

Ensemble-approaches for clustering health status of oil sand pumps

F. Di Maio ^a, J. Hu ^b, P. Tse ^b, M. Pecht ^b, K. Tsui ^b, E. Zio ^{a,c}

^a Energy Department, Politecnico di Milano, Via Ponzio 34/3, 20133 Milano, Italy, enrico.zio@polimi.it

^b Smart Engineering Asset Management Laboratory (SEAM), MEEM, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong

^c Ecole Centrale Paris and Supelec, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France

Abstract

Centrifugal slurry pumps are widely used in the oil sand industry, mining, ore processing, waste treatment, cement production, and other industries to move mixtures of solids and liquids. Wear of slurry pump components, caused by abrasive and erosive solid particles, is one of the main causes of reduction in the efficiency and useful life of these pumps. This leads to unscheduled outages that cost companies millions of dollars each year. Traditional maintenance strategies can be applied, but they provide insufficient warning of impending failures. On the other hand, condition monitoring and on-line assessment of the wear status of wetted components in slurry pumps are expected to improve maintenance management and generate significant cost savings for pump operators. In this context, the objective of the present work is to develop and compare two unsupervised clustering ensemble methods, i.e., fuzzy C-means and hierarchical trees, for the assessment and measurement of the wear status of slurry pumps when available data is extremely limited. The idea is to combine predictions of multiple classifiers to reduce the variance of the results so that they are less dependent on the specifics of a single classifier. This will also reduce the variance of the bias, because a combination of multiple classifiers may learn a more expressive concept class than a single classifier.

Keywords: Degradation, Fault detection, Fuzzy C-means, Hierarchical Tree, Ensembles of classifiers, Slurry pumps.

1. Introduction

Detection of anomalies and faults in hydraulic pumps is an important task with implications for the safe, economical, and efficient operation of hydraulic systems. These systems are often embedded in hazardous plants, and are thus expected to operate with high levels of reliability, availability, and safety although they run in adverse environment that may lead to sudden break down due to excessive wear. For this reason, both equipment manufacturers and owners invest significant resources in maintenance programs designed to ensure that the required hydraulic system performance is maintained at maximum efficiency. In particular, equipment owners strive to avoid unplanned machine downtimes and costly component replacements and repairs that result from unexpected premature failures or gradual performance degradation caused by system wear [Mitchell, 1999; Hancock et al., 2006].

This work evolved from a particular need in industry to monitor the health of a machine prone to sporadic catastrophic breakdowns. Previous maintenance and condition monitoring schemes provided insufficient warning of the impending failure. A system that could distinguish between normal machine operation and an impending mechanical failure was needed, i.e., a fault classifier had to be put in operation.

In general terms, fault classification methods can be divided into two categories [Venkatasubramanian et al., 2003]: model-based and pattern recognition techniques. In model-based methods, faults can be detected by performing some mathematical calculations. For example, in the case of interest here, the state-of-the-practice entails oil pump failures being diagnosed by expert analysis of the parameter values measured during the monitoring time and their comparison with the nominal power curve of every oil pump: drawing the actual power curve according to the measured parameters values, i.e., by manual analysis, allows the analyst to identify whether any fault exists. Indeed, failed pumps often show hollow pumping action and energy waste. Because of the nonlinearity of the wear behavior and the size of the input data and their uncertainties, this way of proceeding requires significant human, material, and financial resources while not guaranteeing the timely detection of faults, thus seriously affecting production [Tian et al., 2007]. On the other hand, pattern recognition methods offer a framework that

can satisfy a number of basic requirements, such as short calculation time, high accuracy, and capability of dealing with large data sets representative of nonlinear wear behaviors [Zio, 2007]. Specially, soft computing approaches (e.g., Artificial Neural Networks and Fuzzy Logic systems) have shown superior robustness, speed, and accuracy compared to model-based methods [Shahrtash et al., 2008].

In pattern recognition methods, the conceptual basis for the detection of failure onset is that different system faults initiate different patterns of evolution of the interested variables, as measured by properly placed sensors [Zio et al., 2006]. Pattern recognition methods entail three different stages: feature extraction, feature selection, and classification [Sheng et al., 2004]. A number of pattern recognition methods have been proposed that differ in the classification stage, e.g., hierarchical trees (H-trees) [Breiman et al., 1984; Ripley, 1996; Loh et al., 1997], artificial neural networks (ANNs) [Rumelhart et al., 1986; Ripley, 1996; Zhang, 2000], and fuzzy logic (FL) systems [Zadeh, 1965; Klir et al., 1995; Zio et al., 2006; Wang et al., 2006; Wang et al., 2007].

H-trees evaluate the contribution of input features in determining the output classes of similarity. Generally, the most effective feature is selected as the first node of the tree, and its border value is used to create two different branches. Then, by the same criterion, the next most effective feature is found in each branch. This process is continued until the final nodes (leaves of the tree) obtained in all of the branches contain only the output classes. After creating the tree, a pruning process is performed to remove unnecessary nodes and decrease the size of the tree. Finally, within a supervised scheme, different sets of input features, obtained from various simulations, are fed to the tree in the test stage, and the outputs are compared with the known classes of each set to determine the accuracy of the tree. It is worth mentioning that there are different procedures for creating an H-tree. Different procedures may be applied to search for the best tree to solve a given problem, and then the best one can be selected by comparing the accuracy of the results and the time required to create the tree [Shahrtash et al., 2008].

ANNs are computing devices inspired by the function of nerve cells in the human brain and provide a powerful way to empirically model and forecast nonlinear systems based on data representative of their behavior [Rumelhart et al., 1986]. They are composed of many interconnected computing units, each performing a few simple

operations and communicating the results to its neighboring units. ANNs can learn to perform the mapping of the input-output relationships underpinning system behavior by a process of training on many different examples of input and corresponding output states [Hancock et al., 2006]. A main limitation of ANNs is that the results they deliver are difficult to interpret physically, and thus the underlying model remains cryptic.

FL modeling is designed to handle imprecise linguistic concepts, such as “small”, “big”, “young”, and “low”, and deal with uncertainties [Zadeh, 1965; Zio et al., 2006]. FL exhibits an inherent flexibility and has proven to be a successful modeling framework in a variety of industrial applications and pattern recognition tasks [Wang et al., 2006; Wang et al., 2007]. One of the main strengths of fuzzy logic modeling compared with other schemes is its capability of dealing with imprecise data [Marseguerra et al., 2004]. As for the limitations of fuzzy logic, the main difficulties stand in the fuzzy partitioning of the input and output spaces and in the establishment of the fuzzy rules that are at the basis of the classification phase and may require a time-consuming, trial-and-error process. Moreover, the elicitation of rules from human experts can be an expensive, error-prone procedure.

H-trees, ANNs, and FL have already been applied to real classification problems in a supervised scheme that entails, first, the classifier to be trained on data from known faults and then to be used to classify new data. In the present work, supervised classification schemes were precluded due to the unavailability of a comprehensive database of failure data, which is necessary for the training stage. Thus, as we shall see, the approaches adopted for detection do not require training. In other words, the classifiers are implemented for fault detection in an unsupervised manner, where the training and test phases collapse into the same clustering phase, and the class assignment is automated from available data of unknown classes. In particular, the adopted unsupervised FL approach, i.e. fuzzy clustering, exploits the advantages of automated generation of fuzzy rules, low computational burden, and benefits from the high-level, human-like rule representation typical of fuzzy systems, which offer an appealingly powerful framework for tackling practical classification problems.

Moreover, because of the shortage of data, the robustness of the classification approaches is augmented by combining multiple classifiers so as to improve upon the

performance of individual classifiers. The idea is to combine the predictions of multiple classifiers (for more details on the methodology, refer to Section 3.3) to reduce the variance of the results and the bias.

The paper is organized as follows. Section 2 presents a case study. The first part of Section 3 presents the procedures developed for the feature selection phase and the method with which the fuzzy rules are generated from the data set (Sections 3.2 and 3.3), respectively. The second part introduces an approach for fault detection based on binary h-Tree (Section 3.4). Section 4 reports the application to the classification of the oil pump into failed or safe status, based on the available vibration data. The motivation for this application comes from the interest of a producer of oil extracted from oil sands in developing a proper monitoring scheme to detect pump failures in a system aimed at moving large amounts of raw oil sand. The reason for the failures is not exactly known, although it has been conjectured that the main contribution to pump wear is the large flow of oil into the vanes and the presence of small particle of dirt and sand in the sucked fluid [LaBour, 1995; Frith et al., 1996]. The monitoring scheme is expected to provide advance warning and lead time to prepare the appropriate corrective actions. Finally, a comparison of the two classification approaches proposed is provided, highlighting the advantages and limitations of the proposed methodologies.

