Tesi di dottorato in :
Condition monitoring and advanced fault detection techniques for engineering systems

Moise Avoci Ugwiri

University of Salerno

Department of Industrial Engineering

Ph.D Thesis

Ph.D. Course in Industrial Engineering
Curriculum in Electronic Engineering - XXXIII Cycle

Condition monitoring and advanced fault detection techniques for engineering systems

---

Supervisor:
Consolatina Liguori

Ph.D. Student:
Moise Avoci Ugwiri

Scientific Referees
Prof.Prof. Claudio De Capua
Prof. Annalisa Liccardo

Ph.D. Course Coordinator:
Francesco Donsì

Academic year 2021-2022

To my loved ones, to my family

Acknowledgments

On completion of this doctoral research, I would like to express my sincere gratitude to Professors Consolatina LIGUORI, Antonio PIETROSANTO and Aime LAY-EKUAKILLE who, despite their multiple occupations, have been constantly available for my scientific and social growth. Their valuable advice, orientations and good mood that I have benefited from during these years spent in the Measurements Laboratory of the Faculty of Engineering of the University of Salerno.

I express my profound gratitude to Prof. Francesco DONSI who has always been very affectionate and constantly available for all the scientific and administrative questions, his advice and coordination have been of great help to me during these years of research.

I will not forget all the team of measurement group of laboratory LT15, people without whose help and friendship, far from my family, to prepare this thesis would have been hard to live. I think of Prof. Vincenzo Paciello, Prof. Paolo Somella, Dr. Marco Carratu, Dr. Salvatore Dello Iacono, Dr. Giuseppe Di Leo, Peppe Ciolino, Laura De Santis, Daniele Buonocore,...

Finally, I would like to thank the University of Salerno for having made this project possible by granting me a doctoral scholarship.

Contents

List of Figures

List of Tables

Publications resulting from this Thesis

Moise Avoci Ugwiri, Marco Carratu, Vincenzo Paciello, Consolatina Liguori, Benefits of enhanced techniques combining negentropy, spectral correlation and kurtogram for bearing fault diagnosis, Measurement, 2021,doi.org10.1016/j.measurement.2021.110013.

M. A. Ugwiri, I. Mpia and A. Lay-Ekuakille, "Vibrations for fault detection in electric machines," in IEEE Instrumentation & Measurement Magazine, vol. 23, no. 1, pp. 66-72, Feb. 2020, doi:10.1109/MIM.2020.8979527.

Aimé Lay-Ekuakille, **Moise Avoci Ugwiri**, Marco Di Paola, Francesco Conversano, Sergio Casciaro, Consolatina Liguori, and Vikrant Bhateja, "Functionalized Nanoparticles for Characterization Through TEM Images: Comparison Between Two Innovative Techniques," in IEEE Instrumentation & Measurement Magazine, vol. 24, no. 4, pp. 29-34, June 2021, doi:10.1109/MIM.2021.9448262.

Lay-ekuakille, A., Djungha Okitadiowo, J., **Avoci Ugwiri, M**., Maggi, S., Masciale, R., Passarella, G. Video-sensing characterization for hydrodynamic features: Particle tracking-based algorithm supported by a machine learning approach (2021) 21 (12), art. no. 4197,. DOI:10.3390/s21124197

Lay-Ekuakille, A., **Ugwiri, M.A.**, Liguori, C., Singh, S.P., Rahman, M.Z.U., Veneziano, D. Medical image measurement and characterization: Extracting mechanical and thermal stresses for surgery (2021) 28 (1), pp. 3-21. DOI:10.24425/mms.2021.135998

Lay-Ekuakille, A., **Ugwiri, M.A.**, Okitadiowo, J.P.D., Telesca, V., Picuno, P.,

Liguori, C., Singh, S. SAR sensors measurements for environmental classification: Machine learning-based performances (2020) 23 (6), art. no.9200877, pp. 23-30. DOI:10.1109/mim.2020.9200877

Lay-Ekuakille, A., **Ugwiri, M.A.**, Okitadiowo, J.D., Chiffi, C., Pietrosanto, A. Computer Vision for Sensed Images Approach in Extremely Harsh Environments: Blast Furnace Chute Wear Characterization (2021) 21 (10), art. no. 9370129, pp. 11969-11976. DOI:10.1109/JSEN.2021.3063264

Lay-Ekuakille, A., **Ugwiri, M.A.**, Telesca, V., Velazquez, R., Passarella, G., Maggi, S. Detection of river flow slow-down through sensing system and quasi-real time imaging (2021) 81, art. no. 102042, . DOI:10.1016/j.flowmeasinst.2021.102042

Ugwiri, Moise Avoci; Carratù, Marco ; Pietrosanto Antonio ; Liguori, Consolatina, Enhanced sensor coverage and fault detection framework for autonomous vehicles, Measurement: Sensors, Volume 18, 2021, 100214, doi.org/10.1016/j.measen.2021.100214.

M. A. Ugwiri, M. Carratù, C. Liguori and V. Paciello, "Bridge dynamic analysis for accurate control and monitoring: Feasibility study,"IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, 2019, pp. 5307-5311, doi:10.1109/IECON.2019.8927692.

A. Lay-Ekuakille, **M. A. Ugwiri**, C. Liguori and P. K. Mvemba, "Enhanced Methods for Extracting Characteristic Features from ECG," 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2019, pp. 1-5, doi:10.1109/MeMeA.2019.8802127.

Hoang, M.L., Carratu, M., **Ugwiri, M.A**., Paciello, V., Pietrosanto, A. A New Technique for Optimization of Linear Displacement Measurement based on MEMS Accelerometer (2020) 2020-October, art. no. 9268038, pp. 155-158. DOI:10.1109/CAS50358.2020.9268038

Ugwiri, M.A., Carratu, M., Pietrosanto, A., Paciello, V., Lay-Ekuakille, A. Vibrations measurement and current signatures for fault detection in asynchronous motor (2020) art. no. 9128433,. DOI:10.1109/I2MTC43012.2020.9128433

Surekha, S., Ur Rahman, M.Z., Lay-Ekuakille, A., Pietrosanto, **Ugwiri, M.A.** Energy detection for spectrum sensing in medical telemetry networks using modified NLMS algorithm (2020) art. no. 9129107, . DOI:10.1109/I2MTC43012.2020.9129107

Lay-Ekuakille, A., **Ugwiri, M.A.**, Liguori, C. Modeling Stresses in Plasmonic Nanosensors (2019) art. no. 9233491, .DOI:10.1109/NanofIM49467.2019.9233491

Ugwiri, M.A., Carratu, M., Monte, G., Espirito-Santo, A., Paciello, V. Edge sensor signal processing algorithms for earthquake early detection (2020) art. no. 9129617, . DOI:10.1109/I2MTC43012.2020.9129617

M. A. Ugwiri, M. Carratù, A. Lay-Ekuakille, V. Paciello and A. Pietrosanto, "Cascade based methods in detecting rotating faults using vibration measurements," 2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2021, pp. 1-5, doi:10.1109/I2MTC50364.2021.9460107.

Ugwiri, M.A., Carratu, M., Paciello, V., Liguori, C. Spectral negentropy and kurtogram performance comparison for bearing fault diagnosis (2020) art. no. IMEKO-TC10-2020-011, pp. 105-110.

Capriglione, D., Carratù, M., Dello Iacono, S., Pietrosanto, A., Sommella, P., **Avoci Ugwiri, M.** Design and Implementation of a Diagnostic Scheme for Stroke Sensors in Motorcycle Semi-active Suspension Systems (2020) 629, pp. 335-341. DOI:10.1007/978-3-030-37558-4_51

M. Avoci Ugwiri, M. Carratù, C. Liguori, V. Paciello, A. Pietrosanto. Calibration technique optimization for articulated arm measuring machines, Congresso Nazionale GMEE 2021, ISBN 978-88-31901-06-2 (pp. 201-202)

M. Avoci Ugwiri, M. Carratù, C. Liguori, V. Paciello,Antonio Espírito-Santo, Gustavo Monte, A. Pietrosanto(Enhanced approach for Early Earthquake Detection using Smart Sensor, Congresso Nazionale GMEE 2020,. ISBN 978-88-31901-06-2 (pp. 349-350)

Abstract

The technological developments are unavoidable paths to ensure the reliability and efficiency of industrial assets, which are requirements for living standard and stable economy both in constant demand. In this context, detecting eventual defect is crucial for ensuring the maximal performance of all the machines within the asset and avoiding their failure and even a complete breakdown. In most of industrial installation nowadays, the strategy is changing from preventive to condition monitoring. The idea consists of scheduling interventions on equipment based on the technical condition of machines, while they are operating. The strategy has tremendous technical and economic advantages. On the other hand, condition monitoring has to be built such that, it gives both the current condition of the machine and an indication on its remaining useful life. The wide community of researchers in condition monitoring concluded that, for signals such as vibrations, they can be modeled as cyclostationary and non-stationary signals, and several mathematical approaches have been developed. Signal processing for a wide class of industrial application is still a subject of serious disagreement both in scientific and industrial community. Indeed, feature extraction and classification are a delicate task, because an error at this step can lead to huge consequences in detection of a potential faulty component and obviously the subsequent failure machine. Another drawback of the signal processing methods currently used in condition monitoring is that they are designed for stationary signals, whereas in real industrial applications, in most of the scenarios, machines operate at varying speed and load, leading in lack of feature extraction performance. The main contributions of this research can inevitably be in methodological development, particularly in the data manipulation in order to improve the results at the detection and diagnosis level, while opening a prognostic window. Applications such as spectral analysis where evaluated, where the development in this thesis suggested an enhancement of the traditional techniques applied to detect electric machines faults. From this enhancement, a particular analysis is done on cascade

methods. An advanced spectral analysis is suggested as well where techniques such as spectral negentropy or spectral correlation are used to extract fault information on bearings. In addition, thesis proposed an experimental study using data mining approaches and the use of digital twins for critical industrial processes monitoring.

Chapter 1

Introduction

1.1 Contextual insight

Nowadays, sensors technology is getting more sophisticated. Sensors can be mounted directly on equipment and built upon Internet of Things (IoT) standard, which makes the condition monitoring future's more interesting. Condition monitoring techniques are standardized through the International Organization for Standardization (ISO) and American Society for Testing and Material (ASTM) [1]. ASTM outlines a variety of standards, mostly dealing with condition monitoring for in-service lubricants, while ISO standards 13372, 18436, 17259 and 13381 [2] specify the guidelines for condition monitoring and diagnostics of machines. Today, industrial devices have unprecedented number of sensing, processing and communication capabilities built to the product itself. In condition monitoring, the data processing stage is a useful step for predictive maintenance. Patterns emerge from the data showing a machine part which may be in a degradation stage. Based on the analysis, maintenance can be schedule to prevent sudden failure and avoid system downtime.

1.2 Motivation

The technological developments are unavoidable paths to ensure the reliability and efficiency of industrial assets, which are requirements for living standard and stable economy both in constant demand. In this context, detecting eventual defect is crucial for ensuring the maximal performance of all the machines within the asset and avoiding their failure and even a complete breakdown. In most of industrial installation nowadays, the strategy is changing from preventive to condition monitoring [3].

The idea consists of scheduling interventions on equipment based on the technical condition of machines, while they are operating. The strategy has tremendous technical and economical advantages. On the other hand, condition monitoring has to be built such as it gives both the current condition of the machine and an indication on its remaining useful life. For managing an industrial installation based on condition based monitoring, three main steps are necessary :

1. Data have to be collected using transducers to measure different physical characteristics of the machines such as : vibrations, temperature, pressure, acoustic emissions, etc...

2. The acquired data must be mapped into a feature space using digital signal processing algorithms

3. Features have to be classified based on the previous history of the machine and the condition is determined.

From the research perspective, two aspects of condition monitoring can be highlighted as the most interesting :

1. Mathematical modeling of the signals to define relevant features that can provide diagnostic information.

2. Signal processing (eventually image processing) methods for extracting features and classify them.

The wide community of researchers in condition monitoring concluded that, for signals such as vibrations, they can be modeled as cyclostaionary and non-stationary signals, and several mathematical approach have been developed [4]. Signal processing for a wide class of industrial application is still a subject of serious disagreement both in scientific and industrial community. Indeed, feature extraction and classification is a delicate task, because an error at this step can lead to huge consequences in detection of a potential faulty component and obviously the subsequent failure machine. Another drawback of the signal processing methods currently used in condition monitoring is that, they are designed for stationary signals, whereas in real industrial applications, in most of the scenarios, machines operate at varying speed and load, leading in lack of feature extraction performance [5].

1.3 Objectives

Objectives pursued in this research can be summarized as follow :

1. Providing a satisfactory elements of evidence on the current state of the development in fault detection and diagnosis
2. The development of a signal pre-processing approach for noise reduction and address the question of the definition and selection of relevant indicators, including in the case where the dimension cannot be neglected.
3. Contribute to the development of approaches for the early faults detection on the system, sensors and actuators.
4. Evaluate different learning-based diagnostic systems. The considered methods are Machine Learning approaches as well as traditional digital signal processing approaches.
5. Finally, fault detection methods based on Digital Twins approaches are evaluated and discussed with suggestions for future implementation in multiphysics systems.

The main contributions of this research can inevitably be in methodological development, particularly in the data manipulation in order to improve the results at the detection and diagnosis level, while opening a prognostic window. Applications such as spectral analysis where evaluated, where the development in this thesis suggested an enhancement of the traditional techniques applied to detect electric machines faults. From this enhancement, a particular analysis is done on cascade methods. An advanced spectral analysis is suggested as well where techniques such as spectral negentropy or spectral correlation are used to extract fault information on bearings. In addition, thesis proposed an experimental study using data mining approaches and the use of digital twins for critical industrial processes monitoring.

1.4 Document organization

In the chapter 2, fundamentals are discussed and general taxonomies of approaches used in fault detection and diagnosis are presented. The chapter elaborate in detail the consideration need to be taken, when deciding which approach is suitable

for the case in question. The chapter 3 discussed the non-learning techniques for fault signature extraction. The effectiveness of extracting useful information on faults and the singularity characterization are discussed as well. The chapter 4 provides major advanced frequency representation and analysis. The interpretability of the most informative frequency band is demonstrated. The chapter also deals with the visual representation of spectrum and proposed approach to dispay vectors of cyclical frequency as sample rate. The experimental validation uses bearings housed in casing allowing shaft to rotate while driven by a variable speed electric motor. Ball Pass Frequency Inner Race (BPFI) was particularly investigated. The chapter 5 focuses on data mining based on machine learning techniques. The application presented in the chapter 3 is used to demonstrate the challenge beside selecting relevant feature by combining the process with the cascade method is deeply discussed. The chapter 6 proposed to take advantage of the digital twin concept in the evaluation of the condition of a given system or subsystem with the goal to predict its future behavior, refine the control or optimize the operation by detecting and identifying faults as early as possible without having to access physically the targeted system or subsystem.

References

[1] ASTM. *2016 Annual Book of ASTM Standard*. 2016.

[2] ISO:13372:2012. *Condition monitoring and diagnostics of machine systems*. ISO, 2012.

[3] Lindley R. Higgins R. Keith Mobley and Darrin J. Wikoff. *Maintenance Engineering Handbook*. The McGraw-Hill Companies, 2008.

[4] Abboud D.; Baudin S. ; Antoni J. ; Rmond D. ; Eltabach M. ; and Sauvage O. The spectral analysis of cyclo-non-stationary signals. *Mechanical Systems and Signal Processing*, 2016.

[5] I. El-Thalji and E. Jantunen. A summary offault modelling and predictive health monitoring of rolling element bearings. *Mechanical Systems and Signal Processing*, 2015.

Chapter 2

Fundamentals of fault detection and diagnosis methods

This chapter depicted fundamentals of commonly used approaches in fault detection and diagnosis. Developed theories are preamble to the contribution provided in this report.

2.1 Introduction

The safety of most technical plants requires a precocious faults detection in the process. A remarkable number of techniques and methods are being developed on daily basis, enabling earlier detection of faults than conventional limit and trend checking based on a single process variable. Most of these methods use information from several variables of the process including non-measurable variables as process state, parameters and characteristics quantities [1]. Some of those methods require accurate process models, while others rely on available historical process data.

2.2 Some useful terminologies

2.2.1 Fault

Based on the definition provided in [2], "a fault is an unpermitted deviation of at least one characteristic property or feature of the system from the acceptable, usual standard condition." The definition proposed by **ISO 26262-1:2018** gives a more extensive view, suggesting that a fault can be seen as an abnormal condition that can cause a system or part of a system including components, hardware, software,

hardware part and software units or an item [3] to fail. The standard highlighted three separate faults type: permanent, intermittent and transient faults.

2.2.2 Failure

A "failure" can be defined as a permanent interruption of a system's to perform a require function under specified operating conditions [4]. A failure is always related to a required function. The function is specified along with the specified requirements. Thus, a failure occurs when a function cannot be performed or has a performance that falls outside of the requirements [4].

2.2.3 Malfunction

A malfunction is an intermittent irregularity of the system to accomplished the required function.

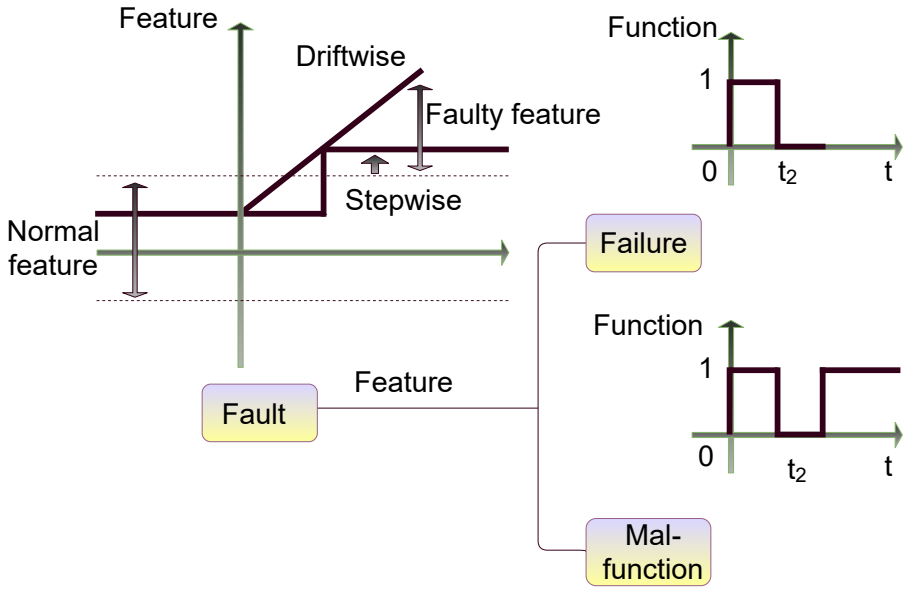


Figure 2.1: The developement of the malfunction

2.3 Procedures

On industrial level, four main procedures depicted in the Fig 2.2, are associated to process monitoring. Fault detection plays the role in establishing the evidence on whether fault has occurred. Fault identification pin points the most relevant variables for the proper diagnosis. Fault diagnosis determines which fault occurred, i.e. it determines the causes of the observed out-of-control status. Intervention, also called process recovery, removes the effect of the fault and it is the procedure needed to close the process monitoring loop.

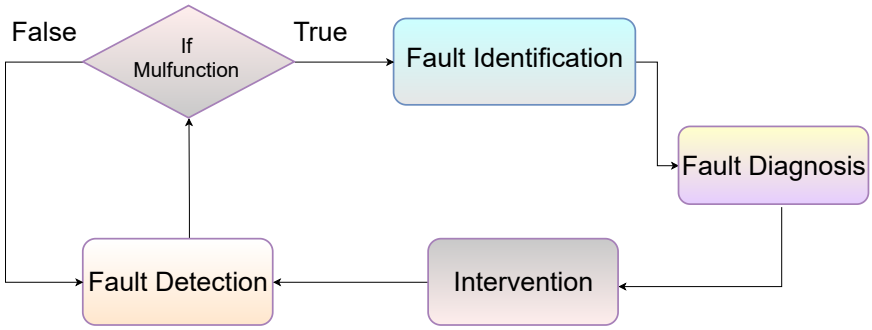


Figure 2.2: General scheme of the process monitoring at industrial level

While all four procedures may be implemented in a given process monitoring, this is not always suitable or necessary. For instance, a fault may be diagnosed without identifying the variable immediately effected by the fault [5]. Moreover, it is not necessary to automate all aforementioned procedures. An automated fault identification procedure may be used to assist the plant operators and engineers to diagnose the fault and recover normal operation. Once the fault has been properly diagnosed, an optimal approach to apply to counteract the fault may not be obvious. In the case of sensor fault for example, a sensor reconstruction technique can be applied to the process.

2.4 Measurement for the process monitoring

A monitoring scheme contains typically one or several measures based on the elaboration from statistical theory, pattern recognition, information theory and system theory. These measures are the complete representation of the behavior of the

process. In Fig. 2.3, a graphical representation of main aspect to be taken for a measurement architecture for a given process.

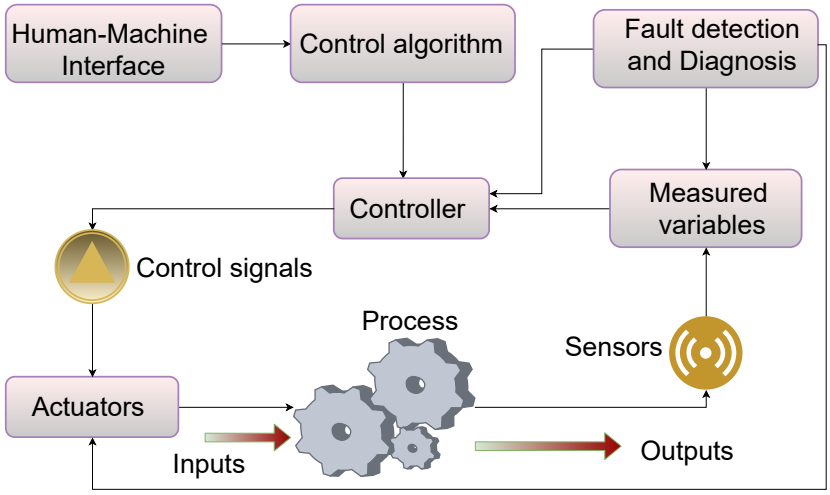


Figure 2.3: Schematic overview of measurement for industrial fault detection and diagnosis

In fault detection practices, limits may be placed on some of the measures and a fault can be detected whenever one of the evaluated measures is outside the limits [6]. In this perspective, the measures are able to define the in-control process behavior and accordingly the out-of-control status. When measurements are accurate, the behavior of each observation variable can be compared against the measured values to determine the variable most affected by the fault. The fault can also be diagnosed by developing and comparing accurate measures data representing different possible faults of the process. It is worth to develop a measurement capability that maximizes the sensitivity and robustness of the process at all possible faults. In practice, faults may occur in several ways, this makes hard to effectively detect and diagnose them with few measurements. Since each measure characterizes a fault in a different manner, one measure will be more sensitive to certain fault and less sensitive to other fault relative to other measures. For this reason, the use of multiple measures with the efficiency of each measure determined for the particular process and the possible faults. Measurements for process monitoring can be classified as being associated with one or more of possible techniques presented in the section 2.5, 2.6, 2.7 and 2.8. Getting hold of flawless measures and constructing the fault models for large-scale system can require a large amount of effort [7]. However, last decades

proved that, software packages are being developed to enable knowledge-based approach to be easily utilized to complex systems.

2.5 Data Based Methods and Signal Models

2.5.1 Data-based Methods

Data-based methods used in condition monitoring rely on available historical data [8] collected from the monitored fault process. The obtained data are usually used first to build an empirical model to be used for the fault detection in the future data. The condition monitoring based on data includes methods such as Partial Least Square (PLS) regression, Principal Component Analysis (PCA), independent component analysis (ICA), Canonical Variate Analysis (CVA) [9], neural networks [10], fuzzy systems [11], pattern recognition [12, 13] and most of the classification/regression algorithms such as Support Vector Machine (SVM), k-Nearest Neighbors (k-NN) etc... [14–16]. Nonlinear systems are often suitable for method such as SVM-based approach because because it offers advantages over conventional nonlinear optimization-based techniques.

2.5.1.1 Limit and trend checking

Limit checking is one of the commonly used methods in fault detection. The idea beside is to monitor and measure a variable $y(t)$ and set an alarm when it exceeds a static threshold [17]. Usually, the upper limit y_{max} and the lower limit y_{min} are known to be the thresholds [18]. If $y(t)$ is less than the y_{min} or greater than y_{max}, the message is clear, the system has probably fault somewhere.

$$y_{min} \leq y(t) \leq y_{max} \tag{2.1}$$

This approach is widely used in process automation systems. Based on the case that is being analyzed, one can be interested in upper, limit, lower limit or both. One of the major drawbacks of limit checking is that it is highly sensitive to noise. In addition, the system might undergo faults and the alarms will be activated only when the thresholds are exceeds and faults already occurred. One can decide to lower the alarm threshold to catch earlier the fault, but this can lead to more false alarms. Setting the alarms thresholds too high might make them less sensitive to noise but

instead make some failures go undetected. This trade-off has to be consider when thinking of implementing such approach for condition monitoring.

Trend checking is another method used in data-based method. This approach measures the first derivative of $y(t)$. As in case of limit checking, the inequality should be satisfied for non-alarming state.

$$\dot{y}_{min} \leq \dot{y(t)} \leq y_{max} \tag{2.2}$$

Trend checking is able to detect faults earlier when the thresholds are small, it can be combined to limit checking for better results. Another notable advantage of both limit and trend checking is their simplicity and reliability. However, they are able to react after relatively large change of feature [18], even if the distribution on non faulty condition data is not always Gaussian, because Gaussian Mixture Models can be used as well in such condition.

2.5.1.2 Binary thresholding [1]

If the method in question aimed to detect changes in the system is binary thresholding, often the monitored variables are stochastic variables ($y_i(t)$) with a probability density function $p(y_i)$, mean value and variance

$$\mu_i = E\{y_i(t)\} \; and \; \overline{\sigma}_i^2 = E\{[y_i(t) - \mu_i]^2\} \tag{2.3}$$

for the nominal values in case of flawless process, changes are expressed as :

$$\mu_i = E\{y_i(t) - \mu_i\} \; and \; \Delta\sigma^2 = E\{[\sigma_i(t) - \overline{\sigma}_i]^2\} \tag{2.4}$$

for $t > t_F$, where t_F is the time of the fault occurrence, which in this case is unknown. In case the mean value and the standard deviations before the change caused by a fault are described by μ_0 and σ_0, and after change appearance, one have μ_1 and σ_1. In the Fig. 2.4, the detection of the change as started above is presented.

Change detection of the random variable $y(k)$ can be done on-line and in real time or off-line. In the case of the off-line change detection, within a sample length N, it has to be determined when some unknown time t_F a change in $y(k)$ occurred from y_0 to y_1. This his only feasible with an historical data. The case of fault detection in

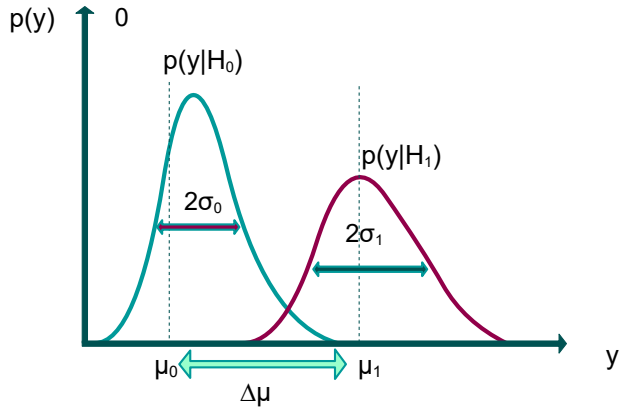


Figure 2.4: The normal probability distribution of the observed variable for nominal and changed states.

real time, the on-line change detection is of more interest. Here for each k time, it has to be decided if a change from y_0 to y_1 has happened.

2.5.1.3 Plausibility checks

For given measured variables, a rough supervision is sometimes performed by checking the plausibility of its indicated values. It means that, measurements are evaluated with regard to credible, convincing values and their compatibility among each other [19]. In practice, a single measurement is analyzed to find a whether the sign is correct and the investigated value belongs to certain limits. This can be seen also as limit check described in the sub-section 2.5.1.1, however with much larger tolerances. If several measurements are available for the same process, then the measurements can be related to each other regarding their normal ranges by using logic rules. The plausibility checks presuppose the ranges of measured process variables under certain operating conditions and represent rough process models. In case the ranges of observed variables are increasingly made smaller [2], many rules would possibly require for the process behavior description. Hence, it is better to use mathematical process models in the form of equations to detect faults.

2.5.1.4 Principal Component Analysis

The Principal component Analysis (PCA) is a among techniques that are used to emphasize variation and bring out strong patterns in a dataset. It is often used to

make data easy to explore and visualize. [21].

2.5.1.4.1 Main Idea

Known as one of the widely used multivariate statistical method, PCA basically replaces the old high-dimensional variables with fewer new variables, and makes the new variables retain most of the information of the original old variables [20–22]. If we consider $x(t) \in \mathbb{R}^m$ denoting the sample of measurement of a vector m sensors at time t. Assuming that N are sample for each sensor, a data matrix :

$$X = [x(1)...x(N)]^T \tag{2.5}$$

The Eq. 2.5 is composed by each row representing a sample $x(t)^T$. The PCA can be performed through the eigenvalue decomposition of the covariance matrix of X. The first step is to standardize the matrix X to zero mean and unit variance, then perform the eigenvalue decomposition of covariance matrix :

$$\Sigma = P \wedge P^T \tag{2.6}$$

In the Eq 2.6, Σ is the covariance matrix of X, $\wedge = diag(\lambda_1, \lambda_2, ..., \lambda_m)$ is the diagonal eigenvalue matrix $(\lambda_1 \geq \lambda_2 \geq ...., \lambda_m)$ are eigenvectors), $P = (p_1, p_2, ..., p_m)$ is the eigenvectors matrix, where $(p_i, i = 1, 2, ..., m)$ are the normalized and mutually orthogonal eigenvectors associated with the eigenvalues. One can represent eigenvalues, eigenvectors and principal components matrices as follow :

$$\wedge = \begin{bmatrix} \wedge_k & 0 \\ 0 & \wedge_{m-k} \end{bmatrix} \tag{2.7}$$

$$P = \begin{bmatrix} P_l & P_{m-k} \end{bmatrix} \qquad T = \begin{bmatrix} T_k & T_{m-k} \end{bmatrix} \tag{2.8}$$

In Eq 2.7, k is the number of the obtained principal components to be kept in the PCA model. The matrix X is then transformed into T as independent variables through :

$$T = X \times P \tag{2.9}$$

where $T = [t(1), ..., t(N)]$ contains the principles components, which are orthogonal to each other. In summary, the PCA model is determined based on an eigen decomposition of the covariance matrix Σ and the selection of the number k of the components to be retained.

