Remaining Useful Lifetime Estimation of Bearings Operating under Time-Varying Conditions

Alireza Javanmardi[1,2], Osarenren Kennedy Aimiyekagbon[3], Amelie Bender[3], James Kuria Kimotho[4], Walter Sextro[3], and Eyke Hüllermeier[1,2]

[1] *Institute of Informatics, LMU Munich, Munich, Germany*

[2] *Munich Center for Machine Learning (MCML), Munich, Germany*

[3] *Chair of Dynamics and Mechatronics, Paderborn University, Paderborn, Germany*

[4] *Department of Mechanical Engineering, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya*

ABSTRACT

This paper investigates the remaining useful lifetime (RUL) estimation of bearings under dynamic, i.e., time-varying, operating conditions (OC). Unlike conventional studies that assume constant OC in bearing accelerated life tests, we introduce a dataset with time-varying OC during run-to-failure experiments, simulating real-world scenarios. We explore data-driven approaches to identify the transition point from a healthy to an unhealthy state and estimate the RUL. Additionally, we examine strategies for integrating OC information to enhance RUL estimations. These methodologies are evaluated through numerical experiments using various machine learning algorithms.

1. INTRODUCTION

Rolling element bearings are extensively used in industrial applications, such as wind turbines, electric motors, and generators. These bearings account for the largest percentage of failures in rotating machinery (Alewine & Chen, 2010) and about 40%-50% of all motor faults (Sharma et al., 2015). Failure of bearings may result in expensive downtime, increased maintenance costs due to failures propagating to other parts, and catastrophic effects if they support critical equipment. Predictive maintenance can be employed to increase the efficiency and reliability of bearings and technical systems in general, as it prevents unexpected failures and maximizes their availability. Predictive maintenance builds on prognostics, which involves the accurate estimation of the remaining useful lifetime (RUL) of technical systems or components, such as bearings.

For developing RUL estimation methods for bearings, existing datasets often focus on accelerated life tests conducted under constant operating conditions (OC) (Lee et al., 2007). In some cases where varying OC were taken into account, the OC only change between different run-to-failure experiments but remain constant within each experiment (Nectoux et al., 2012; Wang et al., 2018). However, this approach falls short of accurately simulating real-world scenarios where bearings may experience time-varying conditions throughout their operational life. Some research has been conducted on time-varying conditions. For example, Du et al. (2022) propose extracting features from the angular domain and RUL prediction based on the unscented particle filter. However, their proposed methodology was evaluated on ball bearing run-to-failure experiments, considering only varying rotating speed in the range [1450, 1550] rpm. Furthermore, the time point of degradation onset was manually determined. N. Li et al. (2019) propose a so-called "two-factor" state-space model based on a Wiener process, where the underlying degradation process is modeled in the state transition function, and the influence of the varying condition on the measured signal is captured in the measurement function of the proposed model. However, the proposed methodology builds on the assumption that the OC are known a priori and follow a known pattern. Furthermore, their methodology was evaluated on ball bearings subjected to cyclic varying speed conditions, taking up two speed values, namely 2200 rpm and 2600 rpm.

To address the presented limitation of existing studies and enhance the relevance to practical applications, we introduce a new dataset of bearing run-to-failure experiments, in which OC can dynamically vary over time, such as in a non-periodic and stochastic manner. Thus presenting new challenges for RUL estimation, as the vibration data not only reflects bearing degradation but is also influenced by changes in OC. Table 1 provides an overview of the existing datasets and their comparison to ours.

Table 1. Comparison between the publicly available datasets and our dataset.

	IMS (Lee et al., 2007)	Pronostia (Nectoux et al., 2012)	XJTU-SY (Wang et al., 2018)	LDM (Aimiyekagbon, 2024)
Bearing type	Roller bearing	Ball bearing	Ball bearing	Ball bearing
Rotating speed [rpm]	2000	1500, 1650, 1800	2100, 2250, 2400	$[1500, 3600]$
Static load [kN]	26.7	4, 4.2, 5	10, 11, 12	$[1.5, 4]$
Dynamic load amplitude [kN]	✗	✗	✗	$[0.5, 1.7]$
Dynamic load type	✗	✗	✗	Sinusoidal and Gaussian noise

Data-driven techniques for estimating the RUL involve establishing a mapping between available information, mainly vibration data, and the RUL. Typically, the initial step is to extract features from vibration data, given its high-dimensional nature and lack of a discernible trend for RUL estimation. Features are typically extracted in the time-, frequency-, and time-frequency-domain. While finding the best feature representation lies beyond the scope of this paper, we primarily adopt the Fast Fourier Transform (FFT) to obtain a frequency-domain representation of the vibration signal. Furthermore, to address the challenge of high dimensionality, we employ a methodology akin to that proposed in (Ren et al., 2018; von Hahn & Mechefske, 2022). This involves segmenting the resulting FFT signal into distinct frequency buckets and subsequently identifying the maximum value within each bucket.

