Impeller fault detection under variable flow conditions based on three feature extraction methods and artificial neural networks

A. Jami* and P. S. Heyns
Center for Asset Integrity Management, Department of Mechanical and Aeronautical Engineering,
University of Pretoria, Pretoria 0002, South Africa

Corresponding author. Tel.: +27 12 420 2096, Fax.: +27 12 420 2096 E-mail address:
amin.jami5@gmail.com

Abstract

Nonstationary flow conditions can introduce complexities and nonlinear characteristics to pumping systems. This paper presents comparative studies of impeller fault detection techniques combined with artificial neural networks (ANNs) to propose the most appropriate diagnosis system. An experimental study, including seven impeller conditions, is performed to further explore the phenomena. Statistical parameters, frequency peaks, and wavelet packet energy present data feature sets, and a three-layer back-propagation ANN is used for fault recognition. The verification of the results proves that the detectability of the wavelet packet transform (WPT)-ANN model is considerably improved by using the energy of the decomposed vibration from WPT. This model can save computational time and provide superior diagnostic information. This study provides two key contributions. First, the feasibility and effectiveness of common monitoring techniques are compared. Second, the results demonstrate the accuracy of the proposed models for impellers operating under variable working conditions, which has not been previously addressed in the literature.

Keywords: ANN; Centrifugal pumps; FDA; Impeller faults; Nonstationary flow conditions; TDA; Vibration signal processing; WPT

1. Introduction

Centrifugal pumps are extensively used in most industrial sectors, with increasing requirements for enhancing reliability, safety, and production capability. Condition-based maintenance (CBM) is an important strategy for dealing with the aforementioned requirements by detecting early signs of failure and allowing efficient maintenance. Damage detection techniques are generally categorized into the following major levels: existing damage detection, localization, damage severity quantification, and remaining life estimation. Vibration monitoring is frequently applied to CBM because of its potential and sensitivity to the early detection of performance degradation. This process is also applicable to pumps. Three major factors that increase vibration levels in centrifugal pumps, namely, peripheral (e.g., nearby machines), mechanical (e.g., mechanical imbalance of rotating components), and hydraulic causes (e.g., flow interaction with internal components), are distinguished [1]. The most appropriate analysis technique can be selected based on the complexity of the system being considered.

Furthermore, automated fault diagnosis techniques may be introduced to create intelligent fault diagnosis systems and are successfully adopted in many applications [2]. These systems aim to interrogate imput data and predict the progression of faults. Among the available methods, artificial neural net-works (ANNs) have been extensively used by researchers. Most of these previous works involve the diagnosis of bearing defects, whereas only a few existing references deal with im-peller fault detection [3, 4]. Zouari and Menad [5] investigated the capability of ANNs to detect and classify specific faults, such as “partial flow rates, loosening of rear pump attach-ments, and misalignment,” over constant flow rates. Mean-while, a combined optimal wavelet packet transform (WPT) and ANN was presented in [6] and used to illustrate the time-varying energy pattern of various signals. In addition, fluctuat-ing operating conditions are observed in many operating sys-tems, which can generate amplitude, frequency, and phase modulation in vibration signals. However, studies regarding the influence of fluctuating flow conditions for monitoring pump element conditions remain limited. The underlying ideas in the current study are as follows: (1) to contribute ef-fective features by introducing nonstationary signal processing and (2) to train a classifier to identify impeller conditions and detect and quantify damage severities under different operat-ing conditions.

The remainder of the paper is organized as follows. Sec. 2

introduces relevant measurement and signal processing techniques. Sec. 3 considers the use of ANNs in condition monitoring applications. Sec. 4 describes the experimental procedure. Sec. 5 discusses feature extraction based on analyses. Sec. 6 presents the performance of ANNs. Sec. 7 concludes the study.

2. Signal processing

Signal information is commonly extracted from simple statistical parameters in a process known as time domain analysis (TDA). Some of these parameters and their corresponding equations are well documented in [7]. Root mean square (RMS), impulse factor, shape factor, skewness, standard deviation (SD), kurtosis, mean value, energy, lower and upper bounds, entropy, central moments, and signal distribution are commonly used in the literature. [8]. Crack, as a common impeller fault, was analyzed in [9] using statistical features, namely, kurtosis, RMS, skewness, and variance. However, these parameters frequently do not show slight changes in a system. Hence, fast Fourier transform has been introduced as an effective approach for dealing with periodic signals. Parameters, such as the arithmetic mean of a frequency spectrum, geometric mean, and RMS amplitude, can provide a quick overview of machine condition [10]. Pump rotational frequency, vane passing frequency [11], and the corresponding harmonics are proposed as indicators that can provide valuable information regarding a centrifugal pump impeller that can be strongly affected by variable conditions. By contrast, wavelet analysis, which simultaneously performs time-frequency analysis (TFA), is an effective approach for dealing with nonstationary signals [12, 13] and an effective tool for rapidly exhibiting variable amplitude and phase changes [14]. Wavelet analysis expands signals in terms of wavelets produced through the scaling and translation of a mother wavelet. Numerous wavelet transform (WT) methods, including continuous wavelet transform, discrete wavelet transform, and WPT, have been developed. A detailed comparison of WT methods and the existing formulation was presented in [15].

