

Signal Complexity and Gaussian Process Models Approach for Bearing Remaining Useful Life Estimation

Pavle Boškoski, Matej Gašperin and Dejan Petelin

Abstract Standard bearing fault detection features are shown to be ineffective for estimating bearings remaining useful life (RUL). In this paper we propose a new approach estimating bearing RUL based on features describing the statistical complexity of the envelope of the generated vibrations and a set of Gaussian process (GP) models. The proposed approach is shown to be sensitive to incipient condition deterioration which allows timely and sufficiently accurate estimates of the RUL. The proposed approach was evaluated on the data set comprising of 17 bearing runs with natural fault evolution.

Keywords Bearing prognostics · Statistical signal complexity · Gaussian process models

1 Introduction

Several surveys show that bearing faults represent the most common cause for failure of mechanical drives [1, 2]. As a result, a plethora of methods have been developed addressing the problems of bearing fault detection and prognostics. Most of the available methods rely on a well-established feature set, based on characteristic bearing fault frequencies. However, these features are shown to be ineffective for estimating bearing's remaining useful life (RUL) [3]. Addressing the problem of bearing fault prognostics, in this paper we propose a combination

P. Boškoski (✉) · M. Gašperin · D. Petelin
Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
e-mail: pavle.boskoski@ijs.si

M. Gašperin
Faculty of Electrical Engineering/RICE, University of West Bohemia,
Plzen, Czech Republic

of new features based on the statistical complexity of the envelope of bearing's vibrations and Gaussian process (GP) models for estimating bearing's RUL.

Majority of the available approaches describe the relationship between the defect growth and the trend of some statistical characteristic of the generated vibrations like energy, peak-to-peak values, RMS, kurtosis, crest factor etc. [3–5]. Usually these values are calculated from the generated vibrations filtered on specific frequency bands. The ratios of these features from various frequency bands are employed for estimating the bearings RUL. The effectiveness of these ratios can be explained through the relation between the time evolution of the excited bearing's natural frequency and the deterioration of the bearing's RUL [6–8]. Based on this assumption, bearing's RUL was estimated using approaches such as: tracking the evolution of the vibration energy using hidden Markov models [9] or tracking the increase of the dimensional exponents of the generated vibrations [10].

Following these two approaches, we propose a set of features that quantify the statistical complexity of the generated vibrations. The concept of statistical complexity is readily applied for analysis of EEG signals [11, 12]. In the context of bearing prognostics, any change in the bearing surface can be treated as a source of additional signal components with complex dynamics, hence increasing the statistical complexity of the generated vibrations. Our results show that the evolution of the statistical complexity of the generated vibrations can be directly related to the bearings RUL. Additionally, the process for calculating the statistical complexity requires no prior information about the operating conditions and no previous knowledge about the physical characteristics of the monitored drive [13, 14].

Using the Rényi entropy based statistical complexity, the bearing's RUL was estimated using GP models, which are probabilistic, non-parametric models. GP models search for relationships among the measured data rather than approximating the modelled system by fitting the parameters of the selected basis functions, which is common for other black-box identification approaches. The predictions of GP models are represented by a normal distribution. Because of their properties GP models are especially suitable for modelling when data is unreliable, noisy or missing. Their uses and properties for modelling are reviewed in [15]. In this paper the GP models are used for two purposes: filtering noisy features and estimating the RUL.

The proposed approach consists of four main steps. In first step three features, in-depth described in Sects. 2 and 3, are extracted from the acquired vibrations. The process of the numerical estimation of these features is presented in Sect. 4. In second step these features are filtered using GP models, described in Sect. 5. Afterwards, GP models are used for the estimation of RUL values based on filtered features. The final RUL estimation is obtained by fusion of all estimated RUL values. The evaluation of the approached is presented in Sect. 6.

2 Signal Complexity

The definitions of the statistical complexity of a signal varies depending of the context. In context of signals one can define two extremes: periodic and purely random signals. Both cases belong to the class of low complexity signals: the former due to its repetitive pattern and the latter due to its compact statistical description [16, 17]. Consequently, “complex” signals should be located somewhere in between, a typical candidates are signals generated by a system with chaotic behaviour.

