

# A Machine Learning Dataset of Artificial Inner Ring Damage on Cylindrical Roller Bearings Measured Under Varying Cross-Influences

Christopher Schnur <sup>1,2,\*</sup> , Payman Goodarzi <sup>1</sup> , Yannick Robin <sup>1</sup> , Julian Schauer <sup>1,2</sup>  and Andreas Schütze <sup>1,2</sup> 

<sup>1</sup> Lab for Measurement Technology, Saarland University, 66123 Saarbrücken, Germany

<sup>2</sup> Centre for Mechatronics and Automation Technology gGmbH, 66121 Saarbrücken, Germany

\* Correspondence: info@lmt.uni-saarland.de; Tel.: +49-681-302-4664

**Abstract:** In practical machine learning (ML) applications, covariate shifts and dependencies can significantly impact model robustness and prediction quality, leading to performance degradation under distribution shifts. In industrial settings, it is crucial to account for covariates during the design of experiments to ensure reliable generalization. The presented dataset of undamaged and artificially damaged cylindrical roller bearings is designed to address the lack of data resources for targeting domain and distribution shifts in this field. The dataset considers multiple key covariates, including mounting position, load, and rotational speed. Each covariate consists of multiple levels optimized for group-based cross-validation. This allows the user to exclude specific groups in the training to validate and test the algorithm. Using this approach, algorithms can be evaluated for their robustness and the effect on the model caused by distribution shifts, allowing their generalization capabilities to be studied under realistic conditions.

**Dataset:** Published on Zenodo. DOI: 10.5281/zenodo.11108503 (MATLAB), 10.5281/zenodo.11108503 (CSV/Python)

**Dataset License:** Creative Commons Attribution 4.0 International (CC-BY)

**Keywords:** machine learning; robust learning; domain shift; bearing dataset



Academic Editor: Juraj Gregaň

Received: 6 April 2025

Revised: 11 May 2025

Accepted: 13 May 2025

Published: 16 May 2025

**Citation:** Schnur, C.; Goodarzi, P.; Robin, Y.; Schauer, J.; Schütze, A. A Machine Learning Dataset of Artificial Inner Ring Damage on Cylindrical Roller Bearings Measured Under Varying Cross-Influences. *Data* **2025**, *10*, 77. <https://doi.org/10.3390/data10050077>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Roller bearings are widely used rotating machine elements that reduce friction and carry loads. Although bearings are considered robust and have a long service life, incorrect usage can lead to unexpected bearing failure and, eventually, machine failure. Typical bearing failures are, e.g., wear, corrosion, or fracture and cracking [1]. In particular, pitting corrosion, which forms small hole-like corrosion pits in the metal, can cause severe damage to the bearing and result in failure [2]. To investigate these damage characteristics using machine learning (ML), multiple datasets are publicly available, e.g.,:

- NASA bearing dataset [3]: The dataset contains acceleration measurements with four bearings that are stressed with a constant load until they reach their wear limit.
- Paderborn University Bearing Dataset [4]: The dataset contains acceleration, rotational speed, load, and torque measurements of 26 damaged (artificial and real) and six undamaged bearings in four scenarios.

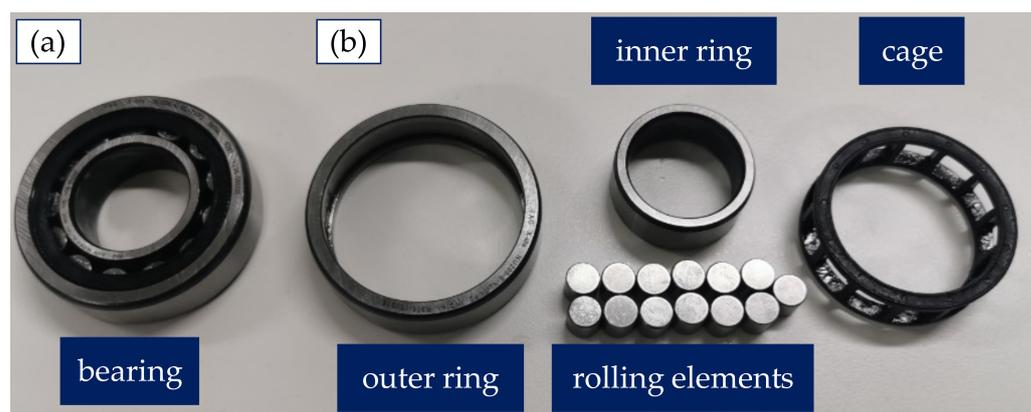
- Case Western Reserve University Bearing Dataset [5]: The dataset contains measurements of an accelerometer for artificially damaged bearings with different damage sizes and loads.

The datasets mentioned above incorporate covariates to a limited extent, such as load or rotational speed. In real-world scenarios, several additional covariates may occur simultaneously and interact with each other. As a result, measurements could be inadequate for developing robust ML models. Therefore, these interactions must be taken into account during data collection. This justifies the need for a new dataset with a focus on covariates during the measurement process. The bearing dataset presented in this study consists of acceleration measurements of three cylindrical roller bearings (B10, B20, and B30) successively measured on a testbed. Each bearing was first measured undamaged and later artificially damaged with a milling cutter. The dataset was designed to address the lack of publicly available data enabling systematic analysis of covariate shift and domain adaptation challenges in machine learning models. By providing controlled variation of multiple influencing factors, this dataset enables more realistic assessment of model robustness under variable operating conditions.

## 2. Methods

### 2.1. Bearing

Figure 1 shows an assembled cylindrical roller bearing of type NU206-E-XL-TVP2 (a), as investigated in the experiments, along with its disassembled components (b).