2.5.1.4.2 Using PCA for fault detection and diagnosis

If one closely inspect the Eq. 2.7 and take into account the first k highest eigenvalues and their corresponding eigenvectors, the matrix X can possibly be decomposed as

$$X = T_k P_k^T + E \tag{2.10}$$

where $T_k = XP_k$, and E is the residual matrix. A sample vector $\mathbf{x}(t) \in \mathbb{R}^m$ can be projected on the principal and residual sub-spaces as

$$\hat{\mathbf{x}}(t) = P_k \mathbf{t}_k(t) \tag{2.11}$$

$$= C_k \mathbf{x}(t) \tag{2.12}$$

where $\hat{\mathbf{x}}(t)$ is the estimation vector of $\mathbf{x}(t)$, $C_k = P_k P_k^T$, the first vector k of the latent variable can be expressed as :

$$\mathbf{t}_t(t) = P_k^T \mathbf{x}(t) \in \mathbb{R}^k \tag{2.13}$$

The vector of $m - k$ last scores of latent variables, which is the projection of measurement data in the residual subspace, is given by

$$\mathbf{t}_{t-k}(t) = P_{m-k}^T \mathbf{x}(t) \in \mathbb{R}^{m-k} \tag{2.14}$$

The residual vector can be represent in measurement space $\mathbb{R}^m$ and express by

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t) \tag{2.15}$$
$$= (I - C_k)\mathbf{x}(t) \tag{2.16}$$
$$= P_{m-k}^T \mathbf{x}(t) \in \mathbb{R}^{m-k} \tag{2.17}$$

In most of fault detection situations, the residual vector e is often small in amplitude when the system under test is flawless. However, it can drastically increases in the presence of fault. While using PCA, the first step is often to build the model using flawless data representing the normal operation of the process, and then PCA model is used to detect fault using one of the detection indices used in statistics. A clear easy to grasp workflow of fault detection using PCA is depicted in the Fig 2.5

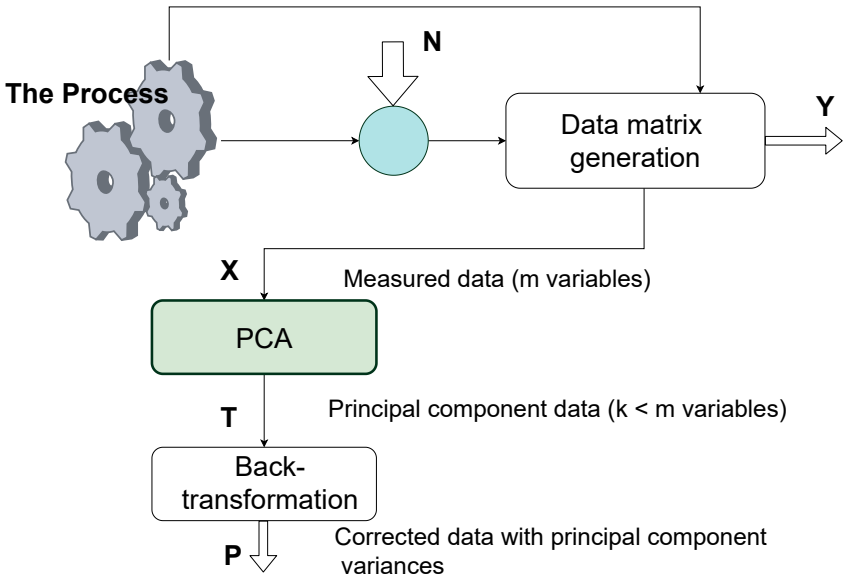


Figure 2.5: Illustrative scheme on the use of PCA for fault detection.

It is worth to mention that, methods based on PCA are suitable for a situation where available process measurements are highly correlated and only small number of faults produce unusual patterns [23]. When the process is highly correlated, the original process can be projected onto smaller number of principal components, consequently reduce the dimension of the variables.

2.5.2 Fault Detection with Signal Models

When changes in signal are related to faults, in a process, a signal analysis can be applied [24]. By assuming mathematical models for measured signal, suitable features are calculated. The signal model is usually better to apply if faults in actuators, process and sensors produce changes. One of the common approach is the use of vibrations, in measuring the position, speed or acceleration to detect faults such as imbalance, bearing faults, knocking in gasoline engines, chattering in metal grinding machines. At the same time, signals from many other sensors like : electrical current, position, speed, force, flow, pressure, etc... frequently contain oscillations with variety of higher frequency than the process dynamics. In the figure 2.6, a summary of fault detection using signal model analysis is depicted.

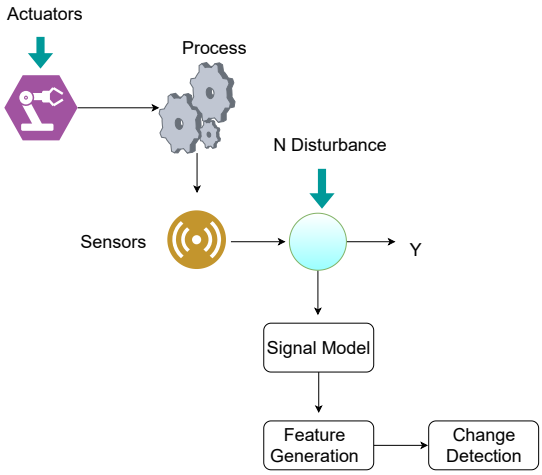


Figure 2.6: Faults detection scheme based on signal models

As aforementioned, if one assume suitable mathematical models for the measured signal, suitable features can be calculated. Those features are often : amplitudes, phases, spectrum frequencies and correlation functions for a certain frequency band of the signal. Often a comparison with the observed features for normal behavior provides changes of the features that are considered as analytic symptoms is analyzed.

2.5.2.1 Spectral Analysis

The extraction of fault-relevant signal characteristics can be restricted to the amplitudes or amplitude densities within a certain bandwidth of the signal. Discrete or continuous time-series can be analyzed in terms of time-domain descriptions and frequency-domain descriptions. The spectral analysis usually reveals some characteristics of a time-series, which cannot be easily seen from a time-domain analysis. The spectral analysis is used for solving a wide variety of practical problems in both engineering and science. While applying spectral analysis, the time series is decomposed into sine wave components using a sum of weighted sinusoidal functions called spectral components. The weighting function in the decomposition is a density of spectral components or spectral density function (SDF). The efficient algorithm used in decomposing a time-series into a sum of weighted sinusoidal functions is the Fourier transform (FT) can be used to calculate frequency content of signal $x(t)$.

2.5.2.1.1 Fourier Transform

During normal operation, components A_i fall within particular range, according to the mathematical theory of Fourier analysis, the time-series can be represented by the following finite FT :

$$x(t) = A_0 + \sum_{i=1}^{N} A_0 sin(\omega_t + \phi_i) \tag{2.18}$$

Most of time-series data recorded in engineering process, are of discrete type and numerical calculation of the FT are often done using digital computers, which can only deal with discrete data and therefore use discrete FT.

2.5.2.1.2 Frequency Spectrum

In general, the FT is a complex function and the plot of $x(t)$ versus ω_t is called frequency spectrum. The modulus $|x(t)|$ and the angle of $x(t)$ is called magnitude or the amplitude spectrum.

2.5.2.1.3 Power Spectrum

The reason to be interested in power spectrum is to answer the question : how much of the signal is at a frequency ω ? If one don't have to worry about the limitation in data, the power spectrum of the signal would be given simply by the FT

$$P_x = \int_{-\infty}^{+\infty} x(t)e^{i\omega_t} \tag{2.19}$$

P_x is the power spectrum, a real-valued function with a zero phase. The power spectrum is an average measure of the frequency-domain properties of the time-series, which shows whether or not a strongly periodic or quasi-periodic fluctuation exists in the time-series.

2.5.2.1.4 Cross Spectrum

The cross-correlation function for two sets of time-series data

$$x_n = \{x_0, x_1, x_2, ..., x_{N-1}\} \text{ and } y_n = \{y_0, y_1, y_3, ..., y_{N-1}\} \tag{2.20}$$

can be defined as

$$C_{xy}(i) = \sum_{n=0}^{N-1-i} (x_n - \hat{x})(y_{n+i} - \hat{y}) \tag{2.21}$$

where its FT is expressed as

$$P_{xy}(t) = \sum_{i=0}^{N-1-i} C_{xx}(i)e^{-j2\omega i/N} \tag{2.22}$$

Where $P_{xy}(t)$ is the cross spectral density function or the cross spectrum which generally is complex-valued function. The cross spectrum represents the common frequencies appearing in both the time-series x_n and y_n.

2.5.2.1.5 Coherence Function

The coherence function is a real-valued frequency function defined as

$$K_{xy}(t) = \frac{|P_{xy}(t)|}{\sqrt{P_{xx}(t)P_{yy}(t)}} \tag{2.23}$$

which is at a particular frequency, is a measure of similarity of strength of components in x_n and x_n at that frequency. The value of K is such that $0 \leq K \leq 1$, and the larger K, the more strongly correlated are the x_n and y_n at a given frequency. Therefore, K behaves like a correlation coefficient for x_n and y_n components at the same frequency. All the above definitions, though given for discrete time-series, are equally applicable to continuous time signals. More details of spectral analysis can be found in [25, 26]

2.5.2.2 Parametric Signal Models

A signal can be represented by a mathematical model which has a predefined structure involving a limited number of parameters. For a given signal, one can decide its representation by choosing a specific set of parameters that results in the model output being as close as possible in some prescribed sense to the given signal. A tremendous publication can be found in [27–29], where a huge elaboration on this topic is carried out. Parametric signal models such as ARMA (auto-regressive moving average) are some of the most used. $ARMA(p, q)$ refers to the model with p autoregressive terms ϕ_i and q moving average terms θ_i, constant c and error terms ϵ_t,ϵ_{t-i}

$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{2.24}$$

Parametric models are very sensitive to small frequency changes. Parametric modeling techniques often find the parameters for a mathematical model describing a signal, system, or process; moreover, the approach is known as information-based. The information help to determine the model. Applications for parametric modeling include condition monitoring by means of vibration, speech and music synthesis, data compression, high-resolution spectral estimation, communications, manufacturing, and simulation.

2.5.2.3 Pattern Recognition : Neural Networks

Neural networks have been successfully used for pattern recognition and as such are suitable for fault detection in most smart manufacturing process [30].

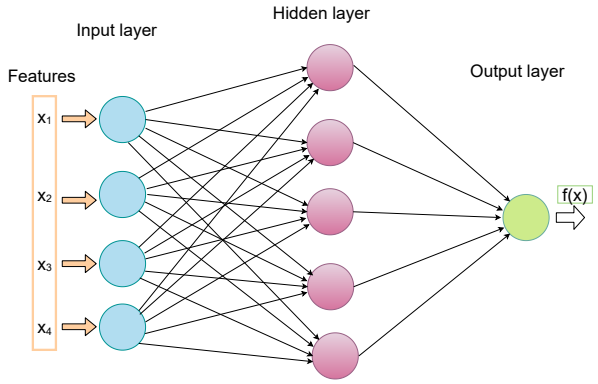


Figure 2.7: Feedforward neural network

In supervised training input-output pairs, both for normal and faulty conditions, are presented to the network. If not enough faults are available in training data, additional training samples can be produced by artificial fault injection. For supervised training a feed forward network is the most common architecture, Fig 2.7, usually trained with some variant of backpropagation algorithm. If unsupervised learning is required, Kohonen selforganizing network is the best choice Fig 2.8.

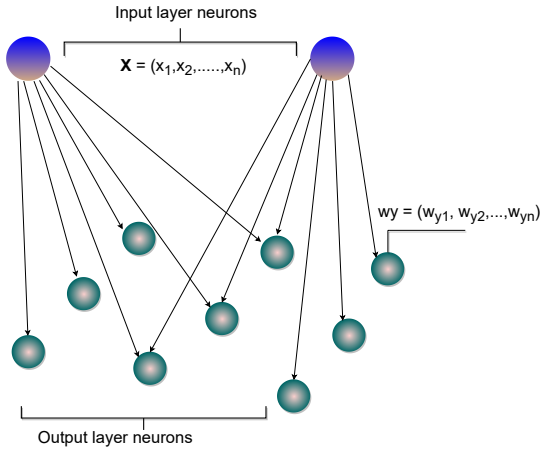


Figure 2.8: Self-organizing map based on the architecture developed in [24]

The neurons of the competitive net learn to recognize groups of similar input vectors, in such a way that each neuron competes to respond to an input vector x_t, the neuron whose value m_c is closest to x_t get the highest net input and therefore wins

the competition and outputs one, all other neurons output zero. Non fault and fault conditions are represented as different subsets of neurons within a map [31].

2.6 Process Model Based Methods

The process model based methods presented in this thesis, are based on concept of analytical redundancy [32]. The essence of this concept is the comparison of the actual outputs of the monitored system with the outputs obtained from a (redundant. i.e. not physical) analytical mathematical model, Fig 2.9 [32].

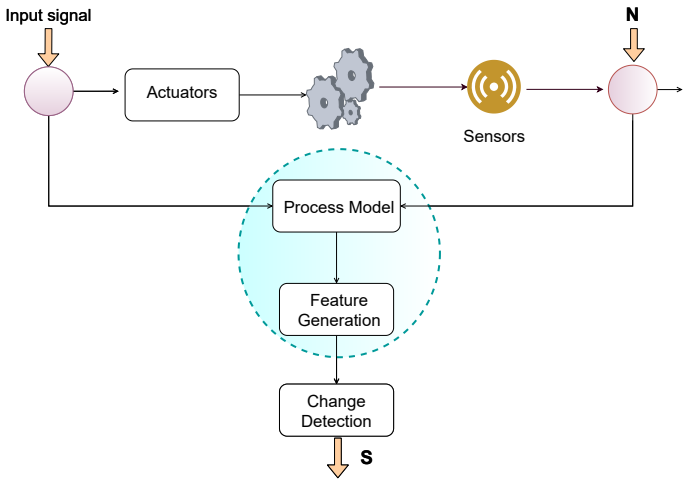


Figure 2.9: Fault detection process based model

It involves two stages: residual generation and residual evaluation. This approach assumes that the structure and the parameters of the model are precisely known. Faults can be modeled as state variable changes. Limiting consideration to linear systems, the actual system may be given in continuous time by state Eq. 2.25 and 2.26.

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.25}$$

$$y(t) = Cx(t) \tag{2.26}$$

where A, B and C are known matrices. Output $y(t)$ is a function.

$$y(t) = f(u(t), \omega(t), x(t), \theta(t)) \tag{2.27}$$

where $u(t)$ denotes measurable outputs and inputs, $x(t)$ and $\omega(t)$ represent (mostly unmeasured) state variables and disturbances, and θ are the process parameters. Process faults cause changes in state variables and model parameters. Based on a process model one can estimate $x(t)$ or $\theta(t)$ by observed $y(t)$ and $u(t)$. Residual evaluation is accomplished by threshold logic and decision function. Beside fixed thresholds, advanced robust adaptive residual evaluators exist [33].

2.6.1 Parity Equations

The method based on Parity Equation compares the behavior of the process with a process model describing normal, i.e non-faulty behavior. The key idea is to check the consistency (parity) [34], of the mathematical equations of the system (analytic redundancy relations) by using the actual measurements. The difference of signals between the process and the model (Fig. 2.10) is expressed by residuals.

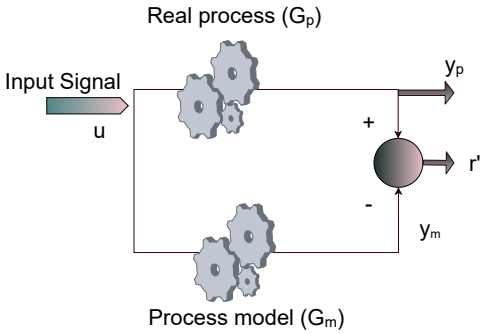


Figure 2.10: Fault detection with parity equations

$G_p(s)$ and $G_m(s)$ are respectively the transfer function which described the process and the pro model. A straightforward model-based method is to take fixed G_M and run it in parallel to the process, thereby forming an output error :

$$r'(s) = [G_p(s) - G_M(s)]u(s) \tag{2.28}$$

2.6.2 State Observers and State Estimation

Changes in the input/output behavior of a process lead to changes of the output error and state variables. The basic idea of the observer approach is to reconstruct the outputs of the system from the measurements with the aid of observers using the estimation error, or innovation, as residual for the detection of the fault [4].

2.6.2.1 States observers

State observer can be applied if the faults can be modeled as state variable changes δx_i. The configuration of linear full order state estimator is shown in Fig. 2.11.

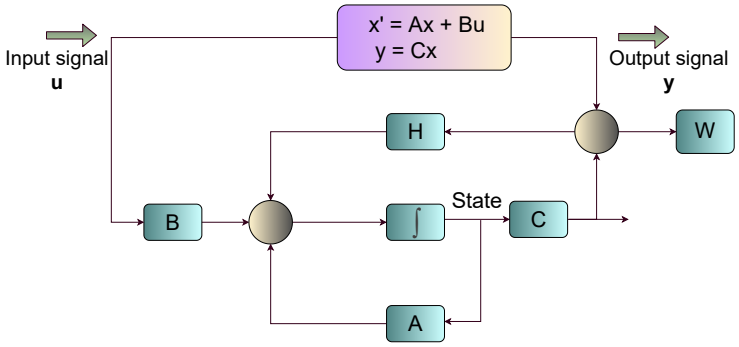


Figure 2.11: Fault detection with state observer

It consists of a parallel model of a process, Eq. 2.29 and Eq. 2.30, with the feedback(matrix H) of the estimation error e [35].

$$\dot{\hat{x}}(t) = A\hat{x}(t) + He(t) \tag{2.29}$$

$$e(t) = y(t) - C\hat{x}(t) \tag{2.30}$$

$$r(t) = We(t) \tag{2.31}$$

e from, is used for calculation of the residual, r, for the purpose of fault detection.

2.6.2.2 Observers

The task of state observers is to reconstruct the states of a process. However, there is generally no such need for diagnostic purpose. It is possible to use output observers if the reconstruction of the state vector $x(t)$ is not of interest.

2.7 Fault Detection with Parameter Estimation

In most practical cases the process parameters are partially not known or not known at all. They can be determined with parameter estimation methods by measuring the input and output signal if the basic model structure is known. Faults of a dynamical system are reflected in physical parameters (friction, mass, resistance, capacitance, inductance etc.). An overview of the procedure adopted in this thesis is depicted in Fig. 2.12

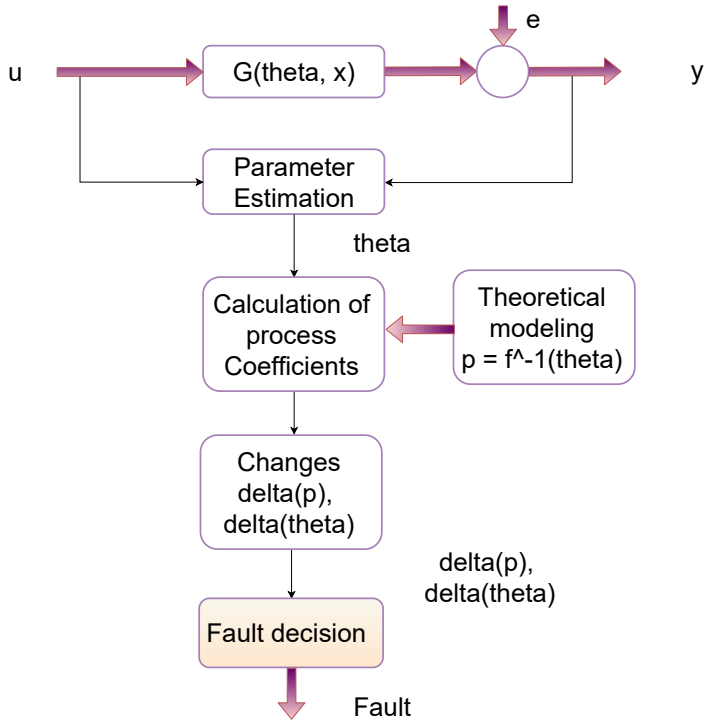


Figure 2.12: A theoretical modeling of fault detection based on parameter estimation

For fully take advantage of the parameter estimation approach, some other

important aspect such as the implementation and robustness, which is not specifically developed in this research, have to be taken into account. The interested reader can find more details in [36–39]. This work proposed a brief presentation of the commonly used approaches for solving mainly the Recursive Least Squares (RLS) algorithm.

2.7.1 Recursive Least Squares algorithms

For a given system which can be represented by a couple input-output such that

$$y(t) = a_1 y(t-1) + ... + a_{na} y(t-na) + b_1 u(t-n_k) + ... + b_{nb} u(t-nb-n_k) \quad (2.32)$$

In the Eq 2.32, a_{na} and b_{nb} are the order model structure and n_k is the delay. The RLS algorithm can be express by :

$$k(t) = \frac{P(t-1)\phi(t)}{1 + \phi^T(t)P(t-1)\phi(t)} \quad (2.33)$$

$$P(t) = P(t-1) - k(t)\phi^T(t)P(t-1) \quad (2.34)$$

$$\epsilon(t) = y(t) - \phi^T(t)\hat{\theta}(t-1) \quad (2.35)$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + k(t)\epsilon(t) \quad (2.36)$$

where $\phi(t) = [-y(t-1)... - y(t-n_a)u(t-nk)...u(t-n_b-nk)]^t$ and $\theta = [a_1, ..., a_{na}, b_1, ..., b_{nb}]$. $\epsilon(t)$ is the innovation error, $P(t)$ is the covariance matrix and $k(t)$ is the innovation gain. Initial values can be provided from the knowledge of the system characteristics or calculated from initial data set using non RLS method. The result minimizes the expression below :

$$V_N(\theta) = \sum_{t-1}^{N} \epsilon^2(t) \quad (2.37)$$

In the Eq. 2.37, N represents the number of the sample. From a practical view, the convergence speed is generally slow which makes the standard RLS estimate method inadequate for real-time fault detection. However, several approaches for modifying the RLS algorithm to make it suitable as a real-time fault detection are

commonly : the forgetting factor, virtual Kalman filter and sliding window data.

2.7.2 Forgetting factor

This technique aimed to modify the loss function to be minimized. If we consider modify the loss function as :

$$V_N(\theta) = \sum_{s=1}^{N} \lambda^{t-s} \epsilon^2(s) \tag{2.38}$$

the measures that are older than $T_0 = 1/(1-\lambda)$ samples are included in the criterion with a weight approximately equal to 36% of that of the most recent measurement. The T_0 means the memory time constant of the criterion and reflects the ratio between time constant of variation in the dynamics and those of the dynamics itself. Several work in the literature [40] made experiences for different values of λ, the results show a decrease in the value of the forgetting factor has two effects :

- The parameter estimates converge to their true valuer faster, consequently, decrease the fault alarm time delay,
- At the expense of increased sensitivity to noise, considering λ is way less than 1, the estimates may even oscillates around their true value.

To overcome this issue, different approaches can be applied, this thesis suggest two options: the Kalman filters and the time-variant forgetting factor. This thesis does not discussed in details those approaches, but the interested ready can find more in [6, 41].

2.8 Fault detection with Nonlinear Model

Most of industrial processes nowadays are not well adapted for conventional modeling approaches due to the lack of precise and explicit knowledge about the system as well as the strong nonlinear behavior. The section 2.7 discussed essentially the mathematical description. But in case the mathematical description is not available, a nonlinear model can be utilized as depicted in Fig. 2.13.

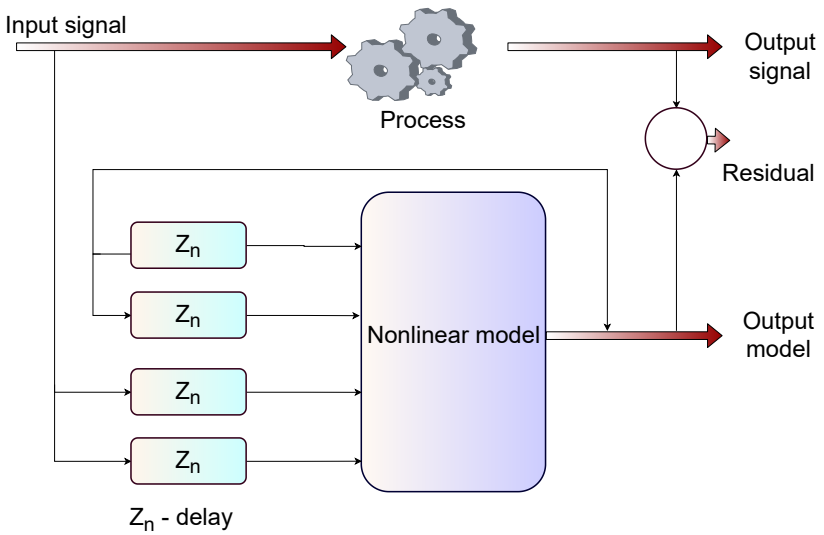


Figure 2.13: Nonlinear model scheme for fault detection in a process.

A nonlinear model can be built using NN (Neural Networks) [42]. Techniques like NN do not necessarily need specific knowledge of the process, they can be considered as a "black-box" models of the general nonlinear, multi-variable static and dynamic system. Despite the number of parameters holding by a NN, they are often not suitable for physical interpretation of the system in question. But, as soon as the process modeling is completed, fault detection with parity equations can be implemented as done in sub-section 2.6.1. Using the NN for model based fault detection with parity equation has been a subject of several research in recent years [43–45]. The pattern recognition using NN can also be combined with several process models and used for residual evaluation [46, 47].

2.9 Knowledge-Based Methods

Elaborations done in previous sections have shown that analytical approaches require a detailed quantitative mathematical model for being effective. However, in most of large scale engineering system, such information may not always be available or may be expensive and time consuming. An alternative to these technique are knowledge-based methods like causal analysis, expert systems and pattern recognition discussed already in the section 2.5. These techniques are mainly based on qualitative models, which can be obtained through causal modeling [48] of the sys-

tem, expert knowledge, a detailed description of the system or fault-symptom relationships [49]. Expert system and fuzzy logic are the most used of these methods in knowledge based. The sub-sections 2.9.1 and 2.9.2 discussed these techniques.

2.9.1 Expert systems

The expert systems are among knowledge based techniques which are closer in style to human problem solving [50]. When well designed, an expert system can fully represent an existing knowledge, accommodate existing databases, accumulate new knowledge, make logical inferences, make recommendations and make decisions with reasoning [51]. The approach offers an efficiency for quasi-static systems operating within a fixed set of rules. Major components of the expert system approach are knowledge and inference engine (Fig. 2.14). One of the advantage of the expert system is that, rules can be added or removed easily. Moreover, explanation of the reasoning process, induction and process deduction are easy as well.

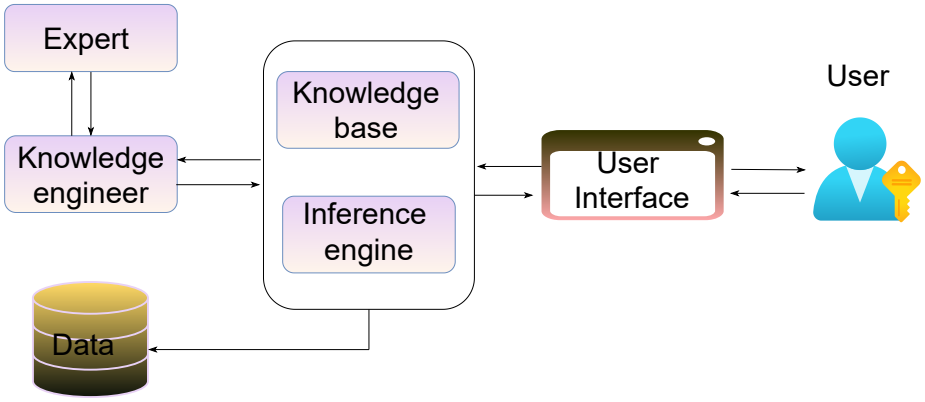


Figure 2.14: Schematic description of main expert system components

In contrast to benefits, things that can be pointed out as drawbacks are lack of generalization [52], poor handling new situations, inability to represent time-varying phenomena to learn from their errors and fault detection architecture may be costly.

2.9.2 Fuzzy logic

The fuzzy logic deals with the more realistic idea beside the detection system's output. In fact, for real applications, a fault detection system does not need to be a

binary output, i.e., fault or no fault. Instead, the fault severity of the system is provided to the operators as the output. A linguistically interpretable rule-based model is formed based on the available expert knowledge and measured data. Fig 2.15 illustrates different possible blocks used in a fuzzy logic architecture.

