iomac
International Operational Modal Analysis Conference

11th International Operational Modal Analysis Conference
20 - 23 May 2025 | Rennes, France

Incipient gear fault detection under varying operating speed and load via Multiple Model vibration time series methods

D.M. Bourdalos, X.D. Konstantinou, J.S. Sakellariou and S.D. Fassois

Stochastic Mechanical Systems & Automation (SMSA) Laboratory, Department of Mechanical Engineering and Aeronautics, University of Patras, 26504 Patras, Greece.
{dimitrios.bourdalos,xkonstantinou}@ac.upatras.gr, {sakj,fassois}@upatras.gr

ABSTRACT
The detection of incipient single-tooth gear faults in gearboxes operating under various speed and load conditions is investigated. The problem is of high importance as the timely detection of critical gear faults may enhance safety and significantly reduce maintenance costs. The effects of the considered faults on the observed gearbox dynamics are minor and largely masked by those effects of the various rotating speeds and loads, thus leading to a challenging fault detection problem. The study focuses on exploring the potential of two methods, both based on Multiple Model framework: A non-parametric Order Spectrum (OS) based, and a parametric AutoRegressive (AR) model-based. Both involve angular resampling of random vibration signals via computed order tracking using the tachometer signal from a reference gearbox shaft. Their performance is assessed with the gearbox in the healthy state, as well as under two distinct levels of incipient single-tooth pinion fault using thousands of experiments run under 21 rotating speeds and 4 different loads. The results demonstrate the methods' effectiveness, as well as their superiority over an alternative approach based on Wavelet Packet Decomposition.

Keywords: Gear faults, robust fault detection, vibration-based methods, statistical time series methods, incipient faults.

1. INTRODUCTION

Gearboxes play a vital role in many industrial systems by ensuring efficient and smooth power transmission. Automated Condition Monitoring (CM) of gearboxes —particularly for detecting incipient gear faults — can greatly enhance safety and reduce maintenance costs [1]. Vibration-based CM has emerged as a leading approach because it leverages vibration signals that are easily captured during normal operation and contain critical information for fault detection. However, the harsh and noisy environments in which gearboxes typically operate, combined with varying operating conditions (such as speed and load) and uncertainty, may diminish diagnostic performance and reliability. This underscores the need for methods capable of achieving high diagnostic performance under varying conditions and uncertainty [1].

The available vibration-data-based methods may be classified into two broad families: *Non-parametric* and *parametric*. The former addresses the problem using non-parametric modeling of the measured signals and the underlying gearbox dynamics, while the latter utilizes parametric modeling.

Most studies in the literature rely on *non-parametric* methods that extract various features from the vibration signals such as peak value, RMS, kurtosis, and crest factor in the time-domain [2, 3], mean frequency in the frequency domain [2], time-frequency representations such as the spectrogram [2], or other specialized features for gear fault detection [2]. In many studies signals are processed using techniques like Wavelet Packet Decomposition (WPD) [2, 4–7] or Empirical Mode Decomposition (EMD) [2] to extract features from wavelet coefficients or intrinsic mode functions, sometimes followed by dimensionality reduction techniques [4, 7]. These features are then either directly used for fault detection, or are fed into classifiers such as Support Vector Machines [3] or k-Nearest Neighbors [4] for early detection of gear cracks under different speeds and loads. More recently Convolutional and Deep Neural Networks [8–10] have also been employed for this purpose. Yet, a key drawback of non-parametric methods is their reliance on extensive data records from both healthy and faulty states; a requirement very hardly met in industrial settings. Moreover, the considered fault types and their effects on the vibration characteristics often need to be also known.

In contrast, *parametric* methods address the problem by first applying filtering and denoising techniques (such as Time Synchronous Averaging) and then identifying parametric models (such as AutoRegressive with eXogenous input, that is ARX type models) to capture the underlying dynamics. Faults are subsequently detected via model residual analysis, by examining characteristics such as variance, kurtosis, and so on. In this direction, AR [11, 12], ARX [13], or Vector AR (VAR) [14, 15] models have been employed for gear fault detection under a limited range of loads. Besides focusing only on load variation, these studies report on run-to-failure experiments without emphasizing early fault detection.