It has been observed that the behavior of bearings does not exhibit a consistent trend from the beginning to the failure time. Instead, a typical scenario involves an initial phase of normal behavior followed by an abrupt shift at some point during the lifespan, indicating the initiation of degradation. These points, marking the transition from a healthy to an unhealthy state, are referred to as transition times. While existing approaches use various engineering techniques to detect these transition times (X. Li et al., 2019), this study employs a 2-means clustering technique on extracted features to define the transition time as the moment when the cluster of a bearing changes with respect to its initial cluster. After identifying the transition time, the subsequent data points can be used to train a RUL estimator model.

The dataset consisting of all features after the transition times, along with their corresponding RUL labels, can be fed into any supervised machine learning or deep learning model for fitting an RUL estimator. The challenge of estimating the RUL under dynamic OC is addressed through various approaches in the literature. Huang et al. (2019) incorporate the OC as an additional input in their deep network model. Fu et al. (2021) and Javanmardi & Hüllermeier (2023) suggest normalizing data according to OC. F. Li et al. (2020) integrate several algorithms into one model and select an optimal algorithm set for different OC to minimize their impact. Numerous studies address this problem by employing transfer learning or domain adaptation to handle the distribution shift between the training (source) and testing (target) domains (Mao et al., 2019; Fan et al., 2020; da Costa et al., 2020; Ding, Jia, Miao, & Huang, 2021; Ding, Jia, & Cao, 2021; Zhang et al., 2021). Ding et al. (2022) consider multi-source adaptation to manage the presence of subdomains in the source caused by multiple OC. To this end, we consider three distinct approaches in this study:

- Firstly, we train a regressor using only the previously attained features without taking the OC into account. This approach serves as a baseline for the subsequent two methods.
- Secondly, we employ the OC to normalize the features, aiming to mitigate its impact on the overall feature set.
- Thirdly, we concatenate the OC with the previously attained features, thereby incorporating them as additional features.

In the following sections, we first formalize the problem statement along with the details of all steps, from feature extraction to transition time determination and RUL estimation. Later, we elaborate on the data generation process and present comprehensive numerical results for the proposed approaches.

2. Problem Statement

Consider a dataset containing N instances of bearing run-to-failure data. Each bearing i in the dataset with a lifetime of T_i is represented as a time series $\mathbf{Z}_i := \left\{z_1^{(i)}, z_2^{(i)}, \ldots, z_{T_i}^{(i)}\right\}$. Here, $z_t^{(i)} := \left(v_t^{(i)}, o_t^{(i)}\right)$, where $o_t^{(i)} \in \mathbb{R}^{d_o}$ contains information about the operating and environmental conditions during the t^{th} measurement cycle, and $v_t^{(i)} \in \mathbb{R}^{d_v}$ represents the vibration signal collected during that measurement. For all $t \in [T_i] := \{1, \ldots, T_i\}$, the RUL $y_t^{(i)}$ of instance i at time t can be computed as follows:

$$y_t^{(i)} = T_i - t\,. \tag{1}$$

2.1. Feature Extraction from the Vibration Data

The vibration signal in the time domain $v_t^{(i)}$ is often high-dimensional, making it unsuitable for direct integration into a machine learning framework. In this study, we employ discrete Fourier transform to convert the signal into its frequency spectrum. This transformation results in $V_t^{(i)}$, a signal with the same dimensionality as the original time signal. Next, we partition the signal into m equally sized buckets $B_1, \ldots, B_m$ (with B_1 corresponding to the lowest frequency bucket and B_m to the highest) and simply extract the maximum ampli-

tude within each bucket to construct the m-dimensional frequency domain features $X_t^{(i)}$, i.e.,

$$X_t^{(i)} = \left(\max_{j_1 \in B_1} V_t^{(i)}(j_1), \ldots, \max_{j_m \in B_m} V_t^{(i)}(j_m) \right), \quad (2)$$

where $V_t^{(i)}(k)$ represents the k^{th} component of the signal $V_t^{(i)}$. Having access to this new feature, in the literature also known as the Spectrum-Principal-Energy-Vector (Ren et al., 2018), resolves the challenge posed by the high dimensionality of the initial vibration signal.

2.2. Transition Time Determination

K-means clustering is an unsupervised machine learning algorithm that clusters similar data points based on their proximity in the feature space. The algorithm initializes K cluster centroids and assigns each data point to the nearest centroid, recalculating the centroid of each cluster based on the mean of the assigned data points until convergence. The goal is to minimize the sum of squared distances between each data point and its assigned centroid. Here, we merely want to divide data points into a healthy or unhealthy cluster, thus K=2. We assume that each bearing starts in a healthy state, and hence, the cluster of the first point is considered healthy. A change in the cluster in the subsequent times is considered the beginning of the degradation. Once trained on the training data, the algorithm can be used in an online fashion for each test data instance to detect its changepoint promptly and initiate RUL prediction.

