International Journal of Engineering

Journal Homepage: www.ije.ir

A Life Clustering Framework for Prognostics of Gas Turbine Engines under Limited Data Situations

A. Mahmoodian[a], M. Durali*[a], M. Saadat[b], T. Abbasian[b]

[a] *Mechanical Engineering Department, Sharif University of Technology, Tehran, Iran*
[b] *ECE Faculty, Tehran University, Tehran, Iran*

P A P E R I N F O

Paper history:
Received 23 August 2020
Received in revised form 01 January 2021
Accepted 04 January 2021

Keywords:
Limited Data
Prognosis and Health Management
Remaining Useful Life Estimation
Robustness

A B S T R A C T

The reliability of data driven prognostics algorithms severely depends on the volume of data. Therefore in case of limited data availability, life estimations usually are not acceptable; because the quantity of run to failure data is not sufficient to train prognostics model efficiently. To board this problem, a life clustering prognostics (LCP) framework is proposed. LCP regenerates the train data at different ages and outcomes to increment of the training data volume. So, the method is useful for limited data conditions. In this research, initially LCP performance is studied in normal situation is; successively robustness of the framework under limited data conditions is considered. For this purpose, a case study on turbofan engines is performed. The accuracy for the proposed LCP approach is 71% and better than other approaches. The prognostics accuracy is compared in various situations of data deficiency for the case study. The prognostic measures remain almost unchanged when the training data is even one third. Successively, prognostics accuracy decreases with a slight slope; so that when the training data drops from 100 to 5%, the accuracy of the results drops 26%. The results indicates the robustness of the proposed algorithm in limited data situation. The main contribution of this paper include: (1) The effectiveness of life clustering idea for use in prognostics algorithms is proven; (2) A step-by-step framework for LCP is provided; (3) A robustness analysis is performed for the proposed prognostics algorithm.

doi*: 10.5829/ije.2021.34.03c.18*

1. INTRODUCTION

In recent years, prognostics and health management (PHM) of complex mechanical systems has become more prominent. Prognostic and Remained Useful Life (RUL) prediction has been initialized in medical field [1]; subsequently attracted much attention in engineering issues due to economical and operational considerations [2]. Predicting future behavior of a complex machine such as a gas turbine is a complicated task. Prognostics is currently at the core of systems' health management to achieve reliable and safe operation of machines. In the framework of PHM, many techniques exist which are basically classified into two principal classes: data-driven and model-based prognostics approach [3-4]. The fact that most researches are focusing on data-driven methods shows the desire to work with easily accessible data as compared to model-based methods, irrespective of the difficulties in gaining accessing statistically significant run-to-failure data. Despite acceptance of data-driven methods, the on-going difficulty with these methods is that they show acceptable RUL estimation only when abundant run-to-failure data are available for training. Although, under the condition of limited failure data, model-based solutions are unsuccessful due to their requirement to large amounts of failure data for validating physical models [5].

In this research, data driven methods are focused and divided mainly in two groups: typical methods and robust methods. A typical prognostics method rely on large amounts of historical failure data (i.e. run-to failure data indicating past degradation patterns) to estimate prognostics model parameters [5]. Otherwise, the predictions may be unreliable and the training can not be

*Corresponding Author Institutional Email: *durali@sharif.edu* (M. Durali)

Please cite this article as: A. Mahmoodian, M. Durali, M. Saadat, T. Abbasian, A Life Clustering Framework for Prognostics of Gas Turbine Engines under Limited Data Situations, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 03, (2021) 728-736

carried out. The fact limits the application of typical prognostics methods in problems with small amount of available training data [6]. Failure data are limited in real industrial cases due to some reasons: (1) rare failures; (2) overprotective maintenance and replacement regimes; (3) incomplete reports [7].

Majority of prognostics researches are developed assuming enough run-to-failure data are available for training. Some researchers have used combined regression techniques, including linear and quadratic models to predict the RUL of gas turbine engines [8]. Other methods like particle filter [9], adaptive-order particle filter [10], Kaplan Meier [11] and support vector machine (SVM) [12] have been used for prognostics. Yu et al. [13] have developed a prognostics system for engine health assessment based on logistic regression and state-space-model. Simon et al. [14] have compared the estimation accuracy and computational effort of variants of the Kalman filter like linearized Kalman filter (LKF), extended Kalman filter (EKF), and unscented Kalman filter (UKF) for aircraft engine health estimation. Lu et al. [15] have presented a nonlinear state estimation method based on EKF and obtained a significant improvement in estimation accuracy and robustness. Ding et al. [16] have fused support vector machine and the genetic algorithm and proposed an intelligent prognostics approach. Goebel et al. [17] compared the results of a relevance vector machine (RVM), a Gaussian process regression (GPR), and an artificial neural network (ANN) approach in prognostics.

On the contrary, with the robust methods, the RUL estimation is acceptable despite the lack of abundant trained data. Recently, some novel prognostics methods based on classic algorithms are suggested, such as fusion of prognostics algorithms [18], multistate structure [19], etc. Xu et al. [18] have integrated the strengths of the experience-based prognostics approach and the data-driven approach. The developed fusion prognostics framework has been employed to predict the RUL of a gas turbine engine as an application example [18]. Moghaddass et al.[19] have demonstrated that deterioration process occurs through different levels of health states before failure, leading to a multistate deterioration process in many real-world cases. Xiang et al. [20] have proposed a probabilistic methodology for fatigue prognostics using an inverse first-order reliability method. However, the robustness tests are rarely reported .

