

## Research Article

# Novel Approach to Rotating Machinery Diagnostics Based on Principal Component and Residual Matrix Analysis

**A. Abouhnik, Ghalib R. Ibrahim, R. Shnibha, and A. Albarbar**

*School of Engineering, Manchester Metropolitan University, Manchester M1 5GD, UK*

Correspondence should be addressed to A. Abouhnik, abouhnik@yahoo.com

Received 6 November 2011; Accepted 21 December 2011

Academic Editor: K. Yasuda

Copyright © 2012 A. Abouhnik et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rotating machinery such as induction motors and gears driven by shafts are widely used in industry. A variety of techniques have been employed over the past several decades for fault detection and identification in such machinery. However, there is no universally accepted set of practices with comprehensive diagnostic capabilities. This paper presents a new and sensitive approach, to detect faults in rotating machines; based on principal component techniques and residual matrix analysis (PCRMA) of the vibration measured signals. The residual matrix for machinery vibration is extracted using the PCA method, crest factors of this residual matrix is determined and then machinery condition is assessed based on comparing the crest factor amplitude with the base line (healthy) level. PCRMA method has been applied to vibration data sets collected from several kinds of rotating machinery: a wind turbine, a gearbox, and an induction motor. This approach successfully differentiated the signals from healthy system and systems containing gear tooth breakage, cracks in a turbine blade, and phase imbalance in induction motor currents. The achieved results show that the developed method is found very promising and Crest Factors levels were found very sensitive for machinery condition.

## 1. Introduction

Detection of faults in rotating machinery remains a big challenge especially for complex mechanical systems. Despite substantial advances in sensing and signal processing technologies, many difficulties remain to the successful detection of faults at an early stage of their development to avoid catastrophic failure [1, 2]. It has been known for some time through both analytical and experimental investigations that some machine faults can be directly related to acoustic and vibration harmonics [3].

With wind turbines, Jüngert [4] used two different acoustic techniques for blade inspection; those were local resonance spectroscopy and audible sound. The results showed that information about the internal structure of the inspected area can be obtained using features extracted from these signals. Sajauskas et al. [5] used secondary longitudinal surface acoustic waves (LSAW II) to detect surface defects on the inaccessible inner surface of sheet products and showed that this method was particularly efficient in the investigation of regular shape defects (cracks)

with predictable orientation. Ghoshal et al. [6] studied four different algorithms for detecting damage on a wind turbine blade based on the vibration response of the blade: transmittance function, resonant comparison, operational deflection shape, and wave propagation. Park et al. [7] studied the linear vibration analysis of a rotating wind-turbine blade to obtain its vibratory characteristics and proposed a computational algorithm based on variations in blade stiffness due to the centrifugal inertia forces.

Parey et al. [8] used empirical mode decomposition (EMD) to analyse nonstationary and nonlinear signals in order to detect localized tooth defects in gears. Kar and Mohanty [9] applied advanced signal-processing techniques, such as the discrete wavelet transform (DWT) and a corrected multiresolution Fourier transform (MFT), to investigate the vibration and current transients in a multistage gearbox under transient load. Endo et al. [10] proposed a differential diagnostic technique to investigate two kinds of localized gear tooth defects, a crack and spall in the tooth of a gear. A set of digital filters filtered the signal to separate the diagnostic information (fault impulses) from the complex

mixture of signals measured with the sensor placed on the gearbox casing. Ibrahim and Albarbar [11] presented a comparison between empirical mode decomposition (EMD) and smoothed pseudo-Wigner-Ville distribution (SPWVD) methods based on the vibration signature and energy calculation procedure, for monitoring gearbox systems. The results showed that the calculation of energy using EMD techniques offers a more effective way to detect early faults than computations using the SPWVD method, and that the computation of energy using the EMD technique is faster than the calculations done using the SPWVD method.

Induction motors are critical components for many industrial processes and frequently integrated in commercially available equipment and industrial processes. Early fault detection would eliminate consequential damages of motors and reduce outage time and costs of repairs.

