# An unsupervised clustering method for assessing the degradation state of cutting tools used in the packaging industry

**F. Cannarile[1,2], P. Baraldi[2], M. Compare[1,2], D. Borghi[3], L. Capelli[3], M. Cocconcelli[4], A. Lahrache[4], E. Zio[1,2,5]**

[1]Aramis Srl, Via Pergolesi 5, 20124, Milano, Italy

[2]Energy Department, Politecnico di Milano, Via La Masa 34/3, 20156 Milano, Italy

[3]Tetra Pak Packaging Solutions S.p.A., 41122 Modena, Italy

[4]Department of Science and Engineering Methods, University of Modena and Reggio Emilia, 42122, Reggio Emilia, Italy

[5]Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, France

## Abstract

*The objective of the present work is to develop a method for the identification of the degradation state of cutting tools (knives) used in the packaging industry. The main difficulties to be addressed are that i) only measurements of a physical quantity indirectly related to the knives degradation are available and ii) only the beginning and the end of operation of the knives are known, whereas no information is available on the component degradation state during its operation life. A method to identify the component degradation state is here proposed. First the general setting for extracting health indicators to measure the amount of knife degradation from a set of signals measured during operation is discussed. Then, an optimal subset of health indicators is selected based on monotonicity and trendability indexes. Finally, the optimal subset of health indicators is fed to a Fuzzy C-Means (FCM) clustering algorithm, which allows assessing the knife degradation state. The application of the proposed method to real condition monitoring knife data is shown to lead to satisfactory results.*

**Key words**: Degradation State Assesment, Health Indicator Developement, Knive, Packaging Industry

## 1. INTRODUCTION

Multi-state degradation modelling can be used to support maintenance strategies by offering a description of the degradation process based on the maintenance inspection practice, of assigning classes of degradation state to the inspected device. Multi-state modelling frameworks have been developed for membranes of pumps operating in Nuclear Power Plants (Baraldi et al., 2011), turbine nozzles for the Oil&Gas industry (Compare et al. 2016), liners of marine diesel engine cylinders (Giorgio et al, 2011), piping of nuclear power plants (Veeramany et al., 2011) (Cannarile et al, 2017). In this work, we present a method for on-line assessing the degradation state of knives installed on Tetra Pak® A3/Flex filling machines used to cut package material. We consider a use case in which we have available measurements of a physical quantity indirectly related to the knife degradation and we know the beginning and end of operation life of the knife, but we do not have any information on the component degradation state during its life. The method to identify the component degradation state is based on the following three steps:

S1. *Extraction of statistical and frequency-based features from the raw data;*
S2. *Selection among the set of extracted features of an optimal subset of Health Indicators (HIs) for the identification of the component degradation*;
S3. *Unsupervised clustering of HI data in order to identify the component degradation state*.

The main novelty of our work lies in the strategy to select the optimal subset of HIs based on monotonicity and trendability indexes and in its application to knives used in the packaging industry. The paper is organized as follows: in Section 2, the general setting for HI development is proposed; in Section 3, Fuzzy C-Means (FCM) clustering for identifying the knife degradation state is discussed. The application of the methodology to Tetra Pak® A3/Flex filling data is described in Section 4, whereas in Section 5 conclusions are drawn.

## 2. HEALTH INDICATORS

In this Section, we describe the general setting for developing HIs capable to catch the degradation evolution of the monitored knives over time. Subsections 2.1, 2.2. and 2.3 present the three steps of the proposed procedure: feature extraction, trend extraction and feature selection, respectively. We assume to have available raw measurements registered during data acquisition sessions performed during the lives of $N$ identical knives, i.e., from their installation up to their replacement. The generic $\tau^{th}$ data acquisition on the $i^{th}$ knife provides the vector $\boldsymbol{x}_i(\tau), i = 1,..,N; \tau = 1,...,n_i$ of the signal values collected during the knife cutting process with $n_i$ indicating the total number of data acquisition performed on the $i^{th}$ knife before its replacement. Notice that during a single cut multiple acquisitions are typically performed, from the beginning of the cut until its end.

