Proceedings of the 6th European Conference of the Prognostics and Health Management Society 2021 - ISBN – 978-1-936263-34-9

Bearings Fault Detection Using Hidden Markov Models and Principal Component Analysis Enhanced Features

Akthem Rehab[1], Islam Ali[2], Walid Gomaa[3], and M. Nashat Fors[4]

[1,4]*Production Engineering Department, Alexandria University, Alexandria, Egypt*
Akthem.Rehab@alexu.edu.eg
Nashat.Fors@alexu.edu.eg

[2]*Department of Industrial and Manufacturing Engineering, Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt*
Islam.Ali@ejust.edu.eg

[3]*Department of Computer Science and Engineering, E-JUST, Alexandria, Egypt*
Walid.Gomaa@ejust.edu.eg

ABSTRACT

The health management of Rolling Element Bearings continues to be of increasing importance to industrial assets' productivity, reliability, and cost reduction. Early fault detection is a key pillar of health management as part of the evolving prognostics and health management philosophy. This paper proposes a fault detection method that starts by segmenting the vibration signal(s) detected from the bearing into overlapping blocks. Principal Component Analysis is then applied to the segmented signal. The combination of data segmentation and Principal Component Analysis is a signal processing approach that captures the second-order structure of the vibration signal. The method proceeds by training a Hidden Markov Model with the processed signal where k-means clustering is applied for setting the Hidden Markov Model's number of states parameter. Finally, the trained Hidden Markov Model is employed together with a goodness-of-fit test to assess the bearing health degradation by processing real-time vibration data. The method is tested on the bearing testbed dataset provided by the center for Intelligent Maintenance Systems, University of Cincinnati, OH. Experimental results show the proposed method outperforms state of the art and benchmark results of this dataset.

1. INTRODUCTION

Rolling Element Bearings (REB) are of great importance to all forms of rotating machinery and are among the most common machine elements. Consequently, REB failure is one of the leading causes of breakdowns in rotating machinery and can develop into a catastrophic failure if its deterioration is not detected and dealt with in time (H. Qiu et al., 2006). REBs are usually managed using a Condition-Based Maintenance (CBM) strategy where vibration signals analysis is widely used for bearings' Fault Detection and Diagnosis (FDD) (Liu et al., 2015; Yu, 2012a).

It is established that CBM reduces downtime and maintenance costs while it increases systems' reliability (Jardine et al., 2006; R Gopinath, CS Kumar, 2018). The evolving PHM philosophy extends CBM to optimize maintenance further; it is defined as an approach for the health management of systems based primarily on fault detection, diagnostics, prognostics, and maintenance decision-making (Soualhi et al., 2018; Vogl et al., 2019). PHM enhances CBM by further boosting detection of faults/degradation, reliability, life expectancy, and operational availability of systems while, at the same time, further decreasing Life-Cycle costs and downtime. Moreover, a key component of PHM is estimating the remaining useful life of systems. In order to realize its full potential, PHM utilizes available real-time sensory data obtained from machinery (Kalgren et al., 2006; Walker & Coble, 2018).

Fault detection oversees the identification of whether the monitored component or process is properly working or not. It is the stepping-stone for diagnostics and prognostics. Data-driven fault detection often requires gathering data for normal operating conditions to compare them with actual operation. Normal operating conditions data is often abundant; hence, fault detection using data-driven methods is usually within reach compared to diagnosis and prognosis that require gathering faulty conditions data, which is difficult in highly reliable and critical systems (Arpaia et al., 2020). Semi-supervised deep learning approaches are being investigated to overcome the issue of faults data scarcity in diagnosis as they allow for effective utilization of datasets when only a small subset of data has labels (Zhang et al., 2019).

Historically, many data-driven methods have been proposed for FDD (Md Nor et al., 2020). Data-driven methods are usually trained on features extracted from the original data. For REBs, time and frequency domain features are usually obtained from the signals generated from the vibration sensors and are used to determine the REBs' condition (Yu, 2012a, 2012b).

On the other hand, Deep Learning (DL) methods have been employed in FDD more recently (Khan & Yairi, 2018; Saufi et al., 2019). While DL largely automates feature extraction and offers potential benefits for FDD, the reviews in (Khan & Yairi, 2018; Saufi et al., 2019) argue that it is encumbered with a number of challenges that hinder its adoption as the modelling technique of choice for FDD in place of data-driven methods.

Subsequently, in this paper, we propose a data-driven method for the REBs fault detection. The method is based on a Hidden Markov Model (HMM) trained by principal component features. These principal components are generated by applying Principal Component Analysis (PCA) to the segmented vibration signal - a method originally proposed for processing Electroencephalogram (EEG) signals in the Brain-Computer Interface (BCI) domain (H. Lee & Choi, 2003). The performance evaluation presented in this paper shows that using Data Segmentation (DS) and PCA for vibration signal processing together with HMMs yields promising results and outperforms some state-of-the-art DL and data-driven methods used for fault detection/degradation propagation in an REB (Hasani et al., 2017; Yu, 2012a).

The remainder of this paper is organized as follows: Section 2 discusses deep learning challenges leading to the decision of using established data-driven models; Section 3 presents the proposed fault detection approach and details the signal processing and training method; Section 4 presents the results of applying the proposed approach to the Intelligent Maintenance Systems' (IMS) bearing dataset and compares it to state of the art methods; and finally, Section 5 concludes the paper and highlights potential future extensions to this work.

2. State of the Art

2.1. Deep learning

Recently, Deep Learning (DL) has advanced to the point of being among the state of the art in Artificial Intelligence (AI). The DL potential benefits of ease of classification and unsupervised feature learning are among the main reasons behind this development. This advancement has led to studying the application of DL in the FDD domain in recent literature and reviews (Khan & Yairi, 2018; G. Qiu et al., 2019; Saufi et al., 2019).

DL models that have been used in FDD systems include Convolutional Neural Network (CNN), Stacked Auto-Encoder (SAE), Restricted Boltzmann Machine (RBM), Deep Belief Network (DBN) and Deep Neural Network (DNN) (G. Qiu et al., 2019). For the advantages and disadvantages of DL models used in FDD, the reader is referred to (Khan & Yairi, 2018). The reviews reveal that, regardless of the DL model used, while DL demonstrates plausible benefits for FDD, it faces several challenges that impede its mass application (Khan & Yairi, 2018). These challenges are linked mainly to the DL architecture and training process and include the choice of hyperparameters, availability of large datasets, input size dimensionality, and data cleanliness (Saufi et al., 2019).

