A sequential hybrid method for full lifetime remaining useful life prediction of bearings in rotating machinery under varying speed conditions

Koen Geurts[1], Kerem Eryilmaz[2], and Ted Ooijevaar[3]

[1] *Flanders Make, Oude Diestersebaan 133, 3920 Lommel, Belgium*
koen.geurts@flandersmake.be

[2,3] *Flanders Make, Gaston Geenslaan 8, 3001 Leuven, Belgium*
kerem.eryilmaz@flandersmake.be
ted.ooijevaar@flandersmake.be

Abstract

Optimal scheduling of the maintenance of bearings in rotating machinery requires accurate remaining useful life (RUL) prediction during the entire lifetime of the bearing. For that reason, this paper proposes a sequential hybrid method that combines the strengths of statistical and data-driven approaches. A statistical model-based approach is preferred before a bearing fault is detected, and a data-driven approach once a bearing fault is detected from the vibration measurements. The method is tested and evaluated on an extensive dataset of accelerated lifetime tests of deep groove ball bearings. It is shown that the method, with a limited amount of training data, delivers accurate RUL predictions during both the healthy stage of the bearing lifetime, as well as during the final stages of increasing degradation under both constant and varying speed conditions.

1. Introduction

Remaining useful life (RUL) prediction is becoming increasingly important for reducing maintenance costs, increasing efficiency in maintenance planning and for companies transitioning from producing products to delivering products as a service. Failure prognosis of bearings can be performed using model-based or data-driven methods, or any combination of these methods (Jammu & Kankar, 2011) (Kan, Tan, & Mathew, 2015) (Mrugalska, 2019). These hybrid combinations can take different forms (Liao & Köttig, 2014), however, the combination of a statistical model-based method combined with a data-driven approach using an autoencoder (AE) as health indicator(HI) to deliver a RUL estimation during the entire lifetime of a rotating component seems to be novel.

Model-based methods are able to deliver RUL predictions during the steady-state operation (Halme & Andersson, 2009) or the healthy stage with stable vibration levels, while data-driven methods based on sensor measurements have difficulties delivering a useful RUL prediction under such conditions (Zhao, Zhang, Wang, Zhou, & Cheng, 2019). Model-based methods consist of developing mathematical or physical models based on historical data that needs to include run-to-failure data to determine trends in the health state of a component (Ferreira & Gonçalves, 2022). They can also be categorized into micro-level models or damage propagation models (such as fatigue life models (Harris & Yu, 1999)) and macro-level models, which represent a system in a simplified way, defining the relations between the various types of variables (input, state, and output) of that system (such as ISO 281 (ISO281, 2007)).

Data-driven algorithms rely on the availability of data to fit a model to the system behaviour. The historical data must include run-to-failure conditions for training in order to predict the RUL or identify faults (Bourgana et al., 2021). Generally, data-driven approaches are based on statistical pattern recognition and machine learning (ML) techniques, utilizing techniques such as particle filters, extended Kalman filters, Gaussian process regression, support vector machines or artificial neural networks (Kan et al., 2015). These data-driven methods can use raw measurement data or features computed from the measurements as input. As these methods make use of sensor measurement data, they can deliver accurate RUL predictions during the degradation stage when pitting and spalling (Halme & Andersson, 2009) can be detected. To increase the applicability of these methods, (Zhao et al., 2019) assume a constant RUL before the anomaly is detected. Other researchers (Aydemir & Acar, 2020) (Kamat, Sugandhi, & Kumar, 2021) aim to include the anomaly detection in the estimator training process.

The goal of this paper is to propose a sequential hybrid RUL prediction method that combines a model-based approach in the healthy stage of the bearing lifetime by using an estimated lifetime based on a bearing basic rating life, and a data-driven method by using an autoencoder (AE) that only requires healthy data for training, to generate a health indicator (HI) from vibration measurements, and then applying Bayesian filtering to determine an accurate RUL prediction during the entire lifetime of the bearing. The switch from the model-based approach to the data-driven method takes place once an anomaly is detected in the vibration measurements.

This approach has multiple benefits compared to others. Firstly, it requires little historical healthy and faulty measurements for training the AE. Secondly, thanks to the use of an AE to generate the vibration HI from statistical and physics-inspired vibration features, the method can cope with changing operating conditions. Finally, this approach delivers RUL predictions throughout the entire lifetime of the bearing, which increases the effectiveness of maintenance planning for a large fleet of rotating machinery.

This paper starts with a description of the developed methodology in Section 2, where the different components of the sequential hybrid method are elaborated. In Section 3, the experimental procedure employed to acquire the large amount of accelerated lifetime data is described. The results of the proposed sequential hybrid method are compared to those from a pure statistical RUL prediction and the PCA-based sequential hybrid method, and discussed in Section 4. Finally, in Section 5, the main conclusions are summarized.

2. METHODOLOGY

The sequential hybrid RUL prediction method proposed in this paper is schematically illustrated in Figure 1 and exploits the strength of both a statistical model-based RUL prediction approach before an anomaly is detected and a data-driven RUL prediction approach once an anomaly is detected from the vibration measurements. Both approaches are subsequently discussed in Section 2.1 and 2.2, while Section 2.3 explains how both methods are joined to deliver a RUL estimation during the entire lifetime of the bearing.

2.1. Statistical model based RUL prediction

The model-based approach leverages prior knowledge of the bearing. This can be in the form of an average lifetime estimate based on a bearing basic rating life (e.g. L_{10} or L_{50}) or one based on statistics if sufficient experimental lifetime data is available.

The bearing basic rating life L_{10} is the fatigue life that 90% of a sufficiently large group of identical bearings operating under identical conditions survive in accordance with ISO 281 (ISO281, 2007). The actual service life in a given application can deviate significantly from the calculated basic rating life. Service life in a particular application depends not only on the load and bearing size, but also on a variety of influencing factors including lubrication conditions, degree of contamination, proper mounting and other environmental conditions, represented by the life modification factor a_{SKF} to supplement the basic rating life. The modified SKF rating life (SKF, 2011) L_{10m} is given by:

$$L_{10m} = a_1 a_{skf} L_{10} = a_1 a_{skf} \left(\frac{C}{P}\right)^p \quad (1)$$

where C represents the basic dynamic load rating, P the equivalent dynamic bearing load and the exponent p is 3 for ball bearings and $\frac{10}{3}$ for roller bearings.

