Comparison of machine learning models based on time domain and frequency domain features for faults diagnosis in rotating machines

Natalia Espinoza Sepulveda[*], and *Jyoti* Sinha

Dynamics Laboratory, School of Mechanical, Aerospace and Civil Engineering (MACE), The University of Manchester, UK

Abstract. The development of technologies for the maintenance industry has taken an important role to meet the demanding challenges. One of the important challenges is to predict the defects, if any, in machines as early as possible to manage the machines downtime. The vibration-based condition monitoring (VCM) is well-known for this purpose but requires the human experience and expertise. The machine learning models using the intelligent systems and pattern recognition seem to be the future avenue for machine fault detection without the human expertise. Several such studies are published in the literature. This paper is also on the machine learning model for the different machine faults classification and detection. Here the time domain and frequency domain features derived from the measured machine vibration data are used separated in the development of the machine learning models using the artificial neutral network method. The effectiveness of both the time and frequency domain features based models are compared when they are applied to an experimental rig. The paper presents the proposed machine learning models and their performance in terms of the observations and results.

1 Introduction

The increased application of vibration-based condition monitoring (VCM) in industry has been reflected in the continued development of technologies and instrumentation, as well data processing and analysis techniques. Intelligent systems and pattern recognition models seem to be the future avenue for machine fault detection without the human expertise.

Artificial neural networks (ANN) have been beneficially applied in mechanical systems due their suitability for complex sensor data processing problems [1]. Nahvi and Esfahanian [2] have evidenced the capability of ANNs of dealing with a vast number of features and providing acceptable results for the fault detection, however the 100% of accuracy is not achieved at any of their conducted experiments. Bearing faults have been also studied by this method [3, 4], demonstrating an ability for damage detection at an early stage. Other researchers have focused their efforts into localise the failures as well into quantify the actual damage [5, 6]. Nevertheless, the revised works still present a high dependency on

---

[*] Corresponding author: natalia.espinoza.sepulveda@gmail.com

elaborated data preparation procedures regarding specialised pre and post signal processing. In addition, most of them are focused in the identification of a unique fault.

This paper is also on the machine learning model for the different machine faults classification and detection. Here the time domain and frequency domain features derived from the measured machine vibration data are used separated in the development of the models using the artificial neutral network method. The effectiveness of both the time and frequency domain features based models are compared when they are applied to an experimental rig. The paper presents the proposed machine learning models and their performance in terms of the observations and results.

2 Experimental rig and data

The schematic of the rig is shown in Figure 1. This study uses pre-existent data acquired in an experimental rig in laboratory conditions. Five different conditions for the rotor are considered and simulated. Firstly, the healthy condition (baseline, with residual misalignment and residual unbalance) and four independent faulty conditions, namely misalignment, bow, looseness and rub. Each fault is introduced in two different points and/or severity level in the rotor, covering a wider range of damage for the same diagnosis.
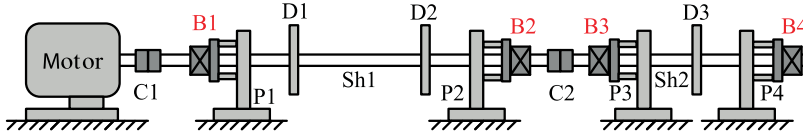
Figure 1. Experimental rig.

The vibration signal is acquired simultaneously from four uniaxial accelerometers, each of them located in one of the four bearings. The sensors are positioned at 45º from the horizontal line in anticlockwise. Data acquisition is done at a sampling frequency of 10^4 Hz. The rotor is running under steady state, at a speed of 1800 rpm (30 Hz) for all the considered faulty conditions and the healthy one.

3 Proposed Model

3.1 Data preparation and features extraction

The measured vibration signals at the 4 bearing housings are analysed in time and frequency domains to extract the required features. Typical vibration signatures in frequency domain for the studied conditions from experimental data are shown in Figure 2.

3.1.1 Time domain features

Four scalar features in the time domain have been selected to give origin to the features vector. The first selected value is root mean square (RMS), typically used on engineering. It is defined in Equation (1), where N is the number of points contained in the sample and z_i is the *i-element* of the sample z. The remaining 3 features are variance, skewness and kurtosis. The variance, V, contains information due the signal power [1] and it is defined as it is shown in Equation (2) [10], with $\bar{z}$ the mean of the data sample. Taking $\bar{z}$ as reference value, the asymmetry of the measurement is represented by skewness, S. Its definition is in Equation (3) [11], where $E()$ denotes mathematical expectation, and σ is the standard

deviation. The last feature selected, kurtosis K, also provides information about the shape distribution of the sample and it is defined in Equation (4) [11].