For the last few decades, the deep architecture of the DL models has been difficult to train. This difficulty is attributed partially to the DL hyperparameters. Manual selection of hyperparameters is difficult, and determining optimal hyperparameters' values is time-intensive and complex (Ma et al., 2018; Saufi et al., 2018).

Another challenge to DL adoption in FDD is the need for large datasets to train the DL models, as the capability of these models to perform FDD with a small number of samples has not been determined. Moreover, DL models' ability to handle more testing than training samples is doubtful. Conversely, most DL models have been trained with 50% more training than testing samples. In addition, obtaining enough data samples in industrial environments is often difficult to achieve, especially for highly reliable and critical systems (Arpaia et al., 2020; Chen et al., 2018; Shriram Ramanathan, 2018).

One more challenge that should be considered is the input dimension size. As per (Shao et al., 2016), the training time of a DL model significantly increases with an increase in input dimensions. Sometimes, the signal length, segmented or raw, equals the input dimension as several DL models are directly fed with a segmented signal in the raw time domain.

Furthermore, data cleaning is critical to ensuring good performance of DL models in practical conditions (Gheisari et al., 2017) as noise in the input signal reduces the accuracy of DL models (Zhong et al., 2019). This reduction in accuracy was observed by (Saidi et al., 2017; Teng et al., 2017) while performing fault diagnosis on real-world wind turbine gearboxes, and it was concluded that it is difficult for DL to achieve satisfactory diagnosis performance in real-world applications.

All the above implies that although DL is advantageous for unsupervised feature-learning, it still needs more development to usurp the more common and established data-driven models like Random Forest, Support Vector Machine (SVM), Bayesian networks, and HMMs for machinery health monitoring. These more established models are very good at generalization and produce high accuracy classification and regression for FDD (Khan & Yairi, 2018).

2.2. Established data-driven models

Data-driven approaches usually apply feature extraction and/or signal processing right after data acquisition. For REBs, the four stages of REB failure are defined by frequency features in the frequency domain processed from vibration signals (Bently Nevada, 2019; Berry, 1996). Frequency domain features attempt to find the characteristic frequencies related to the rotation of REBs like the Ball-Pass Frequency of Outer ring (BPFO) and the Ball-Pass Frequency of Inner ring (BPFI) for REB health degradation monitoring.

The frequency domain analysis shows what frequencies are present in a signal, but it would not show when the frequencies have occurred if this signal was changing with time. Hence, the time-frequency domain analysis is used to resolve this issue and show how frequency changes with time. Wavelet transform is usually the tool of choice for time-frequency domain analysis. Vibration signals are also processed in the time domain. Time domain features are usually sensitive to impulsive oscillations (Yu, 2012a, 2012b).

The feature extraction and signal processing options are plenty, and choosing proper statistical features containing useful information is challenging (Hasani et al., 2017). Many health monitoring methods attempted to address this challenge (Yu, 2012a, 2012b). Furthermore, techniques like PCA-based methods followed by a post-processing stage such as HMM for health degradation monitoring provided a reasonable prediction and an attractive accuracy on the system's status (Hasani et al., 2017; Yu, 2012a). HMM based methods could be generalized, albeit only after subjecting them to modifications, to other test cases (Hasani et al., 2017; Tobon-Mejia et al., 2012).

A review of data-driven FDD methods is provided in (Khan & Yairi, 2018; Md Nor et al., 2020; Saufi et al., 2019), covering the different data-driven FDD frameworks, including machine learning, signal-based, and knowledge-based methods. The review also summarizes the benefits and challenges of data-driven FDD implementations. The review highlights HMMs power in handling sequential data, scalability, ability to model highly non-linear problems, high classification accuracy, and HMMs being one of the modelling approaches adopted for multimode process monitoring due to their robust stochastic and inferential features. Moreover, in (Arpaia et al., 2020), HMMs are considered a promising trend for research in the field of PHM and are exploited for fault detection in real-world conditions where readings of multiple physical measurements available are sparse and asynchronous.

PCA can be used in several ways for feature extraction in combination with HMM. Recent work on HMM-based fault detection (Arpaia et al., 2020) employed PCA for dimension reduction by projecting the variables in a new space and using the principal components to train the HMM. (Yu, 2012a) applied Dynamic PCA (DPCA), an extension of PCA, to apply dimensionality reduction to the processed features and at the same time consider serial correlations in the processed signal.

PCA is also used in the context of other domains like Brain-Computer Interface (BCI). (Ozg, 2010) generated autoregressive (AR) parameters from Electroencephalogram (EEG) data and used PCA for dimension reduction before training the HMM. (H. Lee & Choi, 2003) applied EEG DS and PCA to find principal component features that capture the second-order statistical structure of the data. The proposed method in (H. Lee & Choi, 2003) does not require the manual selection of statistical features, overcoming a key drawback of data-driven approaches compared to DL. While the method still has hyperparameters needing to be set, they are arguably fewer than other DL methods. Furthermore, combining this signal processing method with HMM allows for greater generalization potential and benefits from the HMM scalability and decreases the needed training data compared to DL.

3. Proposed Approach

This paper proposes a fault detection method based on DS, PCA, and HMM (DS-PCA-HMM). First, the vibration signal is segmented into overlapping blocks, and PCA is applied to the segmented signal for the purpose of signal processing and selecting the principal components. Next, the principal components are clustered using k-means clustering to set the parameter of the HMM number of states. After this, the Baum-Welch algorithm is applied to train the HMM and specify the HMM parameters λ. Finally, the trained HMM is employed for fault detection through a goodness-of-fit test. The test utilizes the forward-backward procedure to estimate the online/test sample Log-Likelihood (LL) and then compare it to the training samples LL distribution. The proposed method is sketched in Figure 1.