Building on the knowledge that the fatigue lifetime of a bearing follows a Weibull reliability function with a shape factor of 1.3 (NSWC, 2011), an estimated average lifetime L_{50} can be determined. The Weibull reliability function is given by equation 2 where $k > 0$ is the shape parameter which is 1.3 for ball and roller bearings and $\lambda > 0$ is the scale parameter or the characteristic life where 63.2% of the components will fail.

$$R(t) = \frac{k}{\lambda}\left(\frac{t}{\lambda}\right)^{k-1} e^{-(t/\lambda)^k} \quad (2)$$

With the availability of actual experimental data, the parameters of this Weibull distribution and the associated L_{50} average lifetime can be more accurately determined. From these historical measurements, a second distribution representing the moment of anomaly detection can also be estimated. This distribution is constructed from applying an anomaly detection algorithm on historical that will be discussed in section 2.2.3.

Using these two distributions, both the curve for the time-to-anomaly prediction can be estimated, as well as the average expected lifetime L_{50}, as visualised in Figure 2a and b. Using this information, the average of the distribution is the statistical RUL_{50} prediction used as part of the proposed sequential hybrid method in this paper can also be estimated as is shown in Figure 2b. This estimate actually follows the same trend as the time-to-anomaly curve, but with an offset equal to the average RUL after anomaly detection. At the start of the lifetime of the bearing, the estimate of the RUL_{50} coincides with the L_{50} estimate of the historical data and is equal to the sum of the expected time-to-anomaly in Figure 2a and the average time to End-of-Life after anomaly detection in Figure 2b. With increasing time and revolutions, the RUL_{50} decreases linearly until the first moment an anomaly in the vibrations can be expected, i.e. the left side of the statistical distribution of anomaly detection. As long as no anomaly is detected

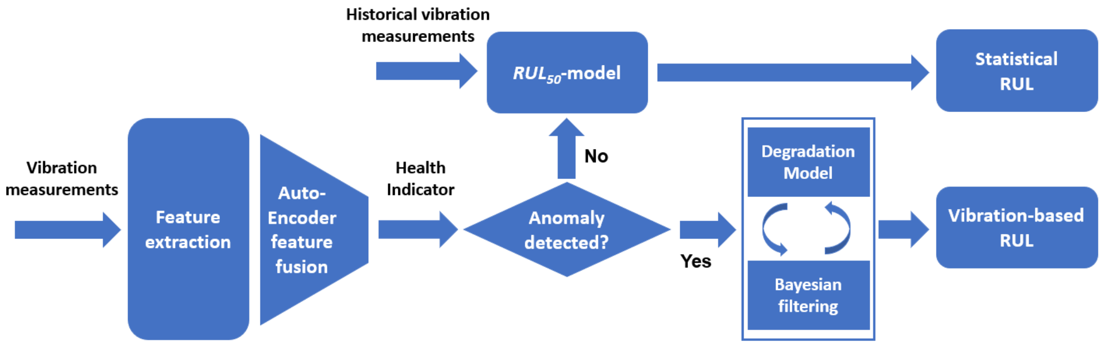

Figure 1. Continuous monitoring of the asset allows for sequential switching from statistical model-based RUL prediction to vibration-based RUL prediction.

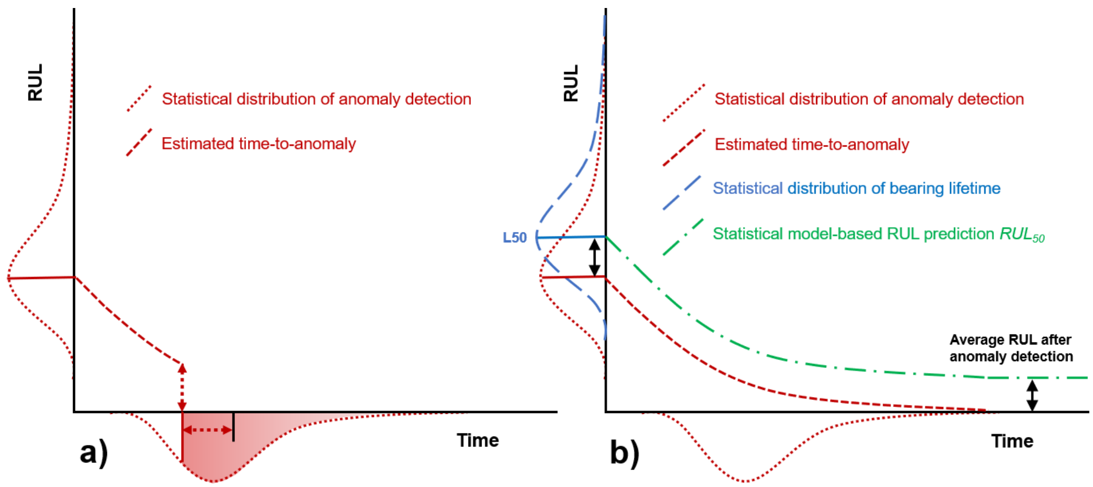

Figure 2. Determination of a) the time-to-anomaly curve based on the expected time of anomaly detection and b) RUL_{50} prediction based on L_{50} and the time-to-anomaly curve.

for the current bearing, it means that this bearing belongs to the remainder of the distribution as can be seen in Figure 2a, where the rate of change of the time-to-anomaly estimation decreases as the center of gravity of the cut-off distribution moves to the right. This process continues until the RUL converges to a horizontal line if the bearing out-survives all bearings in the historical dataset. The RUL prediction at that time is the average time it takes for the bearing to reach its end-of-life after anomaly detection as this anomaly and the increased degradation is expected to happen at any time as shown in Figure 2b. The full statistical RUL_{50} curve is the RUL prediction used for the remainder of this paper as the first part of the sequential hybrid method, until the moment an anomaly in the vibration measurements is detected.

2.2. Data driven RUL prediction

From the moment that a bearing fault is detected based on the vibration measurements, the method relies on a data-driven approach to perform the RUL prediction based on the health indicator (HI) from vibration features. An autoencoder (AE) is chosen to reconstruct vibration features and the reconstruction error is used as HI to represents the current health state of the bearing, as this AE can (and in fact, should) be trained only on healthy-state data [1]. It can also capture non-linear relationships between the vibration features, and can learn to compensate for changing operating conditions at the early stages of the degradation process. Finally, Bayesian filtering with an exponential degradation model is used to predict the future health state of the bearing and the corresponding RUL including an uncertainty estimate. The extraction of the vibration features, the construction of the AE based HI and the fitting of the degradation model by means of Bayesian filtering are described in the following sections.