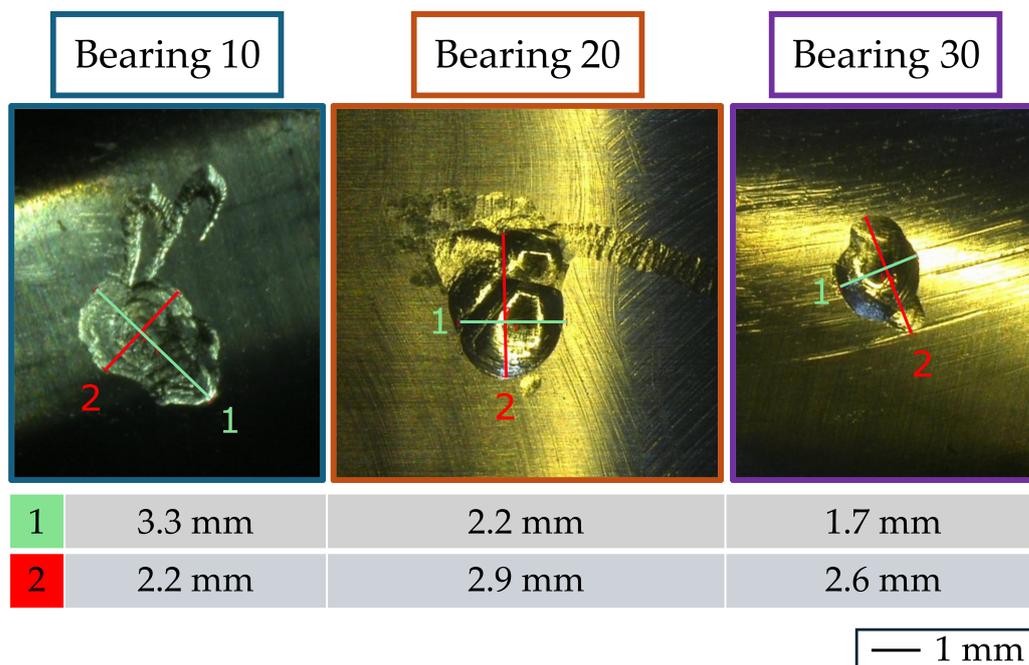
**Figure 1.** (a) Assembled cylindrical roller bearing and (b) disassembled into its components: outer ring, inner ring, rolling elements, and cage.

The inner ring has no ribs to secure the rolling element, so that it can be easily separated. This approach enables the introduction of artificial damage to the inner ring without causing additional damage during the bearing's mounting and dismounting. The corresponding basic frequency factors can be found in Table 1 [6].

**Table 1.** Basic frequency factors of the NU206-E-XL-TVP2 related to 1/s [6].

Basic Frequency Factors	Abbreviation	Factor
Overrolling frequency factor on outer ring	$BPFFO$	5.24
Overrolling frequency factor on inner ring	$BPFFI$	7.76
Overrolling frequency factor on rolling element	$BSFF$	2.49
Ring pass frequency factor on rolling element	$RPFFB$	4.97
Speed factor of rolling element set for rotating inner ring	$FTFF_i$	0.40
Speed factor of rolling element set for rotating outer ring	$FTFF_o$	0.60

Figure 2 presents microscope images of the artificial damage applied to bearings B10, B20, and B30, along with their dimensions, which are also included in the dataset.



**Figure 2.** Microscope images of the artificial damage introduced to bearings 10, 20, and 30, along with their respective dimensions.

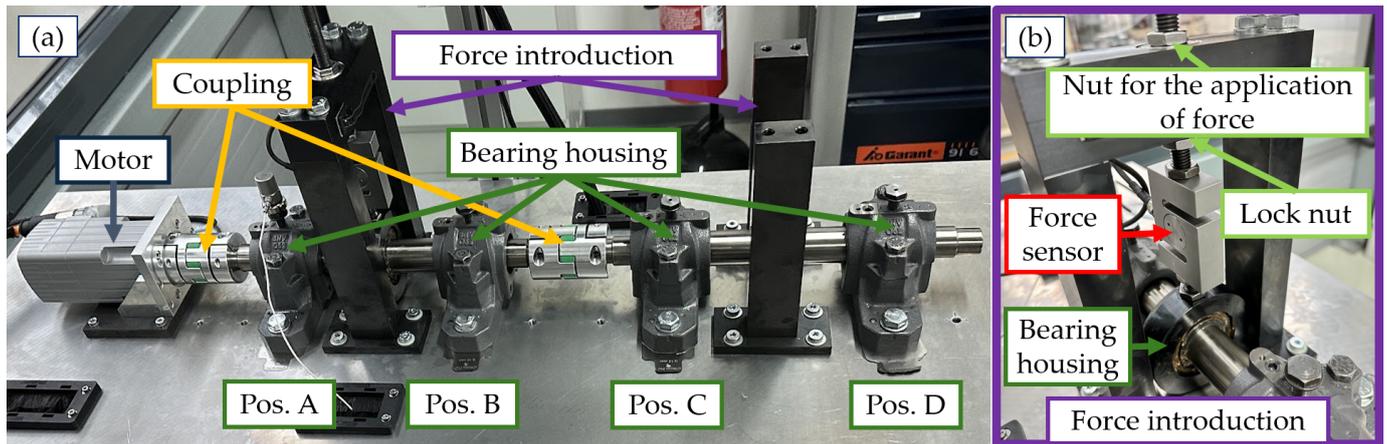
In addition to the NU206-E-XL-TVP2, which was used as a loose bearing in the setup, a 1206-TVH was employed as a fixed bearing to compensate for axial forces and minor shaft misalignments. The NU207-E-XL-TVP2, another loose bearing, was used to apply the pulling force to the rotating shaft. All relevant frequency factors for 1206-TVH and NU207-E-XL-TVP2 are provided in Table A1.

## 2.2. Testbed

Figure 3a shows the mechanical setup of the testbed. The corresponding components of the testbed (I. Mechanical System) and the data acquisition (II. Data Acquisition System) can be found in Table 2.

