Comparison of Test Stand and Helicopter Oil Cooler Bearing Condition Indicators

Paula J. Dempsey
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio
Paula.J.Dempsey@nasa.gov

Jeremy Branning
RMCI, Inc.
U.S. Army Aviation Engineering Directorate
Redstone Arsenal, Alabama

Daniel R. Wade
U.S. Army Aviation Engineering Directorate
Redstone Arsenal, Alabama

Nathan Bolander
Sentient Corporation
Idaho Falls, Idaho

ABSTRACT

The focus of this paper was to compare the performance of HUMS condition indicators (CI) when detecting a bearing fault in a test stand or on a helicopter. This study compared data from two sources: first, CI data collected from accelerometers installed on two UH-60 Black Hawk helicopters when oil cooler bearing faults occurred, along with data from helicopters with no bearing faults; and second, CI data that was collected from ten cooler bearings, healthy and faulted, that were removed from fielded helicopters and installed in a test stand. A method using Receiver Operating Characteristic (ROC) curves to compare CI performance was demonstrated. Results indicated the bearing energy CI responded differently for the helicopter and the test stand. Future research is required if test stand data is to be used validate condition indicator performance on a helicopter.

INTRODUCTION

Helicopter Health Usage Monitoring Systems (HUMS) have potential for providing data to support increased service life of dynamic mechanical components in helicopter transmissions. HUMS use algorithms, referred to as condition indicators (CI), which are generated from fault patterns produced in vibration signatures when damaged components interact with their environment. CI algorithms are developed using data generated in controlled ground test environments. In order to evaluate the performance of an individual CI to detect a fault in a component, a threshold must be defined that differentiates between healthy and faulted components. When defining the threshold, there is a tradeoff between the sensitivity of the limit to indicate damage and the number of false alarms. If a limit is decreased, damage may be detected, but more false alarms may result. If a limit is increased, false alarms may decrease, but the CI will be less sensitive to damage.

To date, the majority of experiments performed to assess CI performance have been in small scale component test stands. Performing seeded fault tests on actual helicopter components in a realistic operational environment has been very limited. In order to validate the performance of a CI to detect a fault, CI data from a healthy helicopter component, faulted helicopter component, healthy test stand component and faulted test stand component is required. For this reason, its performance on a helicopter cannot be completely validated until the fault occurs on aircraft. It is not surprising that for some components, the response of a CI to a fault in a test stand is not representative of the CI response in a helicopter.

This paper begins to look at some of the issues involved when trying to compare the performance of a CI generated in a test stand to a helicopter. Existing test stand and helicopter data was available from previous work that contained test stand and helicopter data on the same faulted component. The component was the oil cooler fan bearings located in the UH-60 Black Hawk helicopter and will be discussed in more detail in the following section. CI data was collected from HUMS when pitting occurred on the races, cages, and rolling elements of the oil cooler fan bearings. CI data was also collected in a test stand on damaged oil cooler bearings removed from aircraft. Data from undamaged bearings were also collected on helicopters and the test stand. Distributions of all four datasets were analyzed. The use of Receiver Operating Characteristic (ROC) curves to compare the performance of different CI threshold values from both the test stand and helicopter was also discussed. Preliminary work published in Reference 1

Presented at the American Helicopter Society 66th Annual Forum, May 11-13, 2010, Phoenix, AZ..

discusses using ROC curves to assess CI performance. Additional details on the helicopter data and test stand data will also be discussed.

OIL COOLER ASSEMBLY FAN BEARINGS

The analysis discussed in the paper is focused on the oil cooler fan bearings located in the UH-60 helicopter. The oil cooler fan is located just aft of the main transmission, and is mounted on and driven by the tail rotor driveshaft. The oil cooler cools oil before it enters the various modules. The fan shaft is supported by two shielded single-row radial deep groove ball bearings as shown in Figure 1.

A HUMS equipped UH-60 oil cooler fan assembly has been identified as a candidate component for maintenance credits in the U.S. Army. The current Time Between Overhauls (TBO) life limit of the Oil Cooler Fan Assembly is 3240 flight hours. Reference 2 outlines an approach for obtaining Condition Based Maintenance credits for the oil cooler fan assembly. Their approach is also relevant to civil applications utilizing the FAA Advisory Circular (AC) 29–2C, Section MG–15, providing guidance for achieving airworthiness approval for installation, credit validation, and instructions for continued airworthiness (ICA) for a full range of HUMS application (Ref. 3). Spalling and wear on the fan shaft bearing was identified in Reference 2 as a failure mode that could be indicated by the HUMS and a candidate for maintenance credit and classified this failure mode as Marginal, enabling indirect evidence to be used to validate the HUMS for maintenance credit. This indirect evidence included models, HUMS database, historical teardown analysis and seeded fault tests. If seeded fault tests are used to measure the performance of condition indicators, the tests must be verified as representative of flight data. The discussion that follows will begin to identify the challenges that need to be addressed when using test stand data to validate condition indicator performance on a helicopter.