2. The Case Study

In this research, experimental data were collected from a number of slurry pumps that are used to deliver a mixture of bitumen, sand, and small pieces of rock from one site to another. For each pump, vibration is monitored as a symptom of system health. Vibration signals have been collected at the inlet and outlet of slurry pumps operating in an oil sand mine. The pump vibration data were collected by the mine staff and one of the authors using the Smart Asset Management System (SAMS) and then further analyzed using the proposed classification methods. SAMS is a PC-based virtual instrument used to perform machine health monitoring [Tse, 2002]. Its measurement platform provides a Graphical User Interface that allows the user to choose from different diagnostic techniques (e.g., higher order statistical analysis and orbit analysis) to conduct machinery fault diagnosis. It can be installed in a notebook PC or desktop computer for portable or

continuous machine health monitoring. SAMS also provides an easy-to-use interface for data management, report generation, trend analysis, etc., to help the maintenance staff in the recording and planning of maintenance activities.

The data acquisition equipment (DAQ) consist of a National Instrument (NI) cDAQ 9172 and a DAQ module NI 9234. Their specifications are listed in Table 1.

Table 1: Measurement Equipment

	Equipment	Model	Specification
1	DAQ Device	NI cDAQ 9172	Max. support module = 8
2	DAQ Module	NI 9234	Resolution = 24 bit Input range = +/- 5V Sampling rate = up to 51.2 kHz per channel
3	Smart Asset Maintenance System (SAMS)	Version 2.3.8	
4	Notebook Computer	IBM T60	Intel Core Duo processor 1.66 GHz Windows XP Professional
5	Accelerometer	PCB 352A60	Mounted on positions Casing Lower and Casing Discharge Sensitivity = 10.2 mV/g
6	Accelerometer	PCB 352C18	Mounted on positions Suction Pipe and Discharge Pipe Sensitivity = 9.7 mV/g

Vibrations were measured by four accelerometers mounted in four different positions so that there were a total of four different vibration signals captured. They are denoted as S1, S2, S3, and S4. As shown in Figure 1, accelerometers S1 and S2 were PCB 352A60 accelerometers (see Table 1) that were mounted on the case of the pump and denoted as ‘Casing Lower’ and ‘Casing Discharge’, respectively. Accelerometers S3 and S4 were PCB 352C18 accelerometers (see Table 1) mounted on the suction and discharge pipes, respectively. All four accelerometers captured the vibration signals from four different positions at a similar sampling frequency rate of 50 kHz.

In Figure 2, the layout of the oil extraction site is represented. It consists of two parallel lines, L1 and L2, each composed of four different pumps. The pumps located in L1 are called G1, G2, G3, and G5, whereas G1, G2, G3, and G4 are those located in L2. Each pump is different in type, size, and working condition, i.e., ground elevation, process fluid, history, and wear.

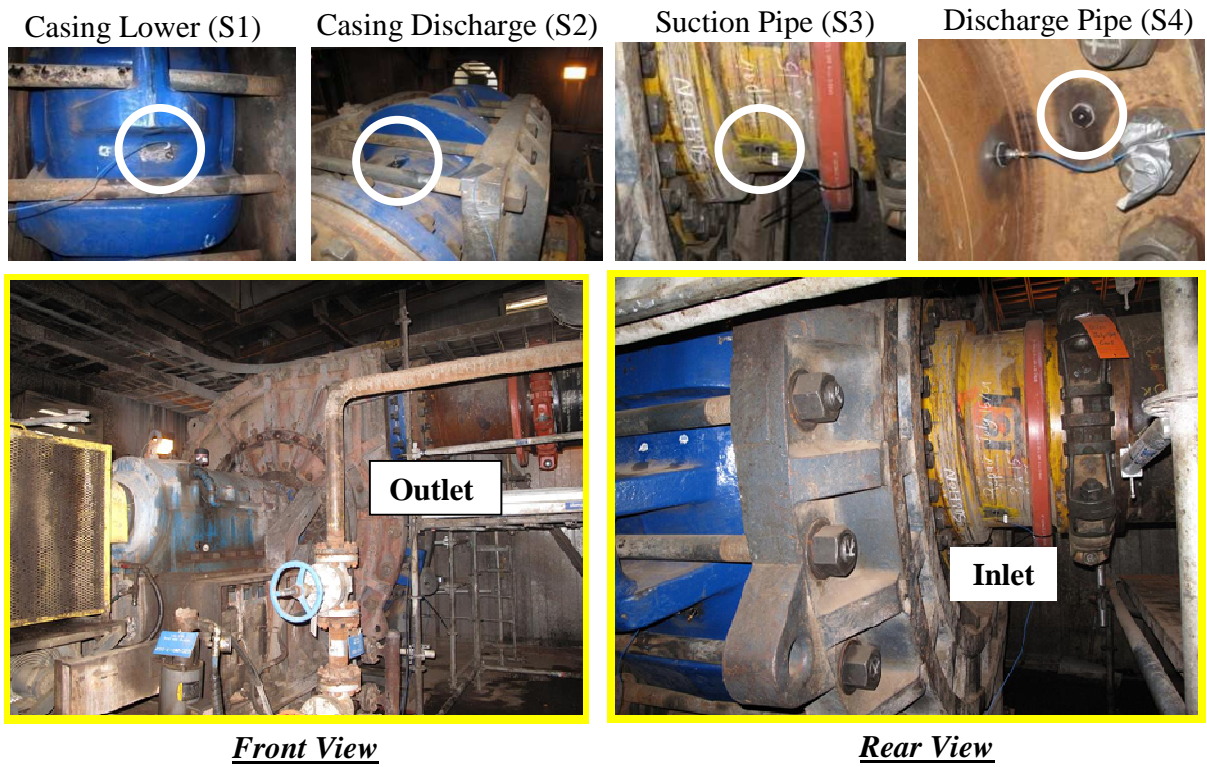


Figure 1: Samples of the measurement locations of a slurry pump.

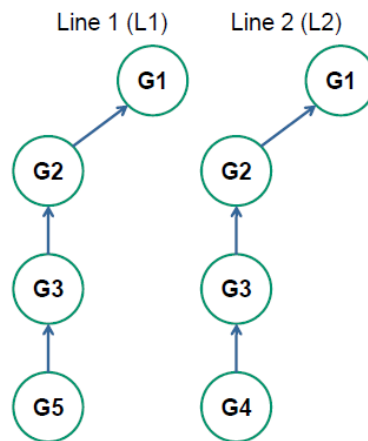


Figure 2: The oil extraction lines layout.

Only 11 batches of 4 sensory vibration signals are available in total. These degradation patterns are representative of different stages of progressive pump deterioration. The number of patterns for each pump is listed in Table 2.

Table 2: List of Available Degradation Patterns for Each Pump in L1 and L2. NA=Not Available

L1	Available degradation patterns	L2	Available degradation patterns
G1	1	G1	2
G2	1	G2	3
G3	2	G3	NA
G5	1	G4	1

Each degradation pattern is composed of 30 intervals of records, each one lasting 1.3 [s], with pauses of 2 [s] in between (Figure 3).

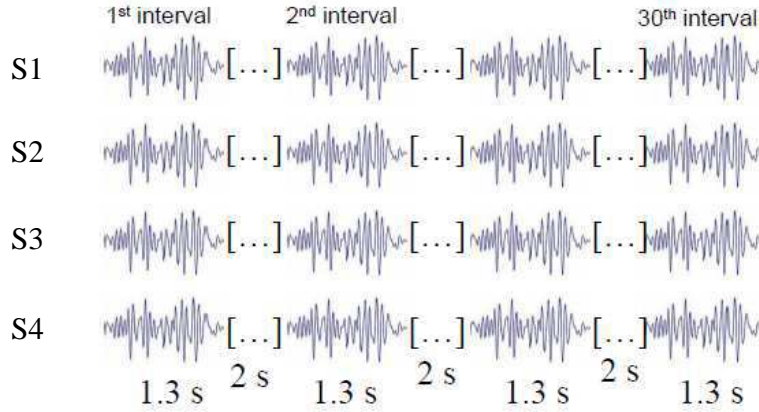


Figure 3: Sketch of the degradation pattern structure.

A preliminary analysis of the data showed that smooth and gradual degradation of the pump performance occurred (except for catastrophic failures), such that there was no significant deviation of the signal along the total 40 seconds of records. Thus, to lighten the computational burden of the data treatment, we have concentrated our analysis only on the records from the first 1.3 [s] (65000 points), discarding the remaining 29 intervals, assuming that the pump is either failed or healthy at time 0 [s].