For a random signal, generated by a random source with probability distribution \mathcal{P} , the statistical complexity $\mathcal{C}(\mathcal{P})$ can be assessed through the information carried by the generated signal [11]. The statistical complexity provides a link between the entropy of the source $H(\mathcal{P})$ and the “distance” between the probability distribution \mathcal{P} and the uniform distribution \mathcal{P}_e as [11]:

$$\mathcal{C}(\mathcal{P}) = Q_0 D_\alpha^w(\mathcal{P}, \mathcal{P}_e) H_\alpha(\mathcal{P}), \quad (1)$$

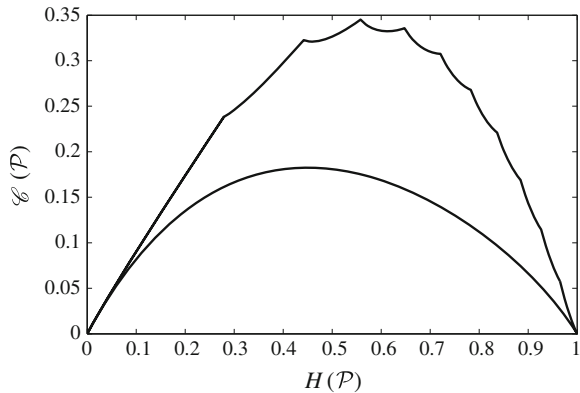
where \mathcal{P}_e is the uniform distribution and Q_0 is a normalisation constant so that $Q_0 D_\alpha^w(\mathcal{P}, \mathcal{P}_e) \in [0, 1]$. The values $H_\alpha(\mathcal{P})$ and $D_\alpha^w(\mathcal{P}, \mathcal{P}_e)$ are the Rényi entropy and Jensen-Rényi divergence respectively, and are defined as [18, 19]:

$$H_\alpha(\mathcal{P}) = \frac{1}{1-\alpha} \ln \sum_{p \in \mathcal{P}} p^\alpha(x), \quad a \geq 0 \quad a \neq 1 \quad (2)$$

$$D_\alpha^w(\mathcal{P}, \mathcal{Q}) = H_\alpha(w\mathcal{P} + (1-w)\mathcal{Q}) - \{wH_\alpha(\mathcal{P}) + (1-w)H_\alpha(\mathcal{Q})\}. \quad (3)$$

The statistical complexity $\mathcal{C}(\mathcal{P})$ is usually plotted versus the entropy $H_\alpha(\mathcal{P})$ [11]. Such a plot always covers a specific pre-defined area, as shown in Fig. 1.

Fig. 1 Signal’s statistical complexity area. The pre-defined shape of the plot outlines the possible time evolution of the signal’s complexity. By trending the evolution of the statistical complexity within the pre-defined area one can perform the prognostics task



3 Complexity of Bearing Vibrations

Healthy bearings produce negligible vibrations. However, in the case of surface damage, vibrations are generated by rolling elements passing across the damaged site on the surface. Each time this happens, impact between the passing ball and the damaged site triggers a system impulse response $s(t)$. The time of occurrence of these impulse responses as well as their amplitudes should be considered as purely random processes. Consequently, the vibrations generated by damaged bearings can be modeled as [20]:

$$y(t) = \sum_{i=-\infty}^{+\infty} A_i s(t - t_i) + n(t), \quad (4)$$

where A_i is the impulse of force that excites the entire structure and t_i is the time of its occurrence. The final component $n(t)$ defines an additive random component that contains all non-modeled vibrations as well as environmental disturbances.

Generally the impulse response $s(t)$ is influenced by the transmission path from the point of impact to the measurement point [21]. As the position of the damaged spot on the bearing surface rotates the transmission path changes in time. However, regardless of its true form, $s(t)$ is characterized by its high-frequency signature. Since this is the only characteristic relevant for our analysis, we will adopt the model (4) as sufficiently accurate one.

