Article

Journal of Vibration and Control
2015, Vol. 21(3) 461–467

DOI: 10.1177/1077546313490778
jvc.sagepub.com
SAGE

Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier

Rohit Tiwari, Vijay K Gupta and PK Kankar

Abstract
The rolling element bearing is among the most frequently encountered component in a rotating machine. Bearing fault can cause machinery breakdown and lead to productivity loss. A bearing fault diagnosis method has been proposed based on multi-scale permutation entropy (MPE) and adaptive neuro fuzzy classifier (ANFC). In this paper, MPE is applied for feature extraction to reduce the complexity of the feature vector. Extracted features are given input to the ANFC for an automated fault diagnosis procedure. Vibration signals are captured for healthy and faulty bearings. Experiment results pointed out that proposed method is a reliable approach for automated fault diagnosis. Thus, this approach has potential in diagnosis of incipient bearing faults.

Keywords
Adaptive neuro-fuzzy classifier, fault diagnosis, multi-scale permutation entropy

1. Introduction

Automated fault diagnosis plays an important role in industries by avoiding catastrophic accidents and machinery malfunction. In order to keep machines safe many fault diagnosis techniques have been developed in which vibration signal analysis is one of the most widely used techniques. Vibration signal analysis can be done in time, frequency and time-frequency domains.

Vibration signals often exhibit nonlinearity due to factors of nonlinear stiffness and clearance of bearing. For this reason, such systems can only be described by nonlinear dynamic models. Commonly used signal processing techniques in time and frequency domains are designed for linear vibration signals. On the other hand, various nonlinear parameter estimation techniques can be a possible substitute for fault associated feature extraction in complex bearing vibration signals. Combined parametric effects have been analyzed to predict the dynamic response of a rotor-bearing system using a response surface method (Kankar et al., 2012a).

Many nonlinear parameter identification schemes have been applied to fault diagnosis. Yang et al. (2007a) have applied correlation dimension in such cases, moreover this technique is not suitable for real time monitoring due to the demand of very long data sets. Appropriate entropy is the another entropy which was exploited to access the status of rotary machines but is not preferred due to its dependency on record length and lower estimation value (Richman and Moorman, 2000). To overcome this shortcoming, a sample entropy is introduced which attracted a lot of attention. Costa et al. (2002, 2005) have carried out a multi-scale entropy (MSE) analysis of biological signals. Then, Zhang et al. (2010) used MSE in bearing fault diagnosis and identification of fault severity and showed the high accuracy of this enhanced method to evaluate the regularity of complex time series. Permutation entropy (PE) was introduced by Bandit and Pompe (2002), for the complexity analysis of time domain data by using the comparison of neighboring values. The PE method was successfully applied in numerous applications like electroencephalography (EEG) signal analysis (Bruzzo et al., 2008), tool breakage detection in end milling (Li et al., 2008)

PDPM-Indian Institute of Information Technology Design and Manufacturing Jabalpur, Jabalpur, India

Received: 2 March 2013; accepted: 6 April 2013

Corresponding author:
PK Kankar, PDPM-Indian Institute of Information Technology Design and Manufacturing Jabalpur, Jabalpur, India.
Email: pavankankar@gmail.com

and complexity of a power electronics based system (El-Mezyani et al., 2012). Time series data captured from mechanical systems are complex in nature and contain multiple temporal scale structures. Therefore, the PE based on a single scale structure shows inefficient analysis results on these complex data. To avoid such a problem, Aziz and Arif (2005) have introduced the concept of multi-scale permutation entropy (MPE) to estimate entropy across multiple scales. Wu et al. (2012) have applied MPE and a support vector machine for bearing fault diagnosis.