Several diagnosis techniques for the identification and localisation of the faults have been proposed. Acosta et al. [12] combined motor current signature analysis (MCSA) and extended Park's vector approach (EPVA) for fault detection and diagnosing induction motor stator and rotor-related faults. Antonino-Daviu [13] have proposed a method for rotor bar failure detection based on the stator current during startup using the discrete wavelet transform (DWT). A novel approach to induction motor current signature analysis based on wavelet packet decomposition (WPD) of the stator current was presented by Ye et al. [14]. The feature coefficients extracted using WPD were then used to detect rotor bar breakage and air-gap eccentricity faults. Liang et al. [15] showed that both current spectra and vibration spectra are able to reveal broken rotor bar faults.

## 2. Principal Component and Residual Matrix Approach

Multivariate statistical analysis (i.e., PCA) has been used in many areas of science and engineering to detect abnormalities in the structure of data sets. Its operation can be thought of as extracting from the data of those parameters which best explains the variability in the data. PCA works by computing a compact and optimal description of the observed or experimental data set, reducing the number of dimensions of the data set while retaining as much relevant information as possible. It is assumed that the different parameters are uncorrelated with each other, that is, the signals from different faults can be discerned from each other. PCA is a useful technique for discerning variables that explain the observed trends in a process.

PCA describes the dispersion of an array of  $n$  points in  $p$ -dimensional space by introducing a set of principal components. The first principal component is chosen in such a way that accounts for as much of the variance in the data set as possible. The second principal component is in mathematical terms said to be orthogonal to the first, uncorrelated with it, that is, in reality it would be an independent fault. This second principal component is chosen to account for as much of the remaining variance as possible. The process then proceeds with each succeeding component accounting for

as much of the remaining variance as possible. The process continues until a suitable termination is reached depending on the accuracy required or limits of computation.

Consider  $n$  samples of actual data collected from  $m$  processes and placed in the matrix  $X \in \mathfrak{R}^{n \times m}$ . The first step in calculating the PCA is to normalise this matrix to unit variance and zero mean using scale parameter vectors  $s$  and  $x$  as the variance and mean vectors, respectively. The next step is to determine the covariance matrix  $R$ :

$$R = \frac{1}{n-1} X^T X. \quad (1)$$

The third step is to perform single-value decomposition (SVD) on  $R$ :

$$R = V \Lambda V^T, \quad (2)$$

where  $\Lambda$  is a diagonal matrix with zero entries everywhere except the leading diagonal, which contains the eigenvalues of  $R$  in descending order ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ ). Columns of matrix  $V$  are the eigenvectors of  $R$ . The transformation matrix  $P \in \mathfrak{R}^{m \times a}$  is generated choosing  $a$  eigenvectors or columns of  $V$  corresponding to  $a$  principal eigenvalues. The purpose of  $P$  is to reduce the dimension space of the measured variables

$$T = XP. \quad (3)$$

The elements of matrix  $T$  are called *scores*, and these are the original data (measured values) after being transformed into the reduced dimension space. The columns  $P$  are called *loadings*. Of course, the relation shown in (3) can be manipulated to transform the scores back into the original space

$$\hat{X} = TP^T. \quad (4)$$

$E$ , the residual matrix, is defined as

$$E = X - \hat{X}. \quad (5)$$

It follows that the original data space can be calculated as

$$X = TP^T + E. \quad (6)$$

In (6),  $E$  is the variability due to process noise, and  $TP^T$  is the principal sources of variability in the process [16]. It is very important to choose the number of principal components  $a$ , because  $TP^T$  represents the principal sources of variability in the process, and  $E$  represents the variability corresponding to process noise.

The proposed novel technique is to determine rotating machinery faults based on the PCRMA algorithm shown graphically in Figure 1.

## 3. Experimental Work

The experimental work was carried out in the Mechanical Laboratory of the Manchester Metropolitan University using several different kinds of rotating machinery: a wind turbine,