In this paper, a novel methodology is presented based on life clustering that allows training datasets to be augmented. Usually, the goal of clustering in literature has been to organize data into homogeneous groups to compact clusters with minimum intra group similarity and to increase separation among clusters with maximized inter group dissimilarity [21]. The proposed method of life clustering prognostics (LCP) is able to increase the train data set samples; in addition to organize data into homogeneous groups.

The case study is to compare the prognostics accuracy for a robust framework in abundant data and little data condition with for turbofan engines Prognostic Health Management (PHM) Challenge data [22]. For 2008 PHM challenge, many authors have reported the RUL estimation for a given data set. Only typical tests using large scale data have been stated. The best results have been obtained by using RULCLIPPER algorithm [23], EVIPRO algorithm [24] and a similarity-instance based approach [25]. These results are used as the reference point of the current study .

The main contributions of this paper can be summarized as follows. First, the effectiveness of the idea of life clustering for use in prognostics algorithms is proven. Although artificial neural network is used as the main prediction tool in this paper, the idea of LCP can be combined and used with other classical methods of prognostics. Second, a step-by-step framework for prognostics based on life clustering is provided. This method significantly improves the reliability of this algorithm while using all the advantages of a predictive algorithm as the core prediction algorithm. Third, a robustness analysis is performed for the proposed prediction algorithm. This study evaluates the performance of the algorithm in different conditions of lack of sufficient data. Based on the evaluations, the LCP algorithm is robust in limited information conditions and has acceptable results. Robustness of a prognostics algorithm is a critical issue in industrial and real-world cases, where predictive maintenance is required against lack of abundant run-to-failure data .

The outline of this paper is as follows. In section 2 layout of the study is presented. Data processing and prognostics method is explained in section 3. In sections 4, implementation of the proposed method on a case study is described. This paper ends with results and conclusions in two last sections.

2. LAYOUT OF THE STUDY

2. 1. Data Description To illustrate the outcomes of this method on prognostics and health monitoring, a case study on turbofan engines from NASA's prognostics Information Repository is performed. The structure of the data set and the effectiveness of the proposed model are presented in this section. The data consists of 21 Measurements, including the measurements listed in Table 1, that are measured during every flight cycle. In the dataset, multiple units operate until failure occurs, providing training set. The other units run to different levels of destruction, forming test set. The challenge is to predict the RUL of test units. This dataset is one of the most widely datasets used for the development and

TABLE 1. List of sensors and Measurements used in this paper [22]

Symbol	Description	Unit
T24	Total temperature at LPC outlet	^{O} R
T30	Total temperature at HPC outlet	^{O} R
P30	Total pressure at HPC outlet	Psia
Nc	Physical core speed	Rpm
Pr	Engine pressure ratio (P50 / P2)	-
Phi	Ratio of fuel flow to Ps30	-
BPR	Bypass ratio	-
BE	Bleed enthalpy	-
T50	Total temperature at LPT outlet	O R
Ps30	static pressure at HPC outlet	Psia
farB	Burner Fuel air ratio	-

validation of prognostics algorithms [2, 18-19]. Figure 1 shows the main components of the aircraft gas turbine engine model.

2. 2. Prognostic Measures In the PHM context, sometimes it is desirable to predict early as compared to predicting late. Therefore, the asymmetric interval *I*= [-10, +13] around the true RUL is considered to evaluate the performance. Accuracy measure is defined as the percentage of correct estimations which falls within the interval *I* [25].

Mean square error (MSE) and mean absolute error (MAE) are two other measures which are used to evaluate the performance of LCP method more accurately,

$$e_{mse} = \frac{1}{N}\sqrt{\sum_{t=1}^{N}(err_t)^2} \quad (1)$$

$$e_{mae} = \frac{1}{N}\sum_{t=1}^{N}|err_t| \quad (2)$$

where error is defined for a given prediction by Equation (2):

$$err = True\ RUL - Estimated\ RUL \quad (3)$$

3. PROGNOSTICS METHOD

In a data-driven PHM process, after data acquisition, the first challenge is how to map the conditions between a complex and interconnected system with its level of the drop; for this purpose, data processing and signal feature extraction should be done. Two general approaches are considered to extract features and design a health indicator (HI). The first approach is to use gas-path parameters such as temperature, vibrations, flow capacity, pressure, compressor efficiency, fan efficiency,

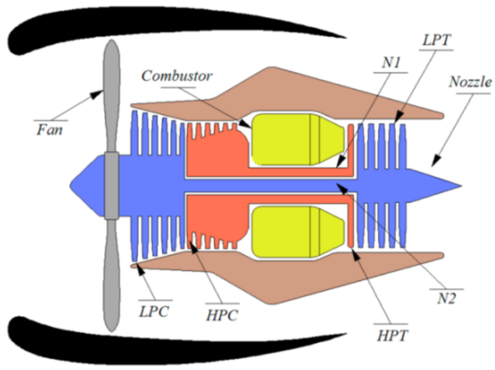


Figure 1. Simplified diagram of the gas turbine engine [22]

etc. Vibration and modal analysis is widely used to estimate life of mechanical systems [26-27]. Mohammadi et al. [28] determined performance deterioration according to efficiency and flow capacity as health indicators. The second approach is the combination of functional and performance sensors. In this field, we can mention the works [18, 29, 30], that combine different sensors with different fusion techniques. Diallo [11] has shown in his research that multi-sensor data fusion approach is more reliable. A step by step methodology is indicated in literature [30] to produce a Health Indicator Feature [*HIF*] vector, which is used in this article.