### 2.1 Feature Extraction

The development of health indicators from the raw measurements $\boldsymbol{x}_i(\tau), i = 1,..,N, \tau = 1,..,n_i$, requires a first phase of feature extraction, where the raw measurements are preprocessed in order to reduce the dimensionality of the vector $\boldsymbol{x}_i(\tau)$ and extract information relevant with respect to the component degradation process. In practice, from the raw measurement vector, which contain thousands of signal values, we extract $F$ different lumped features by considering statistical metrics (e.g., means, standard deviations, etc.) and analytics (e.g., derivatives, elongation, etc.), signal transforms in the frequency domain (e.g., Fourier Transform, Laplace Transform) and/or in the time-frequency domain (e.g., Short Time Fourier Transforms (STFT) (Kaewkongka et al., 2003) and Wavelet Transform (WT) (Baraldi et al., 2016).

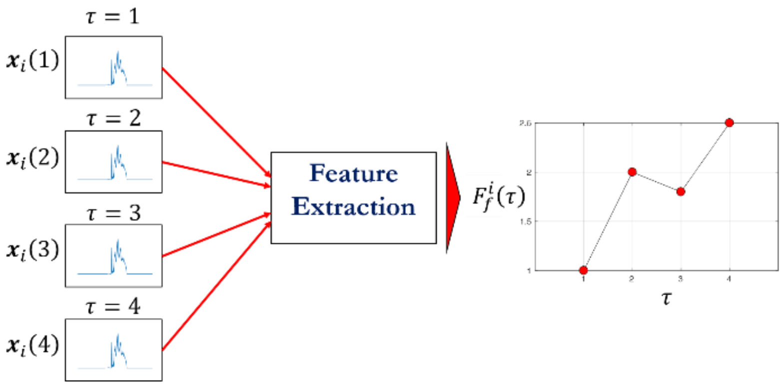


**Figure 1: Feature extraction procedure**

The raw vector $\boldsymbol{x}_i(\tau)$, is thus, transformed into the vector $F_f^i(\tau), f = 1,..,F$ (see Figure 1). Since the obtained features are typically noisy, a further step of data preprocessing is performed by using Empirical Mode Decomposition (EMD) (Huang, 2000). The main idea behind the use of EMD is that a signal can be typically formed by fast oscillations superimposed to slow oscillations (Rilling & Flandrin, 2006). EMD is able to identify a local 'low frequency' component, that is referred to as local trend $m_G[F_f^i](\tau)$, and zero-mean oscillations or Intrinsic Mode Functions (IMFs) $q_g[F_f^i](\tau)$, allowing the decomposition of the signal $F_f^i$ as follows:

$$F_f^i(\tau) = \sum_{g=1}^{G} q_g[F_f^i](\tau) + m_G[F_f^i](\tau) \quad (1)$$

Thus, EMD performs a multi-scale decomposition that is fully data-driven (model-free) and that can be applied to any oscillatory time series, including nonstationary ones and/or those generated by a nonlinear system (Huang & Wu, 2008). As an example, Figure 2 shows the application of EMD to a noisy nonstationary signal $F_f^i$. The first IMF $q_1[F_f^i](\tau)$ contains the details of $F_f^i$, i.e., high-frequency content of the original signal, whereas the second IMF $q_2[F_f^i](\tau)$ contains the approximimation of $F_f^i$, i.e., the low-frequency content of the original signal.

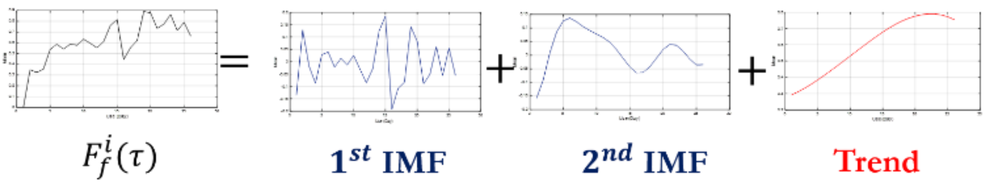


**Figure 2: EMD applied to signal $F_f^i(\tau)$**

Notice that each time $\tau$ a new data is acquired, the local trend and IMFs are modified in the entire signal time domain $[0, \tau)$, as shown in the example of Figure 3.

