Journal of Industrial and Systems Engineering
Vol. 12, No. 4, pp. 269-282
Autumn (November) 2019

JISE

Direct remaining useful life prediction based on multi-sensor information integrations by using evidence theory

Mohammad baharshahi[1], Mohammad Seyedhosseini[1*], Kamran Shahnaghi[1]

[1]*School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*

baharshahi@ind.iust.ac.ir, seyedhosseini@iust.ac.ir, shahanaghi@iust.ac.ir

Abstract

Estimation of remaining useful life (RUL) is one of most interesting subjects in prognostic and health management. Performing an analysis of the results of such estimation can increase the reliability and the safety of the system, and reduce the unnecessary costs. In this paper, a similarity-based combination method is proposed to combine several run-to-failure historical datasets in order to directly estimate the RUL. In this method, reference datasets are clustered and the initial RUL is calculated based on the artificial neural networks trained by the reference datasets. By using the extended Dempster-Shafer, the similarity between the initial RUL and the average RUL for each dataset is obtained. The proposed methodology is tested and validated on Commercial Modular Aero-Propulsion System Simulation (C-MAPSS), test-bed developed by NASA. The results of the evaluation show that the proposed method outperforms other methods in the literature.

Keywords: Dempster-Shafer Theory, information integration, remaining useful life

1-Introduction

Today, traditional maintenance approaches, such as preventive and corrective maintenance, are gradually being replaced by more advanced concepts such as reliability centered maintenance (RCM), condition based maintenance (CBM), structural health monitoring (SHM), and Prognosis and health management (PHM).

The most important achievement of the PHM is the capability of predicting the conditions of equipment degradation and estimating the remaining useful life (RUL). Accordingly, awareness of the RUL prevents practitioners from spending unnecessary costs and increases the availability of equipment (Ben Ali, Chebel-Morello, Saidi, Malinowski, & Fnaiech, 2015). An examination of the literature in this field suggests that prediction is more effective than diagnosis in reducing unexpected maintenance costs and increasing overall reliability diagnosis (Kunche, Chen, & Pecht, 2012; Lei et al., 2018; Van Tung & Yang, 2009). According to ISO-13381, the prognosis is to estimate the time and risk of failure in one or more existing assets and to predict future failures (Kunche et al., 2012). According to the literature of prognostics and health management, three approaches have been proposed to estimate the useful life of the equipment (Van Tung & Yang, 2009): 1) Data-driven; 2) Model-based, and 3) Hybrid approach.

Model-based approach (Physical failure approach) is based on predetermined mathematical models of equipment degradation considering health and degradation.

*Corresponding author
ISSN: 1735-8272,

Generally, in these models, the uptime cycles of the equipment are considered as functions of failure state. Additionally, RUL can be estimated by obtaining information from the failure states (Lei et al., 2018). However, the failure modes do not necessarily provide satisfactory explanations of the equipment degradation. Furthermore, a systematic, accurate, and reliable model for some types of equipment (especially when systems have a level of complexity) and its current conditions is not always available. To address these challenges, the data-driven approaches will be used. In these approaches, RUL is estimated by adjusting one or more characteristics that represent the equipment's degradation status. Subsequent approaches make a trade-off between accuracy and applicability. Often, the transit time, in which an indicator overruns the pre-determined threshold, is considered as the moment in which complete equipment breakdown occurs (Tobon-meja, Medjaher, Zerhouni, & Iso, 2010). The hybrid approach, which combines the outputs of both data-driven and model-based approaches, generally provides more reliable and accurate results (Lee et al., 2014). Despite the development of different methods in the data-driven approach, various challenges remain unsolved. Perhaps, the most important challenge is to determine the failure threshold. Interestingly, the uncertainty of the failure threshold poses a high level of unreliability to the problem. In other words, how long should the equipment feature behavior be predicted? Besides, in machine learning approaches, we generally encounter a large volume of learning data sets, each of which can explain a part of the current state of the equipment. Therefore, the question raised from this argument is: how can we fairly deal with the information obtained from various data sets?

The main contribution of this paper is to provide a framework for online determination of degradation to estimate RUL directly without determining any threshold. There are certain features that make this model distinct from those available in literature and give it an edge over them. In order to introduce some of the novel features of this model, it must be mentioned that this model is the estimations are online, no thresholds are required to be determined, the whole data including the similarities and the differences of the test and learning data are taken into account and considered, and a novel index is defined in the similarity of the dataset which improved the model accuracy. This paper is then organized as follows: In Section 2, the literature of data-driven approaches has been reviewed. A majority of the reviewed studies are related to methods of integrating evidence, identifying similar patterns, measuring the similarity of the routes, and its applications in RUL estimation. In Section 3, the proposed framework is described in detail. The foundations of the proposed method are presented during the description of the model. In Section 4, the proposed method has been implemented. Results and discussions about validation and evaluation are presented in Section 5. Eventually, this paper ends with a conclusion.

2-Related works

Generally, there are three basic steps for calculating RUL in data-driven approaches. 1. Determination of a health index representing the condition of equipment degradation. 2. Determining the condition of equipment degradation. 3. Determining the RUL by predicting health index, equipment degradation status, and time to failure of the learning data sets. Some researches take all three steps to determine RUL; however, others disregard the second step. The following methods are used to reduce the dimensions of data: extracting and selecting features such as PCA (Moghaddass & Zuo, 2014), logistic regression (Yan, Koc, & Lee, 2004), linear regression (Sun, Zuo, Wang, & Pecht, 2012; T. Wang, Yu, Siegel, & Lee, 2008), exponential regression (Saxena, Goebel, Simon, & Eklund, 2008), weighted average methods (Liu, Gebraeel, & Shi, 2013; Ramasso, Rombaut, & Zerhouni, 2013), linear transformation of data forms (Hu, Youn, Wang, & Yoon, 2012; P. Wang, Youn, & Hu, 2012; Xi, Jing, Wang, & Hu, 2014), feature selection based on maximum accuracy of clustering (decision tree) (Ishibashi & Júnior, 2013), and empirical analysis (EMD) (Mosallam, Medjaher, & Zerhouni, 2016).

