This is a post-print version of the article presented at the Seventeenth International Conference on Condition Monitoring and Asset Management (CM 2021). The event website is available at: https://www.bindt.org/events/CM-2021/ and the abstract is available at: https://www.bindt.org/events/CM-2021/abstract-9a7/.

On the applicability of time series features as health indicators for technical systems operating under varying conditions

Osarenren Kennedy Aimiyekagbon, Amelie Bender, and Walter Sextro

Chair of Dynamics and Mechatronics, Paderborn University, Paderborn, Germany
osarenren.aimiyekagbon@uni-paderborn.de

Abstract

Several methods, including order analysis, wavelet analysis and empirical mode decomposition have been proposed and successfully employed for the health state estimation of technical systems operating under varying conditions. However, where information such as the speed of rotating machinery, component specifications or other domain-specific information is unavailable, such methods are often infeasible. Thus, this paper investigates the application of classical time-domain features, features from the medical field and novel features from the highly comparative time-series analysis (HCTSA) package, for the health state estimation of rotating machinery operating under varying conditions. Furthermore, several feature selection methods are investigated to identify features as viable health indicators for the diagnostics and prognostics of technical systems. As a case study, the presented methods are evaluated on real-world and experimentally acquired vibration data of bearings operating under varying speed. The results show that the selected features can successfully be employed as health indicators for technical systems operating under varying conditions.

1. Introduction

In his book, Umbaugh[1] presented some characteristics a suitable feature should possess. In that light, independent of the diagnostics or prognostics task at hand, a suitable feature as a viable health indicator should be *distinctive*, that is, capable of distinguishing between potential fault or failure modes. It should be *reliable*, which implies its value does not differ for similar fault or failure modes. It should be *independent*, meaning it is not derivable from other features. It should be *robust* to noise or variations in operating condition and, it should reflect possible changes in health state or degradation over the life cycle of a technical system. Typically, health indicators are derived from time-, frequency and, time-frequency domain analysis of condition monitoring data acquired over time[2]. In an attempt to find suitable health indicators for reliably estimating the current health state and the remaining useful life (RUL) of technical systems and rotating machinery, in particular, several techniques have been proposed in scholarly literature. A few notable techniques are presented in the following paragraphs.

Classical time-domain features such as crest factor and kurtosis are easily implemented and employed in trend analysis. They have found application as health indicators in rotating machinery and components thereof, such as gears and bearings[3, 4]. The crest factor of a vibration signal, similarly known as the impact index, was developed by the General Electric Company as a health indicator for bearings[5]. The crest factor has found numerous applications due to its simple implementation, simple interpretation and robustness to variations in loading condition[6]. However, in a report by Houser and Drosjack[6] on the investigation of health indicators for helicopter gears and bearings, they found out that the crest factor is not robust to noise and to variations in speed.

The statistical parameter, kurtosis, which denotes the fourth moment of a vibration signal, typically has a value of approximately three for healthy bearings and a value greater than three otherwise[2]. The advantages of kurtosis as a health indicator are that it is independent of the bearing or application type and, its value can easily be tracked and trended[2]. However, Howard[2] additionally pointed out a major limitation of the kurtosis, which is that its value is not consistent, even in characteristic frequency bands, when faults manifest into failure.

The spectrum and order analysis are widely accepted and employed techniques in academia and industry. Faults are not only identified, but faulty components are correspondingly localized irrespective of variations in speed when the kinematics of the components under investigation are known[6]. However, for situations where the shaft speed varies, the shaft speed must be acquired (synchronous) with condition monitoring data. Furthermore, characteristic frequencies of interest can be smeared by neighbouring frequencies and masked by frequencies of other system components, thus introducing difficulties in fault identification and localization. Further limitations are that such techniques are not robust to noise and varying load[6].

Prominent techniques of the time-frequency domain analysis are the wavelet and the Hilbert–Huang transform. They both allow for analysing stationary and non-stationary signals, such as signals involving varying speed and load[2, 7]. According to Peng et al.[7], the wavelet transform is probably the best and widely employed time-frequency domain technique for detecting bearing faults. There exist several wavelet families and variants, each producing different analysis outcomes, depending on the application[2, 7]. A major limitation of the wavelet transform is that there is no consensus on the optimal mother wavelet. A further drawback of the wavelet transform is the poor frequency resolution at high frequencies, which can impair interpretation and fault identification[2, 7].

The key step of the Hilbert–Huang transform is the empirical mode decomposition. It is not as computationally expensive as the wavelet transform, thus making it suitable for signals with high sampling rates[7]. Nonetheless, an underlying shortcoming of the empirical mode decomposition is that it introduces aliasing. Therefore, it becomes difficult if, at all possible, to identify faults or interpret results[7].

Although several attempts have been made to alleviate the shortcomings of the previously discussed techniques, these attempts have either increased complexity or decreased interpretability[7]. Given that the time-domain analysis is simple and this domain contains vast information[2], this paper focuses on the time-domain analysis of vibration data to find suitable health indicators that fulfil the presented characteristics for reliably estimating the current health state and the RUL of technical systems and rotating machinery in particular. To this end, features from the medical field and novel features from the highly comparative time-series analysis (HCTSA) package[8, 9] are investigated in addition to classical time-domain features derived from vibration data.

The remainder of this paper is organized as follows. A brief overview of the employed feature extraction and feature selection techniques is presented in the next section. Subsequently, the presented techniques are evaluated based on two case studies. On a concluding note, the results are discussed, and further research possibilities are proposed.