2.2.1. Feature extraction

Numerous vibration features can be extracted from vibration measurements and can originate from the vibration signal in the time-domain, the frequency-domain or even from specific

[1]The method still requires limited faulty-state data for thresholding and/or scaling as detailed in Section 2.2.2, but not for the model training.

fault frequencies in the frequency-domain. Based on experience and engineering knowledge, a sub-set of all these features is selected to be used in generating the health indicator of for the proposed sequential hybrid method.

Seven time-domain features are selected on the basis of showing high sensitivity to the health state of the bearing, as well as showing a clear dependency or independency on the induced RPM on the signal. The independent features Kurtosis, margin factor, impulse factor and shape factor do not need any scaling, however, the peak-to-peak, RMS and Shannon entropy feature values are linearly scaled by the RPM range. To remove their RPM dependency as much as possible as can be seen in equation 3 where the scaling factor for the entire time series of the feature values is determined. This can be done as long as the low feature values correspond to the low RPM values and similarly for the higher feature values.

$$scale(t) = \frac{max(y(t)) - min(y(t))}{max(rpm(t)) - min(rpm(t))} \tag{3}$$

Furthermore, the maximum peak and the RMS of the frequency spectrum are selected as frequency-domain features, where the maximum peak does not need to be scaled, while the RMS of the frequency spectrum is scaled by the RPM similarly to the time-domain features. Finally, fault frequency features (Ooijevaar et al., 2019) that look at specific frequencies representative for inner and outer race fault modes of the bearing are also computed, to result in a total of 17 features to be selected as input for training and using the autoencoder.

A final step in the feature extraction process is the normalization of the features in order to efficiently train the AE. All features are normalised using Z-score normalization, as shown in equation 4 based on healthy vibration data of the bearing. This normalises the features to fit in the interval $[-1, 1]$ for the healthy state of the bearing.

$$z(t) = \frac{x(t) - \mu}{\sigma} \tag{4}$$

2.2.2. Physics-inspired health indicator

Principle Component Analysis (PCA) (Dong & Luo, 2013) is often used to fuse vibration features into a health indicator. Since PCA outputs the HI as a linear combination of the input features, it is expected that non-linear effects that could arise with non-stationary operating conditions will limit the effectiveness of PCA for the proposed sequential hybrid method and the dataset containing varying operating conditions. In this work, the PCA feature fusion method is applied for the sequential hybrid method, but furthermore replaced by the use of an AE to construct the required health indicator or reduce dimensionality by fusing vibration features. The AE should be able to handle non-linear effects present in the training data. Both methods will be examined and compared.

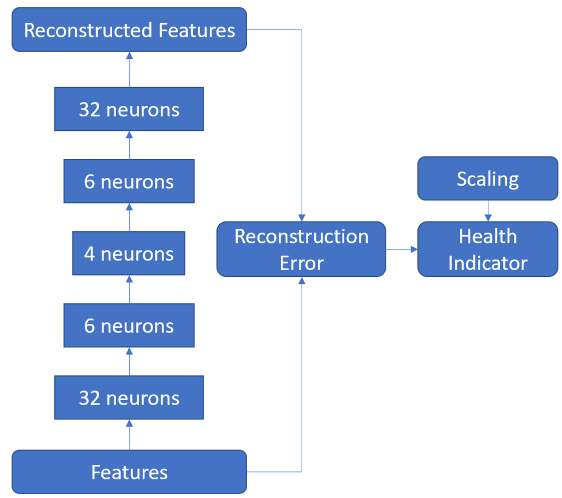

Figure 3. The procedure of obtaining health indicator.

The bottle-neck of the autoencoder can deliver a Virtual Health Index (Hervé de Beaulieu, Shekhar Jha, Garnier, & Cerbah, 2022) that can be used as representation of the health state of the component. However, the reconstruction error of the full AE can also deliver a quantification of the deviation from a healthy state, and thus a suitable HI. The latter is applied in the purpose of this research paper.

As a first step, an AE is trained on healthy data alone. This teaches the AE to reproduce the feature values under healthy conditions, meaning its reconstruction error is expected to increase when the bearing is no longer healthy. Hence, reconstruction errors closer to zero indicate a healthier bearing.

No formal procedure to determine the exact architecture of the AE has been used here, but a good rule of thumb is to keep it as shallow and narrow as possible to avoid overfitting. In this case, the width of the input/output layers are set to accommodate as many features as necessary, and the other layers are kept as few and narrow as possible while managing to reconstruct the input features. The procedure to reduce the feature input values to a health indicator using the AE can be seen in Figure 3.

The use of features computed from the raw vibration signal delivers multiple benefits. First of all, selection of specific vibration features as input for the AE is a useful tool for the maintenance engineer or developer of the AE to implant specific historical experience or physical understanding of the system into the AE. This can keep the AE layers narrow and shallow at the same power of reconstruction, and reduces computational effort and possibly the size of the minimum training set required. Secondly, the analysis of the total reconstruction error, as well as the reconstruction error of the indi-

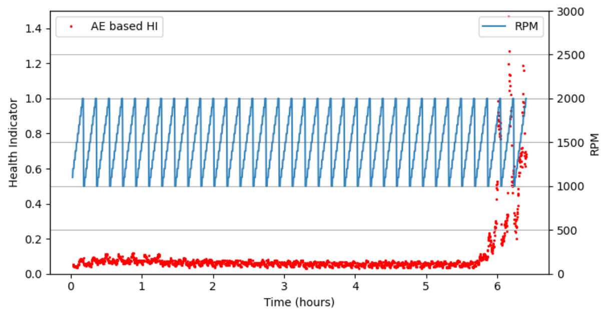

Figure 4. Health indicator based on AE reconstruction error that is independent of operating condition in the healthy stages of the bearing lifetime.