Furthermore, the following solutions have been proposed to determine the degradation status: Operation-regime partitioning (Heimes, 2008; T. Wang et al., 2008), k-means clustering (Ramasso et al., 2016; Ramasso & Denoeux, 2014; Zemouri & Gouriveau, 2010), Gaussian mixture model (Lin, Chen, & Zhou, 2013; Ying Peng & Dong, 2011), evolving extended TS system (exTS) (El-Koujok, Gouriveau, & Zerhouni, 2011; Ramasso & Gouriveau, 2014), Mahalanobis distance based classifier (MD) (Tamilselvan, Wang, & Wang, 2012), KNN (Ramasso et al., 2013), hierarchical clustering (Mosallam et al., 2016), and

Fuzzy Clustering (Ramasso & Gouriveau, 2014). Finally, predictive methods are used to determine the RUL. In some studies, the health index has been predicted (Nie & Wan, 2015; Yu Peng, Wang, Wang, Liu, & Peng, 2012a), and in others, the duration of collapse is directly considered (Jianzhong, Hongfu, Haibin, & Pecht, 2010). Exponential regression methods (T. Wang et al., 2008), artificial neural network of multilayer perceptron (Jianzhong et al., 2010), Sparse Bayesian (P. Wang et al., 2012), Bayesian linear regression (Hu et al., 2012), backup vector machine (Xu, Wang, & Xu, 2014), case-based reasoning (Ramasso, 2014), instance-based learning (Khelif, Malinowski, Chebel-Morello, & Zerhouni, 2014), recurrent radial basis function network (RRBF network) (Zemouri & Gouriveau, 2010), hidden Markov model (Giantomassi et al., 2011), Echo State Network (Yu Peng, Wang, Wang, Liu, & Peng, 2012b), State space model (Sun et al., 2012), Deep Belief Network(Tamilselvan et al., 2012), and Fuzzy Rule-Based System (Ishibashi & Júnior, 2013) are methods that have been considered in the literature for the third step. Researchers use different approaches when facing more than one learning sets. Some researchers use the whole collection of the learning data in a container and use the clustering algorithm to determine the state of deterioration or training and fit the prediction model (Javed, Gouriveau, & Zerhouni, 2015). But in some other studies, there are some ways to integrate information. This integration has been considered at the level of predictors (the combination of weak artificial networks using an approach like AdaBoost (Jianzhong et al., 2010)) or at the level of RUL. Integration at the level of RUL is based on the weight composition of the obtained RULs from each learning set, namely: $\mathrm{RUL_f} = \sum_{i=1}^{N} \mathrm{w_i RUL_i}$, with $\mathrm{RUL_f}$ estimated to be the useful life of the final remainder. $\mathrm{RUL_i}$ is the remaining useful life based on each of N learning sets and $\mathrm{w_i}$ is the assigned weight to the RUL, which is obtained from the i-th learning data set. Various approaches have been proposed to determine $\mathrm{w_i}$. These approaches are shown in Table 1. Considering references (T. Wang et al., 2008) and (Khelif et al., 2014), the Euclidean-based function was used, the most specific path for each learning unit was determined by the unit of testing, and the RUL is calculated according to the sum of RULs obtained from each learning unit. Accordingly, weights will be calculated from the similarity scale. Considering reference (Ramasso, 2014), which is similar to the previous references, RUL is derived from the Euclidean-like weight composition. However, the weight composition of the minimum and maximum RULs obtained from the learning set is used. Considering reference (P. Wang et al., 2012), based on the model trained by each learning set, prediction has been made to a certain time horizon and the weight of the RULs' composition is based on the Euclidean similarity between the predicted path and the achieved learning set. Considering references (Hu et al., 2012) and (Xu et al., 2014), the use of precision prediction and deviation of predictive error, obtained after teaching the prediction model for each learning set, is considered as the weight of each set of Learning method data.

Table 1. Information integration approaches for RUL prediction

Ref.	Information integration on RUL level	Weighting method (RUL coefficients in linear composition)
(Khelif et al., 2014; T. Wang et al., 2008)	Linear combination of RULs	Euclidean distance of the extracted health index with the reference data sets (learning datasets)
(Ramasso, 2014)	Linear combination of Maximum and Minimum RULs	Euclidean distance of the extracted health index with the reference data sets (learning datasets)
(P. Wang et al., 2012)	Linear combination of RULs	Determine the combination weight using the Euclidean distance of the predicted path for the health index with the reference data sets (learning datasets)
(Hu et al., 2012; Xu et al., 2014)	Linear combination of RULs	- Euclidean distance of the extracted health index with the reference data sets (learning datasets) - Prediction using the relevant learning data set, and extraction of prediction accuracy - Prediction using the relevant learning data set, and extraction of prediction diversity error

In the proposed method of this research, PCA was used to determine a health index. By aggregation of learning sets, k-means clustering is used to determine the general degradation status. Moreover, the forecasting process is directly carried out for the duration of the decomposition; as a result, there is no need to determine the threshold of failure. This process is performed for each learning set. Thus, RUL needs to be integrated. In order to determine the weights of the composition, the extended Dempster-Shafer theory is used.

3-The proposed Approach

Figure 1 shows the details of the proposed method. The RUL estimation takes place in two phases: Learning and Online RUL Estimation. In the learning phase, prediction is provided, predictors are taught, and degradation situations are determined. In the online RUL estimation phase, predictors and degradation conditions are used to estimate the RUL. In the following, the steps of each of these phases are discussed in detail according to figure 1.