2. Overview of the employed methods

In this section, a brief overview of the employed feature extraction and feature selection techniques is given while referencing related literature and highlighting key contributions.

2.1 Feature extraction techniques

Given that extensive research and application of classical time-domain features as health indicators for rotating machinery and components thereof exist in scholarly literature, 13 classical time-domain features, as in described by Kimotho and Sextro[(3)], are extracted here for comparison. These include, but are not limited to, crest factor, kurtosis and root mean square.

The electromyography (EMG) feature extraction technique originates from the medical field for the diagnostics of muscles[(10, 11)]. The motivation for considering the EMG feature extraction technique here are the works of Sánchez et al.[(12)], and Nayana and Geethanjali[(13)], who successfully employed EMG features in addition to classical time-domain features for gearbox and bearing fault diagnosis, respectively. Although both works considered various fault types and operating conditions, the operating conditions were kept constant per experiment. Moreover, they only considered vibration data acquired under a controlled experimental setting. In this paper, 17 EMG features were extracted from the MATLAB® EMG feature extraction toolbox as implemented by Too et al.[(10, 11)] Of these, Nayana and Geethanjali[(13)] identified the features wavelength, Willison amplitude, zero crossing and slope sign change as most relevant features for diagnosing bearing faults. Further descriptions of the EMG features can be found in the paper by Phinyomark[(14)].

The HCTSA package is a collection of features from interdisciplinary disciplines such as physics, economics and medicine by Fulcher et al.[(8, 9)]. Depending on the input parameters, the approximately 1000 basic operations yields over 7700 characteristic features. These features can be classified into the following groups as proposed by the authors[(15)]. Distribution; correlation; entropy and information theory; time-series model fitting and forecasting; stationarity and step detection; nonlinear time-series analysis and fractal scaling; Fourier and wavelet transforms, periodicity measures; symbolic transformations; statistics from biomedical signal processing; basic statistics, trend; and others. The HCTSA package has found numerous application in the medical field, such as for heartbeat classification[(16)] and stress level classification based on calcium dynamics[(17)], and in engineering for diagnosing an industrial robotic arm based on movement measurements[(18)]. During the preparation of this paper, there was no published scholarly literature that applied the HCTSA package for diagnostics or prognostics task of rotating machinery and components thereof, while considering vibration data. A total of 3905 features were extracted from the MATLAB® HCTSA package[(8, 9)].

Given the number of generated features, a manual selection of viable health indicators from the set of the above-described features was no longer feasible. Thus, feature selection techniques as described in the following section were utilized.

2.2 Feature selection techniques

Feature selection techniques can generally be categorized in filter, wrapper and embedded methods[19]. The filter methods select a subset of features based on some relevance measure, irrespective of a machine learning algorithm. The wrapper methods employ a machine learning algorithm to rank a subset of features, depending on their predictive performance. The embedded methods are implemented within a given machine learning algorithm, as the name implies, and hence simultaneously rank a subset of features during algorithm training[19, 20]. This paper employs filter methods, as described in the following sections since the goal is to find optimal health indicators independent of a specific machine learning algorithm.

2.2.1 Filter methods for classification

There are several filter methods for classification tasks. Among these, reliefF and chi-square test were employed as filter methods in this paper because they have successfully found application in the diagnostics of a gearbox[12].

- **ReliefF** is a distance-based filter method, such that features are more relevant that have a maximal distance between classes (interclass) and a minimal distance within a class (intraclass). The class membership is typically determined by the k-nearest neighbours algorithm[12, 21, 22].

- **Chi-square test** is a statistical test of the independence of two statistical variables. As a filter method, features that are dependent of the target class are considered more relevant[12].

2.2.2 Filter methods for regression

Monotonicity, prognosability and trendability are three prominent measures by Coble and Hines[23] for determining optimal health indicators for the prognostics of technical systems. Prognosability and trendability both consider a fleet of systems, where run-to-failure data is available. However, in this paper run-to-failure data is not available for a fleet of systems. Thus, monotonicity, as well as an additional method (RReliefF), are considered in this paper for comparison.

- **Monotonicity** generally implies a decreasing or increasing trend. As opposed to the definition proposed by Coble and Hines[23], the general monotonic-relationship can be evaluated with nonparametric statistical tests such as the Spearman's rank correlation coefficient, similarly known as Spearman's rho[24, 25]. This metric was successfully employed as a health indicator to estimate the RUL of ball bearings[25]. Hence, it is considered in this paper.

- **Regressional ReliefF (RReliefF)** is a further development of the ReliefF for regression problems. The relevance measure is based on the equivalent difference of conditional probabilities[22].

3. Case Studies

Two case studies based on publicly available data sets were considered to evaluate the feature extraction and feature selection techniques presented. A case study is considered for detecting seeded faults on bearings under time-varying shaft speed, and the other case study is considered for fault detection on a real-world wind turbine subjected to time-varying wind speed.

3.1 Case Study: Fault detection of seeded bearing faults

The bearing data set considered here is publicly available, and a detailed description of the experimental set-up can be found in the accompanying paper by Huang and Baddour[26, 27]. A total of 60 files, each comprising a 10 s vibration measurement sampled at 200 kHz and synchronously measured shaft speed, are made available. The bearing experiments consist of five health states; namely, **H**: healthy, **I**: inner race fault, **O**: outer race fault, **B**: ball fault, and **C**: combined fault of the preceding three fault cases. Each of these health states was examined considering four different speed variations; namely, **A**: increasing speed, **B**: decreasing speed, **C**: increasing, then decreasing speed and **D**: decreasing, then increasing speed. For each of the 20 different combinations, measurements were acquired three times to mitigate measurement uncertainty.