After feature extraction, a classifier is exploited to achieve automated fault diagnosis. Various classification techniques have been applied to bearing fault diagnosis such as support vector machines (SVM) (Yang et al., 2007b; Kankar et al., 2011a,b; Kankar et al., 2012b) and neural network (NN) (Vyas and Kumar, 2001; Samanta and Al-Balushi, 2003; Kankar et al., 2011a,b, 2012b). Kankar et al. (2013) have presented a feature-recognition system for rolling element bearings fault diagnosis using cyclic autocorrelation of raw vibration signals. Fuzzy logic is also employed for fault diagnosis but it lacks learning ability. To overcome this shortcoming fuzzy logic is combined with NNs to obtain a hybrid model which contains the ability of both fuzzy logic and NNs. The adaptive neuro fuzzy classifier (ANFC) is such a system in which NNs will provide learning ability to fuzzy logic and fuzzy logic will offer a high level IF-THEN rule thinking to deal with system uncertainty. The ANFC is successfully implemented in image recognition, machine fault diagnosis and disease diagnostics.

In the present work, vibration signals for healthy and defective bearings are considered. Three defects have been simulated on bearings which include inner race defect (IRD), outer race defect (ORD) and ball defect (BD). Then a new technique using MPE and ANFC is utilized for bearing fault diagnosis. Firstly the MPE is estimated for 16 scales and these values are partitioned into training and test set for fault classification using ANFC. But before dividing the data in to training and test set the mean value is calculated to reduce the dimension of data. The diagnosis results show that this method is an effective and reliable technique for machine condition monitoring. When compared with a time domain statistical feature method this method displays a better performance in terms of accuracy.

2. Multi-scale permutation entropy

The physical and biological system shows nonlinear and complex behavior, time series complexity analysis of such a system is very useful. The MPE is employed by Aziz and Arif (2005) for the estimation of complexity parameters. The MPE calculates PE over multiple scales to avoid contradictory results by a single scale entropy. In the case of Shannon entropy, the sequential relation between values of the time series is neglected. This is more useful for a linear system while MPE employs the comparison of neighboring values for analysis of complex time domain data. This property of MPE makes it more useful for the analysis of non-stationary signals.

Firstly, time series data $t = \{t_1, t_2, \ldots, t_N\}$ is converted into multiple coarse grained time series. This can be done by taking the average of the data inside non overlapping windows of length t. coarse grain time series shown in Figure 1 can be expressed using following equation.

$$c_n^{(l)} = \frac{1}{l} \sum_{i=(n-1)l+1}^{nl} t_i \tag{1}$$

Now from each coarse grain series the PE is estimated (Costa et al., 2002). PE is the function of l.

Therefore, procedure for PE is given as

- Course grained time series is transformed in to m dimensional space

$$c_n^m = [c(n), c(n+d) \ldots c(n+(m-1))d] \tag{2}$$

where m is embedded dimension and d is delay time.

- Now arranging in increasing order

$$[c(n+(n_1-1))d \le c(n+n_2-1)d \ldots\ldots \le c(n+(n_m-1))d] \tag{3}$$

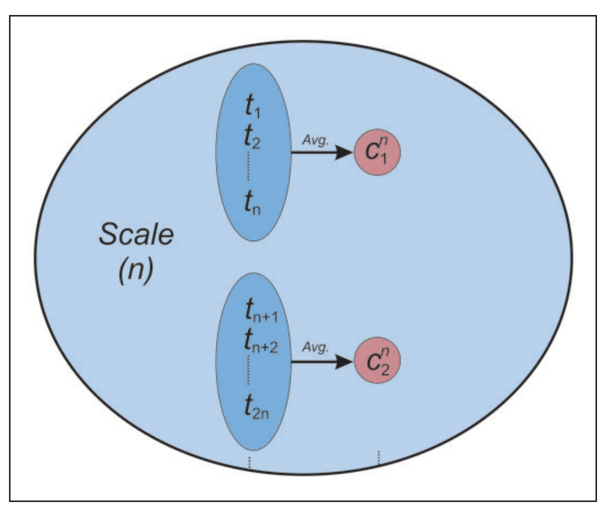

Figure 1. Schematic illustration of the course-grained time series for scale *n*.

Therefore, for m-tuple vector there is $m!$ possible permutations. Probability distribution for m distinct *symbols* $[1, 2 \ldots m]$ *be* $Y_1, Y_2, \ldots Y_k$ *where* $k \leq m!$.

