Multiscale permutation entropy and supervised classifiers for state detection in an ICE

J C Mejía1, J D Echeverry2 and H F Quintero1, 3
1Mechanical Engineering Faculty, Universidad Tecnológica de Pereira, Pereira, 660003 Colombia
2Electrical Engineering Department, Universidad Tecnológica de Pereira, Pereira, 660003 Colombia
Email: hquinte@utp.edu.co

Abstract. Entropy has been positioned as a measurement capable of detecting irregularities and uncertainties in time series. Specifically, Multiscale Permutation Entropy (MPE) has positioned itself as an alternative to classical characterization methodologies, managing to identify non-linear dynamics present in non-stationary signals. MPE has been used successfully to identify states of different systems without relying heavily on factors such as length or sampling frequency. This article presents a robust methodology that combines MPE with a feature selection technique and a supervised classifier (Analysis from Variance (VRA) and K-Nearest Neighbours (KNN) respectively). The robust proposed methodology is validated through two experiments (bearing failures and states of an internal combustion engine (ICE)). The proposed methodology overcomes the experiments with high precision and low computational consumption.

1. Introduction
Nowadays, most of the transport systems are based on internal combustion engine as their primary power source, thanks to the relatively low cost, high performance, and capability to operate on diverse renewable fuels. Reciprocating internal combustion engine works under the influence of burning pressure and inertial loads; its performance, among many factors, is affected by the demands of the piston-crank mechanism, piston slap motion (significantly dependent upon the clearance between the piston and the cylinder wall), inputs from the timing gear system, inputs from engine auxiliaries (that is alternator, steering system pump, etc.), inputs from the drive transmission system.

From the economic standpoint it is necessary for public transport to be reliable, operationally available, and economically maintained. Monitoring, detection, and diagnosis of faults of combustion engines is essential to secure the correct operation and reliability of the road transport means; it also allows to reduce breakdowns and dawn-time, increase safety and profitability.

Engine vibration measurements are non-intrusive techniques suitable for engine diagnostics, yet the mechanism of vibration in them can be correlated with the changes in engine operation behaviour and faults. Different methodologies have been developed for combustion engine fault detection and diagnosis of states in its main components [1]. Generally, these diagnoses are made from the capture and processing of vibration signals since they contain relevant information about the state of the machine [2]. However, these signals have many non-stationary and non-linear characteristics since their capture inevitably takes place with friction and impacts. To overcome this problem, a series of techniques have been developed for processing and classifying these signals. A widely used approach is based on the analysis of temporal and spectral characteristics of vibration signals [3] [4]. However, analysis in time, frequency and time-frequency domains are seriously affected by the signal length and sampling...
frequency of the capture [5]. To solve this problem another approach has been presented in recent years, which is based on entropy of different natures, such as Simple Entropy (ApEn) [6], Approximate Entropy (SampEn) [7], Multiscale Entropy (MSE) [8], Permutation Entropy (PE) [9] and Multiscale Permutation Entropy (MPE) [10]. All the above mentioned have been used successfully for the characterization of signals of different nature. For instance, in [6] the ApEn is used for diagnosis and clinical monitoring in physiology. In [11] PE is used for the classification of patients from EEG signals.

Finally, in [1] and [8] is used the MPE and MSE for the identification and diagnosis of faults in bearing systems based on vibration signals [12]. It should be noted that the MPE is an evolution of the MSE and the PE since it gives a much more complete measure of the nonlinear dynamic parameters of a system. MPE includes a combination of different scales and time delays, which identifies particularities that are not perceptible within other entropies [10]. All the previous methods of characterization allow to obtain a great quantity of information of the system that is analysing. When classifying, many of these characteristics can be redundant or irrelevant, which makes it more difficult to detect significant patterns besides increasing computing time. To solve this problem, feature selection techniques can be implemented, improving the quality and efficiency of the model. In the bearing fault diagnosis application, Variance based on Relevance Analysis (VRA) [1], Laplacian Score (LS) [13] and Relief (REL) [5] are normally used. After the selection of features, a classification process is performed with machine learning algorithms such as Multiple Vector Support Machines (SVMM) [1], [14], [15] and [16], Chains Hidden Markov (HMM) [17] and Neural Networks (ANN) [18], [19] and [20]. However, these classifiers have a high degree of complexity, computation time and initial parameters that must be optimized. Few works have attempted to exploit the potential of less complex conventional classifiers, such as K-Nearest Neighbours (KNN) [21], [22], Decision Trees (TREE) [23] or Naive Bayes (BAYES) [22].