3. Methods

As mentioned in the Introduction, pattern recognition methods entail three different stages: feature extraction, feature selection, and classification [Sheng et al., 2004]. In the

following, these procedural steps are detailed with reference to the particular case study introduced in Section 2.

3.1. Feature extraction

Ten features were selected and extracted from the batches of vibration data collected by the accelerometers. For each of the 11 degradation patterns, the following $M=10$ indexes were evaluated (N is equal to 65000 sampling points):

$$(1) \text{ Peak value: } \max = \max_{j=1, \dots, N} n_j$$

$$(2) \text{ Mean: } u = \frac{1}{N} \sum_{j=1}^N n_j$$

$$(3) \text{ Standard deviation: } \sigma = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (n_j - u)^2}$$

$$(4) \text{ Root mean square: } RMS = \sqrt{\frac{1}{N} \sum_{j=1}^N (n_j)^2}$$

$$(5) \text{ Skewness: } SK = \frac{\sum_{j=1}^N (n_j - u)^3}{(N-1)\sigma^3}$$

$$(6) \text{ Kurtosis: } KU = \frac{\sum_{j=1}^N (n_j - u)^4}{(N-1)\sigma^4}$$

$$(7) \text{ Crest indicator: } CI = \frac{\max_{j=1, \dots, N} |n_j|}{\sqrt{\frac{1}{N} \sum_{j=1}^N (n_j)^2}}$$

$$(8) \text{ Clearance indicator: } CLI = \frac{\max_{j=1, \dots, N} |n_j|}{\left(\frac{1}{N} \sum_{j=1}^N \sqrt{|n_j|}\right)^2}$$

$$(9) \text{ Shape indicator: } SI = \frac{\sqrt{\frac{1}{N} \sum_{j=1}^N (n_j)^2}}{\frac{1}{N} \sum_{j=1}^N |n_j|}$$

$$(10) \text{ Impulse indicator: } MI = \frac{\max |n|}{\frac{1}{N} \sum_{j=1}^N |n_j|}$$

3.2. Feature selection

The objective of feature selection is three-fold: to improve the performance of the classifier, provide faster and more cost-effective classification, and provide a better understanding of the underlying process that generated the data [Guyon et al., 2003]. Depending on the nature of the regression technique, the presence of irrelevant or redundant features can cause the system to focusing attention on the idiosyncrasies of the individual samples while losing sight of the broad relational picture that is essential for generalization beyond the training set. This problem is compounded when the number of observations is also relatively small. If the number of variables is comparable to the number of training patterns (as in our case study), the parameters of the model may become unstable and are unlikely to be replicated if the study were to be repeated. Feature selection seeks to remedy this situation by identifying a small subset of relevant features and using only them to construct the actual model. In this work, the selection of the most relevant features to be used in the classification phase is based on two standard approaches based on commonly used statistical tools: 1) box plots and 2) correlation plots.

3.2.1. Feature selection based on box plots analysis

Box plots provide an excellent visual summary of many important aspects of a distribution and are useful for identifying its outliers [Massart et al., 2005]. The conceptual basis for using box plots in distinguishing the most relevant features for classification is that things can be distinguished from each other based on their inconsistency [Hsiao et al., 2009]. Outliers can in fact be used as a primary method for pattern classification: the more outliers a parameter distribution has, the more that parameter will be useful in defining clusters in the feature space defined by the considered parameter while avoiding cluster overlapping.

There are several steps in constructing a box plot. The first relies on the evaluation of the 25th, 50th, and 75th percentiles in the distribution of the 11 patterns. Figures 4-7 show how these three statistics are used in our case study: for each extracted feature, we draw a box extending from the 25th percentile to the 75th percentile. The

50th percentile is drawn inside the box. We also put “whiskers” above and below each box to give additional information about the spread of data. Whiskers are vertical lines that end in a horizontal. They are drawn from the lowest and upper hinges to the lowest datum still within 1.5 Inter Quartile Range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile, respectively [Massart et al., 2005]. Finally, we represent outliers in our box plots by adding additional crosses beyond the whiskers.

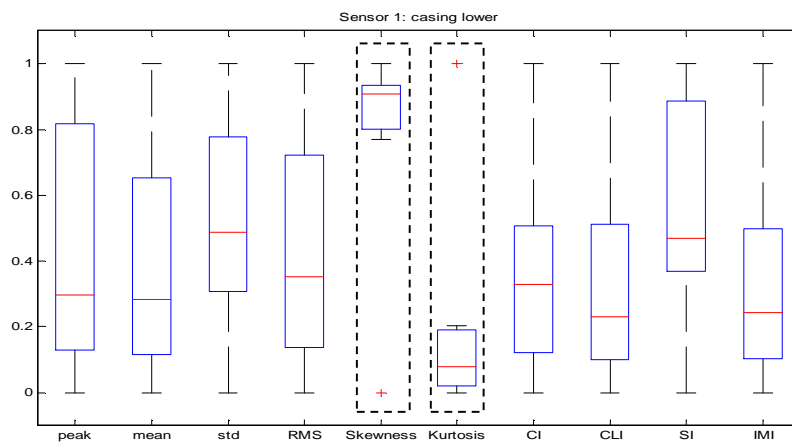


Figure 4: Box plots of the distributions of the 10 extracted features from sensor S1.

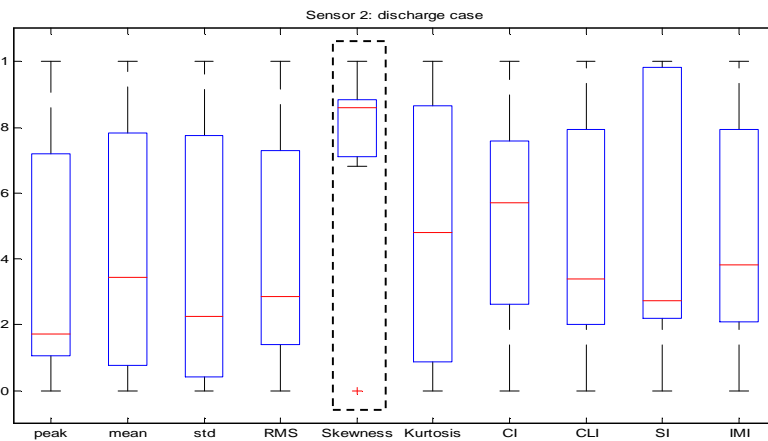


Figure 5: Box plots of the distributions of the 10 extracted features from sensor S2.

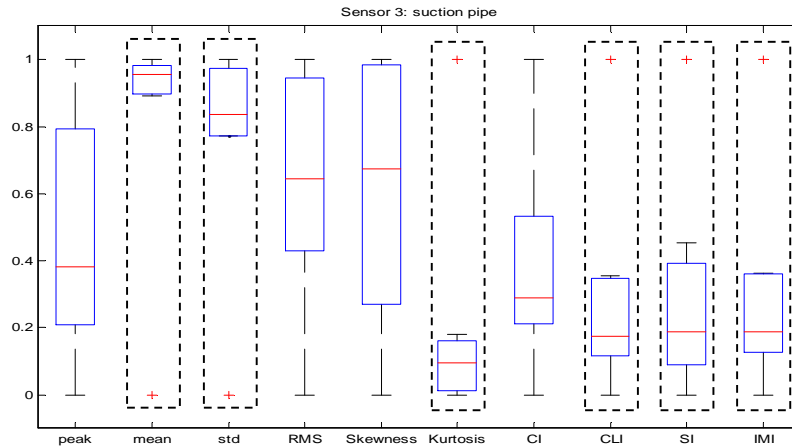


Figure 6: Box plots of the distributions of the 10 extracted features from sensor S3.

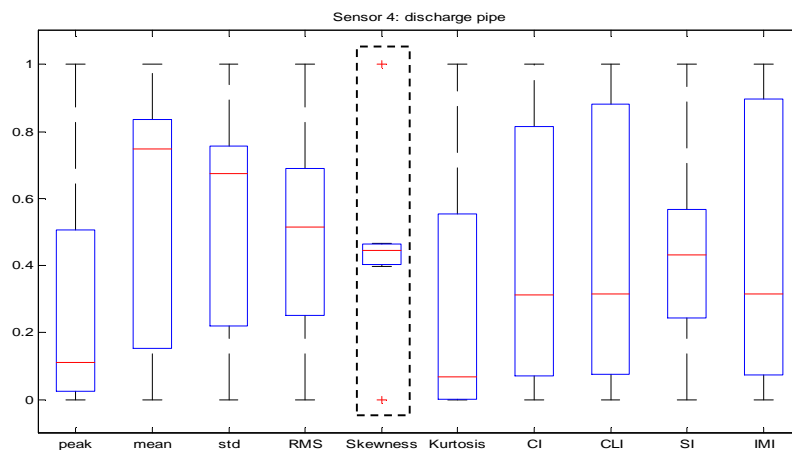


Figure 7: Box plots of the distributions of the 10 extracted features from sensor S4.