- Then PE of m dimension is defined as

$$H_{pe}(m) = -\sum_{i=1}^{k} Y_i \ln Y_i \tag{4}$$

Maximum value of $H_{pe}(m)$ is $log(m!)$. where all possible permutation shows the same probability.

3. Adaptive neuro fuzzy classifier

The neuro fuzzy classifier is an adaptive network based system in which the antecedent parameters are adapted with NNs. This combined system with a fuzzy logic qualitative approach and artificial NN adaptive capabilities named as ANFC. The ANFC explicates a zero order surgeon fuzzy inference model (Cetisli, 2010) in to the framework of a multilayer artificial neural network (ANN) with adaptive and non-adaptive nodes. The ANFC is based on fuzzy rules and to initialize fuzzy rules the k-mean algorithm is used. The ANFC regulates the membership function and other antecedent parameters using the scaled conjugate gradient (SCG) algorithm. Moller (1993) has described SCG as two times faster than the back propagation algorithm because of its super linear convergence rate.

For two inputs $\{x_1,x_2\}$ and one output y fuzzy classification rule is defined as

$$\text{If } X_1 \text{ is } A_1 \text{ and } X_2 \text{ is } A_2 \text{ then } y \text{ is } C_1 \text{ class}$$

where A_1 and A_2 are the linguistic terms that are defined on feature space X_1 and X_2 and C_1 represents class label of the output y.

In this architecture each node in the same layer has the same node function. The structure of the classifier is shown in Figure 3. The first layer generates the membership grade of each input to a specified fuzzy region. In this layer for the membership function (MF) bell shape, Gaussian, triangular and trapezoidal functions can be used. Gaussian function has less parameter and because of smooth partial derivatives of its parameters it is utilized as MF.

The Gaussian MF is described as

$$\gamma_{ij}(x_{cj}) = \exp\left(-0.5\frac{(x_{cj} - \alpha_{ij})^2}{\delta_{ij}^2}\right) \tag{5}$$

where x_{cj} is the input variable and α_{ij} and δ_{ij} are the centre and width of the Gaussian function respectively. The Nnext layer is the rule layer which uses the membership values of input to calculate the firing strength of fuzzy rules. So the θ_{ic} firing strength of the i^{th} rule is

$$\theta_{ic} = \prod_{j=1}^{N} \gamma_{ijc} \tag{6}$$

This layer describes the fuzzy rules for $x_c$sample. N is the number of features. The third layer in the classifier calculates the weighted outputs. The maximum firing strength of the rules decides the output class. If the rule output weight for a class is greatest among the other class weights it means that the particular class region is controlled by that rule. Weighted output ϕ_{ck} for the c^{th} sample in the k^{th} class can be shown as

$$\phi_{ck} = \sum_{i=1}^{M} \theta_{ic}\omega_{ik} \tag{7}$$

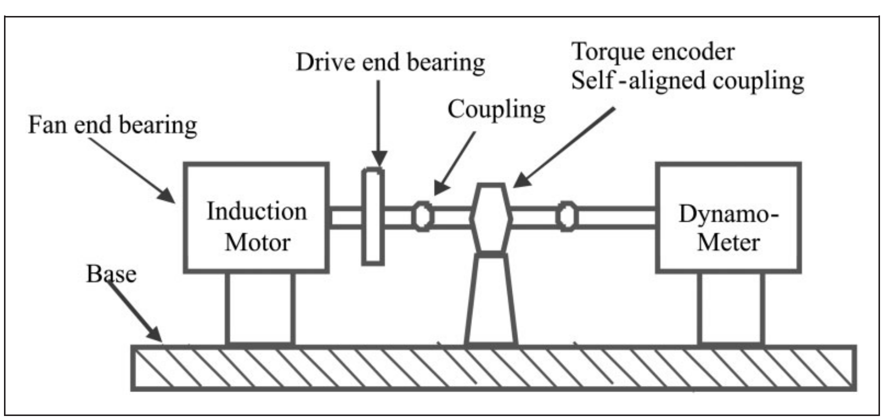

Figure 2. Schematic diagram of the experimental setup.