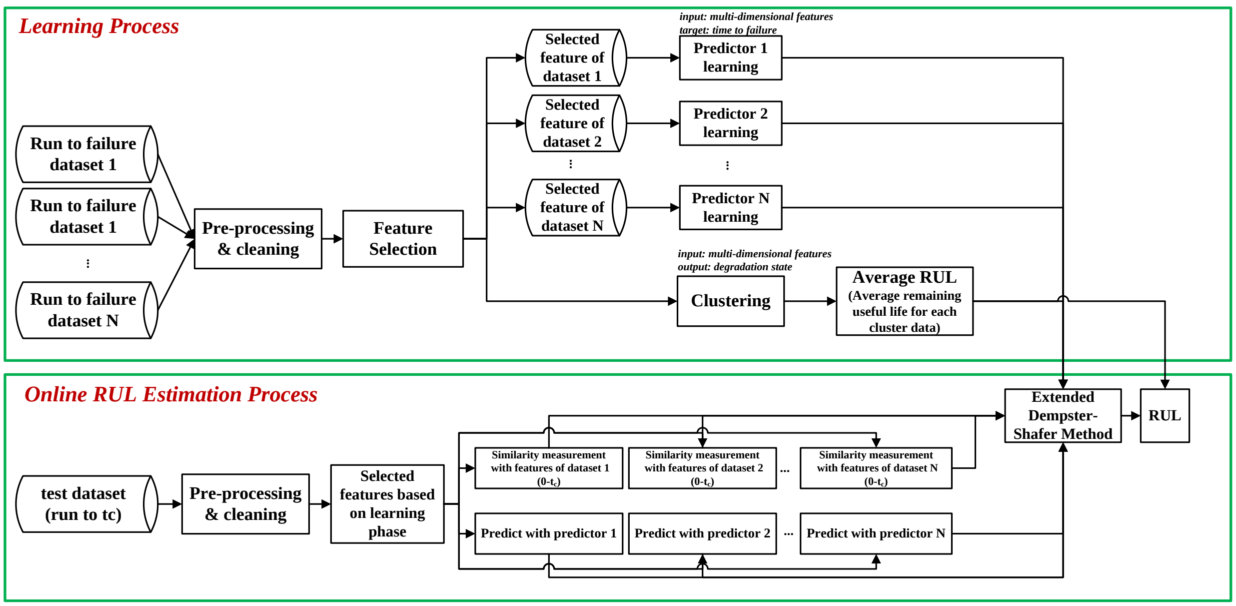

Fig1. The proposed process of RUL estimation

3-1-Learning phase

3-1-1-Acquisition and data collection

In this paper, determination of RUL is based on the acquisition and collection of run to failure data. In other words, the required dataset, including the N implementation that runs to the failure dataset of the monitoring sensors equipment, is related to a specific type of equipment.

The structure of the used data is a dimensional T × D matrix, where T is the time or cycle run to complete failure and D is the number of the sensors that represent the information of the equipment status. In the learning phase, this data set is used to train predictor and cluster, determine the failure states, and to estimate the average RUL of each cluster which is equipment situation here.

3-1-2-Preprocessing data and selecting features

In the preprocessing step, data are normalized and the outlier data are removed. Normalization of data helps us to improve the clustering and prediction process (Zhang, 1994) and eliminates the outlier data caused by unwanted events (e.g. operator error, electric fluctuations, Unexpected atmospheric conditions, etc.) from the analysis set.

The following relation is used for data normalization:

$$y_{d,t}^{i}=\frac{x_{d,t}^{i}-min(x_{d,1}^{i},x_{d,2}^{i},\dots,x_{d,T}^{i})}{max(x_{d,1}^{i},x_{d,2}^{i},\dots,x_{d,T}^{i})-min(x_{d,1}^{i},x_{d,2}^{i},\dots,x_{d,T}^{i})}\ ,\boldsymbol{i}=\mathbf{1},\dots,\boldsymbol{N}\ \boldsymbol{d}=\mathbf{1},\dots,\boldsymbol{D} \tag{1}$$

In the above equation, $x_{d,t}^{i}$ is the data point of the learning set i from the sensor d at time t, and $y_{d,t}^{i}$is the normalized data of the learning set I from the sensor d at time t. After data preprocessing, the core features of the data set are extracted to minimize the error in the clustering process and decrease the computational cost by reducing the dimensions of the dataset. Here, the PCA method is used to determine the core features. The PCA is an orthogonal linear transformation that transfers the data to a new coordinate system; consequently, the largest data variance will be placed on the first coordinate axis, the second largest variance will be placed on the second coordinate axis, and so on.

This transformation is made using the D-dimensional coefficients of weights $W=(w_1,w_2,\dots,w_D)$, where each linear vector $y_t^i=(y_{1,t}^i,y_{2,t}^i,\dots,y_{D,t}^i)$ is mapped into the new vector of the main component points, and D is obtained as a new property. Having transformed the PCA, the components with explanatory variances greater than 1 are obtained, and F (F≤D) core feature is selected.

3-1-2-1-Fitness of predictive model

In this step, a predictive model is fitted to each of the N learning datasets. Each of these datasets can represent one of the failure modes and degradation trends. For this reason, their data can be used to predict the future trends. Here, a multi-layer perceptron neural network model is used for fitting. F, the main component obtained in the previous stage, is considered as the model input and RUL (T-t) is considered as the output (target). Therefore, the N trained neural network (NET_i) is obtained for the N learning set. By using the trained network sets at the testing stage, the useful life of the new data set will be predicted and the details will be described in the testing phase.

3-1-2-2-Clustering and RUL estimation

Each data point in each learning dataset indicates a specific degradation state. At normal conditions, starting points are usually within the predetermined control limits. With increasing equipment degradation in the course of time, data points have been located in different channels. By clustering the set of data points, a tag can be assigned to each of these data points, indicating the state of the equipment at that specific time. In other words, the basic assumption about clustering is that the RUL of data points in a cluster are similar to each other. All F features in N learning datasets were clustered into U degradation mode of equipment. After clustering, the average useful life of the remaining cluster was calculated as follows:

$$ARUL_u^i=\frac{\sum_{l=1}^{L_u^i}(T^i-t_{u,k}^i)}{L_u^i} \tag{2}$$

where $ARUL_u^i$ is the average RUL of U-th cluster for i-th learning dataset, T^i is the failure time of the learning dataset i, $t_{u,l}^i$ is the current time of the l-th data in the U-th shared cluster set in the i-th learning data set, and L_u^i is the number of data points in the u-th cluster in the i-th learning data set.