Evolution of the statistical complexity of the generated bearing vibrations The main diagnostic information regarding bearing faults are the time moments t_i in (4). Therefore, the usual approach is to analyze the envelope of the generated vibrations. In our case, we look for any changes in the statistical characteristics of the envelope [13].

In the case of healthy bearings, due to the lack of impacts, the envelope of the generated vibrations will be without any visible structure. Therefore, the envelope will have low complexity but high entropy, i.e. such a signal would be positioned in the lower right corner in Fig. 1. The occurrence of a surface fault will introduce some “structure” in the envelope of the generated vibrations. Consequently, its statistical complexity will increase while in the same time the entropy will decrease. In the terminal phase, the envelope will contain impulse responses with sufficiently high amplitude. As a result the signal complexity will sharply drop accompanied with a significant decrease in its entropy, hence the final position will be in the lower left corner in Fig. 1. By trending this evolution, one will be able to estimate the bearing’s RUL.

4 Wavelet Based Estimation of the Statistical Complexity of the Signal Envelope

The first step in the calculation of the statistical complexity is the estimation of the PDF \mathcal{P} of the envelope of the generated vibrations. Due to the link between the signal's envelope and its instantaneous power [22], the PDF is estimated through the energy distribution of the wavelet packet transform (WPT) coefficients [23]. WPT is described by a binary tree structure, as shown in Fig. 2. Each node in WPT tree with depth d_M is marked as (d, n) , where depth $d = \{1, 2, \dots, d_M\}$ and $n = \{1, 2, \dots, 2^d\}$ stands for the number of the node at depth d . The wavelet coefficients, in the set of terminal nodes T , contain all information regarding the analysed signal.

Each of the n nodes at level d contains N_d wavelet coefficients $W_{d,n,t}$ $t = 0, \dots, N_d - 1$, $N_d = 2^{-d}N_s$, N_s is the sample length of the signal [24]. Using these coefficients, the portion of the signal's energy $E_{d,n}$ for each node (d, n) reads [25]:

$$E_{d,n} = \sum_{t=0}^{N_d-1} \|W_{d,n,t}\|^2 \quad (5)$$

and total signal's energy becomes:

$$E_{tot} = \sum_{\substack{t=0 \\ d,n \in T}}^{N_d-1} \|W_{d,n,t}\|^2 = \sum_{d,n \in T} E_{d,n}. \quad (6)$$

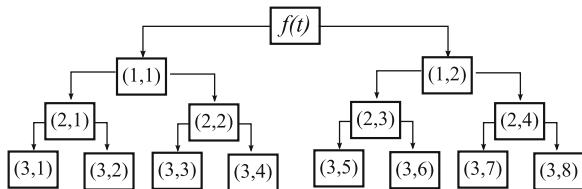
The set $\mathcal{P}^{d,n}$ expresses the contribution of each wavelet coefficient to the energy of the signal within the terminal node (d, n) :

$$\mathcal{P}^{d,n} = \left\{ p_t^{d,n} = \frac{\|W_{d,n,t}\|^2}{E_{d,n}}, t = 0, \dots, N_d - 1 \right\}. \quad (7)$$

A similar set can be defined for the contribution of the energy of each terminal node $(d, n) \in T$ in the total energy of the signal E_{tot} :

$$\mathcal{P}^T = \left\{ p_{d,n} = \frac{E_{d,n}}{E_{tot}}, d, n \in T \right\}. \quad (8)$$

Fig. 2 Example of a full WPT tree with depth $d_M = 3$



The elements contained in both sets $\mathcal{P}^{d,n}$ and \mathcal{P}^T can be treated as realisation of a random process. Based on these realisations one can estimate the corresponding probability distributions and calculate their entropies and statistical complexity according to relations (1–3).