3.1. Signal processing

The DS-PCA-HMM approach is initiated by applying a data segmentation procedure followed by PCA for signal processing, a method originally developed for EEG signals in the BCI domain (H. Lee & Choi, 2003). Since PCA retains maximum variance, the method is expected to provide features that are robust to noise. Additionally, (H. Lee & Choi, 2003) noted that the PCA, when applied to the segmented data, learns basis functions that look similar to wavelet basis functions.

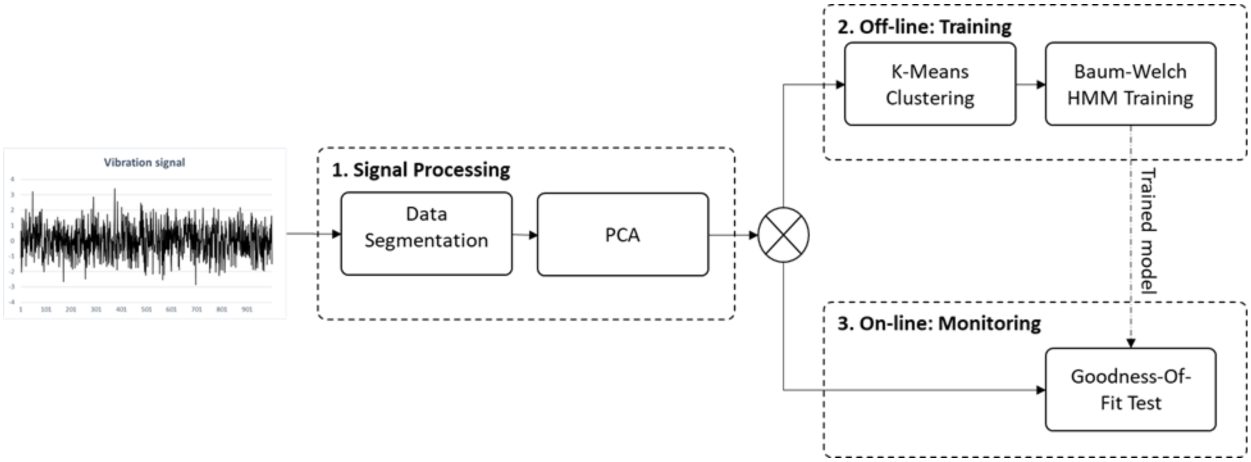

Figure 1 REB Fault detection proposed approach

3.1.1. Data segmentation

The vibration signal is decomposed into N overlapping blocks where principal components are extracted from each block. The blocks are used to construct an M x N data matrix (where M is the data block/window time length). The block window length and the overlap length are parameters to be optimized in the model. The DS procedure is sketched in Figure 2.

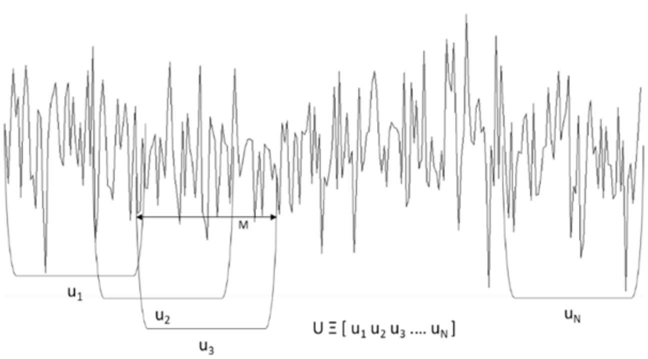

Figure 2 Overlapping windows grouped in matrix U

3.1.2. PCA

PCA is mainly a dimensionality-reduction method. It applies orthogonal linear transform to map the data to a new eigenspace, such that maximum variance is retained in the main components (Jollife & Cadima, 2016; Jolliffe, 2002).

Eigen decomposition is used to calculate eigenvalues and eigenvectors. Letting u be the observation vector and the data covariance matrix be R_u in which

$$R_u = U_u D_u U_u^T \qquad (1)$$

Where U_u is the eigenvector matrix of R_u and D_u is the corresponding diagonal matrix of eigenvalues. Hence, letting $W = U_u^T$, the orthogonal linear transform, v, of u is achieved by

$$v = W u \qquad (2)$$

Dimension reduction is achieved by ordering eigenvalues and selecting the corresponding p columns of U_u where $p < M$. W is reconstructed to the p x M matrix of the p dominant column vectors in U_u.

PCA is applied to the decomposed time series matrix U = M x N (see Figure 2) and is used to find a smaller p by M matrix W for PCA. In the case of this study, W matrix is calculated for each of the directions where vibration is measured, i.e., W_x and W_y. (X and Y denoting vibration measurement directions) for the tests conducted with measurements in X and Y directions. The principal component features vector for the vibration signal in each direction is then computed by $v_n = W_{Vn.} u_n$ (where n stands for X or Y directions). Principal component features extracted from each direction are then concatenated (H. Lee & Choi, 2003).

3.2. Offline training

The concatenated principal component features generated from the signal processing step are utilized in training the HMM. HMMs are used for sequential data in various fields, most notably in speech recognition (Bishop, 2006; Rabiner, 1989). HMMs are constituted of a defined number of states; the state the model is in at a specific point of time generates an observation. The sequence of states of an HMM is not observable directly, i.e., hidden. Five elements characterize an HMM:

1- **S**, A defined number of states;
2- **A**, the state transition probability matrix;
3- **O**, the number of distinct observation symbols per state;
4- **B**, the observation symbol probability matrix;
5- $\boldsymbol{\pi}$, the initial state distribution.

Given an observation sequences $O = \{O_1, O_2, O_3 \ldots O_T\}$, a complete specification of an HMM requires specifying the model parameters $\lambda = \{A, B, \pi\}$, state sequence $S = \{S_1, S_2, S_3 \ldots S_T\}$, and the likelihood of the observation sequence given a model $P(O|\lambda)$.

Before specifying the HMM model parameters λ, the number of hidden states in an HMM should be decided first. There is no straightforward method to optimize the HMM number of states choice (Zec et al., 2018). The traditional model selection methods like the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) have their drawbacks as they often recommend complex models with a large number of states (Pohle et al., 2017). (Zec et al., 2018) used empirical analysis to overcome the challenge of determining the proper number of states. In this paper, k-means clustering is applied to the concatenated principal component features generated from the signal processing step, and the optimum number of clusters is set to be the number of states of the HMM.