Preliminary Analysis

A three-dimensional plot of the vibration data acquired from the first measurement involving increasing speed and all health states is presented in Figure 1(a) to get an overview of the data set. As evident from Figure 1(a), the healthy instance (H-A-1) can hardly be distinguished from the outer ring (O-A-1) and ball fault (B-A-1) instances. However, the inner ring (I-A-1) and combined fault (C-A-1) instances can easily be distinguished from the healthy instance.

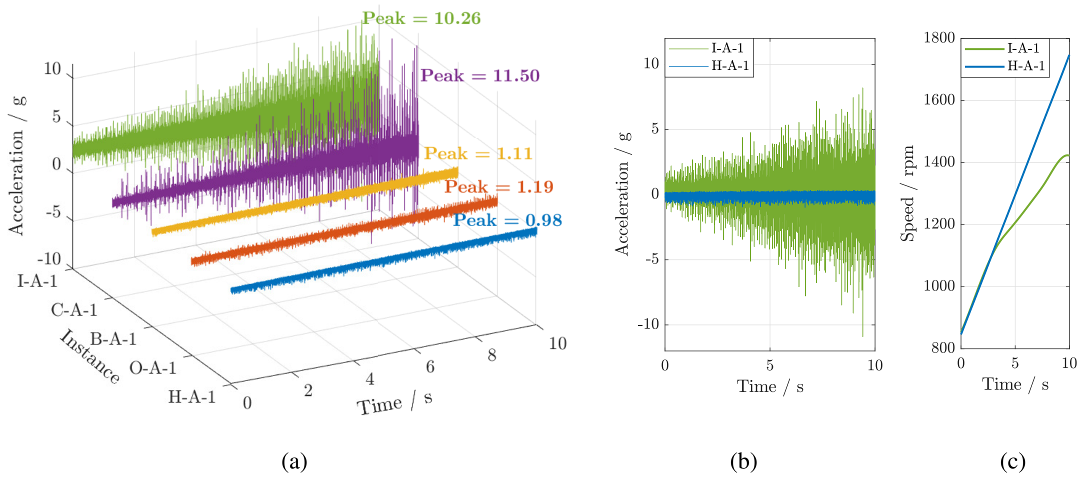


Figure 1: (a) Vibration data for the five health states while considering the first measurement regarding increasing speed, (b) Mean-normalized vibration data for first-measured healthy and inner ring fault instances, (c) Corresponding first-measured shaft speed for healthy and inner ring fault instances

Furthermore, as compared to the healthy instance, the maximum value of the vibration amplitude has increased approximately tenfold for the latter fault instances. This increase in the maximum vibration amplitude is clearly seen in Figure 1(b). Additionally, the vibration intensity for the inner ring fault is higher than that for the healthy instance, even at comparable or lower shaft speed, as can be deduced from Figures 1(b) and 1(c). Hence features extracted in the time domain should capture this impulsiveness.

Given the high sampling rate of $200\,\mathrm{kHz}$ and increased processing time, frequency analysis was conducted to discern if down-sampling was feasible without losing relevant information. The frequency spectrum of the vibration data from the second measurement involving the speed variations and health states is presented in Figure 2. As visible from Figure 2, the vibration amplitudes above $40\,\mathrm{kHz}$ are not significantly distinguishable. Hence, it can be postulated that frequency components above $40\,\mathrm{kHz}$ are no longer relevant for fault detection of the presented instances. Thus, it can be concluded that down-sampling by a factor of two should not impair accurate fault detection. On a closer look at Figure 2, it can be deduced that the inner ring fault instances can clearly be distinguished from all other health states due to the intensity of the vibration amplitude and specifically at about $35\,\mathrm{kHz}$. This frequency component at about $35\,\mathrm{kHz}$ is also slightly visible by the instances of the ball fault. Thus, it can be postulated that this frequency component corresponds to a bearing element's resonance frequency, and it is independent of the speed variations. Although only presented for the second measurement, the results are comparable to the others.

Feature extraction and selection

To evaluate the effectiveness of the presented feature extraction and feature selection methods under a noisy environment, the features described in section 2.1 are extracted from raw and band-passed filtered vibration data. Via sensitivity analysis the bandpass frequency range was set at $[5.7;\,17]\,\mathrm{kHz}$.