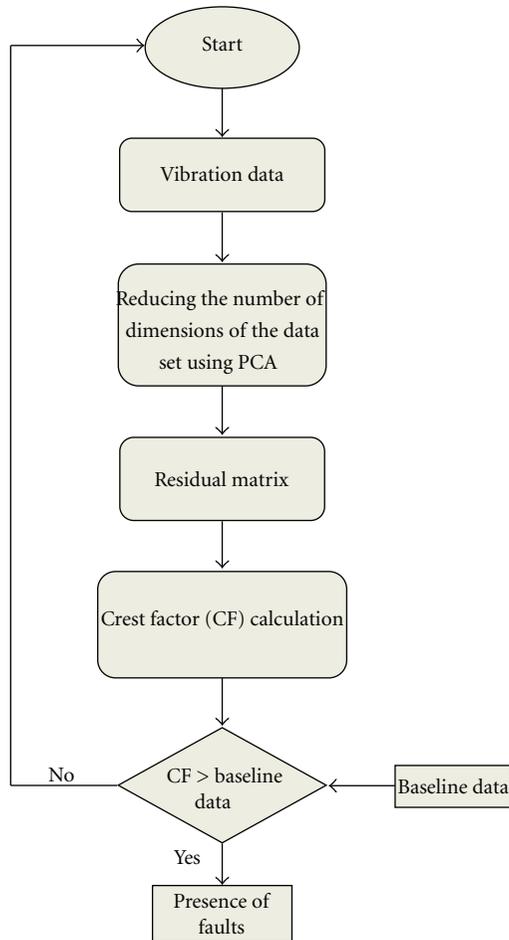


FIGURE 1: Illustration of the PCA-based proposed technique.

a gearbox, and an induction motor. The same instrumentation, for example, accelerometers and associated charge amplifiers were used to collect data from all three systems. The accelerometers were B and K type 4371 with sensitivity of 10 mV/g and suitable for vibration measurements within a range of 1 Hz to 12 kHz. Before the accelerometer signals were fed to the analogue-to-digital converter NI USB 9233 card, they passed through a B and K type 2635 charge amplifier to condition the signal. The charge amplifier converts the accelerometer high-impedance, low-charge (in the range of Pico-coulomb) signal into low impedance and high voltage (in the range of mV), and the cut-off frequency for initialising filter was set to 10 kHz. The developed Labview-based system allowed the user to monitor and store machine variables, and subsequently, a Matlab code was used for further signal processing. The accelerometer was mounted vertically on the gearbox housing, on the nacelle of the wind turbine, and on the induction motor.

Figure 2 represents wind turbine test rig, and it consists of a six airfoil blades; each blade has a length of 32 cm. The wind turbine was located about 1 meter in front of the wind tunnel for two wind speeds: 4.7 and 5.3 m/s. The experimental work was carried out with five healthy blades

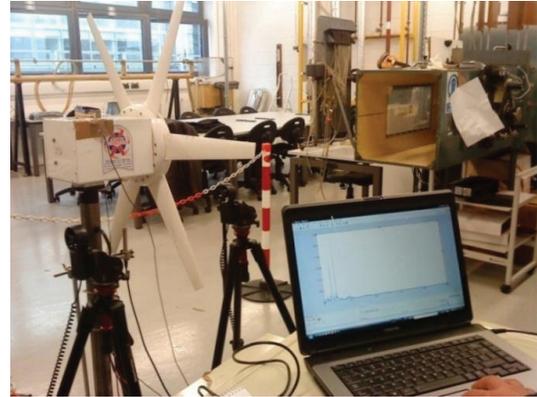


FIGURE 2: Wind turbine test rig.

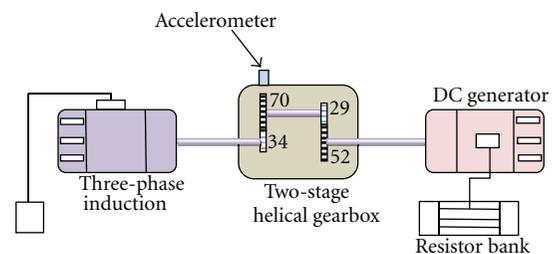


FIGURE 3: Schematic layout of the gearbox test.

and with a single blade with a crack introduced. The seeded crack was of three lengths: 15 mm, 25 mm, and 40 mm, and it had a consistent 3 mm width and 2 mm depth.

The gearbox used as a second machinery test consists of a two-stage, 11 kW, helical gearbox driven by a three-phase induction motor and connected to a DC generator and resistor banks as shown in Figure 3. Tests were carried out using a healthy pair of gears and the first drive pinion suffering from a tooth breakage of three magnitudes: fault 1 (25% tooth removal), fault 2 (50% tooth removal), and fault 3 (75% tooth removal). The drive pinion at the first stage had 34 teeth meshing with a 70-tooth wheel. The pinion gear at the second stage had 29 teeth meshing with 52-tooth wheel. The vibration signals were collected using an accelerometer mounted vertically on the gearbox housing for a drive speed 90% of full speed and a load of 40% cent of maximum load.

