms the proposed approach. LSTM has better results because unlike SVR that predicts the RUL directly considering each time point independently, LSTM is a recurrent neural network designed to learn the long term dependencies and can remember information for long periods of time. The results confirm that the proposed

TABLE XI: Average of RUL estimation accuracies with increasing number of a priori sequences

		Number of a priori sequences								
		CRA (1seq)	CRA (2seq)	CRA (3seq)	CRA (4seq)	CRA (5seq)	CRA (6seq)	CRA (7seq)	CRA (8seq)	CRA (9seq)
Predicted sequences	S1	0.722+0.098	0.702+0.0774	0.71+0.0573	0.702+0.0497	0.715+0.0471	0.736+0.0344	0.745+0.0194	0.749+0.0187	0.76+1.57e-16
	S2	0.719+0.071	0.705+0.0791	0.712+0.0488	0.714+0.0385	0.708+0.0303	0.722+0.0363	0.729+0.0277	0.725+0.0231	0.726+1.11e-16
	S3	0.754+0.0867	0.731+0.0722	0.732+0.0689	0.691+0.066	0.697+0.0711	0.687+0.0667	0.686+0.045	0.685+0.0326	0.695+1.3e-16
	S4	0.775+0.0845	0.82+0.0381	0.797+0.0571	0.803+0.0489	0.809+0.041	0.808+0.0296	0.811+0.0192	0.813+0.0166	0.823+1.76e-16
	S5	0.541+0.32	0.714+0.193	0.697+0.242	0.811+0.0482	0.829+0.023	0.836+0.0181	0.838+0.0153	0.835+0.0128	0.836+1.18e-16
	S6	0.725+0.097	0.701+0.0413	0.701+0.0467	0.68+0.0212	0.681+0.0127	0.682+0.0179	0.682+0.0144	0.683+0.00919	0.686+9.61e-17
	S7	0.76+0.121	0.835+0.055	0.818+0.0909	0.832+0.05	0.839+0.0295	0.835+0.0369	0.835+0.0297	0.842+0.0189	0.853+1.3e-16
	S8	0.608+0.142	0.709+0.0968	0.759+0.0771	0.792+0.0736	0.807+0.0592	0.813+0.0388	0.826+0.0336	0.833+0.0299	0.833+1.92e-16
	S9	0.741+0.149	0.82+0.0485	0.828+0.0617	0.848+0.0184	0.851+0.0136	0.85+0.0122	0.853+0.0108	0.855+0.0103	0.858+1.96e-16
Overall accuracy		0.705+0.13	0.749+0.078	0.751+0.0834	0.764+0.0461	0.771+0.0364	0.774+0.0323	0.778+0.0239	0.78+0.0191	0.786+1.45e-16



(a) Predicted RUL using a priori sequence #9 (b) Predicted RUL using a priori sequence #9, #2, and #6 (c) Predicted RUL using a priori sequence #9, #2, #6, #8, and #1

Fig. 5: RUL prediction for sequence #4 using different available a priori sequences

approach is very efficient for RUL estimation dealing when few historical degradation sequences are available.

Figure 6 shows the adaptability of our proposed approach by comparing with SVR and LSTM methods. The RUL is estimated for the sequence #4 which is a sequence with a medium degradation speed. the RUL is estimated using as a priori one sequence with different degradation speed (i.e., fast, medium, and slow). It can be observed that when the RUL is predicted using an a priori sequence with a degradation speed different than the predicted one, our approach has a good accuracy comparing with SVR and LSTM that show a low accuracy. Our proposed approach can estimate the RUL with good accuracy even when the degradation speed of the a priori sequence is different from the predicted one. This is thanks to the adaptability of our proposed approach that combine different adaptive models. When the degradation speed of the a priori sequence is similar to the new sequence, it can be seen that the LSTM also shows a good accuracy for RUL estimation, SVR shows a low accuracy because CRA accuracy penalizes more the errors of prediction near to the failure.