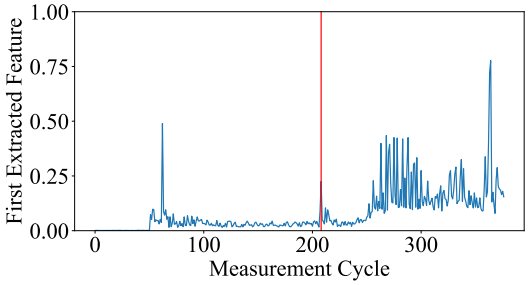
Following the extraction of low-dimensional features $X_t^{(i)}$ from the vibration data, we can utilize a 2-means clustering algorithm to assign a cluster label $\delta_t^{(i)} \in \{0, 1\}$ to each measurement time t for every bearing i. Subsequently, we define $t_{\text{TT}}^{(i)}$, the *transition time*, as the moment when the cluster of the i^{th} bearing differs from its initial cluster. Formally, this is expressed as

$$t_{\text{TT}}^{(i)} = \min \left\{ t : t \in [1 + T_c, T_i] \text{ and } \delta_t^{(i)} \neq \delta_1^{(i)} \right\}, \quad (3)$$

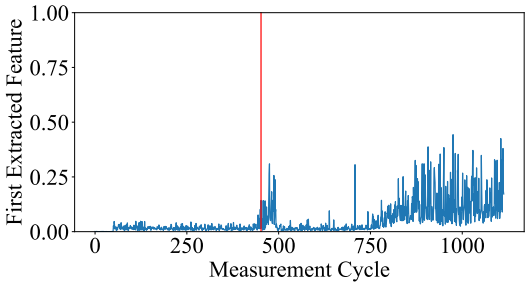
where T_c serves as a hyperparameter, representing the tolerance level. It signifies that a change in the cluster occurring earlier than T_c is not considered in the transition time calculation. The transition times for two bearing experiments are exemplarily depicted in Figure 1. Once the transition times are determined, we can define healthy and unhealthy datasets as follows:

$$\mathcal{D}_{\text{healthy}} = \left\{ (X_t^{(i)}, o_t^{(i)}, y_t^{(i)}) : i \in [N], t < t_{\text{TT}}^{(i)} \right\}, \quad (4)$$

$$\mathcal{D}_{\text{unhealthy}} = \left\{ (X_t^{(i)}, o_t^{(i)}, y_t^{(i)}) : i \in [N], t \geq t_{\text{TT}}^{(i)} \right\}. \quad (5)$$



(a) B01



(b) B04

Figure 1. The first extracted feature (blue) plotted for two bearings alongside their determined transition times (red). Here T_c is set to 150.

2.3. RUL Estimation

After extracting features from the vibration data and determining transition times, the next step is to estimate the RUL of the bearing. The primary focus of this paper is to leverage machine learning algorithms for that purpose. From a machine learning perspective, the problem is framed as a supervised regression setting—finding a mapping from the feature space to the RUL space. However, we have yet to explore how to benefit from OC information. In this context, we consider three distinct scenarios as outlined below and depicted in the flowchart in Figure 2.

- **Scenario 1 (disregarding OC):** In this scenario, OC information is neglected, and training proceeds without considering such contextual data.
- **Scenario 2 (OC for feature scaling):** This approach involves utilizing OC information for data/feature normalization. The methodology employs PCA to reduce the dimensionality of OC data from d_o to 1. Subsequently, a uniform discretization method is applied to bin the resulting one-dimensional feature into B bins. Next, inspired by a prior study (Javanmardi & Hüllermeier, 2023), the data in each bin is normalized to the $[0, 1]$ interval using B distinct MinMax scalers, aiming to mitigate the impact of diverse OC indirectly.
- **Scenario 3 (OC as additional features):** In this method, OC information is treated as an additional set of features, thereby augmenting the feature space. The objective is to