**Table 2.** Components of the testbed.

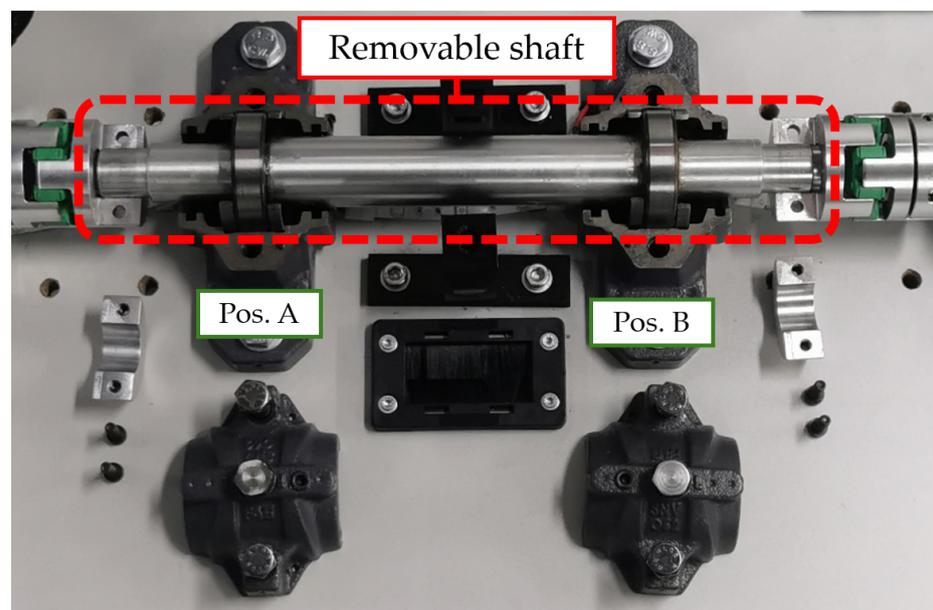
Component	Model	Manufacturer
I. Mechanical System		
Motor	EMMS-AS-70S-LS-RSB	Festo
Motor controller	CMMP-AS-C2-3A-M3	Festo
Coupling	GWE 5106-24-11-25	Ringfeder Power Transmission
Loose bearing (Cylindrical roller bearing)	NU206-E-XL-TVP2	Schaeffler Technologies
Fixed bearing (Self-aligning ball bearing)	1206-TVH	Schaeffler Technologies
Bearing Force introduction (Cylindrical roller bearing)	NU207-E-XL-TVP2	Schaeffler Technologies
II. Data Acquisition System		
Accelerometer	3233a	Dytran Instruments
Force Sensor	K-25	Lorenz Messtechnik
Embedded Controller	cRIO 9040	National Instruments
Vibration Input Module	NI-9232	National Instruments
Voltage Input Module	NI-9215	National Instruments



**Figure 3.** (a) Mechanical setup of the testbed. (b) Enlarged view on the force introduction.

In the testbed, a servo motor (indicated in blue in Figure 3) powered two shafts connected by two couplings (yellow). Each shaft was supported by one fixed and one loose bearing located in the bearing housings (green). Two Force Introductions were constructed to apply an external force on the bearings. Figure 3b shows an enlarged view of one Force Introduction with the components force-application nut and locking nut (light-green), force sensor (red), and bearing housing (green). The Force Introduction applied a pulling force on one shaft at a time by tightening the force-application nut. Using a pulling force improved the signal path of the resulting vibrations to the accelerometer as the sensor was mounted on top of the bearing housing [7].

The design of the testbed allowed for mounting the bearing in other bearing housings without disassembling it from the shaft (Figure 4), as disassembling and reassembling could influence the data.



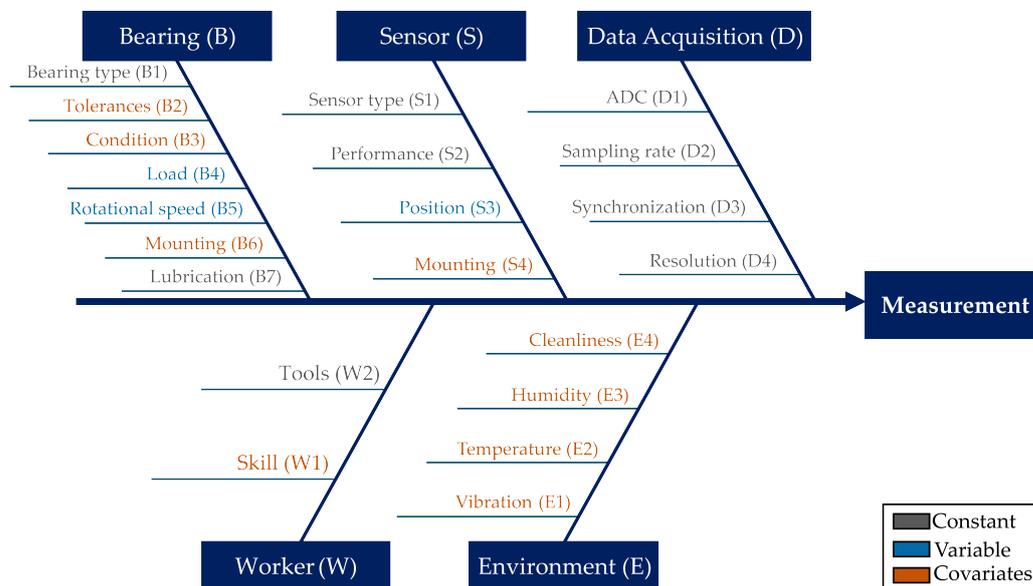
**Figure 4.** Dismantled shaft before a position change.

To change the bearing position, the covers of the bearing houses and the couplings were removed. Next, the shaft with the mounted bearings was lifted out of the bearing housings and rotated by  $180^\circ$  so that the bearing was switched from Pos. A to Pos. B. The two shafts had the same dimensions, so they could be exchanged. This allowed for a bearing to be mounted in all four positions without dismounting it from the shaft. Note that the direction of rotation of the shaft was constant.

Before the experiments, the testbed was aligned using a Fixturlaser EVO (see Figure A1). The resulting measurements are presented in Table A2.

### 2.3. Identification of Influencing Factors

To identify relevant influencing factors on the data, a cause–effect graph [8] was used (see Figure 5).