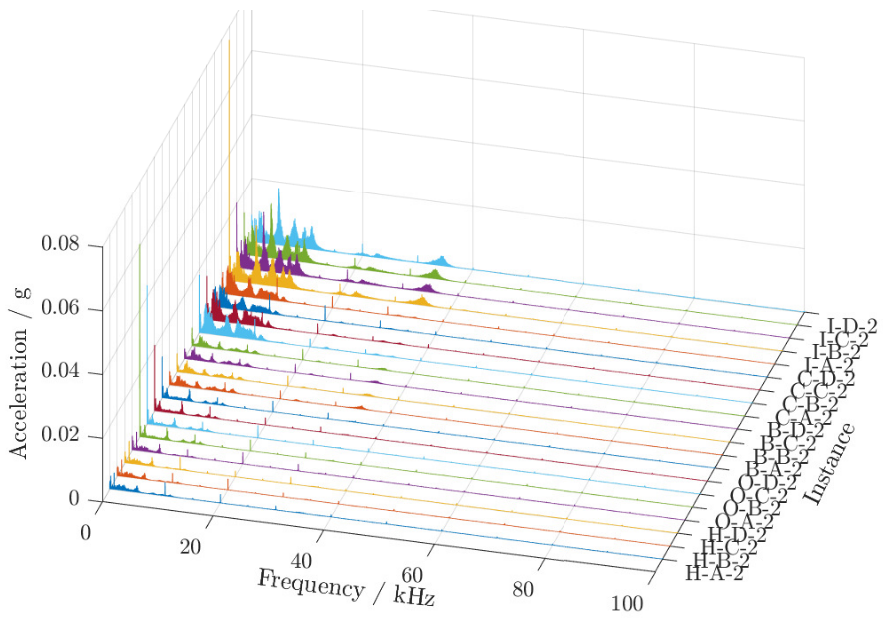
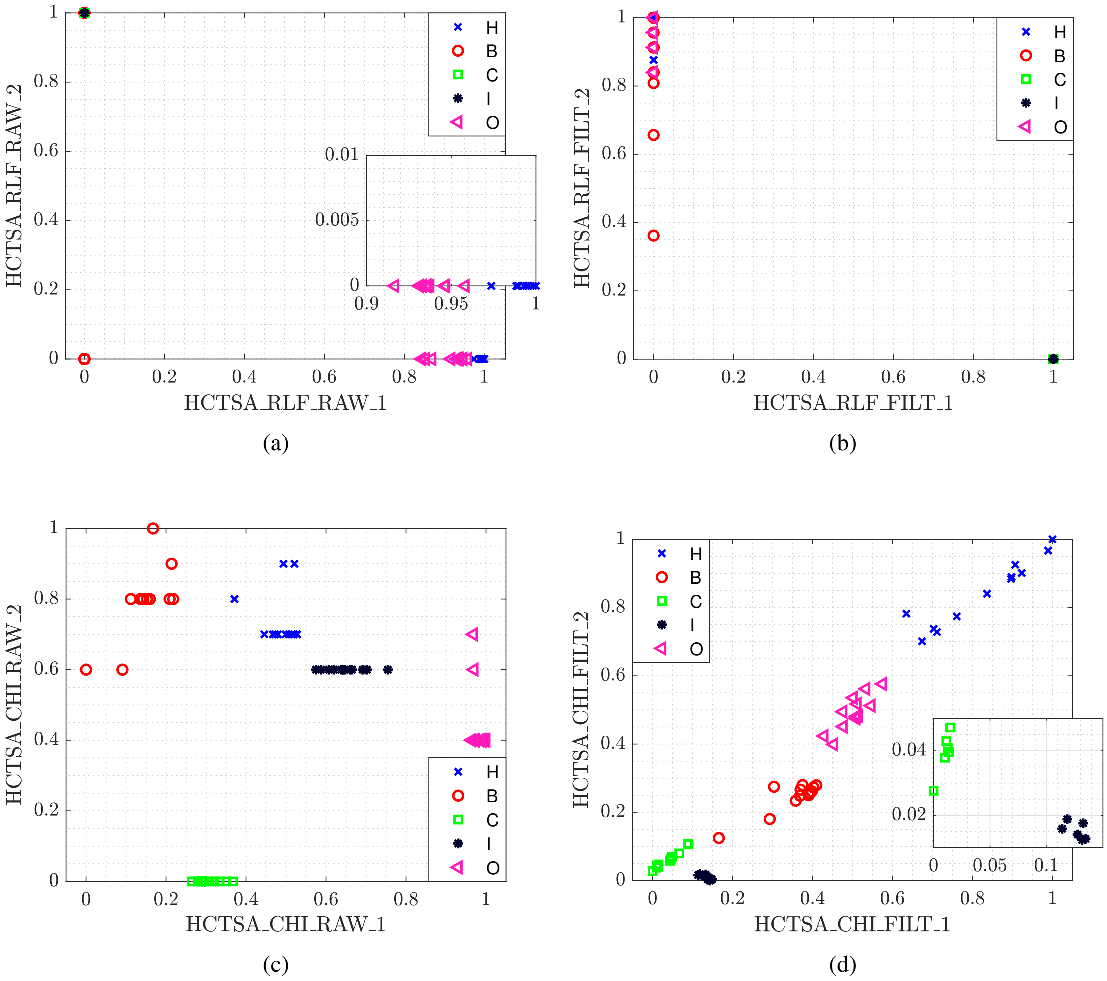


Figure 2: Waterfall plot of the frequency spectrum for the second measurement involving all experimented speed variations and health states

Since the primary goal of this case study is to distinguish between healthy and faulty instances, irrespective of the time-varying shaft speed condition, the task can be generalized as a binary classification problem. Thus, the feature selection methods described in section 2.2.1 for classification are employed to select two top-ranked features. The two top-ranked features for both feature selection methods are exemplarily depicted in Figure 3. As can be inferred from Figures 3(a) and 3(b), except for the feature HCTSA_RLF_RAW_1, the features selected by **reliefF** do not allow a clear distinction between the different health states. On a closer look at the zoomed region of Figure 3(a), it can be seen that HCTSA_RLF_RAW_1 allows at least to distinguish between healthy and faulty instances. On the other hand, as can be seen from Figures 3(c) and 3(d), the features selected by **chi-square test** allows for a distinctive clustering of the different health states. Furthermore, as can be deduced from Figure 3(d), the features selected from the filtered vibration data can be employed independently.



