Early detection of bearing degradation

Nadav Rindler (nrindler3)

Applied Analytics Practicum (MGT 6748) – Final Report

Sponsor: Sandia National Labs

Georgia Tech Online Master's in Analytics

Spring 2024

Georgia Tech

Predictive maintenance

Prognostics

- Apply data science and machine learning techniques to predict in advance when maintenance or replacement will be needed via the detection of failure modes and modeling of degradation.[1]
  - Identify components that need replacing or machines that need servicing before fault or failure occurs.
  - Find the optimal balance between maximizing component lifetime and reducing downtime due to malfunction.

Project scope

- **Classify high-frequency bearing vibration signal data** into 4 stages (healthy, worn, faulty, failed), for early degradation detection.
  - Utilize open-source bearing data from the NASA Prognostics Data Repository.

Method

- Common predictive maintenance problems include modeling **continuous degradation** (e.g., predicting the Remaining Useful Life of a bearing), a regression problem, and **discrete degradation** (e.g. identifying discrete stages of bearing degradation), a classification problem.

- In both approaches, "typical features for bearing diagnostics and prognostics include time-domain features (root mean square, peak value, signal kurtosis, etc.) or frequency-domain features (peak frequency, mean frequency, etc.)" [6].

- Project scope, per sponsor Sandia National Laboratories, is to build a **signal classifier**. Classifiers can include:

  - **Single-class:** model trained using only healthy bearing data, and faults are detected when the signal deviates significantly from the healthy signal.

  - **Two-class (binary):** model trained on data from the beginning (healthy) and end (faulty) from a run-to-failure experiment.

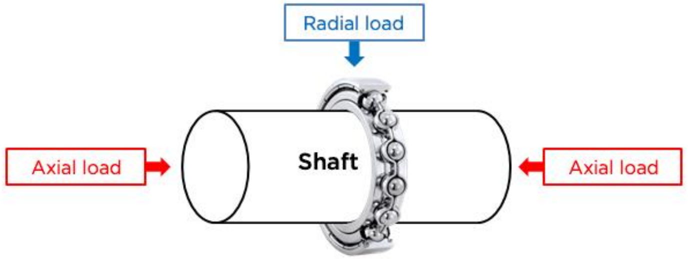
  - **Multi-class:** model trained on data segmented into healthy, degraded, faulty, and failed states, corresponding to the bearing's degradation stage.

- I initially attempted a **semi-supervised learning approach** developed by researchers at the IT University of Copenhagen (ITUC) for use in bearing degradation classification (Juodelyte, et. al. 2022) [3]. I then switched to a new **unsupervised time series segmentation** approach that uses time series segmentation (TSS) for change detection (Ermshaus, et. al. 2023) [11].

*For numbered references, see Appendix.

IMS Bearing Data

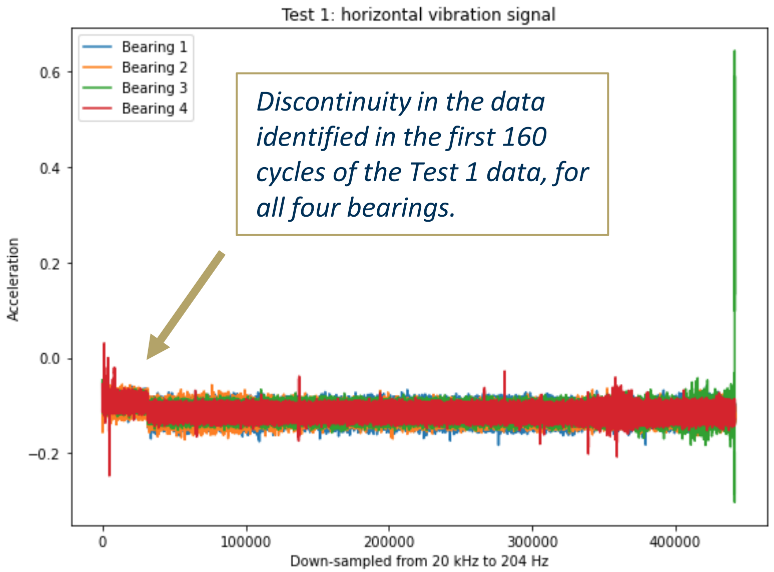
- Data comprised of three run-to-failure experiments, each with four bearings. Experiments were run until one or more bearings failed.
- Files were provided in a >1GB zipped folder. File names indicate the measurement timestamp.
- Each file is a **1-second snapshots of vibration signal (acceleration)** recorded at 5- or 10-minute intervals across multiple days and contains 20,480 observations per second (sampling rate of **20,480 Hz**).
- Experiment #1 included two accelerometers (horizontal and vertical vibration). Experiments #2 and #3 included only one accelerometer.
- Data is **un-labelled**, meaning that we do not know bearings' status at each observation – only the measured acceleration.
- **Data Source:** Center for Intelligent Maintenance Systems (IMS), University of Cincinnati (2007).[12]