**Figure 5.** Cause–effect graph illustrating the influencing factors on the measurement for cylindrical roller bearings.

All identified influences were further assigned to the groups constant, variable, and covariates.

Constant influences (grey) were maintained constant throughout the measurements. These included the bearing type (B1) and its lubrication (B7, defined amount of lubrication), the sensor type (S1) and its performance (S2), the data acquisition (D1–D4), and the tools (W2) used by the workers. To ensure that the influence of the tools remained constant, specific torque wrenches with preset torque were used for each type of screw.

Variable influences (blue) are covariates varied in a controlled manner during the experiment. In this study, the load on the bearing (B4), its rotational speed (B5), and its mounting position (S4) were varied. For the variable influences Load (B4) and Rotational Speed (B5), Latin Hypercube Sampling [9] was used to determine the levels and their order in the design of experiments (DoE), ensuring well-distributed coverage while avoiding temporal correlation. The range of these values was determined based on the maximum values provided in the manufacturer’s specifications, including a safety factor of 2. For rotational speed, the maximum was limited by the coupling and set to 1000 rpm, while the maximum load was restricted by the fixed bearing (1206-TVH) and set to 3700 N. Note that all bearings, especially the loose bearing (NU206-E-XL-TVP2), are designed to handle higher loads and are intentionally oversized to minimize wear effects, which could influence the results during the course of the experiments.

Covariates (orange) could only be tracked or influenced with extended effort. Manufacturing tolerances (B2) were considered by repeating the experiment with three different bearings. Bearing condition (B3) was addressed by using new bearings and measuring the undamaged condition before introducing any damage. The mounting of the sensor (S4) and the bearing (B6) was managed using the concept of *Runs*. In each run, the sensor was mounted three times in the same configuration, following the order: Run 1 (Positions

A, B, C, D), Run 2 (Positions A, B, C, D), and Run 3 (Positions A, B, C, D). The influence of the worker’s skill was minimized through training and monitored via pictures of the configuration after modifications. Environmental influences (E1–E4) were reduced by conducting the measurements in a temperature-controlled laboratory, while temperature and humidity were recorded.

Figure 6 illustrates the variation of covariates in the DoE.

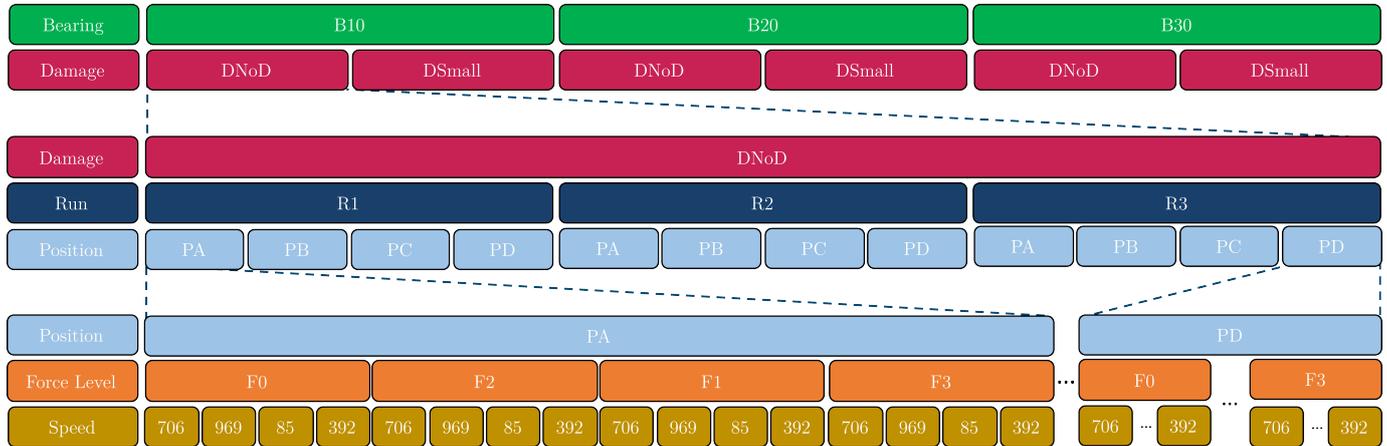


Figure 6. Variation of the covariates in the DoE.

### 3. Data Description

The dataset is published on Zenodo [10] and consists of the three folders Data, Meta-data and +functions as well as the two files info.mat and readdata.m. Figure 7 provides an overview of the dataset structure.