The *goal* of this study is to overcome some of the aforementioned limitations via two statistical time series type methods, one non-parametric and one parametric, both operating on angular resampled vibration signals and using a Multiple Model (MM) framework [16] and a similarity distance metric for fault detection. The non-parametric method relies on the Order Spectrum, while the parametric on AR modeling. Both are designed to maintain detection performance under varying conditions and uncertainty (robustness), while using a single accelerometer and a tachometer.

The performance of the two methods is systematically assessed via thousands of experiments on a single-stage spur gearbox. Experiments include the healthy state and two levels of incipient single-tooth pinion faults across 21 rotating speeds and 4 load levels. Results are presented via similarity distance plots and Receiver Operating Characteristic (ROC) curves [17]. A comparative analysis with an alternative method, within the same MM framework but based on Wavelet Packet Decomposition [6], is also performed. It is finally noted that the study constitutes an extension of our earlier, preliminary, work in which a parametric AR based method was applied to a simpler rotating machinery problem operating under different speeds [18, 19].

2. THE EXPERIMENTAL SET-UP

This section describes the experimental set-up employed, the vibration data acquisition, as well as the effects of different operating conditions and incipient gear faults on the vibration signals.

2.1. The gearbox, the incipient gear faults, and the vibration signals

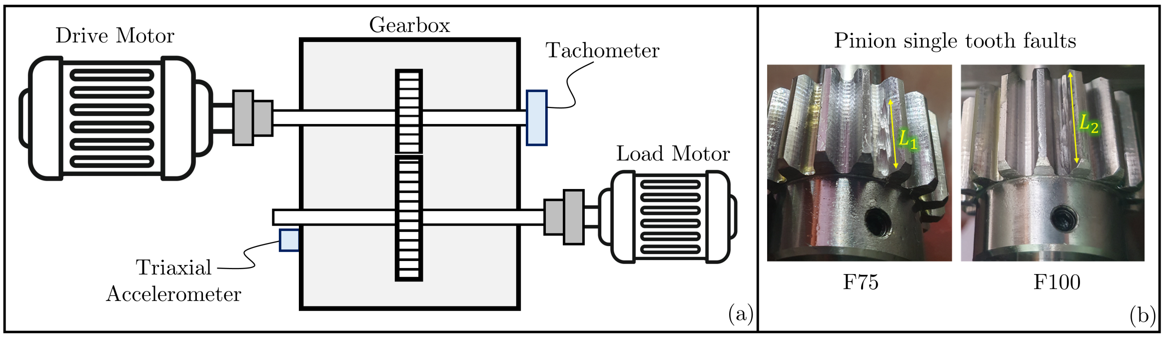

Figure 1: The experimental set-up: (a) Schematic diagram of the gearbox including the sensor (accelerometer and tachometer) locations, (b) pinion single-tooth fault scenarios $F75$ and $F100$.

Table 1: Details on the vibration signals used for the training and assessment of the fault detection methods.

Gearbox state	Nominal rotating speed (Hz)	Load	No. of signals per speed & load	No. of signals per state
Training (Learning) Phase				
Healthy (H)	$\{20, 20.25, \ldots, 25\}$	$\{0, 1, 2, 3\}$	1	84
Inspection Phase				
H / $F75$ / $F100$	$\{20, 20.25, \ldots, 25\}$	$\{0, 1, 2, 3\}$	15	1 260
Sampling frequency: $f_s = 10\ 240$ Hz; Signal length: $N = 51\ 200$ samples (5 s); Frequency bandwidth: $BWD = [0 - 5\ 120]$ Hz; Total No. of signals = 3 864				

The experimental set-up (Fig. 1 (a)) comprises a single-stage spur gearbox with a 17-tooth pinion and 34-tooth gear, driven by an AC motor at 21 nominal speeds ranging from 20 Hz (1 200 rpm) to 25 Hz (1 500 rpm) in increments of 0.25 Hz (15 rpm), using a standard inverter. A DC motor, acting as a generator, provides adjustable loading through four scenarios: Load 0, when the motor is detached, Load 1, with the motor attached but unloaded, Load 2 with a 500 W device connected and Load 3 with 1000 W device connected. An incipient fault is introduced at the base of a single pinion tooth in the gearbox using a standard Dremel-type cutting tool. The fault is implemented at two distinct levels each defined as a percentage of the total tooth face width (w) affected (see Fig. 1 (b)). The first level, corresponds to 75% of the tooth face width affected ($L_1 = 0.75w$) while the second to 100% ($L_2 = w$); these are subsequently referred to as faults $F75$ and $F100$.

