

OPEN ACCESS

Remaining Useful Life Estimation using Time Trajectory Tracking and Support Vector Machines

To cite this article: D Galar *et al* 2012 *J. Phys.: Conf. Ser.* **364** 012063

View the [article online](#) for updates and enhancements.

You may also like

- [A sequence-to-sequence remaining useful life prediction method combining unsupervised LSTM encoding-decoding and temporal convolutional network](#)
Jialin Li, Renxiang Chen and Xianzhen Huang
- [Spatial attention-based convolutional transformer for bearing remaining useful life prediction](#)
Chong Chen, Tao Wang, Ying Liu et al.
- [Research on remaining useful life of rolling bearings using EWT-DI-ALSTM](#)
Runxia Guo and Bo Gong



ECS The Electrochemical Society
Advancing solid state & electrochemical science & Technology

242nd ECS Meeting

Oct 9 – 13, 2022 • Atlanta, GA, US

Early hotel & registration pricing ends September 12

Presenting more than 2,400 technical abstracts in 50 symposia

The meeting for industry & researchers in

BATTERIES
ENERGY TECHNOLOGY
SENSORS AND MORE!

 Register now!

 **ECS Plenary Lecture featuring M. Stanley Whittingham,**
Binghamton University
Nobel Laureate –
2019 Nobel Prize in Chemistry



Remaining Useful Life Estimation using Time Trajectory Tracking and Support Vector Machines

D Galar¹, U Kumar¹, J Lee² and W Zhao²

¹ Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden;

² NSF/UCR Center for Intelligent Maintenance System (IMS), University of Cincinnati, Cincinnati, OH 45221, USA;

E-mail: diego.galar@ltu.se, uday.kumar@ltu.se, jay.lee@uc.edu, zhaow@mail.uc.edu

Abstract. In this paper, a novel RUL prediction method inspired by feature maps and SVM classifiers is proposed. The historical instances of a system with life-time condition data are used to create a classification by SVM hyper planes. For a test instance of the same system, whose RUL is to be estimated, degradation speed is evaluated by computing the minimal distance defined based on the degradation trajectories, i.e. the approach of the system to the hyper plane that segregates good and bad condition data at different time horizon. Therefore, the final RUL of a specific component can be estimated and global RUL information can then be obtained by aggregating the multiple RUL estimations using a density estimation method.

1. Introduction

Fault detection, diagnostics and further prognostics are performed in order to choose different maintenance management actions and scheduling of these actions. These steps correspond to the need, firstly, of perceiving phenomena, next, of understanding them, and finally, of acting consequently. However, rather than understanding a phenomenon which has just appeared like a failure, it seems convenient to anticipate its manifestation and consequences in order to consequently and, as soon as possible, resort to protective actions. This is what could be defined as prognostics, and it is strongly related to the remaining useful life (RUL) of the observed asset.

Prognostics reveal to be a very promising maintenance activity as it should permit to improve the dependability of the whole system. Also, industrials show a growing interest in this thematic which becomes a major research framework. However, considering the benefits that such technology may bring to the security, economics and resource management fields, the research community still doesn't agree in a formal definition, methodology or framework to instrument the prognostic process. That can be explained from different aspects. Firstly, prognostics still is not a stabilized concept: there is no consensual way of understanding it which makes harder the definition of tools to support it in real

¹ Diego Galar; diego.galar@ltu.se

applications. Secondly, many approaches for prediction exist whose applicability is highly dependent of the available knowledge on the monitored system. Thirdly, the vagueness of prognostic process definition impedes to point out the inherent challenges for scientists.

Thus, the purpose of this paper is to analyze and discuss the advantages of information supplied by Support Vector Machines (SVM), used as classifiers, in diagnosis to perform proper prognosis based on it for complex systems. The whole aim is to show a clear difference between diagnostics information, performed many times by anomaly detection tools, and prognostic information, not properly paid attention to perform (and develop) real prognostic systems.

1.1. Concept of prognostic

The European Standard on maintenance terminology does not define prognosis [1]. This reveals that prognostic is a quite new area of interest.

Prognostic is traditionally related to mechanics and fatigue cracks. It started to be brought up by this research community as a field of interest. In this meaning, prognostic is called the prediction of a system's lifetime and corresponds to the last level of the classification of damage detection methods introduced by [2]. Prognostic can also be defined as a probability measure: a way to quantify the chance that a machine operates without a fault or failure up to some future time. This probabilistic prognostic value is all the more an interesting indication as the fault or failure can have catastrophic consequences and maintenance manager needs to know if inspection intervals are appropriate.