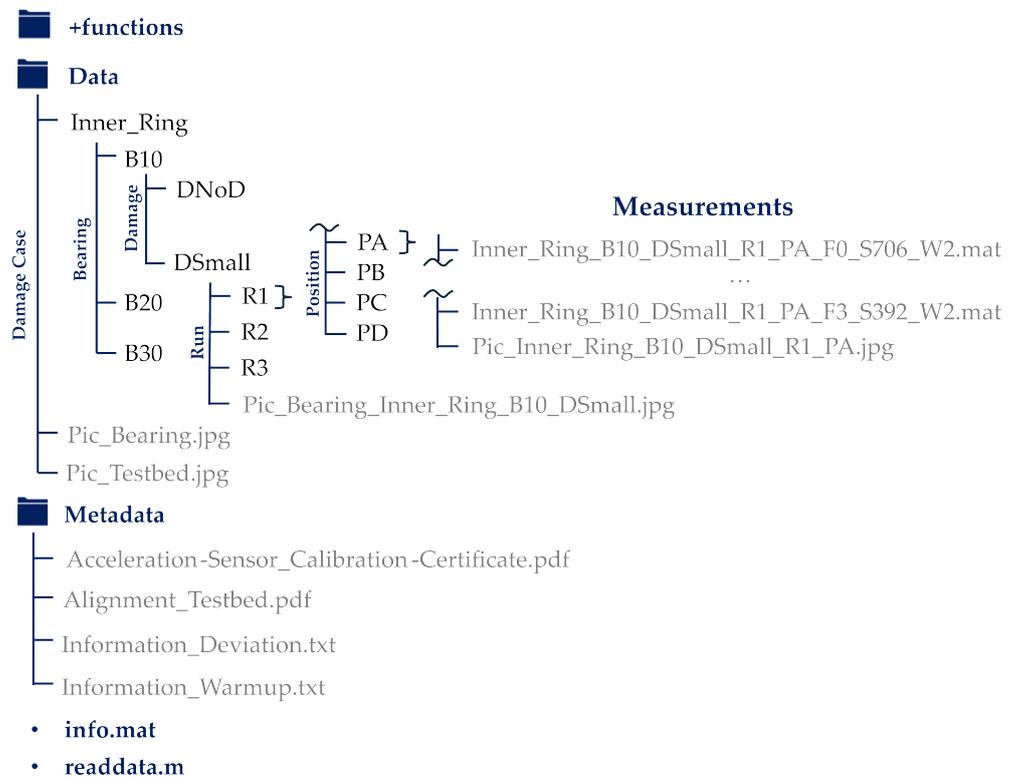


Figure 7. Overview of the dataset’s folder structure.

The Data folder contains 1151 measurement files, where each file contains the measurement of a three-axis accelerometer (20 kHz, 60 s) of a specific combination of covari-

ates, which is named according to the syntax `DamageCase_Bearing_Damage_Run_Position_Force_Speed_Worker`, e.g., `Inner_Ring_B10_DSmall_R1_PA_F0_S706_W2.mat`. Originally, 1152 combinations were planned, but one measurement file was missing due to corruption. The missing combination corresponded to Bearing 30, Damage Case 0, Run 1, Position B, Force Level 1, and Speed 392. In addition, the folders include images showing the testbed configuration and the dimensions of the damage. The folder structure reflects the testbed's specific configurations. Measurement files are, therefore, sorted according to bearing number (B10, B20, B30), damage size (DNoD for no damage, DSmall for small damage), run (R1, R2, R3), and position (PA, PB, PC, PD).

The folder `Metadata` contains the calibration certificate of the accelerometer, the testbed's alignment report, a text file listing all deviations from the initial DoE, and a text file containing further information about the warm-up phases of the testbed.

The file `info.mat` contains all the parameters recorded during the measurements. Table 3 lists the parameters of the measurements. Parameters 1–6 represent the controlled variables, while 7–12 correspond to the tracked covariates.

**Table 3.** Parameters for the design of experiments.

Nr.	Parameter	Quantity	Label	Values
1	Bearing	3	B10, B20, B30	10, 20, 30
2	Damage state	2	No damage, small damage	0, 1
3	Run (Position A to D)	3	R1, R2, R3	1, 2, 3
4	Position	4	PA, PB, PC, PD	1, 2, 3, 4
5	Force level <sup>1</sup> ( $\pm 50$ N)	4	F0 $\approx$ 0 N, F2 $\approx$ 2500 N, F1 $\approx$ 1600 N, F3 $\approx$ 3300 N	0, 2, 1, 3
6	Speed <sup>1</sup> [rpm]	4	706, 969, 85, 392	706, 969, 85, 392
7	Worker	2	W1, W2	1, 2
8	Mounting sensor	2	Normal, flipped	0, 1
9	Mounting coupling	4	Normal, twisted, right-centered, left-centered	0, 1, 2, 3
10	Mounting second shaft	2	Normal, flipped	0, 1
11	Temperature [°C]	-	-	21.6–22.7
12	Rel. humidity [%]	-	-	36.6–49.1

<sup>1</sup> Randomized order of the parameters.

In addition, this file contains the following information:

- **Timestamp:** The measurement start time is automatically recorded using the internal clock of the data acquisition system (NI cRIO 9040).
- **Measurement day and batches:** One measurement day consists of 48 batches. Each batch consists of all speed cycles for a given configuration.
- **Damage dimensions:** Each damage was measured in two dimensions using a microscope. The resulting images are included in the corresponding folders and shown in Figure 2. The `info.mat` contains the dimensions as `Damage_width` and `Damage_length`.
- **Filename:** Name of the measurement file with the corresponding folder path.

The file `readdata.mat` is designed to load the data into MATLAB and internally uses functions from the `+functions` folder. The user can load all data or selectively load measurements with specific annotation, e.g., only measurements from position `Pos. A` and force level `F0`. Furthermore, the user can reshape (split) measurements into segments. For example, if split into eight sub-measurements, each segment will contain at least 10 rotations [7].

After executing the script, the user receives the file `dataset.mat` with the variables `data` and `target`. The variable `data` is a  $1 \times 3$  cell containing the measurements of the three axis

X ({1,3}), Y ({2,3}) and Z ({3,3}), where the measurements are stored row-wise. The variable `trainTarget` contains the corresponding annotation (metadata) of the measurements in data.

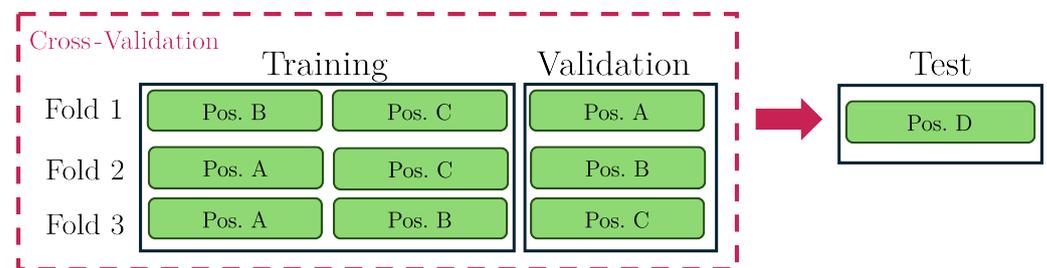
To support broader accessibility and interoperability, the dataset is provided both in MATLAB (.mat) and open CSV format, accompanied by data loading and processing scripts available in both MATLAB (version R2024b) and Python (version 3.13).