3. 1. The LCP Prognostic Framework A prognostics framework based on life clustering is developed as shown in Figure 2. RUL estimation is accomplished through the life clustering of the engines and subsequently construction of a specific prediction module for each cluster. The proposed framework can be implemented as the following phases.

Clustering In the first phase, clustering is performed. The estimation start time of each test engine is shown with symbol *tc*. In the prognostics issue for a fleet of engines, a range of *tc* [min (*tc*): max (*tc*)] exists. The range must be separated to several divisions in the clustering phase. The cluster width (*CW*) is defined as follows:

$$CW_i = t_i - t_{i-1}, \sum_{i=1}^{n} CW_i = t_n - t_0 \quad (4)$$

where t_is are clustering borders, t_0 is the min (*tc*), t_n is the max (*tc*) and *n* is the number of partitions. In the simple form, clusters widths may be assumed equal. To attain more accurate results, the number of partitions and cluster widths can be found by an optimization process.

Reproduction In the second phase, a time step (*ts*) is considered so that train data set is observed several times at each time step. Considering maximum observation resolution, each time step is an observation point (*ts*=1). For example, the i^{th} cluster width is m_i ($CW_i = m_i$), so the train data set is reproduced m_i times. For any

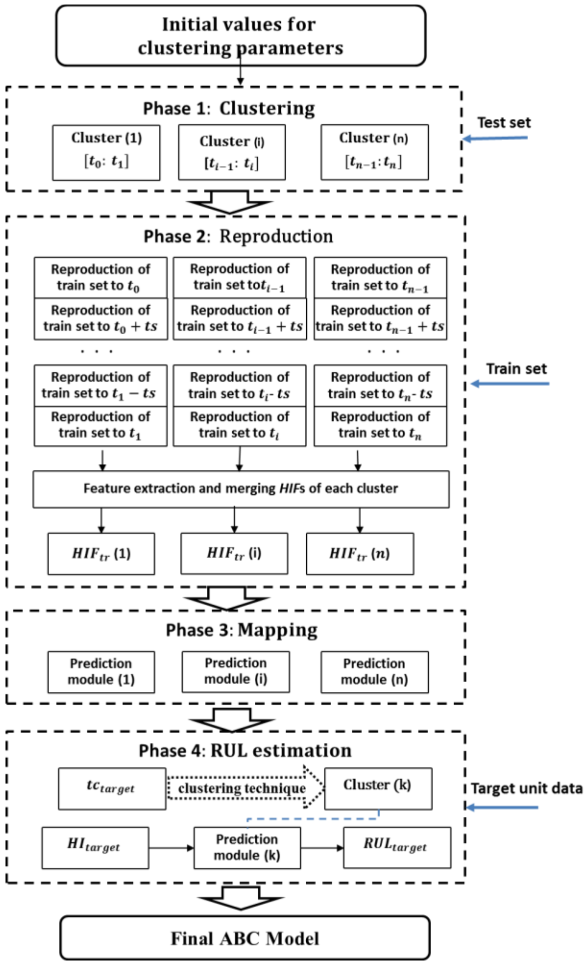


Figure 2. The framework of LCP prognostics

reproduction, units with life lengths less than the relevant observation point are withdrawn; the others are stopped at that point.

Till now, mi reproductions of train data set are produced for each cluster. Now, data processing of units in each reproduction is performed and mi vectors HIF are produced. In the last step of this phase, the HIF vectors for each cluster are merged to one and finally n cumulative HIF vectors are remained for n clusters.

Mapping In the previous phase, train data was reproduced several times and n cumulative *HIF* vectors were resulted. Now in the third phase, a prediction tool such as neural network is used to find a relation between *HIF*s and RULs for train data. The *HIF*s are selected as the input, and the corresponding true RUL data are selected as the target values to train the prediction modules. The results of the third phase are *n* prediction modules for *n* different life clusters.

RUL estimation In the fourth phase, RUL estimation of the target unit is performed. Initially, signal processing for the device is performed and the relevant HIF vector is calculated. Then, cluster selection must be done according to the prediction start age (tc) of the test unit. It can be done in two ways:

1. Classic way, so that each unit belongs to a cluster if its tc is between the minimum and maximum age of that cluster
2. Fuzzy way, in which each unit belongs to a cluster to some degree that is specified by a membership function.

After cluster selection, the *HIF* vector of the test unit is inputted to the related prediction module and the device RUL is estimated.

4. CASE STUDY

The steps taken in this study are based on the framework given in Figure 2. The main parameters of the LCP framework are the number of clusters (*n*), the clusters width (*CW*) and the observation time step (*ts*). The optimum values for these parameters are different for each problem and must be optimized. To achieve this goal, various prognostics measures may be defined as the objective function. Sequential phases of the LCP framework are executed and the best parameters are found through an optimization process.

Phase 1 To determine the clustering parameters, four phases of LCP prognostics are implemented and the prognostics measures were compared. For the current case study, prognostics criterions remain almost unaffected while *n*>2, as shown in Figure 3. Thus, *n* = 4 is chosen to evade costly computations. Sequentially, the optimum clusters widths (*CW*) are found out by genetic algorithm to maximize the accuracy. The optimal clustering parameters are summarized in Table 2.