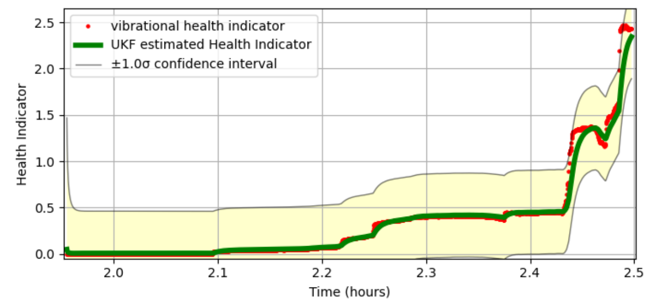

Figure 5. UKF estimated HI resulting from the reconstruction error output of the AE.

vidual features can deliver more insight into the health state of the component, and the failure mode that it endures. The earlier selected features, supplemented with the RPM, are used to train the AE such that it can learn the influence of the RPM on the different features (at least during healthy operation) and non-linear RPM dependencies that are still present in the scaled and normalised feature values.

The final health indicator resulting from the AE can then be used for RUL prediction, however, to do this, one more value is required: the level of the health indicator at the EoL of the specific component or bearing. If historical vibration data is available, the simplest way to determine this, is to compute the reconstruction error of the AE using these vibration measurements at (or just before) the EoL. This value can be used to determine a relevant threshold. When the HI of the bearing reaches this threshold, it should be considered as failed and has reached its EoL. Figure 4 shows the resulting health indicator which clearly is independent of the RPM in the healthy stages of the bearing lifetime.

2.2.3. Anomaly detection

The statistical model RUL_{50} prediction is dependent on the moment in time with respect to the distribution of the moment of anomaly detection as is displayed in Figure 2 a). This distribution is build up from the anomalies detected in the historical data by means of an anomaly detection algorithm. The physics-inspired health indicator, which is independent of the operating conditions, is used to detect the moment of anomaly in the vibration measurements. Once the value of the health indicator exceeds the threshold determined by the moving average of the HI and its standard deviation according to equation 6, the point of anomaly detection is defined. Applying this algorithm on historical data delivers the required distribution for the statistical model RUL_{50}, while detecting an anomaly in new testing data, delivers the moment the sequential hybrid method switches from the statistical model RUL_{50} to the data driven RUL prediction.

$$threshold(t) = SMA_{t,k} + 3 * MSTD_{t,k} \tag{5}$$

$$\text{where: } SMA_{t,k} = \frac{1}{k} \sum_{i=t-k+1}^{t} HI_i$$

$$MSTD_{t,m} = \sqrt{\frac{1}{m} \sum_{i=t-m+1}^{t} (HI_i - \mu)} \tag{6}$$

$$\text{and: } \mu = SMA_{t,m}$$

2.2.4. Bayesian filtering and degradation modelling

In Bayesian filtering, an empirical degradation model is used together with online measurements to estimate the unknown parameters of the model. Process noise (i.e., the inaccuracy of the corrosion model) and measurement noise (i.e., the inaccuracy of the measurements) are both taken into account in Bayesian filtering (Brijder, Helsen, & Ompusunggu, 2023). In cases where the state-transition model and the observation model are non-linear transformations, a Kalman filter is not applicable. In this case, while still assuming normally distributed measurement and process noise, adaptions of Kalman filtering can be used, such as extended Kalman filtering (EKF) or unscented Kalman filtering (UKF) (Wan & Van Der Merwe, 2000).

Various degradation models are published in literature (Gebraeel, Lawley, Li, & Ryan, 2005) and in this paper, the most common exponential growth degradation model according to equation 7 is used as state-transition model. The model parameters are updated with new measurement data as can be seen in Figure 5 where an example is given of the UKF HI resulting from filtering the AE reconstruction error. More information on the applied filtering process can be found in (Brijder et al., 2023) and (Gebraeel, 2006).

$$x_t = x_0(1 + r)^t \tag{7}$$

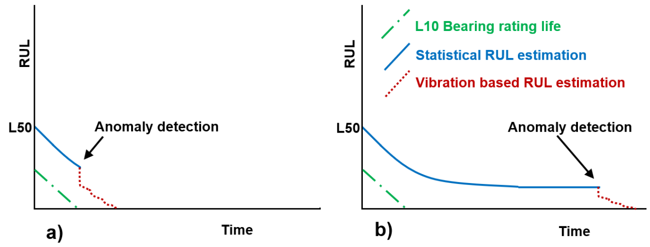

Figure 6. Examples results of the sequential hybrid method for a) short living bearing, and b) long living bearing

2.3. Sequential hybrid RUL prediction

To give an accurate RUL prediction over the entire lifetime of the component, the data-driven HI are computed for each measurement as this HI is used both for anomaly detection and for RUL prediction after the anomaly is detected. The statistical model-based RUL prediction is computed untill the anomaly is detected, and the RUL_{50} prediction is substituted with the data-driven prediction. As shown in Figure 1, a switch from the statistical RUL prediction to the Bayesian filtering of the HI happens once an anomaly is detected in the vibration measurements, indicating the onset of the degradation by pitting or spalling. This switch makes it possible to use to most accurate method available for the current health status. With the availability of more data and the development of early anomaly detection algorithms and accurate RUL prediction techniques, the sequential hybrid method can be improved without the approach being replaced. As mentioned before, the sequential hybrid method can make use of different RUL prediction techniques after the anomaly detection, and for the purpose of this paper both a PCA-based HI (SHM-PCA) and an AE-based HI (SHM-AE) are examined, with the main focus on the AE, as it is deemed more relevant for industrial applications due to it only needing healthy measurement data for training, and very limited faulty data, but close to the End-of-Life, for threshold determination.

Figure 6 schematically shows how applying the sequential hybrid method with the hard switch behaves when an anomaly in the vibrations is detected a) early during the bearing lifetime, or b) even beyond the maximum historical bearing lifetime, when the statistical RUL is converged to the expected degradation time from anomaly to EoL, which is the best estimate for RUL available at that time.

The combination of the model-based statistical RUL prediction and the feature-based data-driven RUL estimation in the proposed SHM-AE approach has multiple benefits:

- The model-based approach can deliver a RUL prediction before the vibration measurements show any signs of the degradation process.
- The use of specific selected statistical and fault frequency features delivers a physically interpretable input for the next steps in the RUL prediction process.
- An AE to reconstruct these vibration features is able to capture non-linear effects that may be present in the physical process and the measurement data.
- The AE gives the possibility the examine the individual reconstructed features and perform diagnostics without adding computational effort.
- The application of Bayesian filtering and a relevant degradation model can smooth out some non-stationary behaviour and gives the opportunity to extrapolate the degradation model to determine the RUL.