Bearing diagram. Source: JTekt.

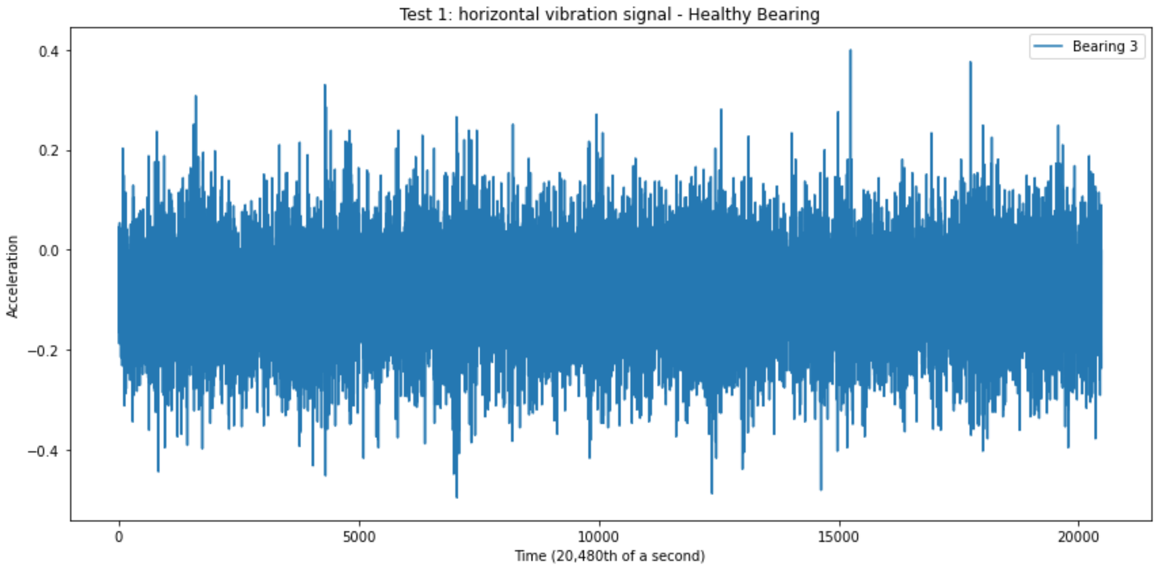
Data Processing

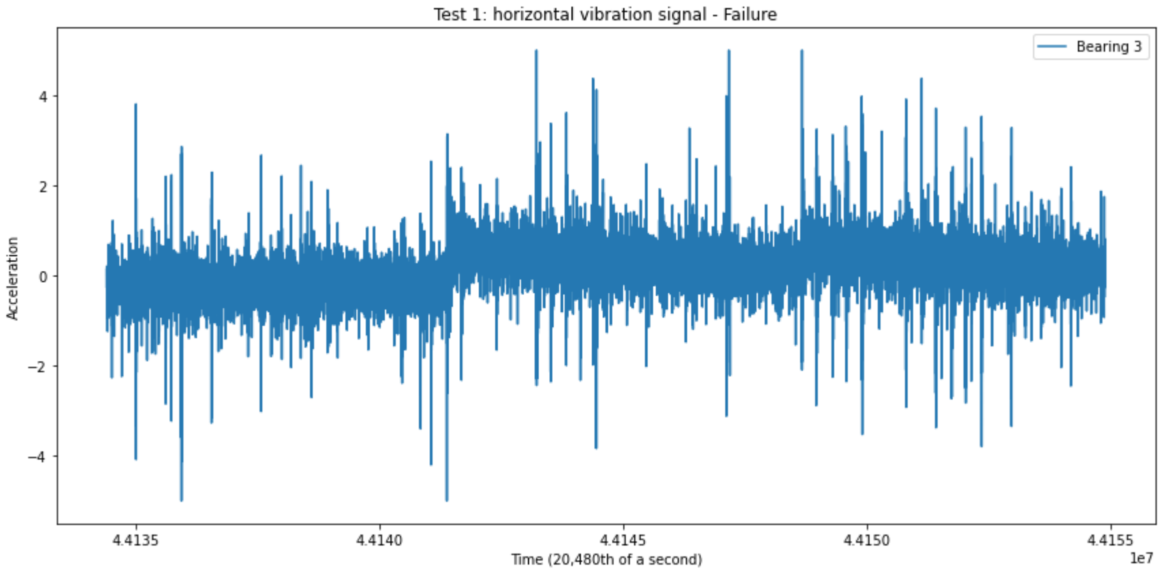
- Data stored in text files (2,156 files in Test 1, each represents a 1-second snapshot of signal data sampled at 20,480 Hz). Files are read into memory and concatenated in Python. Processing included:
    - **Filtering**: excluded vertical (y-axis) signal data in Test 1 because it was not measured in Tests 2 and 3. Excluded the
    - **Sampling**: down-sampled by a factor of 10 from 20 kHz to ~2 kHz to reduce volume of data.
    - **Time-domain features**: calculated 22 measures (e.g., mean, RMS, kurtosis, etc.) for each obs.