A third rotating machinery test used a 3 kW, 220 V induction motor, and a DC generator with a resistor bank to act as a load. A schematic diagram of the test rig is shown in Figure 4. The motor was tested at 0%, 25%, 50%, 75%, and 100% load in a healthy condition and then under two different phase imbalance voltages of 20 and 40 V.

#### 4. Result and Discussion

The collected vibration data from a wind turbine for the healthy blade case is shown in Figure 5. These were analysed using statistical method such as Kurtosis, RMS, and CF. Figure 6 illustrates statistical analysis for wind turbine vibration signal under different conditions: h is healthy,

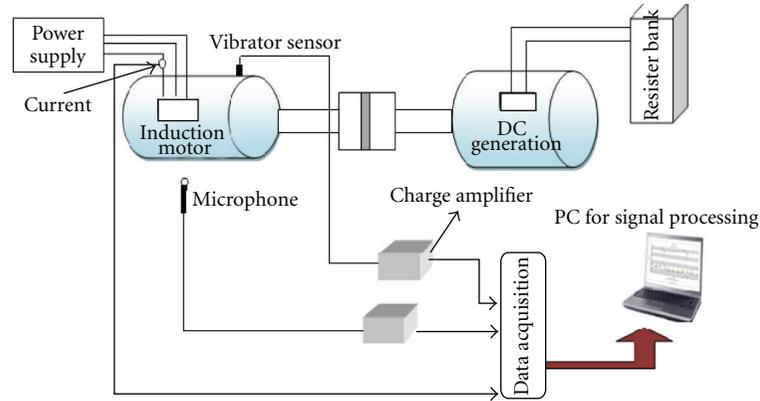


FIGURE 4: Schematic diagram of induction motor test rig.

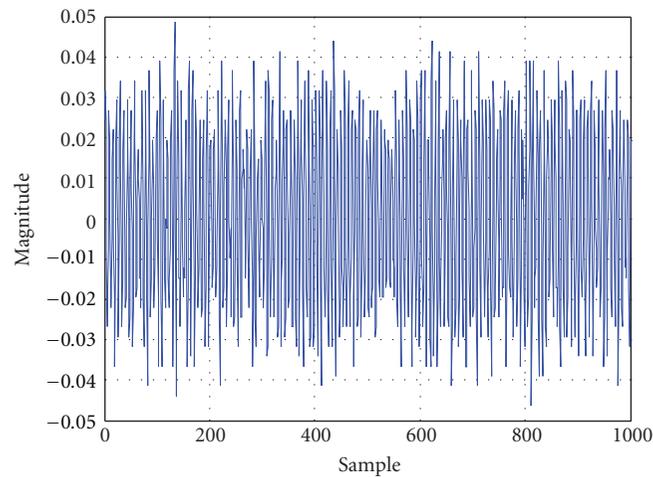


FIGURE 5: Healthy vibration signal from wind turbine.

fault1 ( $f_1$ ) is a crack 15 mm long, fault2 ( $f_2$ ) is a crack 25 mm long, and fault3 ( $f_3$ ) is a crack 40 mm long. From Figure 5, it can be seen that the results show no clear trend, which means that the method cannot be used to diagnose system defects.

PCA is a method for reducing what may be considered a contaminated signal into a series of manageable data sets. Every set contains its principal components (PCs), which are interpretations of the original data. Data were collected and analyzed using PCA method as shown in Figures 7, 8, and 9. Figure 7 shows eigenvalues versus PCs. From theory each PC contains a proportion of the total variance, and the eigenvalues represent that amount. Thus, the eigenvalues are a measure of the relative importance of each PC in reconstructing the real signal.

As shown from the eigenvalues, plot healthy and faulty cases were studied for different fault conditions, but the change in shape of the curves is not sufficiently clear to detect and identify the seeded faults. Figure 8 shows score plot of healthy and faulty cases; in this case, the first plot is a healthy condition and can be used as a guide for the faulty cases. This plot contains certain characteristics such as size, spread, and clusters in the data. From the plot, there is no clear

difference between healthy and faulty cases, which means that this method is not sensitive enough to small changes in the vibration signal to detect initial faults.