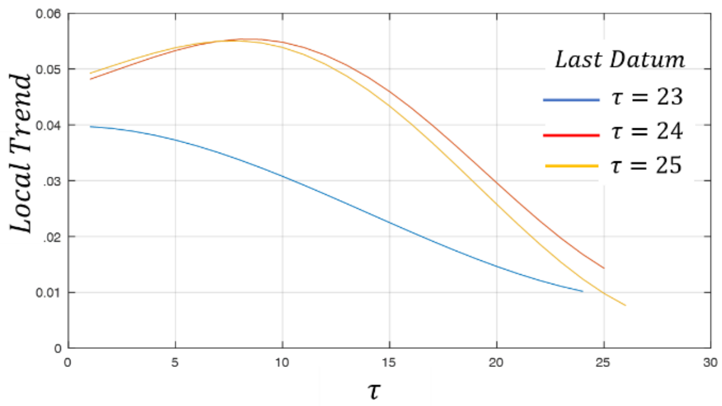


**Figure 3: Variations in the EMD trend when a new data is collected**

According to (Mosallam et al., 2014), variations of the resulting EMD trends can be informative with respect to the component degradation and can be quantified using different statistical indicators, such as the local trend mean value ($mean$), its standard deviation ($std$) and the parameters of its linear polynomial fitting, i.e., the slope and the intercept, which will be referred to as $a$ and $b$, respectively. In practice, four potential health indexes are extracted from each feature:

$$HI_f^i(\tau) = \left[HI_{f,1}^i(\tau)\ HI_{f,2}^i(\tau)\ HI_{f,3}^i(\tau)\ HI_{f,4}^i(\tau)\right] = \left[a_f^i(\tau)\ \ b_f^i(\tau)\ mean_f^i(\tau)\ std_f^i(\tau)\right] \quad \tau = 2, \dots, n_i \tag{2}$$

Figure 4 shows the procedure for HIs development described in this Subsection. Notice that each time a new value of the feature $f$ is acquired, the EMD transform has to be applied, the trends extracted and the four potential HIs computed.