    - **Frequency-domain features**: converted signal to the frequency domain via Fast Fourier Transform.



Exploratory Data Analysis (EDA)

- Healthy vs. failed bearing states can be visually identified by charting the vibration signal (acceleration). Note the **irregular acceleration pattern** and **large magnitude** (the y-axis 10x larger in left vs. right chart).
- Bearing failure modes and lifespan are highly variable, complicating accurate classification.





Manual Labeling

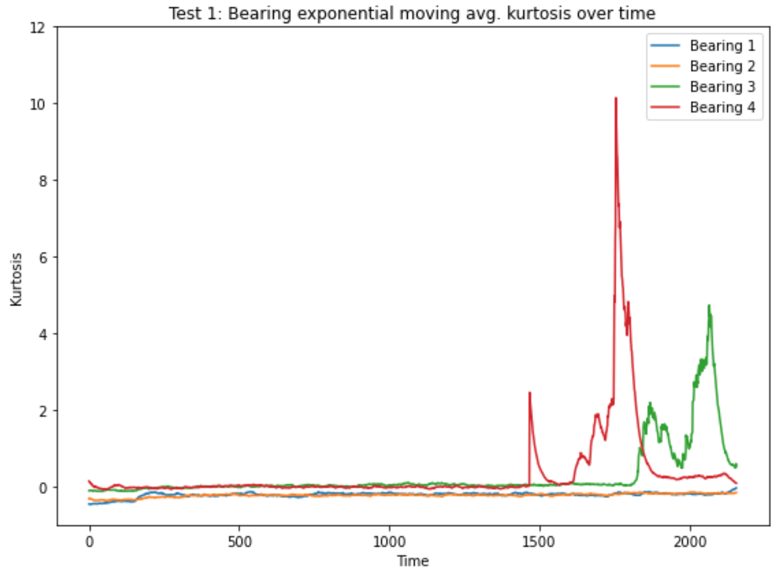
- Healthy vs. failed bearing states can be visually identified by charting the vibration signal.