The main use of PCA is to reduce the dimensionality of a data set, and it is assumed that the first few PCs contain most of, not all, the necessary relevant information contained with the original data. Thus, it follows that the remaining PCs should contain mostly noise from the original data. According to PCA theory, these PCs can be contained in a separate matrix called the residual matrix, which is constructed in the same way as the original data matrix except using only the irrelevant PCs and their respective weightings.

Actually, the relevant scores (PCs) are used to calculate the residual matrix. The residual matrix contains the information which has been removed from the analysis, and the errors are found using this matrix. The sum of the errors in each column of the matrix is calculated and squared to give a positive result which is plotted. The residual errors plots for healthy and faulty cases are showed in Figure 9. From this figure, there is a difference between healthy and faulty cases which means that the condition of the wind turbine was

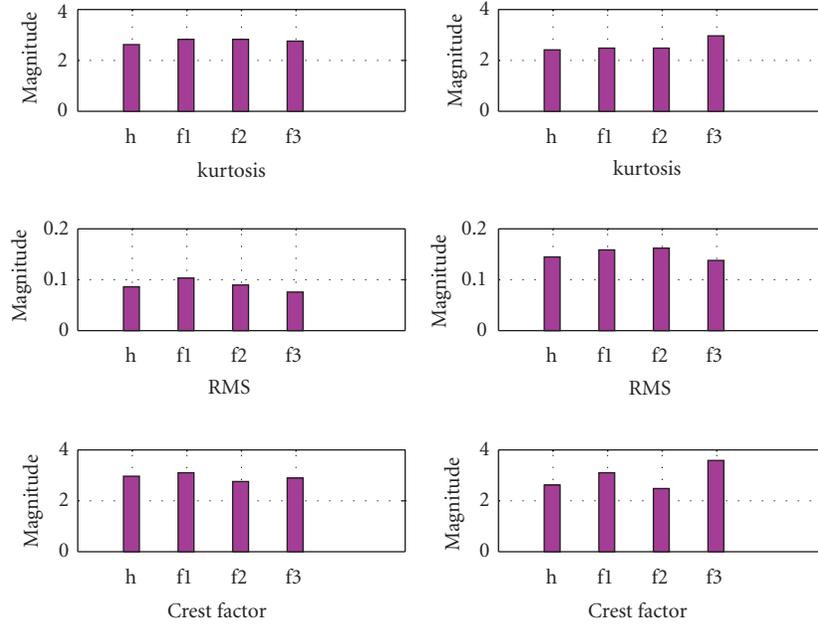


FIGURE 6: Magnitude of kurtosis, RMS, and crest factor vibration signal from wind turbine with healthy blades and one faulty blade with the fault of three different degrees of severity.

varying. But to classify the cases between healthy and faulty was difficult.

Developing a new way of thinking about this analysis could provide an effective method of evaluating the signals obtained from condition monitoring to determine incipient faults. A PCA-based technique was applied to the measured data for each of the three test systems. The CF was calculated for the residual signal for each of the three systems for each fault seeded into the system.

To test this novel approach, the proposed technique was applied to different data sets. Figures 10 and 11 represent crest factor values for the residual error for wind turbine blades signal for six healthy blades and when the three faults were seeded into one of the blades. It can be seen that the value of crest factor increased with fault level which suggests that the proposed method may have advantages over the other statistical techniques when used with PCA. Secondly, the new approach was used to analyse gearbox signals collected under healthy conditions and for the three tooth faults: fault 1 (25% tooth removal), fault 2 (50% tooth removal), and fault 3 (75% tooth removal). The test was carried out for a load that was 40% of the total load, and the speed was 1350 rpm. Figure 12 shows the results using a PCA-based technique applied to the measured vibration signal. The results showed that there was a clear trend with the value of the CF of the residual error plot increasing with severity of the fault. It is concluded that there is advantage to using the proposed technique.

Finally, the new technique was applied to data collected from an induction motor under different conditions. The motor was tested at 0%, 25%, 50%, 75%, and 100% load in a healthy condition and then under two different single-phase imbalance voltages of 20 and 40 V.