Wavelet packets (WPs) combine traditional wavelet functions. Therefore, WPT enables the extraction of high time and frequency resolutions for signals that inherit stationary and nonstationary characteristics. A WP function $\psi^{i}_{j,k}(t)$ was introduced as Eq. (1) in [16] with the first wavelet function, which is known as the mother wavelet $\psi^{1} = \psi(t)$.

$$\psi^{i}_{j,k}(t) = 2^{j/2}\,\psi^{j}\left(2^{j}t - k\right),\ i = 1,2,3,\ldots \tag{1}$$

WPT consists of a set of quadrature filters (H and G), which decompose a signal into j layers with the following expression:

$$f(t) = \sum_{i=1}^{2^{j}} [Hf^{i}_{j}(t) + Gf^{i}_{j}(t)]. \tag{2}$$

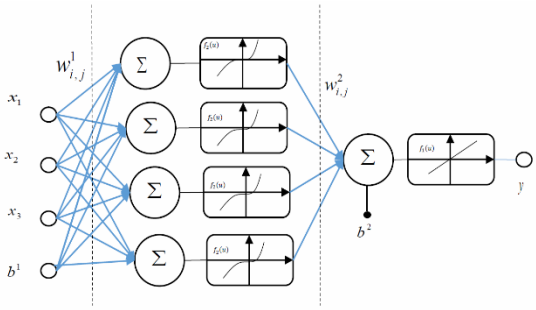

Fig. 1. Typical MLP network architecture.

Wavelet decomposition requires prior information, such as the decomposition level and mother wavelet type. Despite the absence of a universal wavelet that can perform optimally for all types of signals, Daubechies wavelets, such as db2 to db35, are considered ideal [17], whereas a Morlet wavelet can be effective in representing transient vibration signals [18]. Subsequently, WP-based energy is introduced as a reliable indicator that can present signal energy in various frequency bands.

WP energy, such as energy entropy and Shannon energy entropy, has several definitions. WP energy entropy and log-energy entropy at a specified node in wavelet decomposition can be computed using Eqs. (3) and (4), where ? presents the wavelet packet coefficients at time ? and node n.

$$Ent_{n} = \sum_{n} |C_{n,j}|^{2} \tag{3}$$

$$Entl_{n} = \sum_{j} \log(C^{2}_{n,j}) \tag{4}$$

3. ANNs

ANNs are effective processing tools for nonlinear problems with specific purposes, such as pattern recognition or data classification [19] through learning procedures, to achieve fault diagnosis. ANN applications use a simple multilayer perceptron (MLP) as common network architecture. This network is first propagated by inputs to provide the error value that corresponds to each node. The error value is then transmitted back to adjust network weights. A typical MLP network architecture is depicted in Fig. 1. The neuron values are normalized within a specific range by passing an activation function, and network output can be executed in matrix form as Eq. (5) [20], where W denotes the weight matrix, b is the bias vector, and f is the activation function.

$$y = f^{2}\left(w^{1} f^{1}\left(w^{1}x + b^{1}\right) + f^{2}\right) \tag{5}$$

$$y_{i} = \log\left(\frac{1}{1 + e^{-x}}\right) \tag{6}$$

In this study, vibration characteristics are used as network inputs, which are selected through frequency- and time-based characterization approaches. A gradient descent technique is applied to the error value to find the desired value of the

weight in each neuron with the aid of a logarithmic unipolar sigmoid function (Eq. (6)).

In addition, the number of inputs and hidden nodes considerably impacts the accuracy of networks. A large number of inputs can lead to an undertrained or overtrained network. Hence, the number of nodes in a hidden layer with respect to the number of inputs can be estimated using empirical Eq. (7) [6].

$$L = \sqrt{M + N} + \alpha \,, \tag{7}$$

where M and N denote the node numbers in the input and output layers, respectively; and α is a constant value between 1 and 10.

4. Experimental verification

4.1 Experimental rig

An experimental setup was developed to measure impeller vibration. The experiment was performed using seven impeller conditions under fluctuating flow conditions. Impeller crack and imbalance were staged with increasing severity levels. The system was run for approximately 27 s over a range of variable flow rates (0 l/min to 90 l/min) and comprised a centrifugal pump with a closed-loop piping system. A single suction pump with a top discharge operating at 2700 rpm with a single-phase electric motor was used. The pump was a CM-50 Pentax series with the following construction features: a cast-iron body and motor bracket, a Noryl® impeller, and a ceramic-graphite mechanical seal that delivers water at a rate of up to 90 l/min.

4.2. Experimental measurements

The experimental data were collected in accordance with the procedure illustrated in Fig. 2. A triaxial accelerometer (Fig. 3, bottom right) with a sensitivity of 10 mV/g and a Quantum X-4 channel data acquisition system were used to collect the data at 4800 samples per second. The accelerometer was mounted on top of the pump casing due to the following reasons: (1) to keep the accelerometer in a horizontal position and (2) to collect vibration data that are more sensitive to the variation in fluid dynamics derived from impeller defects in the radial and axial directions [8]. Two pressure gauges were installed to measure suction and discharge pressures, whereas a Venturi flowmeter was installed to obtain pump performance curves. Two flow valves were installed at the inlet line of the pump and reservoir discharge to control and modify the flow rate at the suction side and the water height of the reservoir. The closed-loop water system is shown in Fig. 3(top). A tachometer (Fig. 3, bottom left) was positioned along the free end of the rotor shaft to detect the operational influence on the pump (rpm). X and Z denote the radial direction, whereas Y denotes the axial direction based on the coordinate system shown in Fig. 3 (bottom right).