3.2.1. K-means clustering

One of the most prevalent clustering methods is the k-means. The k-means algorithm works as follows:

1. Initialize the algorithm by selecting initial centroids for each cluster;
2. Given an initial clustering of the data, relocate each point to its new nearest center;
3. Update the clustering centers by calculating the mean of the member points;
4. The relocating-and-updating procedure is repeated until a convergence criterion is met (such as there is no further change in the assignments, or a predefined number of iterations is reached).

An example of k-means clustering with K = 2 is shown in Figure 3 and for more details, the interested reader is referred to (Bishop, 2006; Mannor et al., 2011).

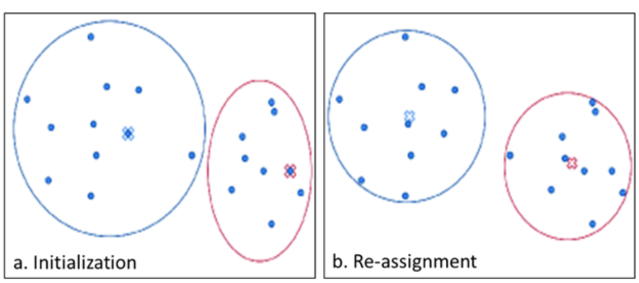

Figure 3 K-means example (K = 2) (Mannor et al., 2011)

The k-means algorithm assumes that the number of clusters K is known/given. This assumption does not hold in this study, and hence a K-value selection algorithm, namely the Elbow Method, is utilized to find the optimum number of K clusters (Thorndike, 1953; Yuan & Yang, 2019). The main idea behind the Elbow Method is to use the square of the distance between the sample points in each cluster and the centroid of the cluster. The squared distances are summed for each K to give a series of Sum of Squared Errors (SSE) that is used as a performance indicator for each K. Iterating over the K-value and the corresponding SSE, smaller values indicate that each cluster is more convergent (Yuan & Yang, 2019).

The SSE rapidly declines when the number of clusters K approaches the real number of clusters in the data while opposite behaviour occurs when K begins to exceed the real number of clusters in the data. SSE will continue to decline but at decreasing pace. The K value can be better determined by plotting the K-SSE curve and finding the inflexion point (Yuan & Yang, 2019).

3.2.2. HMM training (Baum-Welch)

After determining the number of clusters of the principal component features, the HMM parameters λ are specified using the Baum-Welch algorithm, a special case of the Expectation-Maximization (EM) algorithm. The Baum-Welch algorithm solves the problem of HMM training. It is used to find the HMM parameters $\lambda = \{A, B, \pi\}$ that maximize the likelihood $P(O|\lambda)$ (Bishop, 2006; Munro et al., 2011).

Once the model is trained with the training samples and λ is specified, the likelihood of each training sample given the trained model is obtained using the forward procedure, a special form of the forward-backward procedure. Given an HMM model $\lambda = \{A, B, \pi\}$ and input observation sequence $O = \{O_1, O_2, O_3 \ldots O_T\}$, the forward-backward procedure computes the sequence's likelihood given the model $P(O|\lambda)$. The forward-only variant sums the forward recursion algorithm run on the entire sample observation sequence for all states and returns the likelihood of the monitored sample (Bishop, 2006).

The likelihoods of the training samples are then subjected to standardization through calculating their Z-Scores.

3.3. Online monitoring

The online/test samples are subjected to the same signal processing procedure that the training samples went through: DS followed by PCA. Then the sample's likelihood with respect to the trained model $P(O|\lambda)$ is obtained by applying the forward procedure to the processed signal. The Z-score of the online/test sample likelihood is calculated next with respect to the same distribution of the training samples.

3.3.1. Goodness-of-fit test

The Z-score of the online/test sample is compared to the training samples distribution. This comparison is employed to decide if the sample indicates the REB is still operating in normal condition. Assuming the training samples are independent and are (approximately) normally distributed, then 99.73% of the training samples should have a Z-score between +/- 3. Hence, the online/test sample Z-score threshold is set to -3. A threshold of a +3 Z-score is not set as it indicates an improvement of LL. When the online/test samples start to score a Z-score < -3, the sample is considered anomalous and deviating from the training samples distribution. The REB is then considered to be entering into a faulty mode and experiencing health degradation (Arpaia et al., 2020).

4. EXPERIMENTAL RESULTS

This section of the paper evaluates the performance of the DS-PCA-HMM method by employing it in several run-to-failure bearing experiments. The IMS Bearing dataset contents and objective are introduced first. Then the dataset is subjected to signal processing, training, and online monitoring as per the proposed approach. The performance of DS-PCA-HMM is illustrated in various tests, and finally, the results are benchmarked with state of the art fault detection methods applied to the IMS Bearing dataset.

4.1. IMS Bearing dataset

The dataset provided by the center for IMS, University of Cincinnati, OH (J. Lee et al., 2007) is used in this paper. The bearing test rig having four double row bearings mounted on one shaft is sketched in Figure 4. The rotation speed is constant at 2000 RPM, and the shaft is rotated by an AC motor coupled to the shaft through rub belts. A radial load of 6000 lbs. is applied onto the shaft and bearing by a spring mechanism. All bearings are force lubricated.

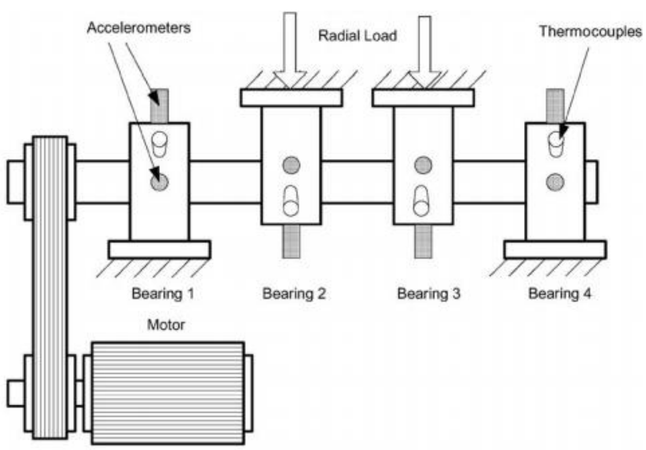

Figure 4 Bearing test rig and sensors placement (J. Lee et al., 2007)

A High Sensitivity Accelerometer is installed on each bearing housing and vibration data have been collected every 10 minutes by a data acquisition card with a data sampling rate of 20 kHz and data length of 20480 points per sample. The Prognostics Center of Excellence (PCoE) of NASA shared through their prognostic data repository the data of three test-to-failure experiments performed independently (H. Qiu et al., 2006).