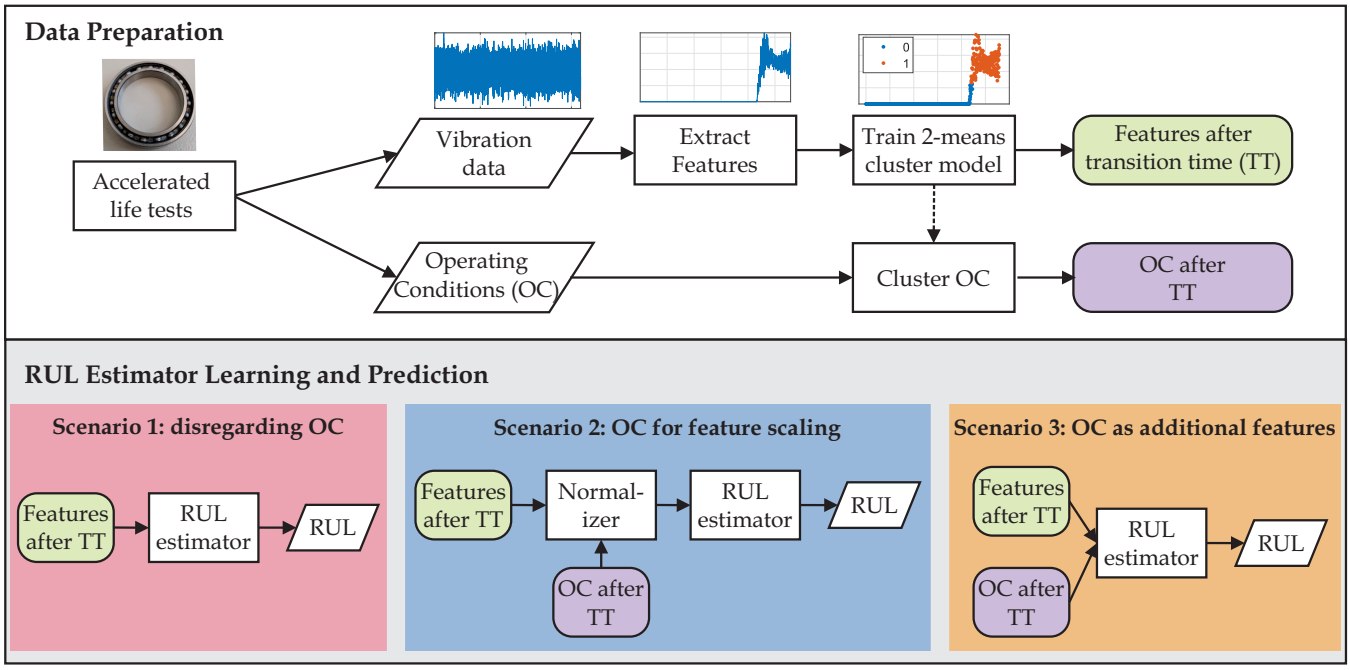


Figure 2. Flow chart of the proposed method.

enable the machine learning model to identify and consider interactions between OC and vibration features during the RUL estimation process.

Note that any machine learning or deep learning model can be used as the underlying RUL estimator for the three proposed scenarios. In this paper, we focus on traditional machine learning models, such as gradient boosting (GB) and random forest (RF).

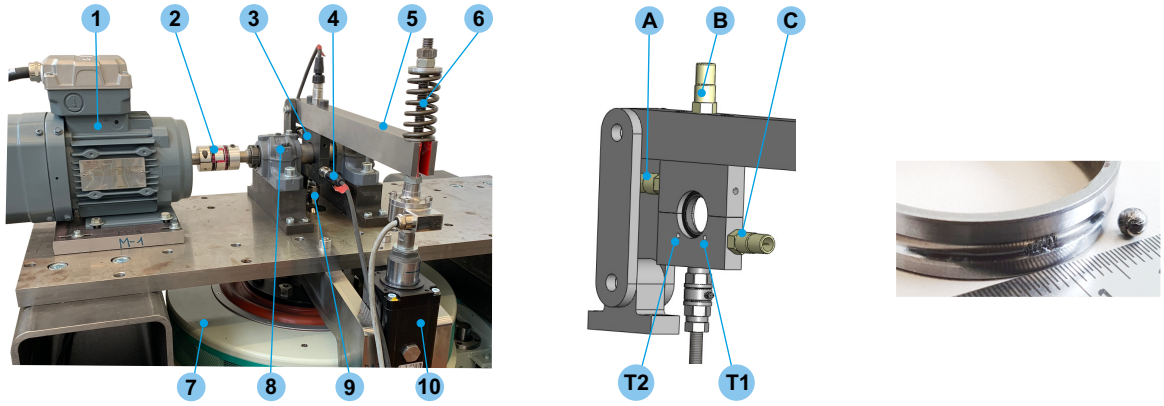
3. Case Study

The experimental dataset, which consists of accelerated life tests of ball bearings subjected to time-varying conditions, is gathered at the Chair of Dynamics and Mechatronics (LDM) at Paderborn University. The specifications of the test bearing allow an experiment with valuable condition monitoring data to take several hours. Specifically, the 61806-2RS rolling element bearing with a basic static load rating $C_0 = 3.15\,\mathrm{kN}$ and a dynamic load rating $C = 4.00\,\mathrm{kN}$ have a basic rating life L_{10} of approximately five hours while considering a constant equivalent load of 4.50 kN, a rotating speed of 2500 rpm and other factors not been considered, such as lubrication.

The bearing test rig with its components is captured in Figure 3(a). The test bearing within its housing (3) is mounted on a shaft. The shaft is coupled with the driving motor (1) via a jaw coupling (2) and supported by two spherical roller bearings (8) within their housing. A static pre-load is exerted on the bearing via a lever structure (5), which is attached to the bearing housing. To this end, the compression spring, mounted on the lever structure, is compressed by the linear actuator (10). A dynamic load is superimposed on the static pre-load by means of an electrodynamic shaker (7), which is connected to the test bearing housing via a stinger (9).

The input signals, namely the exerted forces and shaft rotating speed, are measured synchronously with vibration and temperature as condition monitoring data. Three one-directional accelerometers (4) measure the vibration of the bearing indirectly. Two accelerometers (A and C) measure the vibration horizontally from the housing, and one (B) measures vertically from the lever structure, as illustrated in Figure 3(b). The ambient temperature and bearing temperature are measured with Pt100 resistance thermometers. The bearing temperature is measured indirectly from its housing at the positions (T1 and T2) depicted in Figure 3(b). Measurements were acquired at a sampling duration of 1.6 s and a measurement interval of approximately 12 s. The temperature signals are measured with a sampling rate of 10 Hz. To facilitate high-frequency analysis, vibration data were sampled at 128 kHz for experiments till B09 and due to data storage issues at 64 kHz for experiments from B10. This lower sampling frequency is theoretically sufficient for analysis in the frequency range of interest up to 32 kHz.