This paper proposes a methodology for the detection of ICE states from mechanical vibrations based on characteristics of the MPE. To make it more efficient and effective, it is combined with the feature selection technique VRA and the supervised classifier KNN: The verification of the advantage of the chosen parameters is done by comparison with different ones used in the state of the art.

The article is organized as follows: in section 2.1 it is presented the mathematical formulation of the MPE; in section 2.2 it is detailed the feature selection technique VRA, while the section 2.3 is devoted to the classifier KNN. A novel methodology for the diagnosis of bearing failures based on MPE and machine learning algorithms techniques are in section 3.1. Results are presented in section 3 and the conclusions drawn from this approach are given in section 4.

![Figure 1. Methodology for vibration classification.](image)

2. Supervised classifier
The proposed methodology combines a characterization method, a feature selection technique, and a supervised classifier as show in figure 1. Each part of the methodology is described in the following
sections. The proposed methodology begins with the characterization of vibration signals, then using an automatic classifier combined with a characteristic selection technique, the state of the system is estimated.

2.1. Multiscale permutation methodology

Multiscale Permutation Entropy (MPE) is used in this paper as a signal characterization method, since it is a measure that allows to detect dynamic changes of the time series. It is based on the comparison in neighbouring values without considering the size of the values and, therefore, has a calculation simple and fast [24]. The above, allows to position the MPE as a particularly useful and robust tool in the presence of dynamic noise [1]. In order to describe the MPE measure, it is important to review the entropy Shannon, described as follows: Considering a series of time $x_1, x_2, \ldots, x_T$ in a space of representation of characteristics, where $T$ is the length of the time series. The entropy Shannon is represented by equation (1):

$$H(X) = -\sum_{i=1}^{T} p(x_i) \ln p(x_i) \tag{1}$$

Where $x_i \in \mathbb{R}$ and $p(x_i)$ is the marginal probability. The time series can be represented with a delay of time and dimension given by equation (2):

$$X_i^{m,\tau} = \left(x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\right) \tag{2}$$

Where $i = 1, 2, \ldots, T-(m-1)\tau$, $m$ is the dimension and $\tau$ the delay. To perform the computation of MPE, the signal must be truncated in $N=T-(m-1)\tau$ subvectors. Then, each subvector is calculated the Shannon Entropy mapped in a space of $m!$ different symbols $\left[\pi_i^{m,\tau}\right]_{i=1}^{m!}$ denoted as $\Pi$ by equation (3):

$$H(m,\tau) = -\sum_{i,\pi_i^{m,\tau}\in\Pi} p(\pi_i^{m,\tau}) \ln p(\pi_i^{m,\tau}) \tag{3}$$

Probability $p(\pi_i^{m,\tau})$ it is calculated by equation (4):

$$p(\pi_i^{m,\tau}) = \frac{\sum_{j=1}^{N} \mathbb{1}_{\text{type}(u)=\pi_i^{m,\tau}} X_j^{m,\tau}}{\sum_{j=1}^{N} \sum_{i=1}^{m!} \mathbb{1}_{\text{type}(u)\in\Pi} X_j^{m,\tau}}$$
Where the judgment type denotes the map from pattern space to symbol space. Also, $1A(u) = 1$ if $u \in A$ and $1A(u) = 0$ if $u \notin A$. The MPE can take values between the ranges $[0, \ln(m!)]$ and it is invariant under nonlinear monotonic transformations. The values of $m$ and $\tau$ vary from 1 to 8 and the values, which are values used for calculating MSE [8] and PE [9]. In figure 2 is plotted the behaviour of the MPE, when are varying $m$ and $\tau$ for a vibration signal.

When choosing the delay and dimension parameters, the nature of the signals must be considered. If the parameters are too small, the nonlinear dynamic of the characteristics from signals will not be analysed effectively. If the parameters are too large, useful information will be deleted in consequence, which will result in a wicked analysis.

2.2. Variance based on relevance analysis

It is based on the Principal Component Analysis (PCA) technique and it is used to describe a data set in terms of new uncorrelated variables (“components”). The components are ordered by the amount of original variance they describe, so the technique is useful for reducing the dimensionality of a data set [13]. Its implementation begins with the assumption of a feature vector $X = [x_1, x_2, ..., x_p]$, where $p$ is the length of the vector, $\mu$ is the mean of the vector and the covariance matrix of the data set is described in equation (5).

$$
\sum_{x} = E(x - \mu)(x - \mu)^T
$$

For a symmetric matrix, such as the covariance matrix, it is possible to calculate an orthogonal basis given by its eigenvalues $\lambda_j$ and eigenvectors. The new orthogonal base is created with the first eigenvector that points to the direction of the greatest variance of the data. In this way, you can select the first eigenvectors ($m < p$) and can it build the transformation matrix using equation (6):

$$
y = A^T x
$$

Where $y$ is the vector of projected characteristics in a new sub-space with a lower dimensionality. The projection is linear.