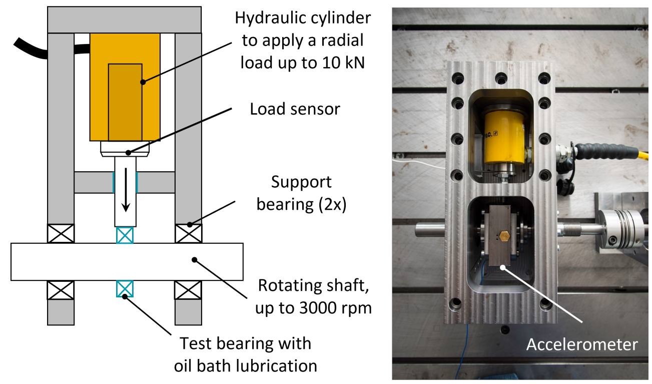

Figure 7. An experimental test rig setup designed to perform accelerated life tests of bearings.

3. FLANDERS MAKE BEARING ACCELERATED LIFE DATA

The bearing datasets used in this study are collected in Flanders Make's Smart Maintenance Living Lab (Ooijevaar et al., 2019). This lab is developed as an open test and development platform and aims to support the adoption of condition monitoring technologies in the industry. The lab consists of seven identical drive train sub-systems. The setups are designed to perform accelerated lifetime testing of bearings and run bearings to their end-of-life. The accelerated lifetime test allows to create surface fatigue faults in bearings and monitor the fault evolution and accumulation during the (accelerated) life.

One of these experimental set-ups to perform the accelerated lifetime test is shown in Figure 7. The set-up comprises of a single shaft with a test bearing. The shaft is supported by a support bearing on each side. A hydraulic cylinder is used to apply a radial load to the test bearing up to a maximum of 10 kN. The test bearing is lubricated by an internal oil bath. The set-up is driven by a motor at a rotation speed up to 3000 rpm. Each set-up is equipped with an accelerometer, temperature sensor, load sensor and speed sensor. The radial accelerations are measured at a sampling frequency of 50 kHz by an accelerometer attached to the bearing housing.

Name	Description	Rotational speed	Radial load	Tests
FMAKE-SMM-1	Single stationary speed and load	2000 rpm	9.0 kN	49
FMAKE-SMM-3	Stepwise varying speed	1000 - 2000 rpm, steps of 100 rpm	9.0 kN	20

Table 1. Overview of the bearing accelerated life tests performed until end-of-life.

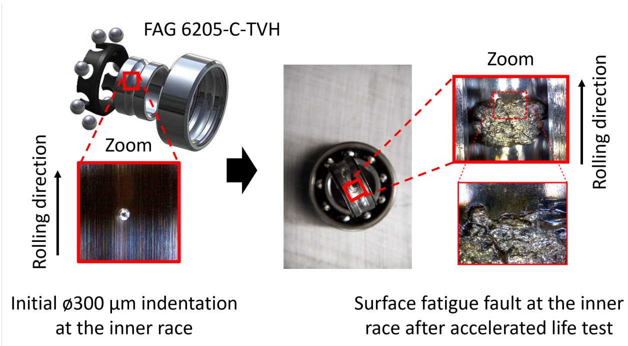

Figure 8. The initial state, a 300 μm, and the final state, a surface fatigue fault at the inner race of the bearing introduced by an accelerated life test.

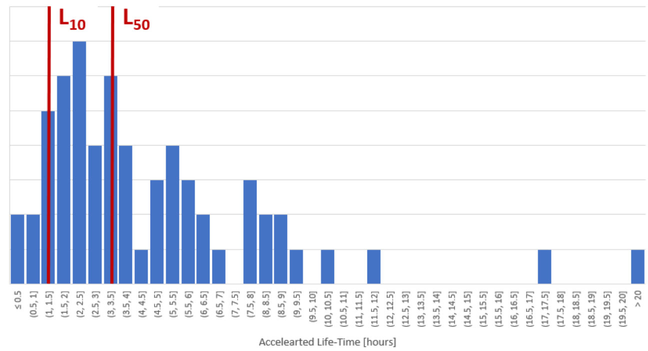

Figure 9. Histogram of the available accelerated lifetime experimental data FMAKE-SMM-1 to determine average expected lifetime L_{50}

The rotational speed and radial load of each set-up can be controlled, such that each set-up can operate at stationary and non-stationary operating conditions. An industrial Beckhoff control platform is used to acquire and store the sensor signals and to control the speed and load of each set-up.

In total more than 80 bearing accelerated life tests are performed on a FAG 6205-C-TVH deep groove ball bearing under 3 different settings in operating conditions and resulting in surface fatigue faults at the inner race, of which 2 are relevant for ths paper as summarized in Table 1. Two acceleration mechanisms are used to accelerate the bearing lifetime:

- A high radial load up to 9 kN (C/P = 1.6) applied to the bearing outer ring.
- Before the start of the test a small initial indentation of approximately 300 μm was created in the bearing inner race using a Rockwell C hardness tester. This indentation is used as a local stress riser and represents a local plastic deformation caused by, for instance, a contamination particle.

The accelerated life time tests are stopped as soon as 20g peak-to-peak accelerations are reached, resulting in severe rolling contact surface fatigue at the inner race (Halme & Andersson, 2009). The start and end condition of the inner race of one of the test bearing are shown in Figure 8.

As the performed lifetime experiments make use of an indentation as stress riser and accelerator of the lifetime of the bearing, the standard bearing rating life L_{10} cannot be computed and is therefore deduced from the histogram of the experimental results of the complete FMAKE-SMM-1 accelerated dataset as shown in Figure 9. Both the (accelerated) bearing rated life L_{10} and the average expected lifetime L_{50} for the remainder of this paper are determined at 1 hour 15 minutes and 3 hours 15 minutes respectively for constant rotational speed of 2000 RPM. With experimental data that contains varying rotational speeds, these values are compensated for by the averaged applied RPM.

4. RESULTS, DISCUSSION AND VALIDATION

In the following sections, the SHM-PCA sequential hybrid method using the PCA HI, the SHM-AE sequential hybrid method using the AE HI and a purely statistical RUL_{50} prediction are compared and discussed. The methods are applied and evaluated for both stationary and varying operating conditions.