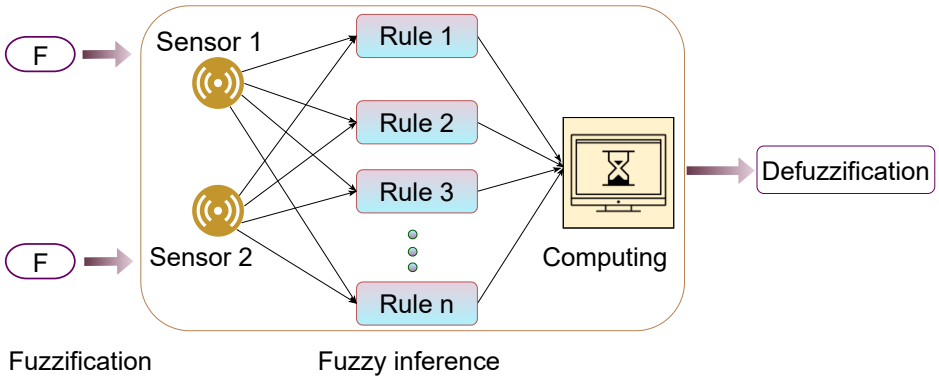


Figure 2.15: Knowledge-based using fuzzy logic

At fuzzification stage, a graphical representation of the magnitude of participation of the inputs called "membership" is passed through membership functions. These function can have different profile such as ramp, linear, etc. [53]. All rules are evaluated using fuzzy reasoning [54]. Ideally, the fuzzy inference process uses membership functions, logical operations and if-then rules. The defuzzification stage is accomplished by combining the results of the inference process and computing the fuzzy centroid of the area [55].

2.10 Conclusion

This chapter established fundamentals and general taxonomies of approaches used in fault detection and diagnosis. The chapter stressed the fact that, when deciding on which approach is suitable for the case in question, a certain number of issues such as type of failures, description of process structure, process dynamics, available signal process, the complexity of the process, quantity, quality and amount of data, etc...have to be considered. This chapter was a preamble for this research report contributions.

References

[1] Rolf Isermann. Model-based fault-detection and diagnosis – status and applications. *Annual Reviews in Control*, 29(1):71–85, 2005.

[2] Rolf Isermann. *Fault-Diagnosis Applications*. Springer, 2008.

[3] 26262-1:2018. Road vehicle engineering (vocabularies) 43.040.10 electrical and electronic equipment. Technical report, ISO, 2018.

[4] G. I. Sainz-Palmero . J. Fuente, D. Garcia-Alvarez and T. Villegas. Fault detection and identification method based on multivariate statistical techniques. In *Proceedings of 2018 Annual IEEE International Systems Conference (SysCon)*, 2009.

[5] Lindley R. Higgins R. Keith Mobley and Darrin J. Wikoff. *Maintenance Engineering Handbook*. The McGraw-Hill Companies, 2008.

[6] Shing-Chow Chan, Jian-Qiang Lin, Xu Sun, Hai-Jun Tan, and Wei-Chao Xu. A new variable forgetting factor-based bias-compensation algorithm for recursive identification of time-varying multi-input single-output systems with measurement noise. *IEEE Transactions on Instrumentation and Measurement*, 69(7):4555–4568, 2020.

[7] D. Shur, A. Kadyshevitch, J. Zelenko, C. Duncan, C. Mata, B. Starr, S. Ventola, and J. Klinger. Wafer current measurement for process monitoring. In *2004 IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop (IEEE Cat. No.04CH37530)*, pages 303–307, 2004.

[8] L. Chen G. Zhang and K. Liang. Fault detection and diagnosis for aerostat sensors based on pca and contribution graph. In *019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2019.

[9] R. Braatz L. Chaing, E. Russel. *Fault Detection and Diagnosis in Industrial Systems*. Springer, 2001.

[10] Krzysztof Patan. Neural network-based model predictive control: Fault tolerance and stability. *IEEE Transactions on Control Systems Technology*, 23(3):1147–1155, 2015.

[11] Linchuang Zhang, Hak-Keung Lam, Yonghui Sun, and Hongjing Liang. Fault detection for fuzzy semi-markov jump systems based on interval type-2 fuzzy approach. *IEEE Transactions on Fuzzy Systems*, 28(10):2375–2388, 2020.

[12] Ethem Alpaydin. *Pattern Recognition*, pages 55–84. 2016.

[13] José Luis Rueda-Torres and Francisco González-Longatt. *Pattern Recognition-Based Approach for Dynamic Vulnerability Status Prediction*, pages 95–117. 2018.

[14] Chen Jing and Jian Hou. Svm and pca based fault classification approaches for complicated industrial process. *Neurocomputing*, 167:636–642, 2015.

[15] Amir Ebrahimifakhar, Adel Kabirikopaei, and David Yuill. Data-driven fault detection and diagnosis for packaged rooftop units using statistical machine learning classification methods. *Energy and Buildings*, 225:110318, 2020.

[16] *Application of the Weighted Nearest Neighbor Method to Power System Forecasting Problems*, pages 41–88. 2017.

[17] Suren Martirosyan. Application of memory scrambling aware multi-level diagnosis flow. In *2017 Computer Science and Information Technologies (CSIT)*, pages 74–77, 2017.

[18] S. Katipamula and M. R. Brambley. *Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems – A Review, Part II*. HVAC and R Research, April 2005.

[19] Sanjay K. Dhurandher, Mohammad S. Obaidat, Amrit Jaiswal, Akanksha Tiwari, and Ankur Tyagi. Vehicular security through reputation and plausibility checks. *IEEE Systems Journal*, 8(2):384–394, 2014.

[20] James V. Stone. *Principal Component Analysis for Preprocessing Data*, pages 179–182. 2004.

[21] Johnson I. Agbinya. *12 Principal Component Analysis*, pages 205–220. 2019.

[22] S Ding, P Zhang, E Ding, A Naik, P Deng, and W Gui. On the application of pca technique to fault diagnosis. *Tsinghua Science and Technology*, 15(2):138–144, 2010.

[23] S Ding, P Zhang, E Ding, A Naik, P Deng, and W Gui. On the application of pca technique to fault diagnosis. *Tsinghua Science and Technology*, 15(2):138–144, 2010.

[24] Lindley R. Higgins R. Keith Mobley and Darrin J. Wikoff. *Maintenance Engineering Handbook*. The McGraw-Hill Companies, 2008.

[25] Chengshi Zheng, Hefei Yang, and Xiaodong Li. On generalized auto-spectral coherence function and its applications to signal detection. *IEEE Signal Processing Letters*, 21(5):559–563, 2014.

[26] Andreas Jakobsson, Stephen R. Alty, and Jacob Benesty. Estimating and time-updating the 2-d coherence spectrum. *IEEE Transactions on Signal Processing*, 55(5):2350–2354, 2007.

[27] A. Rao and R. Kumaresan. A parametric modeling approach to hilbert transformation. *IEEE Signal Processing Letters*, 5(1):15–17, 1998.

[28] Fei Hua, Cédric Richard, Jie Chen, Haiyan Wang, Pierre Borgnat, and Paulo Gonçalves. Learning combination of graph filters for graph signal modeling. *IEEE Signal Processing Letters*, 26(12):1912–1916, 2019.

[29] Jong Won Shin, Joon-Hyuk Chang, and Nam Soo Kim. Statistical modeling of speech signals based on generalized gamma distribution. *IEEE Signal Processing Letters*, 12(3):258–261, 2005.

[30] M. Soualhi, K.T.P. Nguyen, and K. Medjaher. Pattern recognition method of fault diagnostics based on a new health indicator for smart manufacturing. *Mechanical Systems and Signal Processing*, 2020.

[31] R. Finn M. Charest and R. Dubay. Integration of artificial intelligence in an injection molding process for on-line process parameter adjustment. In *Proceedings of 2018 Annual IEEE International Systems Conference (SysCon)*, 2018.

[32] Y. Lounici, Y. Touati, and S. Adjerid. Uncertain fault estimation using bicausal bond graph: Application to intelligent autonomous vehicle. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, pages 1140–1171, 2020.

[33] R.N. Banerjee Anupam Das, J. Maiti. Process monitoring and fault detection strategies: a review. *International Journal of Quality and Reliability Management*, 2012.

[34] S. Cho and J. Jiang. A fault detection and isolation technique using nonlinear support vectors dichotomizing multi-class parity space residuals. *Journal of Process Control*, 2020.

[35] STMicroelectronics. Mems digital output motion sensorultra low-power high performance 3-axes “nano” accelerometer. Technical report, STMicroelectronics group of companies, 2019.

[36] Marcelo G. S. Bruno. 2013.

[37] T. Jiang, K. Khorasani, and S. Tafazoli. Parameter estimation-based fault detection, isolation and recovery for nonlinear satellite models. *IEEE Transactions on Control Systems Technology*, 16(4):799–808, 2008.

[38] Balázs Kulcsar and Michel Verhaegen. Robust inversion based fault estimation for discrete-time lpv systems. *IEEE Transactions on Automatic Control*, 57(6):1581–1586, 2012.

[39] W.H. Chung and J.L. Speyer. Fault detection via parameter robust estimation. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 4, pages 3976–3981 vol.4, 1997.

[40] H. J. Tan, S. C. Chan, and L. Zhang. A variable forgetting factor qrd-based rls algorithm with bias compensation for system identification with input noise. In *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2643–2646, 2016.

[41] Reza Zekavat and R. Michael Buehrer. *Kalman Filter-based Approaches for Positioning: Integrating Global Positioning with Inertial Sensors*, pages 763–838. 2019.

[42] Haiying Qi, Abdelkarim Ertiame, Kingsley Madubuike, Ding-Li Yu, and J Barry Gomm. Nonlinear observer fault detection for a multivariable process using a learning methodology. In *2018 24th International Conference on Automation and Computing (ICAC)*, pages 1–6, 2018.

[43] Kai Wang, Junghui Chen, Zhihuan Song, Yalin Wang, and Chunhua Yang. Deep neural network-embedded stochastic nonlinear state-space models and their applications to process monitoring. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–13, 2021.

[44] Tongli Jia, Yunqing Liu, and Jiaqi Li. Fault detection for mechanical arm systems: An sliding mode observer approach. In *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 1541–1544, 2019.

[45] Gavneet Singh Chadha and Andreas Schwung. Comparison of deep neural network architectures for fault detection in tennessee eastman process. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–8, 2017.

[46] Denis Berdjag, Cyrille Christophe, and Vincent Cocquempot. Nonlinear model decomposition for fault detection and isolation system design. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 3321–3326, 2006.

[47] Zhiwen Chen, Ketian Liang, Steven X. Ding, Chao Yang, Tao Peng, and Xiaofeng Yuan. A comparative study of deep neural network-aided canonical correlation analysis-based process monitoring and fault detection methods. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–15, 2021.

[48] S. Gentil, J. Montmain, and C. Combastel. Combining fdi and ai approaches within causal-model-based diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(5):2207–2221, 2004.

[49] Sujeong Baek and Duck-Young Kim. Fault prediction via symptom pattern extraction using the discretized state vectors of multisensor signals. *IEEE Transactions on Industrial Informatics*, 15(2):922–931, 2019.

[50] Mircea Eremia, Chen-Ching Liu, and Abdel-Aty Edris. *Expert Systems*, pages 731–754. 2016.

[51] Kenneth D. Forbus. *19 Expert Reasoning*, pages 349–365. 2019.

[52] M. Tang, Q. Zhao, H. Wu, and Z. Wang. Cost-sensitive lightgbm-based online fault detection method for wind turbine gearboxes. *Frontiers in Energy Research*, 9, 2021.

[53] Swati Aggarwal, Ranjit Biswas, and A.Q. Ansari. From fuzzification to neutrosophication: A better interface between logic and human reasoning. In *2010 3rd International Conference on Emerging Trends in Engineering and Technology*, pages 21–26, 2010.

[54] Tamalika Chaira. *Fuzzy/Intuitionistic Fuzzy Measures and Fuzzy Integrals*, pages 111–132. 2019.

[55] Mircea Eremia, Chen-Ching Liu, and Abdel-Aty Edris. *Fuzzy Systems*, pages 785–818. 2016.

Chapter 3

Non-learning methods for pre-processing and feature extraction : an enhancement of traditional techniques

This chapter presents most of non learning techniques used in thesis for the pre-processing stage and for faults signature extraction. Most of presented results are peer reviewed and available online in both journals and proceedings.

3.1 Introduction

The effectiveness and relevance of a detection and diagnostic system, lies on the quality of the parameters used as indicators. These parameters need to have a sufficient informational content to be able to discriminate between different conditions equipment operation. Measured data are not often exploited without an adequate pre-processing stage that can be performed in time domain, frequency domain or in time-frequency domain. This process can be seen as a representation of these signals in a space with N dimensions where the discrimination is better. The time domain method uses measurements of the power of the monitored signal, the variation energy in this case is expected to increase. Another technique in using the time domain is to focus on statistical parameters analysis by considering the monitored signal as a random variable and then using higher-order statistical moments and cumulants as features. The frequency domain on the other hand exploits the principle that, the

spectrum of the monitored signal changes when faults occur. The condition of the equipment is assessed through the change in the spectrum or number of discriminating features extracted from the spectrum of the measured signal. In some situations these features can be specific frequency components that depend on the equipment in question as well as on the nature of the fault under investigation.

3.2 Cascades based methods

In the diagnosis process, several aspects, such as the data acquisition process, environmental influence, etc... have a direct impact on the analysis outcomes. Usually fault information carry by the signal tend to be masked by the disturbances [1–4]. Different statistical parameters can be extracted for understanding the behavior of the fault. Among those features, Kurt is the most used due to its sensitivity to the signal impulses [5]. Nevertheless, Root Mean Square (RMS), CF, and skewness (skew) are highly interesting for non-stationary [6–8] signal; their reliability still is known to be an issue in the presence of the noise. The unknown noise sources often make the detection task much harder. In such situations, improving the quality of the signal by removing noise is a required step. Various denoising methods have been studied and validated, i.e., features in both time and frequency domain show more accurate characteristics. Lu, S et al. [9] applied the Stochastic Resonance (SR), a nonlinear filter for denoising a vibration signal to improve the weak signal with the assistance of proper noise. In [10] they applied the same approach by a full-wave signal construction strategy. An impressive result was presented by Gavish and Donoho [11], where they have investigated the use of Wavelet Transform (WT) based on the soft thresholding. The same idea was used by Djebala et al. [12]; they specifically analyzed the technique based on Wavelet Multiresolution Analysis (WMA) optimization for bearing fault diagnosis. Wang et al. [13] came up with a novel denoising method based on multiwavelet denoising for fault detection in bearings of the locomotive. The idea suggested by Dohono and Johnston was used by Sadooghi and Khadem [14]. They applied the DWT, Donoho-Johnston function, and the parameter method thresholding for denoising the jet engine vibration signal. The sub-sections below are discussing the use of Discrete Wavelet Transform (DWT), Empirical Mode Decomposition (EMD) and Ensemble Empirical Mode Decomposition (EEMD) by using the vibration signal. Each technique's performance is evaluated relying on parameters such as the mother wavelet, the decomposition level, the thresholding type, the interpolation model, and so on.

3.2.1 Empirical Mode Decomposition

The EMD is defined by a process called screening (sifting), which consists of decomposing the signal in core contributions called empirical modes or Intrinsic Mode Functions (IMF). The decomposition is local, iterative, sequential, and entirely data-driven [15]. The primary idea is to extract a non-stationary signal from systems that can be non-linear patterns for a time-frequency analysis, where Fourier analysis and wavelets are sometimes ineffective. The method makes the time-frequency representation more readable and physically easy to interpret.

In practice, IMF estimation requires some primary steps consisting of finding the local maxima points of the signal and connecting them using a cubic spline to obtain a lower envelope [1]. After calculating the mean value of the two envelopes, the signal's first component is computed by subtracting the mean calculating from the original signal. This can be express as :

$$h_1(t) = x(t) - m_1(t) \tag{3.1}$$

$x(t)$ is the original signal, and m_1 is the calculated mean value. If h_1 after calculation meets the IMF criteria, it will be taken as the first IMF; otherwise, the iteration will keep running by substituting the original signal $x(t)$ by h_1. Once the criteria are satisfied, the first IMF is stored and subtracted from the original as :

$$x_{new}(t) = x(t) - c_1(t) \tag{3.2}$$

c_1 is the first IMF. The original signal is once again substituted by the new $n_{new}(t)$ and the process also runs for the following IMFs. The last IMF is called the residual and noted $r(t)$. The original signal can be reconstructed using the residual and different IMFs calculated by :

$$x_{new}(t) = \sum_{i=1}^{K} c_i(t) + r(t) \tag{3.3}$$

In the Eq. 3.3, K represents the number of IMFs calculated. Even if EMD is considered as a powerful and reliable technique in non-linear and non-stationary signal analysis, its major drawback is the mode mixing [16]. The mode mixing can

be interpreted as the appearance of the same frequency mode in two different IMF's, which usually results in aliasing. Another challenge when using EMD is that it is unable to address the end effect phenomenon. Thanks to the brilliant researches presented by different scientists, particularly Huang and Wu [17], a new approach called Ensemble EMD was introduced to address those issues carry by the EMD.

3.2.2 Ensemble Empirical Mode Decomposition

EEMD technique solves the problem of mode mixing in the EMD algorithm by adding the white Gaussian noise uniformly in the original signal. If one considers adding noise to given signal M times and decomposing the signal using EMD. Each new signal can then be written as :

$$x_i(t) = x(t) + w_i(t) \tag{3.4}$$

In the Eq.3.4, $w_i(t)$ is the i-th added white noise to the original signal $x(t)$ and $x_i(t)$ is the generated noisy signal. The i-th noisy signal can be expressed as :

$$x_i(t) = \sum_{j=1}^{n} c_{ij} + r_i(t) \tag{3.5}$$

In the Eq.3.5, c_{ij} represents the j-th calculated IMF of noisy signal $x_i(t)$ and $r_i(t)$ is the residual signal. The collection of the noisy signals gives the resulting IMFs. The accuracy of the result obtained by using EEMD is strongly based on the number of ensemble and amplitude of the added noise (Eq.3.6).

$$\epsilon_n = \frac{\epsilon}{\sqrt{N}} \tag{3.6}$$

ϵ is the added noise, ϵ_n the standard deviation of the error, and N the number of realizations. The Fig. 3.1 presents the flowchart of the proposed EEMD algorithm.

3.2.3 Discrete Wavelet Transform

WT has been in the center of most research work since it was formulated in the late 1980s by Daubechies, Mallat, and others [18–20]. The transformation consists

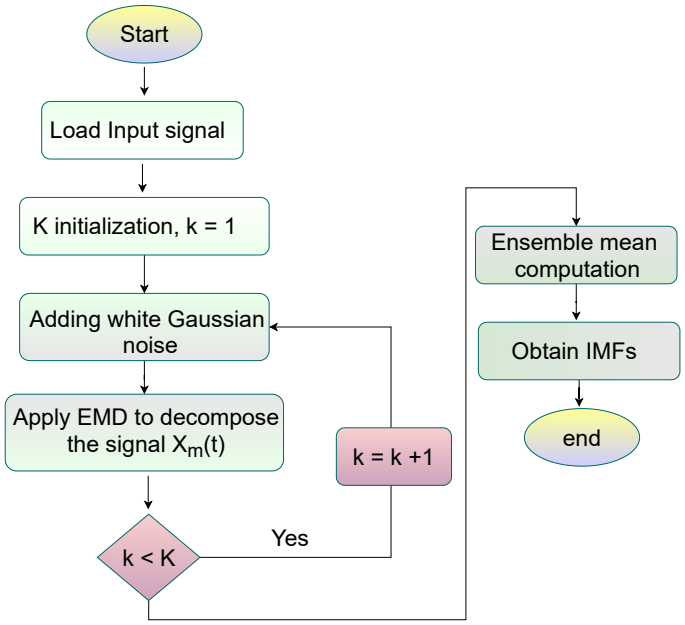


Figure 3.1: The flowchart of the proposed EEMD algorithm

of scaling and translating the "mother wavelet", one can express this as :

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right) \Rightarrow \psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k). \tag{3.7}$$

This approach is currently used as a Continuous Wavelet Transform (CWT), which, unfortunately, has been proven weak in terms of presenting information [18]. In this work, the choice made on DWT is justified because it performs the transformation only for a set of positions and scales. In work done by [21], an effective method is implemented; it uses filter banks and convolutes the signal with low-pass and high-pass filters. The filters are the actual wavelets with the scale adapted to the studied frequencies.

3.3 Spectrum Analysis

As stated above, often, the time domain signals are uncertain and difficult to analyze, on the other hand, their spectra are pretty easy to exploit [22]. The spectral

analysis allows the decomposition of a starting signal easily into independent signals. Different components are built to distinguish specific components delivering the trend with a shallow frequency in variation [23, 24]. Other components represent oscillation content and noise. For a periodic signal of period $T = \frac{1}{f_0}$, the basic equation can be written as :

$$x(t) = \sum_{k=-\infty}^{+\infty} X_k e^{i\frac{2\pi k\Delta t}{T}} \tag{3.8}$$

where X_k are Fourier coefficients of $x(t)$. The spectrum is thus a spectrum of lines with energy only at the frequencies . The spectrum density can be written as :

$$\Gamma_x(\nu) = |X_k|^2\delta(\nu - kf_0), \tag{3.9}$$

,

Generally, spectral estimation problem can be started more formally as a finite-length recorded $[y(1), ..., y(N)]$ of a second oder stationary random process. From the recorded data determine an estimate $\hat{\phi}(\omega)$ of its power spectral density $\phi(\omega)$, for $\omega \in [-\pi\ \pi]$ [25, 26]. In fault detection using vibration signals, spectral analysis is one of the best choices allowing the distribution of vibratory [27, 28] energy according to the frequency and giving the various amplitudes for various frequencies. It can be noted that various types of defects often appear at particular frequencies.

3.4 Envelope analysis

In Fig. 3.2, the primary step of envelope analysis is given. It consists of finding impulsive signals even when they have a deficient energy level up to the point of being hidden by other spectral components.

The envelope analysis fetches given signals in the high-frequency zones by demodulating around a well chosen resonance [29]. The interpretation of envelope signal is often done after FT. The method has advantage of overcomes the interference of non-impulsive signals of low frequency resulting from the motor such as unbalance, misalignment, and so on [30]. For faults related to bearings, envelope analysis is better than raw spectral analysis.

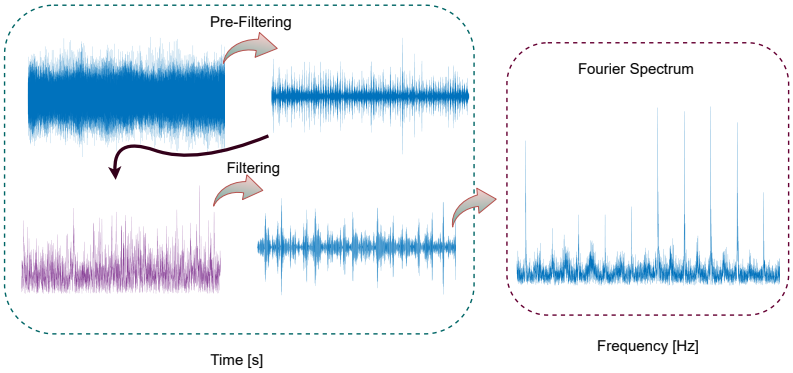


Figure 3.2: Envelope analysis principle

3.5 Park transformation

Diagnosis using Parks vector [28, 31] analysis proposes two possible methods :

- The first uses two-phase quantities i_{ds} and i_{qs}, which are calculated from three supply current for obtaining the curve calls "Lissajous Curve" : $i_q = f(i_d)$. The change in thickness and shape of the curve provides information on the fault in the motor [32, 33].
- The second called Extended Parks Vector (EPV), is based on spectral analysis of the module of Parks vector. EPV has great advantages especially in stator and rotor fault detection.

Park Transformation is composed of three-phase - two-phase transformation follows by a rotation as illustrated in Fig. 3.3.

It allows to shift from the frame abc to the frame $\alpha\beta$ then to dq. The frame $\alpha\beta$ is always fixed respect to the frame abc, however the frame dq can move.

It is important to notice that, the Lissajous curve produce a circular shape with its center at origin [34]. The faults in three phase electric motor, such as rotor fault, lead to the distortions in circular pattern. For fault detection, the circular pattern is compared with reference circle to check any deviation [35]. The Park's vector approach is computationally more efficient but usually used for fault detection in steady state [36]. Since changing current magnitude produce circular pattern of different diameter therefore advanced techniques are applied for fault detection in dynamic state. These techniques implement normalized current values for fault detection [36].

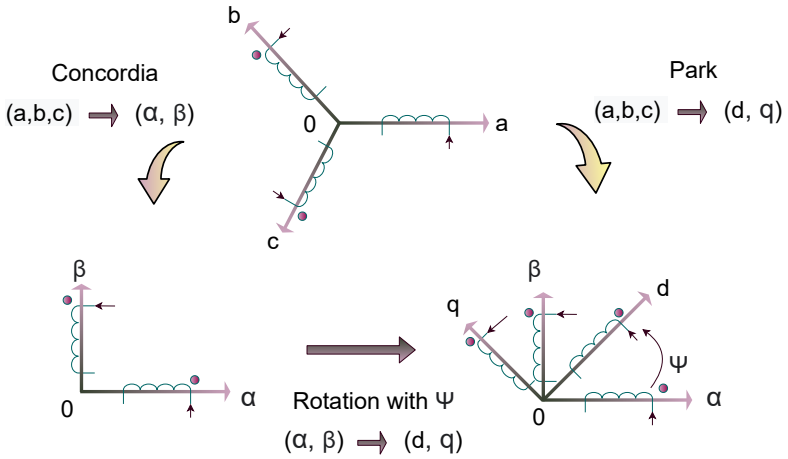


Figure 3.3: Spatial representation of Park transformation

3.6 Application : Electric machines

3.6.1 The test bench

For evaluating the performance of the proposed techniques, several experimental studies have been performed using a squirrel cage induction motor directly coupled to a compressor acting as a load as depicted in Fig. 3.4.

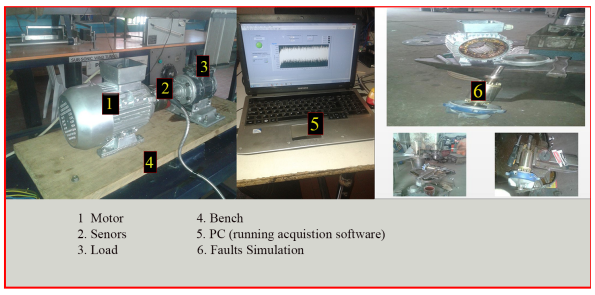


Figure 3.4: Bench of the test

The experiment system consists of $IEC34 - 1$ 3-phase motor (2.2 kW, 2885 rpm, low-voltage squirrel cage). Signals are acquired thanks to National Instruments (NI) materials (Fig.3.4). The acquisition is performed at 5120 Hz sampling rate, using a multi-channel acquisition system. The measurement campaign was carried out in 120 different operating condition of 20 seconds each represented by 4

vibration signals. The acquision system includes :

- Compact DAQ on which NI9233 module was mounted for conditioning
- a laptop Dell Intel Pentium UHD 605, 4 GB Memory, 256 GB SSD
- an industrial PCB piezoelctronics accelerometer SN 7470 with a sensitivity of $10.3mV/(m/s^2)$

The denoising methods, namely EEMD and DWT are applied to the recorded signals. The Fig. 3.5 presents signals for different analyzed scenarios.

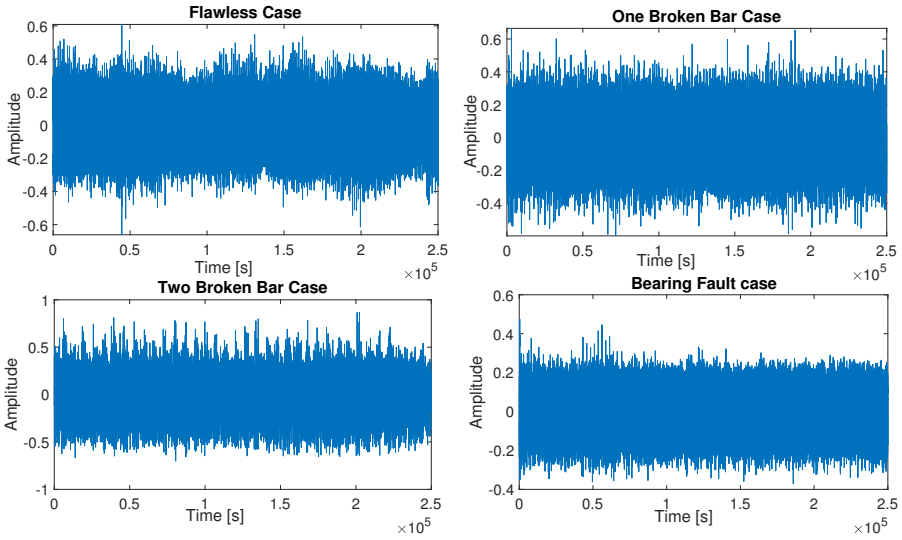


Figure 3.5: Raw signal recorded from different analyzed scenarios.

3.6.2 Results and discussion

For successful outcomes, some crucial steps have to be considered during the analysis; in this particular work, we evaluated the parameter selection based on the factual information they have carried and their effectiveness.

Applying EEMD on signals of the different analyzed cases shows that the method has interesting capabilities. EEMD transforms the vibration signal from healthy case to null signal whatever the selected parameters. The technique highlights only singularities due to instabilities (faults) in the motor. This characteristic

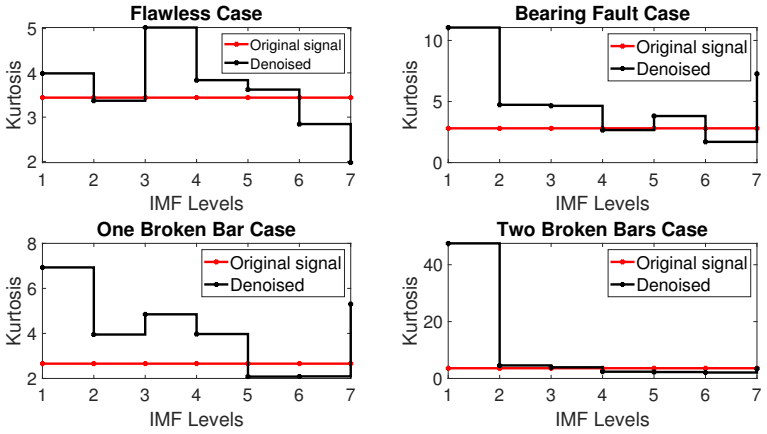


Figure 3.6: Influence of the EEMD decomposition levels on healthy motor and different fault cases ⟶ Kurtosis case

demonstrates that the denoising using EEMD applies primarily to time-domain signals. In Fig. 3.6, the linear interpolation model is used for the determination of IMFs, then Kurt and CF. The results demonstrate that both Kurt and CF(Fig 3.7) perform well in all analyzed cases, only from level 3 to 5 of the decomposition.