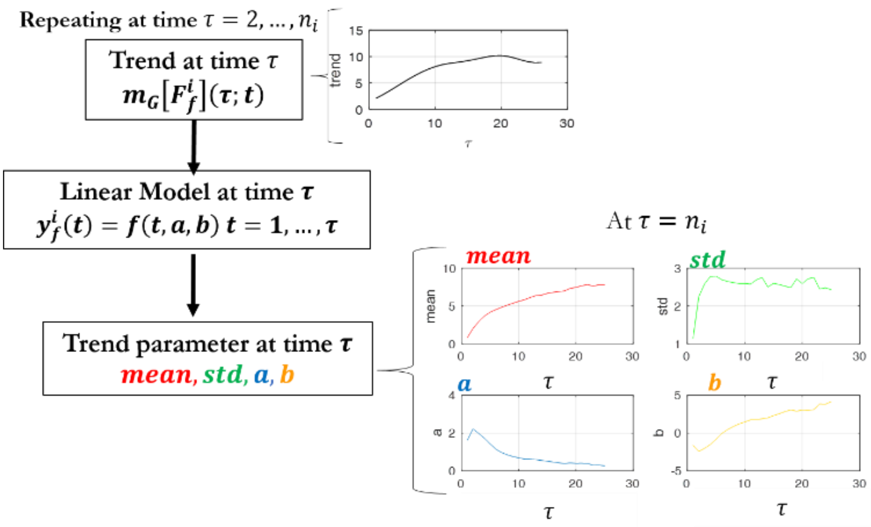


**Figure 4: Sketch of the procedure to extract trend features**

## 2.2 Feature Selection

Feature selection is the process of selecting among several possible candidates a smaller subset of features which satisfy the user desiderata with respect to the specific application under investigation. In this work, we consider as possible candidates the $4 \cdot F$ potential health indicators identified at the end of the feature extraction step and our objective is to select those that are best describing the component degradation evolution. To this purpose, we resort to the monotoncity and trendability indexes (Coble & Wesley Hines, 2009). Monotonicity is a measure of the underlying positive or negative trend of a HI. This is an important feature for the fault diagnostic task, since it is generally assumed that systems do not undergo self-healing, which would be indicated by a non-monotonic HI (Coble & Wesley Hines, 2009). Specifically, the monotonicity of a HI is defined as

$$M(HI_f) = mean\left(M_i(HI_f^i)\right) \qquad (3)$$

where $N$ is the number of degradation trajectories used to compute the index, $\#\frac{d}{d\tau} > 0 (< 0)$ is the number of points such that the first derivative of signal $HI_f^i$ is strictly positive (negative). This index assumes values within the interval [0,1], where 1 indicates the most satisfactory monotonicity level. On the other side, trendability indicates the degree to which the HI of a population of components have the same underlying shape and can be described by the same functional form; it is defined as

$$Tr(HI_f) = 1 - std(tr_i(HI_f^i)) \tag{4}$$

$$tr_i(HI_f^i) = \frac{\# \frac{d}{d\tau} > 0}{n_i - 1} + \frac{\# \frac{d^2}{d\tau^2} > 0}{n_i - 2} \quad i = 1{:}\,N \tag{5}$$

Trendability assumes values within the interval [0,1] where 1 indicates the best case. HIs satisfying user requirements in terms of trendability and monotoniity indexes will be referred to as $HI_{f*}, f^* \in \{1,..,F^*\}$, where $F^*$ is the user selected number of HIs. Figure 5 shows the procedure described above to develop and select an optimal subset of HIs.

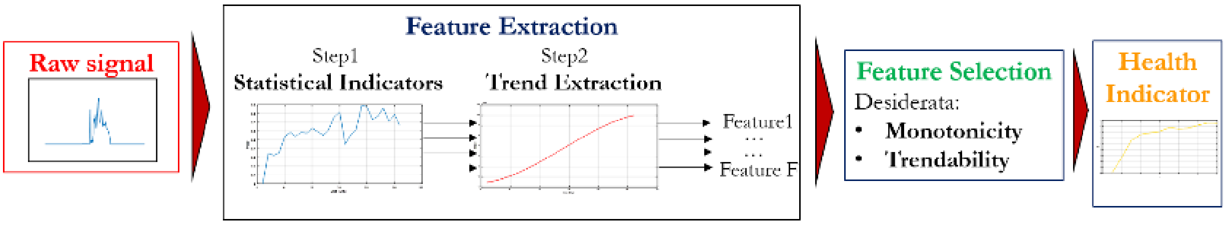


**Figure 5: Sketch of the procedure to develop and select the optimal subset of His**

## 3. UNSUPERVISED CLUSTERING FOR DEGRADATION STATE ASSESSMENT

The objective of this step is $i$) to identify the correct number of degradation states in which the degradation process should be discretized and $ii$) to label the data $HI_{f*}^i(\tau)$ by associating to each pattern its degradation state. This unsupervised clustering problem is tackled by resorting to the Fuzzy C-Mean (FCM) clustering algorithm (Jain et al., 1999). As input features to the FCM algorithm we employ the His selected at the previous step, $HI_{f*}^i$. FCM algorithm partions the data $HI^i(\tau) = [HI_1^i(\tau) \dots HI_{f^*}^i(\tau) \dots HI_{F^*}^i(\tau)]$ in order to minimize the following objective function