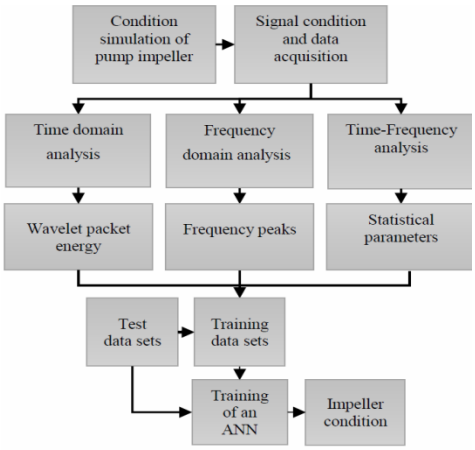

Fig. 2. Scheme of the proposed method.

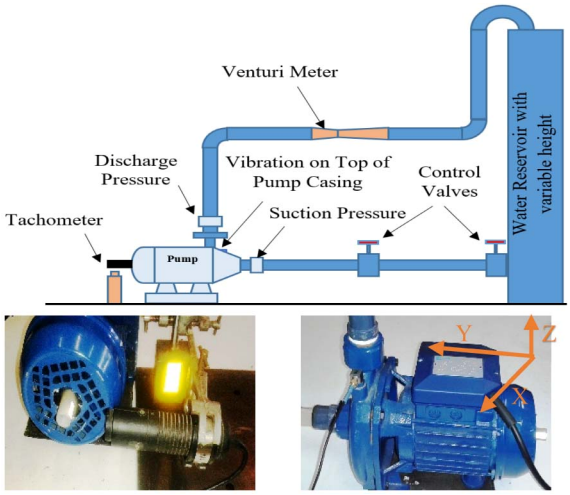

Fig. 3. (top) Closed-loop water system; (bottom left) location of tachometer; (bottom right) accelerometer.

4.3 Fault simulation

Corrosion and external solid materials in fluids are the major factors that contribute to impeller damage [4]. Hence, cracking and imbalance conditions with three severity levels (low, medium and high) were manually simulated using a hacksaw and by delivering a hammer blow. As shown in Fig. 4, the crack location was close to the impeller eye, whereas imbalance was introduced to the external perimeter of the impeller. Closed-type impellers were used in the setup; thus, the damages were induced at the locations of the blades. Table 1 outlines the condition types and the relevant fault dimensions staged on the impeller.

5. Feature extraction

Data samples were collected under a constant flow rate of 60 l/min and a flow variation ranging from 0 l/min to 90 l/min. Then, 4 measurements for each condition (each condition contained a data set of 259200 samples) were performed.

Table 1. Specifications of impeller damage.

Impeller condition	Fault type
Normal (*N*)	Without fault
Fault 1 (*C1*)	Crack depth ≃ 10 mm
Fault 2 (*C2*)	Crack depth ≃ 20 mm
Fault 3 (*C3*)	Crack depth ≃ 30 mm
Fault 4 (*U1*)	Unbalance depth ≃ 4 mm
Fault 5 (*U2*)	Unbalance depth ≃ 8 mm
Fault 6 (*U3*)	Unbalance depth ≃ 12 mm

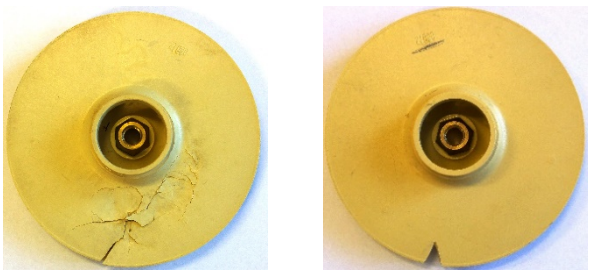

Fig. 4. Damage induced on the impeller: (left) Impeller with crack; (right) unbalanced impeller.

Among the 7 impeller conditions, 12 vibration signals were observed under each condition. Therefore, 84 raw vibration signals were measured and used to train ANNs.

5.1 Time domain features (TDFs)

The following statistical parameters were calculated from the measured vibration: mean μ, RMS, kurtosis K_t, skewness S_n, energy E, crest factor C_r, SD, impulse factor C_r, and peak-to-peak value. The selection of appropriate feature sets is crucial and can enhance the accuracy rate of a diagnosis system. No standard method for selecting the types of features has yet been established; thus, the statistical parameters that are sensitive to flow variation were selected in this study. In this regard, the maximum likelihood estimation (MLE) method was used to map a simple linear model, i.e., $Y(x) = \beta^T X + \int = \sum_{j=0}^{p} \beta_j x_j + \int$, onto each feature set to individually identify the average absolute value of slope (β) for each feature model. The results show that only the following features are sensitive to the operational condition with the lowest uncertainties: RMS, kurtosis K_t, skewness S_n, and SD. Fig. 5 depicts these features and the MLE linear model with regard to their predicted slope parameter.