ANALYSIS METHODS

In order to calculate a CI for a specific component, vibration data is collected from a helicopter or test stand using an accelerometer mounted at a location sensitive to the component frequencies under investigation. Location and mounting can be optimized to obtain the best response, although installation is often limited to space availability on the helicopter. The data acquisition system samples the vibration data at speeds that provide sufficient vibration data for calculating asynchronous and time synchronous averaged data based on the component rotational speed. If the CI is sensitive to environmental conditions, parameters such as torque and speed must be measured while maintaining steady flight regimes. The CI used is determined by its sensitivity for detecting specific component faults using analytical and experimental data. Bearing Energy (BE) was the CI used to detect the oil cooler fan bearing fault. Bearing Energy is calculated as the Root-Sum-Square (RSS) of the asynchronous vibration spectrum filtered around select frequency bands for specific bearings (Ref. 4). For the helicopter data, this CI is calculated while the aircraft is on the ground in flat pitch at full rotor speed. A summary of the data analyzed is listed in Table 1.

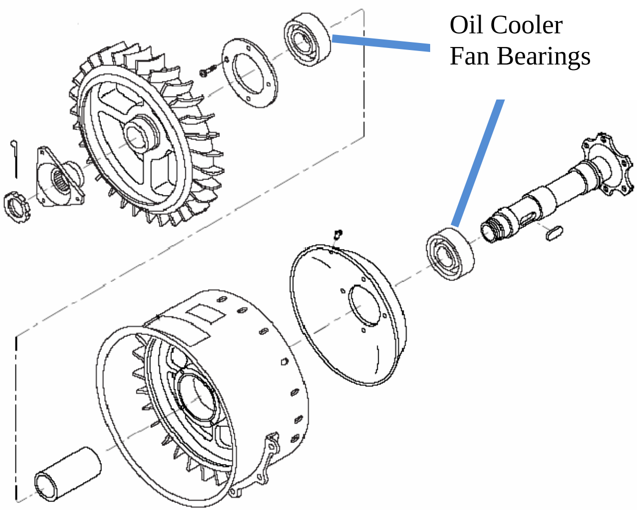

Figure 1: UH-60 Oil Cooler and Fan Bearings.

Table 1. Summary of Datasets

Data Collection Timeframe	Component Fault Type	CI	Sensor
	Honeywell VMEP—Helicopter Data Analysis I		
UH60 874 2/04-12/06	Oil cooler fan **bearings** pitting on race and cages, spalling and pitting on balls	Bearing Energy (BE)	Oil Cooler Bearing Energy
UH60 900 1/04-9/07	Oil cooler fan **bearings** pitting on race and cages, spalling and pitting on balls	Bearing Energy (BE)	Oil Cooler Bearing Energy
UH60 (40 helicopters) 12/99-6//09	NA - Healthy	Bearing Energy (BE)	Oil Cooler Bearing Energy
	Sentient Oil Cooler Test Rig		
4 bearings	corrosion and pitting on races and ball surfaces	Bearing Energy (BE) Shock Pulse Energy SPE) HP	Axial/Radial Oil Cooler Bearing
6 bearings	NA - Healthy	Bearing Energy (BE) Shock Pulse Energy SPE) HP	Axial/Radial Oil Cooler Bearing

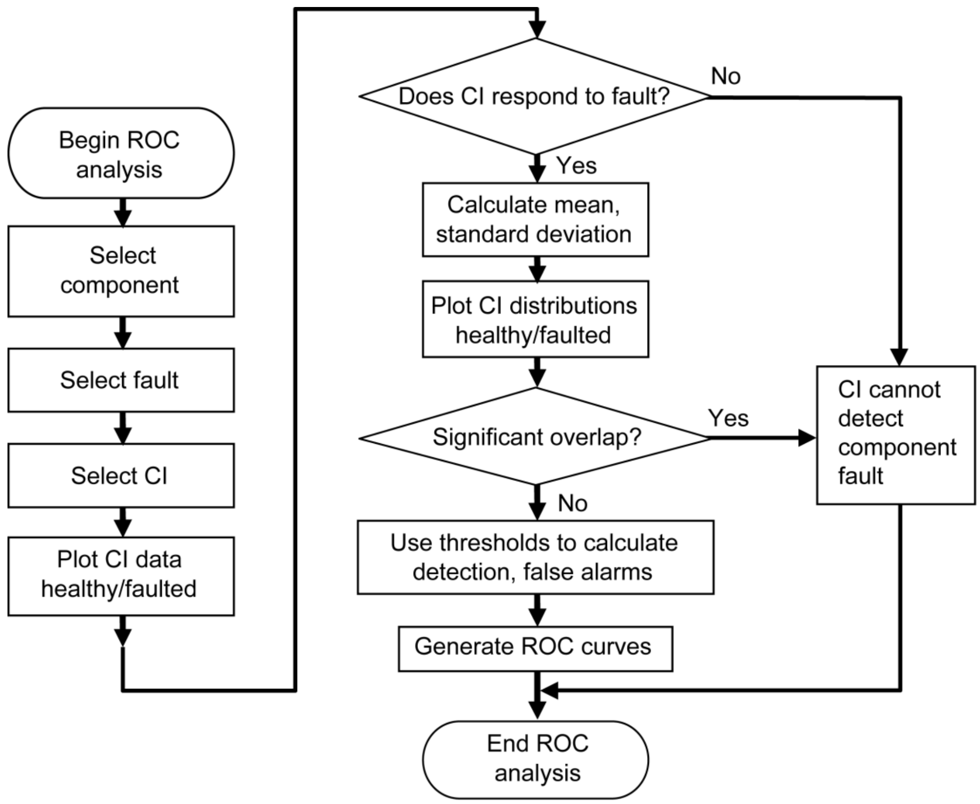

Figure 2: Flowchart of Analysis.