$$J(\Gamma, V) = \sum_{c=1}^{C} \sum_{i=1}^{N} \sum_{\tau=2}^{n_i} \left(\mu_{HI^i(\tau)}^i(c)\right)^{r_m} \left\|HI^i(\tau) - h_c\right\|^2 \quad r_m > 1 \tag{6}$$

$$\mu_{HI^i(\tau)}^i(c) = \frac{1}{\sum_{f=1}^{C} \left(\frac{\|HI^i(\tau) - h_c\|}{\|HI^i(\tau) - h_f\|}\right)^{\frac{2}{r_m - 1}}} \quad i = 1, \dots, N; \tau = 1, \dots, n_i;\ r_m > 1; c = 1, \dots, C \tag{7}$$

where $C$ is a user defined integer number which represents the number of clusters in which data have to be grouped, $h_c$ is the center of cluster $\Gamma_c$, $r_m$ is the fuzzifier which determines the level of cluster fuzziness and $\mu^i_{HI^i(\tau)}(c)$ is the degree to which element $HI^i(\tau)$ belongs to cluster $\Gamma_c$. To choose the most appropriate number of clusters $C^*$ the FCM is run several times changing the candidate number of clusters from 2 to $C_{max}$. Then, the best number of clusters is selected resorting to the silhouette and Davies-Bauldin coefficients (Cannarile et al., 2015). The former measures how similar a given pattern is to patterns in its own cluster with respect to the patterns in the other clusters, where values close to one indicate a good clustering. The latter is based on a ratio of within-cluster and between-cluster distances, and therefore, the smaller the Davies-Bouldin index value, the better the clustering.

## 4. CASE STUDY

This Section presents the results of the application of the proposed method to condition monitoring data collected during the degradation of $N = 23$ identical knives.

### 4.1 Statistical Indicator Feature Extraction

Firstly, we have extracted $F = 27$ features (see Appendix) from each 2600-dimensional cut signal $\boldsymbol{x}_i(\tau), i = 1, \ldots, 23$. Figure 6 shows an example of evolution of one of the extracted features (amplitude of $3^{rd}$ harmonic from DFT) during the life of the 23 knives.

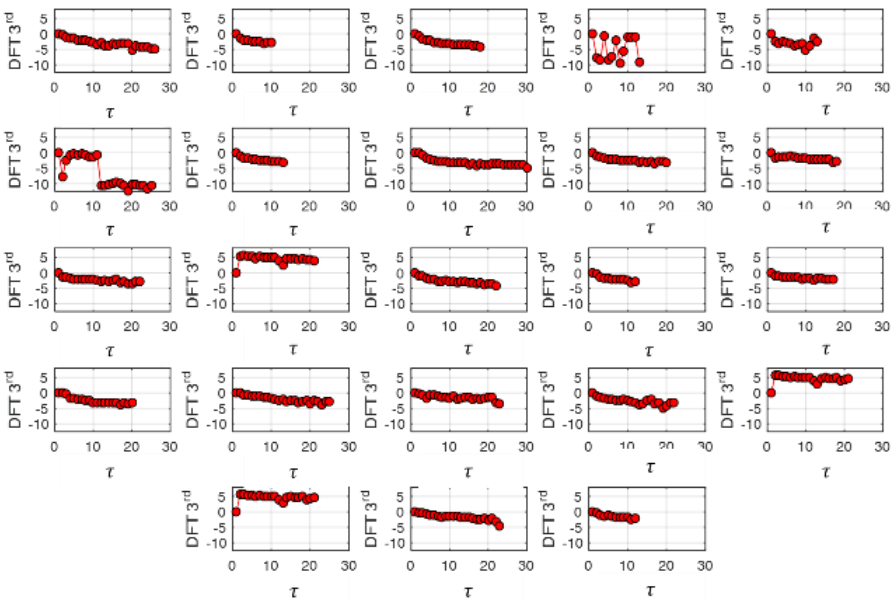


**Figure 6: Time-evolution of amplitude of $3^{rd}$ harmonic from DFT**

## 4.2 Trend Extraction

Applying the methodology based on EMD trend extraction discussed in Subsection 2.3; we have obtained $Fx4 = 108$ different candidate health indicators for each knife. Figure 7 shows an example of evolution of the mean trend obtained from the application of the Empirical Mode Decomposition (EMD) to the of $3^{rd}$ harmonic. Notice that after trend extraction, degradation trajectories in Figure 6 are less noisy and then, more monotone than before.