Condition monitoring based on the statistical characteristics of the sets $\mathcal{P}^{d,n}$ and \mathcal{P}^T . The values of the selected features (1–3) are calculated on short non-overlapping windows, where the initial values are regarded as reference ones. As time evolves, the presence of faults will alter the envelope PDF for particular node, hence changing the feature values. As a result, RUL can be estimated by tracking the evolution of their values. It is important to stress that the window length is usually very short so that one can assume that within its duration the operating condition is almost constant. If speed varies mildly the distribution pattern will not change much as shifted harmonics will remain within the frequency band associated to a particular node. If changes in the operating speed are severe, it might happen that the frequency content from one node moves to the adjacent node, thus fooling entirely the diagnostic reasoning. On the other hand, mild variations in load normally have no significant impact on the frequency distribution pattern.

5 Gaussian Process Models

Features (1–3) based on $\mathcal{P}^{d,n}$ and \mathcal{P}^T are quite noisy. Therefore we filter them using GP models.¹ Afterwards based on these filtered features, GP models are used for estimating RUL.

A Gaussian process is a collection of random variables which have a joint multivariate Gaussian distribution. Assuming a relationship of the form $y = f(\mathbf{x})$ between input \mathbf{x} and output y , we have $y_1, \dots, y_N \sim N(0, \mathbf{K})$, where $\mathbf{K}_{pq} = \text{Cov}(y_p, y_q) = C(\mathbf{x}_p, \mathbf{x}_q)$ gives the covariance between output points corresponding to input points \mathbf{x}_p and \mathbf{x}_q . Thus, the mean $m(\mathbf{x})$ and the covariance function $C(\mathbf{x}_p, \mathbf{x}_q)$ fully specify the Gaussian process.

The value of covariance function $C(\mathbf{x}_p, \mathbf{x}_q)$ expresses the correlation between the individual outputs $f(\mathbf{x}_p)$ and $f(\mathbf{x}_q)$ with respect to inputs \mathbf{x}_p and \mathbf{x}_q . It should be noted that the covariance function $C(\times, \times)$ can be any function that generates a positive semi-definite covariance matrix. Most commonly used covariance function is a composition of the square exponential covariance function and the constant covariance function presuming white noise:

¹ Filtering is simply performed by modelling the data as time-series and then estimating the mean value for whole series. Such a filtering does not introduce any additional lag in the time series, which is not the case with other commonly used filtering methods.

$$C(\mathbf{x}_p, \mathbf{x}_q) = v_1 \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_{dp} - x_{dq})^2 \right] + d_{pq} v_0, \quad (9)$$

where w_d , v_1 and v_0 are the hyperparameters of the covariance function, D is the input dimension, and $\delta_{pq} = 1$ if $p = q$ and 0 otherwise. Hyperparameters can be written as a vector $\Theta = [w_1, \dots, w_D, v_1, v_0]^T$. Hyperparameters w_d indicate the importance of individual inputs. If w_d is zero or near zero, it means that the inputs in dimension d contain little information and could possibly be neglected.

To accurately reflect the correlations presented in the data, the hyperparameter values of the covariance function need to be optimized. Due to the probabilistic nature of the GP models, instead of minimizing the model error, the probability of the model is maximized.

Consider a set of N D -dimensional input vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$ and a vector of output data $\mathbf{y} = [y_1, y_2, \dots, y_N]$. Based on the data (\mathbf{X}, \mathbf{y}) , and given a new input vector \mathbf{x}^* , we wish to find the predictive distribution of the corresponding output y^* . Based on training set \mathbf{X} , a covariance matrix \mathbf{K} of size $N \times N$ is determined. The overall problem of learning unknown hyperparameters q from data corresponds to the predictive distribution $p(y^* | \mathbf{y}, \mathbf{X}, \mathbf{x}^*)$ of the new target y , given the training data (\mathbf{y}, \mathbf{X}) and a new input \mathbf{x}^* . In order to calculate this posterior distribution, a prior distribution over the hyperparameters $p(\Theta | \mathbf{y}, \mathbf{X})$ can first be defined, followed by the integration of the model over the hyperparameters

$$p(y^* | \mathbf{y}, \mathbf{X}, \mathbf{x}^*) = \int p(y^* | \Theta, \mathbf{y}, \mathbf{X}, \mathbf{x}^*) p(\Theta | \mathbf{y}, \mathbf{X}) d\Theta. \quad (10)$$