Phase 2 Considering observation time step =1 cycle, the train data set is reproduced at each cycle. For every cycle, for instance the c^{th} cycle, all engines with age longer than c cycles are stopped at the cth cycle. Then data process is performed and health indicator features $HIF_{tr}(c)$ are extracted. Finally relevant HIFs

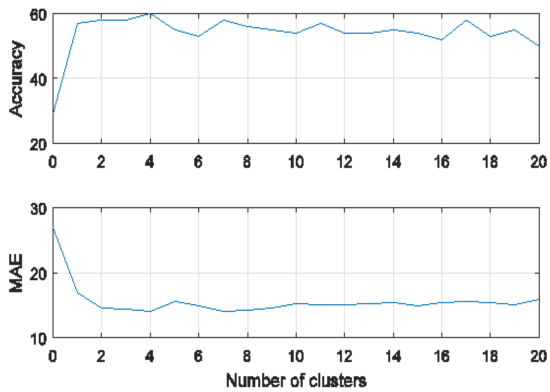


Figure 3. Prognostic measures via. number of clusters

TABLE 2. Summary of optimal clustering scheme

Cluster	Life interval	Cluster width
I	31-75 cycles	45 cycles
II	76-125 cycles	50 cycles
III	126-196 cycles	71 cycles
IV	197-303 cycles	107 cycles

for each cluster are merged to one, for example $\boldsymbol{HIF_{tr}(31)}, \boldsymbol{HIF_{tr}(32)} \dots \& \boldsymbol{HIF_{tr}(75)}$ are merged to one cumulative vector for the 1st cluster, $\boldsymbol{HIF_{tr}(76)}, \boldsymbol{HIF_{tr}(77)} \dots \& \boldsymbol{HIF_{tr}(125)}$ are merged to one cumulative vector for the 2nd cluster and so on for other clusters from Table 2. Summary of reproductions of train data set are represented in Table 3.

Phase 3 In the third phase, the algorithm creates a map between HIFs and RULs for each cluster. For this aim, neural network fitting tool is used [31]. In the present study, a forward multilayer perceptron (MLP) with backward propagation Bayesian training algorithm is applied. The network is composed from 10 hidden layers, 70% of data is used for train, 15% for test and 15% for validation. The summary of the formation of four networks is presented in Table 4.

Phase 4 RUL estimation is performed through the following steps:

1. For the j^{th} test engine, *tc* is considered (tc_j).
2. Depending on its age, the appropriate group from Table 2 is selected (group #k).
3. The relevant ***HIF*** vector of j^{th} test engine is extracted ($\boldsymbol{HIF_{te}(j)}$).
4. $\boldsymbol{HIF_{te}(j)}$ is applied as an input to the k^{th} network.
5. The network output is the estimated RUL of the test engine.

TABLE 3. Summary of train data set reproductions

Cluster	length of cumulative HIF
1	4500
2	5000
3	5971
4	2240

TABLE 4. Summary of the formation of the prediction modules

ANN	Input	Target
I	2x4500 matrix, representing 2 features of 4500 HI signals	1x4500 matrix, representing RULs of engines
II	2x5000 matrix, representing 2 features of 5000 HI signals	1x5000 matrix, representing RULs of engines
III	2x5971 matrix, representing 2 features of 5971 HI signals	1x5971 matrix, representing RULs of engines
IV	2x2240 matrix, representing 2 features of 2240 HI signals	1x2240 matrix, representing RULs of engines

5. RESULTS AND DISCUSSION

The performance of the proposed prognostics method is evaluated in two states:

1. Full train data: In this state, all engines (100 units) of train dataset #1 from turbofan engines of the NASA Prognostic Data Repository are utilized to train the LCP algorithm. This is similar to the condition in which most researches have used for training their algorithms and represented their results.
2. Limited train data: In this state, a portion of train dataset #1 is used for training LCP algorithm. Tests are performed with 50, 30, 20 and 10% of train data (equal to 50, 30, 20 and 10 engines). This state is similar to real world in industries when one should deal with a limited train data.

5. 1. LCP Results in Full Train Data Condition To evaluate the effectiveness of the LCP algorithm, a comparison with other approaches is performed as indicated in Table 5. Full testing dataset is used in few papers to our knowledge: Ramasso et al. [23, 24], Khelif et al. [25] and Wang et al. [32] that achieved the best score in PHM challenge 2008. The accuracy for the proposed LCP approach is better in comparison with other approaches.

5. 2. LCP Results under Limited Train Data Condition It was shown in the previous section that the accuracy of the LCP method is reliable for full data condition in comparison with other methods. In this section a sensitivity analysis is performed under limited data condition. As mentioned earlier, four tests are performed in this section with 50, 30, 20 and 10% of train data. Each test is repeated several times in a way that different portions of the train data are selected. Finally, the mean value of prognostic measures for each test is

TABLE 5. Comparison of accuracy for different methods

Method	Correct %	Early %	Late %
LCP model (using ANN for prediction modules)	71	23	6
Ramasso [24]	67	Nan	Nan
Khelif et al.[25]	54	18	28
Ramasso et al. [23]	53	36	11
Javed et al. [33]	53	27	20
Wang et al. [32]	44	19	37

summarized in Table 6. A comparison of prognostics measures for different sizes of train data is indicated in Figure 4.

While the proposed algorithm reproduces train data several times for each life cluster, the size of the training set is several times larger than the initial train data set. Comparison of results for different numbers of training data showed that when the number of training data decreases, (1) prediction accuracy remains almost unchanged for about 30% of available train data, (2) prediction accuracy decreases about 14% with a slight slope when the available train data falls from 30 units to 10 units and (3) when the available train data goes down from 10 units to 5 units, prediction drops 10% significantly.