## 4. User Notes

### 4.1. Validation

Validation of ML models is a crucial step to assess their generalization ability and prevent overfitting [11]. Leave-One-Group-Out cross-validation (LOGOCV) is considered a more realistic scenario than methods such as k-fold cross-validation [12].

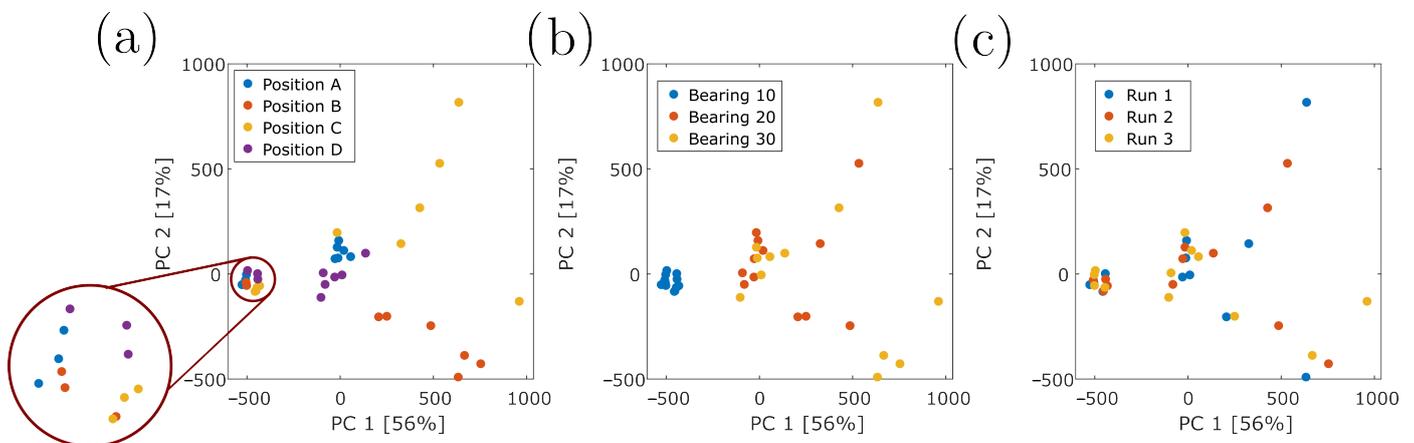
The presented bearing dataset is designed explicitly for LOGOCV. Figure 8 illustrates how LOGOCV can be applied to the dataset by excluding certain bearing positions. This approach reflects real-world scenarios where the ML model should perform robustly regardless of the bearing's mounting position.



**Figure 8.** Example of a LOGOCV for the bearing dataset.

In the example, two bearing positions are used for training, while a third, “unknown” position is used for validation in each fold of the LOGOCV. Subsequently, the model is tested on a fourth position that is entirely excluded from the training process. The resulting accuracy serves as an indicator of the ML model's robustness to new mounting positions. Furthermore, other covariates such as load and speed levels, measurements from different bearings, or combinations of covariates can also be employed in a LOGOCV scenario.

Figure 9 [13] highlights the relevance of considering cross-influences during model building. It shows the principal component analysis (PCA) [14] of features from the frequency domain (Best Fourier coefficients [15]) of the undamaged bearings colored by mounting position (a), bearing (b), and run (c).

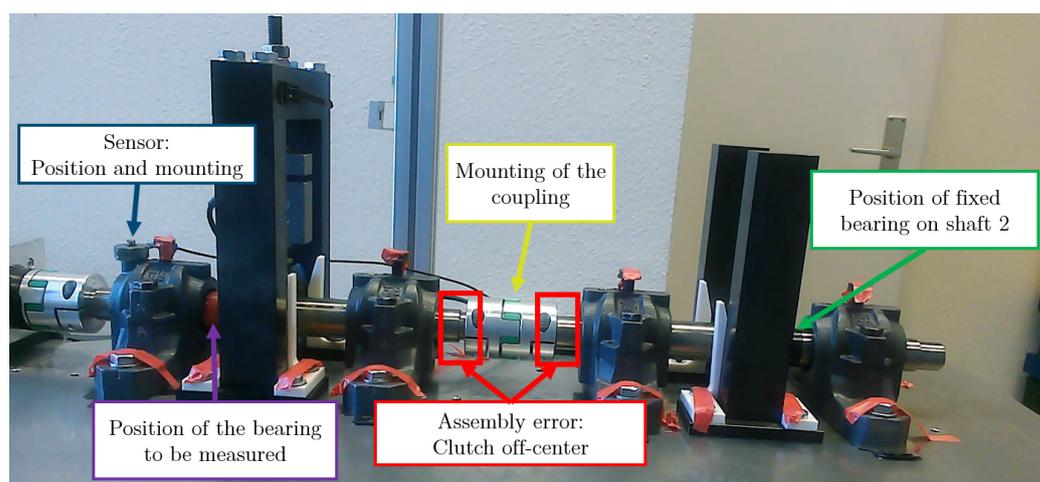


**Figure 9.** PCA of the undamaged measurements colored by the mounting position (a), bearing (b), and run (c). Adapted from [13].

In particular, the measurements of the positions and, to a certain degree, the measurements of the bearings tend to form clusters. This highlights the relevance of considering covariates as, e.g., a different mounting position can significantly change the data distribution and cause the ML model to fail. The presented dataset enables systematic investigation of such covariate influences by providing controlled, documented variation of key operating factors, allowing users to evaluate model robustness under these shifts.

#### 4.2. Assembly Errors

During the measurements, certain assembly errors occurred. By using the pictures of the configuration, these assembly errors (cf. Table 3) could be identified. Figure 10 shows, e.g., a configuration where the coupling is off-center.