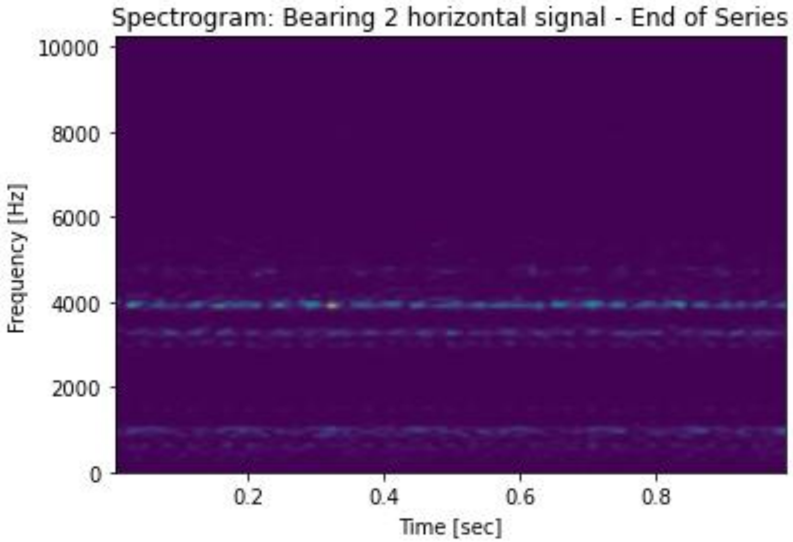
- Bearing failure modes and lifespan are highly variable, complicating accurate classification.

- Sutrisno, Oh, et. al. (2012), find that the **exponential moving average of kurtosis** of the bearing increases roughly monotonically over time, making it a good feature for prognostics and therefore for manual labeling.[5]

- Sahoo and Mohanty (2022) identify two measures as predictive – the highest magnitude frequency (frequency domain) and the **smoothed maximum acceleration** (the average of the five highest absolute acceleration measurements, time domain). Of these two, only smoothed maximum acceleration had predictive value for the IMS data.[4]

Both EMA kurtosis and smoothed maximum acceleration time-domain measures were used to apply manual labels.





(1) Semi-Supervised Learning

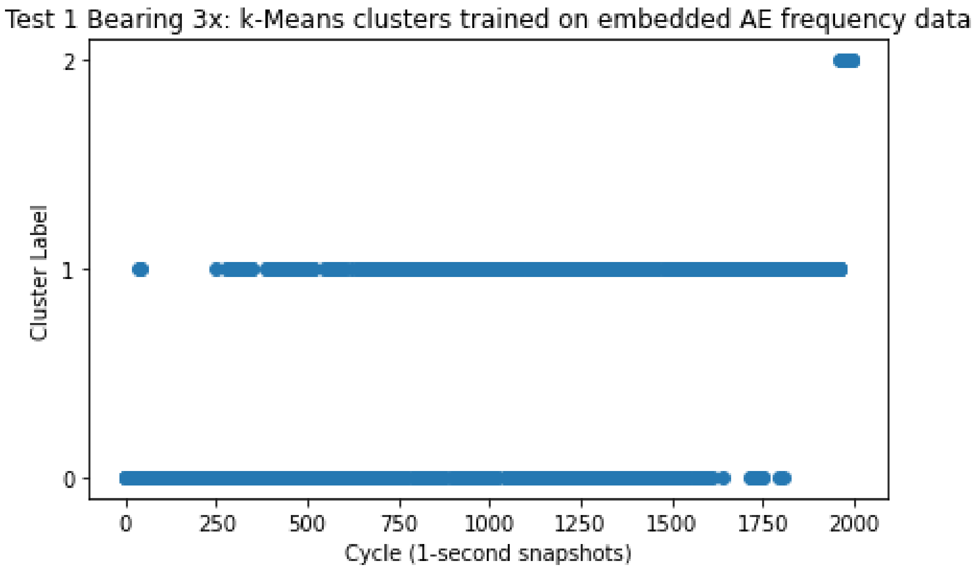
Given desire to classify a large amount of un-labeled data, **semi-supervised learning** can provide a solution.