Further investigations can be made on the robustness of the proposed algorithm under limited training data condition. Prognostics results of engines in different cycles are studied in this paper. The actual RUL value and the RUL estimate with limited train data (10 units for training) are shown in Figure 5 for engines #99-100. Other engines results (#91-98) are shown in Appendix A The results showed that overall (1) the RUL estimate with small data is reasonably close to the actual RUL especially in large ages, (2) as the engine ages, the prediction error for the RUL gets narrower, means that, the prognostics uncertainty declines, and (3) in some cases, lack of enough data especially in the last cluster, results to increase prediction error. In general, the results indirectly support the hypothesis that, life clustering method leads to acceptable results in condition of train data deficiency (although not necessarily the best method).

TABLE 6. Summary of prognostic measures

Number of train data units	Accuracy (%)		MAE	MSE
	Mean	STD		
100	58 [2]	0	14.14	1.97
50	57	3.1	14.5	2
30	56	3.4	14.4	2
20	46.6	3.1	17.3	2.3
10	42	5.7	21.2	2.9
5	32	7.8	31.75	4.4

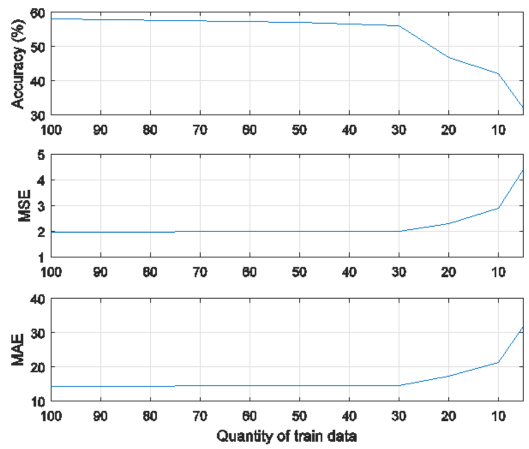


Figure 4. Comparison of prognostics measures for different sizes of train data

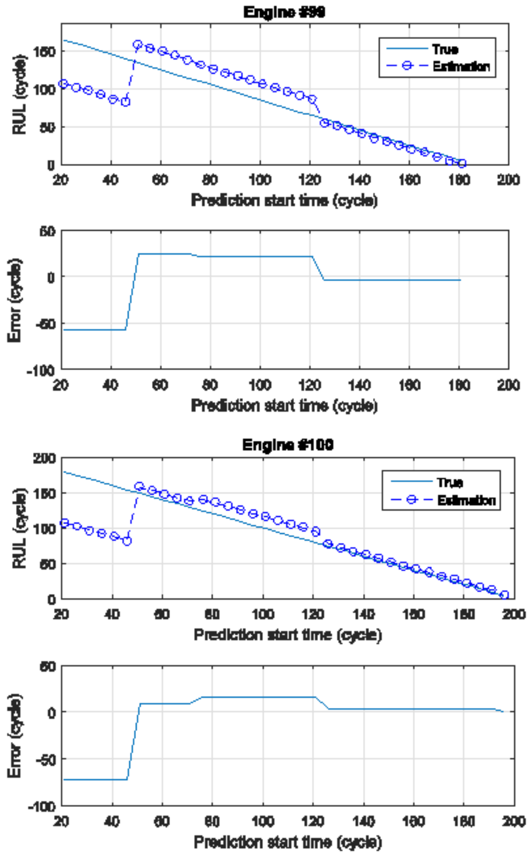


Figure 5. Prognostics results of engines #99-100 in different cycles with limited train data set (10 units for training)

6. CONCLUSION

In this paper, a prognostic algorithm is proposed that first classifies the test units in different age groups, then estimates their RUL using predictive techniques. The

[2] The neural network training multiple times will generate different results due to different initial conditions; Therefore, for sensitivity analysis, the polynomial regression is used for composing prediction modules of the ABC method and the resulted accuracy is different with table 5 (ABC model using ANN)

proposed algorithm uses one of the conventional and available prediction methods (such as ANN as presented in this study) as the core prediction tool and rectifies it with more reliable and robust results .

A case study shows that the results achieved by this method were significantly improved compared to other conventional methods and it was observed that life clustering can be very effective in prognostics. LCP was able to predict with 71% accuracy, a little better than the best published results on the same case study. While the performance of the LCP method was evaluated under normal conditions, its results were examined in the condition of limited training data, which happens frequently in industry .

Comparison of results for different conditions of available training data showed that the prognostic measures remain almost unchanged when the training data is even one third. The reason is that the training data set has been multiplied several times and it compensates the lack of enough training data. Successively, prognostics accuracy decreases with a slight slope; so that when the available training data drops from 30 to 10%, the accuracy of the results drops from 56 to 42%. After that accuracy drops considerably to 32% for 5% of available train data. Although significant accuracy drop is observed below 10% of available train data, it is notable that LCP is using train data of only 5 units to predict remaining life of 100 test units .

In the final stage, more cases were tested and the results of the prognostics algorithm were plotted using a low number of training data (10 engines). The results show that the RUL estimate with small data is rationally close to the actual RUL, although in some cases, severe lack of data especially in the last cluster, results to increase prediction error .

The results of this case study confirmed that the LCP method (1) is a powerful prognostics tool in normal condition and (2) is a robust technique under limited data condition. So the proposed method can be integrated with any classic method to result more accurate and robust RUL estimates for real-world situations. The methodology developed in this paper is not limited to the use with turbojet engine prognostics. It can be extended to other prognostics problems .