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} z_i{}^2} \qquad (1)$$

$$V = \frac{1}{N}\sum_{i=1}^{N}(z_i - \bar{z})^2 \qquad (2)$$

$$S = \frac{E(z-\bar{z})^3}{\sigma^3} \qquad (3)$$

$$K = \frac{E(z-\bar{z})^4}{\sigma^4} \qquad (4)$$

3.1.2 Frequency domain features

The frequency spectrum is calculated from each data set using the fast Fourier transform (FFT). The amplitude values are extracted as the features at 0.5x,1x,1.5x …4x times the rotational speed, 1800RPM (30 Hz). The features are normalised by dividing with the correspondent values for the healthy condition. By this, the selected features became the non-dimensional parameters that represent the variations of the vibration's amplitudes from the baseline condition. The measured vibration acceleration spectra at the bearing B1 are shown in Figure 2.
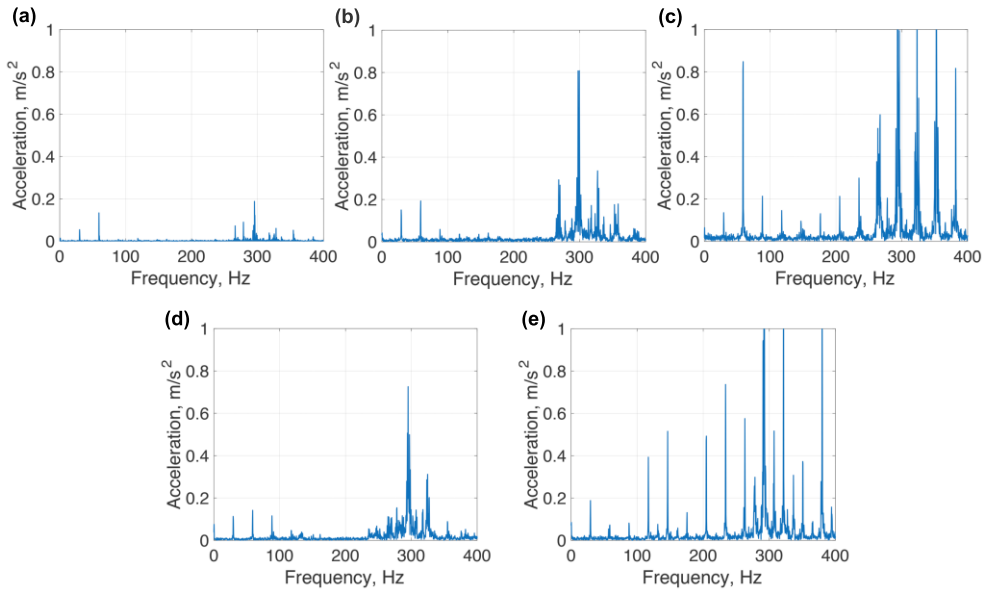
Figure 2. Typical frequency spectra at the bearing B1. a) Healthy (residual misalignment and residual unbalance). b) Misalignment (added). c) Shaft bow. d) Looseness in B3 pedestal. e) Rub.

3.2 Studied scenarios

Two scenarios regarding the measuring locations are proposed in order to compare their performances when processed using features from time and frequency domains, separately. Scenario 1 (S1) considers separately each bearing and the data acquired at them are independently processed. On the other hand, at Scenario 2 (S2), each input vector contains data acquired simultaneously from the four measuring points.

$$\mathbf{input_{S1B1T\_j}} = [RMS1_j \; V1_j \; S1_j \; K1_j] \tag{5}$$

$$\mathbf{input_{S2T\_j}} = [RMS1_j \; RMS2_j \; RMS3_j \; RMS4_j \; V1_j \; V2_j \; V3_j \; V4_j \; S1_j \; S2_j \; S3_j \; S4_j \; K1_j \; K2_j \; K3_j \; K4_j] \tag{6}$$