2.3. K-Nearest neighbors

The method of the K-Nearest Neighbors is a supervised and non-parametric classification method that estimates the posteriori probability that an $x$ element belongs to the class $C_j$ from a set of information provided [25]. A point in space is assigned to the $C$ class if this is the most frequent class among the K-nearest training examples, like is show in the figure 3. Generally, the Euclidean distance is used.

![Figure 3](image_url)

**Figure 3.** Operation of a KNN classifier.

This classifier can be severely affected by the presence of noise or irrelevant characteristics, which is why it is necessary to perform a combination with different characteristics selection techniques.
3. Results and discussion
This paper exposes a methodology to classify states of an internal combustion engine by vibration signals. The classification carried out through a cross validation of \( k = 10 \) and repeated by changing the number of training characteristics. After the validation, the best results were chosen through the shortest distance to the ideal point \([100\% (\text{Acc}), 0(\text{Std}), 0(\text{Ca})]\), where \text{Acc} is the accuracy, \text{Std} is the standard deviation of the accuracy of the cross validation and \text{Ca} is the number of characteristics. The above explained process applies to two different databases.

3.1. Methodology for fault detection
The proposed methodology is validated through two experiments with the methodology in the figure 4. The first experiment is based on the identification of bearing failures from vibration. The second experiment is based on the classification of the states of an internal combustion engine from a vibration signal.

![Figure 4](image)

**Figure 4.** Methodology proposed for the detection faults. Own elaboration.

Each signal is characterized with the MPE and then assigning it into train or test group. VRA technique is applied to the training group to obtain a ranking of relevance and it is applied to the test group. Then, with the reorganized training group, KNN is trained and the test group is evaluated.

3.2. First experiment
The first experiment is carried out by performance evaluation classifying bearing failure signals obtained from the Case Western Reserve [26]. In this data base, signals were collected for the normal bearings (Nor), faults in the internal train (IR1), external train (IR2) and ball (BE). Faults are also found in order of severity, 0.007 inches in diameter to 0.040 inches in diameter and at variable engine speeds of 1720 to 1797 RPM. Each experiment was repeated three times and the data was collected at 12 kHz for 5 seconds. Each signal was divided into 10 sub-signals to have more samples per class and imitating the experimental framework established in the literature [27]. A sub-signal of each of the faults can be seen in the figure 5.
The results are obtained with the KNN classifier combined with the VRA feature selection technique, as shown in Table 1. The quantity of characteristics used for the classification were 9 and 10 for 4 and 10 classes, respectively. It should be noted that the results are obtained thanks to the characterization made with the MPE, which achieves a high level of separability of the classes that allows the classifiers to adapt and solve the proposed application. Finally, a summary of the best classifications can be seen in the Table 1, including the computation time. The methodology proposed with the classifier KNN and VRA, performs a high classification but with greater number of characteristics as shown by the behaviour of the accuracy in the figure 6. However, the methodology is compensated by having a low computation time.

The best results are obtained with the KNN classifier, regardless of the feature selection technique. Specifically, the best accuracy was obtained with the KNN classifier combined with the Relief feature selection technique. The quantity of characteristics used for the classification were 9 and 10 for 4 and 10 classes respectively. It should be noted that the results are obtained thanks to the characterization made with the MPE, which achieves a high level of separability of the classes that allows the classifiers to adapt and solve the proposed application. Finally, a summary of the best classifications can be seen in the Table 2.

Table 1. Characteristics of the best classifications.

| Class | Classifier | Feature Selection | Acc % | No Features | Time (s) |
|-------|------------|-------------------|-------|-------------|---------|
| 4     | KNN        | VRA               | 99.72 | 9           | 16.37   |
| 10    | KNN        | VRA               | 99.57 | 10          | 18.22   |

Figure 5. Vibration signals for different states.
3.3. Second experiment

To verify the practical applicability of the proposed Multiscale Permutation Entropy (MPE) as a characterization methodology for detection of Internal Combustion Engine (ICE) states, an experiment was conducted on an engine test bench.