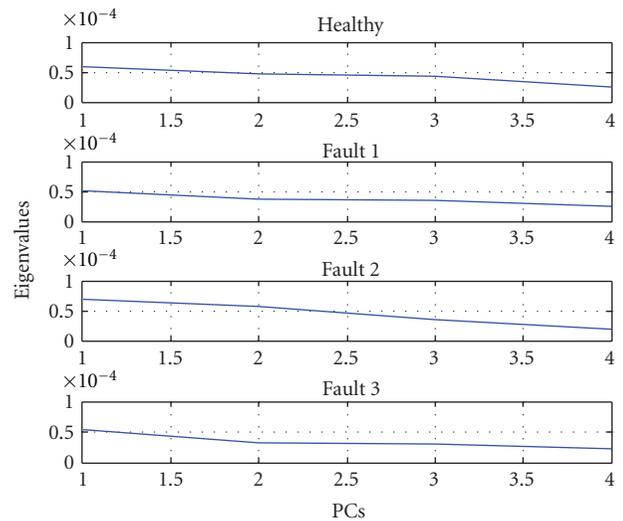


FIGURE 7: Eigenvalues as a function of Principal Component for healthy and faulty blade at wind speed of 4.7 m/s.

In this case, the CF was also used as a statistical feature extracted from the raw vibration signal, but no adequate, reliable information related to the condition of the motor was observed. To detect the phase imbalance, the PCA method was applied to the measured vibration signals. Crest factor values for the residual error based on PCA were calculated, and in Figure 13, it can be seen that there is a clear relationship between the fault severity and the crest factor for all loads tested. The approach developed here was able to detect changes of the phase imbalance.

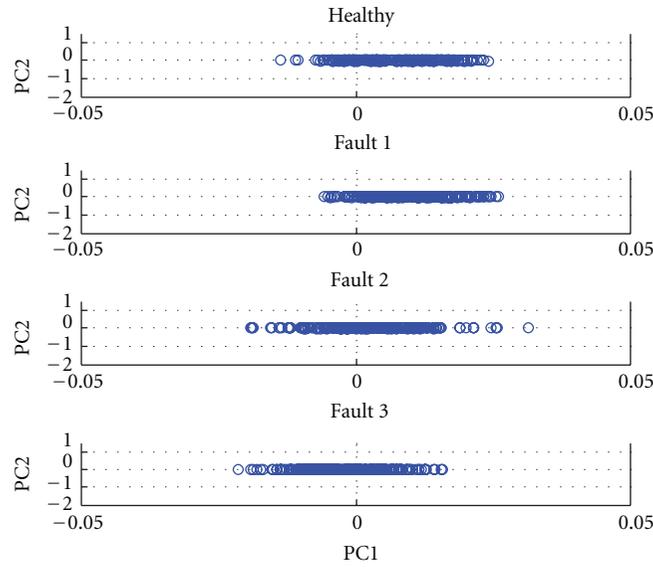


FIGURE 8: Score plot for healthy and faulty blade at wind speed of 4.7 m/s for the first principal component.

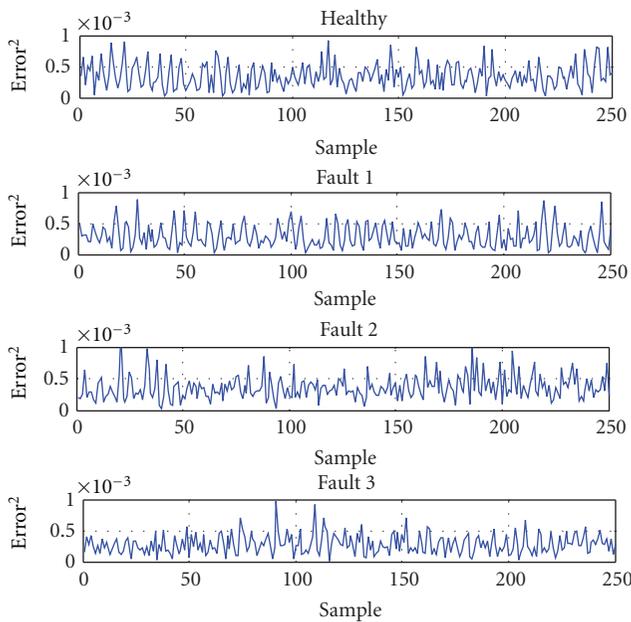


FIGURE 9: Residual error plot for healthy and faulty blade at wind speed 4.7 m/s.