One can also notice that, in the case of the bearing fault, Kurt performs better the early detection over the crest factor. On the other hand, it is interesting to notice that both indicators decrease appreciably for level 4 of the decomposition in the presence of broken bars. In Fig. 3.8, an example of the choice of reconstruction for the kurtosis. Soft-thresholding is the most happen to be the most relevant.

Indeed, in the case of EEMD, the sifting effect has a tremendous influence on the results; in DWT, the thresholding type is a big deal for the reliability and interpretability of the results. In this study, we tested symlet and Daubechies as mother wavelets; however, the results presented are based on Daubechies. As mention above, wavelets are often used for signals which contain some singularities. The Fig. 3.10 shows the influence of the decomposition level on the kurtosis and crest factor.

One can see (in Fig 3.9 and Fig. 3.10 (a) and (b)) that, the values of both Kurt and CF are more significant at the at the level 1 of the decomposition. From level 2, they decrease relatively beyond the spectrum of decomposition level. It is worth to mention here also that the crest factor for the DWT case is more sensitive to a fault, as it can be seen in the Fig 3.10 (a) and (b), the CF value decreases up to 7% from the level 1 to the level 2. The mother wavelet and the decomposition level

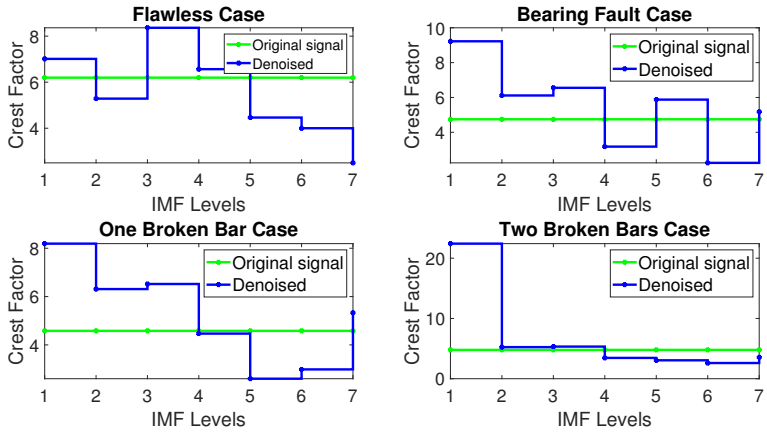


Figure 3.7: Influence of the EEMD decomposition levels on healthy motor and different fault cases ⟶ Crest Factor case

Figure 3.8: The effect of the weighting mode on different analyzed cases.

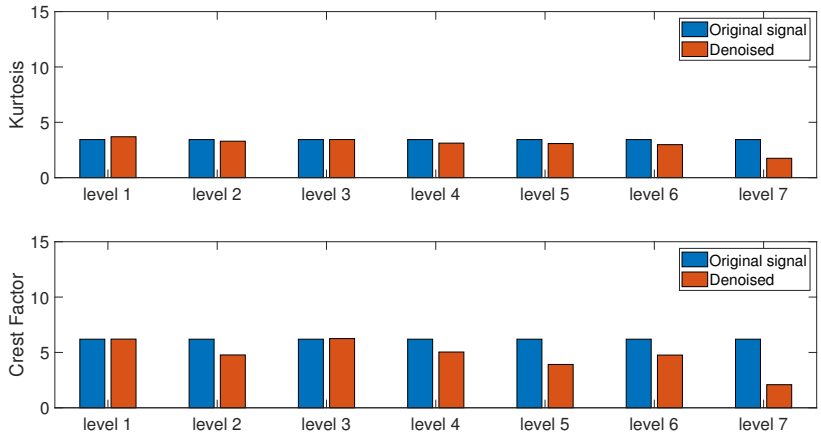
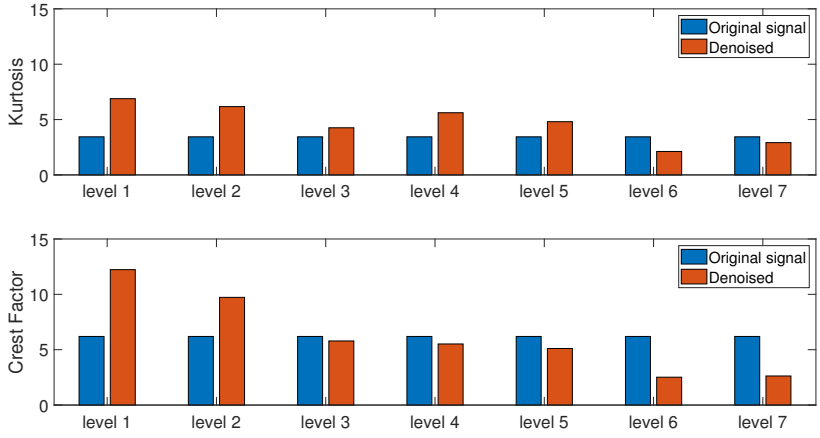


Figure 3.9: The effect of the decomposition level on Kurt and CF ⟶ Healthy Motor

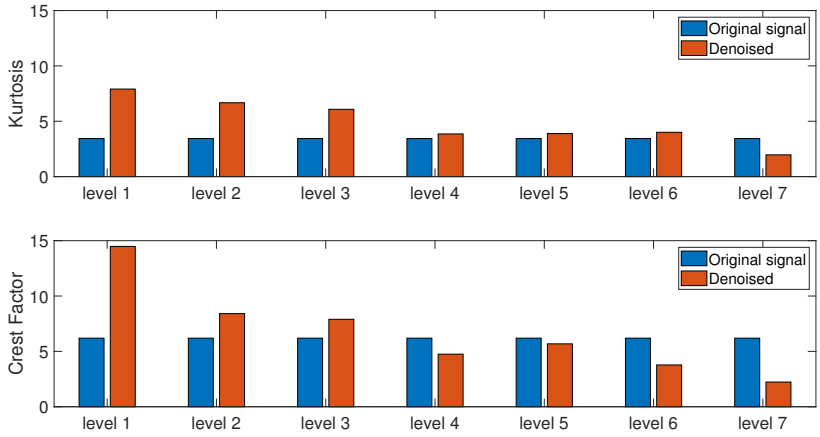
suitable for these scenarios are determined. It is interesting to see that while using EEMD, the Kurt is well related to fault situations up to level 6, whereas, with the DWT, the relation is limited to level 3, hardly to level 4. One can observe that DWT is better than EEMD when we considered the CF. However, this correlation holds only the noise at a low level. Focusing on CF can lead to the conclusion that EEMD provides nonrelevant values. Thus, it is suitable for faults such as a bearing or broken bars to use EEMD and the Kurt indicator. However, DWT remains substantially less efficient; it can be more suitable for the selected indicator.

3.7 Conclusion

In this chapter, the focus was on non-learning methods that can be used for both pre-processing and feature extraction. In addition, the presented methods could be used for detection as well if well elaborated. The experimental investigation targeted two major cascade approaches for vibration signal denoising. It was demonstrated that DWT was primarily challenging to use because it requires in prior a set of parameters well defined. The EEMD, however, could take few parameters thanks to its adaptability to the signal. The parameters selection was limited and had less influence on the denoising performance. Interesting benefits of EEMD were pointed out, such as the transformation of the signal without fault to the null signal. The chapter demonstrates that, regardless of the selected indicator, DWT was discovered to be



(a)



(b)

Figure 3.10: The effect of the decomposition level on Kurt and CF ⟶ (a) Bearing Fault Case and (b) One Broken Bar

of more general purpose. The only issue that should be properly address is that, by revealing singularities, they can skip useful information on faults that do not show up necessarily in the form of singularities.

References

[1] M.A. et al Ugwiri. Vibrations measurement and current signatur ands for fault detection in asynchronous motor. In *I2MTC 2020 - International Instrumentation and Measurement Technology Conference, Proceedings*, number 9128433, 2020.

[2] Y. Liu and A.M. Bazzi. A review and comparison of fault detection and diagnosis methods for squirrel-cage induction motors: State of the art. *Meccanica in press*, 70:400–409, 2017.

[3] D. Capriglione, M. Carratù, S. Dello Iacono, A. Pietrosanto, P. Sommella, and M. Avoci Ugwiri. Design and implementation of a diagnostic scheme for stroke sensors in motorcycle semi-active suspension systems. *Springer*, pages 335–341, 2020.

[4] M. Carratù, A. Pietrosanto, P. Sommella, and V. Paciello. Smart wearable devices for human exposure vibration measurements on two-wheel vehicles. *Acta IMEKO*, 2018.

[5] M.A. Ugwiri, M. Carratu, V. Paciello, and C. Liguori. Spectral negentropy and kurtogram performance comparison for bearing fault diagnosis. In *17th IMEKO TC 10 and EUROLAB Virtual Conference "Global Trends in Testing, Diagnostics and Inspection for 2030"*, 2021.

[6] Lemma T Omar, N. Gebremariam M, and S Ahsan. Anti-friction bearing malfunction detection and diagnostics using hybrid approach. pages 117–131, 2020.

[7] H Yiming Lei, W Jianfeng, Q, and R Hao. A fault diagnosis methodology based on non-stationary monitoring signals by extracting features with unknown probability distribution. (9025196):59821–59836, 2020.

[8] M.A. Ugwiri, I. Mpia, and A. Lay-Ekuakille. Vibrations for fault detection in electric machines. 23(1):66–72, 2020.

[9] S. Lu, Q. He, and J. Wang. A review of stochastic resonance in rotating machine fault detection (2019) mechanical systems and signal processing. 116:230–260.

[10] Siliang Lu, Qingbo He, Haibin Zhang, and Fanrang Kong. Rotating machine fault diagnosis through enhanced stochastic resonance by full-wave signal construction, mechanical systems and signal processing. 85:82–97, 2017.

[11] M. Gavish, Donoho, and D.L. The optimal hard threshold for singular values is 4/*sqrt*3. (6846297):5040–5053, 2014.

[12] A. Djebala, N. Ouelaa, N. Hamzaoui, and L. Chaabi. Detecting mechanical failures inducing periodical shocks by wavelet multiresolution analysis. application to rolling bearings faults diagnosis. 2006.

[13] L. Wang, F. Chen, Z. Li, and S. Zhang. De-noising method of insar data based on empirical mode decomposition and land deformation monitoring application. In *Proceedings of SPIE - The International Society for Optical Engineering*, 2011.

[14] M.S. Sadooghi and S. Esmaeilzadeh Khadem. A new performance evaluation scheme for jet engine vibration signal denoising. *Mechanical Systems and Signal Processing*, 2016.

[15] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Snin, Q. Zheng, N.-C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and the hubert spectrum for nonlinear and non-stationary time series analysis. pages 903–995, 1998.

[16] G. Tang, Y. Wang, Y. Huang, N. Liu, and J. He. Compound bearing fault detection under varying speed conditions with virtual multichannel signals in angle domain. *IEEE Transactions on Instrumentation and Measurement*, 2020.

[17] Z. Wu and N.E. Huang. *Ensemble empirical mode decomposition: A noise-assisted data analysis method*. 2009.

[18] Stéphane Mallat. pages 263–376. Academic Press, 2009.

[19] R. Bouhali, K. Tadjine, H. Bendjama, and M.N. Saadi. Fault diagnosis of bladed disc using wavelet transform and ensemble empirical mode decomposition. *Australian Journal of Mechanical Engineering*.

[20] B. Akhil Vinayak, K. Anjali Anand, and G. Jagadanand. Wavelet-based real-time stator fault detection of inverter-fed induction motor. *IET Electric Power Applications*, 2020.

[21] Y Qin, Mao, Y, Tang, B, Wang, Y, and H Chen. M-band flexible wavelet transform and its application to the fault diagnosis of planetary gear transmission systems. *Mechanical Systems and Signal Processing*, 2019.

[22] D Capriglione, G Cerro, L Ferrigno, and G Miele. *Analysis and implementation of a wavelet based spectrum sensing method for low SNR scenarios*, pages 1–6. Coimbra, 2016.

[23] D Capriglione, G Cerro, L Ferrigno, and G Miele. *Analysis and implementation of a wavelet based spectrum sensing method for low SNR scenarios*, pages 1–6. Coimbra, 2016.

[24] M Ugwiri, M Carratù, C Liguori, and V Paciello. *Bridge dynamic analysis for accurate control and monitoring: Feasibility study*, pages 5307–5311. Lisbon, Portugal, 2019.

[25] A Buonanno, M D'urso, G Prisco, M Felaco, L Angrisani, M Ascione, R Schiano Lo Moriello, and N Pasquino. A new measurement method for through-the-wall detection and tracking of moving targets. *Measurement: Journal of the International Measurement Confederation*, 46(6):1834–1848, 7 2013.

[26] D Capriglione, M Carratù, A Pietrosanto, and P Sommella. *Real-time implementation of an IFD scheme for motorcycle sensors*, pages 1–6. Houston, TX, 2018.

[27] C Cristalli, R Rodriguez, and N Paone. Mechanical fault detection of electric motors by laser vibrometer and accelerometer measurements. mechanical systems and signal processing, 2006.

[28] B Ebrahimi and J Faiz. Magnetic field and vibration monitoring in permanent magnet synchronous motors under eccentricity fault. *IET Electric Power Applications*, 6(1):35–45, 2012.

[29] L Angrisani, F Bonavolontà, A Liccardo, and R Schiano Lo Moriello. Identification and classification of transformers current transients through huang hilbert transform. *MEASUREMENT*, 125:123–132, 2018.

[30] P Tavner. Review of condition monitoring of rotating electrical machines. *IET Electric Power Applications*, 2(4):215–247, 2008.

[31] B Asad, T Vaimann, A Kallaste, A Rassolkin, and A Belahcen. *Hilbert transform, an effective replacement of park's vector modulus for the detection of rotor faults*. 2019.

[32] F El Hammouchi, L El Menzhi, A Saad, Y Ihedrane, and B Bossoufi. Wind turbine doubly-fed asynchronous machine diagnosis defects using stator and rotor currents lissajous curves. *International Journal of Power Electronics and Drive Systems*, 10(2):961–970, 2019.

[33] G Vachtsevanos, F Lewis, M Roemer, A Hess, and B Wu. Intelligent fault diagnosis and prognosis for engineering systems, 2006.

[34] G Angelo, M Laracca, S Rampone, and G Betta. Fast eddy current testing defect classification using lissajous figures. *IEEE Transactions on Instrumentation and Measurement*, 67(4):821–830, 4 2018.

[35] D' Gianni, Marco Angelo, Salvatore Laracca, and Rampone. Automated eddy current non-destructive testing through low definition lissajous figures" proceeding of. *IEEE International Workshop on Metrology for Aerospace*, pages 22–23, 6 2016.

[36] A Hameed, S Gul, and A Khan. *A Park's vector approach using process monitoring statistics of principal component analysis for machine fault detection*. 2017.

Chapter 4

Cyclostationarity diagnosis using advanced spectral techniques

This chapter provides non-exhaustive advanced frequency representation and analysis. The presented methods and techniques aim to demonstrate the selection and interpretation of the most informative frequency band. On the other hand the chapter deals with the visual representation of spectrum and the proposed approach displays vectors of cyclical frequencies as sample rate.

4.1 Introduction

In condition monitoring based on signal processing, one of the typical symptoms is the presence of repetitive transients, which are characterized by impulsive and cyclostationarity signature. The approach quite popular nowadays in the industry for their detection is time-frequency techniques [3, 8]. Those techniques are mainly analysis tools instead of processing tools, and in any case, they are unable to offer a versatile methodology that applies to all mechanical signals in all circumstances. In recent developments, researchers tend to deeply being interested to take advantage of the periodic aspect of the signal components. For instance, a periodically varying filter is used for fault indication in [1]. Actually, the periodicity is one of the reasons that the popular approach is to use multidimensional data representations, particularly the time-frequency representations. Many papers published on cyclostationary analysis of vibration signals in the diagnostics of machines operating in a cyclic way can be

found in [2–6]. Other similar approach based on the fractional moments analysis [6] has been presented to deal with this kind of situation. The fractional lower-order covariance map has been used to observe the periodicity of the fault-related component. Based on the fractional moments analysis, X. Ma and C.L. Nikias [7], proposed a lower order statistics analysis. Unluckily, those approaches as well focus mainly on the fact that the periodicity cannot return the time series of a component as a result, so neither the fault component can be extracted and observed, nor the random impulsive component can be accurately studied. All approaches mentioned above can be considered beneficial in fault detection based on certain conditions if well exploit. This chapter proposed on the one hand, methods and techniques for the right frequency band collection tips and accurate interpretation of the most informative band. On the other hand the chapter deals with the visual representation of spectrum and proposed approach displays vectors of cyclical frequencies in terms of the sample rate.

4.2 Limitations of traditional methods

The traditional time-domain and spectral analysis techniques have two main shortcomings :

1. They do not provide information about the time-dependence of the frequency content of non-stationary signals, and vice versa.

2. The Fourier transform does not accurately and conveniently represent the signal that has non-periodic component such as a transient signal.

For most of rotating machines, the presence of certain frequency components has demonstrated to be indicative of faults. However, since some of these frequencies depend on the rotational speed, it is not possible using spectral analysis to determine them when the equipment runs at a variable rotational speed. Many studies have stressed the usefulness of machines monitoring during the transient states [1, 3, 7], because some machine failures happen during these transition stages. Transient signal are good source of information about machine condition that is not available during steady states [5]. This makes traditional frequency methods not suitable for machine monitoring.

4.3 The cyclostationarity: understanding the informative events

As for other type of signals, information in vibration is often conveyed by non stationarity events [8]. This implies that, during the monitoring of the targeted equipment, any change in the structure of the signal will possibly disclose information about the underlying system's condition. One will observe in this case, unrecognized transients generated by the incipient faults, which will locally change the frequency spectrum and the amplitude of the signal at the time of their occurrence. The energy flow of the signal will unexpectedly experienced considerable fluctuations in some frequency bands, consequently more gain in information. This vision has interesting thermodynamic roots. As a reminder, the concept of entropy which is central both to thermodynamics and the theory of information. For a discrete and stochastic process that is allowed to explore a set of states which can be indexed by i, with the probability p_i, the theory is generally defined as :

$$H = \sum_i p_i ln(p_i) \tag{4.1}$$

This definition is common to statistical thermodynamics known as Bolzmann's entropy [9] and signal processing. It might subsequently differ in how the probabilities p_i are defined in each case. Usually in thermodynamics, the probability p_i is the occupation of the state i in the phase space spanned by the degrees of freedom of the system, for instance the position and the velocities of all particles in the system. In signal processing on the other hand, p_i is the probability that a signal takes a given amplitude A_i. Based on thermodynamics principles, less informative macro-state of a system is, the system is said to be in state of equilibrium. This is known in signal processing as a the stationary signal. The information is produced when, for some reason, the system is driven outside its state of equilibrium and in particular when this is followed by structured patterns. In such situation, the system experiences deviations away from its mean energy level until the perturbation eventually dies out. On the other hand, entropy varies away from its maximum level by an amount that is indicative of the degree of structure that has been produced. Because any departure from the equilibrium must have a cause, the decrease of entropy is also indicative of the action of an external force acting on the system and thus of an increase of knowledge about its status. In signal processing terminology, the system has undergone a

transient and therefore has become nonstationary. This analogy with thermodynamics actually unveils the profound links between nonstationary energy fluctuations, entropy and information.

4.4 Spectral negentropy

The description presented in this section is the one elaborated in [10], whereas the spectral negentropy, was described as a contraction for "negative entropy" which initially has been elaborated by Schrödinger in 1940 [11]. From the thermodynamic perspective, entropy evaluates the level of the disorder in a system from the state of equilibrium; in parallel, negentropy measures the inclination of a system to increase its level of organization. In the science of information, entropy describes how much information contains a signal. For a discrete signal $y(k), k = 0, 1, 2, ..., N-1$, where N is the length of the signal. If one denotes its complex envelope as $y(k, f, \Delta f)$ in a defined frequency band Δf, its squared envelope (SE) can be expressed as :

$$e(k, f, \Delta f) = |y(k, f, \Delta f)|^2 \tag{4.2}$$

In Fourier domain, the squared envelope spectrum (SES) can be express as :

$$E(\lambda, f, \Delta f) = \sum_{k=0}^{N-1} e(k, f, \Delta f) e^{-j2\pi k \frac{\lambda}{F_s}} \tag{4.3}$$

λ is the cyclic frequency and F_s is the sampling frequency.

From the SE expressed in the time domain, spectral negentropy can be defined as :

$$\Delta I_e(f, \Delta f) = \sum_{k=0}^{N-1} \left(\frac{e(k, f, \Delta f)^2}{\frac{1}{N} \sum_{k=0}^{N-1} e(k, f, \Delta f)^2} ln \left(\frac{e(k, f, \Delta f)^2}{\frac{1}{N} \sum_{k=0}^{N-1} e(k, f, \Delta f)^2} \right) \right) \tag{4.4}$$

Known as a weighted sum of the spectral kurtosis, the SE negentropy is hardly sensitive to the cyclostationarity of the repetitive transients even though it has a remarkable capability in measuring the impulsiveness of a signal. From observations

made on SES, the spectral negentropy can be rewritten as

$$\Delta I_E(f,\Delta f) = \sum_{k=0}^{N-1}\left(\frac{E(k,f,\Delta f)^2}{\frac{1}{N}\sum_{k=0}^{N-1}E(k,f,\Delta f)^2} ln\left(\frac{E(k,f,\Delta f)^2}{\frac{1}{N}\sum_{k=0}^{N-1}E(k,f,\Delta f)^2}\right)\right) \tag{4.5}$$

The spectral negentropy is essentially a function of frequency-bandwidth/frequency resolution [12–14]. One can consider to calculate spectral negentropy for frequency-frequency resolution plan by partitioning finely and continuously; the infogram which is the image representation of negentropies, can therefore be obtained by cascading and displaying all the spectral negentropies in a single diagram. Based on the Eq. 4.5, one can use this image representation previously defined as infogram to generate its SE. The peak location in the infogram links to the frequency band where the repetitive transients are found.

In the case of bearing fault, the vibration generate often features repetitive impulses. The more impulses the fault induces, the larger is the spectral negentropy.

4.4.1 Signature of the cyclostationarity events

4.4.1.1 The signature of transients

The discussion done in sections 4.3 and 4.4 shows that, a transient corresponds to a departure from equilibrium of a system away from its state of balance or equipartition of energy and maximum entropy. In order to assess the production assocated with it, the measurement of the energy fluctuation in the signal makes it full sense. In a formal manner, let consider a discrete-time signal $x(n), n = 0, ..., L-1$ of length L and let $x(n; f, \Delta f)$ denotes its complex envelope in a frequency band $[f - \Delta/2; f + \Delta f/2]$. The instantaneous flow of energy in the band is returned by the SE

$$\epsilon_x(n; f, \Delta f) = |x(n; f, \Delta f)|^2 \tag{4.6}$$

and its average value is

$$\hat{\epsilon_x}\left(f,\Delta\right) = \langle \epsilon_x\left(n;f,\Delta f\right)\rangle = \frac{1}{L}\sum_{n=0}^{L-1}\epsilon_x\left(n;f,\Delta f\right) \tag{4.7}$$

In thermodynamic analogy, is $k_B T_S$, the system temperature T_S times the Boltzmann constant k_b. Therefore, the strength of the energy fluctuations may be measured by the variance of the energy flow, that is

$$Var\{\epsilon_x(n;f,\Delta f)\} = \langle \epsilon_x(n;f,\Delta f)^2\rangle - \bar{\epsilon_x}(f,\Delta)^2 \tag{4.8}$$

the normalized version of Eq. 4.8, which is the squared coefficient of variation is given by

$$V_x(f;\Delta) = \frac{Var\{\epsilon_x(n;f,\Delta f)\}}{\bar{\epsilon_x}(f,\Delta)^2} = \frac{\langle \epsilon_x(n;f,\Delta f)^2\rangle}{\bar{\epsilon_x}(f,\Delta)^2} = 1 \tag{4.9}$$

So, when the system reached the thermodynamic equilibrium, the principle of maximum entropy [15] started that the energy fluctuations are distributed according to Chi-squared probability low or equivalent complex envelope distributed according to a circular complex Gaussian law. Hence, $V_x(f;\Delta f) = 1$. Alternatively, in the non-equilibrium regime resulting from the presence of a transient, the energy fluctuations are characterized by $V_x(f;\Delta f) > 1$. This is actually the signature to be tracked in the signal as a function of frequency f and frequency resolution Δf

4.4.1.2 Spectral Kurtosis (SK)

In order to localize transients, traditional kurtosis can be applied to real and imaginary part of the Short Time Fourier Transform (STFT). The SK was initially defined as the kurtosis of the frequency components and was compared to the variability in the amplitude of different spectral frequencies. This is an evident indication of how the impulsiveness of a signal varies with the frequency [12]. For zero-mean signal, the normalized variance of the energy flow is closely linked to the SK, this can be expressed as

$$K_x(f;\Delta f) = V_x(f,\Delta f) - 1 = \frac{\langle |x(n;f,\Delta f)|^4 \rangle}{\langle |x(n;f,\Delta f)|^2 \rangle} - 2 \tag{4.10}$$

This equality draws a correspondence between nonstationarity on the one hand which is measured by the variance of the energy flow $V_x(f;\Delta f)$ and nonliearity on the other hand which is measured by the SK $K_x(f,\Delta)$. One can certify that, the basic idea behind the SK is to get a quantity that can ideally take the high values when the signal is transient, and will be zero when the signal is stationary Gaussian.

4.4.1.3 The squared envelope spectrum

As started in the sub-section 4.4.1.2, SK is ideally suited in the detection of a single or a few transients, but not necessarily to the situation where their rate of the repetition is high. If we consider as an illustration a series of transients with a relaxation time $1/\Delta f$ and mean rate of occurrence equal to p per second. Then it can be shown that the SK is

$$K_x(f;\Delta f) \simeq \frac{\Delta f}{p} \tag{4.11}$$

The Eq. 4.11 shows that SK falls to zero as p increases. Knowing that repetitive faults are not periodic in general, but rather cyclostationary, which means their energy flow instead of their waveform is periodic. The Discrete Fourier Transform (DFT) of the energy flow is the Squared Envelope Spectrum (SES). This evidences a spectrum of harmonics equispaced by the fault frequency.

$$E_x(\alpha;f,\Delta) = \sum_{n=0}^{L-1} \epsilon_x(n;f,\Delta) e^{-j2\pi\frac{\alpha n}{F_s}} = \sum_{i\in\mathbb{Z}} E_{x,i}(f,\Delta f)\delta(\alpha - i\alpha_0) \tag{4.12}$$

The description above, paws the way for cyclostationarity techniques based on the exploitation of energy distribution $E_x(\alpha;f,\Delta)$, where δ is the discrete Dirac symbol.

4.5 Kurtogram

In the literature, most of the authors including Antoni [12] have agreed on the fact that kurtogram has been drawn from spectral kurtosis. Technically, the use of SK for transients detection in a signal, focuses on measurement of distance between random and Gaussian process. Following the calculation for each frequency component of the signal, one can obtain the SK value. More than indicate transient components, SK can also track the locations of those components in the frequency domain.

From the Wold-Gramer [16] decomposition of a given non-stationary signal, one can define $y(t)$ as the response of system with time varying impulse response $h(t, s)$ exited by signal $x(t)$. The signal $y(t)$ can be written as :

$$y(t) = \int_{-\infty}^{+\infty} e^{j2\pi ft} H(t, f) dX(f) \tag{4.13}$$

$H(t, f)$ is the complex envelope of the considered signal, and $dX(f)$ is the spectral process in $x(t)$ investigation. For a given frequency, the calculation of the SK values is implemented using STFT of the analyzed signal. It can be expressed by :

$$F_y(\tau, f) = \int_{-\infty}^{+\infty} y(t) w^*(t - \tau) e^{j2\pi ft} dt \tag{4.14}$$

Even if kurtogram outcomes are satisfactory, the kurtosis-based estimator tends to point out misleading frequency [16]. One of the solution to overcome this is the synchronous averaging, but this as well may cause loss of information.

4.6 Spectral correlation

In condition monitoring, the spectral correlation analysis of cyclostationary signal such vibrations plays an important role in any situation for which the cyclostationarity autocorrelation function plays a central role, because the latter is completely characterized by the former [17]. Analytically, the cyclostationarity autocorrelation with period T_0 is defined for a time series $x(t)$ by the average such that :

$$\widehat{R}_x(t,\tau) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(t+nT_o+\tau/2).x(t_nT_o-\tau/2) \tag{4.15}$$

$$= \widehat{R}_x(t+T_0,\tau) \tag{4.16}$$

The periodic function obtain in the Eq. 4.16 can be expanded into Fourier series as :

$$\widehat{R}_x^{\alpha}(t,\tau) = \sum_{\alpha} \widehat{R}_x^{\alpha}(\tau) e^{-i2\pi\alpha t} dt \tag{4.17}$$

For α ranging over all integer multiples of the fundamental frequency of periodicity $1/T_o$. If the we express the Fourier Transform of the situation described above as :

$$\widehat{S}_x^{\alpha}(f) = \lim_{\Delta f\to 0} \lim_{\Delta t\to\infty} \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} \Delta f X_{1/\Delta f}(t, f+\alpha/2).X^*_{1/\Delta f}(t, f-\alpha/2)dt \tag{4.18}$$

where $X_{1/\delta f}(t, f-\alpha/2)$ is the complex envelope of the narrow band-pass component of x(t) with center frequency $\nu = f-\alpha/2)$ and approximate bandwidth δf.

The Eq. 4.18 represents the limit, as spectral resolution becomes infinitesimal at frequencies $(f+\alpha/2)$ and $(f-\alpha/2)$. A deep elaboration of theory related to the spectral correlation of stationary time series can be found in [18, 19]. This paper take advantage of the fact that the spectral correlation analysis of cyclostationary time signal can point out hidden frequencies holding useful information crucial for mechanical fault detection and diagnosis.