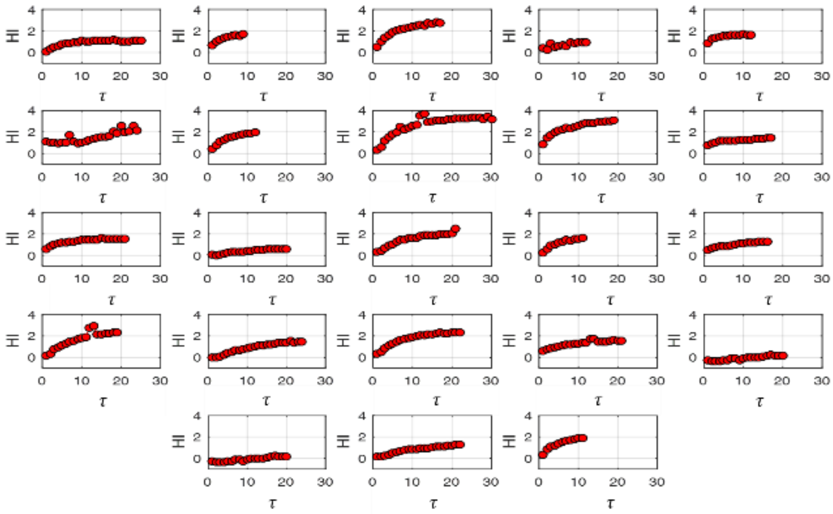

**Figure 7: Time-evolution of the mean trend obtained from the application of the Empirical Mode Decomposition (EMD) to the of $3^{rd}$ harmonic**

## 4.3 Feature Selection

To select an optimal subset of HIs, we have computed the trendability index and the monotonicity index for the 27 signals, based on statistical indicators and the 118 obtained possible health indicators. Results shows that the choice of constructing HIs based on feature extracted from EMD trends was successful: one can observe that there is a considerable improvement in the monotonicity index (best monotonicity index for HIs based on statistical indicators is approximatively 0.32, whereas for HIs based on EMD trend features, it is greater than 0.7). As HIs to be considered by the FCM clustering, we have chosen those reported in Table 1.

| Selected HIs | Statistical Indicator | After Trend Extraction? | Trend Feature | Monotonicity | Trendability |
|---|---|---|---|---|---|
| $HI_1^i$ | $\Delta x$ | Yes | $a$ | 0.7015 | 0.8783 |
| $HI_2^i$ | $\Delta x$ | Yes | $b$ | 0.7001 | 0.8683 |

**Table 1: Selected HIs**

where feature $\Delta x$ corresponds to the difference between two particular values of the cut signal $\boldsymbol{x}_i(\tau)$ (more details are not provided for confidentiality reasons).

## 4.4 FCM clustering

The FCM algorithm has been run several times by changing the number of candidate clusters from 2 to 10. Figures 8 and 9, respectively, show the values of the silhouette and Davies-Bauldin coefficients in correspondence to the number of clusters varying from 2 to 10.