**5. Conclusion**

The vibration signals collected from rotating machinery are often so contaminated that simple statistical parameters are not sufficient for fault detection. This study has introduced a promising new approach to detect faults in different rotating machines. PCA was used to extract the residual matrix which contains the information removed from the analysis and the errors from the vibration signal. Then the crest factor was applied to the residual matrix. It has been found that such

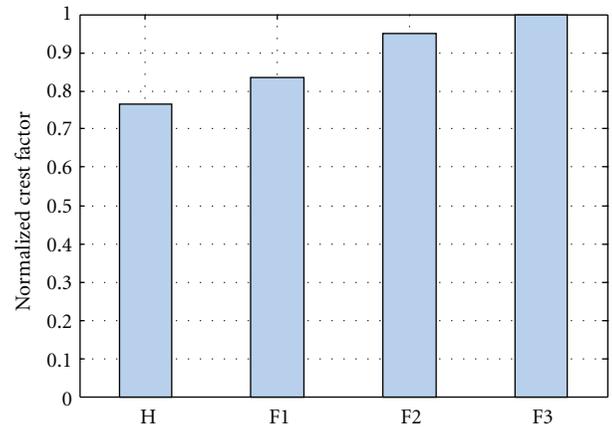


FIGURE 10: Crest factor for healthy and faulty turbine blade at wind speeds of 4.7 m/s.

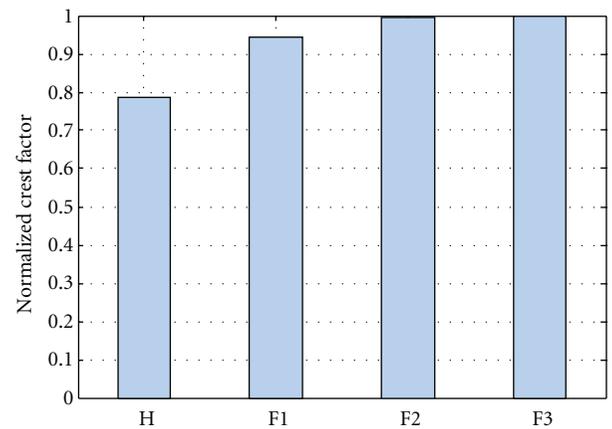


FIGURE 11: Crest factor for healthy and faulty turbine blade at wind speeds of 5.3 m/s.

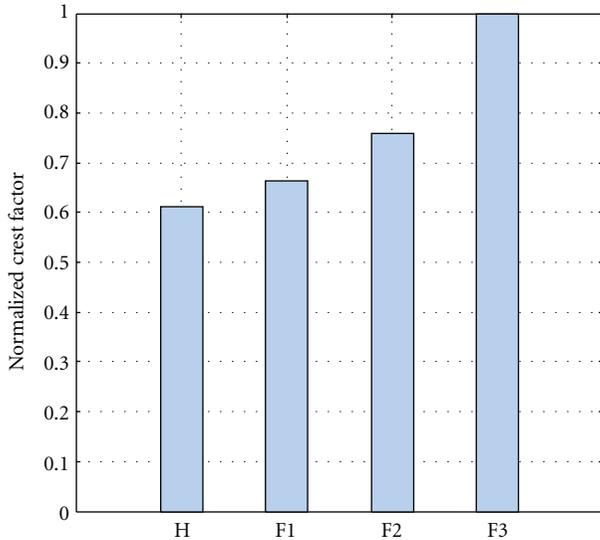


FIGURE 12: Crest factor plot for healthy gearbox and gearbox with three tooth faults introduced (40% full load and 90% full speed).

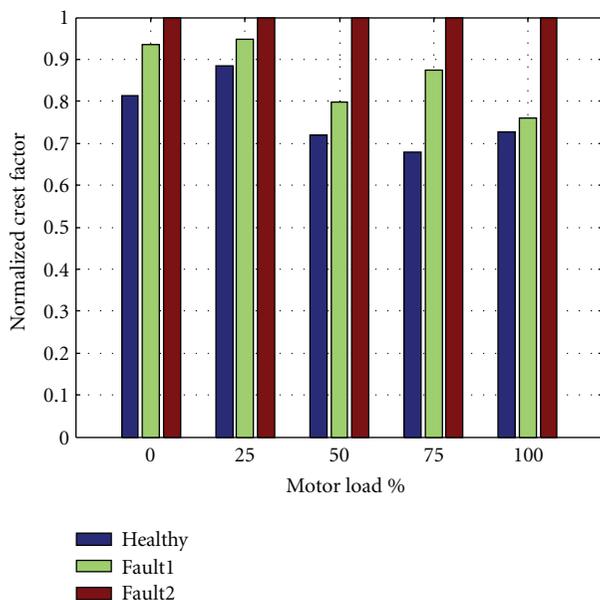


FIGURE 13: Crest factor of residual error for healthy and faulty induction motor.

a technique has shown great promise in detecting voltage imbalance in an induction motor, gear teeth breakage in a gearbox, and cracked blades on a wind turbine.