The first experiment utilized two accelerometers for each bearing, while the second and third experiments utilized only one accelerometer. Table 1 summarizes the properties of the collected data in each experiment.

Table 1 Specification of the IMS Bearing data

Experiment	# of Samples	Sample Size	Faulty Bearing(s)
T1	2156	4 X 20480	B3 and B4
T2	984	4 X 20480	B1
T3	6324	4 X 20480	B3

The proposed approach is only trained on the faulty bearings. It is worth noting that T3 number of samples in the downloaded dataset (J. Lee et al., 2007) is found to be differing from that of other benchmark studies and the readme file provided with the datasets (Hasani et al., 2017; J. Lee et al., 2007), and hence the proposed approach performance will not be applied to the T3B3 bearing. This study has three different simulated experiment cases as follows:

1- Dataset 1 bearing 3 (T1B3)

2- Dataset 1 bearing 4 (T1B4)

3- Dataset 2 bearing 1 (T2B1)

4.2. Signal processing

The DS-PCA-HMM approach is initiated by applying signal processing to each bearing. Starting with DS, the window size and overlap are optimizable parameters that need to be set. In this study, a window size of 512 readings and overlap every 32 readings is set. The window size and overlap were chosen as factors of the 20480 data length to avoid having overlapping windows that are partially filled with data at the end of the sample. The PCA is run next, and the first ten eigenvectors corresponding to the largest ten eigenvalues, i.e., setting p = 10, are kept. It is worth noting that the first two Eigenvalues were always retaining more variance, but due to the large size of variables, ten principal components have been chosen to represent each window. The ten main components retain about 20% of the vibration signal variance in each direction and compress the original vibration signal at the same time. More options and permutations are possible for the number of eigenvalues p, window size, and overlap, and their optimization could be the subject of a follow-on study.

4.3. Offline training

Transitioning to the HMM training, the first one-third of the bearing's whole life is chosen as the healthy dataset used for training. It is further assumed that the first one fifth of this portion of data to be early-stage operation and is excluded from the training data. Healthy state readings from T1B3, T1B4, and T2B1 selected for training are further detailed in Table 2.

Table 2 Healthy state readings used for training

	T1B3	T1B4	T2B1
Readings	2156	2156	984
1/3rd the readings	Readings 1-719	1-719	1-328
Omit early 20%	Readings 144-719	144-719	65-328

An HMM is then trained for each bearing experiment on its selected main components. The training data is clustered, and the HMM number of states is set to equal the number of clusters. Figure 5, Figure 6, and Figure 7 show the elbow graph used to determine the optimum number of clusters for T1B3, T1B4, and T2B1, respectively.

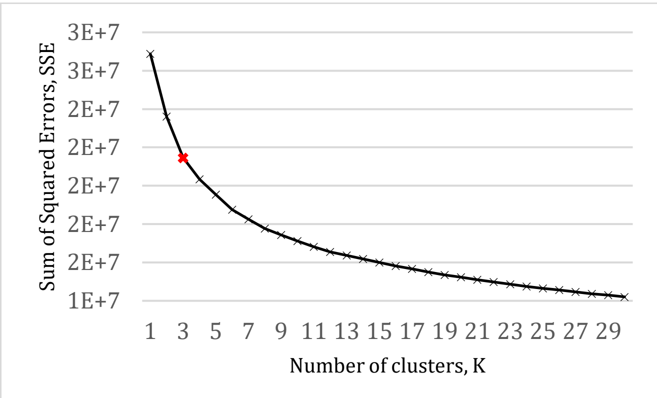

Figure 5 T1B3 SSE-K circling optimal k at 3 clusters

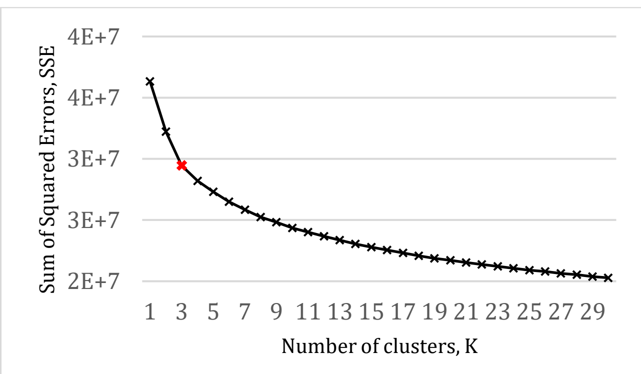

Figure 6 T1B4 SSE-K circling optimal K at 3 clusters

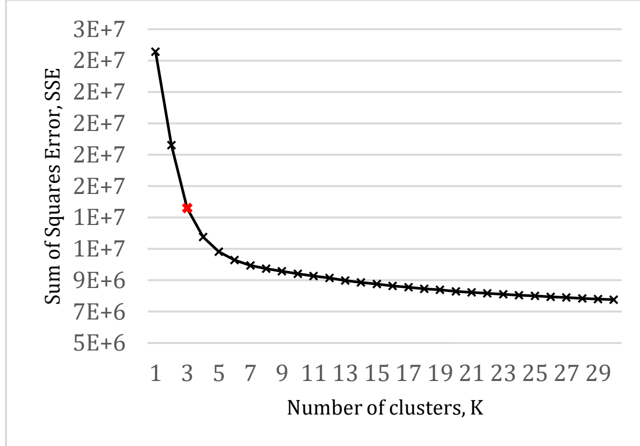

Figure 7 T2B1 SSE-K circling at 3 clusters

The Baum-Welch algorithm is initialized for each bearing, and the model parameters λ are specified after determining the optimum number of clusters for each bearing dataset.