V. CONCLUSION

In this paper, a new data-driven approach for RUL estimation is proposed. This approach can estimate the RUL when few historical degradation sequences are available. It can predict several RULs with several HI-model pairs (combinations), while these pairs are selected in offline by using the few available a priori sequences. Then, the predicted RULs

are merged by using a weighted mean in order to obtain the final RUL. The approach is validated using a degradation scenario selected from an aircraft engine degradation dataset (C-MAPSS dataset). The proposed approach has proven their effectiveness for an accurate RUL estimation when few a priori sequences are available comparing to the SVR and LSTM methods. The accuracy of prediction is improved when more a priori sequences are collected in offline. The proposed approach is also more robust when more a priori sequences are collected, because the RUL prediction becomes consistent with the number of a priori sequences (as shown by the decreasing of the STD). Moreover, the proposed approach has shown its adaptability by predicting the RUL of a sequence with a degradation speed different than the a priori one, this is thanks to the use of adaptive models.

One future work is to provide a confidence interval for the predicted RUL with respect to the chosen model. It is also planned to improve the proposed approach in order to estimate the RUL with a higher accuracy when multiple a priori sequences are available offline using efficient techniques such as deep learning.

ACKNOWLEDGMENT

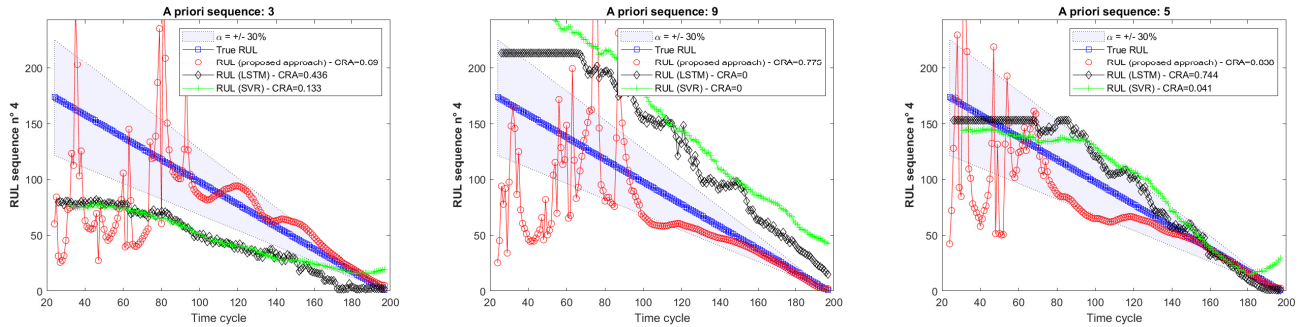
This paper is the result of the research work supported by the European Union, European Regional Development Fund.

REFERENCES

- [1] R. Veikko, "Safety and reliability technology review," in *VTT symposium*, vol. 592. VTT; 1999, 2006.

TABLE XII: Comparison of the RUL estimation overall accuracies

	CRA (1seq)	CRA (2seq)	CRA (3seq)	CRA (4seq)	CRA (5seq)	CRA (6seq)	CRA (7seq)	CRA (8seq)	CRA (9seq)
SVR	0.167	0.325	0.408	0.498	0.559	0.620	0.663	0.702	0.742
LSTM	0.362	0.448	0.540	0.615	0.657	0.749	0.794	0.851	0.9120
Best HI-model	0.487	0.635	0.586	0.647	0.648	0.659	0.661	0.670	0.676
Proposed approach	0.7050	0.7490	0.7510	0.7640	0.7710	0.7740	0.7780	0.7800	0.7860



(a) Predicted RUL using a priori sequence #3 (b) Predicted RUL using a priori sequence #9 (c) Predicted RUL using a priori sequence #5

Fig. 6: RUL prediction for sequence #4 using a priori sequences with different degradation speed