The vibration signals (details in Table 1) are acquired using a single triaxial accelerometer mounted on the ball bearing housing of the secondary shaft (Fig. 1 (a)) with a sampling frequency of $f_s = 10\,240$ Hz, while a laser tachometer simultaneously measures the rotational speed of the drive motor. A total of 3 864 signals are collected per accelerometer direction, yet only the vertical one is employed. The training of the fault detection methods is based on 84 signals from the experiments with the healthy gearbox corresponding to the 2.17% of the total dataset, while the remaining 3 780 signals (97.83%) are exclusively used in the inspection phase for the methods assessment and comparison. It is worth noting that the faults are introduced directly on the gears within the gearbox without any disassembly avoiding any additional uncertainty in the measurements.

2.2. Effects of different operating conditions and incipient gear faults on the vibration signals

The effects of the different rotating speeds, loads, and faults on the vibration signals are initially analyzed using commonly employed for fault detection time domain features [4], including signal RMS, kurtosis and crest factor. The values of these features, derived from 84 vibration signals per health state (one for each combination of rotating speed and load level), are presented in Fig. 2. A distinct color and marker

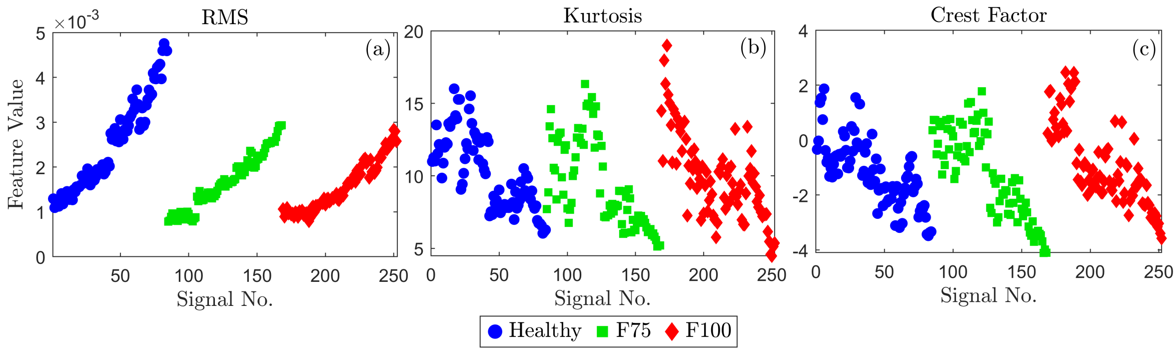

Figure 2: Time domain features obtained under the Healthy and Faulty ($F75$ and $F100$) states under all 21 speeds and 4 loads: (a) RMS, (b) Kurtosis, and (c) Crest Factor. (Each color represents a different health state; 252 signals per subplot, 84 per health state, 1 per each speed and load combination).

correspond to each health state, with the x-axis indicating the signal index and the y-axis showing the corresponding feature value. In particular, Fig. 2 displays feature values across subplots: (a) RMS, (b) Kurtosis, and (c) the Crest Factor. Notably, the feature values for both faulty and healthy gearboxes exhibit significant overlap, making fault detection particularly challenging and emphasizing the early stages of the faults.

In addition, the effects of the different rotating speeds, loads and faults on the vibration signals are examined in the order domain by constructing order spectrum zones using the Fast Fourier Transform (FFT). These zones are derived from the 84 angular vibration signals of each health state, with one signal corresponding to each speed and load combination. It is noted that the angular resampling is based on computed order tracking with a constant resampling frequency of $f_{res} = 402$ (samples per rotation) across all different rotating speeds considered (see details in Section 4.). In particular, Figs. 3(a) and (b) show a comparison of the healthy spectrum zone with those of fault level 75% ($F75$) and 100% ($F100$), respectively. Fault detection proves to be particularly challenging, as the variations in rotating speeds and loads influence the healthy dynamics in a way that obscures the fault-induced effects, causing the faulty spectrum zones to almost entirely overlap with the healthy ones across the considered bandwidth.