Finally, prognostic can be defined as recently proposed by the International Organization for Standardization: prognostic is the estimation of Time To Failure (TTF) and risk for one or more existing and future failure modes [3]. In this acceptance, prognostic is also called the prediction of a system's lifetime as it is a process whose objective is to predict the remaining useful life (RUL) before a failure occurs given the current machine condition and past operation profile [4]. These approaches are grounded on the failure notion, i.e., termination of the ability to perform a required function [1], which implies that the prognosis is associated with a degree of acceptability. Following that, prognostic should be based on assessment criteria, whose limits depend on the system itself and on performance objectives [5], [6], and prognostic could be split into two sub-activities: a first one to predict the evolution of a situation at a given time, and a second one to assess this predicted situation with regards to an evaluation referential (figure 1).

A central problem can be pointed out: the accuracy of a prognostic system is related to its ability to approximate and predict the degradation of equipment; the prediction phase is a critical one. A look at prognostic metrics enables to point it out. Prognostic is in essence an uncertain process. In this way, prognostic system performance measures can be constructed. The main prognostic measure pursued is the predicted time to failure (TTF), also called the remaining useful life (RUL).

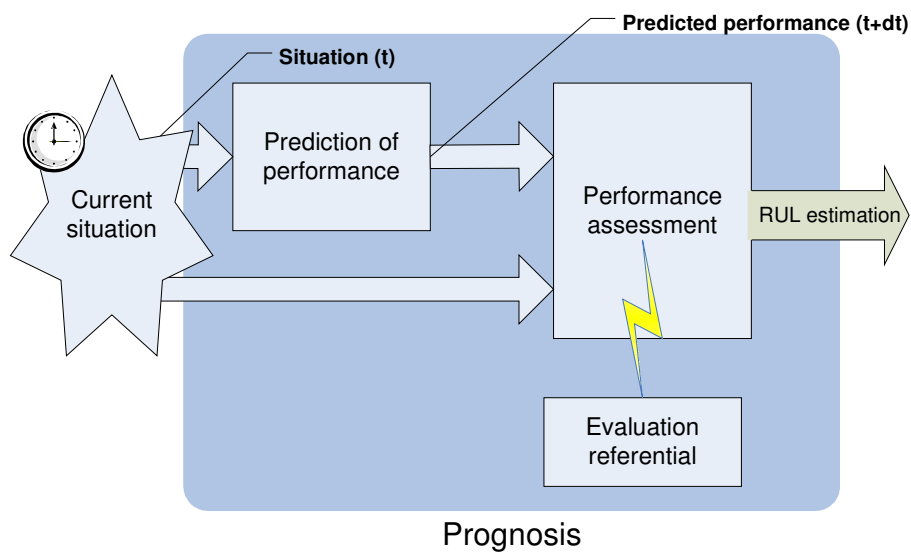


Figure 1. Prognostic as an assessment performance process.

The paper is organized as follows. First of all, the concept of prognostic has been briefly defined and positioned within the maintenance strategies. So, the next part is dedicated to the analysis of the diagnosis tools used in diagnostics and the information produced by them to predict RUL in a later stage. Finally SVM classifiers in feature maps are shown as successful diagnosis tool and promising prognosis tool with the proposed methodology.

2. Feature Extraction

Feature extraction and selection are the first steps after raw data acquisition to perform diagnosis and further prognosis. They help to extract the useful information from the raw data to perform more accurate, faster and easier defect diagnosis. Selected features should be sensitive to machine faults and robust to background noise. Another important consideration in feature selection is that computation complexity for extracting features should be low to be applied for real time diagnosis. The selection of a small set of features will be fused in a self-organizing feature map (SOFM) or in a n-dimensional space where the current status of the analyzed element will be represented according to the selected parameters.

In rotary machinery diagnosis, the features are mainly extracted from two different sources: time domain signals and frequency domain. The most popular ones are focused on the frequency domain. However, in the early stage of failure development, the damage is not significant and the defective signal is masked by the noise present in the acquired signal. Therefore, spectral analysis may not be effective. On the other side, using the time domain feature is recommended because normal and defective signals differ in their statistical characteristics in time domain where calculation is simple and complexity is low. Some time domain features used in literature are listed in table 1.