The experiment consists on the classification of the states of a single-cylinder air-cooled diesel ICE settled to operate under a set of operating points defined by three load states, three speed regimes and three different injector fuel pressure to simulate an artificial fault event on the engine. The engine is coupled to an eddy current dynamometer and is instrumented to measure the in-cylinder pressure along with the vibration signals in the engine block, as shown in figure 7.
Figure 7. Mounting for the acquisition of vibration signals.

The noise spectra obtained when the engine was in fired condition were used to follow engine events and faults. The different states are achieved by changing different ICE working conditions such as speed (1800 RPM, 2400 RPM, 3000 RPM), injector (18 MPa, 20 MPa, 22 MPa) and load (High, Medium, Low). The vibration signals are captured by accelerometers positioned on the X and Y axis of the motor. The accelerometers are conditioned by the National Instruments DAQ cards and the MatLab development software. The signals have a duration of 10 seconds and 51 kHz of sampling frequency. For the classification task, the signals are segmented into segments of 80 ms (i.e. 4096 samples per segment) and 90% of the segments were designated for training and 10% for test (a 10-fold cross validation was used). The results of the classification with MPE, KNN and VRA can be seen in figure 8.

Figure 8. Results for the classification of states of an ICE.

The proposed experimental methodology achieved a mean accuracy of 93.82% for the classification of different speed regimes. Table 3 shows the confusion matrix for this classifier.

| Labeled | Classified | 1800 RPM | 2400 RPM | 3000 RPM |
|---------|------------|----------|----------|----------|
| 1800 RPM| 254        | 9        | 5        |
| 2400 RPM| 11         | 252      | 7        |
| 3000 RPM| 6          | 7        | 259      |
This characterization methodology makes it possible to classify the different speeds of an ICE with a high accuracy rate, without the performance of the classifier being affected by changes in the other ICE operating regimes. The proposed experimental methodology also achieved a mean accuracy of 89.51% for the classification of different load regimes. Table 4 shows the confusion matrix for this classifier.

Table 4. Confusion matrix for load classifier.

| Labeled | Classified | High  | Medium | Low  |
|---------|------------|-------|--------|------|
| High    |            | 255   | 16     | 6    |
| Medium  |            | 23    | 243    | 12   |
| Low     |            | 8     | 20     | 227  |

As can be seen in Table 4, MPE allows to extract discriminant characteristics that leads to a high accuracy rate in the classification of the ICE load. However, these results are lower than those obtained in the speed regimes classification since the different load regimes do not influence the engine vibrations to the same extent.

In the last configuration of the second experiment, was achieved a mean accuracy of 91.37% for the classification of different pressures in the injector. Table 5 shows the confusion matrix for the last classifier.

Table 5. Confusion matrix for pressure classifier.

| Labeled | Classified | 18 MPa | 20 MPa | 22 MPa |
|---------|------------|--------|--------|--------|
| 18 MPa  |            | 255    | 25     | 11     |
| 20 MPa  |            | 11     | 245    | 4      |
| 22 MPa  |            | 10     | 8      | 241    |

Changes in pressure create differences in combustion, which in turn are reflected in engine vibrations. These dynamic changes have a direct impact on the ability of the proposed methodology to reveal relevant characteristics through MPE, which in turn leads to a high success rate in the classification of different pressures.

4. Conclusion
This article presents a methodology for classifying internal combustion engine states (ICE) and bearing failures. This methodology is based on measurements of Multiscale Permutation Entropy (MPE), which is capable of detecting non-linear dynamics present in non-stationary signals thanks to the incorporation of delay and scale parameters. Specifically, in this research it was used to search for characteristics in vibration signals, which would not be detected by classical techniques. MPE is combined with the VRA feature selection technique to identify the most relevant features and thus improve the effectiveness and efficiency of the proposed model. Lastly, the MPE and VRA are combined with the KNN algorithm to consolidate a robust methodology capable of classifying ICE states and bearing failures. The results allow observing an accuracy of 99.72% in the classification of 4 bearing failures with a time of 16.37 seconds, positioning itself as one of the best results in the literature. In addition, regarding the states of an ICE, it generated an accuracy of 3.82% for the identification of the speed, 89.51% for the classification of the state of charge and 91.37% for the identification of the pressure in the injector.

Acknowledgements
The authors thank to Ministerio de Ciencias y Tecnología de Colombia for supporting the project entitled: “Desarrollo de un sistema de monitoreo para el análisis energético y de condición de emisiones en
motores de combustión interna diésel con base en técnicas no destructivas” with code 1110-776-57801, through which the research described in this article was developed.