The computation of such integrals can be difficult due to the intractable nature of the non-linear functions. Therefore the general practice for estimating hyperparameter values is minimising the following negative log-likelihood function:

$$\mathcal{L}(\Theta) = -\frac{1}{2} \log(|\mathbf{K}|) - \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{N}{2} \log(2\pi). \quad (11)$$

GP models can be easily utilised for regression calculation. Based on training set \mathbf{X} , a covariance matrix \mathbf{K} of size $N \times N$ is calculated. The aim is to find the distribution of the corresponding output y^* for some new input vector $\mathbf{x}^* = [x_1(N+1), x_2(N+1), \dots, x_D(N+1)]$. The predictive distribution of the output for a new test input has normal probability distribution with mean and variance

$$\mu(y^*) = \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{y}, \quad (12)$$

$$\sigma^2(y^*) = \kappa(\mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*), \quad (13)$$

where $\mathbf{k}(\mathbf{x}^*) = [C(\mathbf{x}_1, \mathbf{x}^*), \dots, C(\mathbf{x}_N, \mathbf{x}^*)]^T$ is the $N \times 1$ vector of covariances between the test and training cases, and $k(x^*) = C(\mathbf{x}^*, \mathbf{x}^*)$ is the covariance between the test input itself.

As can be seen from (13), the GP model, in addition to mean value, also provides information about the confidence in prediction by the variance. Usually, the confidence of the prediction is depicted with 2σ interval which is about 95% confidence interval. This confidence region can be seen in the example in Fig. 3 as a grey band. It highlights areas of the input space where the prediction quality is poor, due to the lack of data or noisy data, by indicating a wider confidence band around the predicted mean.

6 Results

The proposed approach was evaluated on the data set for the IEEE PHM 2012 Data Challenge [26]. Provided data consist of three batches, each corresponding to different speed and load conditions. The generated vibrations were sampled with 22 for duration of 100, repeated every 5 min. The experiments were stopped when the RMS value of the generated vibrations surpassed 20. The available vibration signals were analysed using WP tree with depth $d_M = 4$, which results into 16 terminal nodes. All features are filtered using GP models. The filtered statistical complexity $\mathcal{C}(\mathcal{P})$ for one particular node is shown in Fig. 4.

The time evolution of the statistical complexity $\mathcal{C}(\mathcal{P})$ has similar shape as the theoretical one. In time the statistical complexity evolves from the area of low complexity and high entropy towards the area of low complexity and low entropy. During this evolution the value of $\mathcal{C}(\mathcal{P})$ passes through the apex of the predefined area shown in Fig. 1.

Prediction results Using experimental runs from the first and the third batch as training set, 16 GP models were defined, one for each of the 16 WP nodes. Each GP model describes the most probable evolution of the three features $D_x^w(\mathcal{P}^{d,n}, \mathcal{P}_e)$, $H_x(\mathcal{P}^{d,n})$ and $H_x(\mathcal{P}^T)$ in respect to the RUL normed in the interval $[0, 1]$. At each time moment, we calculate the likelihood of the bearings RUL based on the estimated GP model estimates.

Fig. 3 Modelling with GP models: in addition to mean value (prediction), we obtain a 95% confidence region for the underlying function f (shown in grey)