Table 1. Time domain features.

Feature	Definition	Feature	Definition
1 Peak value	$Pv=(1/2)[\max(x_i)-\min(x_i)]$	6 Clearance factor	$Clf = \frac{Pv}{(\frac{1}{n} \sum_{i=1}^n \sqrt{ x_i })^2}$
2 RMS	$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i)^2}$	7 Impulse Factor	$Imf = \frac{Pv}{\frac{1}{n} \sum_{i=1}^n x_i }$
3 Standard Deviation	$Std = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$	8 Shape factor	$Shf = \frac{RMS}{\frac{1}{n} \sum_{i=1}^n x_i }$
4 Kurtosis Value	$Kv = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{RMS^4}$	9 Normal Negative Likelihood value	$NNL = -\ln L$; $L = \prod_{i=1}^N f(x_i, u, \sigma)$
5 Crest factor	$Crf = Pv / RMS$		

One has to be careful when features are selected due to different amount of information embedded in them. Some features may not contribute to the failure diagnosis and even degrade the performance of the diagnosis and others can have redundant information. For that purpose, a Separation Index is used to define the significance of features [7]. For two signals presented to be compared, let \bar{m}_d and \bar{m}_h represent mean of the samples from each signal respectively. S_d and S_h represent standard deviation. One Separation Index (SI) is defined as [9]

$$SI = \frac{|\bar{m}_d - \bar{m}_h|}{S_d + S_h} \quad (1)$$

For trajectory prediction purposes, it is important to select features that exhibit some predictable trends that relate to the health of the system. This feature follows an upward trend over time, which would be useful during the prediction through measurement updates. If selected features are extremely sensitive to changes in degradation process when updated data are collected, then clear trajectories will be displayed in the n-dimensional space, where these features are fused, to predict future behavior.

3. Diagnosis using SVM Classifier

The n-dimensional map, where features are displayed, requires the existence of boundaries to create clusters of different faulty classes. SVM is proposed as a feasible technique to provide these boundaries. SVM was initially developed to classify two classes of objects. One decision function needed to be found for such binary classification. However, there are lots of applications, fault diagnosis among them, where number of classes is more than two, as the example illustrated in figure 2. To accommodate the multi-classes problems, one solution is to merge several binary SVMs together. One-Against-All multi-class SVM is one of them, discussed in [8].

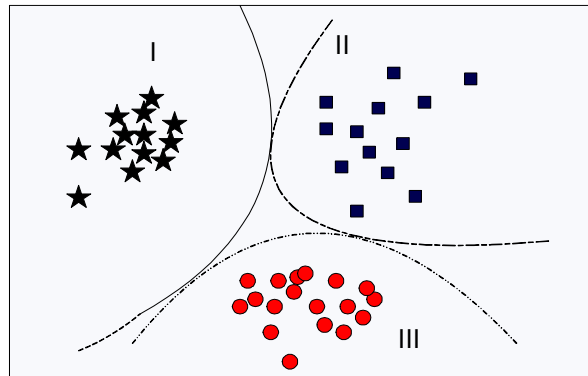


Figure 2. Three-Class Classification Problem.

Suppose there are k classes of data sets to be separated. Given l training data $(x_1, y_1), \dots, (x_l, y_l)$, which is the input of SVM. The input of the SVM is a feature vector in this paper. The $y_i \in \{1, 2, 3, \dots, k\}$ is the output of the SVM and it is the indicator of the category (class) of a data set belonging to. The One-Against-All method transformed the multi-class problem into k sub binary classification problems. The i^{th} sub binary classification problem label the indicator of data sets belong to the i^{th} class with 1 and label all the remaining data sets with -1. The mathematical formula for this i^{th} binary SVM is [10]:

$$\begin{aligned} \min \quad & \frac{1}{2}(\omega^i)^T \omega^i + C \sum_{j=1}^l \xi_j^i (\omega^i)^T \\ & (\omega^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i, \text{ if } y_j = i \\ & (\omega^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i, \text{ if } y_j \neq i \\ & \xi_j^i \geq 0, \quad j = 1, \dots, l \end{aligned} \quad (2)$$