4.7 Spectrogram

The spectrogram is known as one of the most common time-frequency analysis map. The core idea is to illustrate how the density of the signal varies with time. Thanks to this technique, one can select a range of frequency where the damped impulses from local faults occur. A spectrogram may be calculated from a sequence

of band-bass filters. This approach is pretty old and obsolete in signal processing area [20]. Most of scientific works tend to use the STFT by breaking up time into intervals and calculating the power spectrum for each part. The window location is slided through the entire data to obtain STFT coefficients. This can be represented mathematically as :

$$X(n,\omega) = \sum_{m=-\infty}^{\infty} x(m)\omega(n-m)e^{-j\omega n} \tag{4.19}$$

Where $X(n)\omega(n-m)$ is a short-time part of the input signal x(n) at time n. In addition, a discrete STFT is defined as :

$$X(n,k) = X(n,\omega)|_{\omega=2\pi k/N)} \tag{4.20}$$

where N shows the number of discrete frequencies. One can then define the spectrogram in logarithmic scale as[14] :

$$S(n,k) = log|X(x,k)|^2 \tag{4.21}$$

The main drawbacks of using this approach are [21], the prior knowledge of the number of intervals into which the signal has to be divided, complex and time consuming calculation, frequency challenging interpretation of the result color map.

4.8 Protrugram

In the late 2009 [22], Barszcz T et al. proposed protrugram as another tool for optimal center frequency selection, as a complement to other techniques such those aforementioned. The approach shows the kurtosis values of the spectral amplitudes of narrowed band envelope signal, which thanks to the Hilbert transform is calculated. Basically, the protogram algorithm proposes the shifting of the selected bandwidth by a defined step and the narrowband envelope spectrum is calculated again. The kurtosis from all spectral amplitudes is calculated and store in a vector as a single scalar value at each center frequency calculation step. Obtain kurtosis values are then plotted as a function of the center frequency. The known drawback of this approach

is the negative influence of harmonic components on the final estimator's value. The protrugram is computed based on the envelope spectrum. It can be written in theory as [23] :

$$\bar{k}_n\left[\bar{x}_{wna}\right] = \frac{\sum_{k=0}^{N-1} |\bar{x}_{wna}[k] - \langle \bar{x}_{wna} \rangle|^4}{\left(\sum_{k=0}^{N-1} |\bar{x}_{wna}[k] - \langle \bar{x}_{wna} \rangle|^2\right)^2} - 2 \tag{4.22}$$

Where $n = 0, 1, 2, ..., [\frac{F_s - M}{2a}]$, $\bar{x}_{wna}$ is the Fourier sequence of the signal. Often, results of SK are used as output of a series of band-pass filters covering a wide range of frequency.

4.9 Cyclostationarity in practice : analysis using experimental data

This section primarily establish a comparison between methods presented in the current chapter, based on technical interest beside each approach. On the other hand, the analysis done helped in establishing the benefits of each technique. The analysis proposes to display frequency versus frequency resolution which is known as Kurtogram. Technically, the kurtogram indicates a strong impulsivity in the corresponding frequency band. From wavelets to tree duel wavelets, enormous research work has been carried out to find approaches for a reliable implementation of the kurtogram to improve its efficiency.

It is also reported in the literature that the kurtogram is unable to recognize whether a series of transients is repetitive or not [24] despite its performance, because the kurtosis value decreases as the transient repetition rate increases. Furthermore, this elaboration crystallizes the analysis on spectral negentropy. The negentropy presented here analyzes the quantities when they are displayed as a function of frequency and frequency resolution as elaborated in the section 4.4 and 4.5, this is called the infogram. Finally this section demonstrate the visual representation of the spectrum known as spectrogram [25] presented earlier in the section 4.7. The proposed approach displays vectors of cyclical frequencies in terms of the sample rate.

4.9.1 Test setup overview

A brief schematic illustration of the measurement test bench helped in data collection is described in the Fig.4.1. The used motor frequency is 48.7 Hz with a speed of 2855 rpm and a sampling frequency of 50 000 Hz.

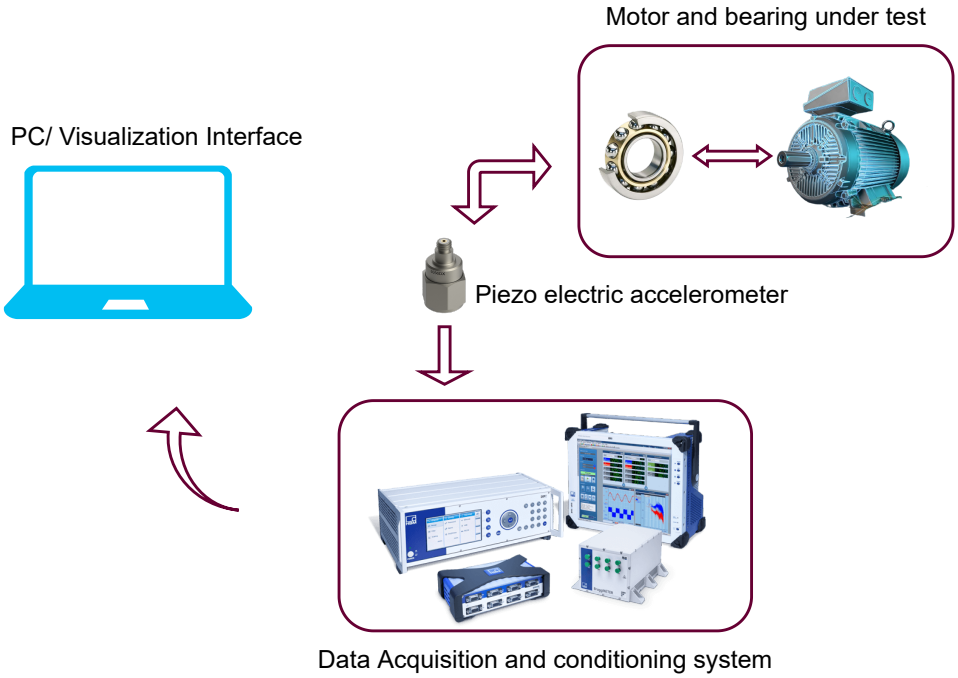


Figure 4.1: Schematic representation of the used test bench

The accelorometer with specifications shows in the table 4.1 is connected to the outer ring of the motor end bearing of type $NU206$.

Three parameters were alternatively modified : the load (L), the speed (SP) and the fault size (FS) in the bearing inner race. Three loads (200 N,500 N, and 700 N) were combined alternatively during the experimentation. Four main situations were analyzed, the first when the bearing is flawless (fault size = $0.00mm^2$), three cases with faulty bearing (fault size = $0.192mm^2$, $1.50mm^2$ and $3.00mm^2$). Faults size were created using an electric pen. The table. 4.2 displays different characteristic frequency for the bearing under test.

In the table 4.2, N is the total number of rolling elements, D_p is the primary diameter, d is the diameter of the rolling element, α is the contact angle and F_r is the

Table 4.1: Used accelerometer specification [24]

Performance	SI
Sensitivity (±5%)	$10.2mV/(m/s^2)$
Measurement Range	$\pm 490m/s^2$
Frequency Range(± 10 %) 300k	0.42 to 5000 Hz
Resonant Frequency	20 kHz
Broadband Resolution(1 to 10k Hz)	$490\mu m/s^2$
Non-linearity	±1%
Transverse Sensitivity	≤ 5%

Table 4.2: Corresponding characteristic frequency for each used speed.

Motor speed [rpm]	2000	1500	1000	Mathematical representation
BSF [Hz]	13.4	10.1	6.7	$F_{rol} = (\frac{D_p}{2d})(1 - \frac{(dcos\alpha)^2}{(D_p}})^2)F_r$
BDF [Hz]	165.8	124.3	82.9	$F_{Housing} = (\frac{1}{2})(1 - \epsilon\frac{dcos\alpha}{D_p})F_r$
BPFO [Hz]	174.7	131.0	84.7	$F_{be} = (\frac{N}{2})(1 - \frac{dcos\alpha)^2}{D_p})F_r$
BPFI [Hz]	258.6	194.0	129.3	$F_{bi} = (\frac{N}{2})(1 + \frac{dcos\alpha}{D_p})F_r$

rotation frequency based on the assumption that the outer ring is fixed. It's important to note that $\epsilon = 1$ if the outer ring is fixed and -1 in case of the inner ring.

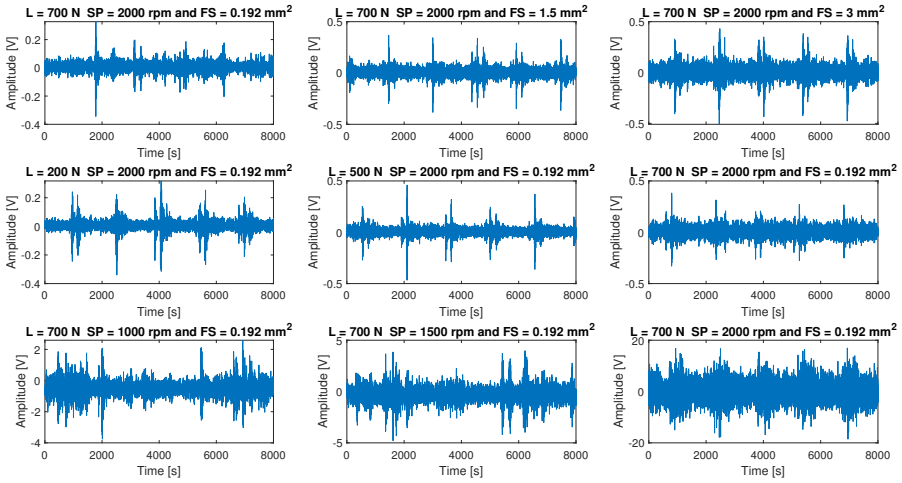


Figure 4.2: Analyzed cases : From the top row to the bottom, raw signal depending on combined parameters.

4.9.2 Results and analysis

In Fig 4.3, useful information related to the probable faults are clearly present in contrast with the results presented in the Fig 4.4, whereas fault presence can be seen as well, but almost submerged by noise. The common and interesting point for both cases (Infogram and Kurtogram)is that using Hilbert transform for the demodulation significantly improves the recognition of the fault characteristic frequency, even if in certain situations that might not be easily noticeable.

On the other hand, spectral correlation is used with the length signal's window and the maximum cyclic frequency with $\alpha = 620$ Hz. Results based on the case where the fault size in inner ring is quantified to be $0.192mm^2$ and $3mm^2$ are presented in the Fig 4.5. A quick observation shows different non-zero values at the cyclic frequencies of the BPFI and obviously its multiples.

In theory, the interest beside computing the cyclic coherence in condition monitoring in addition to other relevant techniques such those presented above is that,

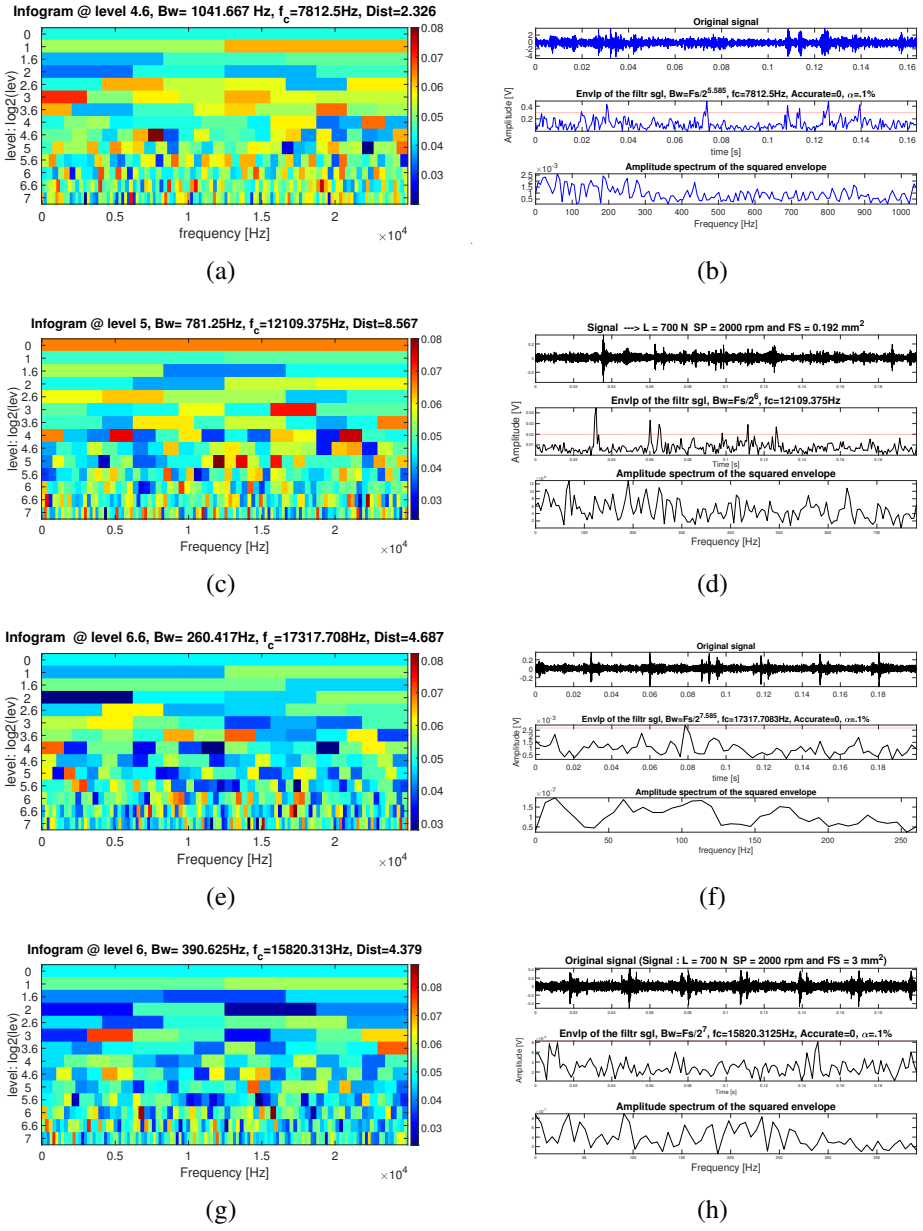


Figure 4.3: From the left row to the right : the Infograms and their envelope representation with fault size respectively :(a)and (b) $\Rightarrow 0mm^2$, (c) and (d) $\Rightarrow 0.192mm^2$,(e) and (f) $\Rightarrow 1.50mm^2$, (g) and (h) $\Rightarrow 3mm^2$

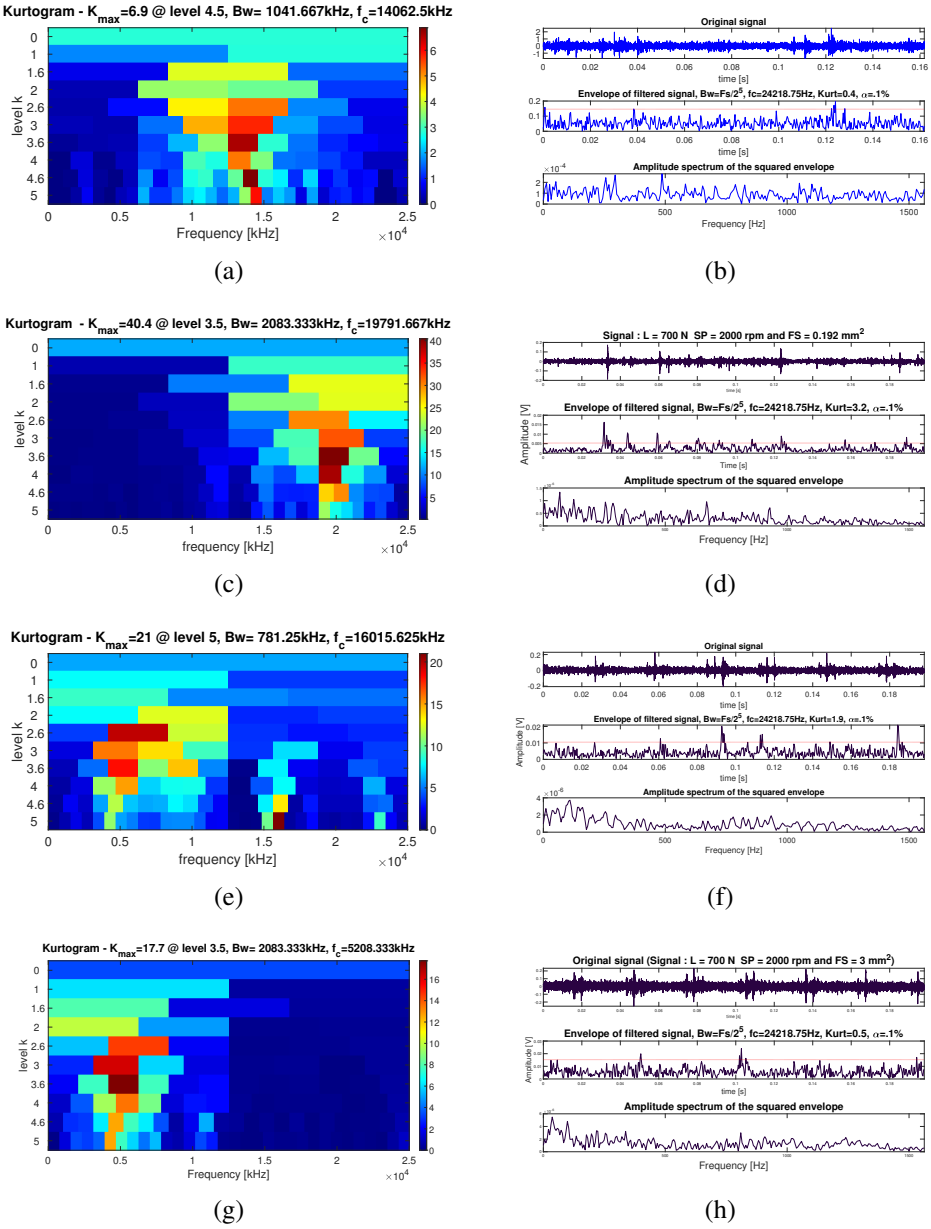


Figure 4.4: From the left row to the right : the Kurtograms and their envelope representation with fault size respectively :(a)and (b) $\Rightarrow 0mm^2$, (c) and (d) $\Rightarrow 0.192mm^2$,(e) and (f) $\Rightarrow 1.50mm^2$, (g) and (h) $\Rightarrow 3mm^2$

cyclic frequencies coincide with the expected characteristic fault frequency. In practice, the fault frequency estimation based mainly on the rotation speed and the angle of contact as presented in the table 4.2, can be approximated and the cyclic coherence will be calculated in a narrow cyclic frequency band around the expected fault frequency, so that its greatest value will be considered.

The presented techniques, whether it is spectral negentropy, spectral kurtosis or spectral correlation, have natural idea, the diagnostic based on the detection of signatures. In order to those strategy to be reliable and valid to be implemented in an automated monitoring system,their sensitivity have to be assessed. Several other strategies have been proposed in the literature with the purpose to detect the cyclostationary signature such as bearings faults [26, 27]. Some of them are based on a direct and in-situ visual inspection of the magnitude of the cyclic power spectral density, in certain situations displayed as images over the frequency domain. The approach developed in this paper involves mainly the cyclic coherence, just one cyclic frequency is taken into account and the complete spectral fault signature, i.e fundamental, harmonics and side-bands is not exploited. Some difficulties might rise when the question related to fault type discrimination is asked. If so, the cyclic frequency resolution $\Delta\alpha$ is set too coarse to correctly resolve between different characteristic frequencies.

The observation made in Fig 4.2 clearly shows that, the temporal analysis of the vibration is not likely to help to provide a diagnosis, because, at first sight, all signals look stationary. In Fig. 4.3 and 4.4, a plot of kurtogram and infogram, as well as their corresponding envelopes, are presented for the case where the fault size is around $0.192mm^2$ and $3mm^2$ respectively. It is obvious to see that both kurtogram and infogram show some interesting information. One can observe high values of the kurtogram and infograms in several regions on the figure. It can be said that the kurtogram displays an evident presence of impulsive behavior in the signal. On the other hand, the SES infogram confirms the corresponding events cyclostastionarity extension in several bands. One can then look at different regions that are likely to contain repetitive transients. Various bands to likely to contain useful information on the fault (as can be seen actually on the Fig 4.3 and 4.4, left column) are in the table 4.3.

According to the table 4.2, the BPFI is expected at a cyclic frequency of 258.6 Hz as the fault progresses. The vibration energy distributed around 3.16 kHz is modulated with the peak at a frequency around 253 Hz, which slightly deviates from the calculated characteristic frequency. Several research results such Qiu et al [28]

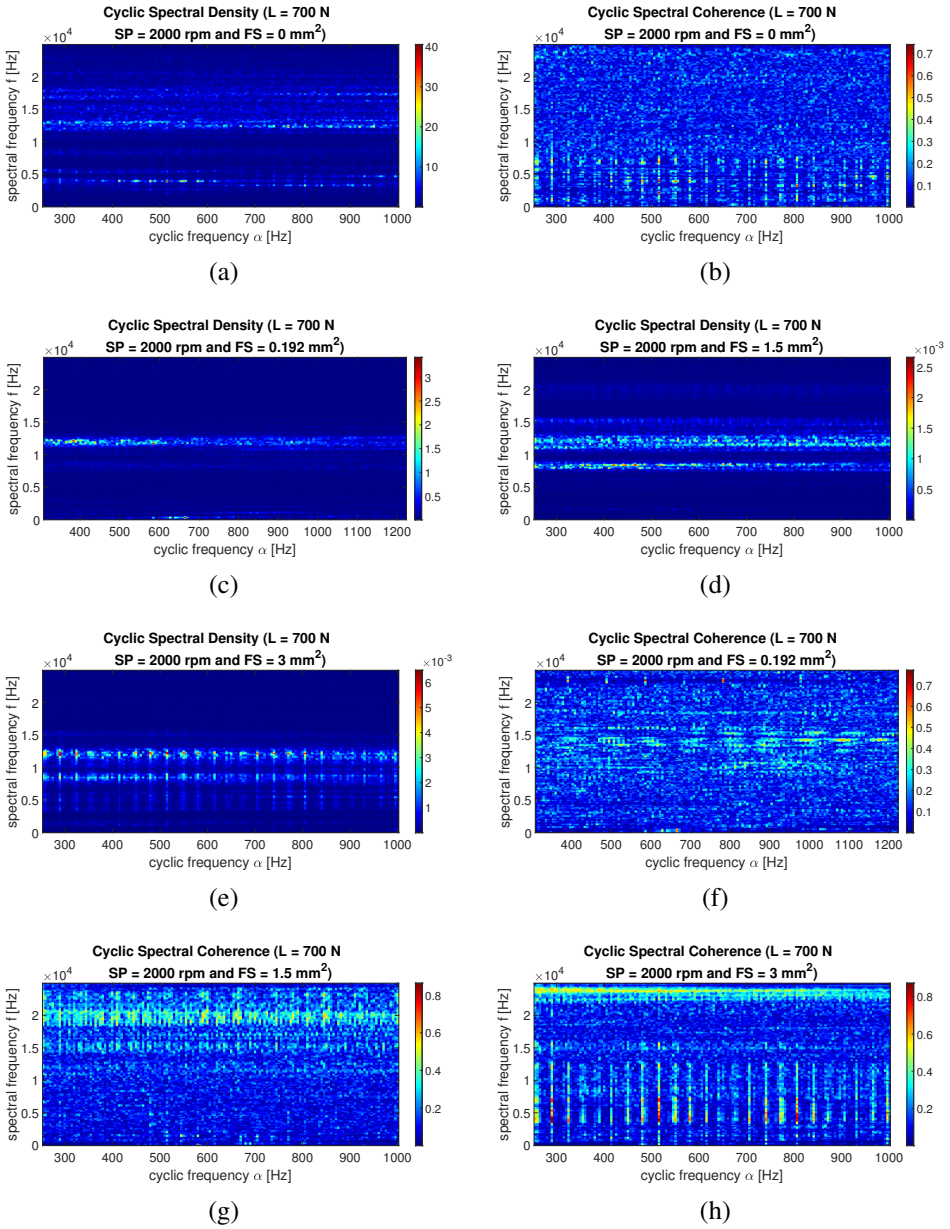


Figure 4.5: Cyclic Spectral Density, Cyclic Spectral Coherence and Squared Magnitude Cyclic Spectral Coherence based on fault size

Table 4.3: Frequency band likely to contain meaningful information on the faults

[15.66 - 16.18] kHz	[15.66 - 16.64] kHz	[3.16 - 6.22] kHz
[2.94 - 4.66] kHz	[4.19 -5.18] kHz	[15.76 - 16.54] kHz

or Ali Mahvash & Aouni A. Lakis [29] reported similar observations. Ideally, the frequency band would be obtained such that the cyclic coherence values are maximized, i.e., signal to noise ratio. Fortunately, the solutions provided by computing the spectral kurtosis or the negentropy using infograms automatically help in finding the best demodulation band without prior knowledge of the signal. In this paper, the cyclic coherences were calculated with parameters presented in the table 4.4

Table 4.4: Used parameter in computing cyclic coherence

Sampling Frequency	50×10^3 Hz
Record length	50×10^3
Window type	Hanning
Overlap	2/3
Window length	256 samples
Frequency resolution (Δf)	200 Hz
Frequency resolution ($\Delta\alpha$)	51 Hz
Variance reduction factor (δ)	0.00125

One of the fundamental assumptions in bearing fault diagnosis is that healthy bearing is not supposed to produce significant cyclostationarity. The faulty bearing, on the other hand, has a cyclic coherence technically above the statistical threshold. The Fig 4.5 presents illustrative cyclic coherence(squared-magnitude) for the three cases presented in the Eq 4.18

In the Fig 4.6, the cyclic coherence exhibits a fundamental spectral line at $\alpha = 253.16\ Hz$ (which is close to 258 Hz as shown in the table 4.2) as well as it is harmonics, this is the clear proof of the apparition of a BPFI. The motor speed significantly modulates the faulty signal due to the periodic passing of the fault in the load zone; this can be explained by symptomatic sidebands around the fault frequencies. It can be noted that, when reading as a function of f, the cyclic coherences, such bands are located mainly within [15.66 – 16.18] kHz and [15.66 – 16.64] kHz.

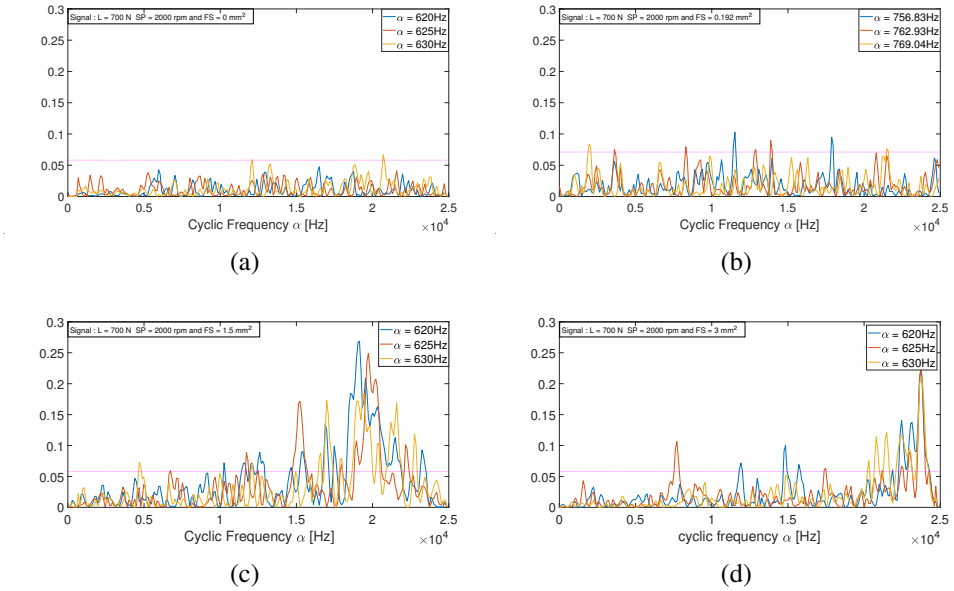


Figure 4.6: The estimated cyclic coherence for signal with the bearing loaded with 700 N a shaft speed of 2000 rpm and fault sizes of: (a) 0 mm^2, (b) 0.192 mm^2, (c)1.50 mm^2 and (d)3 mm^2 respectively.

4.10 Conclusion

Commonly used advanced frequency representation and analysis were discussed in this chapter. It was demonstrated that these techniques could tremendously extend the precision in the selection and interpretation of the most informative frequency band. The usefulness of some relevant and advanced frequency representation used in the analysis of signal in condition monitoring was presented. This was motivated by observations made on traditional time-frequency analysis which concluded that, these techniques were mainly analysis tool as opposed to processing tools. In addition, traditional methods do not offer a versatile methodology that can be apply to all mechanical signals in all situations. The validation was done on experimental data using enhanced methods such as kurtogram, spectral negentropy, or spectral correlation analysis. Ball Pass Frequency Inner Race (BPFI) of bearings with characteristics presented in section 4.9 was particularly investigated. The bearings are housed in a casing allowing the shaft to rotate while driven by a variable speed electric motor. The study outcomes demonstrate the capability of kurtogram, the spectral correlation, and spectral negentropy. The later is fundamentally a concept borrowed from the field of thermodynamics to identify nonstationary events of different nature by locating more precisely and specifically demodulation frequencies, which might be of considerable use for several fault detection and diagnosis scenarios. It was found that high values appear in a set of frequency bands. The processing effort grows then proportionally with the number of frequency bands where the signal has to be demodulated; easily, one can use them to compute the envelope spectrum. In the case of fault like BPFI, it was shown that the negentropy gives more information even for a shallow size of fault and allows the localization of meaningful frequency bands. The analysis done using spectral correlation, on the other hand, suggested that the approach may not always be indicative of the severity of the bearing faults. However, it can be used as a monitoring tool because its value always reads higher for a faulty bearing than for a healthy one.

References

[1] C Yiakopoulos, M Koutsoudaki, and K Gryllias. Antoniadis, i. improving the performance of univariate control charts for abnormal detection and classification, 2017.