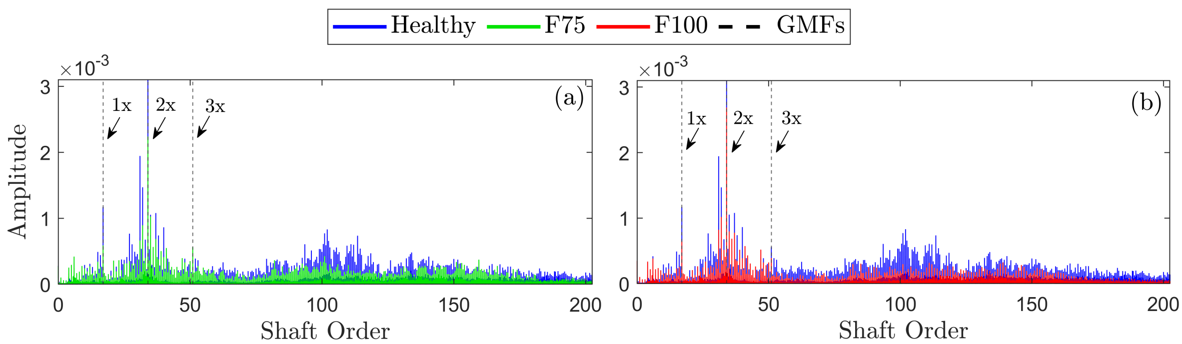

Figure 3: Order spectrum zones based on Fast Fourier Transform (FFT) estimates using a single angular vibration signal from each rotating speed and load of the healthy gearbox and comparisons per fault scenario zones (84 angular signals for each health state): (a) Healthy vs $F75$, (b) Healthy vs $F100$; the black dashed lines indicate harmonics of Gear Meshing Frequency (GMFs)

3. THE ROBUST MULTIPLE MODEL TIME SERIES FAULT DETECTION METHODS

Two robust and unsupervised Multiple Model (MM) time series methods are employed: The first is a parametric Multiple Model AutoRegressive (MM–AR) model based method and the second a non–parametric Multiple Model Order Spectrum (MM–OS) based method. Both methods are based on angular resampled vibration signals and accurate representation of the gearbox dynamics via the Order Spectrum (OS) in the first (MM–OS) method and AutoRegressive (AR) modeling in the second (MM–AR). The OS, estimated through the Fast Fourier Transform (FFT), and the AR parameter vector, estimated using standard identification procedures [20, pp.81-83], constitute the feature vector of each method, respectively. Both robust fault detection methods operate within a Multiple Model framework, which implies that multiple models, say $M_{o,i}$, $i = 1, 2, \ldots, n$ (with n designating the MM dimensionality) are used [16]. These collectively constitute the MM representation, say $\mathbb{M}_o$, of the healthy gearbox dynamics under different rotating speeds and load levels. The MM representation is built in an initial training (learning) phase using signals obtained from the healthy gearbox state operating under all considered speeds and loads.

In the inspection (real-time) phase of the method, a fresh vibration signal along with the corresponding tachometer signal are acquired with the gearbox in an unknown health state. Based on it, a corresponding model, say M_u, is estimated, and fault detection is based on determining whether or not the current model M_u belongs to MM representation $\mathbb{M}_o$. The decision-making mechanism employs a distance metric Q between the current model M_u and the representation $\mathbb{M}_o$. This is defined as the minimum of individual distances between M_u and all elements of $\mathbb{M}_o$:

$$Q := \min_i (d(M_{o,i}, M_u)), \quad i = 1, 2, \ldots, n \tag{1}$$

where $d(M_{o,i}, M_u)$ designates either the Euclidean (MM–OS method) or the Mahalanobis distance (MM–AR method). Fault detection is then declared if and only if Q is greater than a user-specified threshold l_{lim}:

$$\begin{aligned} &Q \leq l_{\text{lim}} \;\; \Rightarrow \text{Healthy Gearbox} \\ &\text{otherwise} \Rightarrow \text{Faulty} \end{aligned} \tag{2}$$

4. EXPERIMENTAL RESULTS

All vibration signals are angular resampled via computed order tracking based on the corresponding once-per-revolution tachometer pulses [21, pp. 148–151]. This process removes non-synchronous frequency modulation in the time-domain signal caused by shaft speed variations, which occur even when the nominal inverter speed is constant [22]. The angular resampling frequency, f_{res} (samples per revolution), is chosen to be the same for all rotating speeds. This selection ensures that the maximum observable order remains within the time-domain Nyquist limit while the ratio between the maximum and minimum speed does not result in significant loss of high-order spectral components at lower speeds [21, pp. 156–159]. The parameters of angular resampling, as well as those of the fault detection methods, have been selected based on preliminary analysis and are summarized in Table 2.