3-2-Testing phase

3-2-1-RUL online estimation procedure

At this stage, the procedure for obtaining information in the learning process, which is used for estimation of RUL of the equipment, is described. It is assumed that the data set contains the same data as the available learning process; this information is recorded by sensors up to the current time t_n and the goal is to calculate the RUL of the equipment. The obtained data is called “test dataset”. Similar to the first step of the learning phase, pre-processing operation is performed on the test data. Then, using the coefficients obtained from the PCA algorithm in the learning process and the dimensions of the extracted features, in which the explanatory variance is greater than 1, the new core feature to the testing data is selected.

3-2-2-Choosing the reference dataset

After selecting the test data feature from the N learning sets, R sets are selected as the reference (controls) sets to determine the initial RUL of the r-th reference: $\boldsymbol{IRUL_r}$.The selection criterion is based on a similarity of the Euclidean distance between the test data features and each of the N learning sets. This criterion is calculated as follows:

$$\boldsymbol{S^i} = \frac{\mathbf{1}}{\frac{\sum_{t=1}^{min(T^i,t_n)} \sqrt{\left(V_{1t}^i - P_{1t}\right)^2 + \left(V_{2t}^i - P_{2t}\right)^2 + \dots + \left(V_{ft}^i - P_{ft}\right)^2}}{min(T^i,t_n)}} \tag{3}$$

In the above equation, $\boldsymbol{S^i}$ is a measure of the similarity between the test data and the i-th learning dataset, $\boldsymbol{V_{ft}^i}$ is the f-th feature of the i-th learning data at time t, $\boldsymbol{P_{ft}}$ is the f-th feature of the test data in time t, $\boldsymbol{T^i}$ is the learning lifetime of the i-th learning data and $\boldsymbol{t_n}$ is the current time or the lifetime of the test data. After determining the value of $\boldsymbol{S^i}$ for each learning data, the R learning set that has the highest $\boldsymbol{S^i}$ value is selected as the reference set.

3-2-3-Information integration and RUL estimation

After choosing R reference sets, the networks associated with the selected reference sets $(\boldsymbol{NET_r}, \boldsymbol{r} = \mathbf{1}, \dots, \boldsymbol{R})$ are extracted and the RULs are estimated at time $\boldsymbol{t_n}$.

Therefore, R remaining useful lives will be obtained from the test data. In order to integrate the information obtained from the selected references, the improved Dempster Schaffer method, presented by Yager in 1987, was used (Yager, 1987). Evidence theory was introduced by Dempster (Dempster, 1967) and expanded by Schafer (Shafer, 1976). This theory is important in discussing current ideas about a situation or a system of situations. Although ideas about the events are not the same, one can examine and combine the current evidence of the situation with the help of this theory. Moreover, Dempster Schaeffer's theory is based on a belief that originates from evidence and relates to the classical probability model.

Yager defines a new function, called the ground probability mass assignment, in which q suppresses the null value basic assumption of the mass function, namely: q (∅) ≥0. A positive value for the probability of occurrence of null state means that the probability that a witness does not choose any state make mistake or be opposed to others, is greater than zero.

Then, by introducing the parameter $\boldsymbol{\alpha_i}$, as an importance factor, the weight of the confidence to i-th witness was determined against others. Therefore, the Yager's improved mass function was defined as $\boldsymbol{m(A)} = \boldsymbol{\alpha_i} \times \boldsymbol{O_i(A)}$, where $\boldsymbol{O_i(A)}$ is the estimate of the i-th evidence from event A. Additionally, Yager collected the possible errors and the contradiction between the evidence in a set called $\boldsymbol{\Theta}$, such that $\boldsymbol{\Theta_i} = \mathbf{1} - \boldsymbol{\alpha_i}$. Based on this parameter, there is no indication of the possibility of states and the estimate Θ of each control being shared with all estimates of other states. With respect to the improved mass function, the combination of the mass function is obtained from the following relationship:

$$\boldsymbol{m(A)} = \frac{\boldsymbol{q(A)}}{\mathbf{1} - \boldsymbol{q}(\emptyset)} = \frac{\sum_{\cap A_i - A}[\boldsymbol{m_1(A_1)} \times \boldsymbol{m_2(A_2)} \times \dots \times \boldsymbol{m_i(A_i)} + \boldsymbol{\Theta_i m_i}]}{\mathbf{1} - \boldsymbol{q}(\emptyset)} \tag{4}$$

According to modified Dempster Shafer theory and the rules of its combination, it is assumed that the M learning set is evident and determines the overall conditions of the equipment degradation. In this case, $\boldsymbol{O_i(U)}$ is the normalized estimation of the i-th evidence (learning dataset i) from the event u (cluster u) based on the inverse distance of the initial RUL interval (IRUL) from average RUL (ARUL). This is shown in the below equation:

$$O_i(cl_u) = \frac{\frac{1}{|IRUL_i - ARUL_u^i|}}{\sum_{u=1}^{U} |IRUL_i - ARUL_u^i|}, i = 1, \ldots, N, u = 1, \ldots, U \tag{5}$$

Moreover, in the proposed method, S^i as a similarity measure is equal to importance factor α_i that indicates the level of confidence in the i-th witness (Learning Data Set i) and $\Theta_i = 1 - \alpha_i = 1 - S^i$ represents the probable errors and the contradiction between the ideas of learning sets about the state of equipment degradation. In other words:

$$m(cl_u) = \alpha_i \times O_i(u) = S^i \frac{1/D_u^N}{\sum_{u=1}^{U} D_u^N} \tag{6}$$