4.4. Online degradation monitoring

Fault detection follows the training of the model. The Z-score is calculated for the samples representing the entire life of the bearing to assess each sample's goodness-of-fit compared to the training samples distribution. The Z-score online monitoring results based on the DS-PCA-HMM method are illustrated in Figure 8, Figure 9, and Figure 10 for T1B3, T1B4, and T2B1 runs to failure, respectively.

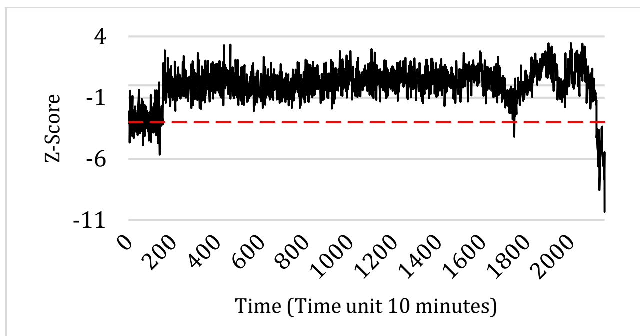

Figure 8 Health monitoring for the full lifetime of T1B3 (fault incipience at 2120)

The T1B3 bearing starts in an abnormal behaviour which could be attributed to the early-stage operation. Then it operates almost smoothly except for one outlier sample at the reading 1747 before returning to normal Z-Scores. The degradation of the bearing starts to occur at the reading 2120 and continues till its failure.

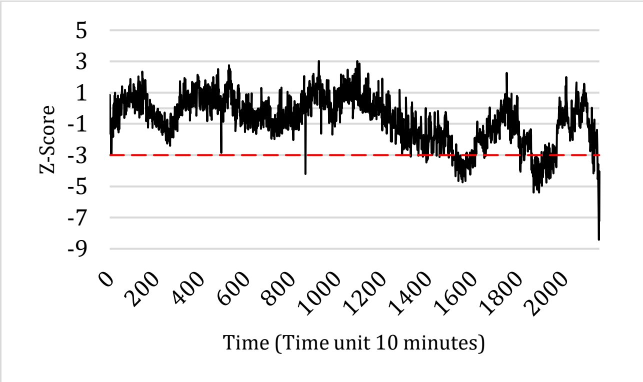

Figure 9 Health monitoring for the full lifetime of T1B4 (Fault incipience at 1508)

T1B4, on the other hand, does not show abnormal startup behaviour. However, the bearing Z-Scores show a cyclic/seasonal behaviour that does not appear in the other experiments. Bearing T1B4 also has more outliers, like the readings 862, 1326, 1360, 1379, 1494, and 1395-1453 before the degradation of the bearing starts at 1508. The bearing then goes back to a suspicious state from 1609-1860 before experiencing a second fault/degradation, which could be related to the cyclic behaviour of the bearing.

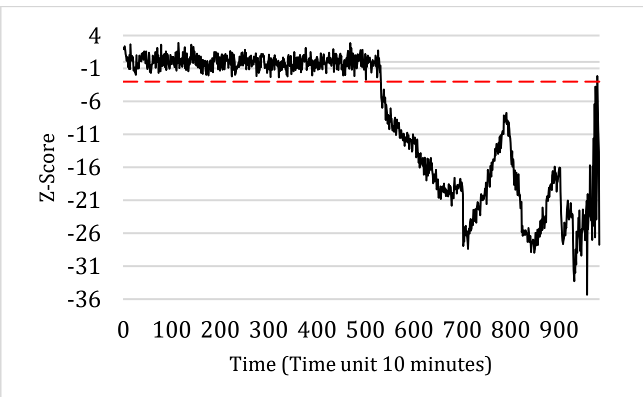

Figure 10 Health monitoring for the full lifetime of T2B1 (Fault incipience at 533)

T2B1 Shows smooth startup and operation behaviour. Its degradation starts at the reading 533. Its degradation behaviour could also be of interest in a fault diagnosis study.

4.5. Results evaluation

The DS-PCA-HMM proposed approach signalled the incipience of degradation and its gradual propagation. Detection performance is determined by the sample time/number at which the method detects the initiation of degradation. Therefore, early fault detection is an indication of better performance. DS-PCA-HMM's performance is compared to benchmark methods performance, including the HMM with DPCA (DPCA-HMM), PCA-HMM, and Variable Replacing Contribution Analysis (VRCA) methods proposed in (Yu, 2012a), an automated Auto-Encoder (AEC) method proposed in (Hasani et al., 2017), and others. Table 3 summarizes the comparison between the DS-PCA-HMM's detection performance and other benchmark health monitoring approaches from the literature for the IMS dataset.

Table 3 Detection performance. DPCA-HMM, PCA -HMM, VRCA (Yu, 2012a). AEC (Hasani et al., 2017). MAS-Kurtosis: Moving average spectral kurtosis (Kim et al., 2016).

Method	**T1B3**	**T1B4**	**T2B1**	**T3B3**
	Degradation starting point			
DS-PCA-HMM	*2120*	*1508*	*533*	*Not tested*
AEC	2027	1641	547	2367
DPCA-HMM	2120	1760	539	Not tested
PCA-HMM	Not tested	1780	538	Not tested
RMS	2094	1730	539	No detection
Kurtosis-MAS	1910	1650	710	No detection
VRCA	Not tested	1727	Not tested	No detection

It is observed that the proposed approach is on par with the DPCA-HMM approach in T1B3 while it performs better than all other approaches, including the AEC approach, in T1B4 and T3B3.

5. Conclusion

The health monitoring method, DS-PCA-HMM, jointly employing PCA to REB segmented vibration data for signal processing and HMM for degradation monitoring was presented. The DS and PCA procedure demonstrated to be an effective automated signal processing technique. It eliminated the need to manually select signal processing features to represent the vibration signal, yet it utilizes the established DS and PCA data-driven methods. The HMM showed strong performance albeit it was trained with a relatively small dataset.

When tested on the IMS Bearing dataset and compared to fault detection benchmark methods, DS-PCA-HMM outperformed state-of-the-art methods from the literature, including the AEC approach. However, the method's expansion to perform diagnosis with limited datasets needs to be further investigated in a follow-up study.