4.1. Stationary speed conditions

Two bearings are selected from the FMAKE-SMM-1 dataset described in table 1 in order to show the RUL prediction of a bearing that reaches EoL before the average statistical lifetime L_{50} of the entire dataset (bearing A22) and a bearing that survives almost twice the average lifetime (bearing A81). The stationary operating conditions of this dataset fall within the range of operating conditions from the varying operating conditions datsets on which the AE is trained, but clearly positioned at the boundary of the operation conditions.

The results of the RUL prediction for bearings under stationary operating conditions are shown in Figures 10 and 11. The linear decreasing red line represents the true RUL for the specific bearing and is only known after the test. The EoL for the FMAKE-SMM accelerated lifetime datasets are reached when the vibrations reach a level of 20g peak-to-peak.

At the start (t = 0 hours) the estimated RUL equals the average

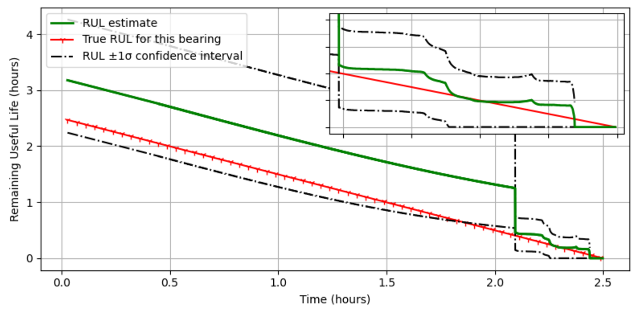

Figure 10. Sequential hybrid RUL prediction for test bearing A22 under stationary operating conditions

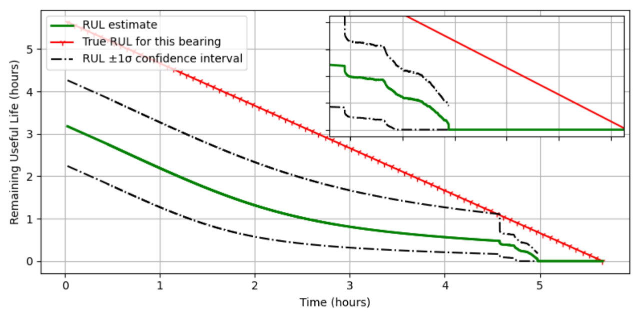

Figure 11. Sequential hybrid RUL prediction for test bearing A81 under stationary operating conditions

expected lifetime L_{50}. This statistical estimate of the RUL decreases with time and starts to deviate slightly from the linear decrease, and converges towards the RUL value representing the average time between the detection of an anomaly and the moment that End-of-Life is reached as seen in Figure2b. However, once an anomaly is detected, and the sequential hybrid method switches from statistical model-based RUL_{50} prediction to the data-driven based approach, the RUL becomes more accurate as the RUL prediction gets closer to the true RUL. Figure 10 displays the result of the sequential hybrid RUL prediction for bearing A22 whose lifetime is less than the average lifetime in the dataset, clearly seen from the statistical estimation being higher than the true RUL before the anomaly detection from the vibration measurements. On the top right, a close-up of the RUL prediction after the anomaly detection is given to visualize the RUL prediction using the data-driven HI. As a second example, Figure 11 shows the result for bearing A81 whose lifetime is 1.5 times the average expected lifetime, high-lighting the benefit of the accurate RUL prediction, as this bearings useful life is almost 5 times the bearing rating life L_{10} of 1 hour 15 minutes, which is often used as maintenance interval when following periodic maintenance strategies.

Table 2 and 3 give an overview of the classical metrics known as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) (Berghout & Benbouzid, 2022) used to compare

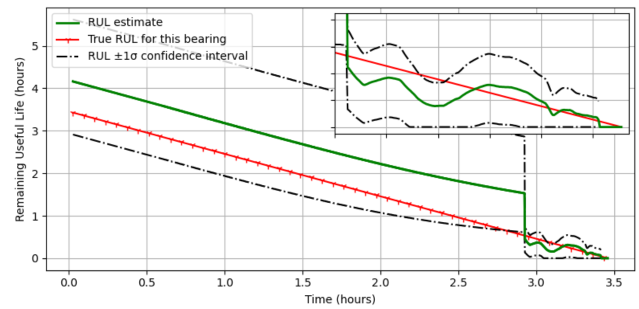

Figure 12. RUL prediction by sequential hybrid method based on an AE based HI for test bearing A150 under varying speed conditions

the performance of the 3 the different approaches. It can be seen from these tables that the sequential hybrid method, both with a PCA-based HI (SHM-PCA) and an AE-based data-driven HI (SHM-AE) show similar results for the stationary operating conditions.

It should be noted however, that the determination of the PCA components requires historical data from the faulty state of the components, whereas the proposed sequential hybrid method using a data-driven AE for fusing vibration features into a health indicator requires only limited data from the healthy state of the component and some faulty data points near the end-of-life of the component to determine the threshold for the AE HI that indicates failure of the bearing.

4.2. Varying speed conditions

The benefit of the proposed sequential hybrid RUL prediction method, using the AE as feature fusion method, is more clearly shown when applied to bearings under varying operating conditions. The current available data consists of non-stationary rotational speeds and constant load. The results of the sequential hybrid method on three bearings of the FMAKE-SMM-3 dataset are shown in Figures 12, 13 and 14. Bearing A150 in Figure 12 reaches its EoL before the expected average lifetime L_{50}, and the hard switch from the statistical model-based RUL estimation to the data-driven RUL_{50} prediction clearly increases the accuracy at the later stages of the bearings lifetime. The oscillations visible during the final degradation stage are a result of the speed variations, which influence the RUL. Logically, a slower rotating bearing can survive longer because of the fact that the rotations left in the bearings lifetime are spread out over a longer time.

The lifetime of bearing A155, shown in Figure 13, falls nicely in the expected spread for this bearing under the applied load and speed conditions as can be seen by the had switch taking place before the statistical RUL_{50} prediction converges to a horizontal level. Similar to the previous bearing, the switch to the data-driven RUL prediction results in a RUL prediction

RUL Method	A22	A81	A150	A155	A156
	Constant operating conditions		Varying operating conditions		
RUL_{50}	0.732045	1.915349	0.799972	1.734906	5.874663
SHM-PCA	0.661791	1.850988	0.714877	1.677779	5.75963
SHM-AE	0.661932	1.852214	0.713457	1.677557	5.759003

Table 2. RMSE over the entire lifetime of the bearings subjected to accelerated life tests until end-of-life.