Therefore, the composition of the mass function is calculated from the following relation:

$$m_{1,..,M}(cl_u) = \frac{q(cl_u)}{1-q(\emptyset)} = \frac{\sum_{\cap cl_u - cl_U}[m_1(cl_1) \times m_2(cl_2) \times \ldots \times m_R(cl_u) + \Theta_r m_r]}{1 - \sum_{\cap_{i=1}^{M} E_i = \Phi} m_1(cl_1).m_2(cl_2) \ldots m_R(cl_u)} \tag{7}$$

The value of $m_{1,..,M}(cl_u)$ is used as the u-*th* cluster weight to determine the final RUL

$$m_{1,..,M}(cl_u) = \text{weight}_u$$

Therefore, weighted average of the RUL of the M selected reference is inserted to improve the Shafer Dempster method and calculate the final RUL:

$$\text{RUL} = \sum_{u=1}^{U} \left(\text{weight}_u \times \frac{\sum_{i=1}^{M} ARUL_u^i}{M} \right) \tag{8}$$

Figure 2 schematically demonstrates how the (weight_u) is calculated.

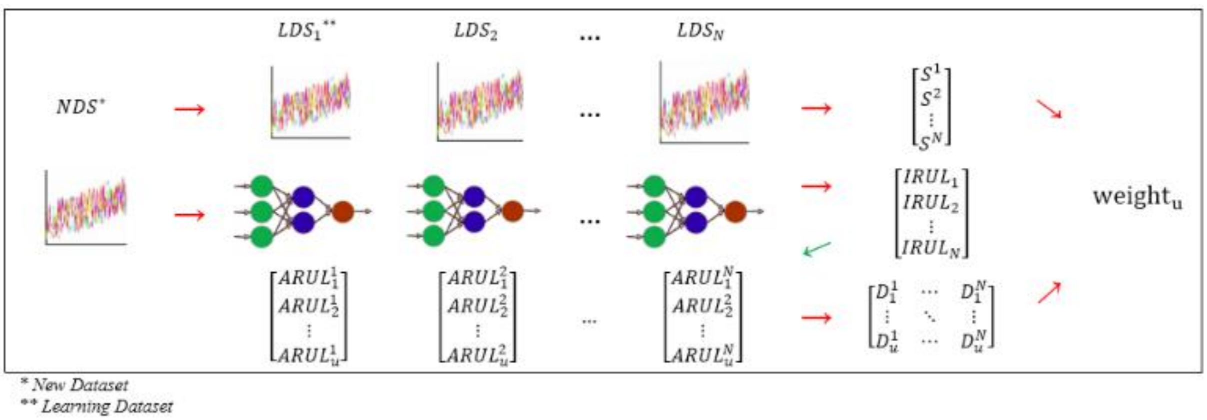

Fig 2. Weighting method of RUL combination

In the following, a case study of the obtained results and a comparison of the proposed method with other classical methods are presented.

4-Application and results

The proposed method was implemented in MATLAB 2016 software and the turbofan engine data set located on the site of the NASA's prediction data base was used. In the next section, the referred datasheet is described first, and then the results of implementing the method of clustering on this data are presented.

4-1- Datasets

The dataset consists of several multi-dimensional signal time series that are obtained by a simulation model developed on the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS). A total of 26 signals have been generated. 21 signals consist of sensor data. The next 3 signals show the specifications of the operating conditions. The remaining 2 signals indicate the motor ID and the number of cycles. Table 2 shows the details of the 21 sensor's signal.

Table 2. Name and unit of the sensors

symbol	description	Unit
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
Epr	Engine pressure ratio (P50/P2)	---
Ps30	Static pressure at HPC outlet	psia
Phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	---
farB	Burner fuel-air ratio	---
htBleed	Bleed Enthalpy	---
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s

In each time, the series relates to a different engine from the same complex system. Each engine contains various components such as compressor, turbine, etc. Figure 3, illustrates the main components of the engine model and indicates how sub-parts are assembled. For more information about this system, refer to (Saxena et al., 2008). Engines worked normally at first; however, as time passed, their degradation progressed and eventually, they failed completely. The dataset is characterized by four operating conditions that are presented in 9 text files.

For each of the conditions, three train files, a test, and an actual RUL were provided. Each learning and testing dataset contains several similar units. In the learning data, the recorded signals (measurements) start at a similar degradation level, which is considered to be a healthy state, and stop when a failure occurs. The test data is incomplete, meaning that the time series was cut before the complete failure. The goal is to predict the useful life of the remaining units of this test data set. The real RUL data includes the actual RUL values of the test units. The details of the dataset are shown in table 3:

Table 3. C-MAPSS dataset

Data type	Operating condition 1	Operating condition 2	Operating condition 3	Operating condition 4
Learning dataset	TRAIN-FD001: 259 learning unit	TRAIN-FD002: 260 learning unit	TRAIN-FD003: 100 learning unit	TRAIN-FD004: 249 learning unit
Testing dataset	TEST-FD001: 100 test unit	TEST-FD002: 259 test unit	TEST-FD003: 100 test unit	TEST-FD004: 248 test unit
Actual RUL	RUL-FD001: 100 test unit	RUL-FD002: 259 test unit	RUL-FD003: 100 test unit	RUL-FD004: 248 test unit