Integration and modification of more prediction methods with life clustering idea are to be investigated in future works. However, there are some potential limitations existing in the prediction for the last cluster (in case of severe lack of data) which could be improved with fusion of previous clusters predictions in the future study.

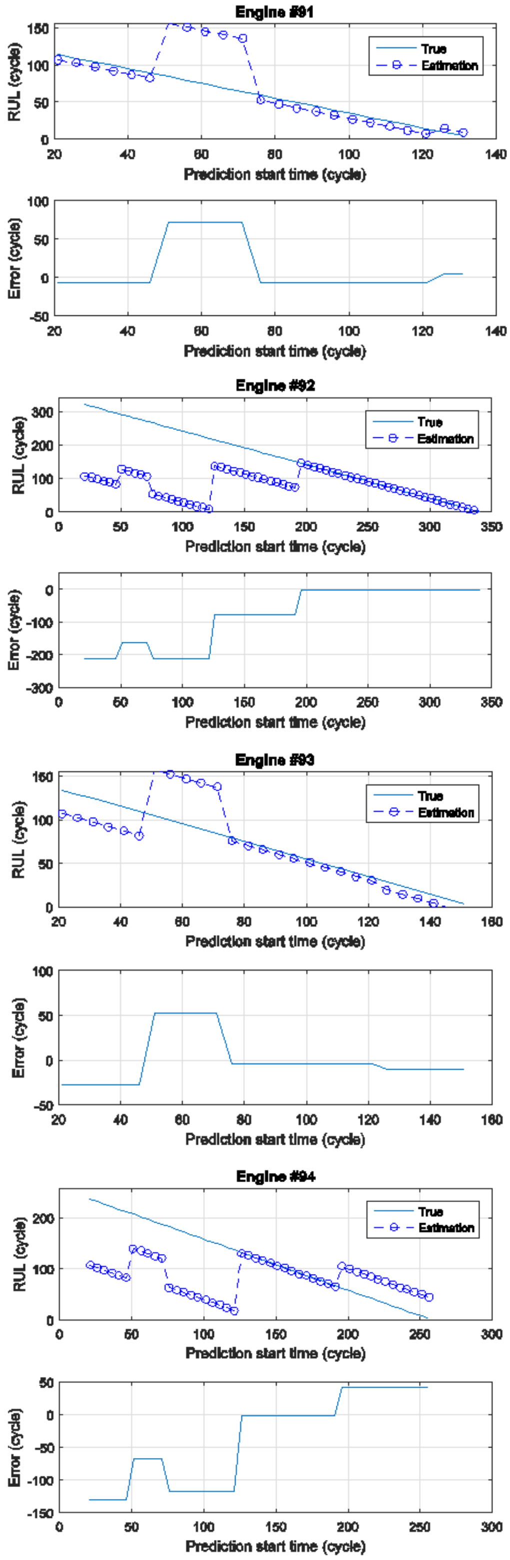
7. REFERENCES

1. Khezri R, Hosseini R, Mazinani M. "A fuzzy rule- based expert system for the prognosis of the risk of development of the breast cancer." ***International Journal of Engineering, Transactions A: Basics***. Vol 27, No 10, (2014), 1557-1564. doi: 10.5829/idosi.ije.2014.27.10a.09
2. Hu C, Zhou Z, Zhang J, Si X. "A survey on life prediction of equipment." ***Chinese Journal of Aeronautics***. Vol 28, No 1, (2015), 25-33. doi.org/10.1016/j.cja.2014.12.020
3. Hamidi H, Daraee A. "Analysis of pre-processing and post-processing methods and using data mining to diagnose heart diseases." ***International Journal of Engineering, Transactions A: Basics*** . Vol 29, No 7, (2016), 921-930. doi: 10.5829/idosi.ije.2016.29.07a.06
4. Amini M, Moharrami A, Poursaeidi E. "Failure probability and remaining life assessment of reheater tubes." ***International Journal of Engineering, Transactions B: Applications***. Vol 26, No 5, (2013), 543-552. doi: 10.5829/idosi.ije.2013.26.05b.11
5. Peng Y, Dong M, Zuo MJ. "Current status of machine prognostics in condition-based maintenance: a review." ***The International Journal of Advanced Manufacturing Technology***. Vol 50, (2010), 297-313. doi.org/10.1007/s00170-009-2482-0
6. Mao R, Zhu H, Zhang L, Chen A. "A new method to assist small data set neural network learning. InIntelligent Systems Design and Applications." ***IEEE ISDA'06.*** Vol 1, (2006), 17-22. DOI:10.1109/ISDA.2006.67
7. Ranasinghe GD, Lindgren T, Girolami M, Parlikad AK. "A Methodology for Prognostics Under the Conditions of Limited Failure Data Availability". ***IEEE Access.*** Vol 17, (2019), 173-179. DOI: 10.1109/ACCESS.2019.2960310
8. Li YG, Nilkitsaranont P. "Gas turbine performance prognostic for condition-based maintenance." ***Applied Energy***. Vol 86, No 10, (2009), 2152-2161. doi.org/10.1016/j.apenergy.2009.02.011
9. Xiongzi CH, Jinsong YU, Diyin TA, Yingxun WA. "A novel pf-lssvr-based framework for failure prognostics of nonlinear systems with time-varying parameters." ***Chinese Journal of Aeronautics***. Vol 25, (2012), 715-724. doi.org/10.1016/S1000-9361(11)60438-X
10. Tongyang LI, Shaoping WA, Jian SH, Zhonghai MA. "An adaptive-order particle filter for remaining useful life prediction of aviation piston pumps." ***Chinese Journal of Aeronautics***.Vol 31, (2018), 941-948. doi.org/10.1016/j.cja.2017.09.002
11. Diallo ON. "A data analytics approach to gas turbine prognostics and health management" (Doctoral Dissertation, Georgia Institute of Technology).
12. Caesarendra W, Widodo A, Yang BS. "Combination of probability approach and support vector machine towards machine health prognostics." ***Probabilistic Engineering Mechanics***. Vol 26, (2011), 165-173. doi.org/10.1016/j.probengmech.2010.09.008
13. Huang HZ, Wang HK, Li YF, Zhang L, Liu Z. "Support vector machine based estimation of remaining useful life: current research status and future trends." ***Journal of Mechanical Science and Technology***. Vol 29, (2015), 151-163. doi.org/10.1007/s12206-014-1222-z
14. Simon D. "A comparison of filtering approaches for aircraft engine health estimation." ***Aerospace Science and Technology***.Vol 12, (2008), 276-284. doi.org/10.1016/j.ast.2007.06.002
15. Lu F, Ju H, Huang J. "An improved extended Kalman filter with inequality constraints for gas turbine engine health monitoring." ***Aerospace Science and Technology***. Vol 30, (2016), 36-47. doi.org/10.1016/j.ast.2016.08.008
16. Ding C, Xu J, Xu L. "ISHM-based intelligent fusion prognostics for space avionics." ***Aerospace Science and Technology***. Vol 29, (2013), 200-205. doi.org/10.1016/j.ast.2013.01.013
17. Goebel K, Saha B, Saxena A. "A comparison of three data-driven techniques for prognostics." In 62nd Meeting of the Society for Machinery Failure Prevention Technology (Mfpt), (2008), 119-131.