As each sub binary SVM has one decision function, one can obtain k decision functions for a k -classes SVM:

$$\begin{aligned} & (\omega^1)^T \phi(x) + b^1 \\ & \dots \\ & (\omega^k)^T \phi(x) + b^k \end{aligned} \quad (3)$$

The predicted class for data set x is the class with largest decision function value, as illustrated in figure 3. It is:

$$i = \arg \max((\omega^j)^T \phi(x) + b^j) \quad (4)$$

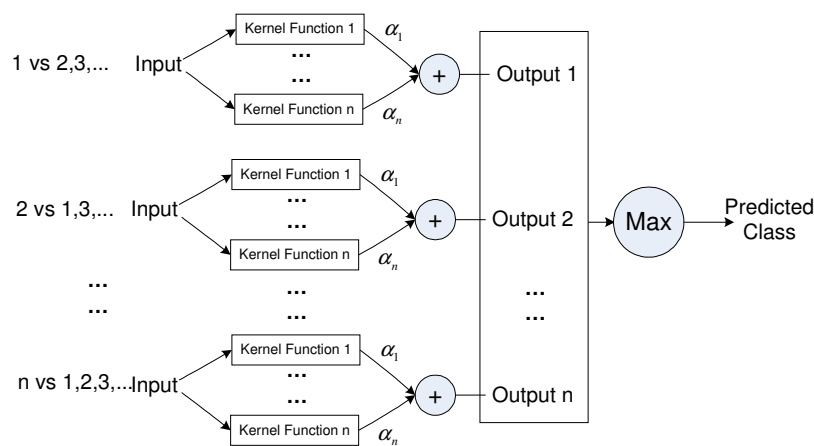


Figure 3. Multi-Class SVM.

Therefore, SVM is a flexible classifier of different faults. When the kernel function is nonlinear, the decision function and boundaries in n -dimensional feature map are also nonlinear, so SVMs provide accurate separation of points maximizing the distance between separating planes in a higher dimensional space.

Final result is an accurate classification, getting the clustering of given data to a specific class. This is the basis of diagnosis, i.e. the identification of the existing failure; but it does not provide any information about the magnitude of the present damage. It just provides a static picture of the current condition but nothing about its evolution and degradation. For that purpose, prognosis is required.

4. Trajectory Estimation

Considering a degradation process involving no or limited maintenance, the process may compose of a sequence of irreversible stages (either discrete or continuous) from new to worn out, which can be implicitly expressed by the trajectory of the measured condition data or features. Therefore the RUL of the system can be estimated if its future degradation trend can be projected from those historical instances that have failed. The similarity between the degradation trajectories of different instances can be computed first; and then the failure time of one instance can be estimated based on the actual failure time of similar instances; finally the RULs estimated from multiple historical instances can be aggregated to generate the final prediction of RUL.

Trajectory tracking and prediction involve systems with nonlinear and time-varying models, unknown and changing noise statistics, and non-uniform sampling intervals. Both tracking and predictions are utilized to develop time to failure (TTF) estimations for the mechanical system device.

4.1. Trajectories in Feature Space

Trajectory tracking involves the estimation and prediction of one or more parameters of a system using observations taken over time. These observations are usually measurements of physical properties taken from a particular sensor, such as vibration for CM applications. These physical properties could include position, velocity and acceleration. There are practical instances however where the measurements do not provide enough or clear information to track and predict the behavior of the system. Therefore it is much easier to track the set of featured previously selected and extracted, since they usually provide more clear and useful information for tracking.

Proposed prognostic method involves the tracking of actual trajectory of the system state in the feature space. Obviously, this is useful when it is known a priori that selected features are strongly

linked to the degradation process. In this case, the trajectory of the current condition in the feature map may be indicative of the real evolution of the health of the system.

This is particularly useful in the CBM application, where the goal is to predict when a mechanical system is going to fail. If a feature is found to describe the current “health” of the system being monitored, the trajectory of the feature may be tracked to estimate the current health and to predict ahead to determine when the system is going to fail.

Traditionally, single feature has been used to predict the behavior of the system. Fusion of two or more features, with no redundant information among them, can contribute to create a much more accurate trajectory of degradation in the n-dimensional feature space. The accuracy of this prediction is based on the physical constraints of the trajectories due to the information embedded into them.