From the analysis of Figures 4–7, the relevant features for each sensor were:

- S1: skewness and kurtosis
- S2: skewness
- S3: mean, standard deviation, kurtosis, clearance indicator, shape indicator and impulse indicator
- S4: skewness

Most of the four signal box plots highlight skewness and kurtosis spread distributions. Thus, these two features were considered to be key features on which the classification of the degradation patterns would be based.

3.2.2. Feature selection based on correlation plots analysis

Correlation plots are suitable tools for feature selection, because it is well known that good feature subsets contain features uncorrelated with each other [NIST, 2010]. The number of correlation coefficients to be evaluated and compared was equal to 45, which follows from the definition of k combinations from a given set of M elements, wherein our case M is equal to 10 (i.e., the number of extracted features) and k is equal to 2.

The relevant features for each sensor are those with the smallest value of correlation among the 45 combinations. From the analysis of Figures 8–11, it can be seen that, for each sensor, the selected features are (highlighted in the Figures with dotted circles):

- S1: standard deviation and skewness
- S2: mean and root mean square
- S3: root mean square and skewness
- S4: skewness and clearance indicator

A comparison of the different features suggested for selection by the two different approaches (box plots and correlation plots) reveals the substantial qualitative agreement on the importance of skewness values in determining the cluster boundaries. In conclusion, both analyses, e.g., box and correlation plots, seem to be useful tools applicable to systematically provide information on the relevant parameters to be selected for optimal performance of the successive classification tasks.

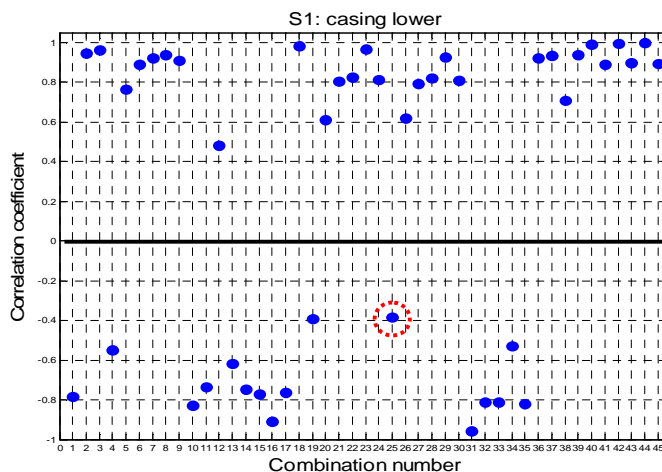


Figure 8: Correlation coefficients for 45 combinations of the 10 features relative to sensor S1.

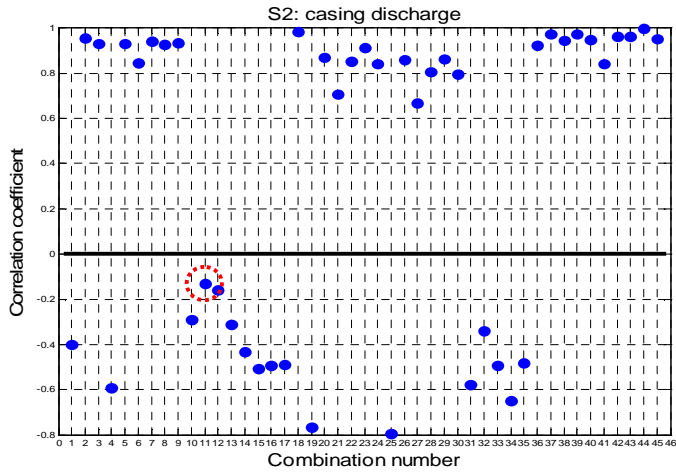


Figure 9: Correlation coefficients for 45 combinations of the 10 features relative to sensor S2.

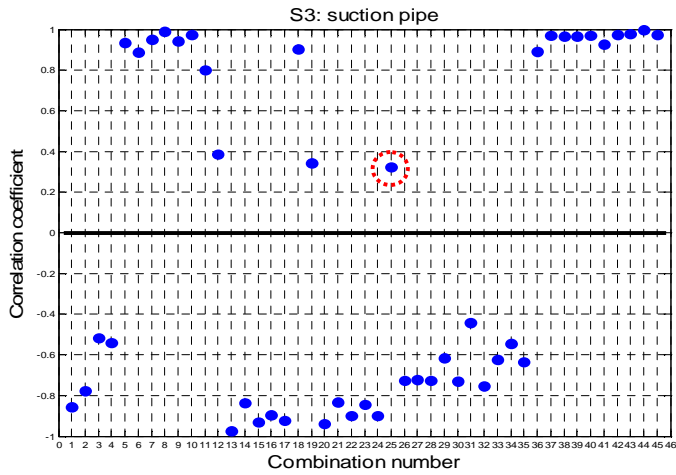


Figure 10: Correlation coefficients for 45 combinations of the 10 features relative to sensor S3.

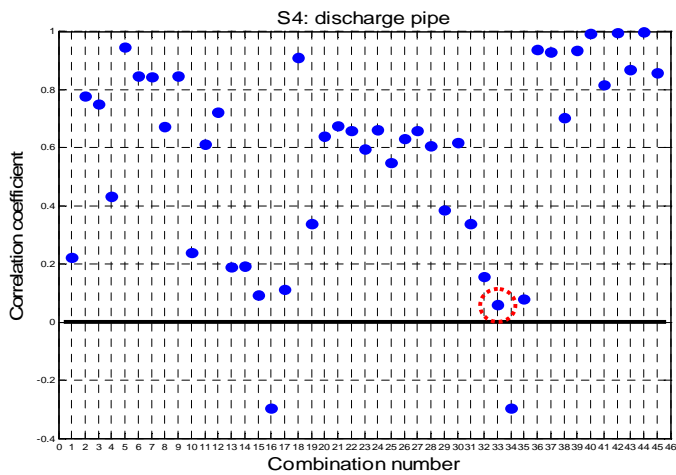


Figure 11: Correlation coefficients for 45 combinations of the 10 features relative to sensor S4.

3.3. Classification

Fault detection may pose difficulties, because it entails the implementation of a classifier for labeling the component status as healthy or failed. In our application, the shortage of data forces us to resort to a combination of classifiers in an attempt to improve the detection-classification performance [Freund et al., 1996; Schapire, 1999; Friedman, 2000]. Figure 12 illustrates the basic framework for the ensemble scheme adopted and that was applied with two different classification approaches (Sections 3.3.1 and 3.3.2). The key step was the formation of an ensemble of diverse classifiers from a single data set. In this work, four different classifiers were fed with different inputs taken from different sensors (S1, S2, S3 and S4), but all relative to the same degradation pattern. The single classifier results were then combined by two different methodologies [Friedman et al., 2000]:

- Majority voting
 - Each ensemble member votes for one of the classes.
 - Predicts the class with the highest number of vote.
 - In case of equal number of votes, the class is labeled as uncertain.
- Weighted voting
 - Make a weighted sum of the votes of the ensemble members.
 - Weights depend on the performance of each independent classifier.

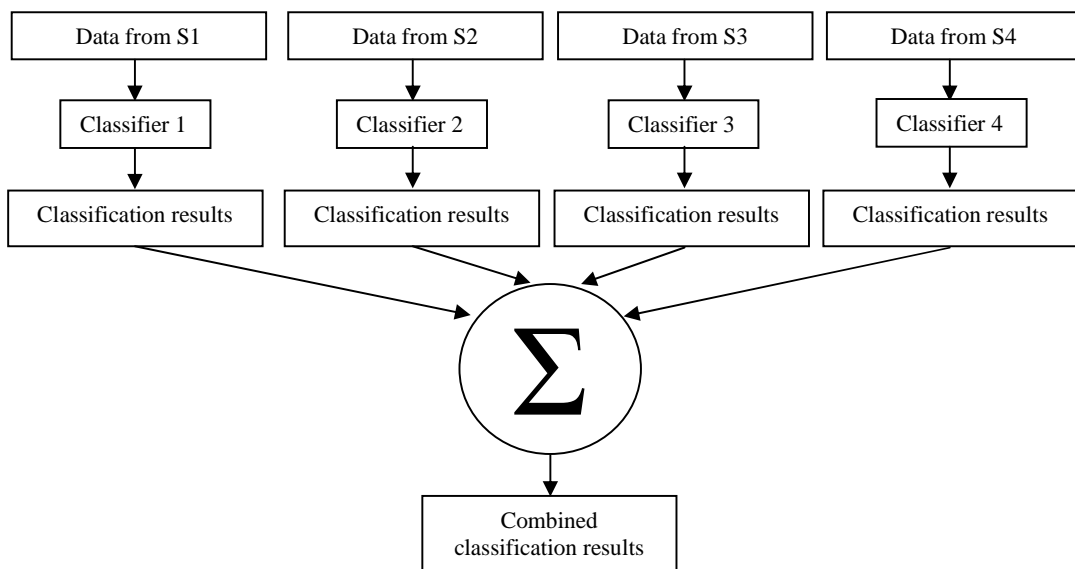


Figure 12: Scheme of combination of different classifiers.