Acknowledgement

Thanks are due to Prof. M. A. Younes (Faculty of Engineering, Alexandria University, Alexandria, Egypt) for his valuable inputs on handling experiment results.

This work was supported by computational resources provided by the Bibliotheca Alexandrina (hpc.bibalex.org) on its High-Performance Computing (HPC) infrastructure.

REFERENCES

Arpaia, P., Cesaro, U., Chadli, M., Coppier, H., De Vito, L., Esposito, A., Gargiulo, F., & Pezzetti, M. (2020). Fault detection on fluid machinery using Hidden Markov Models. *Measurement: Journal of the International Measurement Confederation*, *151*, 107126. https://doi.org/10.1016/j.measurement.2019.107126

Bently Nevada. (2019). *Rolling Element Bearing Methodology - Application Guide*.

Berry, J. E. (1996). High-Frequency Enveloping and Demodulation Techniques. *Technical Associates of Charlotte, PC Charlotte, NC, 28204*. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q="High-Frequency+Enveloping+and+Demodulation+Techniques&btnG=

Bishop, C. (2006). Pattern Recognition and Machine Learning. In *Journal of Organic Chemistry*. https://doi.org/10.1021/jo01026a014

Chen, Y., Peng, G., Xie, C., Zhang, W., Li, C., & Liu, S. (2018). ACDIN: Bridging the gap between artificial and real bearing damages for bearing fault diagnosis. *Neurocomputing*, *294*, 61–71. https://doi.org/10.1016/j.neucom.2018.03.014

Gheisari, M., Wang, G., & Bhuiyan, M. Z. A. (2017). A Survey on Deep Learning in Big Data. *Proceedings - 2017 IEEE International Conference on Computational Science and Engineering and IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, CSE and EUC 2017*, *2*, 173–180. https://doi.org/10.1109/CSE-EUC.2017.215

Hasani, R. M., Wang, G., & Grosu, R. (2017). An Automated Auto-encoder Correlation-based Health-Monitoring and Prognostic Method for Machine Bearings. In *arXiv*. http://arxiv.org/abs/1703.06272

Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 374, Issue 2065). Royal Society of London. https://doi.org/10.1098/rsta.2015.0202

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. In *Encyclopedia of Statistics in Behavioral Science* (Vol. 30, Issue 3). https://doi.org/10.2307/1270093

Kalgren, P. W., Byington, C. S., Roemer, M. J., & Watson, M. J. (2006). Defining PHM, a lexical evolution of maintenance and logistics. *AUTOTESTCON (Proceedings)*, 353–358. https://doi.org/10.1109/AUTEST.2006.283685

Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. In *Mechanical Systems and Signal Processing* (Vol. 107, pp. 241–265). Academic Press. https://doi.org/10.1016/j.ymssp.2017.11.024

Kim, S., Park, S., Kim, J. W., Han, J., An, D., Kim, N. H., & Choi, J. H. (2016). A new prognostics approach for bearing based on entropy decrease and comparison with existing methods. *Proceedings of the Annual Conference of the Prognostics and Health Management Society, PHM*, *2016-Octob*, 154–161. https://click.endnote.com/viewer?doi=10.36001/phmconf.2016.v8i1.2540&route=6

Lee, H., & Choi, S. (2003). PCA+HMM+SVM for EEG pattern classification. *Proceedings - 7th International Symposium on Signal Processing and Its Applications, ISSPA 2003*, *1*(2), 541–544. https://doi.org/10.1109/ISSPA.2003.1224760

Lee, J., Qiu, H., Yu, G., Lin, J., & Rexnord Technical Services. (2007). Bearing Data Set, IMS, University of Cincinnati. In *NASA Ames Prognostics Data Repository (http://ti.arc.nasa.gov/project/prognostic-data-repository)*. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/#bearing

Liu, W. Y., Tang, B. P., Han, J. G., Lu, X. N., Hu, N. N., & He, Z. Z. (2015). The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. *Renewable and Sustainable Energy Reviews*, *44*, 466–472. https://doi.org/10.1016/j.rser.2014.12.005

Ma, M., Sun, C., & Chen, X. (2018). Deep Coupling Autoencoder for Fault Diagnosis with Multimodal Sensory Data. *IEEE Transactions on Industrial Informatics*, *14*(3), 1137–1145. https://doi.org/10.1109/TII.2018.2793246

Mannor, S., Jin, X., Han, J., Jin, X., Han, J., Jin, X., Han, J., & Zhang, X. (2011). K-Means Clustering. In *Encyclopedia of Machine Learning* (pp. 563–564). Springer US. https://doi.org/10.1007/978-0-387-30164-8_425

Md Nor, N., Che Hassan, C. R., & Hussain, M. A. (2020). A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. In *Reviews in Chemical Engineering* (Vol. 36, Issue 4, pp. 513–553). De Gruyter. https://doi.org/10.1515/revce-2017-0069

Munro, P., Toivonen, H., Webb, G. I., Buntine, W., Orbanz, P., Teh, Y. W., Poupart, P., Sammut, C., Sammut, C., Blockeel, H., Rajnarayan, D., Wolpert, D., Gerstner, W., Page, C. D., Natarajan, S., & Hinton, G. (2011). Baum-Welch Algorithm. In *Encyclopedia of Machine Learning* (pp. 74–74). Springer US. https://doi.org/10.1007/978-0-387-30164-8_59

Ozg, A. (2010). *AR-PCA-HMM Approach for Sensorimotor Task Classification in EEG-based Brain-Computer Interfaces*. 113–116. https://doi.org/10.1109/ICPR.2010.36

Pohle, J., Langrock, R., van Beest, F. M., & Schmidt, N. M. (2017). Selecting the number of states in hidden markov models - Pitfalls, practical challenges and pragmatic solutions. In *arXiv*. https://click.endnote.com/viewer?doi=arxiv%3A1701.08673&token=WzE3NDgyMDUsImFyeGl2OjE3MDEuMDg2NzMiXQ.ZXTAwKkl1oGTfLo65edlV4RzOt8

Qiu, G., Gu, Y., & Cai, Q. (2019). A deep convolutional neural networks model for intelligent fault diagnosis of a gearbox under different operational conditions. *Measurement: Journal of the International Measurement Confederation*, *145*, 94–107. https://doi.org/10.1016/j.measurement.2019.05.057

Qiu, H., Lee, J., Lin, J., & Yu, G. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, *289*(4–5), 1066–1090. https://doi.org/10.1016/j.jsv.2005.03.007

R Gopinath, CS Kumar, K. R. (2018). Scalable fault models for diagnosis in a synchronous generator using feature mapping and transformation techniques. *International Journal of Prognostics and Health Management*, *9*(2), 11.