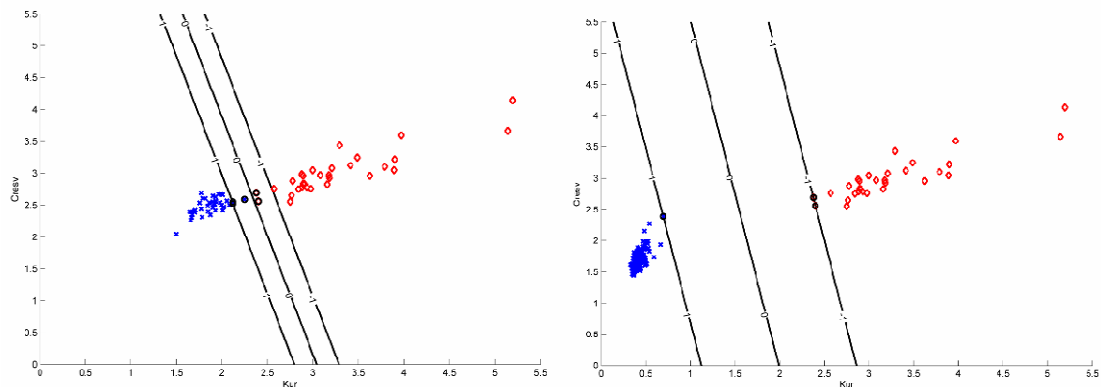


Figure 4. Two stages of bearing degradation.

Feature trajectories are governed by a dynamics model which describes how the feature changes with time. Since features are extracted from data that obeys the laws of physics, some features have constraints that restrict their dynamics over time, examples being how position, velocity and acceleration are restricted by the laws of dynamics. Features that follow a consistent and predictable trend are useful for tracking, while features that have abrupt and unpredictable changes may be more useful for classification purposes. Figure 4 shows an example of how two different features behave over time.

Top figure shows blue points with specific failure size in inner race. Simultaneously red points are the training set corresponding to a bigger failure size. Figure below shows a later stage where cloud of points is moving towards bigger failure sizes. These two dimensional feature map uses Crest Factor and Kurtosis value in the fusion process.

4.2. Time to failure (TTF)

Time to failure (TTF) is the amount of time left before a system reaches mechanical failure. In this model, failure is a theoretical event that occurs when a feature value crosses some predetermined failure threshold. The TTF is calculated anytime the feature value crosses a particular value of interest called the detection threshold. The TTF estimate is set to zero before the feature value crossed the detection threshold. After the detection threshold is crossed, the TTF is calculated by propagating the current feature track forward n time steps based on the methodology further discussed. The TTF estimate is the difference between the current time step and the time of the predicted threshold crossing. This process is shown in figure 5.

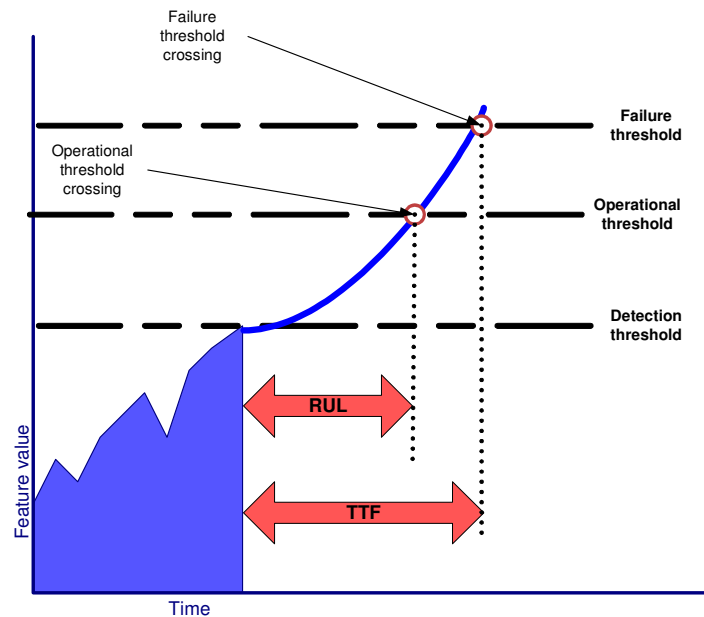


Figure 5. Difference between RUL and TTF.