**Figure 10.** Testbed configuration with the coupling mounted off-centered as assembly error.

Users can interpret these pictures as follows:

- The shaft with the bearing to be measured is indicated with a red off-centered ring (purple). A black off-centered ring (green) on the second shaft indicates the position of the fixed bearing. Due to the colored rings, the positions of all bearings can be tracked.
- The mounting of the sensor (blue) can be tracked by comparing the mounted position with the label in the dataset. In some measurements, the sensor is mounted upside down, which can be seen as a black surface on the top of the sensor (indicated in the data as `sensor_flipped`).
- The coupling in the middle (red) can be controlled on a centered mounting. Furthermore, it can be controlled if the coupling itself is mounted correctly, e.g., through the gap dimensions. Each side of the coupling has a corresponding engraving “R” for the right side and “L” for the left side, which are not visible in most of the pictures due to the camera’s low resolution. The coupling on the left side is always mounted on the motor side (screws covered), and only the shaft side is dismounted.
- The bearing housings have an engraving (e.g., A for Pos. A) on the cover and the body to check that the covers are mounted on the correct body in the correct orientation. All screws that are not used to change the testbed configuration are covered (red tape).

#### 4.3. Limitations

The following limitations apply to the bearing dataset:

- Despite numerous countermeasures, such as employee training, multiple assembly errors occurred during the measurements that were not part of the DoE. These assembly errors did not influence the function of the testbed, but might cause changes in the

data distribution. Therefore, they are transparently labeled in the data. As assembly errors also occur in real applications, users can try to identify those errors with their ML model and investigate their influence on the data.

- The damages on the inner ring of the bearing are artificial, meaning that the ML model is only valid for this specific error type. Artificially damaging bearings is an established method in bearing diagnostics research to simulate pitting corrosion, as it enables controlled, reproducible defect sizes and locations while avoiding the time and variability associated with natural fault development [16]. However, artificially induced defects can only approximate real-world damage to a certain extent and are not directly transferable to all operational conditions. Representative images of real pitting corrosion can be found in [17].
- The dataset includes only artificially induced damage on the inner ring, while the other bearing components remain undamaged. This design choice enabled isolation of inner ring defect effects, but simplifies real-world conditions, where damage can occur simultaneously on multiple components, progressively worsen over time, and arise from more complex degradation mechanisms than those represented by the artificially introduced defects [17]. Consequently, machine learning models trained on this dataset exhibit limited sensitivity to early-stage, progressive, or multi-component faults.
- The metadata are provided in descriptive form but not in a standardized schema, which may limit automated indexing.

## 5. Conclusions

The presented bearing dataset focuses on the controlled variation of multiple covariates and allows users to evaluate ML models under realistic validation scenarios to systematically test their robustness. Assembly errors that occurred during the measurements are well documented, enabling investigation of their influence on the data.

A first investigation of the dataset [13] demonstrated successful detection of bearing damage at positions excluded from the training data (LOGOCV), achieving a validation error of 4.3 % (969 rpm, F1–F3). This result shows that ML models trained on this dataset can generalize to previously unseen mounting positions despite covariate shifts, confirming the dataset's applicability for evaluating model robustness under more realistic scenarios. In addition to this validation, exploratory PCA analysis revealed clustering effects based on the mounting position and bearing, further highlighting the significant influence of covariates on the data distribution and the potential impact on model performance.

These findings emphasize the need to systematically account for varying operational conditions during model development and validation to ensure reliable fault detection in practical applications. Further studies could investigate the combined effects of multiple covariates, such as unknown speed and unknown mounting position in LOGOCV, as well as explore the transferability of ML models.

**Author Contributions:** Conceptualization, C.S., P.G. and Y.R.; methodology, C.S., Y.R. and P.G.; software, C.S.; validation, C.S. and P.G.; formal analysis, P.G.; investigation, C.S. and P.G.; resources, C.S., J.S. and P.G.; data curation, C.S. and P.G.; writing—original draft preparation, C.S.; writing—review and editing, P.G., J.S., Y.R. and A.S.; visualization, C.S. and P.G.; supervision, A.S.; project administration, A.S.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was performed in the context of project VProSaar (“Verteilte Produktion für die saarländische Automotivindustrie: Nachhaltig, Vernetzt, Resilient”) carried out at the Centre for Mechatronics and Automation Technology gGmbH and funded by the Ministry of Economic Affairs, Innovation, Digital and Energy (MWIDE) and the European Funds for Regional Development (EFRE).

**Data Availability Statement:** The dataset is available on Zenodo. MATLAB: <https://doi.org/10.5281/zenodo.11108503>. CSV/Python: <https://doi.org/10.5281/zenodo.15376390>.

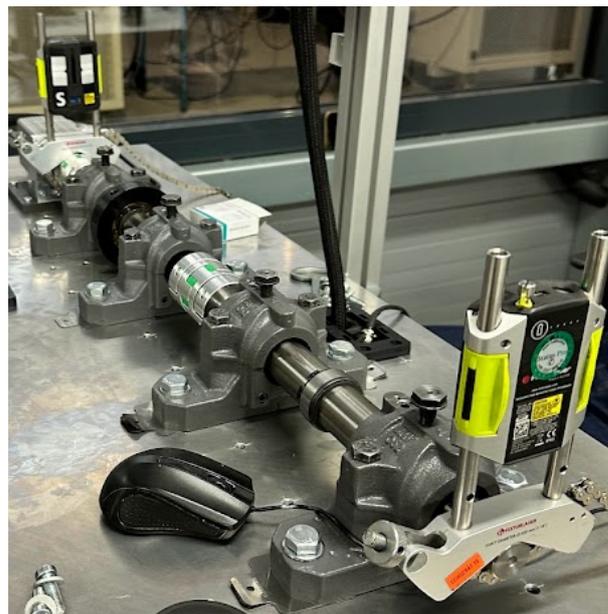
**Acknowledgments:** The authors sincerely thank Houssam El Moutaouakil for his valuable technical support during the measurements and Jannis Morsch for testing the dataset, providing insightful feedback on the paper and improving the source code. Furthermore, they would like to thank the research assistants Yage Zhang and Ali Ali Ahmad.