The results indicate an apparent difference in the vibration level in the axial direction (*Y*-axis). Such difference is attributed to the applied defects on the impellers, which lead to an asymmetric flow in the system and generate large flow impacts with a high vibration level in the axial direction. Fig. 6(left) shows the selected features over constant flow for 60 s, whereas Figs. 6(middle) and (right) present these features as

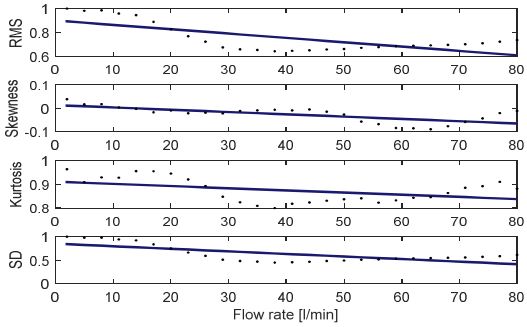

Fig. 5. Selected statistical features and their predicted linear model using the MLE method.

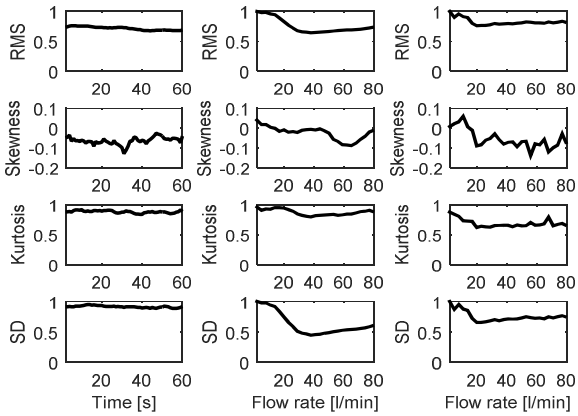

Fig. 6. TDFs of the normal impeller under (left) A constant flow rate; (middle) a variable flow rate; (right) TDFs of the damaged impeller (C1) under a variable flow rate.

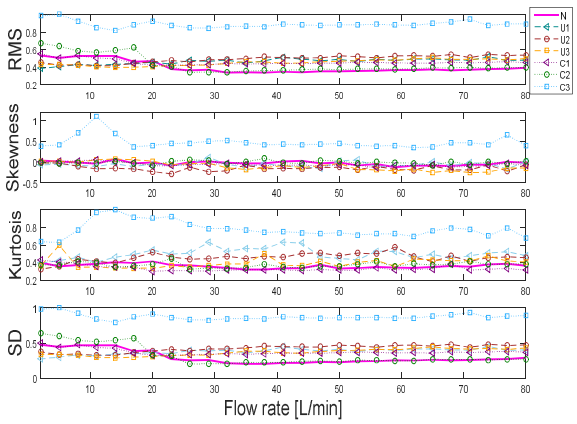

Fig. 7. Behavior of TDFs under a variable flow rate.

functions of the flow variation under normal condition and crack damage level 1 (C1), respectively. A careful consideration of these figures shows a pattern at the vibration level related to the water flow rate and the applied defects. Fig. 7 illustrates the features of all the conditions against one another over flow variation.

Evidently, the vibration parameters are generally high for flow rates below 20 l/min due to cavitation. By contrast, cer-

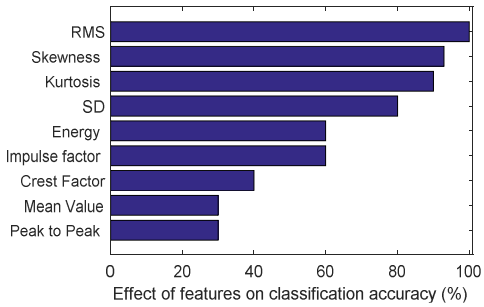

Fig. 8. Comparison of the effects of features on classification accuracy.

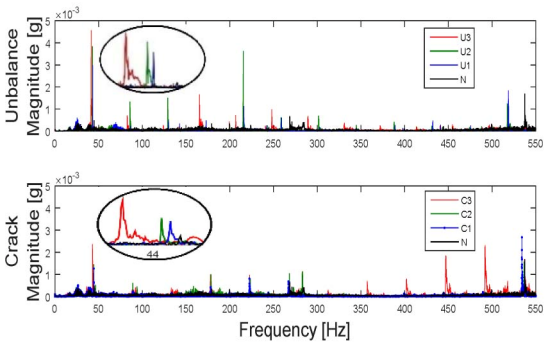

Fig. 9. Vibration spectra of low-frequency range: (top) Unbalanced impeller vs. normal impeller; (bottom) impeller with crack vs. normal impeller.

tain features (e.g., crest factor and mean value) do not exhibit a clear pattern. The crack has a more significant impact on vibration magnitude, whereas C3 exhibits a higher energy value compared with the imbalance. This finding is attributed to changes in impeller geometry, which considerably alter the flow path. Therefore, RMS and SD are more reliable than the other features. Meanwhile, fault classification performance was evaluated for the aforementioned features. The RMS value exhibits high sensitivity to flow variation and high performance in fault classification; thus, the other features are normalized according to this value. The results are presented (in %) in Fig. 8.