- [2] A. Helle, "Development of prognostic concepts and tools," in *VTT symposium*, vol. 243. VTT; 1999, 2006, p. 5.
- [3] K. Abid, M. Sayed-Mouchaweh, and L. Cornez, "Fault prognostics for the predictive maintenance of wind turbines: State of the art," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2018, pp. 113–125.
- [4] H. Toubakh, M. Sayed-Mouchaweh, M. Benmiloud, M. Defoort, and M. Djemai, "Self adaptive learning scheme for early diagnosis of simple and multiple switch faults in multicellular power converters," *ISA transactions*, 2020.
- [5] J. B. Coble, "Merging data sources to predict remaining useful life—an automated method to identify prognostic parameters," 2010.
- [6] L. Saidi, J. B. Ali, E. Bechhofer, and M. Benbouzid, "Wind turbine high-speed shaft bearings health prognosis through a spectral kurtosis-derived indices and svr," *Applied Acoustics*, vol. 120, pp. 1–8, 2017.
- [7] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," *Neurocomputing*, vol. 240, pp. 98–109, 2017.
- [8] K. Javed, R. Gouriveau, N. Zerhouni, and P. Nectoux, "Enabling health monitoring approach based on vibration data for accurate prognostics," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 647–656, 2015.
- [9] J. Yan, M. Koc, and J. Lee, "A prognostic algorithm for machine performance assessment and its application," *Production Planning & Control*, vol. 15, no. 8, pp. 796–801, 2004.
- [10] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: a neural network approach," *IEEE Transactions on industrial electronics*, vol. 51, no. 3, pp. 694–700, 2004.
- [11] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 9, pp. 2046–2054, 2016.
- [12] K. Abid, M. Sayed-Mouchaweh, and L. Cornez, "Adaptive machine learning approach for fault prognostics based on normal conditions-application to shaft bearings of wind turbine," in *Proceedings of the Annual Conference of the PHM Society*, vol. 11, no. 1, 2019.
- [13] W. Ahmad, S. A. Khan, and J.-M. Kim, "A hybrid prognostics technique for rolling element bearings using adaptive predictive models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1577–1584, 2017.
- [14] T. Benkedjough, K. Medjaher, N. Zerhouni, and S. Rechak, "Remaining useful life estimation based on nonlinear feature reduction and support vector regression," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1751–1760, 2013.
- [15] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2015.
- [16] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network," *Mechanical Systems and Signal Processing*, vol. 56, pp. 150–172, 2015.
- [17] Z. Tian, "An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring," *Journal of Intelligent Manufacturing*, vol. 23, no. 2, pp. 227–237, 2012.
- [18] A. Malhi, R. Yan, and R. X. Gao, "Prognosis of defect propagation based on recurrent neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 703–711, 2011.
- [19] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5695–5705, 2018.
- [20] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla lstm neural networks," *Neurocomputing*, vol. 275, pp. 167–179, 2018.
- [21] B. Zhang, S. Zhang, and W. Li, "Bearing performance degradation assessment using long short-term memory recurrent network," *Computers in Industry*, vol. 106, pp. 14–29, 2019.
- [22] L. Liao, "Discovering prognostic features using genetic programming in remaining useful life prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2464–2472, 2013.
- [23] A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and M. Schwabacher, "Metrics for evaluating performance of prognostic techniques," in *2008 International Conference on Prognostics and Health Management*. IEEE, 2008, pp. 1–17.
- [24] K. Goebel, A. Saxena, S. Saha, B. Saha, and J. Celaya, "Prognostic performance metrics," in *Machine Learning and Knowledge Discovery for Engineering Systems Health Management*, A. N. Srivastava and J. Han, Eds. Chapman and Hall - CRC, 2011, ch. 5.
- [25] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in *2008 international conference on prognostics and health management*. IEEE, 2008, pp. 1–9.