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, *77*(2), 257–286. https://doi.org/10.1109/5.18626

Saidi, L., Ben Ali, J., Bechhoefer, E., & Benbouzid, M. (2017). Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR. *Applied Acoustics*, *120*, 1–8. https://doi.org/10.1016/j.apacoust.2017.01.005

Saufi, S. R., Ahmad, Z. A. Bin, Leong, M. S., & Lim, M. H. (2018). Differential evolution optimization for resilient stacked sparse autoencoder and its applications on bearing fault diagnosis. *Measurement Science and Technology*, *29*(12). https://doi.org/10.1088/1361-6501/aae5b2

Saufi, S. R., Ahmad, Z. A. Bin, Leong, M. S., & Lim, M. H. (2019). Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review. In *IEEE Access* (Vol. 7, pp. 122644–122662). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2019.2938227

Shao, S., Sun, W., Wang, P., Gao, R. X., & Yan, R. (2016). Learning features from vibration signals for induction motor fault diagnosis. *International Symposium on Flexible Automation, ISFA 2016*, 71–76. https://doi.org/10.1109/ISFA.2016.7790138

Shriram Ramanathan. (2018). 5 Challenges for Deep Learning. *EE Times*. https://www.eetimes.com/5-challenges-for-deep-learning/

Soualhi, A., Hawwari, Y., Medjaher, K., Clerc, G., Hubert, R., & Guillet, F. (2018). PHM survey: Implementation of signal processing methods for monitoring bearings and gearboxes. *International Journal of Prognostics and Health Management*, *9*(2), 1–14.

Teng, W., Zhang, X., Liu, Y., Kusiak, A., & Ma, Z. (2017). Prognosis of the remaining useful life of bearings in a wind turbine gearbox. *Energies*, *10*(1). https://doi.org/10.3390/en10010032

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*(4), 267–276. https://doi.org/10.1007/BF02289263

Tobon-Mejia, D. A., Medjaher, K., Zerhouni, N., & Tripot, G. (2012). A data-driven failure prognostics method based on mixture of gaussians hidden markov models. *IEEE Transactions on Reliability*, *61*(2), 491–503. https://doi.org/10.1109/TR.2012.2194177

Vogl, G. W., Weiss, B. A., & Helu, M. (2019). A review of diagnostic and prognostic capabilities and best practices for manufacturing. *Journal of Intelligent Manufacturing*, *30*(1), 79–95. https://doi.org/10.1007/s10845-016-1228-8

Walker, C., & Coble, J. (2018). Wind turbine bearing fault detection using adaptive resampling and order tracking. *International Journal of Prognostics and Health Management*, *9*(2).

Yu, J. (2012a). Health condition monitoring of machines based on hidden markov model and contribution analysis. *IEEE Transactions on Instrumentation and Measurement*, *61*(8), 2200-2211. https://doi.org/10.1109/TIM.2012.2184015

Yu, J. (2012b). Local and nonlocal preserving projection for bearing defect classification and performance assessment. *IEEE Transactions on Industrial Electronics*, *59*(5), 2363–2376. https://doi.org/10.1109/TIE.2011.2167893

Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, *2*(2), 226–235. https://doi.org/10.3390/j2020016

Zec, E. L., Mohammadiha, N., & Schliep, A. (2018). Statistical Sensor Modelling for Autonomous Driving Using Autoregressive Input-Output HMMs. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-November*, 1331–1336. https://doi.org/10.1109/ITSC.2018.8569592

Zhang, S., Ye, F., Wang, B., & Habetler, T. G. (2019). Semi-supervised learning of bearing anomaly detection via deep variational autoencoders. In *arXiv*. http://www.merl.com

Zhong, J. H., Zhang, J., Liang, J., & Wang, H. (2019). Multi-Fault Rapid Diagnosis for Wind Turbine Gearbox Using Sparse Bayesian Extreme Learning Machine. *IEEE Access*, *7*, 773–781. https://doi.org/10.1109/ACCESS.2018.2885816

Akthem Rehab is a seasoned management consultant with solid experience in operations and maintenance strategy and transformation. He is now concentrating on industry 4.0 and its applications. Akthem holds a BSc. in Industrial Engineering, Alexandria University, Egypt and is pursuing his MSc. in Industrial Engineering, Alexandria University. His research is focused on AI and digital twin practical applications in predictive maintenance.

Islam Ali, ***PhD.*** Assistant Professor, Industrial and Manufacturing Engineering Department, School of Innovative Design Engineering, E-JUST, Egypt. He holds PhD degree in Industrial Engineering from Purdue University, USA. He has written and co-authored research papers and articles on systems optimization and digital twin applications.

Walid Gomaa, ***PhD.*** Professor of Computer Science and Engineering, E-JUST, Egypt. He holds PhD degree in Computer Science from University of Maryland College Park, USA. He directs the Cyber Physical lab in E-JUST and leads Machine Learning research and its applications. He has written and co-authored research papers and articles on numerous AI applications.

M. Nashat Fors, ***PhD.*** Professor Emeritus of Industrial Engineering, Faculty of Engineering, Alexandria University, Egypt. He has written and co-authored research papers and articles on Mathematical Programming & Simulation Techniques Applications, Supply Chain Management, Operations Management, Maintenance Planning, Water Management, and Facilities Layout. He has many joint projects, consultations, and training programs with the industry and other universities in the areas of Operations Planning & Scheduling, Maintenance Planning, and Water Management. Dr. Fors is the advisor or co-advisor of more than 50 earned M.Sc. and 10 Ph.D. in the different areas of Industrial Engineering. He participated and advised in designing, innovation, and development of educational programs and curricula at graduate and undergraduate levels at Alexandria University, Cairo University, Ain Shams University, E-Just University, and AAST&MT.