During an experiment, the test bearing is subjected to dynamic load superimposed on a static pre-load. To accommodate different dynamic load types, the dynamic load is sinusoidal with a constant frequency of 2 Hz for some experiments. The amplitude of the sinusoidal load is stationary per measurement and takes on a random value from a station-



(a) (b) (c)

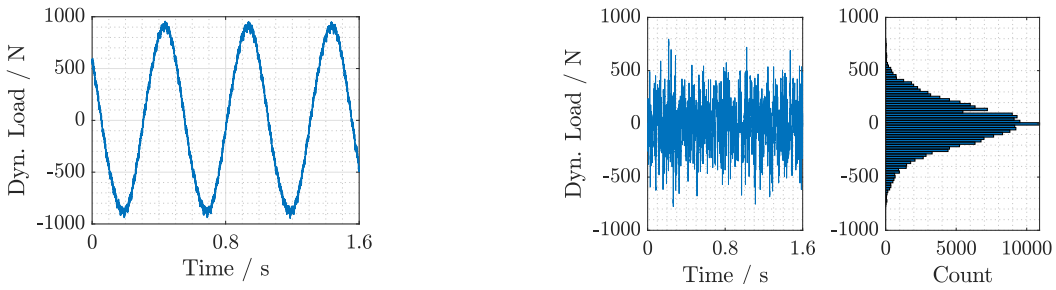
Figure 3: (a) Ball bearing test-rig with the following components: (1) motor, (2) jaw coupling, (3) bearing housing, (4) accelerometers, (5) lever structure, (6) electrodynamic shaker (DFG, 2017), (7) support bearing housing, (8) stinger connected to a quartz force sensor, and (9) linear actuator.
(b) Accelerometer and temperature placement on the test bearing housing without the shaft
(c) A dismantled test bearing with a surface defect on the inner ring raceway and spalls on a rolling element.

ary uniform distribution within a predefined interval between measurements. For other experiments, the dynamic load is Gaussian white noise with a maximum excitation frequency of 200 Hz and truncated to remain within a predefined interval between measurements. A measurement of the dynamic load types is exemplarily shown in Figure 4. The shaft rotating speed is also set to be constant per measurement and takes on a random value from a stationary uniform distribution within a predefined interval between measurements. The predefined range of values per experiment can be found in Table 2.

To avoid failure of other components, except the test bearing, an experiment ends when the test bearing fails. Bearing failure is determined by two predetermined failure threshold criteria. If one threshold is exceeded, the experiment is stopped. On the assumption that the test bearing is in a normal state, without previous loading history, and the ranges of speed and force are restrained, the first criterion builds on the equivalent energy content of the vibration data and is formulated as:

$$F_{vibration} = O \cdot \frac{1}{m} \sum_{t=1}^{m} \text{RMS}\left(v_t^{(i)}(1), \ldots, v_t^{(i)}(n)\right), \quad (6)$$

where $F_{vibration}$ denotes the vibration threshold value, $O = 8$ is a predetermined constant value, m is the number of vibration signals to consider, and $v_t^{(i)}(k)$ represents the k^{th} index of a vibration signal $v_t^{(i)}$ of length n. For the provided experiments, this threshold value lies approximately between 6 g and 10 g. Since improper lubrication of the bearing raceway leads to increased friction and subsequently to increased temperature, the second criterion builds on the bearing temperature. According to the data sheet and to avoid melting the bearing seal made up of nitrile butadiene rubber (NBR), the threshold value, based on the bearing housing temperature $F_{temperature}$, is set as 110 °C.



(a) (b)

Figure 4. Dynamic load: (a) Sinusoidal load and (b) Gaussian white noise.

Table 2. Set range of OC for experiments.

Experiment	Rotating speed [rpm]	Static load [N]	Dynamic load [N]	Dynamic load type
B01 - B03	$[2400, 3000]$	$[3300, 3800]$	$[500, 1400]$	Sinusoidal
B04 - B05	$[1500, 3000]$	$[2500, 3800]$	$[500, 1500]$	Sinusoidal
B06	$[1500, 3600]$	$[3300, 3800]$	$[750, 1700]$	Sinusoidal
B07	$[1500, 3000]$	$[3250, 4000]$	$[250, 750]$	Gaussian noise
B08	$[1500, 3000]$	$[3250, 4000]$	$[500, 1000]$	Gaussian noise
B09	$[1500, 3000]$	$[2500, 3800]$	$[750, 1000]$	Gaussian noise
B10	$[1500, 2700]$	$[2000, 3250]$	$[750, 1500]$	Gaussian noise
B11 - B13	$[1500, 2700]$	3000	1000	Gaussian noise
B14 - B15	2700	2500	$[750, 1500]$	Gaussian noise
B16	2700	2500	$[1000, 1500]$	Gaussian noise
B17	2700	$[1500, 2500]$	1500	Gaussian noise

In this paper, 17 run-to-failure experiments are utilized, with the temperature threshold value $F_{temperature}$ being exceeded for four experiments (B03, B07, B08, and B15) and the vibration threshold value $F_{vibration}$ for others. The failure types are not predetermined, but several single or combined failure types, such as an outer ring raceway defect or rolling element fault combined with an inner ring raceway defect, could occur during an experiment. Figure 3(c) is an image of a dismantled test bearing with spalls on a rolling element and an inner ring raceway defect after an experiment (B06). Due to the proximity to the bearing and for brevity, only the horizontal accelerometer (labeled A) is exemplarily considered in the following analysis. Also, three-dimensional OC information is considered, including peak dynamic load [N], mean absolute static load [N], and mean absolute rotating speed [rpm].