The performance of the fault detection methods is assessed based on 1 260 signals from the healthy gearbox and 1 260 from each fault level, with the gearbox operating under 21 different speeds and 4 different load levels (see Table 1). It is noted that the tachometer signal is considered available during the inspection (real-time) phase of the method, while the load level is considered as completely unknown. The Q distance metric of both methods for all considered inspection signals (3 780) are presented in Fig. 4 via scatter type plots and Receiver Operating Characteristic (ROC) curves [17]. Each ROC curve represents the True Positive Rate (TPR, that is the correct fault detection rate) versus the False Positive Rate (FPR, that is the false alarm rate) for varying decision threshold.

Based on the values of the Q distance metric (Fig. 4 (a)) of the MM–AR based method, it is evident that signals corresponding to the healthy gearbox are almost fully separated from those obtained from the considered faults indicating excellent fault detection performance. These results are also validated by the corresponding ROC curves in Figs. 4 (d) and (e) which indicate 100% TPR at less than 1% FPR for both fault levels.

The fault detection results based on the values of the Q distance metric of the MM–OS method (Fig. 4 (b)) indicate this method's also high performance, yet, slightly inferior compared to that of the MM–AR, as there is minor overlap among the values corresponding to the healthy and faulty gearbox. This is also confirmed by the corresponding ROC curves in Figs. 4 (d) and (e) that shows 98% and 99% TPR at 5% FPR for fault level 75% and 100%, respectively.

Finally, an interesting comparison with an alternative method is presented, which is motivated by a study showing promising performance on gear fault detection under varying speeds and loads [6]. The method employs the coefficients' Energy Distribution (ED) of the Wavelet Packet Decomposition (WPD) within a certain node as feature, and the Jensen-Rényi Divergence metric to quantify the distance between two distinct EDs. For a fair comparison, the above method is slightly modified to operate within the presented MM framework (and is thus subsequently referred to as the MM-WPD based method). The level and the node of the WPD are selected to be to 4 and 13, respectively, while the α parameter of the Jensen-Rényi Divergence is set to 0.4 (see [6] for definitions and details). The fault detection results based on the values of the Q distance metric of the MM–WPD method are depicted in Fig. 4 (c) and based on them it is evident that it offers no improvement in detection performance, as it significantly lags behind both MM–AR and MM–OS methods, which is further confirmed by the corresponding ROC curves in Figs. 4 (d) and (e) that shows only 63% and 90% TPR at 5% FPR for fault level 75% and 100%, respectively.

5. CONCLUSIONS

The robust detection of two levels of incipient single-tooth gear fault in a gearbox operating under 21 distinct rotating speeds and 4 distinct load levels has been explored via two unsupervised vibration-based statistical time series methods. Both employ angularly resampled vibration signals and operate within a Multiple Model framework with fault detection being based on a similarity distance metric. The first method (MM-OS) employs the Order Spectrum as feature, while the second (MM-AR) the AR model parameter vector.

Based on the fault detection results obtained from 3 780 experiments, it is concluded that the MM-AR method achieves excellent detection reaching 100% TPR for 1% FPR and outperforming its MM-OS counterpart (98% TPR, 5% FPR). The MM-WPD based alternative methods lags behind (TPR $\leq$ 90%, 5% FPR). Both methods are trained using data exclusively from the healthy gearbox, without any knowledge of the considered fault, or its effects on the vibration characteristics, being thus suitable for unsupervised detection of any type of fault in rotating machinery under different speeds and loads. On the contrary, the MM-WPD method's hyperparameter selection depends upon determining the frequency bandwidth where the fault manifests (see [6] for details). Finally, as all methods operate based on a single accelerometer and a tachometer, and are easy to implement, with low computational complexity in the Inspection Phase, renders them promising candidates for automated real time implementation.

Table 2: Details on the fault detection methods.