(a) (b) (c) (d)

Figure 3: Scatter plot of two top-ranked features: selected by **reliefF** from (a) raw and (b) filtered vibration data, respectively; and selected by **chi-square test** from (c) raw and (d) filtered vibration data, respectively. (Actual names of selected features can be found in Table 1)

Hence, it can be concluded from this case study that at least two independent features from the HCTSA package can successfully be utilized as health indicators to detect bearing faults, irrespective of the time-varying shaft speed condition.

3.2 Case Study: Fault detection on a real-world wind turbine

The data set evaluated in this case study involves real-world vibration data acquired from six same-model wind turbines in a wind farm located in northern Sweden. The data set is publicly available, and a detailed description of the wind turbine schematics and sensor placement can be found in the accompanying paper by Martin del Campo Barraza et al.[28, 29]. Each provided $1.28\,\mathrm{s}$ vibration data is sampled at $12.8\,\mathrm{kHz}$ and axially measured from the output shaft's housing of the wind turbine gearbox. Additionally available is the mean output shaft speed corresponding to each vibration measurement. The measurements were recorded for approximately four years, with a 12-hour mean interval. The turbines are labelled Turbine 1 through 6 as proposed by the authors[29]. During the measurement period, Turbine 2 had a possible electrical failure, Turbine 5 had two recorded mechanical component failures, while the others had none. The possible electrical sensor failure of Turbine 2 was detected and rectified shortly after commissioning. The first failure recorded within Turbine 5 was an inner race fault of the output shaft bearing. Consequently, this led to the replacement of this bearing after about 1.2 years after commissioning. The second failure recorded within Turbine 5 was likewise an inner race fault but on one of the planetary gear bearings near the main shaft. As a consequence, the entire gearbox was replaced after about two years after commissioning. Given the relevance of the time of bearing and gearbox replacement, the corresponding timestamps are $\mathbf{1.5673}$ and $\mathbf{2.3708}$ years, respectively. In this paper, Turbine $\mathbf{5}$ is considered in detail, while the other wind turbines are analysed for comparison.

Preliminary Analysis

Before the vibration measurements analysis, the measured speed at the output shaft is analysed to investigate the underlying operating condition. As can be seen in Figure 4(a), the output shaft speed has a lower and upper bound at approximately $700\,\mathrm{rpm}$ and $1200\,\mathrm{rpm}$, respectively. Although the output shaft speed is stationary per measurement, it is nonstationary over the entire measurement period since a random speed sets in from the operating range. Furthermore, as be inferred from Figure 4(a), Turbine $\mathbf{5}$ had a significant downtime shortly after commissioning for reasons unfortunately not disclosed.

The shaft speed distributions over the presented turbines are analysed to investigate if the turbines operated under similar operating conditions. As seen in Figure 4(b), all wind turbines possess a trimodal distribution, whereby the local mode at approximately $700\,\mathrm{rpm}$ is not as pronounced as at about $800\,\mathrm{rpm}$ and $1150\,\mathrm{rpm}$. Furthermore, as can be inferred from Figure 4(b), all turbines have more or less similar speed distribution, with Turbines $\mathbf{3}$, $\mathbf{4}$ and $\mathbf{5}$ subjected to roughly the same speeds on average and are possibly geographically nearer than Turbines $\mathbf{1}$, $\mathbf{2}$ and $\mathbf{6}$, which also have more or less the same average speed. Thus, in summary, the wind turbines are not only of the same model but were subjected to similar operating conditions as well.

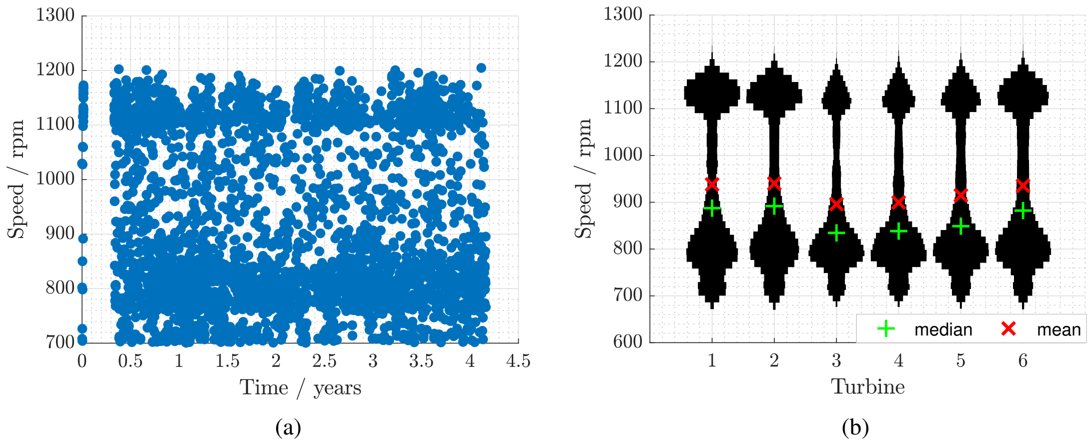


Figure 4: (a) Scatter plot of the mean output shaft speed over the measurement period for Turbine 5, (b) Violin plot[30] of the mean output shaft speed distribution over Turbines $\mathbf{1}$ through $\mathbf{6}$