RUL Method	A22	A81	A150	A155	A156
	Constant operating conditions		Varying operating conditions		
RUL_{50}	0.73025	1.734932	0.793397	1.564567	5.252493
SHM-PCA	0.611975	1.66768	0.669223	1.475042	5.090453
SHM-AE	0.613295	1.671588	0.664556	1.475999	5.090577

Table 3. MAE over the entire lifetime of the bearings subjected to accelerated life tests until end-of-life.

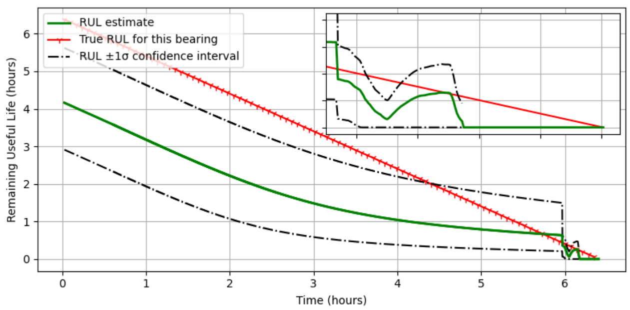

Figure 13. RUL prediction by sequential hybrid method based on an AE based HI for test bearing A155 under varying speed conditions

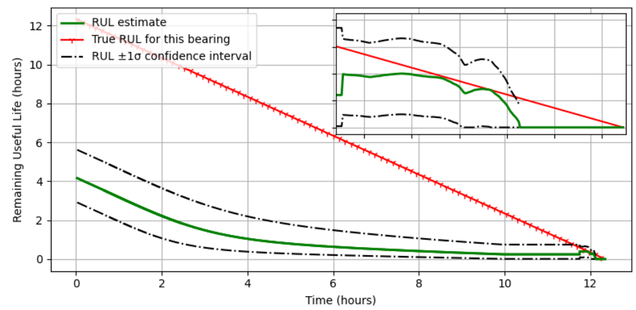

Figure 14. RUL prediction by sequential hybrid method based on an AE based HI for test bearing A156 under varying speed conditions

that is closer to the true RUL as can be seen in the close-up of Figure 13.

Finally, bearing A156 in Figure 14 survives significantly longer than any of the bearings in the training datasets, but the sequential hybrid method is still able to give a useful RUL prediction for planning its replacement. Initially the predicted RUL is significantly underestimated, but converges to the true RUL after the anomaly is detected. Even if the bearing would remain in use long after the statistically expected lifetime, the switch to the data-driven HI delivers a RUL prediction close to the true RUL after anomaly detection, seen by the increase in the RUL resulting from the data-driven HI.

In order to give an quantitative comparison, two reference methods that are applicable for the entire lifetime of the bearing are defined. First of all, the use of the statistical RUL_{50} prediction based on the expected average lifetime L_{50} and the expected uncertainty distribution without any vibration information as shown in Figure 2b, and secondly the same sequential hybrid method but with a PCA-based HI in stead of the AE-based HI. The results of the comparison can be found in tables 2 and 3, showing the RMSE and MAE computed over the entire lifetime of the bearing. Similarly to the results for the constant operating conditions, both the RMSE and MAE show similar results for the PCA-based and AE-based

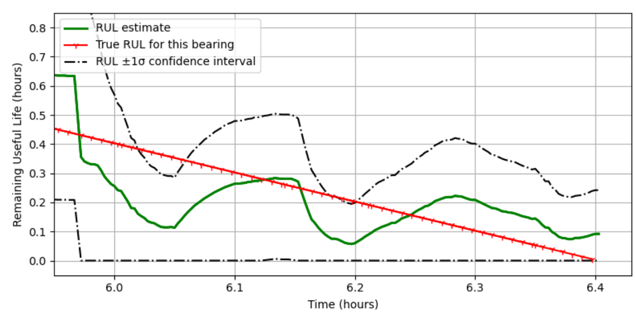

Figure 15. Sequential hybrid RUL prediction for long living test bearing A155 applying PCA health indicator after anomaly detection

sequential hybrid method, while outperforming the statistical RUL_{50} prediction.

However, when looking at the close-up of both the PCA result in Figure 15 and the AE result in Figure 16 during the degradation of bearing A155, it can be seen that the PCA HI does not cross the set threshold, and does not estimate the EoL of the bearing before the actual failure during the experiment.

Overestimation is very costly in the application of RUL methods in industrial applications, as they can lead to unplanned stops of production lines or machines. In order to punish

RUL Method	A22	A81	A150	A155	A156
	Constant operating conditions		Varying operating conditions		
RUL_{50}	4.368598	0.82255	11.2866	2.27442	0.919807
SHM-PCA	0.605474	0.658486	0.958878	0.894308	0.823996
SHM-AE	0.614277	0.66341	0.494484	0.469929	0.828339

Table 4. MAPE of the entire lifetime of the bearings subjected to accelerated life tests until end-of-life.

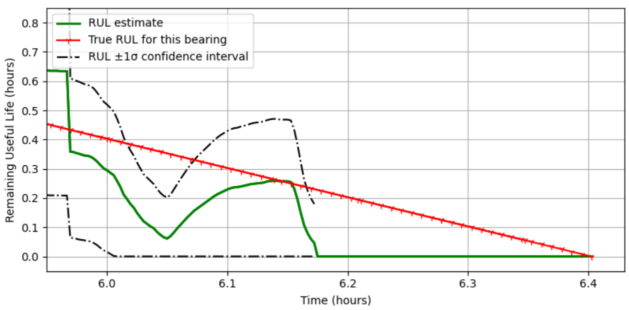

Figure 16. Sequential hybrid RUL prediction for long living test bearing A155 applying AE health indicator after anomaly detection

any overestimation in the later stages of the components lifetime, more weight can be put on the error towards EoL where more accuracy is required. Computing the comparison metric Mean Absolute Percentage Error (MAPE) (Berghout & Benbouzid, 2022) (Kamat et al., 2021) in equation 8 automatically delivers this, as the true RUL is a linear decreasing value and using this true RUL to compute the error percentage will lead to a higher importance of errors near the EoL. This make the MAPE very applicable for RUL prediction method comparison.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{EoL} \left| \frac{RUL_t - RUL_{est}}{RUL_t} \right| \tag{8}$$

Comparing the previous tables to table 4 gives a better overview of the accuracy of the sequential hybrid method. As expected, for stationary operating conditions, both the PCA method as well as the AE method perform well, however for non-stationary speeds and bearings that have lifetimes within the expected distribution the AE clearly outperforms PCA as feature fusion method.