[2] P Kruczek, J Obuchowski, A Wylomanska, and R Zimroz. Multiple local damage detection in gearbox by novel coherent bi-frequency map and its spatial, cycle oriented enhancement, 2019.

[3] F Combet and L Gelman. Optimal filtering of gear signals for early damage detection based on the spectral kurtosis, 2009.

[4] Z.-H Lai and Y.-G Leng. Weak-signal detection based on the stochastic resonance of bistable duffing oscillator and its application in incipient fault diagnosis, 2016.

[5] N Nikolaou and I Antoniadis. Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets, 2002.

[6] J Obuchowski, A Wyłomańska, and R Zimroz. Selection of informative frequency band in local damage detection in rotating machinery, 2014.

[7] X Ma and C Nikias. Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics. *IEEE Transactions on Signal Processing*, 44(11):2669–2687, 1996.

[8] J Antoni. Cyclostationarity by examples (2009) mechanical systems and signal processing.

[9] C.G. Chakrabarti and I.Chakrabarty. Boltzmann entropy : Probability and information. *Journal of Physics*, 52:559 – 564, 2007.

[10] M Ugwiri, M Carratu, V Paciello, and C Liguori. Spectral negentropy and kurtogram performance comparison for bearing fault diagnosis (2020) art.

[11] Y Xu, J Chen, C Ma, K Zhang, and J Cao. Negentropy spectrum decomposition and its application in compound fault diagnosis of rolling bearing. *Entropy*, 21(5), 2019.

[12] S Udmale and S Singh. Application of spectral kurtosis and improved extreme learning machine for bearing fault classification. *IEEE Transactions on Instrumentation and Measurement*, 68(11):4222–4233, 2019.

[13] L Song, H Wang, and P Chen. Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery. *IEEE Transactions on Instrumentation and Measurement*, 67(8):1887–1899, 2018.

[14] M Carratù, A Pietrosanto, P Sommella, and V Paciello. Semi-active suspension system for motorcycles: From the idea to the industrial product, 2018.

[15] E.T. Jaynes. Information theory and statistical mechanics. *Physics Review*, 106:620 – 630, May 15, 1957.

[16] M Ugwiri, I Mpia, and A Lay-Ekuakille. Vibrations for fault detection in electric machines (2020) ieee instrumentation and measurement magazine.

[17] J Li, Q Yu, X Wang, and Y Zhang. An enhanced rolling bearing fault detection method combining sparse code shrinkage denoising with fast spectral correlation. *ISA Transactions*, 102:335–346, 2020.

[18] L Dhamande and M Chaudhari. Compound gear-bearing fault feature extraction using statistical features based on time-frequency method. *Measurement: Journal of the International Measurement Confederation*, 125:63–77, 2018.

[19] Structural characterization of locally optimum detectors in terms of locally optimum estimators and correlators. *IEEE Trans. Inform. Theory*, pages 924–932, 11 1982.

[20] X Song, X Li, W.-G Zhang, and W Zhou. The new measurement algorithm of the engine speed base on the basic frequency of vibration signal (2010) 5, art.

[21] G Manhertz and A Bereczky. Stft spectrogram based hybrid evaluation method for rotating machine transient vibration analysis (2021) 154, art.

[22] A Klepka, T Barszcz, and A Jabłoński. Comparison of advanced signal analysis techniques for bearing fault detection, 2009.

[23] Y Wang, J Xiang, R Markert, and M Liang. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications, 2016.

[24] M Ugwiri, M Carratu, A Pietrosanto, V Paciello, and A Lay-Ekuakille. Vibrations measurement and current signatures for fault detection in asynchronous motor (2020) i2mtc 2020 - international instrumentation and measurement technology conference, proceedings, art.

[25] P Gangsar and R Tiwari. A support vector machine based fault diagnostics of induction motors for practical situation of multi-sensor limited data case. *Measurement: Journal of the International Measurement Confederation*, 135:694–711, 2019.

[26] A Jablonski, Z Dworakowski, K Dziedziech, and F Chaari. Vibration-based diagnostics of epicyclic gearboxes - from classical to soft-computing methods. *Measurement: Journal of the International Measurement Confederation*, 147(106811), 2019.

[27] C Capdessus, M Sidahmed, and J Lacoume. Cyclostationary processes: application in gear faults early diagnosis (2000) mechanical systems and signal processing.

[28] H Qiu, J Lee, Lin, and J. Wavelet filter-based weak signature detection method and its application on roller bearing prognostics. *Journal of Sound and Vibration*, 289(4-5):1066–1090, 2006.

[29] Ali Mahvash, Aouni, and Lakis A. Application of cyclic spectral analysis in diagnosis of bearing faults in complex machinery. *Tribology Transactions*, 58(6):1151–1158, 2015.

Chapter 5

Detection and diagnosis based on data mining: an application of machine learning techniques

The chapter two provided non learning techniques for the pre-processing and for faults signature extraction. The focus in this chapter is on machine learning methods that can be used for detection and diagnosis. The chapter discusses the critical issue of selecting relevant indicators by combining the process with the cascade method introduced earlier in this work in chapter 3, and address the problem of evaluating the performance of a diagnostic system.

5.1 Introduction

As described in precedent chapters, using vibration signature is a hot topic in condition monitoring. This chapter proposed to take advantage of the cascade method, particularly the Ensemble Empirical Mode Decomposition (EEMD) in order to enhance the sensitivity of scalar indicators such as Kurtosis and Crest Factor. The improved indicators are then selected using Principal Component Analysis technique. A profound discussion on the use of EEMD prior to classification by means of algorithms such as SVM and k-NN using the is elaborated. Each technique's performance is evaluated relying how good it achieves in both fault detection and classification.

5.2 Data mining and machine learning : Cursory definition

In a larger context, Data mining techniques and Machine Learning (ML) algorithms are only part of the greater process to solve a particular problem. This section does not provide a deeper technical detail on these subjects, the ready is encouraged to exploit reference [1] and [2] for deeper technical aspects.

5.2.1 Data mining

Historically, data mining is not a new invention that came with digital age [3]. The concept is actually well known and has been around for century. However, it came into greater public in the early 20 century. One of the first data mining instances occurred in 1936, when Alan Turning introduced the idea of a universal machine that could perform computations similar to those of modern age [4]. Data mining combines statistics and artificial intelligence to analyze large data sets to discover useful information [5]. Data mining techniques basically extract hidden patterns from a a collection of data, the knowledge extracted from the process has to be interpretable and directly exploitable in an automatic process of analysis or decision process. Today, data mining is used in many fields of business and research, including sales and marketing, product development, healthcare, education, fault detection and diagnosis, etc... Some of the common techniques used in data mining are [6] :

- **Association rules**: An association rule is a rule-based method for finding relationships between variables in a given dataset. These methods are frequently used for market basket analysis, allowing companies to better understand relationships between different products. Understanding consumption habits of customers enables businesses to develop better cross-selling strategies and recommendation engines.
- **Neural networks**: Primarily leveraged for deep learning algorithms, neural networks process training data by mimicking the inter-connectivity of the human brain through layers of nodes. Each node is made up of inputs, weights, a bias (or threshold), and an output. If that output value exceeds a given threshold, it “fires” or activates the node, passing data to the next layer in the network. Neural networks learn this mapping function through supervised learning, adjusting based on the loss function through the process of gradient descent. When the cost function is at or near zero, we can be confident in the model’s accuracy to yield the correct answer.

- **Decision tree**: This data mining technique uses classification or regression methods to classify or predict potential outcomes based on a set of decisions. As the name suggests, it uses a tree-like visualization to represent the potential outcomes of these decisions.

- **K-nearest neighbor (KNN)**: K-nearest neighbor, also known as the KNN algorithm, is a non-parametric algorithm that classifies data points based on their proximity and association to other available data. This algorithm assumes that similar data points can be found near each other. As a result, it seeks to calculate the distance between data points, usually through Euclidean distance, and then it assigns a category based on the most frequent category or average.

5.2.2 Machine learning

A easy way to look at Machine Learning (ML) is from computer science view. Often, scientists classify ML as a subfield of computer science, and its concerned of designing algorithms that, in order to be useful rely on a collection of data of some phenomenon. These data can be from nature, be handcrafted by humans, or generated by another algorithm. ML can also be defined as the process of solving a practical problem by :

- Collecting a dataset

- Algorithmically training a statistical model based on that dataset

The aforementioned statistical model is assumed to be used somehow to solve a practical problem.

5.3 The learning strategy adopted in this thesis

The strategy adopted in this thesis uses the cascade method, namely the Ensemble Empirical Mode Decomposition (EEMD) discussed in chapter 3. As presented in the Fig.5.1, prior to apply the classification algorithms, the EEMD is applied on the raw vibration signal to obtain the Intrinsic Mode Functions (IMFs). Based on the correlation values, the best IMFs, which will be referred as representatives are selected and used for feature extraction/feature selection. The PCA is used for dimensionality reduction before applying SVM and k-NN which are used in this paper.

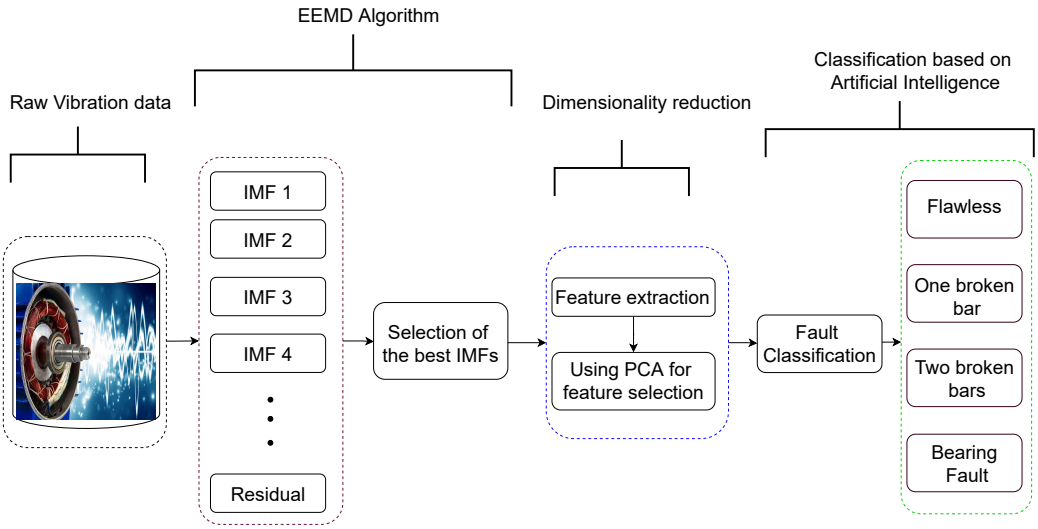


Figure 5.1: The proposed and adopted approach : from raw vibration to fault classifiction

5.4 The used framework

The machine learning framework proposed in this paper for fault classification is presented in the Fig.5.2.

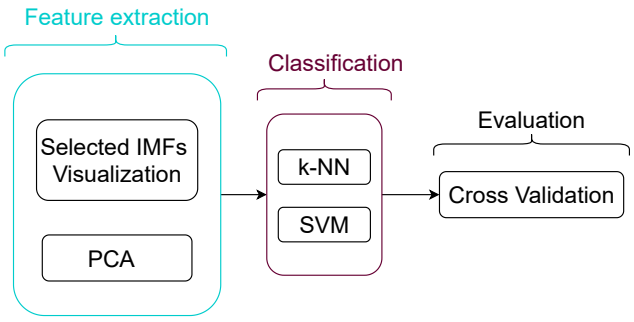


Figure 5.2: The proposed Machine Learning Framework

The framework suggests to select futures using PCA after implying EEMD, the classify extracted future by means of k-NN and SVM. The evaluation is done using cross validation

5.4.1 Feature extraction

The idea beside this stage is the reduction of data by extracting useful subset of data that present clear characteristic of the machine dynamic. In this work, the Principal Component Analysis (PCA) [7] is used. The PCA is chosen for its outstanding dimentionality reduction capabilities. To apply the PCA, dataset of size $n \times 1$ is split into m chunks, such that $p \times m = n$, where p is the length of each chunk. The resulting data is then organized into a $p \times m$ matrix, denoted as M. The standard step follow in this paper can be summarized in the scheme of the Fig.5.3

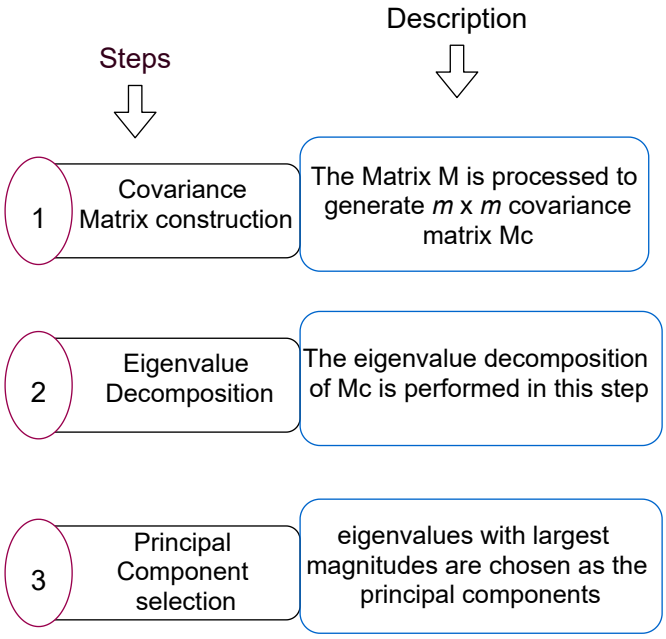


Figure 5.3: Steps follwed for feature extraction using PCA

5.4.1.1 Classification algorithms

Two classification algorithms are exploited in this paper, the k Nearest Neighbors (k-NN) and the Support Vector Machine (SVM).

5.4.1.1.1 k-NN The k-NN algorithm behaves like a majority vote classifier [8]. By specifying some metric as a typical Euclidean distance measurement. Data point displayed in feature space is assigned to a particular class according to the majority

class of its k nearest neighbors. A class i is assigned to the data sample according to the probability equation (to be written).

5.4.1.1.2 Support Vector Machine (SVM) It happens often in practice to hardly obtain enough samples on the faults [9]. SVM can solve this issue by using as small sample as possible with high accuracy and good generalization capability. This paper proposes to use the multi-class classification method. The approach constructs k SVM model where k represents the number of classes.The *nth* SVM is trained with all of examples in the *nth* class with positive labels, and the remaining examples are labeled as negative. For a given τ training data $(x_1, y_1), ..., (x_\tau, y_\tau)$, where $x_n \in R^N, n = 1..., \tau$ and $y_n \in (1, ..., k)$ is the class of x_n, the *nth* SVM presented in this paper solve the minimization problem started as :

minimize :

$$\frac{1}{2} \parallel \mathbf{w}^n \parallel^2 + C \sum_{n=1}^{\tau} \zeta_j^n (\mathbf{w}^n)^T \tag{5.1}$$

subject to :

$$case1 : y = n, (\mathbf{w}^n)^T \Phi(\mathbf{x}_j) + b^n \geq 1 - \zeta_j^n \tag{5.2}$$

$$case2 : y \neq n, (\mathbf{w}^n)^T \Phi(\mathbf{x}_j) + b^n \leq 1 - \zeta_j^n \tag{5.3}$$

With $\zeta_j^n \geq 0$, $j = 1, ...\tau$

$\mathbf{x}_n$, the training data is mapped to a higher-dimensional space by the function Φ and C is the penalty parameter. The minimization of the Eq. 5.3 implies to maximize $\frac{2}{\|\mathbf{w}_n\|}$, which is the margin between two groups of data. In the case where the data is not separable, one will notice the presence of the penalty term $C \sum_{n=1}^{\tau} \zeta_{n,n}$, which considerably reduce the number of training errors. In practice, the statistical learning concept presented in this paper considers the training algorithm of an SVM fault classifier for the recorded vibration dataset, and builds a model able to map features where other features are categorized by transparent gab. The mapped categories are separated by a gap which is called hyper plane. The test dataset are also mapped into the same hyper plane and predicted by the trained SVM model. All features lying near the boundary are called support vectors. In other to efficiently determined the loss function a parameter ν discussed in-depth in [10] is utilized. This parameter

makes also easy to control the number of support vectors and training errors. It is a lower bounded on fraction of support vectors and upper bound on fraction of training errors parameter.

5.5 Analysis on experimental data

The bench use for collecting data is the same as the one presented in chapter 3. The first acquisition is made when the motor is running on a no-loaded stage and without fault. The second measurement phase is done when a fault is introduced in the bearing. The defect was achieved using an electric pen, where the cage was partially destroyed. The last measurement stage was done after creating defects successively in two rotor bars. The bleeding was sealed with epoxy to reduce the unbalanced effect caused by material removal. The accelerometer described in chapter 4 section 4.9, was alternatively placed in the axial and radial position under the same operating conditions of the machine. The Fig. 5.4 presents signals for different analyzed scenarios.

The test focuses on 127 operating conditions, where each of them is represented by 4 vibratory signals recorded with the aforementioned sampling frequency (50 000 Hz). The faults were successively simulated in two rotor bars and outer ring of the motor's front bearing. Different level coded as case 0 to case 3 is analyzed, where the index 0 is the flawless level and the 1 and 2 are respectively one broken bar and two broken bars and the 3 is the bearing fault.

5.5.1 Best IMFs selection

Once the decomposition has been achieved successfully, each IMF contains unique characteristic. One of the challenging task in this step is to recognize the IMF which carry the most relevant information on the fault signature. In the experimentation done by Rai et al [11], FFT of IMFs has been proposed, which technically leads to errors since IMFs are selected manually. One of the interesting approach to overcome manual selection proposed by Rai et al, is the mathematical/merit index developed by Ricci and Pennacchi [12] which requires addressing issues related to the mathematical model in prior. Jianhua Yand et al [13] used Spectral Amplification Factor (SAF) and Kurtosis as parameters for selection on a study done for extracting the unknown faults of rolling bearings in fixed effects shape condition. From several similar trial done in this work, first of all in the conditions close to [13, 14], but more

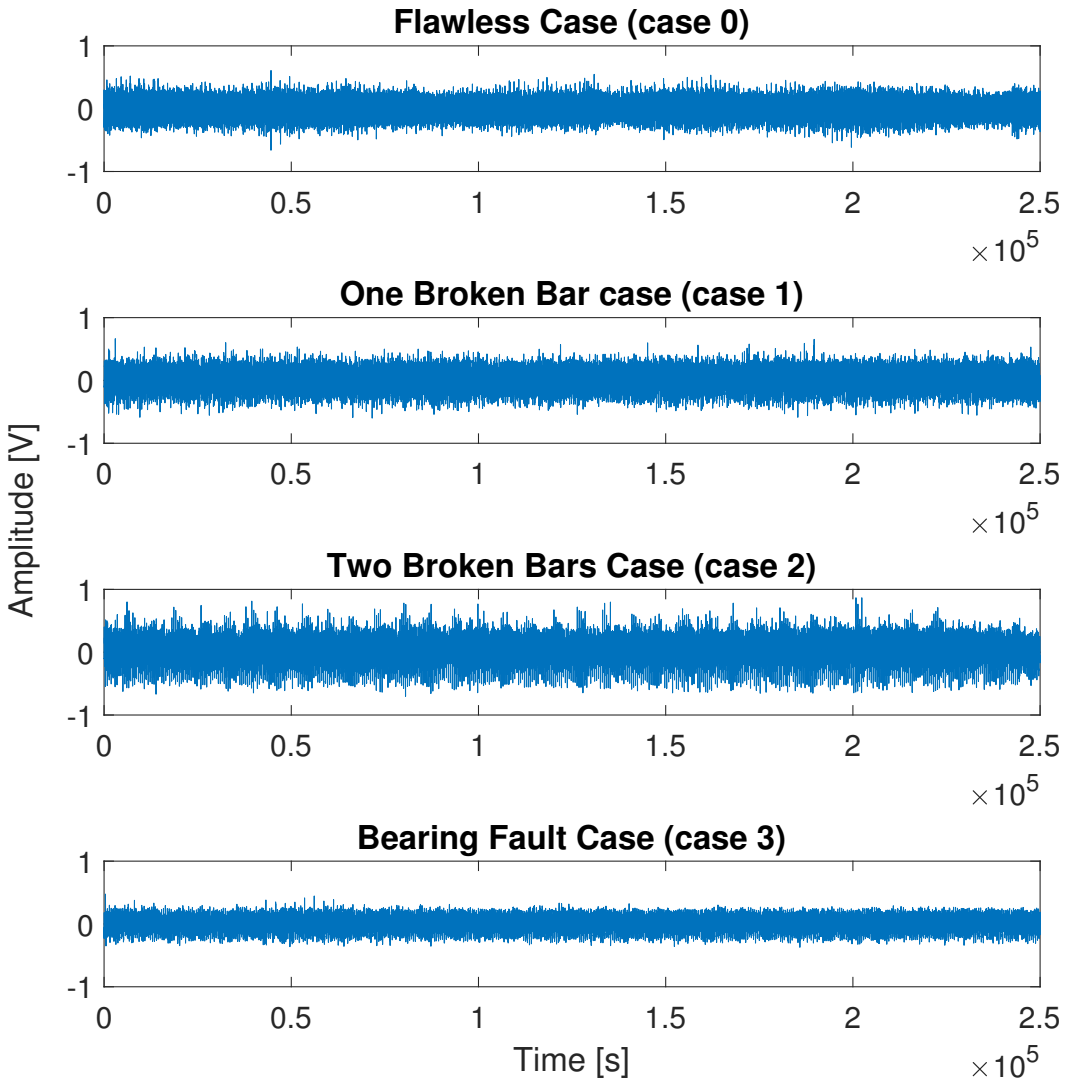


Figure 5.4: Raw signal recorded from different analyzed scenarios.

with known fault approach, the kurtosis was demonstrated to be effective in pointing out the IMF contains non-Gaussian distribution(table 5.1), which is the behavior that most of electric machines subject to faults follow. In table 5.1, the kurtosis value for IMFs of different fault cases investigated in this paper is presented.

Table 5.1: IMFs Kurtosis values

Index	Case 0	Case 1	Case 2	Case 3
Original signal	4.869	2.689	31.312	2.692
IMF 1	6.546	5.347	8.107	8.336
IMF 2	4.597	4.023	3.587	4.000
IMF 3	3.597	3.546	3.587	3.450
IMF 4	3.028	3.974	3.296	4.001
IMF 5	2.701	3.843	3.334	2.568
IMF 6	2.945	2.043	2.254	2.688
IMF 7	2.835	2.044	2.688	2.844
IMF 8	1.98	3.022	1.869	2.303
IMF 9	3.576	6.212	5.831	4.458
IMF 10	3.114	2.524	3.500	3.079
IMF 11	2.526	5.740	2.952	3.923
IMF 12	2.825	3.186	3.148	2.863
IMF 13	2.922	2.583	3.736	2.658
IMF 14	2.482	2.915	3.265	2.212
IMF 15	2.794	2.017	3.265	2.515
IMF 16	1.623	2.153	3.265	2.039
IMF 17	2.298	1.930	2.834	2.013
IMF 18	2.461	1.594	1.474	1.504
IMF 19	1.574	2.682	3.659	1.520
IMF 20	1.696	-	-	2.722
IMF 21	2.103	-	-	-

The kurtosis of each IMF component is calculated separately, as can be seen in the table 5.1. The largest values are found in IMF1, IMF2, IMF9 and IMF14. IMF1 and IMF9 are selected as optimal vibration components (Fig. 5.5) for further analysis.

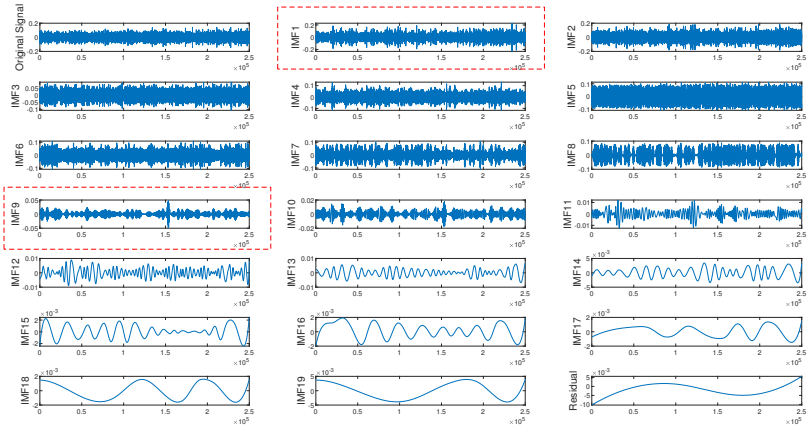


Figure 5.5: Typical EEMD decomposition for the case 2 (the IMFs carrying significant information are surrounded by the dashed rectangle)

5.5.2 Features extraction

Feature extraction as a significant step in all pattern recognition problems, is carefully done in this study. Follow up is the normalization, where each of the computed features is normalized between $[0, 1]$ for each class, except for skewness normalized in the range of $[-1, +1]$. While solving a classification problem, this stage is critical in the discrimination against any bias due to different scales among features. The table5.2 displayed statistical features used for classification in this paper.

5.5.3 Dimensionality reduction

The features selection is one of the core concepts in preventing machine learning's algorithm degradation due to non-relevant features. Good feature selection has an considerable impact on the obtained results and on the performance of the classifier [14]. The approach adopted in this paper is based on the Principal Component Analysis (PCA). The motivation beside using PCA is that, most of features coming from vibration signals when analyzed for condition monitoring tends to be correlated, which drastically compromised redundant information.

Table 5.2: Used features for classification

Feature	Mathematical expression	Description
RMS (root mean square)	$\sqrt{\frac{1}{N}[\sum_{i=1}^{N}(x_i)^2]}$	This can be defined in terms of an integral of the squares of the instantaneous values during a cycle of a continuously varying function
STD (Standard deviation)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^2}$	Is the central value of a discrete set of numbers
Skew (Skewness)	$\frac{\sum_{i=1}^{N}(x_i-\overline{x})^3}{N\sigma^3}$	This quantifies the asymetric behavior of the signal through its Probability Distribution Function (PDF)
Kurt (Kurtosis)	$\frac{\sum_{i=1}^{N}(x_i-\overline{x})^4}{N\sigma^4}$	Quantifies the peak value of the PDF
CF (Crest Factor)	$\frac{max\lvert x_i\rvert}{x_{rms}}$	By definition, CF is the difference between the peak and the average level of a signal, it is used in this work to measure the impact of the contact between rolling elements of bearing with raceways.
Impulse factor (IF)	$\frac{max\lvert x_i\rvert}{\overline{x}_{abs}}$	Evaluates the ratio of absolute peak amplitude to the mean of absolute signal.

5.5.4 Training and testing

5.5.4.1 Using SVM

One of the considerations taken in this paper is that, good classification by using support vector machine (SVM) approach, requires a suitable kernel function selection [14]. The adopted logic used in this paper, for both training and testing is summarized in the Algorithm 1. SVM-based fault classification often uses linear, polynomial or Gaussian RBF kernel functions based on the type of separation needed between classes. The selected kernel function defines the type of separation surface. The extracted and the selected features obtained from the signal presented in the Fig. 5.4 is used for classifying the motor faults using SVM classifier with different aforementioned functions. The kernel functions are evaluated to observe their suitability for the given classification problem. The one-versus-one method is used to train the SVM classifier where the total number of the classes is four. Accuracy for each combination of features in order of relevance are computed. A 5-fold cross-validation scheme is used for SVM classifier evaluation.

5.5.4.2 Using k-NN

In the training stage, subsets of vectors are selected from the original data set as a training set. If one assign to the selected subset τ as name, one can construct a

Algorithm 1: Pseudo code for SVM

Pre-processed data: f^*

1 `// the dataset with variables, time and status`

Ouput **:** Aplication of PCA

2 `// the pca is applied, the prediction or function`

Input **:** Ranked list of variables according to their relevance

3 `/* Finding the optimum values for the tuning parameters of the SVM model` $f \Longleftarrow f^*$ `*/`

4 **while** $f \geq 2$ **do**

5 Fit the SVM with the optimized tuning parameters for the f variables and observations in the dataset

6 $W_f \Longleftarrow$ calculate the weight vector of the SVM

7 $rank.criteria \Longleftarrow (W_{f1}^2, \ldots, W_{pp}^2)$

8 get rid of $min.rank.criteria$ from **data**

9 $rank_f \Longleftarrow min.rank.criteria$

10 $f \Longleftarrow f - 1$

11 **if** $Method = Predict$ **then**

12 Calculate the decision value for each training point `/* The decision value vector will be incorporated in this stage as a variable, then the PCA space is calculated with all` $(f+1)$`, the projected into PCA space for the first tree components */`

13 **else**

14 $Method = Function$

15 calculate the PCA space with all f variables `// Project the SVM decision function into PCA space`

16 **end if**

17 **end while**

18 **return** $(rank_1, rank_2, ..., rank_{f^*})$

pseudo metric and write it as follow :

$$d\left(\overrightarrow{x}_i, \overrightarrow{x}_j\right) = \left(\overrightarrow{x}_i - \overrightarrow{x}_j\right) M \left(\overrightarrow{x}_i - \overrightarrow{x}_j\right)^{\tau} \tag{5.4}$$

Instances are represented by $\overrightarrow{x}_i$ and $\overrightarrow{x}_j$. In some rare case where M is identity, the metric is the same as Euclidean distance. The algorithm developed in this paper makes a distinction between different type of instances called target neighbors and impostors. The test phase specifically is well summarized in the Algorithm 2.