For instance, the input in Equation (5) correspond to one made from data taken at B1 with features from time domain, and similarly are built the ones at B2, B3 and B4. Inputs with this configuration are used at Scenario 1 and processed separately for each location. The input shown in Equation (6) corresponds to time domain features at Scenario 2, when components from the four locations make the vectors to feed the network. In an analogue manner, vectors with features from frequency domain are built. Thus, in Equation (7) there is an example when data come from B1 in Scenario 1, whilst Equation (8) corresponds to an input in Scenario 2 within frequency domain features.

$$\mathbf{input_{S1B1F\_j}} = [FF1_{0.5x,\,j} \; FF1_{1x,\,j} \; FF1_{1.5x,\,j} \; FF1_{2x,\,j} \; FF1_{2.5x,\,j} \; FF1_{3x,\,j} \; FF1_{3.5x,\,j} \; FF1_{4x,\,j}] \tag{7}$$

$$\mathbf{input_{S2F\_j}} = [FF1_{0.5x,\,j} \; ...FF1_{4x,\,j} \; FF2_{0.5x,\,j} \; ...FF2_{4x,\,j} \; FF3_{0.5x,\,j} \; ...FF3_{4x,\,j} \; FF4_{0.5x,\,j} \; ...FF4_{4x,\,j}] \tag{8}$$

3.3 Artificial Neural Network: architecture and generalisation

In order to perform the pattern recognition and classification of vibration data from different rotor conditions, a multilayer perceptron (MLP) is used in this study. This is a feedforward network with a variable number of nonlinear neurons, *m1*, *m2*, *m3*, *m4* in each hidden layer. Signal components pass the input layer and move forward along the hidden layers, finishing with the result delivery from the decision layer. The number of neurons varies according the experiments conducted. The proposed architecture is in Figure 3, with m_i the size of the input vectors and 5 classes at the output layer, one per rotor condition.
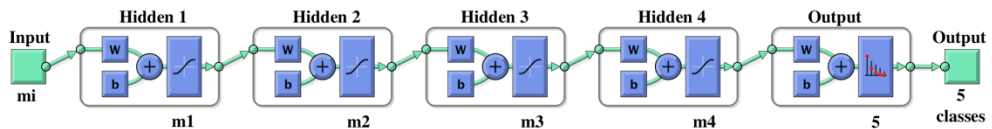
Figure 3. Neural network architecture proposed.

Supervised learning is used for training the network, thus the input data have already set to the known outputs. The weights are adjusted according to the actual error of the outputs against the desired ones during this process. The error is calculated through the performance function, being used in this study the mean squared error or cross-entropy, depending on the performances observed at each proposed scenario.

The *early stopping* method is used to avoid an overfitting during the training process and to get a proper generalisation when processing. By this procedure, a cross-validation of independent data is carried out [1], dividing the total available in 3 sets, of which 70% of samples are used for training, 15% of samples for validation and the last group of unknown data (15%) is tested, providing the generalisation of the network [12].

4 Results

Overall performances for all the proposed scenarios are shown in Figure 4, and their details regarding the learning and testing stages are summarised in Table 1. The poorest performances, for both domains, are observed for Scenarios 1. There, while the best result for time domain features is at B1, for frequency features it is achieved at B2, and the p̶ ̶ ̶ ̶t̶ ̶B4̶
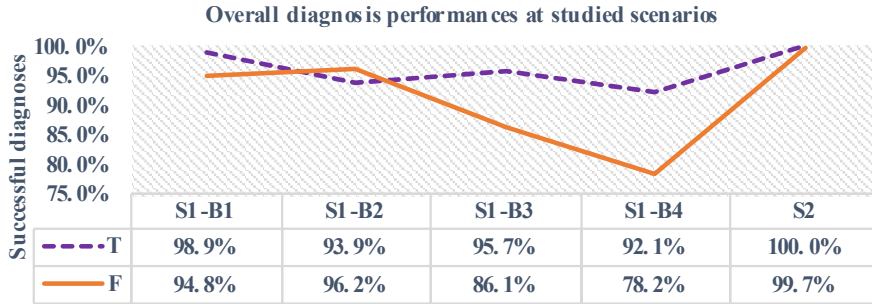

	S1-B1	S1-B2	S1-B3	S1-B4	S2
- - T	98.9%	93.9%	95.7%	92.1%	100.0%
— F	94.8%	96.2%	86.1%	78.2%	99.7%

Figure 4. Overall performances achieved in time and frequency domains.

Table 1. Performances achieved (%) at Scenarios 1 and 2, with features extracted from time (T) and frequency (F) domains.