Manually Label	Automatically Label	Supervised Classifier
Manually label a segment of the data based on pre-defined threshold values. This is expensive due to the high time and cost for human intervention, so is only performed on a small portion of the data.	**Automate labeling** of un-labelled data by training an **unsupervised clustering algorithm** to infer the correct labels for the unlabeled data, evaluating its accuracy against the manually-generated labels.	Finally, train a **supervised classifier** algorithm to predict the degradation stage given a new bearing vibration signal. This classifier is trained on the automatically-labeled data, allowing it to utilize <u>all</u> available data.

Challenges

- Initially attempted to replicate the ITUC team's semi-supervised learning approach.
  - Inputted frequency data to an unsupervised AutoEncoder neural network to reduce dimensionality from 1,025 columns to only 8 after encoding.
  - Used the exact same AutoEncoder architecture as the ITUC team from their code repository.
  - Trained a k-means clustering algorithm on the AutoEncoder output.
- K-means generated cluster labels that were almost entirely overlapping and thus did not distinguish between subsequent bearing life stages.
- As a result, changed strategy to a new time series segmentation approach.

K-means clustering output. Note the near-total overlap between cluster 0 and cluster 1 over time (x-axis). Consecutive, non-overlapping cluster labels could provide information about the bearing's life stage over time.



(2) Time series segmentation (TSS)

- **Key insight**: unsupervised clustering does not account for time dimension of signal data, which is crucial for identifying **sequential, non-overlapping bearing states**, (healthy → faulty and faulty → failed).
  - TSS solves this problem by splitting a time series into segments.
  - New **Classification Score Profile (ClaSP)** TSS implementation with associated Python package. [11]
  - ClaSP automatically partitions a time series into sequential, semantically meaningful segments that are fully interpretable to human inspection.
- TSS can function as a **standalone change detector** that automatically "learns" a bearing state change from **real-time signal data**, without need for a supervised classifier.

Manually Label

Manually label a segment of the data based on pre-defined threshold values. This is expensive due to the high time and cost for human intervention, so is only performed on a small portion of the data.

→

Automatic Segments

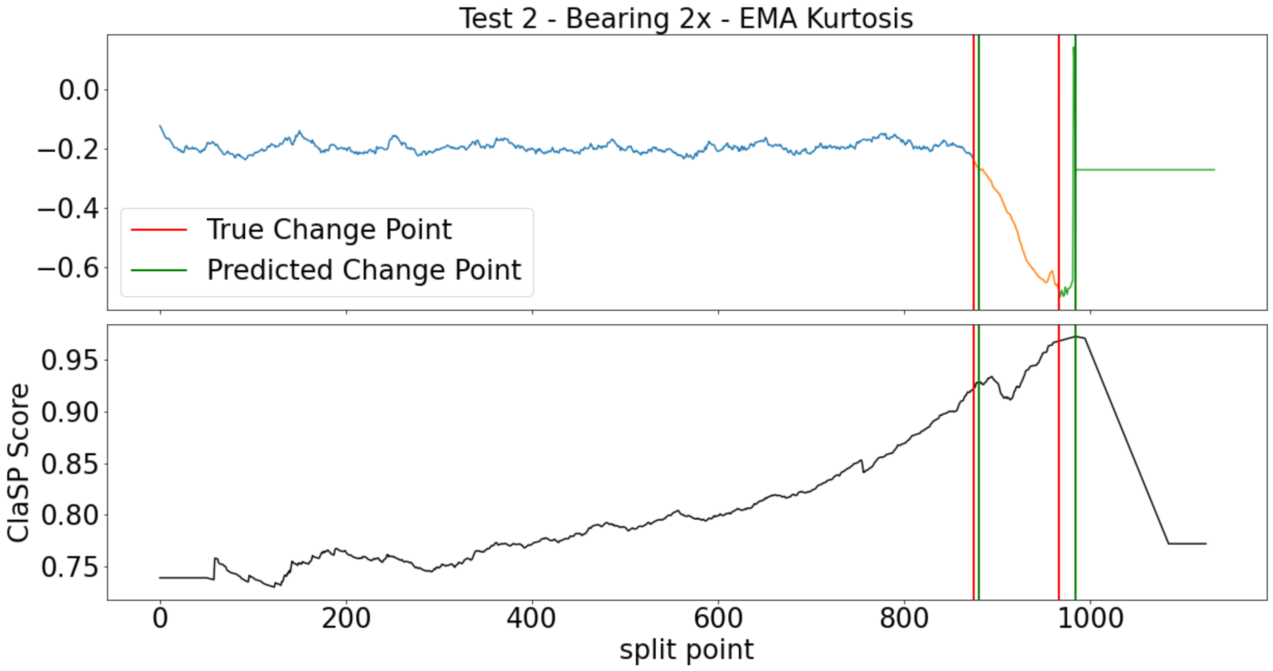
Unsupervised TSS automatically "learns" bearing state change from real-time data. When change is detected as a new predicted TS segment, the bearing's status changes from healthy to faulty, triggering maintenance.