We proceeded to generate results using the two different approaches. We varied the type of single classifier implemented and compared their performances. In this Section, we focus on the two unsupervised classification approaches employed for fault detection: 1) unsupervised Fuzzy C-Means (FCM) algorithm and 2) Binary Hierarchical Clustering Tree (H-Tree).

3.3.1. The Unsupervised Fuzzy C-Means algorithm

Fuzzy C-Means (FCM) is one of the most popular fuzzy clustering methods [Bezdek, 1981; Leguizamon et al., 1996; Alata et al., 2008]. The FCM method originated from hard C-Means clustering, allowing data points to belong to two or more clusters [Klir et al., 1995]. The clusters emerged from the minimization of the following objective function:

$$J(N, C) = \sum_{i=1}^N \sum_{j=1}^C m_{ij}^{\varphi} d_{ij}^2(x_i, c_j) \quad (1)$$

where $J(N, C)$ is the sum of the square errors of the distance of each individual data point x_i , $i=1,2,\dots,N$, to the center c_j , $j=1,2,\dots,C$, of the given cluster (class) j . The minimization is done with respect to the membership m_{ij}^{φ} and the centers c_j . More specifically, $d_{ij}^2(x_i, c_j)$ is the square of the distance between x_i and c_j , whereas m_{ij}^{φ} is the degree of membership of x_i to cluster j . The value φ is any real number greater than 1, and it modulates the fuzziness of the clusters.

Fuzzy partitioning is carried out through an iterative optimization of $J(N, C)$, with the update of memberships m_{ij}^{φ} and the cluster centers c_j by:

$$m_{ij}^{\varphi} = \frac{d_{ij}^{-2(\varphi-1)}}{\sum_{j=1}^c d_{ij}^{-2(\varphi-1)}} \quad (2)$$

$$c_j = \frac{\sum_{i=1}^n m_{ij}^{\varphi} x_i}{\sum_{i=1}^n m_{ij}^{\varphi}} \quad (3)$$

For further details, the interested reader may refer to [Bezdek, 1981].

3.3.2. Binary Hierarchical Tree

Hierarchical clustering groups data over a variety of scales by creating a cluster tree, also called a dendrogram. The binary hierarchical tree is not a single set of clusters, but rather a multilevel hierarchy where crisp clusters at one level are joined as crisp clusters at the next level, i.e., the pattern belongs to one and only one cluster. This allows us to decide the level or scale of clustering that is most appropriate for the application of interest. Although many other clustering methods exist, hierarchical clustering is one of the most commonly used methods in crisp clustering [Breiman et al., 1984; Ripley, 1996; Loh et al., 1997].

A hierarchical clustering algorithm initially places each data point in a cluster by itself, and then it recursively and merges the clusters that are closest to each other according to some distance or similarity criterion until a single dendrogram is created.

4. Results

The classification systems introduced in Section 3 were developed using some of the data from a database of 11 degradation patterns of oil vane pumps, as described in Section 2. These degradation patterns are representative of different stages of progressive pump deterioration. Moreover, they refer to pumps that are different in type, size and working conditions i.e., ground elevation, process fluid, history, and wear. Thus, in order to analyze only pumps subjected to similar working conditions, we only selected the degradation patterns relative to G1 and G2 from lines L1 and L2. Finally, the total number of available degradation patterns to be classified is 7. The real health status of the pumps relative to the 7 degradation patterns is listed in Table 3. This list is only used at the end of the unsupervised fault detection process for comparison with the obtained results and performance evaluation of the classifiers.

Table 3: Actual Health Status of the Pumps Relative to the 7 Degradation Patterns Considered (H=healthy, F=failed)

Degradation pattern	1	2	3	4	5	6	7
Actual health status	H	H	F	H	H	H	H

The attributes for the pump vibrational data were extracted from the current signals, as introduced in Section 3.1. A total of four feature sets, each composed of 7×10 values, were collected. Then, among these features, the most important were identified according to box plots and correlation plots. Finally, the classification tools of FMC and H-tree were fed with the values of selected features for building the clusters in an unsupervised manner.

In what follows, the results for all the considered combinations of approaches for feature selection and classification are reported.

Method I:

Feature Selection: Box Plot Analysis

Classifier: Fuzzy C-Means

The features to be fed to the FCM classification algorithm were selected by box plot analysis. In Section 3.2.1 we justified the choice of skewness and kurtosis as important features. The classification phase identified 2 clusters that can be useful for labeling the degradation patterns as relative either to failed or to healthy pumps. In this case, the classification results are shown in Figure 13. By analyzing the skewness and kurtosis values of the considered degradation patterns plotted on the scatter plot of Figure 12, it turns out that the main differences between the two identified clusters (represented by circles and crosses) are:

- Circles have skewness values close to zero and lower kurtosis values.
- Crosses have skewness values far from zero and higher kurtosis values.

Based on this consideration we have decided that:

- Degradation patterns with skewness values close to zero, i.e., vibrational data normally distributed, are relative to healthy pumps (the flow of abrasive and erosive particles can only generate white noise on the measurements).
- Degradation patterns with skewness values far from zero, i.e., vibrational data not normally distributed, are working in anomalous conditions (failed components highly deform the parameter distributions).

Thus, hereafter, circles are labels of degradation patterns for safe pumps, whereas crosses indicate the class of failed pumps.

Method II:

Feature Selection: Box Plot Analysis

Classifier: Hierarchical Clustering Tree

In this application, the same features selected by the box plot approach and used in Method I have been fed into the h-Tree algorithm. The classification results are shown in Figure 14.