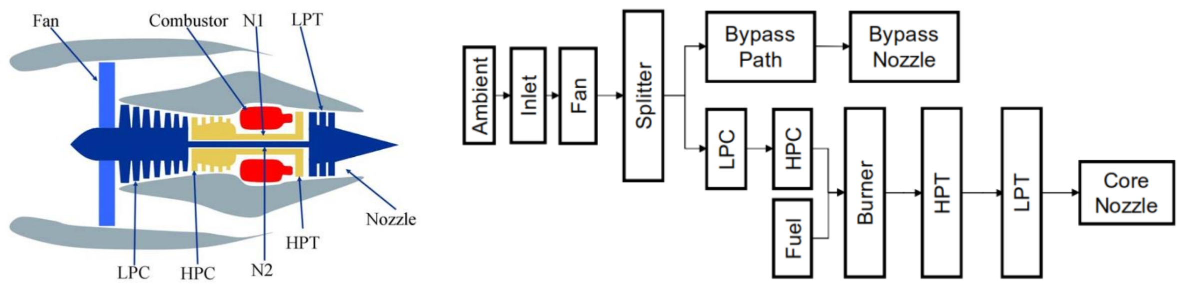

Fig 3. Turbo fan engine (left), diagram of engine component (right)

The proposed process of this research is used for the learning phase from the learning data, and the test data is used to evaluate the RUL estimation. Sensor data is nonlinear and noisy, and the life cycle of the units is in a wide range (between 128 and 362 cycles). Therefore, these conditions make the prediction of RUL more difficult.

4-2-Models under comparison

To evaluate the method proposed in this paper, two models are compared below.
1. The proposed integrated model of research
2. The proposed model of research, regardless of the integration approach and prediction using the network trained for the most similar dataset to the test dataset. The latter method is called “non-combinational” in the rest of the paper. (section 3-2-1)

Furthermore, the results of these two models have been compared with the results of other researches using the first-state data set of C-MAPSS (first condition of CMAPSS dataset).

4-3-Evaluation criteria

In order to evaluate and compare the performance of the proposed method, the following criteria have been used:

1) Score: This criterion is proposed by the data provider. This criterion assigns, for any RUL prediction, the s_n criterion formulation as follows:

$$s_n = \begin{cases} e^{-\frac{(r_n - \hat{r}_n)}{10}} - 1 & \text{if } (r_n - \hat{r}_n) \leq 0 \\ e^{+\frac{(r_n - \hat{r}_n)}{13}} - 1 & \text{if } (r_n - \hat{r}_n) > 0 \end{cases} \tag{9}$$

In the formula above, r_n is the RUL of the n-th unit and $\hat{r}_n$ is the estimated RUL of the n-th unit. The total score is obtained from the total s_n of all test units:

$$S = \sum_{n=1}^{N} s_n \tag{10}$$

Smaller values for S are better.

2) Performance: This criterion evaluates performance as a percentage of the correct predictions. According to this criterion, a prediction is correct if the forecast error $E = r_n - \hat{r}_n$ is in the interval $I = [-10,13]$. In other words, the forecast can be sooner or later with respect to the real RUL. Therefore, if P is the number of the predictions, the performance criterion is computed as follows:

$$\text{performance} = \frac{P}{N} \times 100 \tag{11}$$

5-Results and discussion

In figures 4 and 5, the actual RUL, predicted RUL, and histogram of the predictive error are presented for conditions 1. The left forms are from the proposed method and the right forms belong to the non-combinational method. As observed in the histogram, the distribution of predictive errors in the non-combinational method shows a skewness to the left, according to the error equation $E = r_n - \hat{r}_n$. This asymmetry indicates the tendency of model to overestimate the RUL.

Generally, an over-estimation for RUL (error -E) is worse than an underestimation (E + error), which is presented in relations (18) and (20). However, histograms of the proposed method are more symmetric. Moreover, the negative tail of the histogram for the non-combinational method is higher than the histograms of the proposed method and shows larger errors in the overestimate.

Due to the consideration of more possibilities for predictive references and the combination of degradation states in these references in the proposed method, the predicted results are more symmetric. The above mentioned results are shown by the score and performance points.

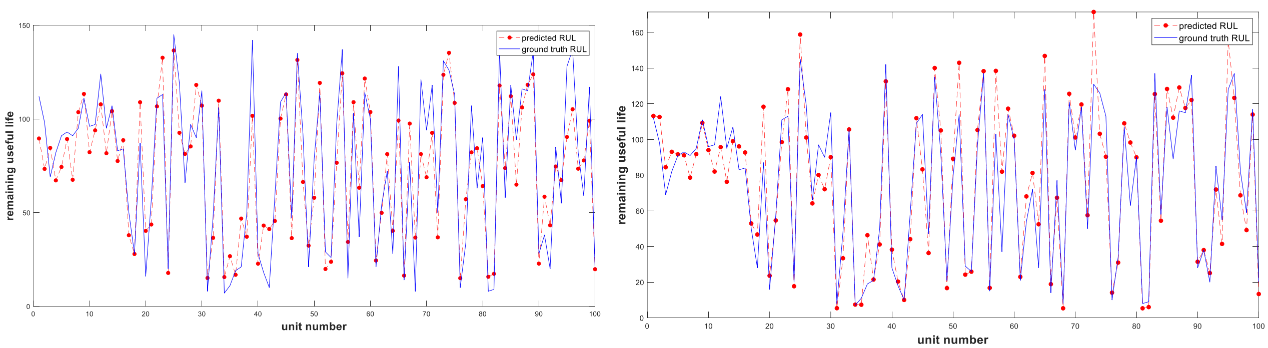

Fig 4. Ground truth Vs. predicted RULs for operational condition 1 proposed method (left) and no-combination (right)

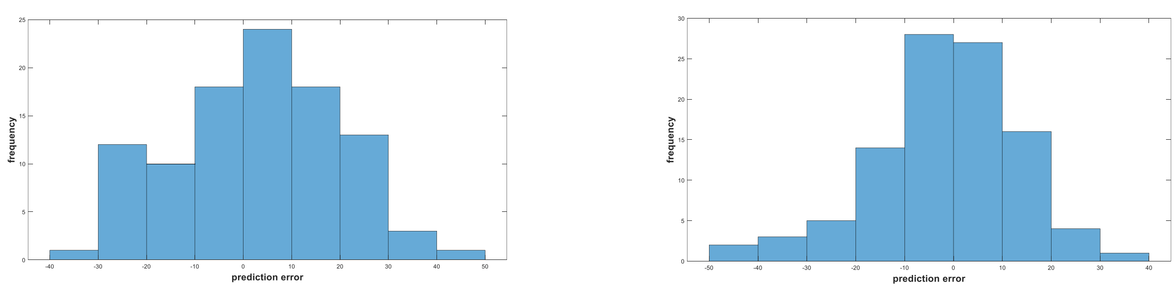

Fig 5. Histogram of errors for operational condition 1 - proposed method (left) and no-combination (right)

In table 4, the results of scoring and performance criteria for the proposed method, the non-combinational method, and the reference methods (by using the C-MAPSS data set) are presented. Additionally, in each of the four data set conditions described in section 4.1, the proposed method has a better result in terms of both scoring and performance criteria. The first conditions of data set were used to evaluate the proposed model in references mentioned in table 5. The proposed method has acceptable performance in two benchmark scores and performance.