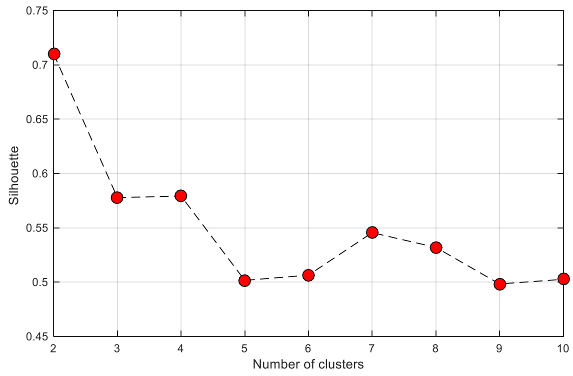


**Figure 8: Silhouette coefficient (to be minimized) increasing the number of clusters from 2 to 10**

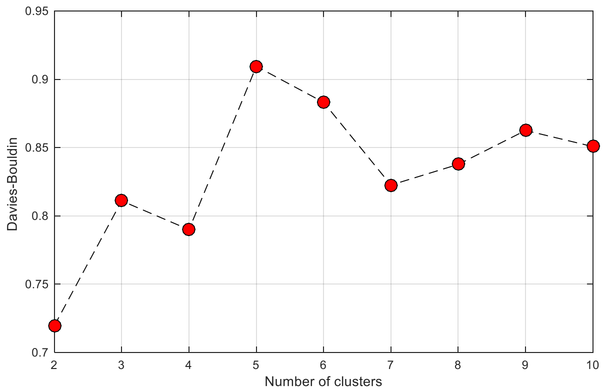


**Figure 9: Davies-Bauldin coefficient (to be maximized) increasing the number of clusters from 2 to 10**

Excluding the case in which only two clusters are identified, from both Figures 8 and 9 we can conclude that the best compromise solution according to the two coefficients is given when $C^* = 4$. Figure 10 shows for each pattern $HI^i(\tau)$ the corresponding associate cluster.

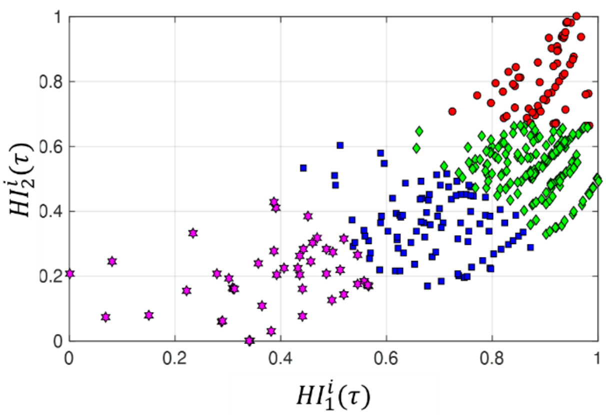

**Figure 10: Data pattern $HI^i(\tau)$ marked according to the assigned cluster (hexagram normal, square low degraded, diamond mildly degraded and circle very degraded)**

From Figure 10, it seems reasonable to assume that the degradation evolution of the knives corresponds to the following marker progression: hexagram (normal)-square (low degraded)-diamond (mildly degraded)-circle (very degraded). Finally, Figures 11 and 12 show degradation trajectories of the $10^{th}$ knife and its corresponding labelling provided by the FCM.

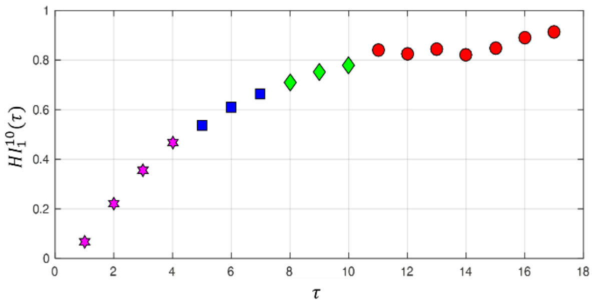


**Figure 11: Data pattern $HI_1^{10}(t)$ marked according to the assigned cluster (hexagram normal, square low degraded, diamond mildly degraded and circle very degraded)**