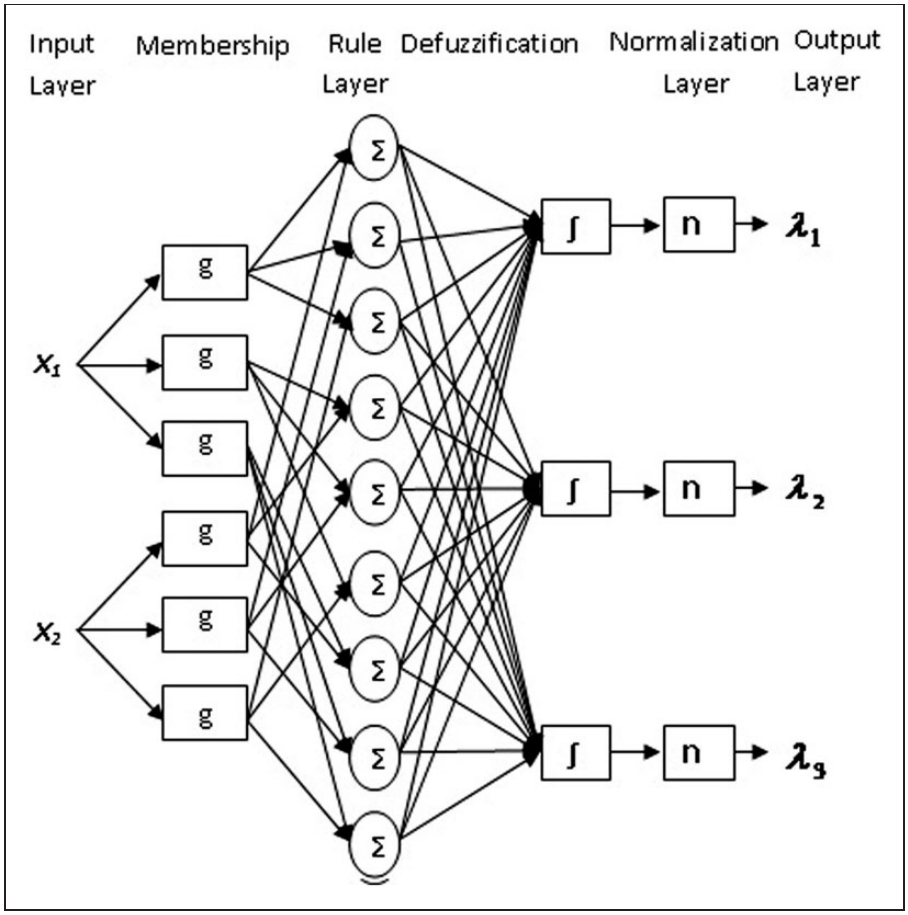

Figure 3. Architecture of neuro fuzzy classifier.

where ω_{ik} denotes the degree of association to k^{th} class that is controlled with the i^{th} rule, M represents the number of rules. The next layer is known as the normalization layer. Its function is to normalized network output because sometimes the summation of the weight can be larger than *1*.

$$\eta_{ck} = \frac{\theta_{ck}}{\sum_{i=1}^{K} \theta_{cl}} = \frac{\theta_{ck}}{\sigma_c}, \sigma_c = \sum_{l=1}^{K} \theta_{cl,} \tag{8}$$

where η_{ck} is the degree of c^{th} sample that belongs to k^{th} class. K is the number of classes. Then, λ_c class label can be calculated by the maximum of η_{ck}

$$\lambda_c = \max_{k=1,2....K} \{\eta_{ck}\} \tag{9}$$

4. Experimental setup

All the bearing vibration data used in this paper are obtained from Case Western Reserve Lab (Loparo, 2013). As shown in Figure 2 the test stand consists of a 2 hp three phase induction motor (left), a dynamometer (right) and a torque transducer (centre). The motor shaft is supported by the test bearing at the drive end. Electro-discharge machining is utilized to introduce single point faults into the test bearing.