3.1. Numerical Experiments

We set the bucket size for feature extraction at 20. To conduct fair experiments, we repeat the proposed methods 17 times, reserving one bearing for testing each time while using the data of the remaining bearings for training. Without loss of generality, take the j^{th} bearing as the test bearing. The training and test data are defined as follows:

$$\mathcal{D}^{\text{train}} = \left\{ (X_t^{(i)}, o_t^{(i)}, y_t^{(i)}) : i \in [N] \setminus \{j\}, t \in [T_i] \right\}, \quad (7)$$

$$\mathcal{D}^{\text{test}} = \left\{ (X_t^{(j)}, o_t^{(j)}, y_t^{(j)}) : t \in [T_j] \right\}. \quad (8)$$

We then define X_{train} as the collection of features from all bearings in the training data. This data is normalized and fed into the 2-means clustering algorithm. We found out that utilizing only the first ten features is sufficient for clustering, yielding stable transition times. For the sake of fair comparison, the same transition times are used in all three RUL estimation scenarios. The resulting transition points lead to the creation of a new training dataset, which consists of only the data after the transition times, aka *unhealthy* points:

$$\mathcal{D}^{\text{train}}_{\text{unhealthy}} = \left\{ \left(X_t^{(i)}, o_t^{(i)}, \tilde{y}_t^{(i)} \right) : i \in [N] \setminus \{j\}, t \geq t_{\text{TT}}^{(i)} \right\}, \quad (9)$$

where

$$\tilde{y}_t^{(i)} = \frac{T_i - t}{T_i - t_{\text{TT}}^{(i)}} \times 100\% \quad (10)$$

is the RUL percentage after the transition time.

Different RUL estimation scenarios require different input data, as illustrated in Figure 2. The details are provided as follows.

- Scenario 1: For this approach, we simply use the training data in the form

$$\left\{ \left(X_t^{(i)}, \tilde{y}_t^{(i)} \right) : i \in [N] \setminus \{j\}, t \geq t_{\text{TT}}^{(i)} \right\} . \quad (11)$$

- Scenario 2: Here, we utilize the training OC for PCA and divide its one-dimensional output space into 20 bins. This way, each $o_t^{(i)} \in \mathbb{R}^3$ is replaced with its discretized counterpart $\tilde{o}_t^{(i)} \in [20]$. Next, for each region $r \in [20]$, a distinct normalizer MinMax_r is applied to the features with the same operating region, i.e.,

$$X_{\text{train}}^r := \left\{ X_t^{(i)} : i \in [N] \setminus \{j\}, t \geq t_{\text{TT}}^{(i)}, \tilde{o}_t^{(i)} = r \right\} . \quad (12)$$

  Let $\tilde{X}_t^{(i)} := \text{MinMax}_{\tilde{o}_t^{(i)}}(X_t^{(i)})$ be the normalized counterpart of $X_t^{(i)}$. The training data for this method can be written as

$$\left\{ (\tilde{X}_t^{(i)}, \tilde{y}_t^{(i)}) : i \in [N] \setminus \{j\}, t \geq t_{\text{TT}}^{(i)} \right\} . \quad (13)$$

- Scenario 3: We simply concatenate feature vectors and OC to create 23-dimensional features. The training data would be in the form

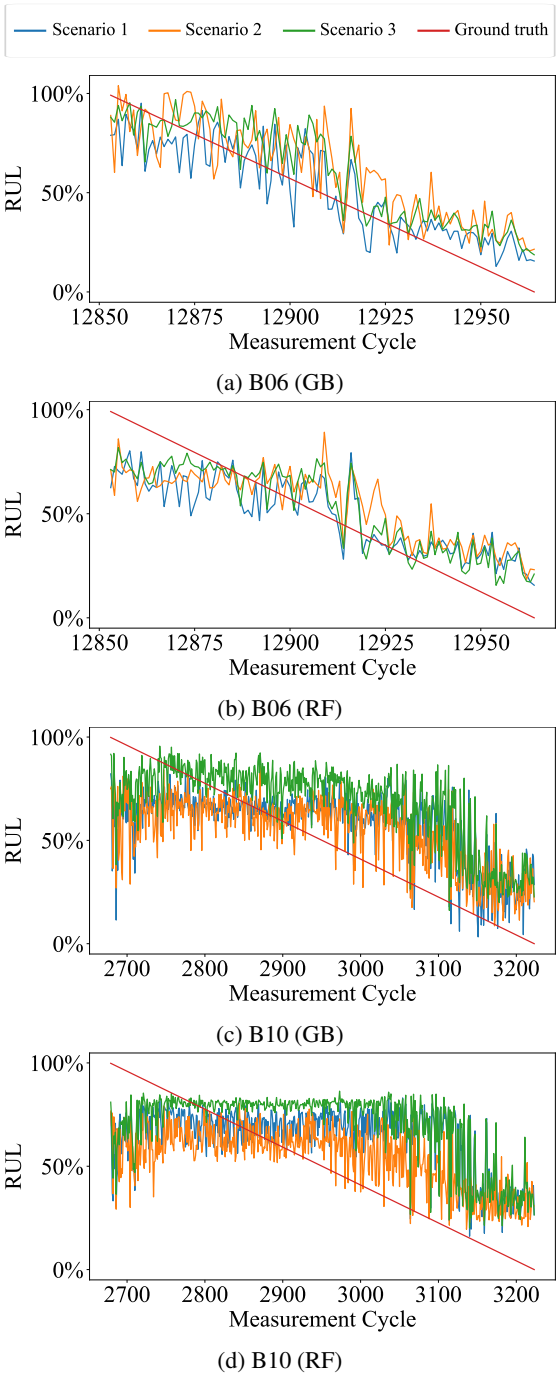
$$\left\{ \left([X_t^{(i)}; o_t^{(i)}], \ \tilde{y}_t^{(i)} \right) : i \in [N] \setminus \{j\}, t \geq t_{\text{TT}}^{(i)} \right\} . \quad (14)$$