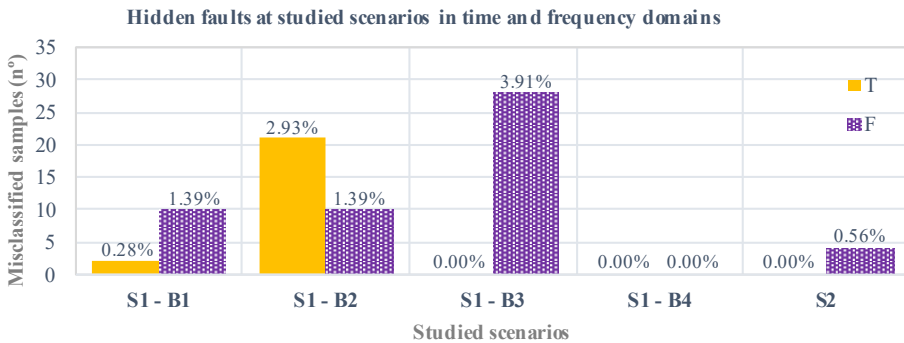|	S1								S2	
	B1		**B2**		**B3**		**B4**			
	T	**F**	**T**	**F**	**T**	**F**	**T**	**F**	**T**	**F**
Training	99.2	96.8	93.6	97.2	96.2	86.2	90.6	80.6	100.0	100.0
Validation	98.1	89.8	93.5	94.4	95.4	90.7	96.3	76.9	100.0	99.1
Testing	98.1	90.7	95.4	93.5	93.5	80.6	94.4	68.5	100.0	99.1
Overall	98.9	94.8	93.9	96.2	95.7	86.1	92.1	78.2	100.0	99.7

Figure 5. False healthy diagnoses by scenario, in time and frequency domains.

The performance of the machine learning models for Scenario 2 is found to much better for both cases (time or frequency features separately) but vibration data from all 4 bearings are considered simultaneously. Thus, the inputs in Scenarios 2 (S2) do not represent just a point of the machine, but the effects in the whole mechanism which allow better differentiation & distinction among the data sets, consequently much more accurate

classifications. The highest accuracy achieved is 100% for the time domain features compared to the frequency domain features (99.7% accuracy only). The results are also shown in Table 1. The detailed information of the misclassified samples for both S1 and S2 scenarios are shown in Figure 5.

5 Conclusions

The machine learning models based on the ANN method for the machine faults classification using the time and frequency domain features are compared. The use of the vibration data simultaneously compared to just one bearing vibration seems to be better option as expected. The data from all bearings represent the machine dynamics more accurately and hence more accurate fault identification. The machine learning model based on the time domain features is observed to be 100% accuracy compared to the frequency domain features. This simply means that the much smaller computational cost and time are required for the time domain based machine learning model. This is definitely a positive outcome for the industrial application but the proposed model needs to be tested more rigorously on different industrial machines.

References

[1] K. Worden, W. J. Staszewski, J. J. Hensman, Mech. Syst. Signal Proc., **25**, 4–111 (2011)

[2] H. Nahvi, M. Esfahanian, Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci., **219**, 141–158 (2005)

[3] N. Vyas, Mech. Mach. Theory, **36**, 157–175 (2001)

[4] J. Ben, N. Fnaiech, L. Saidi, B. Chebel-morello, F. Fnaiech, Appl. Acoust., **89**, 16–27 (2015)

[5] R. B. Walker, R. Vayanat, S. Perinpanayagam, I. K. Jennions, Mech. Mach. Theory, **75**, 54–66 (2014)

[6] A. A. Mohammed, R. D. Neilson, W. F. Deans, P. MacConnell, Meccanica, **49**, 255–266 (2014)

[7] A. D. Nembhard, J. K. Sinha, A. Yunusa-Kaltungo, J. Sound Vib., **337**, 321–341 (2015)

[8] A. D. Nembhard, J. K. Sinha, Measurement, **89**, 120–136 (2016)

[9] A. Yunusa-kaltungo, J. K. Sinha, Struct. Heal. Monit., **15**, 555–567 (2016)

[10] M. Roberts, R. Russo, *A Student's Guide to Analysis of Variance* (Routledge, 2014)

[11] R. Shanmugam, R. Chattamvelli, *Statistics for Scientists and Engineers* (Hoboken, New Jersey: John Wiley & Sons, Inc., 2015)

[12] L. Tarassenko, *A Guide to Neural Computing Applications* (Arnold, 1998)