To collect the vibration data from the bearing, an accelerometer with bandwidth up to 5000 Hz is mounted on the bearing housing at the drive end of the motor. Healthy bearing data is considered as baseline data. Specific bearing faults considered are inner race fault, outer race fault at 6 o'clock position and rolling element fault having defect size 7 mils, 14 mils and 21 mils (1 mil = 0.001 inch) in diameter. The speeds of the motor are 1730, 1750, 1772 and 1797 rpm and sampling frequency is 48,000 Hz per channel.

5. Results and discussion

In the present paper, a 2048-point width non-overlapping window is selected to divide the vibration data captured from different fault conditions. An MPE over 16 scales corresponding to each window is calculated. A sample training vector is shown in Table 1 contains average values of MPE over different scales.

Table 1. Average value of MPE over different scales for sample input.

Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8	Scale 9	Scale 10	Scale 11	Scale 12	Scale 13	Scale 14	Scale 15	Scale 16	Class
2.7736	3.5975	3.9391	4.2895	4.5480	4.5854	4.5490	4.5121	4.4474	4.4370	4.3923	4.3112	4.3171	4.3213	4.2705	4.1868	Ird
2.7369	3.5678	3.9197	4.2818	4.5209	4.5420	4.4933	4.4455	4.3964	4.4176	4.3880	4.3361	4.3385	4.3149	4.2610	4.2129	Ird
2.6812	3.5224	3.9389	4.3222	4.5454	4.5418	4.4693	4.4395	4.4065	4.4081	4.3884	4.3631	4.3300	4.2891	4.2488	4.2079	Ird
2.4145	3.1756	3.4182	3.8576	4.2356	4.3606	4.2977	4.2270	4.2973	4.3825	4.3415	4.2398	4.2650	4.2806	4.2755	4.2339	Ball
2.3593	3.1073	3.3423	3.7835	4.1261	4.2045	4.0929	4.0376	4.1537	4.3059	4.3340	4.2361	4.1655	4.1892	4.2063	4.2155	Ball
2.3362	3.0697	3.3304	3.8037	4.1319	4.1654	4.0430	4.0017	4.1584	4.3223	4.3459	4.2167	4.1164	4.1193	4.1496	4.1738	Ball
2.2852	2.9996	3.3162	3.8059	4.0925	3.9114	3.7001	3.6105	3.9380	4.2173	4.3112	4.2253	4.1434	4.1502	4.1875	4.0784	Ord
2.2880	3.0039	3.3159	3.8076	4.0879	3.8545	3.5973	3.5029	3.9096	4.2109	4.3102	4.1552	4.0023	3.9741	4.0635	3.9585	Ord
2.2300	2.9309	3.2701	3.7802	4.0676	3.8064	3.5155	3.4372	3.8531	4.1605	4.2497	4.1312	3.9790	3.9305	3.9940	3.8878	Ord
3.4549	4.5383	4.6158	4.5420	4.5801	4.5150	4.4796	4.4131	4.3219	4.2519	4.1993	4.1558	4.1286	4.1409	4.1241	4.1079	Hb
3.4769	4.5117	4.5990	4.5159	4.5506	4.4540	4.4607	4.3867	4.2501	4.1622	4.0998	4.0519	4.0369	4.0475	4.0513	4.0641	Hb
3.4730	4.4156	4.5419	4.4929	4.5192	4.4084	4.4351	4.3649	4.1704	4.0762	4.0121	3.9451	3.9218	3.9466	3.9811	4.0001	Hb

MPE: multi-scale permutation entropy**;** Ird: inner race defect
Hb: healthy bearing; Ord: outer race defect.

Table 2. Confusion matrix of ANFC classification with MPE.