5.2 Frequency domain features

Pump frequency (rpm), vane pass frequency (VPF), and their corresponding harmonics are considered valuable and indicative parameters. The analysis of an impeller with normal (N) condition operating at constant and varying flow rates implies high amplitude for 1 rpm, VPF, 2VPF, 3VPF and 6VPF, and a significant increase in noise level for measurements under flow variation. Moreover, the addition of damages with relevant severity levels (Table 1) to the experiment show that crack C influences the high harmonics of rpm, whereas the unbalanced condition U affects the first rpm and VPF.

Hence, spectral behavior was investigated at low- (1 Hz to 1 kHz) and high-frequency ranges (1 kHz to 2.4 kHz).

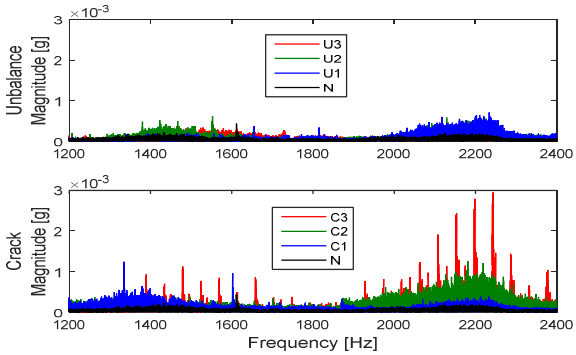

Fig. 10. Vibration spectra of high-frequency range: (top) Unbalanced impeller vs. normal impeller; (bottom) impeller with crack vs. normal impeller.

5.2.1 Spectrum characteristics of low-frequency range

Fig. 9 shows the frequency spectra that correspond to impeller conditions (Table 1). The low-frequency range plot depicts nearly the same broadband, which is influenced by flow turbulence. Fundamental frequency components are visible with high amplitudes, which are attributed to the interaction between flow and impeller vanes. A comparison of the U and C conditions with the normal condition indicates that the 1 rpm amplitude increases as the severity of damage increases. However, imbalance effects occur with higher amplitude for 2 rpm to 5 rpm, whereas crack excitation appears in high harmonics. Furthermore, a considerable difference in frequency location can be observed for each damage, where rpm and damage severity appear to be inversely proportional. The 1 rpm for the normal and U3 conditions appears at 45 Hz (2700 rpm) and 41.33 Hz (2480 rpm), respectively.

5.2.2 Spectrum characteristics of high-frequency range

Fig. 10 shows that high orders of rpm are excited under crack conditions. Spectral energy is also increased significantly at frequencies over 2000 Hz, which is due to a pressure drop in the pump suction. Fig. 10(bottom) depicts a high amplitude of U3, thereby indicating that the vibration energy of the rpm harmonics becomes stronger as the severity of defects increases.

The system operates under fluctuating flow conditions; hence, the fundamental frequency amplitudes and broadband energies are studied as functions of flow. Fig. 11 shows the behavior of five selected frequency components against a varying flow rate. Despite the considerable effects of imbalance at low frequencies (1 rpm and VPF), as shown in Figs. 11(a) and (b), respectively, no specific behavior pattern related to defect severity and flow rate is illustrated. By contrast, Figs. 11(c)-(e) explicitly show the influence of crack damage as the number of harmonic orders increases.

5.3 Time-frequency features

Vibration signals were decomposed into numerous fre-

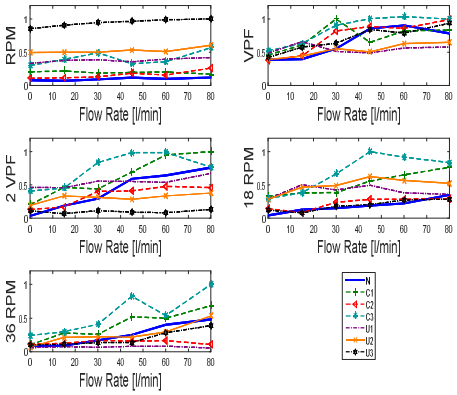

Fig. 11. Amplitude evolution of the frequency components: (a) 1 × rpm; (b) VPF; (c) 2 × VPF; (d) 18 × rpm; (e) 36 × rpm.

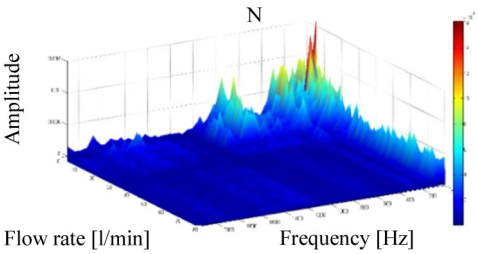

Fig. 12. Time-frequency plot of an impeller with normal condition.

quency bands with WPT to extract time-frequency features. Wavelet basis and decomposition level were selected based on the achieved performance in terms of classification accuracy using different types of wavelets and decomposition levels. Therefore, a 5-layer wavelet packet decomposition using Daubechies wavelet (db20) was performed, and 32 frequency bands, with each band covering a frequency interval of 75 Hz, were obtained. Fig. 12 simultaneously shows the feature maps of impeller vibration in the frequency and time domains. The flow rate evidently affects vibration energy, in which the fundamental frequency components 1 × rpm and VPF are clearly below 300 Hz.

The rpm, VPF frequencies, and corresponding harmonics are excited and increased in low flow rates. In addition, the minimum noise energy is observed in flow rates higher than 20 l/min. The C and U conditions, with their corresponding severities, are presented in Fig. 13. The influence of cracks is evident at high rpm harmonics, whereas the imbalance effects are clearly shown to be higher at 1 rpm.