18. Xu J, Wang Y, Xu L. "PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data." ***IEEE Sensors Journal***. Vol 14, (2014), 1124-1132. DOI: 10.1109/JSEN.2013.2293517
19. Moghaddass R, Zuo MJ. "An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process." ***Reliability Engineering & System Safety***. Vol 124, (2014), 92-104. doi.org/10.1016/j.ress.2013.11.006
20. Xiang Y, Liu Y. "Application of inverse first-order reliability method for probabilistic fatigue life prediction." ***Probabilistic Engineering Mechanics***. Vol 26, (2011), 148-156. doi.org/10.1016/j.probengmech.2010.11.001
21. Javed K. "A robust & reliable Data-driven prognostics approach based on extreme learning machine and fuzzy clustering" (***Doctoral Dissertation***).
22. Saxena A, Goebel K, Simon D, Eklund N. "Damage propagation modeling for aircraft engine run-to-failure simulation." ***IEEE InPrognostics and Health Management***, (2008). 1-9. DOI:10.1109/PHM.2008.4711414
23. Ramasso E, Rombaut M, Zerhouni N. "Joint prediction of observations and states in time-series: a partially supervised prognostics approach based on belief functions and knn. Networks". (2013). DOI: ff10.1109/TSMCB.2012.2198882
24. Ramasso E. "Investigating computational geometry for failure prognostics." ***International Journal of Prognostics and Health Management***. Vol 5, (2014).
25. Khelif R, Malinowski S, Chebel-Morello B, Zerhouni N. "RUL prediction based on a new similarity-instance based approach." ***IEEE InIndustrial Electronics (ISIE),*** (2014), 2463-2468. DOI: 10.1109/ISIE.2014.6865006
26. Yakout M, Elkhatib A, Nassef MG. "Rolling element bearings absolute life prediction using modal analysis." ***Journal of Mechanical Science and Technology***. Vol 32, (2018), 91-99. doi.org/10.1007/s12206-017-1210-1
27. Prasad SR, Sekhar AS. "Life estimation of shafts using vibration based fatigue analysis." ***Journal of Mechanical Science and Technology.*** Vol 32, (2018), 4071-4078. DOI: 10.1007/s12206-018-0806-4
28. Mohammadi E, Montazeri-Gh M. "Simulation of full and part-load performance deterioration of industrial two-shaft gas turbine." ***Journal of Engineering for Gas Turbines and Power***.Vol 136, (2014), 26-35. DOI: 10.1115/1.4027187
29. Dabaghi E, Kashanian H. "Feature dimension reduction of multisensor data fusion using principal component fuzzy analysis." ***International Journal of Engineering, Transactions A: Basics,*** Vol 30, No. 4, (2017), 493-499. doi: 10.5829/idosi.ije.2017.30.04a.06
30. Mahmoodian A., Durali M., Abbasain T., Saadat M., "Optimized Age Dependent Clustering Algorithm for Prognosis: A Case Study on Gas Turbines", ***Scientiairanica Transaction B: Mechanical Engineering***, (2020), (Articles in Press, Available Online) DOI:10.24200/SCI.2020.53863.3459
31. An D, Kim NH, Choi JH. "Statistical aspects in neural network for the purpose of prognostics." ***Journal of Mechanical Science and Technology***. Vol 29, (2015), 1369-1375. DOI: 10.1007/s12206-015-0306-8
32. Wang T, Yu J, Siegel D, Lee J. "A similarity-based prognostics approach for remaining useful life estimation of engineered systems." ***IEEE InPrognostics and Health Management***, (2008). 1-6. DOI: 10.1109/PHM.2008.4711421
33. Javed K, Gouriveau R, Zerhouni N., "Novel failure prognostics approach with dynamic thresholds for machine degradation." IEEE InIndustrial Electronics Society IECON 2013-39th Annual Conference of the IEEE, (2013) 4404-4409. DOI: 10.1109/IECON.2013.6699844

8. APPENDIX A

Prognostics results of engines #81-96 in different cycles with limited train data set are shown in Figure A-1.