## References

- [1] T. Reilly and T. Proulx, *A Review of Signal Processing and Analysis Tools for Comprehensive Rotating Machinery Diagnostics, Rotating Machinery, Structural Health Monitoring, Shock and Vibration*, vol. 5, Springer, New York, NY, USA, 2011.
- [2] Y. Lu, J. Tang, and H. Luo, "Wind turbine gearbox fault detection using multiple sensors with feature level data fusion," in *Proceedings of the ASME Turbine Technical Conference & Exposition*, 2011, GT2011-46538.
- [3] G. K. Singh and S. A. S. Al Kazzaz, "Induction machine drive condition monitoring and diagnostic research—a survey," *Electric Power Systems Research*, vol. 64, no. 2, pp. 145–158, 2003.
- [4] A. Jüngert, "Damage Detection in wind turbine blades using two different acoustic techniques," *NDT Database & Journal of Nondestructive Testing*, 2008.
- [5] S. Sajauskas, A. Valinevičius, and L. Miežutavičiūtė, "Non-destructive testing of sheet product inner surfaces using longitudinal surface acoustic waves," *Ultrasound*, pp. 12–16, 2005.
- [6] A. Ghoshal, M. J. Sundaresan, M. J. Schulz, and P. F. Frank Pai, "Structural health monitoring techniques for wind turbine blades," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 85, no. 3, pp. 309–324, 2000.
- [7] J. H. Park, H. Y. Park, S. Y. Jeong, S. I. Lee, Y. H. Shin, and J. P. Park, "Linear vibration analysis of rotating wind-turbine blade," *Current Applied Physics*, vol. 10, no. 2, supplement, pp. S332–S334, 2010.
- [8] A. Parey, M. El Badaoui, F. Guillet, and N. Tandon, "Dynamic modelling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect," *Journal of Sound and Vibration*, vol. 294, no. 3, pp. 547–561, 2006.
- [9] C. Kar and A. R. Mohanty, "Vibration and current transient monitoring for gearbox fault detection using multiresolution Fourier transform," *Journal of Sound and Vibration*, vol. 311, no. 1-2, pp. 109–132, 2008.
- [10] H. Endo, R. B. Randall, and C. Gosselin, "Differential diagnosis of spall vs. cracks in the gear tooth fillet region: experimental validation," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 636–651, 2009.
- [11] G. R. Ibrahim and A. Albarbar, "Comparison between Wigner-ville distribution- and empirical mode decomposition vibration-based techniques for helical gearbox monitoring," *Journal of Mechanical Engineering Science*, vol. 225, no. 8, pp. 1833–1846, 2011, Proceedings of the Institution of Mechanical Engineers C.
- [12] G. G. Acosta, C. J. Verucchi, and E. R. Gelso, "A current monitoring system for diagnosing electrical failures in induction motors," *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 953–965, 2006.
- [13] J. Antonino-Daviu, M. Riera-Guasp, J. Roger-Folch, F. Martínez-Giménez, and A. Peris, "Application and optimization of the discrete wavelet transform for the detection of broken rotor bars in induction machines," *Applied and Computational Harmonic Analysis*, vol. 21, no. 2, pp. 268–279, 2006.
- [14] Z. Ye, B. Wu, and A. Sadeghian, "Current signature analysis of induction motor mechanical faults by wavelet packet decomposition," *IEEE Transactions on Industrial Electronics*, vol. 50, no. 6, pp. 1217–1228, 2003.
- [15] B. Liang, B. S. Payne, A. D. Ball, and S. D. Iwnicki, "Simulation and fault detection of three-phase induction motors," *Mathematics and Computers in Simulation*, vol. 61, no. 1, pp. 1–15, 2002.
- [16] D. García-Álvarez, "Fault detection using principal component analysis (PCA) in a wastewater treatment plant (WWTP)," in *Proceedings of the International Student's Scientific Conference*, 2009.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