Feature extraction and selection

Since Turbine $\mathbf{5}$ was monitored continuously from commissioning till at least one failure case occurred, the accompanying data can be considered run-to-failure data. Thus, the main objective is to map each vibration measurement to a continuous value that represents the possible degradation of the respective turbine. To this end, the presented feature extraction and feature selection methods described in section 2.1 and 2.2.2, respectively, are employed to extract and select a top-ranked feature from the raw vibration measurements. The vibration measurements are not filtered, given the relatively low sampling rate in comparison to the previous case study. Figures 5(a) and 5(b) depicts the top-ranked feature selected by Monotonicity and RReliefF, respectively, from raw vibration data over the measurement period of Turbine $\mathbf{5}$.

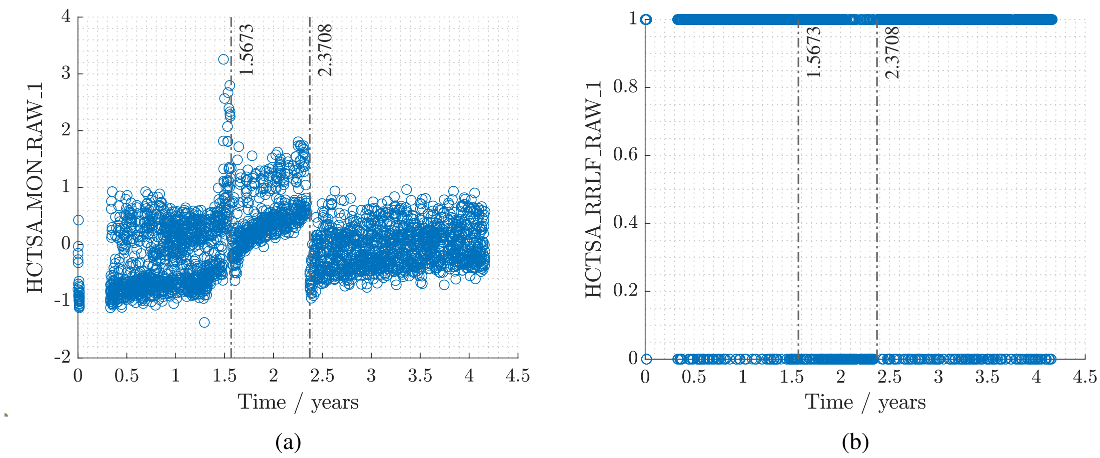
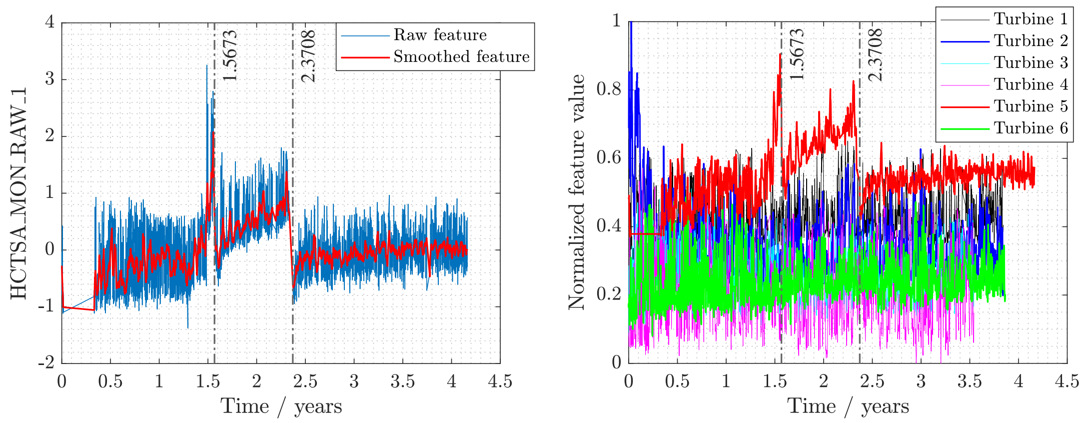


Figure 5: Scatter plot of the top-ranked feature selected from raw vibration data of Turbine $\mathbf{5}$ by (a) Monotonicity and (b) RReliefF, respectively. Actual names of selected features can be found in Table 1. The timestamps $\mathbf{1.5673}$ and $\mathbf{2.3708}$ years correspond to bearing and gearbox replacement, respectively

Given that the actual mapping of health states to vibration data is unavailable, the selected features cannot be validated by comparing them to a baseline or ground-truth. However, significant regions within Figure 5(a) justifies that the top-ranked feature (HCTSA_MON_RAW_1) determined by Monotonicity reflects the components' degradation that occurred within Turbine 5, and thus reflects its continuous health state over the measurement period. Firstly, as can be inferred from Figure 5(a), there is a monotonic increase in the feature value from approximately 1.4 years, that is before the bearing replacement (1.5673 years), which can be interpreted as possible degradation. Furthermore, soon after the bearing replacement, there is a minor decline in the feature value, which suggests that multiple faults had occurred, and unfortunately, only one was localised. Secondly, in a period after the bearing replacement (1.5673 years) and before the gearbox replacement (2.3708 years), there is a monotonic trend that ceased right after the gearbox was replaced. On the other hand, the top-ranked feature selected by RReliefF, as seen in Figure 5(b), is not sensitive to the degradation that occurred within Turbine 5. Thus, this feature is not further considered.