Georgia Tech

Results

- Change first detected on avg. **85% of the way through the bearing's lifetime**.
- Best result on **full data** with **Smoothed Maximum Acceleration** measure. TSS detected change in 9 bearings, vs. all 12 identified by manual labeling – a **sensitivity rate of 75%**. Correctly detected change in 3 of the 4 bearings that ended in failure.
- Best result on **truncated data** (excl. final 3% failure state in each time series) with **EMA of Kurtosis** measure. Manual labels identified change in 6 of the 12 bearings, and TSS achieved **50% sensitivity and 83% specificity**. Correctly detected change in 2 of the 4 bearings that ended in failure.
- Also **simulate real-world deployment** by training TSS on **successively larger chunks** of data. TSS still detected change but with **higher likelihood of early detection** (too early).

*ClaSP **predicted change points** close to actual change points identified by **manual labeling**.*



Conclusions

- Further accuracy improvements are likely possible with (1) **additional research** into which time-domain features are most predictive of bearing failure; (2) **ClaSP parameter optimization**, and (3) **more data**.
- Further research required to avoid false positives in a real-world update scenario.
- Ultimately, TSS produced **automatic, interpretable, and accurate** time series segments for change detection.
- This analysis shows that early detection of bearing degradation is possible and that TSS can improve predictive maintenance.

Appendix

References

1. https://en.wikipedia.org/wiki/Prognostics
2. Lacey, Dr. S J. An overview of Bearing Vibration Analysis. Schaeffler UK Limited.
3. Dovile Juodelyte, Veronika Cheplygina, Therese Graversen, and Philippe Bonnet. 2022. Predicting Bearings Degradation Stages for Predictive Maintenance in the Pharmaceutical Industry. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22). Association for Computing Machinery, New York, NY, USA, 3107–3115. Article: https://doi.org/10.1145/3534678.3539057 and GitHub repository: https://github.com/DovileDo/BearingDegradationStageDetection
4. Biswajit Sahoo and A. R. Mohanty. 2022. Multiclass Bearing Fault Classification Using Features Learned by a Deep Neural Network. R. Karim et al. (Eds.): IAI 2021, LNME, pp. 405–414, 2022. https://doi.org/10.1007/978-3-030-93639-6_35
5. E. Sutrisno, H. Oh, A. S. S. Vasan and M. Pecht, "Estimation of remaining useful life of ball bearings using data driven methodologies," 2012 IEEE Conference on Prognostics and Health Management, Denver, CO, USA, 2012, pp. 1-7, doi: 10.1109/ICPHM.2012.6299548. https://ieeexplore.ieee.org/document/6299548
6. Condition Monitoring and Prognostics Using Vibration Signals. https://www.mathworks.com/help/predmaint/ug/condition-monitoring-and-prognostics-using-vibration-signals.html
7. Identification of Size and Location of Bearing Damage via Deep Learning. https://www.ijrrs.com/article_137575_6971fa2f93db00fa550cf9ab25f964cf.pdf
8. Unsupervised electric motor fault detection by using deep autoencoders. https://ieeexplore.ieee.org/document/8651897
9. Vibration Analysis of an Industrial Motor with Autoencoder for Predictive Maintenance. https://link.springer.com/chapter/10.1007/978-3-031-19496-2_19
10. Bearing Health monitoring based on Hilbert-Huang Transform, Support Vector Machine and Regression. https://ieeexplore.ieee.org/document/6847199
11. Ermshaus, A., Schäfer, P. & Leser, U. ClaSP: parameter-free time series segmentation. Data Mining and Knowledge Discovery 37, 1262–1300 (2023). Article: https://doi.org/10.1007/s10618-023-00923-x and GitHub repository: https://github.com/ermshaua/claspy
12. J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services (2007). Center for Intelligent Maintenance Systems (IMS), University of Cincinnati. "Bearing Data Set", NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA. Download: https://data.nasa.gov/download/brfb-gzcv/application%2Fzip