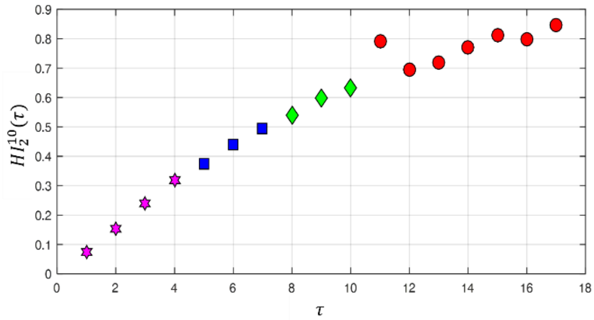


**Figure 12: Data pattern $HI_2^{10}(t)$ marked according to the assigned cluster (hexagram normal, square low degraded, diamond mildly degraded and circle very degraded)**

## 5. CONCLUSION

In this work, we have developed a method for assessing the degradation state of knives used in the packaging industry. We have proposed a general setting to develop HIs and considered the best subset of HIs for Fuzzy C-Means clustering to 1) infer the correct number of degradation states (which was not known a priori) and 2) label on-line monitored data with their correct degradation state. The developed method has been applied with success in practice, to real data from the packaging industry.

## 6. REFERENCE


Baraldi, P., Canese, R., Zio, E., Seraoui, R., Chevalier, R., 2011. Genetic algorithm-based wrapper approach for grouping condition monitoring signals of nuclear power plant components. Integrated Comuter-Aided Engineering, 18(3): 225-241.

Baraldi, P., Cannarile, F., Di Maio, F., Zio, E., 2016. Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions. Engineering Applications of Artificial Intelligence, 56: 1-13.

Cannarile, F., Compare, M., Di Maio, F., Zio, E., 2015. Handling reliability big data: A similarity-based approach for clustering a large fleet of assets. Safety and Reliability of Complex Engineered Systems - Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015: 891-896.

Cannarile, F., Compare, M., Rossi, E., Zio, E., 2017. A Fuzzy Expectation-Maximization algorithm for estimating the parameters of a multi-state degradation model for imprecise maintenance outcomes. Annals of Nuclear Energy (under review).

Coble, J., Wesley Hines, J., 2009. Identifying optimal prognostic parameters from data: A genetic algorithms approach. Annual Conference of the Prognostics and Health Management Society, PHM 2009.

Compare, M., Martini, F., Mattafirri, S., Carlevaro, F., Zio, E, 2016. Semi-Markov Model for the Oxidation Degradation Mechanism in Gas Turbine Nozzles. IEEE Transactions on Reliability, 65(2): 574-581.

Giorgio, M., Guida, M., Pulcini, G., 2011. An age- and state-dependent Markov model for degradation processes. IIE Transactions (Institute of Industrial Engineers), 43 (9): 621-632.

Huang, N.E., 2000. New method for nonlinear and nonstationary time series analysis: Empirical mode decomposition and Hilbert spectral analysis. Proceedings of SPIE - The International Society for Optical Engineering, 4056: 197-209.

Kaewkongka, T., Au, Y.H.J., Rakowski, R.T., Jones, B.E., 2003. A comparative study of short time fourier transform and Continuous Wavelet Transform for bearing condition monitoring. International Journal of COMADEM, 6 (1): 41-48.

## APPENDIX

Figure 13 shows the extracted features.

1. Mean
2. Maximum point
3. Minimum value
4. Standard deviation
5. Kurtosis
6. Impulse indicator
7. Shape indicator
8. 5th central moment
9. Shannon entropy
10. Coefficient of variation
11. Skewness
12. Crest indicator
13. Clearance indicator
14. Peak value
15. 3rd quartile
16. Maximum point of the first derivative
17. Maximum value of the first derivative
18. Amplitude of the first harmonic (FFT)
19. Amplitude of the second harmonic (FFT)
20. Amplitude of the third harmonic (FFT)
21. Amplitude of the fourth harmonic (FFT)
22. Amplitude of the fifth harmonic (FFT)
23. Amplitude of the sixth harmonic (FFT)
24. Norm $l^2$
25. Integral of the signal
26. Energy
27. $\Delta x$

**Figure 13: List of extracted features (acronym FFT refers to Fast Fourier Transform)**