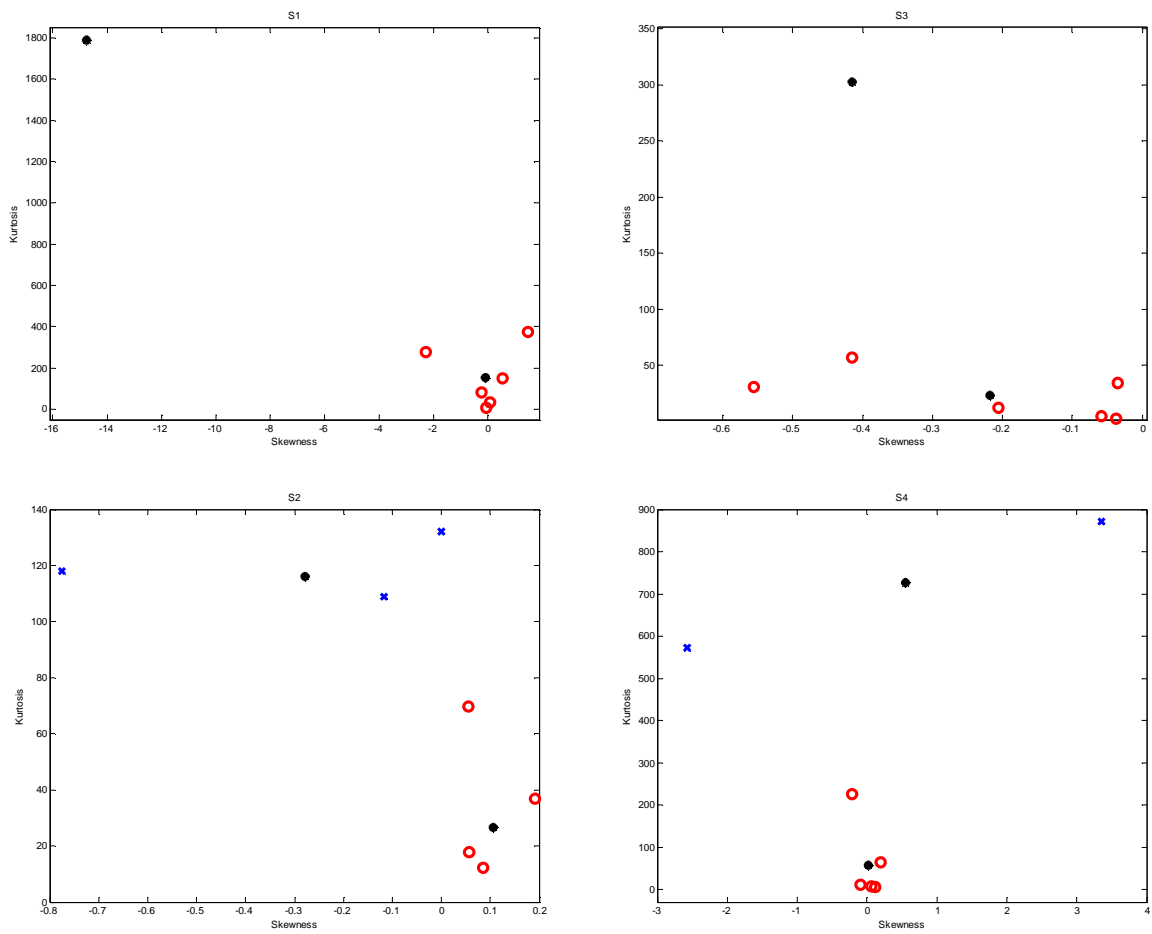


Figure 13: Classification results using box plots and an unsupervised FCM classifier algorithm: circles are healthy pumps; crosses are failed pumps; dots are clusters centers.

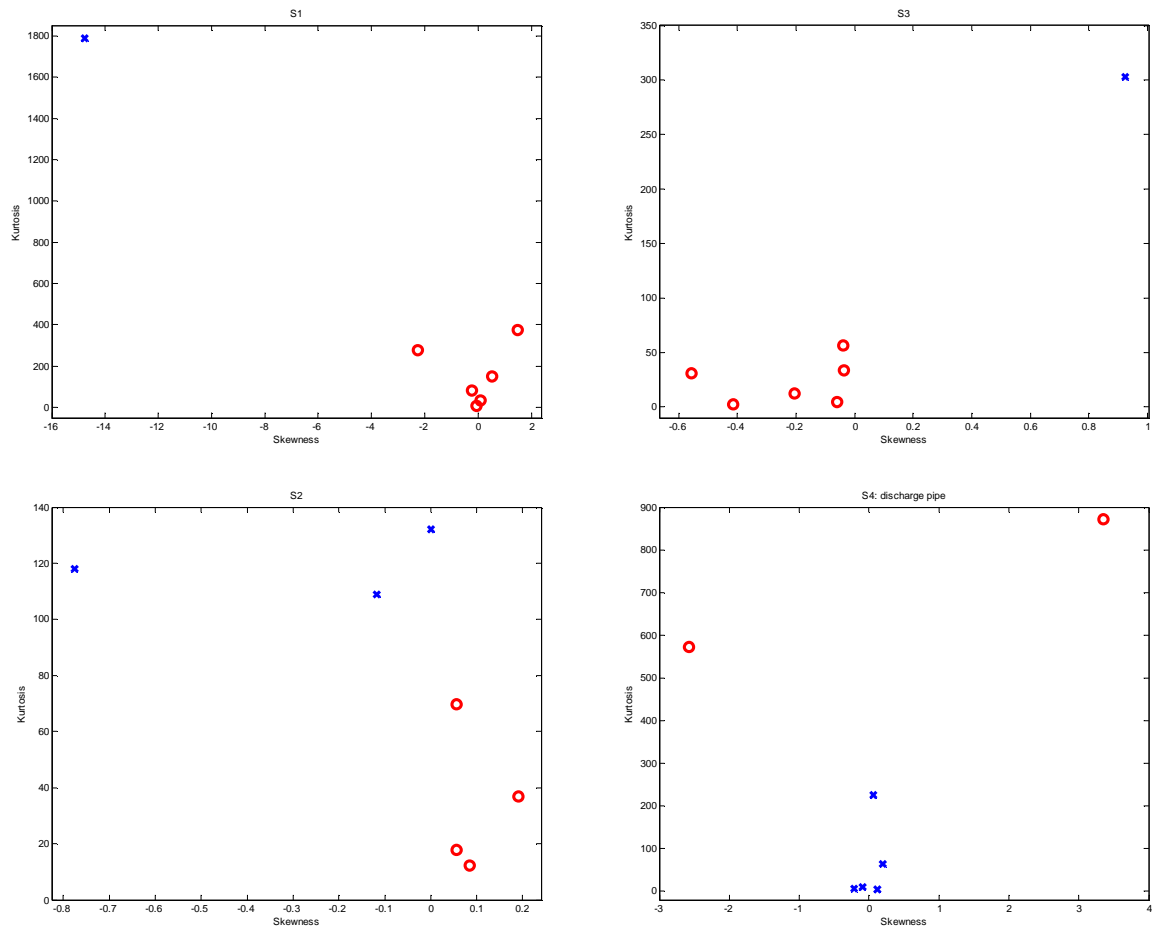


Figure 14: Classification results using box plots and h-Tree classifier algorithm: circles are healthy pumps, whereas crosses are failed pumps.

To compare the effectiveness of Method I and Method II, the classification results based on the batch of four sensors, S1, S2, S3, and S4, from the 7 degradation patterns have been listed in Table 4. Based on the FCM classifier algorithm, when using majority voting, the correctness of the classification was 86%, with an uncertain assignment percentage equal to 14%; whereas, using weighted voting, the correctness of the estimations reached 100%. Resorting to h-Tree, only 57% of the data was classified correctly and 43% was uncertain, when using majority voting; the weighted voting approach raised the correct estimation percentage to 86%, but 15% of the data were misclassified.

Table 4: The Comparative Classification Results Using Unsupervised FCM and H-tree, with Box Plots as Selecting Methods for the Input Features (H=healthy, F=failed, ?=uncertain)

Degradation pattern	1	2	3	4	5	6	7	Correct	Uncertain
FCM (majority voting)	H	?	F	H	H	H	H	86%	14%
FCM (weighted sum)	H	H	F	H	H	H	H	100%	0%
h-Tree (majority voting)	?	H	?	H	?	H	H	57%	43%
h-Tree (weighted sum)	H	H	F	H	F	H	H	86%	0%

The analysis of Table 4 highlights that:

- Fault detection based on FCM performed better than h-Tree in all cases because it allowed building clusters with uncertain boundaries accommodating for different pump locations and different pump types and sizes.
- Cluster centers identified only by the FCM turned out useful during on-line fault detection for classifying a new developing degradation pattern into healthy/failed clusters according to the distances of the feature values from the centers.

Method III:

Feature Selection: Correlation Coefficients

Classifier: Fuzzy C-means

The features to be fed into the single FCM algorithm have been selected by correlation plot analysis. In Section 3.2.2 we justified the choice of different input features for each single classifier. Classification results are shown in Figure 15 and summarized in Table 5.