Georgia Tech

Results

*Fig. 11. TSS results on **full data**: Comparison of predicted change points vs. manually labeled change points.*

				Manual Labels			TS Segmentation: Smoothed Max (5) Acceleration				
Test	Bearing	# Obs.	Failed?	# Segments	Manual Change Points	First Change Detection as % of Total Time*	# Segments	CPs	First Change Detection as % of Total Time*	Classification Outcome**	% in Advance of Manual CP
1	1	1,996	N	2	[1973]	99%	1	[]	N/A	FN	N/A
1	2	1,996	N	2	[1961]	98%	2	[1911]	96%	TP	3%
1	3	1,996	Y	3	[1667, 1905]	84%	2	[1591]	80%	TP	4%
1	4	1,996	Y	3	[1307, 1594]	65%	2	[1307]	65%	TP	0%
2	1	984	Y	2	[975]	99%	1	[]	N/A	FN	N/A
2	2	984	N	3	[875, 966]	89%	2	[939]	95%	TP	-7%
2	3	984	N	3	[860, 981]	87%	2	[984]	100%	TP	-13%
2	4	984	N	3	[861, 982]	88%	3	[769, 910]	78%	TP	9%
3	1	6,324	N	4	[350, 5967, 6323]	6%	1	[]	N/A	FN	N/A
3	2	6,324	N	3	[6174, 6323]	98%	3	[3931, 5608]	62%	TP	35%
3	3	6,324	Y	3	[6174, 6323]	98%	2	[6084]	96%	TP	1%
3	4	6,324	N	3	[6174, 6323]	98%	2	[6125]	97%	TP	1%

*When was state change first detected, as a percentage of the experiment's total duration?

**Classification outcome for the Time Series segmentation models were evaluated against the manual labels.

TP	9
FN	3
FP	0
TN	0
Sensitivity	**75%**
Specificity	N/A

Results, cont'd.

*Fig. 12: TSS results on **truncated data**: Comparison of predicted change points vs. manually labeled change points.*

				Manual Labels			TS Segmentation: Exponential Moving Avg. of Kurtosis			
Test	Bearing	# Obs.	Failed?	# Segments	Manual Change Points	First Change Detection as % of Total Time*	# Segments	CPs	First Change Detection as % of Total Time*	Classification Outcome**
1	1	1,996	N	1	[]	N/A	1	[]	N/A	TN
1	2	1,996	N	1	[]	N/A	1	[]	N/A	TN
1	3	1,996	Y	2	[1667]	84%	1	[]	N/A	FN
1	4	1,996	Y	3	[1307, 1594]	65%	2	[1199]	60%	TP
2	1	984	Y	1	[]	N/A	2	[876]	N/A	FP
2	2	984	N	2	[875]	89%	2	[880]	89%	TP
2	3	984	N	2	[860]	87%	2	[867]	88%	TP
2	4	984	N	2	[861]	88%	1	[]	N/A	FN
3	1	6,324	N	3	[350, 5967]	6%	1	[]	N/A	FN
3	2	6,324	N	1	[]	N/A	1	[]	N/A	TN
3	3	6,324	Y	1	[]	N/A	1	[]	N/A	TN
3	4	6,324	N	1	[]	N/A	1	[]	N/A	TN

*When was state change first detected, as a percentage of the experiment's total duration?

**Classification outcome for the Time Series segmentation models were evaluated against the manual labels.

TP	3
FN	3
FP	1
TN	5
Sensitivity	**50%**
Specificity	**83%**