Method	Feature vector	Feature Dimensionality	Distance type
MM-AR	AR parameter vector	250	Mahalanobis
MM-OS	Order Spectrum	19 900	Euclidean
MM-WPD	ED[1] of Wavelet Coefficients	19 217	Jensen-Rényi

AR modeling: Estimation via OLS [20, p. 204] (*Matlab function:* `ar.m`); Selected order: $n_a = 250$; BIC = -17.23; Condition Number = 1.92×10^5

Angular Resampling: $f_{res} = 402$ samples per shaft rotation[2]; Number of shaft rotations = 99; Length = 39 798 samples; Order spectrum resolution = 0.01

Wavelet Packet Decomposition: Mother wavelet: db10; WPD level = 4; Selected terminal node: 13th; Matlab function: `wpdec.m`; Jensen-Rényi divergence: $\alpha = 0.4$

[1]*Energy Distribution (see [6] for details).* [2]*Constant across all different speeds.*

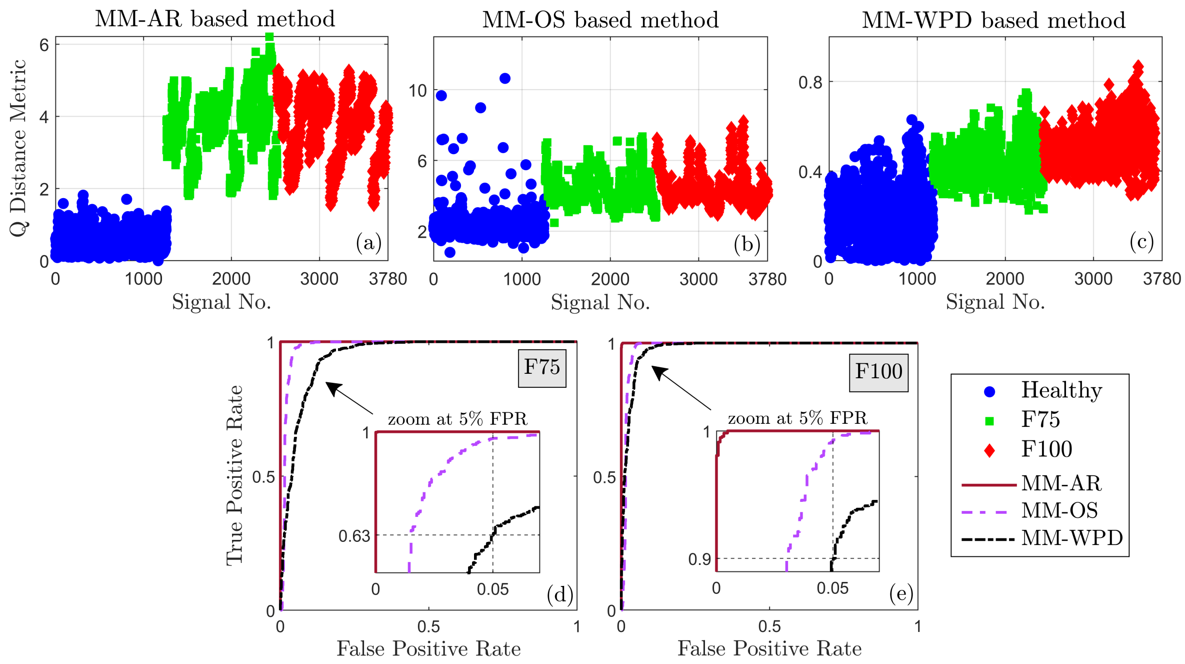

Figure 4: Fault detection results (Inspection Phase): Q distance metric values for the (a) MM-AR, (b) MM-OS, and (c) MM-WPD methods (upper row). Corresponding ROC curves for the (d) $F75$ and (e) $F100$ fault scenarios (lower row). (3 780 Inspection experiments per method, uniformly obtained under all combinations of 21 rotating speeds and 4 loads.)

ACKNOWLEDGMENT

This study has been carried out as part of the 'EEDRIVEN' project implemented in the framework of the H.F.R.I call 'Basic Research Financing (Horizontal support of all Sciences)' under the National Recovery and Resilience Plan 'Greece 2.0' funded by the European Union – NextGenerationEU (H.F.R.I. Project Number 16381).

REFERENCES

[1] K. Feng, J. Ji, Q. Ni, and M. Beer. A review of vibration-based gear wear monitoring and prediction techniques. *Mechanical Systems and Signal Processing*, 182:109605, 2023.