Moreover, the comparison of the Shannon energy distribution of all the sub-bands and log-energy distributions indicates explicit information regarding log-energy distribution. Fig. 14 depicts the normalized log-energy distribution of the 32 frequency sub-bands. As shown in the figure, the amount of energy for C conditions is observed at high frequencies, whereas the imbalance effects are observed at low frequencies. The energy of frequency sub-bands from WPT varies under each impeller condition; hence, it can represent features that exhibit

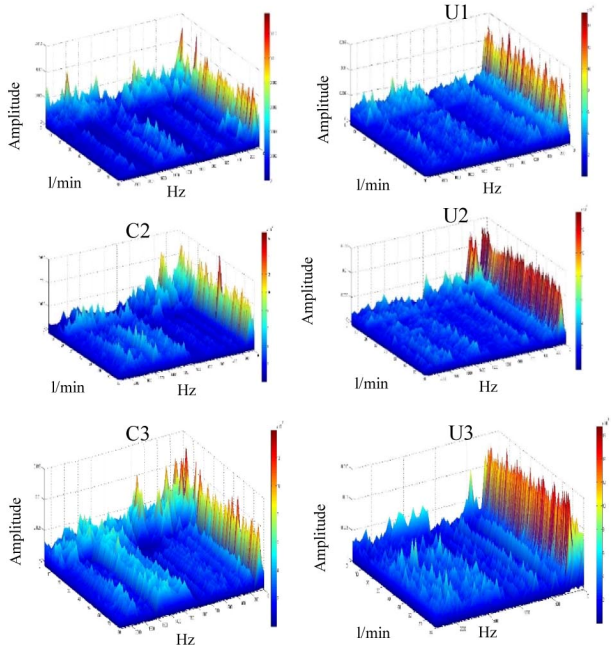

Fig. 13. Time-frequency plot of impellers with faulty conditions: impeller with crack (left column); unbalanced impeller (right column).

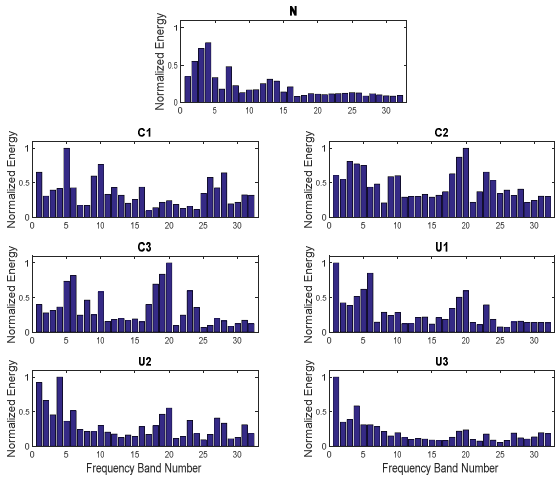

Fig. 14. Normalized log-energy distribution of the frequency sub-bands of impellers under different operational conditions.

the characteristics of diagnostic faults.

6. ANN Performance

A three-layer feedforward ANN with an input layer, processing units (hidden layer), and an output layer was adopted in this section. Seven sets of features formed the inputs and were divided into three independent subsets to be used for training, validating, and testing the networks. Three ANN structures were designed based on the extracted features from TDA, FA and TFA. Four features were selected from TA, namely RMS, kurtosis, skewness, and SD. Five frequency peaks that correspond to the pump rpm represent the spectral features. Finally,

Table 2. ANN classification accuracy rates for different transfer functions.

Feature set	Transfer function			
	(tansig, tansig)	(tansig, purelin)	(logsig, tansig)	(logsig, purelin)
TDA features	0.05158	0.0822	0.1028	0.0839
FA features	0.0621	0.1009	0.1501	0.0915
WPT features	0.0969	0.0885	0.1185	0.0617

Table 3. ANN classification accuracy rates for different numbers of hidden layer nodes.

No. of nodes	4	6	8	10	12	14	16	18
TDA features%	55.33	59.5	61.15	**63.45**	60.01	61.01	52.8	56.2
FA features%	33	58.64	**73.84**	68.22	65.45	65.56	58.25	60.9
WPT features%	51.74	55.29	59	61.30	90.51	63.82	75.81	**96**

32 features, including frequency sub-band energy, were selected from TFA.

The preliminary number of nodes in the hidden layer is estimated using Eq. (7) as follows. For the TDA feature sets (4 input features), the network must include 4 to 14 hidden nodes; for the FA feature sets (5 input features), 5 to 15 hidden nodes should be included; and for the WPT feature sets (32 input features), the ANN must have 8 to 18 hidden nodes. Therefore, a set of training tests (within the estimated range of the hidden layer's nodes) with three popular transfer functions and learning algorithms, namely, the Levenberg-Marquardt (LM), gradient descent, and quasi-Newton algorithms, were performed to select the most appropriate ones. This process was repeated 10 times for new testing data sets, and the average results demonstrated that the logarithmic sigmoid function for the hidden layer, the linear function for the output layer, and the LM algorithm achieved the lowest mean square error (MSE) and could provide satisfactory results in the training and testing steps. Table 2 shows the average MSE of the results. Afterward, a comparison of classification accuracy was performed for different numbers of nodes in the hidden layer, as shown in Table 3. The best node numbers that can stabilize accuracy and reduce the overfitting problem are 10, 8 and 18 for each type of feature set. Moreover, the accuracy of ANNs was examined for a number of nodes from the given range and similar results were obtained.