Ird	Ball	Ord	Hb	Classified as
12	0	0	0	Ird
0	12	0	0	Bd
0	0	12	0	Ord
0	0	0	4	Hb

ANFC**:** adaptive neuro fuzzy classifier; MPE: multi-scale permutation entropy
Bd: ball defect; Hb: healthy bearing**;** Ird: inner race defect; Ord: outer race defect

Table 3. Confusion matrix of 10-fold cross validation ANFC classification with MPE.

Ird	Ball	Ord	Hb	Classified as
11	0	0	1	Ird
2	10	0	0	Bd
0	0	12	0	Ord
0	0	0	4	Hb

ANFC: adaptive neuro fuzzy classifier; MPE: multi-scale permutation entropy
Bd: ball defect; Hb: healthy bearing; Ird: inner race defect; Ord: outer race defect

Table 4. Confusion matrix of ANFC classification with statistical features.

Ird	Ball	Ord	Hb	Classified as
9	3	0	0	Ird
1	11	0	0	Bd
1	3	8	0	Ord
0	0	0	4	Hb

ANFC**:** adaptive neuro fuzzy classifier; Bd: ball defect; Hb: healthy bearing
Ird: inner race defect; Ord: outer race defect.

Table 5. Confusion matrix of 10-fold cross validation ANFC classification with statistical features.

Ird	Ball	Ord	Hb	Classified as
9	2	0	1	Ird
0	10	1	1	Bd
4	0	8	0	Ord
0	0	0	4	Hb

ANFC: adaptive neuro fuzzy classifier; Bd: ball defect; Hb: healthy bearing;
Ird: inner race defect
Ord: outer race defect

The MPE parameter delay time d is taken as one and the number of possible permutation m is set at three through empirical observation. Bandit and Pompe (2002) suggested $m = (3 \ldots 7)$ because a larger value of possible permutations increase the computational time and memory requirement. After extracting MPE as feature vectors, this data is divided in to training and testing data sets for automated fault diagnosis. However, computation complexity will increase with the high dimensional feature vectors. Therefore, the arithmetic mean of the MPE values is calculated to reduce the feature vectors.

A total of 40 cases are considered for testing in which 12, 12, 12, 4 cases IRD, ORD, BD and healthy bearing (HB) have been taken. For accurate estimation of these models ability to unseen data 10-fold cross

Table 6. Evaluation of the success of the numeric prediction.

Instances	ANFC with MPE	10-fold cross validation ANFC with MPE	ANFC with statistical features	10-fold cross validation ANFC with statistical features
Correctly classified	40 (100%)	37 (92.50%)	32 (80%)	31 (77.50%)
Incorrectly classified	0 (0%)	3 (7.50%)	8 (20%)	9 (22.50%)

ANFC: adaptive neuro fuzzy classifier; MPE: multi-scale permutation entropy

validation is carried out in the study. Ten-fold cross validation is the standard method of testing classifiers. Thus, the results reported in this study are statistically unbiased. From Table 2, it is inferred that the ANFC has correctly classified all the cases. It has shown 100% accuracy while using 10-fold cross validation, as shown in Table 3, classification accuracy is 92.50% for the ANFC.

This result of MPE and ANFC is compared with the most widely used time domain statistical features kurtosis, skewness, mean, max, min and standard deviation, which have been calculated from time domain vibration data. These features are given as input to ANFC for fault diagnosis. Table 4 demonstrates the confusion matrix of such a method which has correctly predicted 9, 11, 8 and 4 cases out of 12, 12, 12, 4 cases of IRD, BD, ORD and healthy bearing. 10-fold cross validation of ANFC with statistical features displays 9, 10, 8, 4 correctly classified cases as shown in the Table 5. Table 6 indicates classification accuracy of these methods 80% and 77.50% respectively.

6. Conclusions

An automated fault diagnosis approach has been developed using MPE and an ANFC. To account for the dynamic nonlinearity and coupling effect between mechanical parts MPE over 16 scales are calculated. Mean value for each scale value is extracted to reduce the dimensionality of the feature vector. Then data is divided into training and test sets. This experimental result demonstrates that ANFC with MPE can achieve 100% accuracy better than ANFC with statistical features and can predict defects at an early stage. This technique can be used as a reliable on-line fault diagnosis system and will avoid machinery breakdown.