The top-ranked feature determined by Monotonicity is also extracted from the raw vibration data of the other wind turbines to investigate the effectiveness of the proposed methods across the turbines. For a better comparison, the features are smoothed, normalized in the range $[0, 1]$ and presented in a line plot. Figure 6(a) exemplarily shows the raw and smoothed feature values for Turbine 5 over the measurement period. As can be seen, the feature characteristics are not smoothed out. Hence the same smoothing was applied to the features of the other wind turbines. As can be seen in Figure 6(b), the feature values of Turbine 5 generally tends to be above the feature values of the other wind turbines. This implies that a direct comparison of the feature values is not feasible over the presented turbines. However, the feature values tend to be effective and consistent per turbine. As can be deduced from Figure 6(b), the feature values of Turbines $1, 3, 4$ and 6 lies at a constant level, which concurs with the fact that they were healthy over the measurement period.



(a) (b)

Figure 6: Top-ranked feature selected from raw vibration data by Monotonicity: (a) Raw and smoothed feature value over the measurement period for Turbine 5, (b) Normalized and smoothed feature values over the respective measurement periods of Turbines 1 through 6. (Actual names of selected features can be found in Table 1)

Furthermore, the possible electrical sensor failure that occurred in Turbine 2 soon after commissioning is reflected by the distinctively high feature values up to about 0.3 years. Hence, it can be concluded from this case study that the top-ranked feature selected from the HCTSA package can successfully be utilized as an indicator to monitor the health of wind turbines operating under real-world conditions.

4. Conclusion

The focus of this paper is finding distinctive, reliable, independent, and robust health indicators for detecting faults in technical systems. To this end, classical features, features from the EMG toolbox and from the HCTSA package were extracted from raw and filtered vibration data. Furthermore, to select top-ranked features from the set of extracted features, several feature selection methods were also presented for prospective diagnostics and prognostics tasks. Two case studies were employed to evaluate the proposed techniques. The first case study focused on fault detection of seeded bearing faults operating under time-varying shaft speed, and the second case study focused on detecting fault on a real-world wind turbine subjected to time-varying wind speed. From both case studies, it can be concluded that features from the time-domain and specifically from the HCTSA package can act as viable health indicators for technical systems and specifically for rotating machinery irrespective of the time-varying operating conditions. However, to reliably localize the faulty component, for example, which bearing components are exactly damaged, either the frequency, the time-frequency analysis or adequate labelled quality data for training a pattern recognition or machine learning algorithm is indispensable.

As a future outlook, the embedded and wrapper feature selection methods can be compared to the presented filter methods to evaluate their strengths and weaknesses for diagnostics and prognostics of technical systems. Further case studies from other technical systems are planned and inevitable to further evaluate the proposed methods.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 451737409.

The Authors are indebted to Huang and Baddour(26, 27), and Martin del Campo Barraza et al.(28, 29) for making the data sets publicly available in the light of promoting research.

References

1. S. Umbaugh. *Digital Image Processing and Analysis: Human and Computer Vision Applications with CVIPtools, Second Edition*. Taylor & Francis, 2010. ISBN: 9781439802052.
2. I. Howard. “A Review of Rolling Element Bearing Vibration’Detection, Diagnosis and Prognosis’”. In: (1994).
3. J. K. Kimotho and W. Sextro. “An approach for feature extraction and selection from non-trending data for machinery prognosis”. In: *Proceedings of the second european conference of the prognostics and health management society*. Vol. 5. 4. 2014, pp. 1–8.
4. P Večeř, M. Kreidl, and R Šmíd. “Condition indicators for gearbox condition monitoring systems”. In: *Acta Polytechnica* 45.6 (2005).
5. K. Smith. “Crest Factor Analysis for Rolling Element Bearing”. In: *Proceedings of the Meeting of the Mechanical Failures Prevention Group (18th), Gaithersburg, Maryland*. 1972, pp. 9–20.
6. D. R. Houser and M. J. Drosjack. *Vibration signal analysis techniques*. Tech. rep. Ohio State Univ Columbus Dept of Mechanical and Aerospace Engineering, 1973.
7. Z. Peng, W. T. Peter, and F. Chu. “A comparison study of improved Hilbert–Huang transform and wavelet transform: application to fault diagnosis for rolling bearing”. In: *Mechanical systems and signal processing* 19.5 (2005), pp. 974–988.
8. B. D. Fulcher, M. A. Little, and N. S. Jones. “Highly comparative time-series analysis: the empirical structure of time series and their methods”. In: *Journal of the Royal Society Interface* 10.83 (2013), p. 20130048.
9. B. D. Fulcher and N. S. Jones. “hctsa: A computational framework for automated time-series phenotyping using massive feature extraction”. In: *Cell systems* 5.5 (2017), pp. 527–531.
10. J. Too, A. R. Abdullah, and N. M. Saad. “Classification of hand movements based on discrete wavelet transform and enhanced feature extraction”. In: *International Journal of Advanced Computer Science and Applications* 10.6 (2019), pp. 83–89.
11. J. Too et al. “EMG feature selection and classification using a Pbest-guide binary particle swarm optimization”. In: *Computation* 7.1 (2019), p. 12.
12. R.-V. Sánchez et al. “Multi-fault diagnosis of rotating machinery by using feature ranking methods and SVM-based classifiers”. In: *2017 International conference on sensing, diagnostics, prognostics, and control (SDPC)*. IEEE. 2017, pp. 105–110.
13. B. Nayana and P Geethanjali. “Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal”. In: *IEEE Sensors Journal* 17.17 (2017), pp. 5618–5625.
14. A. Phinyomark, C. Limsakul, and P. Phukpattaranont. “A novel feature extraction for robust EMG pattern recognition”. In: *arXiv preprint arXiv:0912.3973* (2009).
15. B. D. Fulcher, M. A. Little, and N. S. Jones. *HCTSA Documentation*. Last viewed in: 03.2021. URL: `https://hctsa-users.gitbook.io/hctsa-manual/`.