Algorithm 2: Pseudo code for k-NN testing phase

M : Output of the training phase algorithm
actual_class: get class
// get the actual class of the test instances
predicted_class: {}
/* Initialize the set of predicted classes as empty set */
W : 0
// Initialize the set of weights to zero
for *each instance* $\overrightarrow{x}_i$ ***in*** *test set* $\left\{\overrightarrow{x}_i\right\}$ **do**
  k-Nearest = generate k-NN from traning set
  for *each* $\overrightarrow{x}_i$ ***in*** *k-NN* **do**
    C_j = class($\overrightarrow{x}_i$) $W(C_j) = W(C_j) + \frac{1}{d(\overrightarrow{x}_i, \overrightarrow{x}_j)}$
  end for
  predict_class[i] = $arg - max(W)$ /* Choose the index of W with highest value as the class label */
  W = 0 // re-initialize set of weights to zero
end for
while $i <$ *test set* **do**
  if ***actual_class[i]*** == ***predicted_class[i]*** **then**
    Acc = Acc + 1
  end if
end while
Acc = Acc /(test set size)
return (conf. Matrix(**predicted_class**, **actual_class**))

The first type are instances which are selected before the learning process. Each $\overrightarrow{x}_i$ possesses exactly k different target neighbors which share the same class label $\overrightarrow{y}_i$. The target neighbors are the data points that should become nearest neighbors

under the learned metric. The second, i.e, the impostors of a data point $\overrightarrow{x}_i$ are another data points $\overrightarrow{x}_j$ with a different class label, where $\overrightarrow{x}_j$ are nearest neighbors of $\overrightarrow{x}_i$. During the learning, the algorithm performs the minimum number of impostors for all the data instances in the training set. It is important to notice here that, the large margin nearest neighbor algorithm which is particularly adopted in this paper optimizes the matrix M using semi-definite programming [14]. The objective function in this case has two components, for all data observation $\overrightarrow{x}_i$, the target neighbors should be as close as possible and the impostors should be as far away as possible simultaneously. The leaned pseudo metric causes the input vector $\overrightarrow{x}_i$ to be surrounded by training instances of the same class. In case a test point is flawlessly classified, it would be done under the new metric distance scheme, because it will be surrounded by its target neighbors. The testing phase however, focuses on determining the weighted k-nearest neighbors and choosing by the majority voting which class has highest representation in the k-nearest neighbors through the evaluation of the instances weight.

5.5.5 Discussion

In this sub-section, the effectiveness of the proposed approach is discussed. The data points are visualized in scatter plots and colored according to their true labels. For better visualization, plots are scaled in the range from 0 to 7. One can observe that pair of feature group are comfortably distribute different fault category in clusters, making them easier for rule-based classifiers to generate prediction.

The Fig. 5.7 shows the selected features by their importance, which helped in reducing the number of features to be used both in training and testing. Crest factor and Kurtosis happened to be the most informative features, and were selected.

The features used in model utilized PCA to keep enough components, to explain 95% variance, 2 components were kept after training, the explained variance considered per component is then 55,8% and 44,2%.

Fig. 5.9 and 5.8 present the confusion matrices using in k-NN an SVM. The diagonal elements represents the samples that have been correctly classified, whereas the off-diagonal are the numbers of test samples that have not been classified properly. The True Positive Rate (TPR) and False Negative Rate (FNR) are presented as well. Several work in the literature [15] point out that techniques such as k-NN are not always the best in mechanical fault classification (bearings or rotor bars in this case). However, the approach presented in this paper demonstrate that this simple

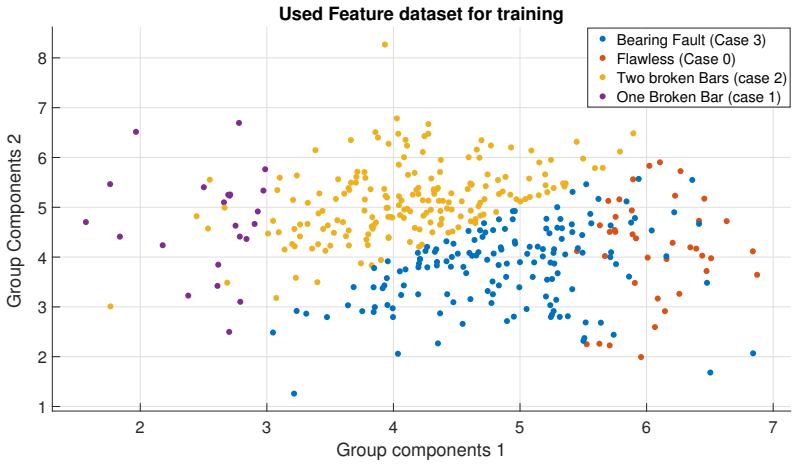


Figure 5.6: Overview of data used for training

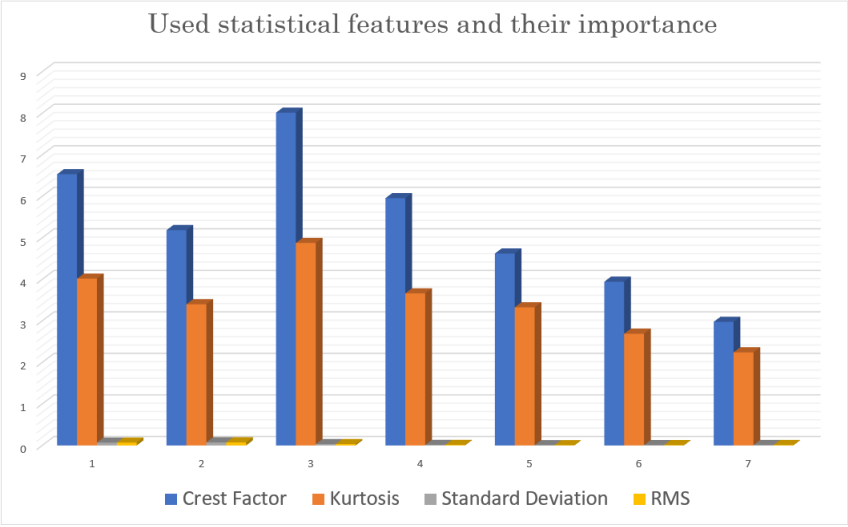


Figure 5.7: Used feature for training and testing

Table 5.3: Overview of used models performance

	EEMD - SVM	EEM - k-NN
Model type	Preset : Fine Gaussian Kernel function : Gaussian Kernel scale : automatic Multicalss method : One-vs-One	Preset : Weighted Number of neighbors : 10 Distance : Euclidean Distance weight : squared inverse Standardize data : true
Training Phase	Accuracy : 87% Total cost 52 Prediction speed : 6000 obs/sec	Accuracy : 85.5% Total cost : 58 Prediction speed : 7700 obs/sec
Testing Phase	Accuracy : 91.3% Total cost : 35	Accuracy : 100% Total cost : 0

classification methods can be enhanced up to 94% in the training phase.

From the confusion matrices presented in Fig 5.9 and Fig 5.8, one can observe that, applying SVM demonstrate a high generalization capability comparing to most of recent learning approaches discussed in the literature [10, 16, 17], whereas k-NN scores the highest training and classification accuracy, 91% and up to 100% in the testing phase. Both algorithms SVM and k-NN show that one can achieve stunning results with small sample analysis ability.

5.5.6 Comparison with recent similar studies in the literature

The classification models developed in this paper is based on data similar to the one used by Christian Lessmier et al. [18]. Their work proposed to train various machine learning algorithms using features of both artificial damages and real damages. In total, 7 ML algorithms were used in their work including SVM and k-NN which are discussed in this paper. They used particle swam optimization (PSO) as turning parameter for SVM, the performance obtained was about 60.9% of accuracy, whereas the k-NN presented an accuracy of 45.5%. Yan et al [19] utilized the same PSO-SVM approach as Christian Lessmier et al. They extracted roughly 37 features in time, frequency and time-frequency domain. They reported an accuracy of 93.3% when considering only time domain features, and almost 97% as the average of accuracies for time, frequency and time-frequency domain. The case using just the time domain features, is consistent with the one used in this work. A summary of recent

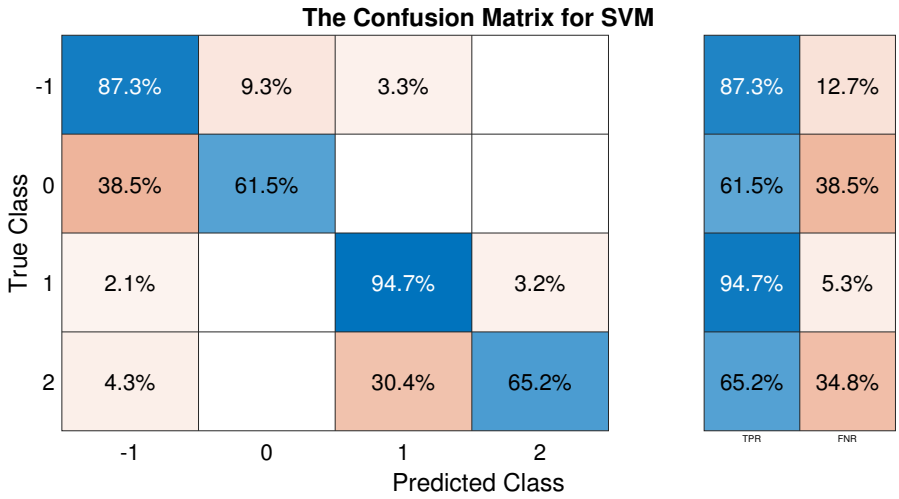


Figure 5.8: Confusion matrix for Support Vector Machine (-1 : Two Broken Bars, 0: Flawless, 1: Bearing Fault, 2: One Broken Bar)

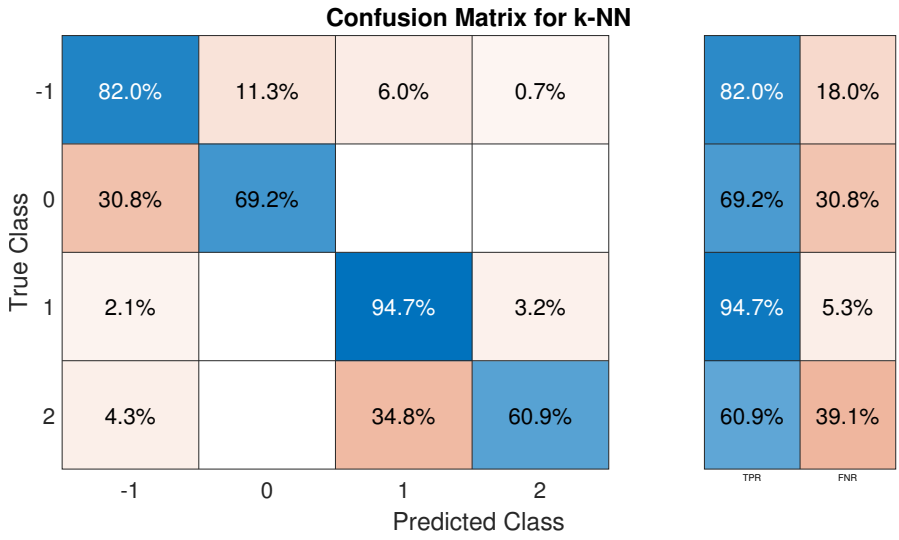


Figure 5.9: Confusion matrix for k Nearest Neighbors (-1 : Two Broken Bars, 0: Flawless, 1: Bearing Fault, 2: One Broken Bar)

and relevant aforementioned research outcomes are presented in the table 6.1.

Table 5.4: Summary of a qualitative comparison of recent works in the literature

Authors	Used Approach	Features	Obtained results
Yan X et al.(2018)	PSO-SVM	16, 13, 8 respectively Time, Frequency and Time-frequency domain	93.33% Accuracy
Ben Ali et al.(2015)	ANN	10 Statistical and over 8 EMD energy entropy	93% Accuracy
Chhaya Grover et al.(2020)	Rule-based classifiers	3 Hjorth Parameters	93.82% Accuracy
Qingfeng Wang et al.(2021)	Weighted k-NN	13 time domain, 4 frequency-domain and 17 entropy	80% - 97% Accuracy
Christian Lessmeier et al.(2016)	SVM-PSO	Time domain	75% Accuracy
This paper	EEMD-SVM, EEMD-weighted k-NN	Time-domain features	94.7% Accuracy

The results carried out in this paper presents a better accuracy with few number of features. The time taken by CPU for fault classification was about 1.319 s for SVM and 54.059 s for k-NN whereas the method proposed in [19] in tested set was reported to be 94.55 s, which is much higher comparing to the the result presented in this paper.

5.6 Conclusion

This chapter elaborate on data mining approach and demonstrate how machine learning algorithm can be used in fault detection and classification. The cascade methods were combining to some of the most effective machine learning algorithms used for classifications, namely SVM and k-NN. Knowing as one of the best cascade methods, EEMD was used on original data signal. Intrinsic Mode Functions (IMFs) of each signal and its residual were calculated. IMFs holding the most informative features were selected to obtained features vector. We utilized PCA for feature selection, so that the most informative can be used to train the model. The proposed models were tested for different operating conditions using front bearing's outer ring of the motor and broken bars faults. Based on the experimental outcomes, combining EEMD and ML algorithms drastically improves the result in terms of effectiveness and accuracy, whereas most of classification algorithms hardly exceed 90% of accuracy, the result obtained in this easily reached 94.3% of accuracy. On the other hand,

EEMD is a adaptation signal processing method with anti-aliasing effect, which is usually suitable for transient impact. Combining this with the presented ML algorithms demonstrates the superiority of the approach on previous studies which hardly performed well on limited data set.

References

[1] Mehmed Kantardzic. *Computers - Algorithms and Data Structures*. Wiley-IEEE Press, 2020.

[2] Andriy Burkov. *Machine Learning Engineering*. Leanpub, 2020.

[3] Majid Ramzan and Majid Ahmad. Evolution of data mining: An overview. In *2014 Conference on IT in Business, Industry and Government (CSIBIG)*, pages 1–4, 2014.

[4] Xin Yang, Raula Gaikovina Kula, Norihiro Yoshida, and Hajimu Iida. Mining the modern code review repositories: A dataset of people, process and product. In *2016 IEEE/ACM 13th Working Conference on Mining Software Repositories (MSR)*, pages 460–463, 2016.

[5] David He, Ruoyu Li, and Junda Zhu. Plastic bearing fault diagnosis based on a two-step data mining approach. *IEEE Transactions on Industrial Electronics*, 60(8):3429–3440, 2013.

[6] Mark Foster. Building the cognitive enterprise: A blueprint for ai-powered transformation." ibm institute for business value, 2 2020.

[7] Mallat Stéphane. pages 263–376. Academic Press, 2009.

[8] Y Qin, Y Mao, B Tang, Y Wang, and H Chen. M-band flexible wavelet transform and its application to the fault diagnosis of planetary gear transmission systems (2019) mechanical systems and signal processing.

[9] Y Zhan, Z Shi, and M Liu. *The Application of Support Vector Machines (SVM) to Fault Diagnosis of Marine Main Engine Cylinder Cover*, pages 3018–3022. 2007.

[10] I Steinwart. *On the optimal parameter choice for /spl nu/-support vector machines*, volume 25, pages 1274–1284. 10 2003.

[11] R Ricci and P Pennacchi. Diagnostics of gear faults based on emd and automatic selection of intrinsic mode functions. *Mech. Syst. Signal Process*, 25:821–838, 2011.

[12] J Yang, D Huang, D Zhou, and H Liu. Optimal imf selection and unknown fault feature extraction for rolling bearings with different defect modes. *Measurement: Journal of the International Measurement Confederation*, 157, 2020.

[13] S Orhan, N Aktürk, and V Çelik. Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool. *Comprehensive case studies*, 39(4):293–298.

[14] M Kang, J.-M Kim, A Kim, E Tan, B.-K Kim, and Choi. Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis. *IEEE Trans. Power Electron*, 30(5):2786–2797, 5 2015.

[15] Y, Lei Jia, J Lin, S Xing, and S Ding. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans. Ind. Electron*, 63(5):3137–3147, 5 2016.

[16] Q Wang, S Wang, B Wei, W Chen, and Y Zhang. Weighted k-nn classification method of bearings fault diagnosis with multi-dimensional sensitive features. *IEEE Access*, 9:45428–45440, 2021.

[17] Q, Wang Liu, B Wei, and W Chen. Online incipient fault detection method based on improved ell_1 trend filtering and support vector data description. *IEEE Access*, 9:2021.

[18] C Lessmeier, J Kimotho, D Zimmer, and W Sextro. *Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification*, page 17. 7 2016.

[19] X Yan and M Jia. A novel optimized svm classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing. *Neurocomputing*, 313:47–64, 2018.

Chapter 6

Digital twin-driven artificial intelligence for fault detection and identification

This chapter proposed to take advantage the digital twin concept in the evaluation of the current condition of a given system and more importantly to predict future behavior, refine the control, or optimize operation by detecting and identifying faults as early as possible without necessarily have to access the system physically.

6.1 Introduction

Lately, interest in Digital Twin (DT) and artificial intelligence applications for fault detection and classification became greater. The concept "twin" is originally derived from NASA's Appollo Project [1], where the aircraft's twin body is a real physical system [2]. In the aforementioned program, astronauts and NASA staffs used twin model to make decision under emergency situations. In 2002, Wang et al [3] proposed the use of twins for the formation of a product lifecycle management center. Back then, Digital Twin was only a digital representation of a physical object provide an enhanced and rich 3D view. With the actual development of the Internet of Things (IoT), DT model becomes able to support new intelligent services to connect and interact with physical object. In condition monitoring, the DT model runs in parallel with the real targeted equipment and immediately flags operational behavior that deviates from the expected behavior (6.3.2).

The DT usage can be extended in condition monitoring to applications such as

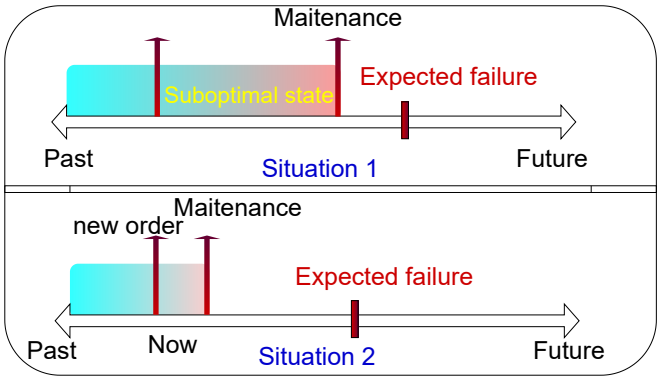


Figure 6.1: Illustration of possible scenarios and expected failure

the determination of remaining useful life and the most opportune time to service or replacement of an equipment.

6.2 The chapter's goal and limitations

This chapter focuses on providing useful patterns on how to approach digital twin model design. On the other hand, the chapter proposes a reference model for rotating machines as well as for application such autonomous vehicles fault detection. A solid evidence-based detail on the requirements of building a digital twin model for fault detection is discussed. The illustrative models presented in this chapter do not use physically collected data, simply because there are often restricted by the condition of acquisition and the specific component for which they were recorded. In addition, this chapter is mainly focus on providing non-exhausting dynamic model entirely based on a co-simulation runs thanks to MATLAB and SolidWORKS without dealing with the integration of data acquired from the twin of the targeted equipment under test.

6.3 Benefits of using a digital twin model

Using digital twins in equipment maintenance has tremendous benefits including:

- Visualize the component in question and understand the fault behavior by real users in real time

- Building a digital thread and promoting the tractability
- Fault correction far away from the equipment
- Managing complexities and linkage within systems of systems
- Asset history
- Future fault scenario simulation

The sub-section 6.3.1 and 6.3.2 provide more details on the two later benefits.

6.3.1 Asset history

One of the greatest benefit of adopting a digital twin in a maintenance policy is that, it captures the real asset's history. As stated in section 6.1, a DT model is updated periodically to represent the real current state of the real asset. Over time, this become the typical asset history. Thereby, one can compare for example data from a given equipment to the digital histories of the other equipment belonging to the same category to understand how they behaved under similar faults and the effects such the on of the fleet efficiency.

6.3.2 Future fault scenario simulation

As aforementioned, a DT model can enable the understanding of the history of the assets, at the same time, they can also help to plan for the future decisions regarding maintenance activities(Fig. 6.1). One can use the model to simulate thousands of possible situations that might happen in the future during equipment's lifecycle to see how factors like weather, speed, vibrations, or all other possible operating conditions might affect the performance of the targeted equipment. This approach help in managing the asset and optimize operations by providing valuable information to the maintenance staff about the expected failure in advance, so they can schedule the future intervention if needed. The modeling method need to be used however, is driven by the intended use. When the DT model can be use to predict failure in advance and reduce the downtime, manage spare part inventory and do the what if simulations, etc.. it is important to carefully decide on modeling methods [4]

6.4 The architecture developed in this work

The reason why one has to investigate how to design a suitable twin model, is because digital twin is actually more than just "twin" of a physical system. In fact, it becomes a key enabling technology for cybernetic manufacturing. DT can not only be used for modeling and simulation of system development to support design or to validate system properties [5–7], but can also support the operation, manufacturing service for operation optimization and failure prediction. The DT paradigm shift is suggested for on-board integrated automated vehicle health management system enabling safety [5]. A customer-tailor model developed in this work is based on the reference twin construction model as presented in the Fig 6.2

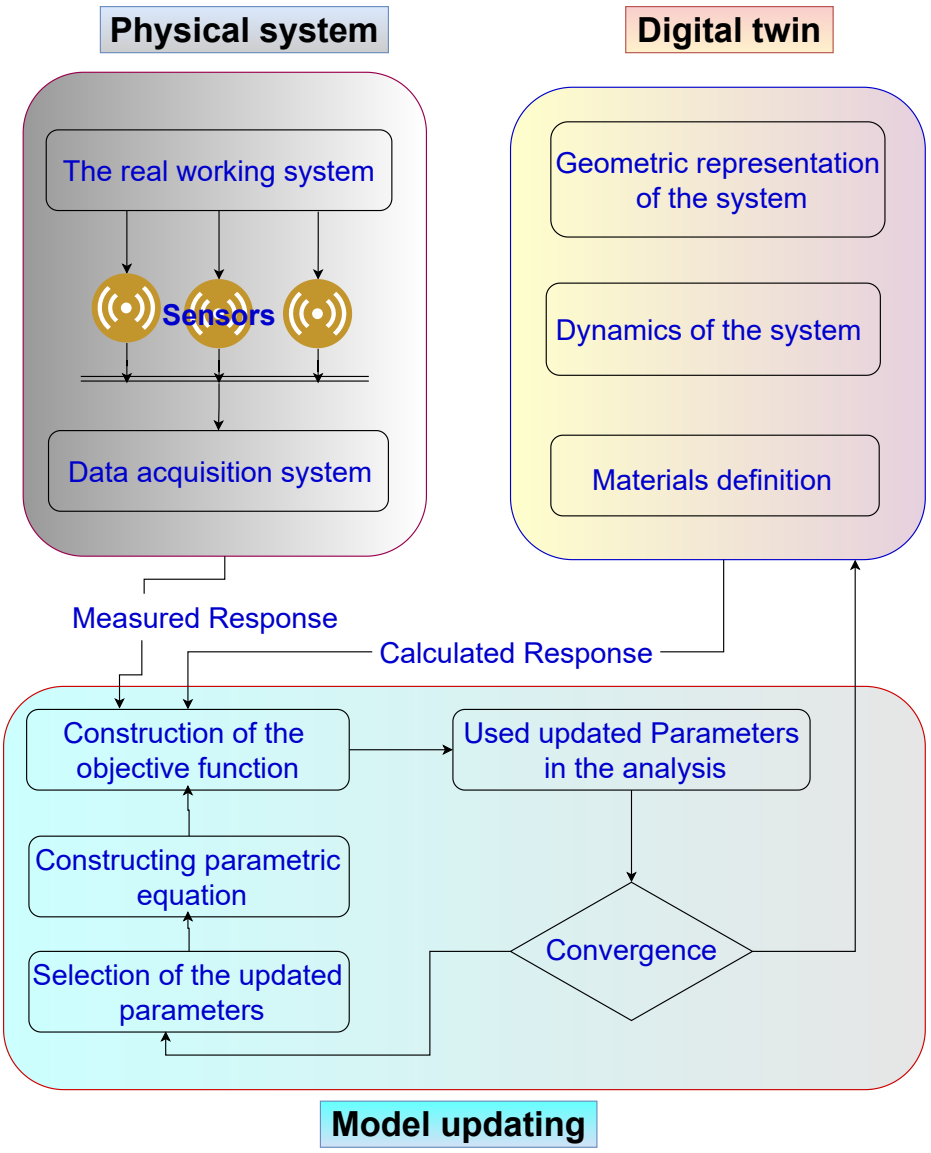


Figure 6.2: Reference twin model for dynamic system

In the Fig 6.3, a reference model is particularly design for rotating machinery fault detection and diagnosis. The model is essentially a unique living model of the physical system with the support of enabling technologies including multi-physics simulation, machine learning, AR/VR, etc... This makes possible a continuous adaptation to the changes in the environment or operations, hence, delivers the best possible result. On the physical side, more information on the characteristic of the targeted equipment is collected to get real working conditions. One the virtual side, several characteristic behavior is added to the DT, thus, designers can not only visualize the equipment, but also test its performance for fault diagnosis.

The designed model is updated based on sensible parameters and its dynamic is updated according to the work status and operating conditions of the physical system. Well elaborated readings on this specific point can be found in [8, 9]

6.5 Use case example 1: Integration of a fault detection scheme for autonomous vehicles

The use case illustrated in this chapter, utilizes the automated driving toolbox of MATLAB, aiming to assess the vision processing in the scenario where the highway lane is considered and the vehicle traveling within a marked lane. The Fig. 6.4 describes the 3 main blocks (scenario, vehicle and sensors) used in the analysis. All selected elements are configured and tested in a 3D environment that includes the radar and camera sensor model.

The model allows to steer the car to travel within a marked lane; it maintains a set of speeds at a safe distance to a forward vehicle driving in the same lane. The vehicle detection from the camera is fused with detection from radar to improve the robustness of perception. The controller uses lane detection, vehicle detection, and a set speed to control steering and acceleration. The thresholding approach for fault identification is adopted. Cameras are mounted on the front and rear bumper, left and right mirror, rear-view mirror, Hood and roof center. The radial distortion and tangential distortion are set to zero as well as the axis skew. For the radar, the azimuthal resolution is initially set to 0.8 degree with a range resolution of 1 m, the range rate resolution of $1\,\mathrm{m\,s^{-1}}$. The radar fractional range bias component is set to 0.05 and the detection probability is about 0.9.. The complete test bench model is presented in Fig. 6.5. This model is inspired by the one provided by MathWorks 2021 [10]. Different configurations were done during the analysis; in some scenarios, for instance,

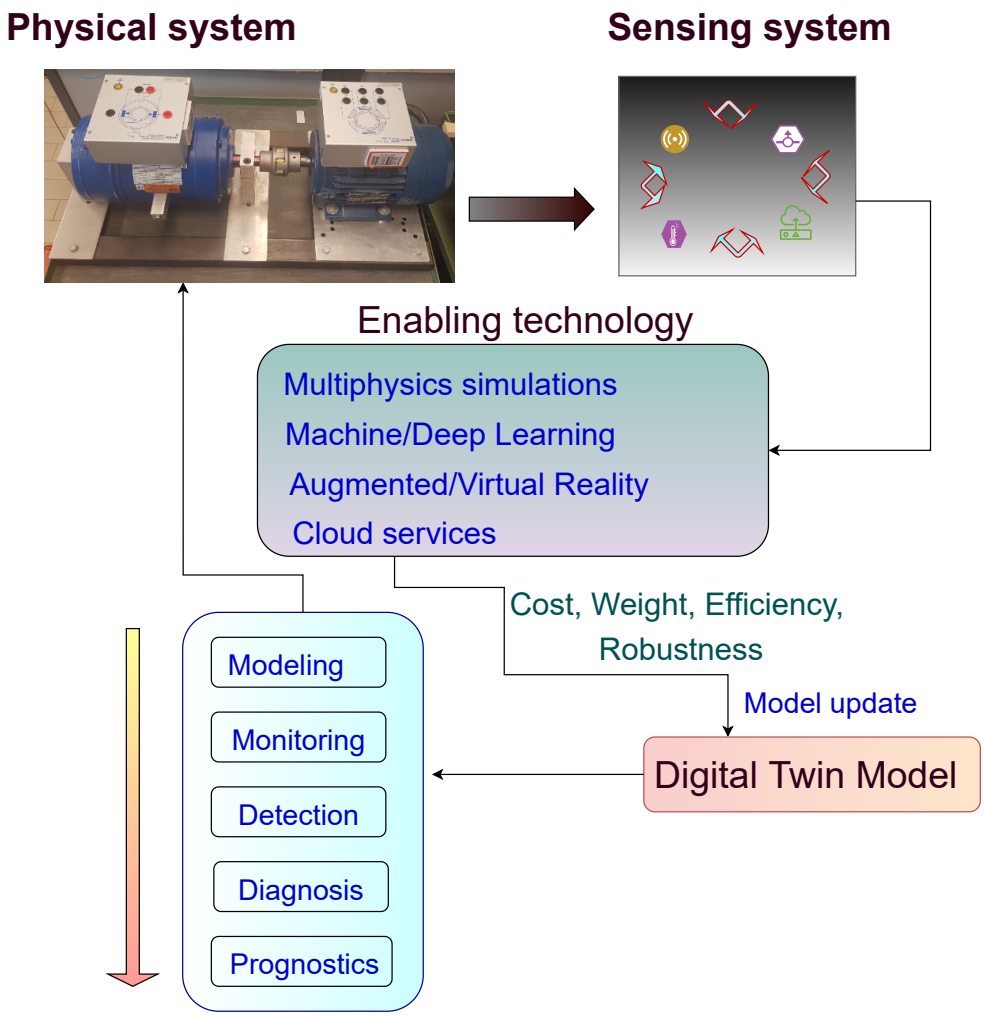


Figure 6.3: A fault diagnosis framework based on digital twin

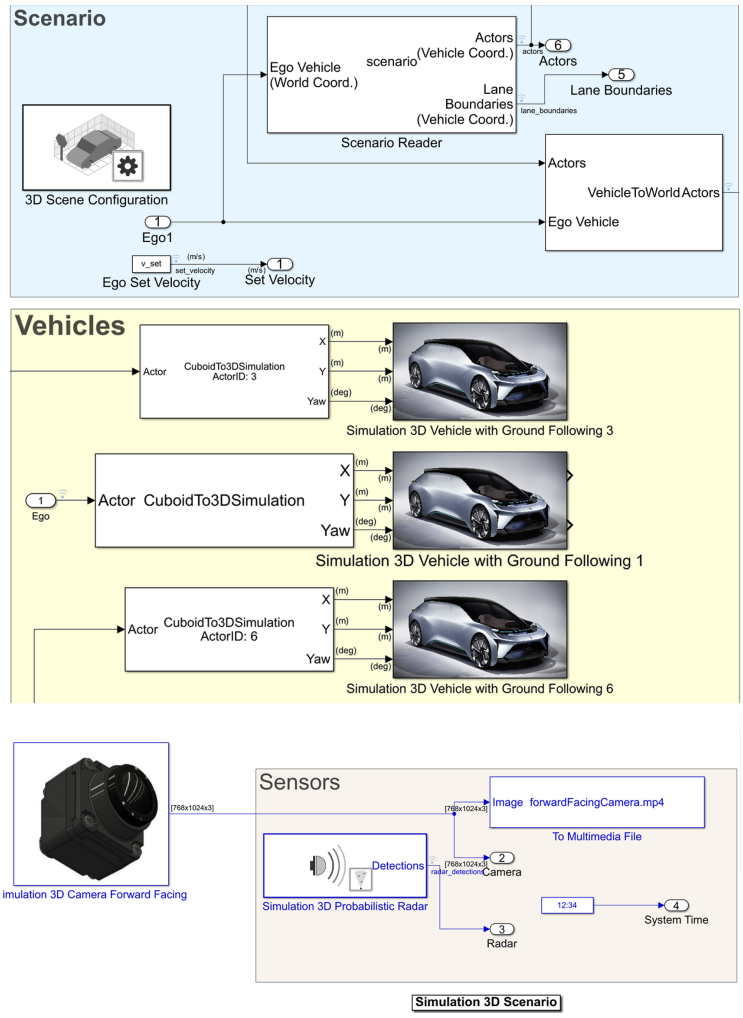


Figure 6.4: The simscape scheme of the used model.

the test was about following the lanes and avoid collisions with other vehicles.