**Conflicts of Interest:** Authors Christopher Schnur, Julian Schauer and Andreas Schütze were employed by the company Centre for Mechatronics and Automation Technology gGmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Appendix A

**Table A1.** Basic frequency factors of the 1206-TVH [18] and NU207-E-XL-TVP2 [19] related to 1/s.

Basic Frequency Factors [1/s]	1206-TVH	NU207-E-XL-TVP2
$BPF_{FO}$	5.79	5.70
$BPF_{FI}$	8.21	8.30
$BSFF$	2.76	2.61
$RPFFB$	5.52	5.21
$F_{TF_i}$	0.41	0.41
$F_{TF_0}$	0.59	0.59



**Figure A1.** Alignment process with the Fixturlaser EVO.

**Table A2.** Results of the alignment measurement.

Position	Measurement	Unit
Vertical Angle	−0.011	°
Vertical Offset	−0.079	mm
Horizontal Angle	−0.021	°
Horizontal Offset	0.063	mm

## References

1. ISO 15243:2017; Rolling Bearings—Damage and Failures—Terms, Characteristics and Causes. International Organization for Standardization: Geneva, Switzerland, 2017.
2. Xu, F.; Ding, N.; Li, N.; Liu, L.; Hou, N.; Xu, N.; Guo, W.; Tian, L.; Xu, H.; Lawrence Wu, C.M.; et al. A review of bearing failure Modes, mechanisms and causes. *Eng. Fail. Anal.* **2023**, *152*, 107518. [CrossRef]
3. Tyagi, V. NASA Bearing Dataset. 2007. Available online: <https://www.kaggle.com/datasets/vinayak123tyagi/bearing-dataset> (accessed on 24 February 2025).
4. Lessmeier, C.; Kimotho, J.K.; Zimmer, D.; Sextro, W. Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification. *PHM Soc. Eur. Conf.* **2016**, *3*, 1–17. [CrossRef]
5. Case Western School of Engineering. Case Western Reserve University Bearing Data Set. Available online: <https://www.engineering.case.edu/bearingdatacenter> (accessed on 24 February 2025).
6. Schaeffler Technologies AG & Co. KG. Zylinderrollenlager NU206-E-XL-TVP2. 2025. Available online: <https://medias.schaeffler.de/de/produkt/rotary/waelz--und-gleitlager/rollenlager/zyylinderrollenlager/nu206-e-xl-tvp2/p/368765>. (accessed on 24 February 2025).
7. Schaeffler Monitoring Services GmbH. *Condition Monitoring Praxis: Handbuch zur Schwingungs-Zustandsüberwachung von Maschinen und Anlagen*, 1. auflage ed.; Vereinigte Fachverlage: Mainz, Germany, 2019.
8. Ishikawa, K.; Ishikawa, K. *Guide to Quality Control*, 13 print ed.; Asian Productivity Organization: Tokyo, Japan, 1996.
9. Loh, W.L. On Latin hypercube sampling. *Ann. Stat.* **1996**, *24*, 2058–2080. [CrossRef]
10. Schnur, C.; Goodarzi, P.; Robin, Y.; Schauer, J.; El Moutaouakil, H.; Ahmad, A.A.; Zhang, Y.; Schneider, T.; Schütze, A. A Cylindrical Roller Bearing Dataset with Varying Speed, Force and Position for Robust and Trasferable Machine Learning. 2025. Available online: <https://zenodo.org/records/11108503> (accessed on 24 February 2025).
11. Maleki, F.; Muthukrishnan, N.; Ovens, K.; Reinhold, C.; Forghani, R. Machine Learning Algorithm Validation: From Essentials to Advanced Applications and Implications for Regulatory Certification and Deployment. *Neuroimaging Clin. North Am.* **2020**, *30*, 433–445. [CrossRef] [PubMed]
12. Goodarzi, P.; Schütze, A.; Schneider, T. Comparing AutoML and Deep Learning Methods for Condition Monitoring using Realistic Validation Scenarios. *arXiv* **2023**, arXiv:2308.14632. [CrossRef]
13. Schnur, C. Methodisches Vorgehen zur Realisierung von maschinellen Lernprojekten im Mittelstand. Ph.D. Thesis, Naturwissenschaftlich-Technische Fakultät, Universität des Saarlandes, Saarbrücken, Germany, 2025.
14. Jolliffe, I.T. *Principal Component Analysis*; Springer Series in Statistics; Springer: New York, NY, USA, 2002.
15. Schneider, T.; Helwig, N.; Schütze, A. Automatic feature extraction and selection for classification of cyclical time series data. *Tech. Mess.* **2017**, *84*, 198–206. [CrossRef]
16. Ye, L.; Saxena, K.K.; Qian, J.; Reynaerts, D. Micro-EDM Drilling/Milling as a Potential Technique for Fabrication of Bespoke Artificial Defects on Bearing Raceways. *Micromachines* **2022**, *13*, 483. [CrossRef] [PubMed]
17. Schaeffler Technologies AG & Co. KG. Rolling Bearing Damage—Recognition of Damage and Bearing Inspection. 2022. Available online: [https://www.schaeffler.com/remotemedien/media/\\_shared\\_media/08\\_media\\_library/01\\_publications/schaeffler\\_2/publication/downloads\\_18/wl\\_82102\\_3\\_de\\_en.pdf](https://www.schaeffler.com/remotemedien/media/_shared_media/08_media_library/01_publications/schaeffler_2/publication/downloads_18/wl_82102_3_de_en.pdf) (accessed on 5 May 2025).
18. Schaeffler Technologies AG & Co. KG. Pendelkugellager 1206-TVH. 2025. Available online: <https://medias.schaeffler.de/de/produkt/rotary/waelz--und-gleitlager/kugellager/pendelkugellager/1206-tvh/p/365837>, (accessed on 24 February 2025).
19. Schaeffler Technologies AG & Co. KG. Zylinderrollenlager NU207-E-XL-TVP2. 2025. Available online: <https://medias.schaeffler.de/en/product/rotary/rolling-and-plain-bearings/roller-bearings/cylindrical-roller-bearings/nu207-e-xl-tvp2/p/368768>. (accessed on 24 February 2025).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.