Remaining useful life (RUL) is the amount of time left before a system fails to operate within acceptable limits. RUL is calculated very similar to TTF, except that instead of a failure threshold, there is an upper operating limit threshold. As shown in Figure 5, the operating threshold for RUL is lower than the failure threshold, which results in a lower value for RUL than TTF. This is reasonable since it is desirable to repair the system before failure, while still utilizing any useful life before maintenance. However RUL is very application dependent as the specified tolerances must be defined for a given system. Because of this, the results are given for TTF instead of RUL.

4.3. Trajectories calculation

Let us suppose that M different failure categories have been successfully trained to recognize the M distinct sequential states of interest for a failure mechanism. Presentation of temporally ordered (by life/usage) observation sequences from such a process would yield the sorts of trajectories illustrated in figure 6. For simplicity three known states are considered: normal bearings, one acceptable failure size in inner race and unacceptable failure size.

If coordinates of measured state are in one of M defined boundaries for any given observation sequence, then, one would declare the machine to be in this state.

The coordinates of the points of intersection of the profiles for the different states along the life/usage axis in figure 6 represent the estimated state-transition time instants. It is these state transition points that would allow us to extend the use of known failure modes/states for prognostics.

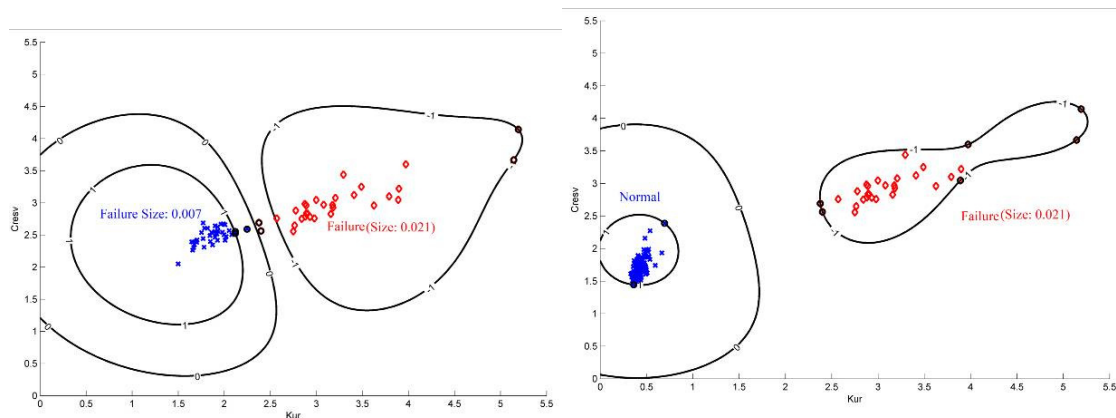


Figure 6. Trajectory evolution with Gaussian kernel SVM.

The objective of the diagnostics process is to recognize the M distinct failure states of a machine. However if one presents the temporally sorted observation sequences from several units, as illustrated in figure 6, it will result in a cloud of points moving towards boundaries with different transition times. Let us suppose that observation sequences are available from several similar units, e.g. identical bearings in a paper mill, collected for the purpose of developing diagnostics and prognostics models. This will result in estimated vectors of state transition times. These vectors of transition times will provide the information necessary to carryout prognostics.