Table 4. Score and performance comparison with and without integration in 4 condition scenarios

Proposed Method	**Score**				**Performance**			
	C 1	C 2	C 3	C 4	C 1	C 2	C 3	C 4
With integration	398.1	1530	593.40	1163.9	73.1	38.22	42	26.69
Without integration	832.6	1491.7	1278.2	1411.3	53.7	42.7	24.6	19.9

Similar to this research, reference (Khelif et al., 2014) provides a similarity-based approach. In (Khelif et al., 2017), the RUL has been directly calculated by fitting the SVR model to learning data set and has better results than other researches as well. The proposed method of this study employs an integrated approach based on direct similarity between estimated RUL and the actual RUL of clusters to demonstrate the final probability of the current test present in each of the clusters. Moreover, the method considers the similarity between the test data set and the choice of the most similar data set and determines the state of degradation based on clustering.

Table 5. score and performance comparison for proposed method with and without integration

Method	Ref. [61]	Ref. [62]	Ref. [51]	Ref. [63]	Proposed method	Proposed method without integration
Score	1046	N/A	N/A	448.7	398.1	832.6
Performance	48	53	54	70	73.1	50.7

6-Conclusion

In this paper, a data-driven method is proposed based on a new approach to the integration of sensor data and the condition of equipment. The proposed method can be implemented for samples that have a significant number of learning run to failure data. The PCA method was used to select features, k-means method for clustering, and ANN method for prediction. After selecting the features, an artificial neural network was fitted and stored on each learning dataset. On the other hand, the average RUL for each dataset was estimated by clustering all the learning data sets.

After selecting the attributes, the similarity between acquired data set and each learning set was determined to estimate the RUL of the equipment. Then, a subset of the most similar collection of learning data is selected and the RUL is estimated by using a related network. Based on the distance of the obtained RUL to the calculated average RUL for each learning cluster, the probability of belonging to each cluster and each learning set is calculated as a reference. Based on the Yager modified algorithm for the rules of integrating, the final probabilities of each cluster are determined and used as the weight for the average RUL in each cluster. Finally, the RUL is estimated based the weighted RUL value.

The proposed method has been implemented on C-MAPSS turbofan data set and the results show that the performance of this method in case of integration outperforms the prior researches. The advantages of the proposed method are integration of a variety of information and utilizing collections of datasets to predict RULs directly, without the need to consider a numerical threshold.

In addition, the other results of the proposed method culminate in the ability to determine the state of equipment degradation based on the integration of sensor information and clustering. Interested researchers can use other fitness models such as SVR, other similarity-determining methods such as non-spatial spacing criteria, and other rules of evidence combination such as fuzzy integration laws of Inagaki (Inagaki, 1991) and zhang (Zhang, 1994) for future development.

References

Ben Ali, J., Chebel-Morello, B., Saidi, L., Malinowski, S., & Fnaiech, F. (2015). Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing*, *56*. https://doi.org/10.1016/j.ymssp.2014.10.014

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 325–339.

El-Koujok, M., Gouriveau, R., & Zerhouni, N. (2011). Reducing arbitrary choices in model building for prognostics: An approach by applying parsimony principle on an evolving neuro-fuzzy system. *Microelectronics Reliability*, *51*(2), 310–320. https://doi.org/10.1016/j.microrel.2010.09.014

Giantomassi, A., Ferracuti, F., Benini, A., Ippoliti, G., Longhi, S., & Petrucci, A. (2011). Hidden Markov model for health estimation and prognosis of turbofan engines. In *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 681–689). American Society of Mechanical Engineers.

Heimes, F. O. (2008). Recurrent neural networks for remaining useful life estimation. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on* (pp. 1–6). IEEE.

Hu, C., Youn, B. D., Wang, P., & Yoon, J. T. (2012). Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, *103*, 120–135.

Inagaki, T. (1991). Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. *IEEE Transactions on Reliability*, *40*(2), 182–188.

Ishibashi, R., & Júnior, C. L. N. (2013). GFRBS-PHM: A genetic fuzzy rule-based system for phm with improved interpretability. *PHM 2013 - 2013 IEEE International Conference on Prognostics and Health Management, Conference Proceedings*. https://doi.org/10.1109/ICPHM.2013.6621419

Javed, K., Gouriveau, R., & Zerhouni, N. (2015). A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. *IEEE Transactions on Cybernetics*, *45*(12), 2626–2639.

Jianzhong, S., Hongfu, Z., Haibin, Y., & Pecht, M. (2010). Study of ensemble learning-based fusion prognostics. In *Prognostics and Health Management Conference, 2010. PHM'10.* (pp. 1–7). IEEE.

Khelif, R., Chebel-Morello, B., Malinowski, S., Laajili, E., Fnaiech, F., & Zerhouni, N. (2017). Direct Remaining Useful Life Estimation Based on Support Vector Regression. *IEEE Trans. Industrial Electronics*, *64*(3), 2276–2285.

Khelif, R., Malinowski, S., Chebel-Morello, B., & Zerhouni, N. (2014). RUL prediction based on a new similarity-instance based approach. In *Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium on* (pp. 2463–2468). IEEE.