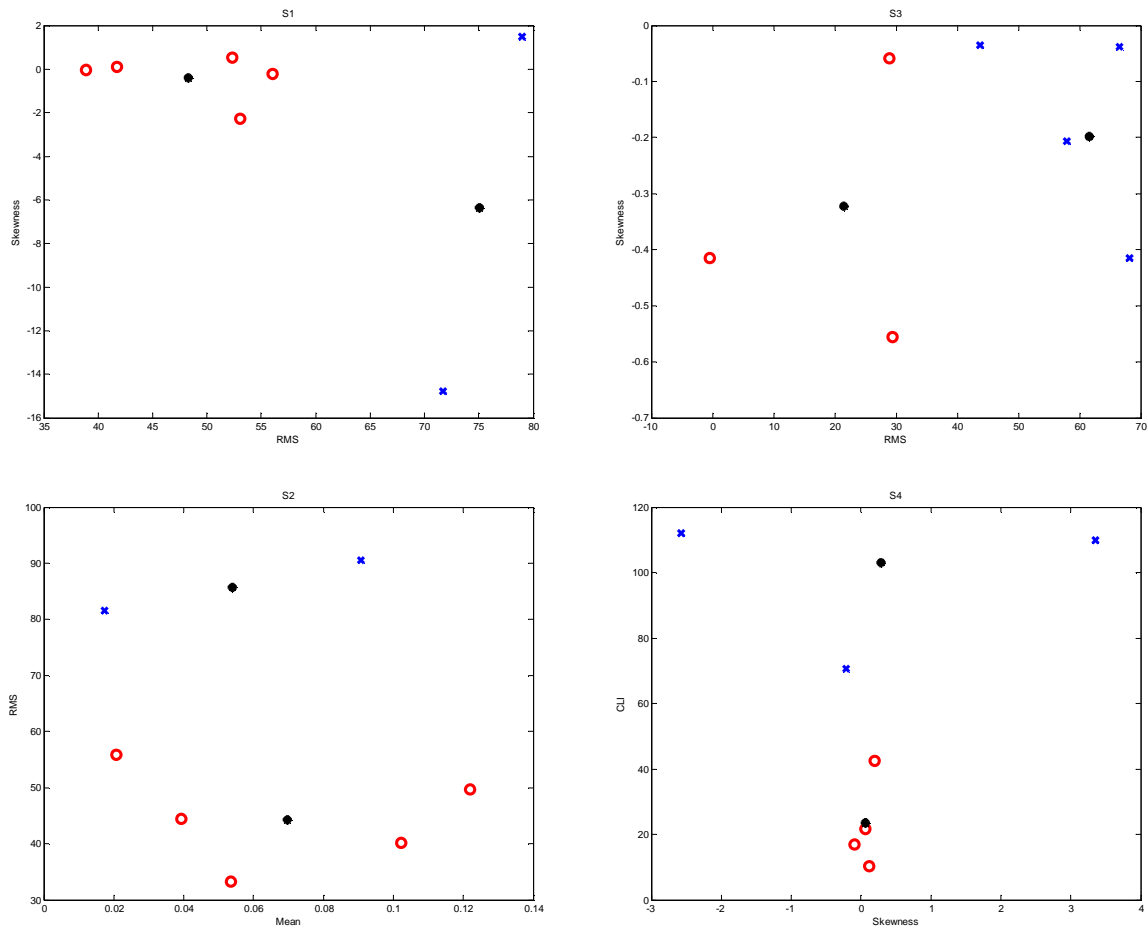


Figure 15: Classification results using correlation plots and unsupervised FCM classifier algorithm: circles are healthy pumps, crosses are failed pumps, and dots are cluster centers.

Method IV:

Feature selection: Correlation coefficients

Classifier: Hierarchical Clustering Tree

The features to be fed to the single FCM algorithm have been selected by correlation plot analysis. The classification results are shown in Figure 16 and summarized in Table 5, together with those obtained using Method III. By comparison with Table 4, it can be seen that:

- When relying on FCM, correct assignment performance is almost constant, irrespective of the feature selection approach used.

- When relying on FCM, the percentage of uncertain estimations increases when the combination of single classifier results is made by a majority voting strategy.
- In general, fewer degradation patterns are assigned to the wrong class.

In conclusion, FCM performed better than H-tree regardless of which features were used as input for the single classifiers, because of the higher percentage of correct class assignment for all the classification methods analyzed. Indeed, h-Trees are not very accurate because they can only ideally deal with imperfect data, outliers, different scales, and irrelevant attributes [Friedman, 2000].

Table 5: The Comparative Classification Results Using Unsupervised FCM and PC+HT, with Correlation Plots as the Selecting Methods for the Input Feature Variables (H=healthy, F=failed, ?=uncertain)

Degradation pattern	1	2	3	4	5	6	7	Correct	Uncertain
FCM (majority voting)	?	H	F	H	H	?	H	71%	29%
FCM (weighted sum)	H	H	F	H	H	H	H	100%	0%
HT (majority voting)	H	H	F	H	?	H	H	86%	14%
HT (weighted sum)	H	H	F	H	F	H	H	86%	0%

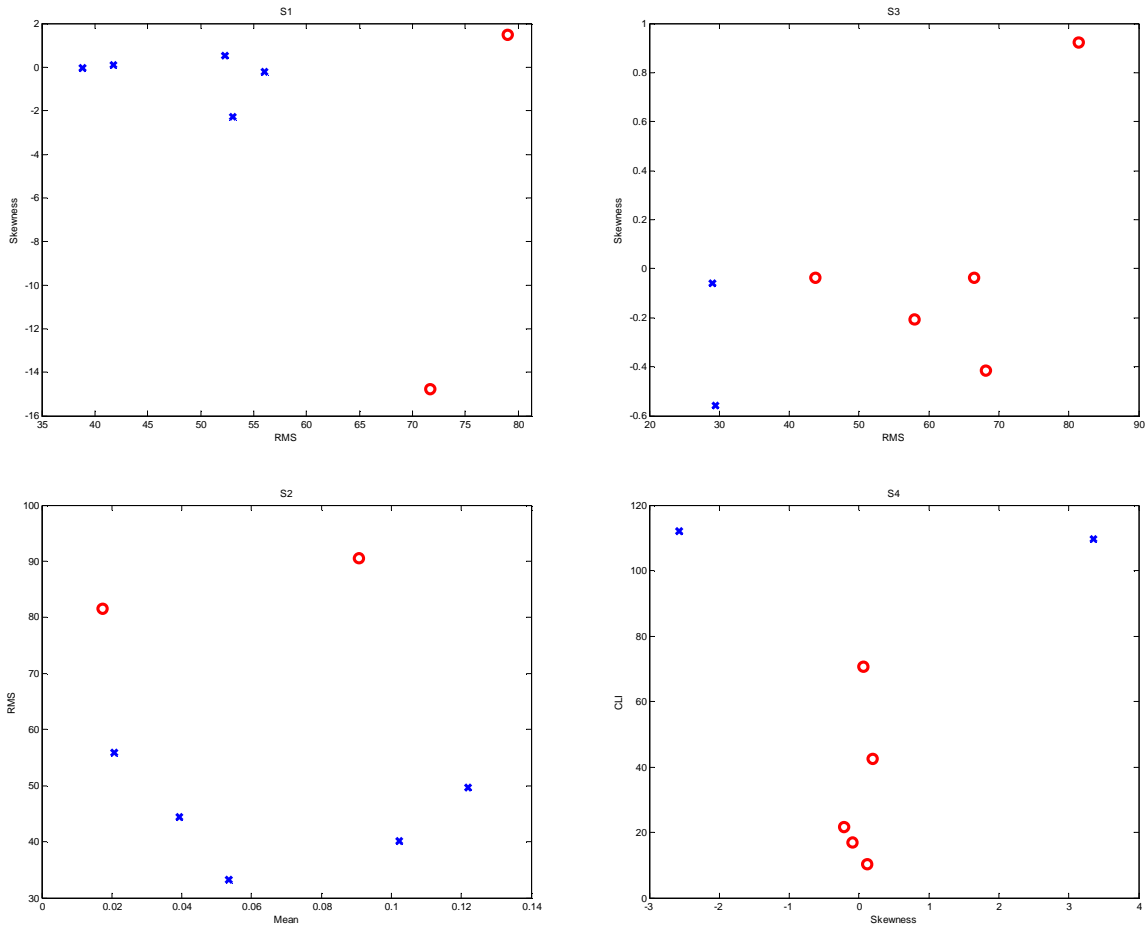


Figure 16 Results using correlation plots and H-tree: circles are healthy pumps, and crosses are failed pumps.

5. Estimation of the Pump Wear Function (PWF)

When components are well-maintained, it allows downtime to be reduced for the sake of plant safety and overall performance efficiency. Since machines often go through degradation before failure, monitoring and predicting the trend of their degradation and condition may allow for correction before failure. Indeed, when the conditions of a component or structure can be monitored, maintenance can be planned dynamically [Williams et al., 1994; Marseguerra et al., 2002]. By predicting the future evolution of the degradation state of a component or structure, it is possible to verify whether it can continue performing the required function. In practice, the estimate of the system state in terms of wear function may be difficult to obtain because the degradation state may not

be directly observable and/or the measurements may be affected by noise and disturbances.

In this Section, a method for constructing dynamic models for system wear estimation is proposed. Instead of assuming a structure for the wear model and identifying its parameters, only an approximate linear model is used. This results in a simple model that can be used as a monitoring tool for online application. Addition of noise, ε , during state estimation is used to reflect inherent process variability. According to verified superior classification performance guaranteed by the features skewness (S) and kurtosis (K) (Section 4), these two features relative to the same 7 degradation patterns under analysis are considered as inputs for modeling the Pump Wear Function PWF. Their values have been fitted to a linear model for estimating PWF, whose function turns out to be equal to:

$$\text{PWF}(S, K) = -0.9 + 5.6 \times S + 0.05 \times K - 0.003 \times S \times K + \varepsilon \quad (4)$$

The fitted plot is shown in Figure 17. As expected, the larger the value of skewness, the larger the pump wear; similarly, the larger the kurtosis value, the larger the pump wear. This confirms the adopted hypotheses for labeling the classified degradation patterns as either healthy or failed.