Acknowledgement

The authors would like to thank Prof. K.A. Loparo and Case Western Reserve University for their effort to make bearing data set available.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Aziz W and Arif M (2005) Multiscale permutation entropy of physiological time series. In *Proceedings of 9th IEEE International Multitopic Conference*, Karachi, Pakistan, Dec. 24–25.

Bandit C and Pompe B (2002) Permutation entropy: A natural complexity measure for time series. *Physical Review Letters* 88: 174102–1–174102–4.

Bruzzo AA, Gesierich B, Santi M, Tassinari CA, Birnaumer N and Rubboli G (2008) Permutation entropy to detect vigilance changes and preictal state from scalp EEG in epileptic patients: A preliminary study. *Neurological Science* 29: 3–9.

Cetisli B (2010) Development of adaptive neuro-fuzzy classifier using linguistic hedges: Part 1. *Expert Systems with Applications* 37: 6093–6101.

Costa M, Goldberger AL and Peng CK (2002) Multiscale entropy analysis of complex physiological time series. *Physical Review Letters* 89: 068102–1–068102–4.

Costa M, Goldberger AL and Peng CK (2005) Multiscale entropy analysis of biological signals. *Physical Review E* 71: 021906.

El-Mezyani T, Wilson R, Sattler M, Srivastava SK, Edrington CS and Cartes DA (2012) Quantification of complexity of power electronics based system. *IET Electrical Systems in Transportation* 2(4): 211–222.

Kankar PK, Sharma SC and Harsha SP (2011a) Fault diagnosis of ball bearing using machine learning methods. *Expert Systems with Applications* 38(3): 1876–1886.

Kankar PK, Sharma SC and Harsha SP (2011b) Rolling element bearing fault diagnosis using autocorrelation and continuous wavelet transform. *Journal of Vibration and Control* 17(14): 2081–2094.

Kankar PK, Sharma SC and Harsha SP (2012a) Vibration signature analysis of a high speed rotor supported on ball bearings due to localized defects. *Journal of Vibration and Control* http://dx.doi.org/10.1177/1077546312448506.

Kankar PK, Sharma SC and Harsha SP (2012b) Vibration–based fault diagnosis of a rotor bearing system using artificial neural network and support vector machine. *International Journal of Modelling, Identification and Control* 15(3): 185–198.

Kankar PK, Sharma SC and Harsha SP (2013) Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform. *Neurocomputing* 110: 9–17.

Li X, Ouyang G and Liang Z (2008) Complexity measure of motor current signals for tool flute breakage detection in end milling. *International Journal of Machine Tools and Manufacture* 48: 371–379.

Loparo KA (2013) Bearing vibration data set Case Western Reverse University. Available at: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file (accessed January 2013).

Moller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* 6(4): 525–533.

Richman JS and Moorman JR (2000) Physiological time series analysis using approximate entropy and sample entropy. *American Journal of Physiology - Heart and Circulatory Physiology* 278(6): H2039–H2049.

Samanta B and Al-Balushi KR (2003) Artificial neural networks based fault diagnostics of rolling element bearing using time domain features. *Mechanical System and Signal Processing* 17: 317–328.

Vyas NS and Kumar SD (2001) Artificial neural network design for fault identification in a rotor bearing system. *Mechanism and Machine Theory* 36: 157–175.

Wu SD, Wu PH, Wu CW, Ding JJ and Wang CC (2012) Bearing fault diagnosis based on multiscale entropy and support vector machine. *Entropy* 14: 1343–1356.

Yang JY, Zhang YY and Zhu YS (2007a) Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. *Mechanical System and Signal Processing* 21(5): 2012–2024.

Yang Y, Dejie Y and Chang J (2007b) Fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM. *Measurement* 40: 943–950.

Zhang L, Xiong G, Liu H, Zou H and Guo W (2010) Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. *Expert Systems with Applications* 37: 6077–6085.