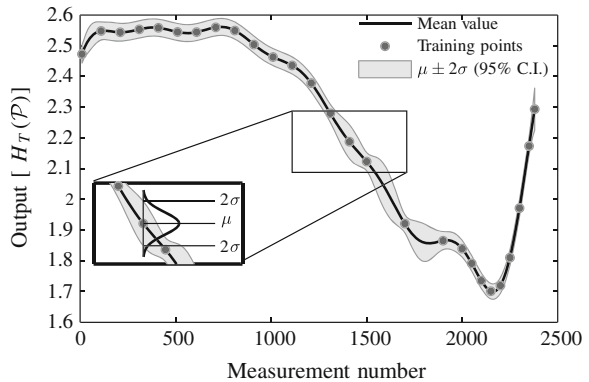


Fig. 4 Evolution of envelope complexity where shading describes the time. In the beginning the generated vibrations have low complexity $\mathcal{C}(\mathcal{P})$ and high entropy $H^{3,3}(\mathcal{P})$. In time, the statistical complexity slightly increases, where as the value of $H^{3,3}(\mathcal{P})$ decreases. At the end, both $\mathcal{C}(\mathcal{P})$ and $H^{3,3}(\mathcal{P})$ significantly decrease

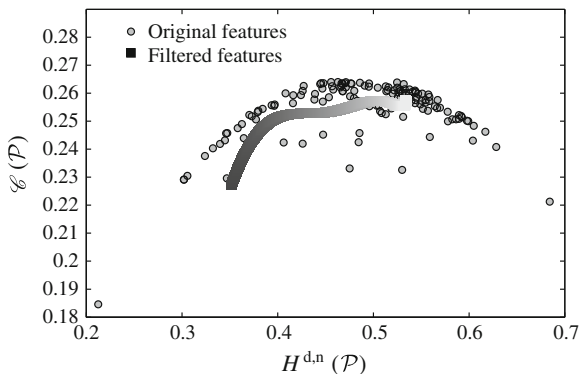
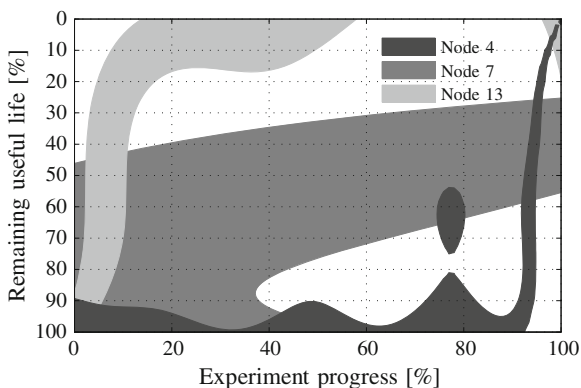


Fig. 5 RUL estimates for different WP nodes. Initial signs of deterioration are firstly detected by the high frequency nodes, like the 13 node. Low frequency nodes become sensitive towards the very end of the experiment. A proper RUL estimate should be based on the information carried by these ratios



These likelihood estimates for one bearing run are shown in Fig. 5, which shows that RUL sensitivity differs among the WP nodes. For instance, high frequency nodes, such as node 13, give first indications that the bearing has reached the end of its useful life. Conversely, RUL estimates of low frequency nodes, for instance node 4, are over-optimistic during the majority of the experiment. The actual bearing condition becomes visible only towards the end of the experiment. Such an observation leads to a conclusion that the first signs of bearing condition deterioration become visible in the high frequency parts of the signal. As the condition deteriorates sensitivity shifts towards features extracted from the lower frequency bands.

7 Conclusions

The combination of statistical complexity features coupled with Gaussian process models provides a suitable solution for estimating bearing RUL. The proposed approach is generally applicable, as it requires no prior knowledge neither about

the bearing physical characteristics nor about the bearing's operating condition. Therefore, the generated GP models were calculated using features extracted from bearings operating under different operating conditions and were evaluated on vibrations generated by bearings that operated under previously "unseen" conditions. These GP models describe the evolution of the selected features in respect to the bearing RUL. The results show that decrease in the bearing condition shifts the sensitivity of the features, making the features extracted from high frequency bands sensitive to initial damage and features from low frequency bands sensitive to severe damage. Consequently, this relation between condition deterioration and frequency dependent feature sensitivity can be employed for estimating the bearing RUL.

Acknowledgments We like to acknowledge the support of the Slovenian Research Agency through the Research Programme P2-0001, the Research Project L2-4160 and project EXLIZ CZ.1.07/2.3.00/30.0013, which is co-financed by the European Social Fund and the state budget of the Czech Republic.