We made the data, as well as all the implementations, publicly available on Zenodo[1] and GitHub[2] to encourage further development of RUL estimation models in dynamic operating conditions.

3.2. Results and Discussion

We employed GB and RF as models for estimating RUL. For each test bearing, separate models were trained for every estimation scenario. Figure 5 compares the performance of the three RUL estimation scenarios for two different bearings. To mitigate random effects, model training was repeated ten times for each test bearing and each scenario, using ten different random seeds for both GB and RF, and the resulting average mean absolute error (MAE) is reported in Table 3. Notably, the standard deviation of the MAE values was negligible and, therefore, not included in the table. The best MAE value for each bearing scenario is highlighted in bold.

The findings showcased in Table 3 shed light on the potential benefits of integrating OC information in the context of RUL estimation. Both learning models exhibited a decrease in MAE for more than 50% of the bearings when OC details were taken into account (scenarios 2 and 3 combined), with



(a) B06 (GB)

(b) B06 (RF)

(c) B10 (GB)

(d) B10 (RF)

Figure 5. RUL estimation performance comparison for three scenarios.

RF benefiting more compared to GB. It should be noted that preprocessing steps, such as feature extraction and transition time identification, can also affect the final outcomes. Despite this, the primary focus here is to compare the performance of different scenarios under fixed preprocessing steps.

[1] https://doi.org/10.5281/zenodo.10805042
[2] https://github.com/alireza-javanmardi/bearing-RUL

Table 3. MAE of the predictions in different scenarios.

			GB			RF		
Bearing	Total lifetime	Transition time	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
B01	377	208	**18.14**	22.77	19.96	18.00	21.51	**17.17**
B02	1116	998	29.43	**25.25**	27.41	28.74	**26.10**	27.24
B03	614	562	24.05	**21.52**	25.83	25.24	**21.90**	24.17
B04	1114	452	15.23	19.25	**15.02**	14.90	19.56	**14.07**
B05	572	560	44.40	41.11	**36.34**	42.20	37.28	**34.75**
B06	12965	12853	**11.73**	16.00	13.55	13.50	15.64	**11.85**
B07	6393	6205	44.36	**42.54**	42.82	43.63	44.78	**43.09**
B08	1827	1219	**15.42**	17.98	15.70	18.18	19.08	**17.51**
B09	1813	253	20.85	23.19	**17.19**	21.41	22.95	**19.67**
B10	3224	2679	19.48	**18.62**	23.13	23.87	**18.79**	26.11
B11	1953	931	23.94	23.85	**22.79**	25.22	**24.28**	24.66
B12	767	154	**15.80**	17.26	17.43	16.90	**16.27**	16.93
B13	19417	18022	26.91	**25.69**	27.36	**24.77**	26.78	29.13
B14	12317	12050	30.71	**26.75**	30.10	30.14	**29.10**	30.36
B15	22567	21051	16.74	19.88	**14.17**	14.20	21.32	**13.66**
B16	5891	5400	**20.17**	21.42	24.74	20.28	**18.41**	21.74
B17	2733	2323	22.19	**21.88**	24.47	22.37	24.15	**21.58**

4. Summary and Outlook

To address the limitation of existing studies and enhance the relevance to practical applications, a new ball bearing run-to-failure dataset, considering time-varying operating conditions (OC), is introduced. Specifically, during an experiment, the test bearing is subjected to a sinusoidal load or Gaussian white noise superimposed on a static pre-load. Furthermore, the rotating speed takes on a random value from a stationary uniform distribution within a predefined interval between measurements. Owing to the degradation path of the ball bearings, a 2-means clustering algorithm is employed to partition the features extracted from raw vibration data into two states, namely healthy and unhealthy states. To estimate the remaining useful lifetime (RUL) for the unhealthy state, even under such time-varying OC, three different scenarios are considered, namely, **Scenario 1**, where the measured OC are disregarded, **Scenario 2**, where the OC are employed for feature scaling, and **Scenario 3**, where the OC serve as auxiliary features. Different machine learning techniques, such as gradient boosting and random forest, are employed as the RUL estimator for each scenario. The results of the presented case study suggest that the usefulness of incorporating OC information depends on the individual case: in some scenarios, it is clearly advantageous, and in others, it does not yield significant benefits.