**References**

[1] Zheng P and Yang H 2018 Generalized composite multiscale permutation entropy and laplacian score based rolling bearing fault diagnosis *Mech. Syst. Signal Process* **99** 229-43

[2] Tiwari R, Gupta V, and Kankar PK 2015 Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier *J. Vib. Control* **21**(3) 461-7

[3] Martin H and Honarvar F 1995 Application of statistical moments to bearing failure detection *Appl. Acoust.* **44** 67-77

[4] Heng R B W and Nor M J M 1998 Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition *Appl. Acoust.* **53** 211-26

[5] Wei Z, Wang Y, He S and Sao J 2017 A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection *Knowl-Based Syst.* **116** 1-12

[6] Richman J S and Moorman J R 2000 Physiological time-series analysis using approximate entropy and sample entropy *American Journal of physiology Heart and circulatory physiology* **278** 39-49

[7] Pincus M 1991 Approximate entropy as a measure of system complexity *Proc. National Academy of Sciences* **88**(6) 2297-301

[8] Costa M, Goldberger A and Chung P 2001 Multiscale entropy analysis of physiologic time series. *Phys. Rev. Lett.* **89** 220-32

[9] Bandt C and Pompe B 2002 Permutation entropy: A natural complexity measure for time series *Phys. Rev. Lett.* **88** 74–102

[10] Aziz W and Arif M 2005 Multiscale permutation entropy of physiological time series 2005 *Pakistan Section Multitopic Conference* Karachi **53** 1–6 Zeng K, Ouyang G, Chen H, Gu Y, Liu X and Li X 2018 Characterizing dynamics of absence seizure EEG with spatial-temporal permutation entropy *Neurocomputing* **275** 577-85

[11] Quintero H and López J 2016 *Vibraciones mecánicas un enfoque teórico-práctico* (Universidad Tecnológica de Pereira Pereira Colombia) pp 294-8

[12] Daza S, Arias J and Llorente J 2009 Dynamic feature extraction: an application to voice pathology detection *Intell. Autom. Soft. Co.* **15** 667-82

[13] Zhang X, Liang Y, Zhou J and Zang Y 2015 A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement* **69** 164-179 Zheng J, Cheng J and Yang Y 2013 A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy *Mech. Mach. Theory* **70** 441-53

[14] Han M and Pan J 2015 A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings. *Measurement* **76** 7-19

[15] Mitchell Y, Qin Y, Zhou J, Guo Y, Celler B and Su S 2016 Automatic bearing fault diagnosis using particle swarm clustering and hidden Markov model *Engin. Appl. of Artif. Intell.* **47** 88-100

[16] Muruganatham B, Krishnakumar S and Murty S 2013 Roller element bearing fault diagnosis using singular spectrum analysis *Mech. Syst. Signal Process.* **35**(1) 150-66

[17] Jaouherl B, Fnaiech N and Saidi L 2015 Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals *Appl. Acoust.* **89** 16-27

[18] William P and Hoffman M 2011 Identification of bearing faults using time domain zero-crossings *Mech. Syst. Signal Process.* **25**(8) 3078-88

[19] Liang L, Liu F and Li M 2016 Feature selection for machine fault diagnosis using clustering of non-negation matrix factorization *Measurement* **94** 295-305

[20] Vencalek O and Pokotylo O 2018 Depth-weighted bayes classification *Comput. Stat. Data. An.* **123** 1-12

[21] Penny B and Geoff J 2017 Classification trees for poverty mapping *Comput. Stat. Data. An.* **115**
53-66
[22] Cao Y, Tung W, Gao J, Protopopescu VA and Hively L M 2004 Detecting dynamical changes in time series using the permutation entropy Phys. Rev. E 70 174-95
[23] Philipp M, Katri S and Ville N 2018 Scent classification by k nearest neighbors using ion-mobility spectrometry measurements Expert Syst. Appl. 15 199-210
[24] Bearing Data Center. https://csegroups.case.edu/bearingdatacenter/hom
[25] Nayana R and Geethanjali P 2017 Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal IEEE Sensors J. 17 5618-25
[26] Zhu K, Song X and Xue D 2014 A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm Measurement 47 669-75
[27] Liu Z, Cao H, Chen X, He Z and Shen Z 2013 Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings Neurocomputing 99 399-410
[28] Ocak H and Loparo K 2005 HMM-based fault detection and diagnosis scheme for rolling element bearings J. Vibr. Acoust. 127 2-15
[29] Shao H, Jiang H, Wang F and Zhao H 2017 An enhancement deep feature fusion method for rotating machinery fault diagnosis Knowl-Based Syst. 11 200-20