References

1. Albrecht PF, Appiarius JC, Shrama DK (1986) Assessment of the reliability of motors in utility applications. *IEEE Trans Energy Convers* **EC-1**:39–46
2. Crabtree CJ (2010) Survey of commercially available condition monitoring systems for wind turbines. Tech. rep., Durham University, School of Engineering and Computing Science
3. Camci F, Medjaher K, Zerhouni N, Nectoux P (2012) Feature evaluation for effective bearing prognostics. *Qual Reliab Eng Int*. doi:[10.1002/qre.1396](https://doi.org/10.1002/qre.1396)
4. Li R, Sopon P, He D (2012) Fault features extraction for bearing prognostics. *J Intell Manuf* **23**:313–321. doi:[10.1007/s10845-009-0353-z](https://doi.org/10.1007/s10845-009-0353-z)
5. Lybeck N, Marble S, Morton B (2007) Validating prognostic algorithms: a case study using comprehensive bearing fault data. In: aerospace conference, 2007 IEEE, pp 1–9
6. Qiu J, Seth BB, Liang SY, Zhang C (2002) Damage mechanics approach for bearing lifetime prognostics. *Mech Syst Signal Process* **16**(5):817–829
7. Randall RB (2011) The challenge of prognostics of rolling element bearings. In: wind turbine condition monitoring workshop
8. Wang W (2008) Autoregressive model-based diagnostics for gears and bearings. *Insight* **50**(5):1–5
9. Ocak H, Loparo KA, Discenzo FM (2007) Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: a method for bearing prognostics. *J Sound Vib* **302**(4–5):951–961
10. Janjarasjitt S, Ocak H, Loparo K (2008) Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal. *J Sound Vib* **317**(1–2):112–126
11. Kowalski AM, Martin MT, Plastino A, Rosso OA, Casas M (2011) Distances in probability space and the statistical complexity setup. *Entropy* **13**(6):1055–1075
12. Martin M, Plastino A, Rosso O (2006) Generalized statistical complexity measures: geometrical and analytical properties. *Phys A* **369**(2):439–462
13. Boškosi P, Juričić Đ (2012) Fault detection of mechanical drives under variable operating conditions based on wavelet packet rényi entropy signatures. *Mech Syst Signal Process* **31**:369–381

14. Bošković P, Juričić Đ (2012) Rényi entropy based statistical complexity analysis for gear fault prognostics under variable load. In: condition monitoring of machinery in non-stationary operations. Springer Berlin Heidelberg, pp 25–32
15. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press
16. Adami C (2002) What is complexity? *BioEssays* 24(12):1085–1094
17. Crutchfield JP, Young K (1989) Inferring statistical complexity. *Phys Rev Lett* 63(2):105–108
18. Basseville M (2010) Divergence measures for statistical data processing. Tech. rep., IRISA
19. Rényi A (1960) On measures of information and entropy. In: 4th Berkeley symposium on mathematics, statistics and probability
20. Randall RB, Antoni J, Chobsaard S (2001) The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mech Syst Signal Process* 15:945–962
21. Antoni J, Randall RB (2003) A stochastic model for simulation and diagnostics of rolling element bearings with localized faults. *J Vib Acoust* 125(3):282–289
22. Antoni J (2009) Cyclostationarity by examples. *Mech Syst Signal Process* 23:987–1036
23. Mallat S (2008) A wavelet tour of signal processing, 3rd edn. Academic Press, Burlington
24. Percival DB, Walden AT (2000) Wavelet methods for time series analysis. Cambridge University Press, Cambridge
25. Blanco S, Figliola A, Quiroga RQ, Rosso OA, Serrano E (1998) Time-frequency analysis of electroencephalogram series iii. Wavelet packets and information cost function. *Phys Rev E* 57(1):932–940
26. Nectoux P, Gouriveau R, Medjaher K, Ramasso E, Morello B, Zerhouni N, Varnier C (2012) Pronostia: an experimental platform for bearings accelerated life test. In: IEEE international conference on prognostics and health management. Denver, CO, USA