Kunche, S., Chen, C., & Pecht, M. (2012). A review of PHM system's architectural frameworks. In *The 54th Meeting of the Society for Machinery Failure Prevention Technology, Dayton, OH*.

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, *42*(1), 314–334.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, *104*, 799–834.

Lin, Y., Chen, M., & Zhou, D. (2013). Online probabilistic operational safety assessment of multi-mode engineering systems using Bayesian methods. *Reliability Engineering & System Safety*, *119*, 150–157.

Liu, K., Gebraeel, N. Z., & Shi, J. (2013). A Data-Level Fusion Model for Developing Composite Health Indices for Degradation Modeling and Prognostic Analysis, *10*(3), 652–664.

Moghaddass, R., & Zuo, M. J. (2014). An integrated framework for online diagnostic and prognostic health monitoring using a multistate deterioration process. *Reliability Engineering and System Safety*, *124*, 92–104. https://doi.org/10.1016/j.ress.2013.11.006

Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, *27*(5), 1037–1048.

Nie, Y., & Wan, J. (2015). Estimation of remaining useful life of bearings using sparse representation method. In *Prognostics and System Health Management Conference (PHM), 2015* (pp. 1–6). IEEE.

Peng, Y., & Dong, M. (2011). A hybrid approach of HMM and grey model for age-dependent health prediction of engineering assets. *Expert Systems with Applications*, *38*, 12946–12953. https://doi.org/10.1016/j.eswa.2011.04.091

Peng, Y., Wang, H., Wang, J., Liu, D., & Peng, X. (2012a). A modified echo state network based remaining useful life estimation approach. In *Prognostics and Health Management (PHM), 2012 IEEE Conference on* (pp. 1–7). IEEE.

Peng, Y., Wang, H., Wang, J., Liu, D., & Peng, X. (2012b). A modified echo state network based remaining useful life estimation approach. *PHM 2012 - 2012 IEEE Int. Conf.on Prognostics and Health Management: Enhancing Safety, Efficiency, Availability, and Effectiveness of Systems Through PHM Technology and Application, Conference Program*. https://doi.org/10.1109/ICPHM.2012.6299524

Ramasso, E. (2014). Investigating computational geometry for failure prognostics. *International Journal of Prognostics and Health Management*, *5*(1), 5.

Ramasso, E., & Denoeux, T. (2014). Making use of partial knowledge about hidden states in HMMs: an approach based on belief functions. *IEEE Transactions on Fuzzy Systems*, *22*(2), 395–405.

Ramasso, E., & Gouriveau, R. (2014). Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions. *IEEE Transactions on Reliability*, *63*(2), 555–566. https://doi.org/10.1109/TR.2014.2315912

Ramasso, E., Rombaut, M., & Zerhouni, N. (2013). Joint prediction of continuous and discrete states in time-series based on belief functions. *IEEE Transactions on Cybernetics*, *43*(1), 37–50.

Ramasso, E., Saxena, A., Ramasso, E., Saxena, A., Benchmarking, P., & Meth-, A. P. (2016). Performance Benchmarking and Analysis of Prognostic Methods for CMAPSS Datasets . To cite this version : HAL Id : hal-01324587 Performance Benchmarking and Analysis of Prognostic Methods for CMAPSS Datasets.

Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on* (pp. 1–9). IEEE.

Shafer, G. (1976). *A mathematical theory of evidence* (Vol. 1). Princeton university press Princeton.

Sun, J., Zuo, H., Wang, W., & Pecht, M. G. (2012). Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance. *Mechanical Systems and*

Signal Processing, *28*, 585–596. https://doi.org/10.1016/j.ymssp.2011.09.029

Tamilselvan, P., Wang, Y., & Wang, P. (2012). Deep belief network based state classification for structural health diagnosis. *IEEE Aerospace Conference Proceedings*. https://doi.org/10.1109/AERO.2012.6187366

Tobon-meja, D., Medjaher, K., Zerhouni, N., & Iso, T. (2010). The ISO 13381-1 Standard ' s failure prognostics process through an example.

Van Tung, T., & Yang, B.-S. (2009). Machine fault diagnosis and prognosis: The state of the art. *International Journal of Fluid Machinery and Systems*, *2*(1), 61–71.

Wang, P., Youn, B. D., & Hu, C. (2012). A generic probabilistic framework for structural health prognostics and uncertainty management. *Mechanical Systems and Signal Processing*, *28*, 622–637. https://doi.org/10.1016/j.ymssp.2011.10.019

Wang, T., Yu, J., Siegel, D., & Lee, J. (2008). A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on* (pp. 1–6). IEEE.

Xi, Z., Jing, R., Wang, P., & Hu, C. (2014). A copula-based sampling method for data-driven prognostics. *Reliability Engineering and System Safety*, *132*, 72–82. https://doi.org/10.1016/j.ress.2014.06.014

Xu, J., Wang, Y., & Xu, L. (2014). PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data. *IEEE Sensors Journal*, *14*(4), 1124–1132. https://doi.org/10.1109/JSEN.2013.2293517

Yager, R. R. (1987). On the Dempster-Shafer framework and new combination rules. *Information Sciences*, *41*(2), 93–137.

Yan, J., Koc, M., & Lee, J. (2004). A prognostic algorithm for machine performance assessment and its application. *Production Planning & Control*, *15*(8), 796–801.

Zemouri, R., & Gouriveau, R. (2010). Towards accurate and reproducible predictions for prognostic: an approach combining a RRBF Network and an AutoRegressive Model. In *1st IFAC Workshop on Advanced Maintenance Engineering, Services and Technology, IFAC A-MEST'10.* (pp. 163–168).

Zhang, L. (1994). Representation, independence, and combination of evidence in the Dempster-Shafer theory. In *Advances in the Dempster-Shafer theory of evidence* (pp. 51–69). John Wiley & Sons, Inc.