The simulation scenario presented in the Fig. 6.4 is included in the model as a subsystem that specifies the road, the vehicle, the camera sensor, and the radar sensor used for the simulation.

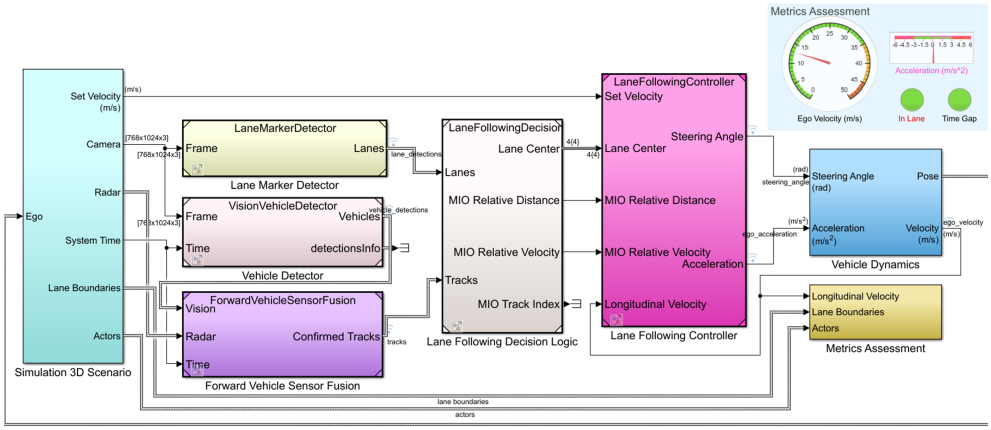


Figure 6.5: The test bench model holding all simulation modules

In Fig. 6.6 (a), one can see the lateral offsets of the detected left and right lane boundaries. The detected value is close to the ground truth, which happens to be around 2.5 meter in this simulation. The deviation is slightly tiny, about 3%, which strengthens the reliability of the simulation outcome, therefore the use of the camera lonely for lane detection.

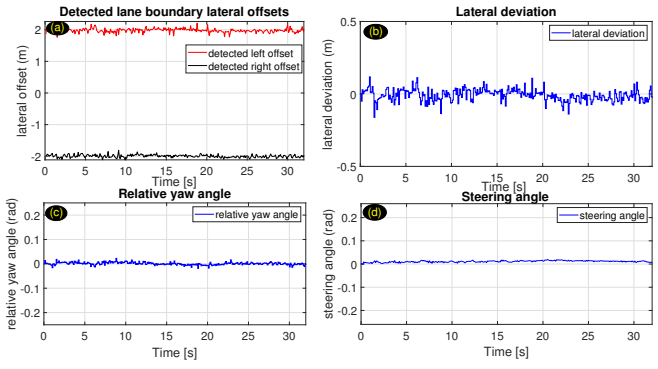


Figure 6.6: lateral results

Fig. 6.6 (b) shows the lateral deviation of the ego car where the start point is

the centerline. In the optimal situation, the lateral deviation is supposed to be equals to 0 m, which implies that the ego vehicle perfectly follows the centerline. However, slight deviations might appear if the vehicle changes the velocity to avoid probable collision with another vehicle on the road. Fig. 6.6 (c) presents the relative yaw angle between the centerline of the lane and the ego vehicle. The relative yaw is near zero, which implies that the car's heading angle matches the centerline's yaw angle. One can see in Fig. 6.6 (d) the steering angle of the vehicle, which presents a pretty smooth profile.

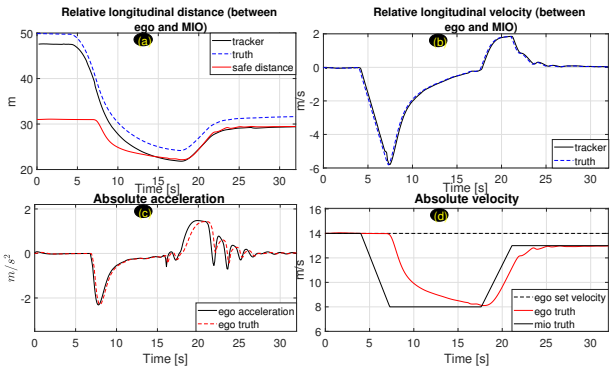


Figure 6.7: Longitidunal result

The distance between the MIO and the ego vehicle can be seen in the Fig.6.7 (a). In fact, the ego vehicle moves towards the MIO and gets close to it. In some situations, it is possible to observe that the ego vehicle exceeds the safe distance. In the relative longitudinal velocity as presented in the Fig.6.7 (b), the vehicle detector only manage to detect positions, the tracking algorithm should be improved at this point, however, it is interesting to see how one can still estimate the velocity. The Fig.6.7 (d) points out an critical aspect that autonomous cars still dealing with namely speed regulation based on tracking objects. The plot shows the vehicle ego at the beginning follows the set velocity. When the MIO slows down to avoid a collision, the ego vehicle slows down as well.

6.6 Use case example 2: Building a parameter estimation model for centrifugal pumps fault detection and classification

6.6.1 The used strategy

For this study the measurements are taken from the nominal operation of the machine, the parameters of the model are estimated and their uncertainty is quantified [11]. The acquired measurements are used to re-evaluate the parameters and the resulting estimates are compared to their nominal values. The Fig. 6.8, presents a schematic representation of the used pump where inputs and outputs can be visualized.

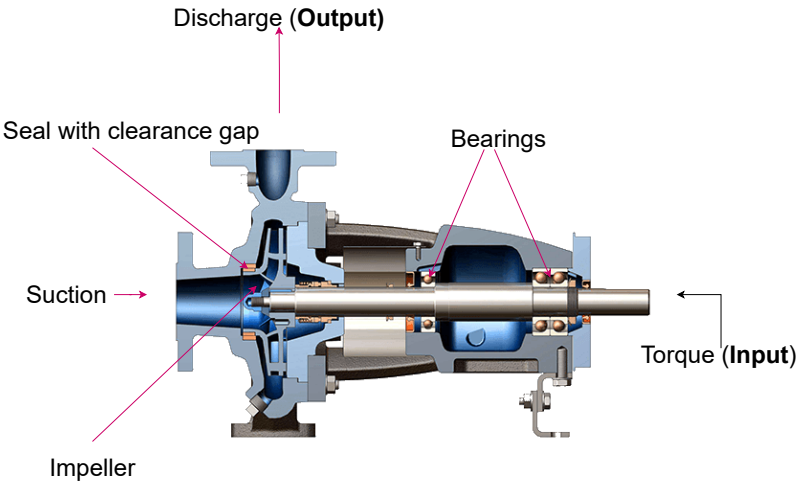


Figure 6.8: Schematic representation of the pump model investigated where input and output are shown

The study uses the common practice used for pump calibration and supervision of operating pumps at a constant speed and recording the static head of the pump and the fluid flow rate. An increase in flow rate results in a decrease in pump head. The pump head characteristics are then compared to the values provided by the manufacturer. Differences may indicate the probability of faults.

The observed changes is suggested to be turned to quantitative information. A reasonable approach is to fit a parameterized curve to the head-flow characteristic. By utilizing the governing equations of pump-pipe dynamics, and applying a simplified

torque relationship, the following equations are obtained:

$$H(t) = h_{nn}\omega^2(t) - h_{nv}\omega(t)\dot{V}(t) - h_{vv}\dot{V}^2 \tag{6.1}$$

$$J_P\frac{d\omega}{dt} = \phi I_2(t) - Mf_0 - \rho g h_{th1}\omega(t)\dot{V}(t) \tag{6.2}$$

$$L_2\frac{dI_2(t)}{dt} = -R_2 I_2(t) - \phi\omega(t) + U_2(t) \tag{6.3}$$

and the least square estimation parameter can be expressed as:

$$M_p \approx k_0\omega V - k_1 V^2 + k_2\omega^2 \tag{6.4}$$

The measured signals are naturally U_2,I_2 respectively the armature voltage and the armature current. $\dot{V}$, the volume flow rate, ω the angular velocity, and H pump total head. In the Eq. 6.4, parameters to be estimated are depicted. $h_{\mathrm{nn}}, h_{\mathrm{nv}}, h_{\mathrm{vv}}, k_0, k_1, k_2$ are parameters to be estimated. If one measures ω, Q, H and M_p the parameter can be estimated by linear least squares. These parameters are the features that can be used to develop a fault and diagnosis algorithm. The Fig. 6.9 presented the flowchart which summarized the suggested diagnosis approach.

6.6.2 Used data and preliminary comparisons

The data used are retrieved from Mathworks Fault Diagnosis dataset [11]. The data was recorded while the driver runs at the nominal speed of 2900 RPM, where the pump uses a coupling power transmission. The table 6.1 illustrates how the values of k_0 and h_{nn} decreases when the measured clearance gap is larger than nominal value. However for a small clearance, one can notice a larger value of k_0 and h_{nn}.

Table 6.1: Parameters and investigated faults

	k0	k1	k2
Healthy	3×10^{-4}	0.017	2.821×10^{-7}
Large Clearance	3×10^{-3}	0.016	3.029×10^{-7}
Small Clearance	3×10^{-4}	0.016	2.667×10^{-7}
	hnn	hnv	hvv
Healthy	5.11×10^{-6}	8.615×10^{-5}	0.011
Large Clearance	4.849×10^{-6}	8.362×10^{-5}	0.011
Small Clearance	5.368×10^{-6}	8.476×10^{-5}	0.009

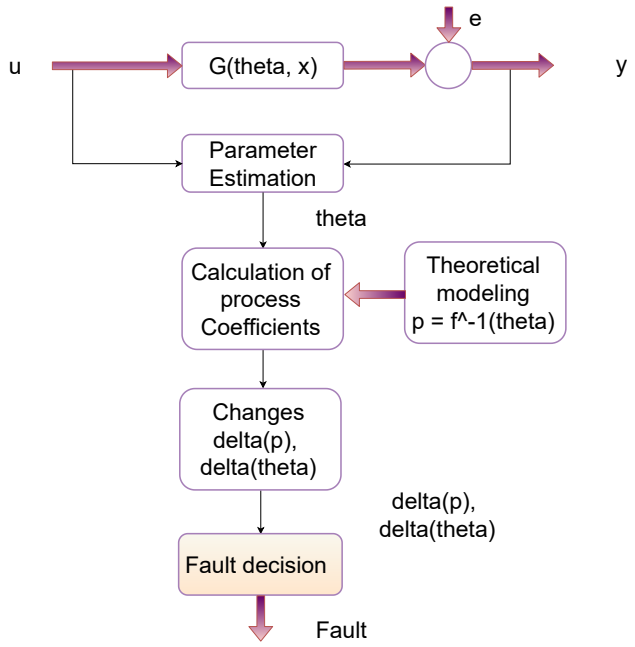


Figure 6.9: The suggested strategy flowchart

Furthermore, one can observe that k_2 and h_{vv} decrease for small clearance gap and increase for large gap. To diagnose the faulty pump, the measurement of speed, torque and flow rate were collected for the different invested conditions. For instance, when the introduced fault is in the clearance ring, the measured head characteristics for the pumps indicate a significant change in the profile.

Based on initial observations, one can notice changes in parameters which can be interpreted as sign of possible faults. But, even a CP in healthy condition may exhibit variations in the measurements due to the measurement noise, fluid contamination and viscosity changes and eventually the torque and slip characteristics. These variations introduce uncertainty in estimated parameters, $k_0, k_1, k_3, h_{nn}, h_{nv}$ and h_{vv} in the case of this paper. The Fig. 6.10 displays the parameters uncertainty as 85% confidence regions.

The ellipsoid is the actual region of confidence of healthy pump parameters. This allow to establish the comparison between values corresponding to healthy state and faulty states. In the table 6.2, measured signals are divided in chunks of 50 runs. One can visualize the repeatability of all the estimated parameters in different fault stages.

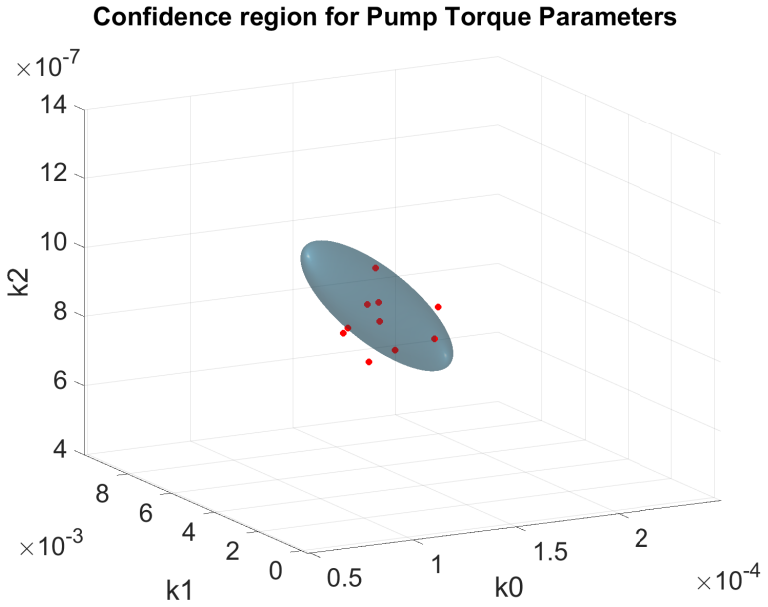


Figure 6.10: Torque parameters where the ellipsoid indicates the confidence regions of healthy pump

Table 6.2: Estimated parameters repeatability

		Chunk 1	Chunk 2	Chunk 3	Chunk 4	Chunk 5	Chunk 6	STD
Healty	h_{nn}	5.081×10^{-6}	5.095×10^{-6}	5.095×10^{-6}	5.158×10^{-6}	5.132×10^{-6}	5.094×10^{-6}	2.943×10^{-8}
	h_{nv}	1.071×10^{-4}	1.224×10^{-4}	1.022×10^{-4}	1.205×10^{-4}	1.102×10^{-4}	1.076×10^{-4}	3.135×10^{-8}
	h_{vv}	0.008	0.007	0.009	0.007	0.008	0.008	4.991×10^{-8}
Small Clearance Gap	h_{nn}	4.360×10^{-6}	4.385×10^{-6}	4.326×10^{-6}	4.330×10^{-6}	4.297×10^{-6}	4.360×10^{-6}	8.029×10^{-6}
	h_{nv}	1.080×10^{-4}	1.141×10^{-4}	9.778×10^{-5}	8.748×10^{-5}	1.037×10^{-4}	9.093×10^{-5}	1.0196×10^{-5}
	h_{vv}	0.009	0.008	0.009	0.008	0.009	0.010	8.385×10^{-6}
Big Clearance Gap	h_{nn}	4.785×10^{-6}	4.838×10^{-6}	4.839×10^{-6}	4.853×10^{-6}	4.844×10^{-6}	4.939×10^{-6}	7.527×10^{-4}
	h_{nv}	8.493×10^{-5}	9.481×10^{-5}	9.923×10^{-5}	1.103×10^{-4}	9.239×10^{-5}	9.581×10^{-5}	7.527×10^{-4}
	h_{vv}	0.010	0.010	0.009	0.010	0.009	0.009	5.477×10^{-4}

6.6.3 Discussion and observations

One of the challenge in CP monitoring, is to accurately estimate how much height that CP can lift the fluid. The Fig 6.11 presents a plot where apparent variation in the pump characteristics (the discharge and head) can be seen. This is critical to evaluate the CP capacity to properly operate.

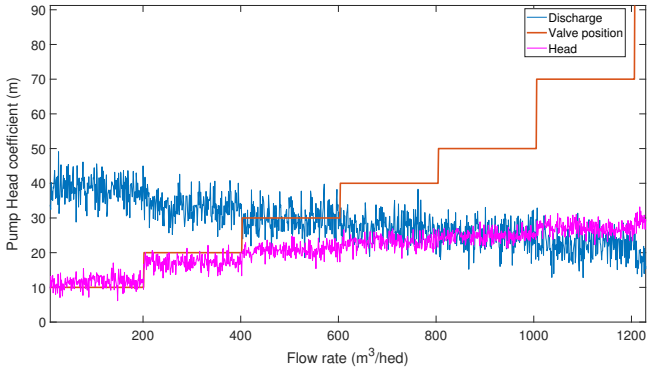


Figure 6.11: The measured discharge and head with alternative valve position

The discharge rate can be seen based on alternative valve position. One can observe a progressive increase in the head parameters whereas the discharge decrease with high values of the valve position. In the Fig 6.12, faults mode histograms for different used parameters (h_{nn}, h_{nv}, h_{vv}) are presented. Changes in clearance gap can be split into large and small. One can see that the parameter h_{nn} presents a better separation respect to the rest.

For torque parameters presented in the Fig 6.13, it can be seen that, individual separability is not good. A persistent variation in mean values and variances can be be observed.

Given that the Probability Distribution Function (PDF), can exibit a useful information on the separation in the mean or variance, we utilized likelihood ratio test to rapidly assign a test dataset to most likely mode. If we consider that available parameter sets as test samples for mode prediction, assign the predicted mode as belonging to one of which the joint PDF has the highest value.One can display the results and comparing the true and predicted mode in a confusion matrix.

As expected, the confusion matrix shows a perfect separation between the tree modes as seen in the histograms for the parameters h_{nn} with an accuracy roughly about 99%. The result is not the same for the torque parameters, but still the success

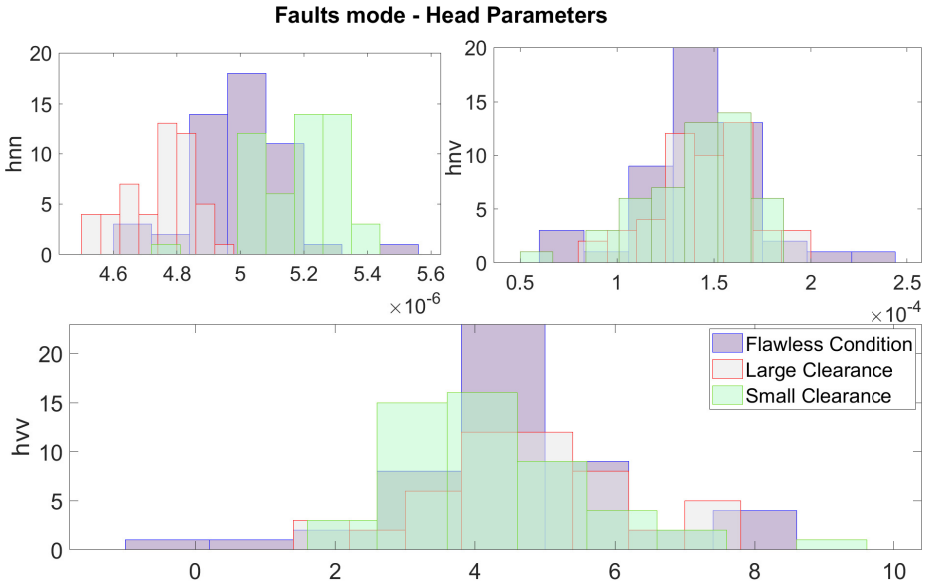


Figure 6.12: Histograms for head parameters

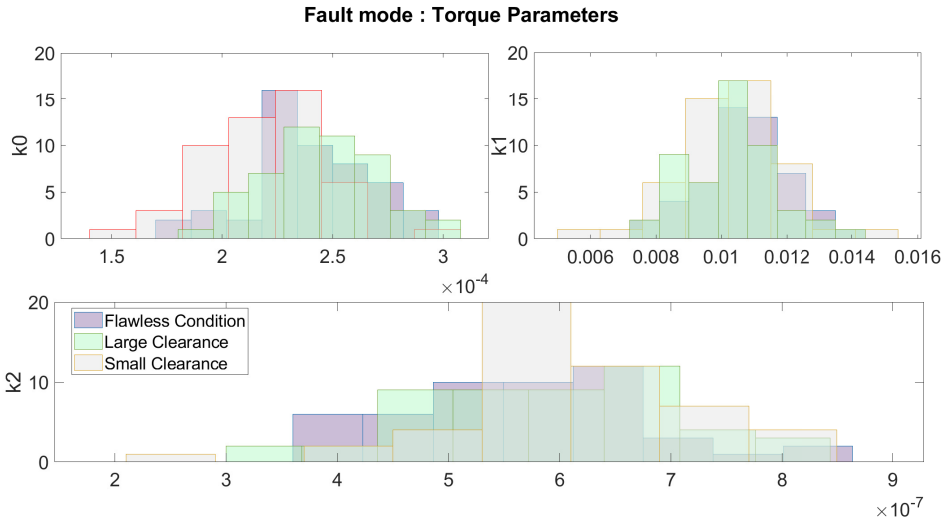
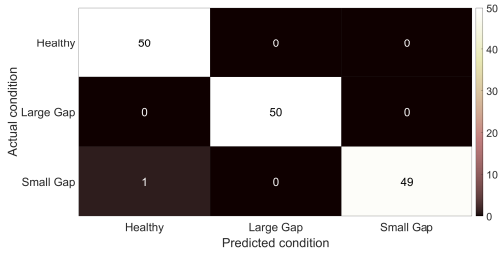
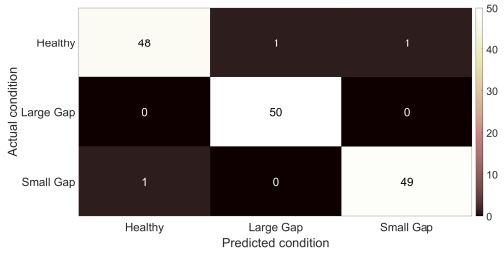


Figure 6.13: Histograms for torque parameters



(a) Pump head confusion matrix



(b) Pump torque confusion matrix

Figure 6.14: Confusion matrices for the torque and head

rate is high about 97%. This is justified by the fact that the PDF value calculation is affected by mean and the variance.

6.7 Conclusion

This chapter paws paths for the use of techniques based data mining and artificial intelligence in order to incorporate on a the digital twin model and help in the evaluation of the current condition of a given system and more importantly to predict future behavior, refine the control, or optimize operation by detecting and identifying faults as early as possible without necessarily have to access the system physically.

References

[1] R Rosen, G Wichert, G Lo, and K Bettenhausen. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3):567–572, 2015.

[2] Jinjiang Wang, Lunkuan Ye, Robert X. Gao, Chen Li, and Laibin Zhang. Digital twin for rotating machinery fault diagnosis in smart manufacturing. *International Journal of Production Research*, 57(12):3920–3934, 2019.

[3] X Wang and L Wang. Digital twin-based weee recycling, recovery and remanufacturing in the background of industry 4. *International Journal of Production Research*, 55:1–11, 2018.

[4] Yassine Qamsane, James Moyne, Maxwell Toothman, Ilya Kovalenko, Efe C. Balta, John Faris, Dawn M. Tilbury, and Kira Barton. A methodology to develop and implement digital twin solutions for manufacturing systems. *IEEE Access*, 9:44247–44265, 2021.

[5] F Tao, F Sui, A Liu, Q Qi, M Zhang, B Song, and A Nee. Digital twin-driven product design framework. *International Journal of Production Research*, 56:1–19, 2018.

[6] T Uhlemann, C Schock, C Lehmann, S Freiberger, and R Steinhilper. The digital twin: Demonstrating the potential of real time data acquisition in production systems. *Procedia Manufacturing*, 9:113–120, 2017.

[7] Q Zhang, X Zhang, W Xu, A Liu, Z Zhou, and D Pham. *Modeling of Digital Twin Workshop Based on Perception Data*, pages 3–14. Springer, Cham, 8 2017.

[8] R Söderberg, K Wärmefjord, J Carlson, and L Lindkvist. Toward a digital twin for real-time geometry assurance in individualized production. *CIRP Annals*, 66(1):137–140, 2017.

[9] Chenyu Wang and Yingshu Li. Digital twin aided product design framework for iot platforms. *IEEE Internet of Things Journal*, pages 1–1, 2021.

[10] MathWorks Technical support. *Automated Driving Toolbox, User's Guide*. The MathWorks, Inc, 2021.

[11] Rolf Isermann. Fault-diagnosis applications. model-based condition monitoring: Actuators, drives, machinery, plants, sensors, and fault-tolerant system, 2011.

Chapter 7

Conclusions and outlook

This PhD thesis aimed to explore possibilities offered by advanced engineering methods used in condition monitoring. The first step was to circumscribe fundamentals of fault detection and diagnosis before tackling aspects related to signal pre-processing. Having established that detection and diagnosis are precise as the quality of the data on which they are based is better, the specific objectives of this thesis were to contribute to the elaboration of adequate pre-processing approach to be applied to measurable signals with potentially great capability to extract information from the targeted equipment. The main focus in this work was rotating machines. The contributions of this research are both methodological and experimental. Through the an evidence-based approach, we demonstrate that condition monitoring by means of signals on a large class of industrial application still a subject of disagreements both in scientific and industrial community due to cyclostationariy and non-starionarity characteristic of most of these signals, whereas most of signal processing methods are designed for stationary signals. The work demonstrate also that feature extraction is a delicate task, because an error at this level can be misleading in detecting a potential fault threat.

At the experiment level, measurement data from experiment campaign, numerical simulation and from open repositories where used to point out challenges besides different approaches. A particular stress has been done on the measurement obtain in chapter 3, where the spectral analysis of the envelope signal reveals itself to be a powerful and accurate tool for bearing fault detection because of lesser complexity, computational time, and lesser cost. This research's contribution at methodological level can be summarize follow :

- Extension of traditional pre-processing approach by applying enhanced cascades-based methods
- Visual representation of spectrum by displaying them as vectors of cyclical frequencies.
- Addressing the evaluation question of the diagnostic system by combining the process with cascades methods before implementing a classification algorithm.
- Establish a foundation on requirements and a realistic scheme for building a digital twin model for fault detection.

It was demonstrated that, most presented cascade methods primarily require a set of parameters well defined. However, some enhancement at implementation level made them take few parameters, hence, provide suitable adaptability to the signal. The parameters selection was limited and had less influence on the denoising performance. Interesting benefits of using cascade methods were pointed out, such as the transformation of the signal without fault to the null signal.

Developing visual representation was motivated by observations made on traditional time-frequency analysis which concluded that, these techniques were mainly analysis tool as opposed to processing tools. In addition, traditional methods do not offer a versatile methodology that can be apply to all mechanical signals in all situations. The idea here was fundamentally borrowed from the field of thermodynamics to identify nonstationary events of different nature by locating more precisely and specifically demodulation frequencies, which might be of considerable use for several fault detection and diagnosis scenarios. It was found that high values appear in a set of frequency bands. The study demonstrated that, the processing effort grows then proportionally with the number of frequency bands where the signal has to be demodulated. In the case of fault like BPFI, it was shown that the negentropy gives more information even for a shallow size of fault and allows the localization of meaningful frequency bands. The analysis done using spectral correlation, on the other hand, suggested that the approach may not always be indicative of the severity of the bearing faults. However, it can be used as a monitoring tool because its value always reads higher for a faulty bearing than for a healthy one.

Another thing done is this research was to utilize PCA for feature selection, so that the most informative futures could be used to train the model. The proposed

models were tested for different operating conditions using front bearing's outer ring of the motor and broken bars faults. Based on the experimental outcomes, combining cascades methods and machine learning algorithms drastically improves the result in terms of effectiveness and accuracy, whereas most of classification algorithms hardly exceed 90% of accuracy, the result obtained in this easily reached 94.3% of accuracy. On the other hand, method such as EEMD is an adaptation of signal processing method with anti-aliasing effect, which is usually suitable for transient impact. Combining this with the presented machine learning algorithms, demonstrates the superiority of the approach on previous studies which hardly performed well on limited data set. A non-exhaustive insight providing useful patterns on how to approach digital twin model design was lastly discussed as well as a reference model for rotating machines.

The future of condition monitoring lies in the integration of innovative methods as presented in this work. The physical models of degradation, reliability models and methods based on digital twins are ways to prevent the development of faults. However, given the uncertainties of most of models, it can happen that fault development gets hard to prevent and faults might happen earlier than predicted by these model. In these situations, measurement-based methods must be used to update these models and take into account and take into account the fact that the defect has already been initiated.

Condition monitoring and advanced fault detection techniques for engineering systems

Moisè Avoci Ugwiri

Pubblicazione edita a cura del
Dipartimento di Ingegneria Industriale

ISBN 88 – 7897-128-6