However, it has to be pointed out that at this research stage the estimated PWF (Eq. 4) can only be considered as a trial approximation of the real PWF. Additional vibrational data coming from the same type and size of pumps will confirm the accuracy of the PWF estimation of wear. Finally, experimental implementation of the wear monitoring system will be necessary to license the proposed monitoring scheme for tracking oil pump wear.

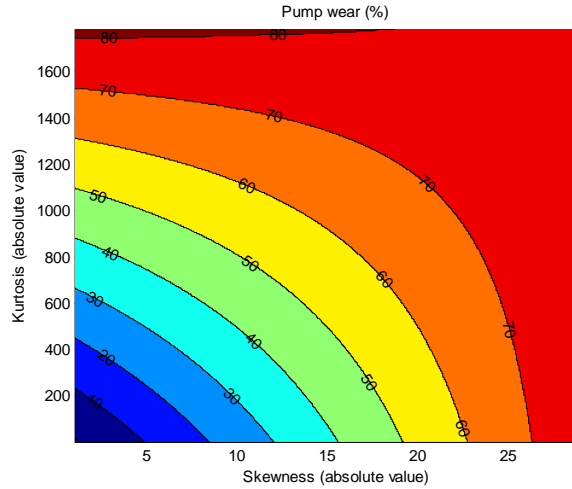


Figure 17: Pump wear function.

5. Conclusions

For effective maintenance, industry needs to monitor the health of a machine prone to degradation and sporadic catastrophic breakdowns. To achieve this aim, a system is needed that can distinguish between normal machine operation and an impending mechanical failure, i.e., a fault classifier.

In this application, the focus is on the wear of slurry pump components, caused by abrasive and erosive solid particles. Unscheduled outages of pumps cost oil and gas companies millions of dollars each year. Traditional maintenance strategies provide insufficient warning of an impending failure.

In this work, we have presented two unsupervised clustering ensemble methods (h-Tree and FCM) and compared their performance in the assessment of the wear status of pumps when available data is extremely limited.

In particular, the adopted unsupervised FCM approach exploits the advantages of the automated generation of fuzzy rules, low computational burden, and the high-level, humanlike thinking and reasoning of fuzzy systems, which offer an appealingly powerful framework for tackling practical classification problems.

Finally, for prognostics, the future evolution of the degradation state of the component or structure is predicted by a simple dynamic model for system wear estimation to verify whether it can continue performing the required function.

References

- (Alata et al., 2008) M. Alata, M. Molhim, and A. Ramini, 2008, "Optimizing of Fuzzy C-Means Clustering Algorithm Using GA", World Academy of Science, Engineering and Technology, p.224-229.
- (Bezdek, 1981) J.C.; Bezdek, 1981, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum, New York.
- (Breiman et al., 1984) L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, 1984, "Classification and regression trees", Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.
- (Freund et al., 1996) Y. Freund, R.E. Schapire, 1996, "Experiments with a New Boosting Algorithm", Machine Learning: Proceedings of the Thirteenth International Conference, pages 148-156.
- (Friedman, 2000) J. Friedman, 2000, "Predictive Learning through Gradient Boosting", Keynote Address, Seventeenth International Conference on Machine Learning, Stanford University.
- (Friedman et al., 2000) J. Friedman, T. Hastie, R., Tibshirani, 2000, "Additive Logistic Regression: a Statistical View of Boosting", Ann. Statist., Volume 28, Number 2, 337-407.
- (Frith et al., 1996) R.H. Frith, W. Scott, 1996, "Comparison of an External Gear Pump Wear Model with Test Data", Wear, Volume 196, Issues 1-2, Pages 64-71.
- (Guyon et al., 2003) I. Guyon, A. Elisseeff, 2003, "An Introduction to Variable and Feature Selection", Journal of Machine Learning Research 3, 1157-1182.
- (Hancock et al., 2006) K. M. Hancock, Q. Zhang. 2006, "A Hybrid Approach to Hydraulic Vane Pump Condition Monitoring and Fault Detection", Transactions of the American Society of Agricultural and Biological Engineers, Vol. 49(4), 1203-1211.
- (Hsiao et al., 2009), C. Hsiao, H. Chen, 2009, "On Classification from Outlier View", <http://arxiv.org/ftp/arxiv/papers/0907/0907.5155.pdf>
- (Klir et al, 1995) G.J., Klir, B. Yaun, 1995, "Fuzzy Sets and Fuzzy Logic: Theory and Applications", Prentice Hall.
- (LaBour, 1995) LaBour Pump Company, 1995, "LaBour Taber: Pump Wear", <http://www.peerlesspump.com/PumpWear.pdf>
- (Leguizamon et al., 1996) S. Leguizamon, H. P. and S. Azzali. 1996, "Unsupervised Fuzzy C-Means Classification for the Determination of Dynamically Homogeneous Areas", p.851-856.
- (Loh et al., 1997) W.Y. Loh, Y.S. Shih, 1997, "Split selection methods for classification trees", Statistica Sinica, 7, 815-840.
- (Massart et al., 2005) D.L. Massart, J. Smeyers-Verbeke, X. Capron, K. Schlesier, 2005, "Visual Presentation of Data by Means of Box Plots", Practical Data Handling, LC•GC Europe 18(4) 215-218 (2005).
- (Marseguerra et al., 2004) M. Marseguerra, E. Zio, P. Avogadri, 2004, "Model Identification by Neuro-fuzzy Techniques: Predicting the Water Level in a Steam Generator of a PWR", Progress in Nuclear Energy, Volume 44, Issue 3, 2004, Pages 237-252.
- (Mitchell, 1999) J.S. Mitchell, 1999, "The History of Condition Monitoring and Condition-based Maintenance", Sound and Vibration, 33(11), pp. 21-28.
- (NIST) NIST/SEMATECH, 2010, "e-Handbook of Statistical Methods", <http://www.itl.nist.gov/div898/handbook/>
- (Walsh et al., 2003) M. Walsh, L.W. Lake, 2003, "A Generalized Approach to Primary Hydrocarbon Recovery", Elsevier.
- (Wang et al., 2006) J.P. Wang, H. T. Hu. 2006, "Vibration-based Fault Diagnosis of Pump Using Fuzzy Technique, Measurement, Vol. 39, 176-185.
- (Wang et al., 2007) H.Q. Wang and P. Chen. 2007, "Fault Diagnosis of Centrifugal Pump Using Symptom Parameters in Frequency Domain, Agricultural Engineering International: the

- CICR Ejournal. Vol. 9, 1-14.
- (Schapire, 1999) R.E. Schapire, 1999, "Theoretical Views of Boosting and Applications", in Tenth International Conference on Algorithmic Learning Theory.
- (Ripley, 1996) B.D. Ripley, 1996, "Pattern Recognition and Neural Networks", Cambridge University Press.
- (Rumelhart et al., 1986) D.E. Rumelhart, J.L. McClelland, 1986, "Parallel Distributed Processing", Vol. 1. MIT Press, Cambridge, MA.
- (Shahrtash et al., 2008) S. M. Shahrtash, A. Jamehbozorg, 2008, "A Decision Tree-based Method for Fault Classification in Transmission Lines", IEEE.
- (Sheng et al., 2004) Y. Sheng, S. M. Rovnyak, 2004, "Decision Tree-Based Methodology for High Impedance Fault Detection", IEEE Transactions on Power Delivery, Vol. 19, No. 2.
- (Tian et al., 2007) J. Tian, M. Gao, K. Li, H. Zhou, 2007, "Fault Detection of Oil Pump Based on Classify Support Vector Machine", 2007 IEEE International Conference on Control and Automation WeD5-4, Guangzhou, China, May 30–June 1, 2007.
- (Tse, 2002) P. W. Tse, "Maintenance Practices in Hong Kong and the Use of the Intelligent Scheduler", Journal of Quality in Maintenance Engineering, Vol. 8 Iss: 4, pp. 369–380, 2002.
- (Venkatasubramanian et al., 2003) V. Venkatasubramanian, R. Rengaswamy, K. Yin, K., S.N. Kavuri, "A Review of Process Fault Detection and Diagnosis: Part I: Quantitative Model-based Methods", Computers & Chemical Engineering, Vol. 27, 293-311, 2003.
- (Zhang, 2000) G.P. Zhang, 2000, "Neural Networks for Classification: A Survey", IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, Vol. 30 (4).
- (Zio et al., 2006) E. Zio and G. Gola, 2006, "Neuro-fuzzy Pattern Classification for Fault Diagnosis in Nuclear Components", Annals of Nuclear Energy, Vol. 33, Issue 5, Pages 415-426.
- (Zio, 2007) E. Zio, "Soft Computing Methods Applied to Condition Monitoring and Fault Diagnosis for Maintenance", Reliability: Theory and Applications, Vol. 3, 2007.