New data sets were accordingly provided to train, validate, and test the ANNs to verify the superiority of the ANN based on WPT features. The network performance and classification accuracy of each feature set from each methodology were examined through the confusion matrix and the network convergence error while the maximum number of training epochs was set to 1000 for each training process. The convergence error shows the convergence process of the network to the lowest output error over the training epochs. The training process stops at the point before the validation error starts to increase, which indicates overfitting of the training data. Finally, the confusion matrix tabulates the proportion of correct predictions for each possible impeller condition.

Table 4. Confusion matrix of an ANN trained using TDA features (C = Condition).

Condition	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**Accuracy**
C1	**14.3 %**	0	0.5 %	0	0	0.5 %	0	93.1 %
C2	0	**13.8 %**	2.6 %	1.1 %	0	0.5 %	0	76.5 %
C3	0	0	**14.3 %**	1.1 %	0	0.5 %	0	87.5 %
C4	0	0	0	**10.6 %**	0	0	0	100 %
C5	0	0.5 %	0	1.6 %	**14.3 %**	0	0	87.1 %
C6	0	0	0	0	0	**12.7 %**	0	100 %
C7	0	0	0	0	0	0	**14.3 %**	100 %
Accuracy	100 %	96.3 %	77.8 %	74.1 %	100	88.9 %	100 %	91 %

Table 5. Confusion matrix of an ANN trained using FA features.

Condition	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**Accuracy**
C1	**14.3 %**	0	0	0	0	0	0	100 %
C2	0	**14.3 %**	0	0	2.4	0	0	85.7 %
C3	2.4 %	0	**9.5 %**	0	0	0	0	100 %
C4	0	0	0	**14.3 %**	0	0	0	100 %
C5	0	0	4.8	0	**11.9 %**	0	0	71.4 %
C6	0	0	0	0	0	**14.3 %**	0	100 %
C7	0	0	0	0	0	0	**14.3 %**	100 %
Accuracy	100 %	100 %	66.7 %	100 %	83.3	100 %	100 %	**92.9 %**

Table 6. Confusion matrix of an ANN trained using WPT features.

Condition	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**Accuracy**
C1	**14.3 %**	0	0	0	0	0	0	100 %
C2	0	**11.9 %**	0	0	0	0	0	100 %
C3	0	0	**14.3 %**	0	0	0	0	100 %
C4	0	0	0	**14.3 %**	0	0	0	100 %
C5	0	0	0	0	**14.3 %**	0	0	100 %
C6	0	0	0	0	0	**14.3 %**	0	100 %
C7	0	2.4 %	0	0	0	0	**14.3 %**	85.7 %
Accuracy	100 %	83.3 %	100 %	100 %	100 %	100 %	100 %	**97.6 %**

From the first ANN trained using the TDA feature sets, MSE was measured at the lowest validation error as 0.017 for the 26th iteration. Table 4 includes the individual accuracy of each class and the percentage of the correct and incorrect classifications.

For the second ANN, five frequency peaks that corresponded to the rpm of the impeller and VPF were selected. The lowest validation error was found within an iteration range of 12 to 20, with a minimum value of 0.0165. Table 5 shows that the accuracy of the overall network performance is 92.9 %.

Moreover, the behavior of the ANN network error trained using the energy values from WPT achieved the lowest value (0.001). Table 6 indicates an overall accuracy of 97.6 % with only one misclassification, i.e., Class 7 (unbalanced impeller with a severity level of 3) is misclassified as Class 2 (crack level 2).

7. Results and discussion

Nonstationary flow conditions can produce nonlinear characteristics in pumping. Thus, the use of signal processing techniques and ANNs in diagnosing centrifugal pump impellers, which operate under fluctuating flow rates, was examined in this work. An adequate technique was introduced to reduce uncertainties regarding the occurrence of flow fluctuations. An experiment was performed under impeller fault conditions, such as cracking and imbalance. The vibration signals were analyzed in the time, frequency, and time-frequency domains. The results proved that fluctuating flow conditions for an impeller without damage might lead to a slight change in pump rotational speed. However, a considerable change can be observed on pump rpm as a defect starts to grow. Furthermore, the features collected from nondimensional training data sets were used to train ANNs. This online fault detection system requires prior off-line work to train the classifier using data from the experimental setup. It can help identify impeller conditions and quantify the severity of damages. Training algorithms, network hidden nodes, and the effectiveness of different transfer functions were compared, and the results proved that the accuracy of ANN prediction was considerably improved by using decomposed vibration signals and energy-based features. The network accuracy results indicate that the WPT-ANN model achieves low MSE, high correlation coefficients, and fast training time, which can save computational time. This model can be used effectively to classify impeller defects under nonstationary conditions.

The effectiveness of the proposed method was demonstrated through the induced impeller conditions under unsteady flow rates. However, this model can be extended and improved to deliver promising results in fault detection for extensive applications. To further improve the model, ANN performance should be maximized using optimization algorithms, such as genetic and evolutionary algorithms, and threshold impact on the classification of new damages with sizes between the ones used in this study should be investigated.