As a future work, one may delve deeper into other ways of incorporating OC information into RUL estimation. For instance, a hybrid model consisting of a physics-based model and a machine learning model can be an interesting extension. The physics-based model could capture the relationship between the varying OC and the system state, while the machine learning model could capture the relationship between the measured system parameters and the RUL. Moreover, more advanced learning algorithms, such as deep learning techniques, along with tools from domain adaptation and transfer learning, can be employed on this dataset to determine whether they can enhance the results.

Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the Project number 451737409.

References

Aimiyekagbon, O. K. (2024). *Run-to-failure data set of ball bearings subjected to time-varying load and speed conditions.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.10805042` (Data set)

Alewine, K., & Chen, W. (2010). Wind turbine generator failure modes analysis and occurrence. In *Wind power 2010 conference, dallas.*

da Costa, P. R. d. O., Akçay, A., Zhang, Y., & Kaymak, U. (2020). Remaining useful lifetime prediction via deep domain adaptation. *Reliability Engineering & System Safety*, *195*.

DFG. (2017). *Schwingungsanalysesystem für automatisierte Schwingungsmessungen. Major research instrumentation supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the Project number 391178551*.

Ding, Y., Ding, P., Zhao, X., Cao, Y., & Jia, M. (2022). Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation. *IEEE/ASME Transactions on Mechatronics*, *27*.

Ding, Y., Jia, M., & Cao, Y. (2021). Remaining useful life estimation under multiple operating conditions via deep subdomain adaptation. *IEEE Transactions on Instrumentation and Measurement*, *70*.

Ding, Y., Jia, M., Miao, Q., & Huang, P. (2021). Remaining useful life estimation using deep metric transfer learning for kernel regression. *Reliability Engineering & System Safety*, *212*.

Du, W., Hou, X., & Wang, H. (2022). Time-varying degradation model for remaining useful life prediction of rolling bearings under variable rotational speed. *Applied Sciences*, *12*. Retrieved from https://www.mdpi.com/2076-3417/12/8/4044

Fan, Y., Nowaczyk, S., & Rögnvaldsson, T. (2020). Transfer learning for remaining useful life prediction based on consensus self-organizing models. *Reliability Engineering & System Safety*, *203*.

Fu, S., Zhong, S., Lin, L., & Zhao, M. (2021). A novel time-series memory auto-encoder with sequentially updated reconstructions for remaining useful life prediction. *IEEE Transactions on Neural Networks and Learning Systems*, *33*.

Huang, C.-G., Huang, H.-Z., & Li, Y.-F. (2019). A bidirectional lstm prognostics method under multiple operational conditions. *IEEE Transactions on Industrial Electronics*, *66*.

Javanmardi, A., & Hüllermeier, E. (2023). Conformal prediction intervals for remaining useful lifetime estimation. *International Journal of Prognostics and Health Management*, *2*.

Lee, J., Qiu, H., Yu, G., Lin, J., & Services, R. T. (2007). IMS, university of cincinnati. bearing data set. *NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA*.

Li, F., Zhang, L., Chen, B., Gao, D., Cheng, Y., Zhang, X., … Huang, Z. (2020). An optimal stacking ensemble for remaining useful life estimation of systems under multi-operating conditions. *IEEE Access*, *8*.

Li, N., Gebraeel, N., Lei, Y., Bian, L., & Si, X. (2019). Remaining useful life prediction of machinery under time-varying operating conditions based on a two-factor state-space model. *Reliability Engineering & System Safety*, *186*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0951832018313024

Li, X., Zhang, W., & Ding, Q. (2019). Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. *Reliability engineering & system safety*, *182*, 208–218.

Mao, W., He, J., & Zuo, M. J. (2019). Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning. *IEEE Transactions on Instrumentation and Measurement*, *69*.

Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., & Varnier, C. (2012). Pronostia: An experimental platform for bearings accelerated degradation tests. In *Ieee international conference on prognostics and health management, phm'12.* (pp. 1–8).

Ren, L., Sun, Y., Wang, H., & Zhang, L. (2018). Prediction of bearing remaining useful life with deep convolution neural network. *IEEE access*, *6*, 13041–13049.

Sharma, S., Abed, W., Sutton, R., & Subudhi, B. (2015). Corrosion fault diagnosis of rolling element bearing under constant and variable load and speed conditions. *IFAC-PapersOnLine*, *48*.

von Hahn, T., & Mechefske, C. K. (2022). Knowledge informed machine learning using a weibull-based loss function. *Journal of Prognostics and Health Management*, *2*.

Wang, B., Lei, Y., Li, N., & Li, N. (2018). A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability*, *69*(1), 401–412.

Zhang, W., Li, X., Ma, H., Luo, Z., & Li, X. (2021). Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions. *Reliability Engineering & System Safety*, *211*.