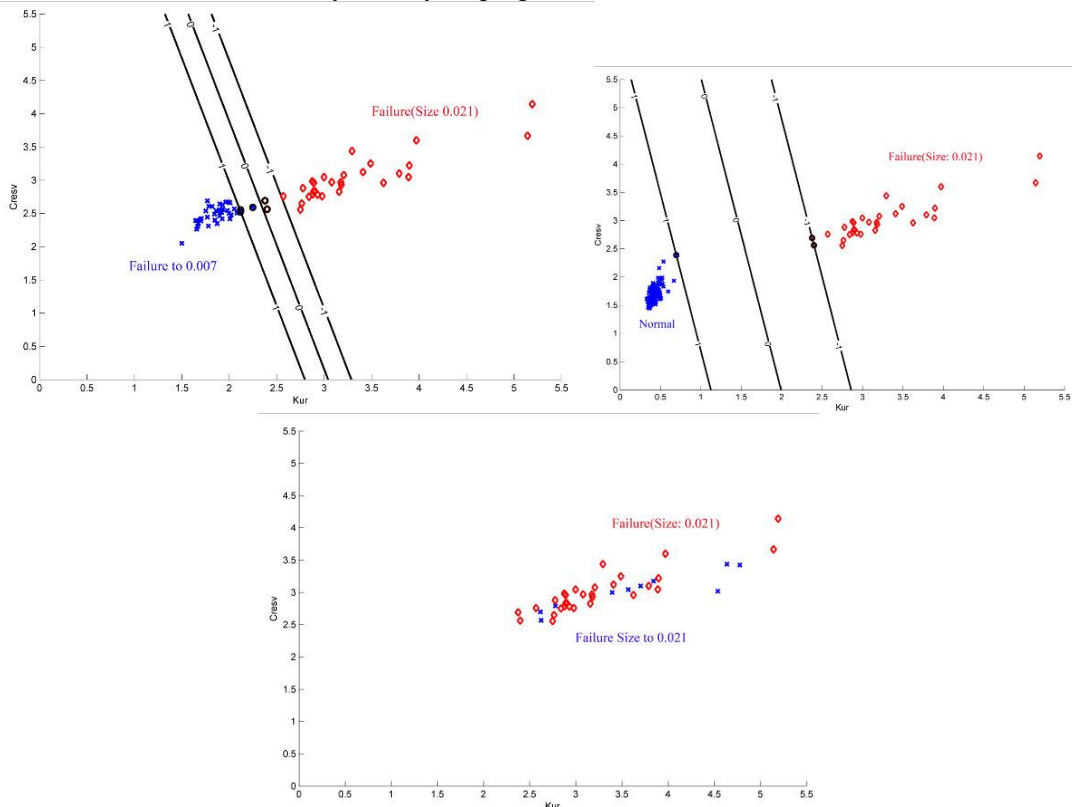


Figure 7. Trajectory evolution with Polynomial kernel SVM.

The procedure is as follows. One can assume that these times follow some multivariate distribution. Once the distribution is assessed, the conditional probability distribution of a distinct state transition given the previous state transition points for any 'individual' unit under investigation can be estimated. The process of constructing any necessary confidence intervals is rather straightforward as well. The process can be iterated in a recursive manner to make predictions regarding several sequential state transition times. As expected, the larger the number of state transitions already witnessed for a unit the tighter the prediction intervals associated with the final states of the unit.

The goal here is to model the state-transition time vectors resulting from the diagnostics models. For the current case study, shown in figure 7, one can observe the evolution of feature clouds of identical bearings.

Average distance in first stage between normal bearings and faulty bearings is around 2.8 in terms of kurtosis values. This distance is reduced to 0.1 in second stage and around 0 when all bearings have failed. Time distance in terms of one feature is useful. However fusion of several features evolutions reveals more accurate in order to process degradation speed of multidimensional vectors, where more information is considered in the degradation process.

5. Conclusion

Proposed model develops an effective RUL prediction method that addresses multiple challenges in complex system prognostics where many parameters are unknown. Similarities between degradation trajectories can be checked in order to enrich existing methodologies in prognostics applications. Existing CM data for bearings will be used to verify the model.

References

- [1] EN 13306, Maintenance terminology. European Standard, CEN/TC 319 – AFNOR, June 2001.
- [2] Charles R Farrar, Nick A J Lieven, *Damage Prognosis: The Future of Structural Health Monitoring* Philosophical Transactions: Mathematical, Physical and Engineering Sciences. Vol. 365, No. 1851, Structural Health Monitoring (Feb. 15, 2007), pp. 623-632A.
- [3] ISO, 13381-1, Condition monitoring and diagnostics of machines - prognostics - Part1: General guidelines. Int. Standard, ISO, 2004.
- [4] A K S Jardine, D Lin and D Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mech. Sys. & Sig. Pro.*, Vol. 20, pp. 1483-1510, 2006.
- [5] O Dragomir, R Gouriveau, N Zerhouni and F Dragomir, "Framework for a distributed and hybrid prognostic system", in: 4th IFAC Conf. on Manag. and Control of Prod. and Logistics, 2007.
- [6] O Dragomir, R Gouriveau and N Zerhouni, "Adaptive neuro-fuzzy inference system for mid term prognostic error stabilization", *Inter. Jour. of Comp. Comm. and Control*, Vol. 3, pp. 271-276, 2008
- [7] B Sreejith, A K Verma, A Srividya. Fault diagnosis of rolling element bearing using time-domain features and neural networks. 2008.
- [8] Hsu CW, Lin CJ A comparison of methods for multiclass support vector machines. *IEEE Transaction Neural Network*. 2002;13:415-25.