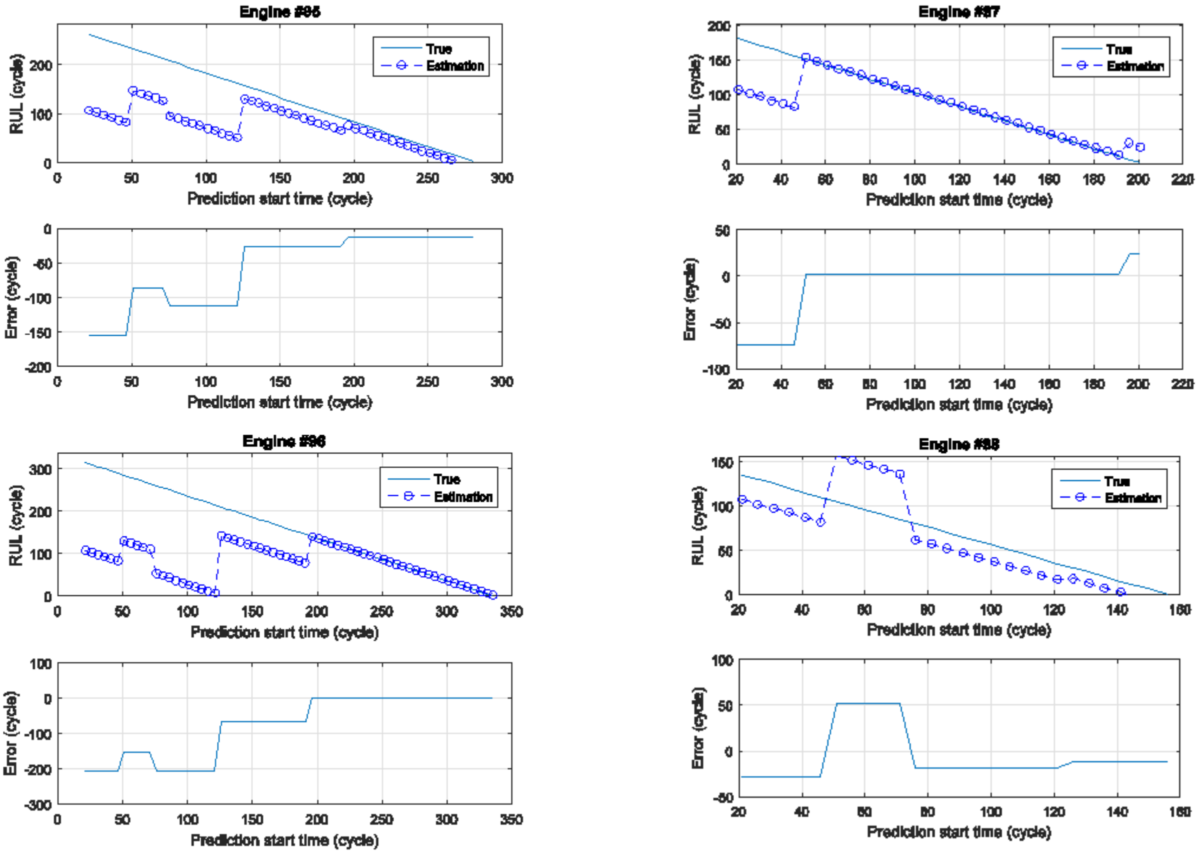


Figure A-1. Prognostics results of engines #91-98 in different cycles with limited train data set

Persian Abstract

چکیده

قابلیت اطمینان الگوریتم‌های پیش آگهی داده–پایه به شدت به حجم داده‌ها بستگی دارد. بنابراین در صورت محدودیت داده، برآورد عمر معمولاً قابل قبول نیست. برای حل این مشکل ، یک ساختار پیش آگهی مبتنی بر خوشه بندی عمر پیشنهاد شده است. این ساختار داده های آموزش را در سنین مختلف بازسازی کرده و در نتیجه حجم این داده‌ها را افزایش می‌دهد. از این جهت این روش برای مسائلی که با داده محدود مواجه هستند می‌تواند کارآمد باشد. در این تحقیق ، ابتدا عملکرد الگوریتم پیشنهادی در شرایط عادی بررسی می شود. متعاقبا عملکرد الگوریتم در شرایط محدودیت داده مطالعه می‌شود. برای این منظور، یک مطالعه موردی روی موتورهای توربوفن انجام می شود. نتایج حاصل نشان می‌دهد دقت روش پیشنهادی در شرایط عادی 71٪ و بهتر از روشهای دیگر بوده است. هنگامی که داده های آموزش به میزان یک سوم کاهش یافته، دقت پیش آگهی تقریباً بدون افت باقی مانده است. وقتی داده های آموزش از 100 به 5٪ کاهش یافته، دقت نتایج 26٪ افت کرده است. در مجموع، نتایج بدست آمده حاکی از مقاوم بودن الگوریتم پیشنهادی در شرایط محدودیت داده است.