Acknowledgment

This research is supported by the Rand Water Company, and their grant is greatly acknowledged.

References

[1] R. Birajdar, R. Patil and K. Khanzode, Vibration and noise in centrifugal pump sources and diagnosis methods, *3rd International Conference on Integrity, Reliability and Failure*, no. July (2009) 20-24.

[2] L.-W. Lee and I.-H. Li, Design and implementation of a robust FNN-based adaptive sliding-mode controller for pneumatic actuator systems, *J. Mech. Sci. Technol.*, 30 (1) (2016) 381-396.

[3] V. Muralidharan, V. Sugumaran and V. Indira, Fault diagnosis of monoblock centrifugal pump using SVM, *Eng. Sci. Technol. an Int. J.*, 17 (3) (2014) 152-157.

[4] P. P. Harihara and A. G. Parlos, Sensorless detection of impeller cracks in motor driven centrifugal pumps, *Volume 5: Design, Analysis, Control and Diagnosis of Fluid Power Systems* (2008) 17-23.

[5] R. Zouari and S. S. Menad, Fault detection system for centrifugal pumps using neural networks and neuro-fuzzy technique, *CMSM 2005* (2005).

[6] Y. F. Xing, Y. S. Wang, L. Shi, H. Guo and H. Chen, Sound quality recognition using optimal wavelet-packet transform and artificial neural network methods, *Mech. Syst. Signal Process.* (2015) 1-18.

[7] J. Patel, V. Patel and A. Patel, Fault diagnostics of rolling bearing based on improve time and frequency domain features using artificial neural networks, *Int. J. Sci. Res. Dev.*, 1 (4) (2013) 781-788.

[8] B. Kamiel, G. Forbes, R. Entwistle, I. Mazhar and I. Howard, Impeller fault detection for a centrifugal pump using principal component analysis of time domain vibration features, *Surveill. 7 Int. Conf.* (2013) 1-12.

[9] X. M. Zhao, Q. H. Hu, Y. G. Lei and M. J. Zuo, Vibration-based fault diagnosis of slurry pump impellers using neighbourhood rough set models, *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, 224 (4) Apr. (2010) 995-1006.

[10] A. K. Nandi, C. Liu and M. L. D. Wong, Intelligent vibration signal processing for condition monitoring, *Proceedings of the International Conference Surveillance,* 7 (2013) 1-15.

[11] E. Chen, Z. Ma, G. Zhao, G. Li, A. Yang and G. Nan, Numerical investigation on vibration and noise induced by unsteady flow in an axial-flow pump, *J. Mech. Sci. Technol.*, 30 (12) (2016) 5397-5404.

[12] S.-G. Park, H.-J. Sim, H.-J. Lee and J.-E. Oh, Application of non-stationary signal characteristics using wavelet packet transformation, *J. Mech. Sci. Technol.*, 22 (11) (2008) 2122-2133.

[13] H. Yuan, J. Chen and G. Dong, An improved initialization method of D-KSVD algorithm for bearing fault diagnosis, *J. Mech. Sci. Technol.*, 31 (11) (2017) 5161-5172.

[14] D. Wang and Q. Miao, Smoothness index-guided Bayesian inference for determining joint posterior probability distributions of anti-symmetric real Laplace wavelet parameters for identification of different bearing faults, *J. Sound Vib.*, 345 (2015) 250-266.

[15] K. Shibata, A. Takahashi and T. Shirai, Fault diagnosis of

rotating machinery through visualisation of sound signals, *Mech. Syst. Signal Process.*, 14 (2) (2000) 229-241.

[16] R. Shao, W. Hu, Y. Wang and X. Qi, The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform, *Meas. J. Int. Meas. Confed.*, 54 (2014) 118-132.

[17] Y. Wang, G. Xu, L. Liang and K. Jiang, Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis, *Mech. Syst. Signal Process.*, 54-55 (2015) 259-276.

[18] L. Toth, On finding better wavelet basis for bearing fault detection, *Acta Polytech. Hungarica*, 10 (3) (2013) 17-35.

[19] E. Kim, Y. Lee and S. Lee, Heath monitoring of a glass transfer robot in the mass production line of liquid crystal display using abnormal operating sounds based on wavelet packet transform and artificial neural network, *J. Sound Vib.*, 331 (2012) 3412-3427.

[20] M. Unal, M. Onat, M. Demetgul and H. Kucuk, Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network, *Meas. J. Int. Meas. Confed.*, 58 (2014) 187-196.

Amin Jami obtained his B.Eng. in 2013 in Iran, his B.Sc. (*Hons*) in 2015, and his M.Eng. in 2016 (*cum laude*) from the University of Pretoria in South Africa. He is currently a Ph.D. student in the Center for Asset Integrity Management of the University of Pretoria. His research interests include condition monitoring, vibration analysis, and machine and structural health diagnostics.

Stephan Heyns obtained his B.Sc. in Mechanical Engineering in 1978 (*cum laude*) and his Ph.D. in 1988 from the University of Pretoria in South Africa. In 1982, he joined the Department of Mechanical and Aeronautical Engineering at the University of Pretoria. He is currently a director of the Center for Asset Integrity Management in the university. His research interests include machine and structural health diagnostics and prognostics, vibration analysis, vibration measurement techniques, and model updating.