Further observations from the MAPE comparison show large values when there is a large overestimation near the EoL for short living bearings A22 and A150. When MAPE is around or below a value of one, the estimation can be considered to be accurate, as values above one can only result from large overestimation, mainly near EoL.

5. CONCLUSION

The proposed sequential hybrid method using the reconstruction error of the autoencoder as health indicator only requires limited healthy and even less faulty measurements to be developed, which is a large benefit for industrial applications. Even with the limited training data, the autoencoder trained on healthy data delivers a health indicator that can cope with varying operating conditions. By combining the data-driven RUL prediction based on the AE for the later stages of the bearings lifetime with a statistical model-based RUL_{50} before any anomaly is detected in vibration measurements, the sequential hybrid method delivers a RUL prediction during the entire lifetime of the bearing which increases the effectiveness of maintenance planning for a large fleet of rotating machinery.

When comparing the different RUL prediction methods over the entire lifetime of the bearing, it is shown that computing the MAPE is more useful than the computation of the RMSE and MAE, because more accuracy is required the closer the component gets to its EoL. By comparing the MAPE between a pure statistical model-based RUL_{50} prediction based on the expected average lifetime L_{50}, and the sequential hybrid method both with PCA HI and AE HI, it is clearly shown that the sequential hybrid method delivers more accurate results by switching to the data-driven HI once an anomaly is detected in the vibration measurements. Furthermore, training an AE to generate the data-driven HI is more promising and delivers less overestimation near the EoL than applying the linear combination of vibration features in the PCA HI for applications with varying speeds.

When more data is available, better anomaly detection algorithms are developed, or better performing and monotonic health indicators see the light-of-day, they can easily be deployed within the sequential hybrid method framework to deliver even more accurate RUL predictions. The proposed method is developed for the industrial use-case of bearings in rotating machinery, but by using different statistical methods and health indicators for different applications, the sequential hybrid method can also be deployed for RUL predictions of other components or machines.

.

ACKNOWLEDGMENT

This research was supported by Flanders Make, the strategic research centre for the manufacturing industry, and more precisely the Smart Maintenance Proeftuin project and the DGTwinPrediction SBO research project.

REFERENCES

Aydemir, G., & Acar, B. (2020). Anomaly monitoring improves remaining useful life estimation of industrial machinery. *Journal of Manufacturing Systems*, *56*, 463-469.

Berghout, T., & Benbouzid, M. (2022). A systematic guide for predicting remaining useful life with machine learning. *Electronics*, *11*(7).

Bourgana, T., Brijder, R., Ooijevaar, T., & Ompusunggu, A. (2021). Wavelet scattering network based bearing fault detection. PHM Society European Conference.

Brijder, R., Helsen, S., & Ompusunggu, A. (2023). Switching kalman filtering-based corrosion detection and prognostics for offshore wind-turbine structures. *Wind*, *3*.

Dong, S., & Luo, T. (2013). Bearing degradation process prediction based on the pca and optimized ls-svm model. *Measurement*, *46*(9), 3143-3152.

Ferreira, C., & Gonçalves, G. (2022). Remaining useful life prediction and challenges: A literature review on the use of machine learning methods. *Journal of Manufacturing Systems*, *63*, 550-562.

Gebraeel, N. (2006). Sensory-updated residual life distributions for components with exponential degradation patterns. *IEEE Transactions on Automation Science and Engineering*, *3*(4), 382-393.

Gebraeel, N., Lawley, M., Li, R., & Ryan, J. (2005, 06). Residual-life distribution from component degradation signals: A bayesian approach. *IIE Transactions*, *37*, 543-557.

Halme, J., & Andersson, P. (2009). Rolling contact fatigue and wear fundamentals for rolling bearing diagnostics - state of the art. *Journal of Engineering Tribology*, *224*, 377–393.

Harris, T. A., & Yu, W. K. (1999, 01). Lundberg-palmgren fatigue theory: Considerations of failure stress and stressed volume. *Journal of Tribology*, *121*(1), 85-89. doi: 10.1115/1.2833815

Hervé de Beaulieu, M., Shekhar Jha, M., Garnier, H., & Cerbah, F. (2022). Unsupervised prognostics based on deep virtual health index prediction. In *Proceedings of the european conference of the phm society 2022* (Vol. 7, p. 193–199).

ISO281. (2007). *Rolling bearings – Dynamic load ratings and rating life* (Standard). ISO: International Organization for Standardization.

Jammu, N., & Kankar, P. (2011, 10). A review on prognosis of rolling element bearings. *International Journal of Engineering Science and Technology*, *3*.

Kamat, P., Sugandhi, R., & Kumar, S. (2021). Deep learning-based anomaly-onset aware remaining useful life estimation of bearings. *PeerJ Computer Science*.

Kan, M. S., Tan, A. C., & Mathew, J. (2015). A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mechanical Systems and Signal Processing*, *62-63*, 1-20.

Liao, L., & Köttig, F. (2014). Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Transactions on Reliability*, *63*(1), 191-207.

Mrugalska, B. (2019). Remaining useful life as prognostic approach: A review. In (pp. 689–695). Springer International Publishing.

NSWC, N. S. W. C. (2011). *Handbook of reliability prediction procedures for mechanical equipment*. West Bethesda, Maryland.

Ooijevaar, T., Pichler, K., Di, Y., Devos, S., Volckaert, B., Van Hoecke, S., & Hesch, C. (2019). Smart machine maintenance enabled by a condition monitoring living lab. *IFAC-PapersOnLine*, *52*(15), 376-381.

SKF. (2011). *SKF bearing maintenance handbook*.

Wan, E., & Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Proceedings of the ieee 2000 adaptive systems for signal processing, communications, and control symposium* (p. 153-158).

Zhao, S., Zhang, Y., Wang, S., Zhou, B., & Cheng, C. (2019). A recurrent neural network approach for remaining useful life prediction utilizing a novel trend features construction method. *Measurement*, *146*, 279-288.