[2] Y. Lei, M. J. Zuo, Z. He, and Y. Zi. A multidimensional hybrid intelligent method for gear fault diagnosis. *Expert Systems with Applications*, 37(2):1419–1430, 2010.

[3] J. Xie, L. Zhang, L. Duan, and J. Wang. On cross-domain feature fusion in gearbox fault diagnosis under various operating conditions based on transfer component analysis. In *Proc. IEEE International Conference on Prognostics and Health Management*, Ottawa, Canada, 2016.

[4] D. Wang. K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: Revisited. *Mechanical Systems and Signal Processing*, 70:201–208, 2016.

[5] S. Guo, T. Yang, H. Hua, and J. Cao. Coupling fault diagnosis of wind turbine gearbox based on multitask parallel convolutional neural networks with overall information. *Renewable Energy*, 178: 639–650, 2021.

[6] P. Boškoski and D. Juričić. Fault detection of mechanical drives under variable operating conditions based on wavelet packet rényi entropy signatures. *Mechanical Systems and Signal Processing*, 31: 369–381, 2012.

[7] X. Wan, D. Wang, P.W. Tse, G. Xu, and Q. Zhang. A critical study of different dimensionality reduction methods for gear crack degradation assessment under different operating conditions. *Measurement*, 78:138–150, 2016.

[8] J. Koutsoupakis, P. Seventekidis, and D. Giagopoulos. Machine learning based condition monitoring for gear transmission systems using data generated by optimal multibody dynamics models. *Mechanical Systems and Signal Processing*, 190:110130, 2023.

[9] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang. Deep model based domain adaptation for fault diagnosis. *IEEE Transactions on Industrial Electronics*, 64(3):2296–2305, 2017.

[10] M. Zhang, W. Lu, J. Yang, D. Wang, and L. Bin. Domain adaptation with multilayer adversarial learning for fault diagnosis of gearbox under multiple operating conditions. In *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, pages 1–6, 2019.

[11] W. Wang and A. K. Wong. Autoregressive model-based gear fault diagnosis. *Journal of Vibration and Acoustics*, 124:172–179, 2002.

[12] Y. Zhan and C. K. Mechefske. Robust detection of gearbox deterioration using compromised autoregressive modeling and kolmogorov-smirnov test statistic. part ii: Experiment and application. *Mechanical Systems and Signal Processing*, 21(5):1983–2011, 2007.

[13] M. Yang and V. Makis. Arx model-based gearbox fault detection and localization under varying load conditions. *Journal of Sound and Vibration*, 329(24):5209–5221, 2010.

[14] C. Lin and V. Makis. Application of vector time series modeling and t-squared control chart to detect early gearbox deterioration. *International Journal of Performability Engineering*, 10:105, 2014.

[15] X. Li, H. Zuo, P. Hao, Y. Su, H. Liu, and C. Xue. Early fault detection of gearbox using tsa and var model considering load variation. In *2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing)*, pages 1–6, 2021.

[16] K.J. Vamvoudakis-Stefanou, J.S. Sakellariou, and S.D. Fassois. Vibration-based damage detection for a population of nominally identical structures: Unsupervised multiple model (mm) statistical time series type methods. *Mechanical Systems and Signal Processing*, 111:149–171, 2018.

[17] R. Duda, P. Hart, and D.G.Stork. *Pattern Classification*, pages 34–35. John Wiley and Sons, 2001. ISBN 0-471-05669-3.

[18] D.M. Bourdalos, I.A. Iliopoulos, and J.S. Sakellariou. On the detection of incipient faults in rotating machinery under different operating speeds using unsupervised vibration-based statistical time series methods. *Lecture Notes in Civil Engineering LNCE*, 254:287–296, 2023.

[19] D.M. Bourdalos and J.S. Sakellariou. Vibration-based unsupervised detection of common faults in rotating machinery under varying operating speeds. In *Proc. Surveillance, Vibrations, Shock and Noise*, Toulouse, France, July 2023.

[20] L. Ljung. *System Identification: Theory for the User*. Prentice Hall information and system sciences series. Prentice Hall PTR, 2nd edition edition, 1999.

[21] R.B. Randall. *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications*. John Wiley & Sons, Ltd, Hoboken, NJ, 2021.

[22] H. Endo and R.B. Randall. Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. *Mechanical Systems and Signal Processing*, 21(2):906–919, 2007.