16. J. Bogatinovski, D. Kocev, and A. Rashkovska. “Feature Extraction for Heartbeat Classification in Single-Lead ECG”. In: *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. 2019, pp. 320–325.

17. C. Murphy-Royal et al. “Stress gates an astrocytic energy reservoir to impair synaptic plasticity”. In: *Nature communications* 11.1 (2020), pp. 1–18.

18. M. Hennig et al. “Comparison of Time Series Clustering Algorithms for Machine State Detection”. In: *Procedia CIRP* 93 (2020). 53rd CIRP Conference on Manufacturing Systems 2020, pp. 1352–1357. ISSN: 2212-8271.

19. I. Guyon and A. Elisseeff. “An introduction to variable and feature selection”. In: *Journal of machine learning research* 3.Mar (2003), pp. 1157–1182.

20. K. Tadist et al. “Feature selection methods and genomic big data: a systematic review”. In: *Journal of Big Data* 6.1 (2019), pp. 1–24.

21. M. F. Redondo and C. H. Espinosa. “How to select the inputs for a multilayer feedforward network by using the training set”. In: *International Work-Conference on Artificial Neural Networks*. Springer. 1999, pp. 477–486.

22. M. Robnik-Šikonja and I. Kononenko. “Theoretical and empirical analysis of ReliefF and RReliefF”. In: *Machine learning* 53.1 (2003), pp. 23–69.

23. J. Coble and J. W. Hines. “Identifying optimal prognostic parameters from data: a genetic algorithms approach”. In: 2009.

24. D. Berryman et al. “Nonparametric tests for trend detection in water quality time series”. In: *JAWRA Journal of the American Water Resources Association* 24.3 (1988), pp. 545–556.

25. J. A. Carino et al. “Remaining useful life estimation of ball bearings by means of monotonic score calibration”. In: *2015 IEEE International Conference on Industrial Technology (ICIT)*. 2015, pp. 1752–1758.

26. H. Huang and N. Baddour. *Bearing Vibration Data under Time-varying Rotational Speed Conditions*. 2019. DOI: `https://doi.org/10.17632/v43hmbwxpm.2`.

27. H. Huang and N. Baddour. “Bearing vibration data collected under time-varying rotational speed conditions”. In: *Data in Brief* 21 (2018), pp. 1745–1749. ISSN: 2352-3409. DOI: `https://doi.org/10.1016/j.dib.2018.11.019`.

28. S. Martin del Campo Barraza, F. Sandin, and D. Strömbergsson. *Dataset concerning the vibration signals from wind turbines in northern Sweden*. Last viewed in: 03.2021. URL: `http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-70730`.

29. S. Martin-del Campo, F. Sandin, and D. Strömbergsson. “Dictionary Learning Approach to Monitoring of Wind Turbine Drivetrain Bearings”. In: *International Journal of Computational Intelligence Systems* 14 (1 2020), pp. 106–121. ISSN: 1875-6883. DOI: `https://doi.org/10.2991/ijcis.d.201105.001`.

30. Jonas. *Violin Plots for plotting multiple distributions (distributionPlot.m)*. MATLAB Central File Exchange. Retrieved March. 2021. URL: https://www.mathworks.com/matlabcentral/fileexchange/23661-violin-plots-for-plotting-multiple-distributions-distributionplot-m.

A. Appendix

The actual names of features selected in this paper can be found in Table 1, as proposed by the Authors of the respective toolboxes. Please see section 2.1 for further details.

Table 1: Actual names of selected features

Feature	**Actual name**	**Category**
HCTSA_RLF_RAW_1	MF_arfit_1_8_sbc.minper	Modelfit,arfit
HCTSA_RLF_RAW_2	MF_arfit_1_8_sbc.hasInfper	Modelfit,arfit
HCTSA_RLF_FILT_1	EX_MovingThreshold_1_01.mediankickf	Outliers
HCTSA_RLF_FILT_2	FC_Surprise_T1_20_2_udq_500.lq	Information,symbolic
HCTSA_CHI_RAW_1	AC_40	Correlation
HCTSA_CHI_RAW_2	IN_AutoMutualInfoStats_diff_20_gaussian.pcrossmedian	Information,correlation
HCTSA_CHI_FILT_1	SY_StdNthDer_5	Entropy
HCTSA_CHI_FILT_2	SY_StdNthDerChange.fexp_rmse	Entropy
HCTSA_RRLF_RAW_1	SY_VarRatioTest_24682468_00001111.IIDperiodminpValue	Vratiotest
HCTSA_MON_RAW_1	SP_Summaries_welch_rect.linfitloglog_mf_a2	FourierSpectrum