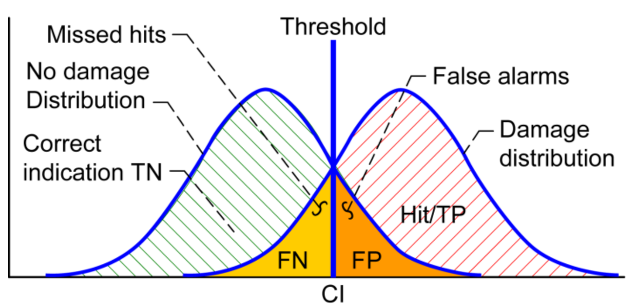

Figure 3: Probability Curves for no damage and damage.

Comparing the oil cooler bearing CI data from the helicopter and the test stand requires several steps to be completed. Details on this analysis can be found in Reference 1. A flowchart of this process is shown in Figure 2.

Per the process listed in Figure 2, statistical distributions of the healthy and faulted CI data can be used to generate ROC curves. The method of generating ROC curves to assess CI performance was taken from signal detection theory, developed to detect weak signals in a noisy environment, with applications to analyzing how decisions are made in uncertain situations (Ref. 5). ROC curves are plotted with the false alarm rate (probability of false alarm or false positive rate) on the horizontal axis (*x*) versus the hit rate (probability of detection-true alarm or true positive rate) on the vertical axis (*y*). The ROC curves can be used to evaluate thresholds since they provide a visual comparison of two or more tests on common scales at all possible thresholds independent of the test scale. Figure 3 provides a visual graph of two normal distributions used to represent a no damage response and a damage response of a CI. The threshold line separates the graph into correct indication/TN-true negative (no damage—no indication), FN-false negative (damage present—no indication), false alarms/FP-false positive (no damage—indicated) and hits/TP-true positive (damage—indicated). The probability of detection would equal the area under the damage distribution curve to the right of the threshold line. The false alarm rate would equal the area of the no damage distribution to the right of the threshold line. Interpreting the overlapping region illustrated in Figure 3 is the challenge to setting reliable thresholds based on the CI.

Per the process identified in Figure 2, a hypothesis test can be used to determine if significant overlap exists between the two datasets. A hypothesis test is defined by testing the value of a population parameter, the null hypothesis, H_o. The alternative hypothesis is the statement that must be true if the null hypothesis is false. A test statistic is used to determine if the hypothesis meets the criteria for rejecting the null hypothesis. Table 2 illustrates a hypothesis test used to assess the distributions of the two datasets. The Kolmogorov-Smirnov (KS) Test is a statistical test that can be used to test for significant overlap and determine if "no damage" and "damage" distributions significantly differ. The KS test makes no assumption about the data distributions and provides a graphical distribution of the data. The cumulative distribution function (CDF) is plotted for each dataset. The KS test uses the maximum vertical deviation between the two curves as test statistic D.

Table 2. Hypothesis of Damage/No Damage Decisions

		True State of System—Health of Component	
		H_o is True $CI_{damage} = CI_{no\ damage}$ No Damage	H_o is False $CI_{damage} \neq CI_{no\ damage}$ Damage
Decision	Reject H_o	False Positive	True Positive
	Indicate Damage	False Alarms	Hits
	Fail to reject H_o	True Negative	False Negative
	Indicate No Damage	Correct Indication	Missed Hits

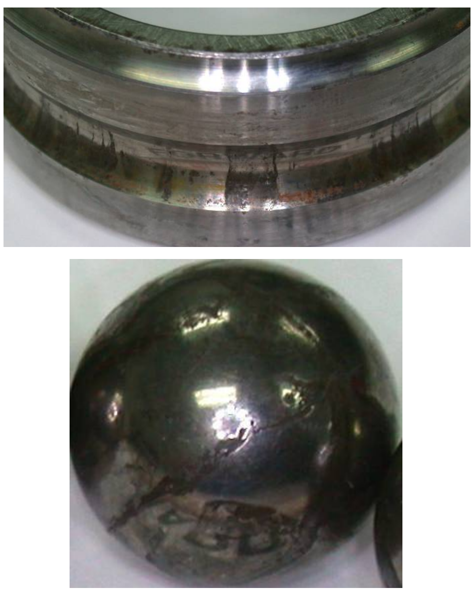

Figure 4: Oil Cooler Fan Bearing Damage (Ref. 4).

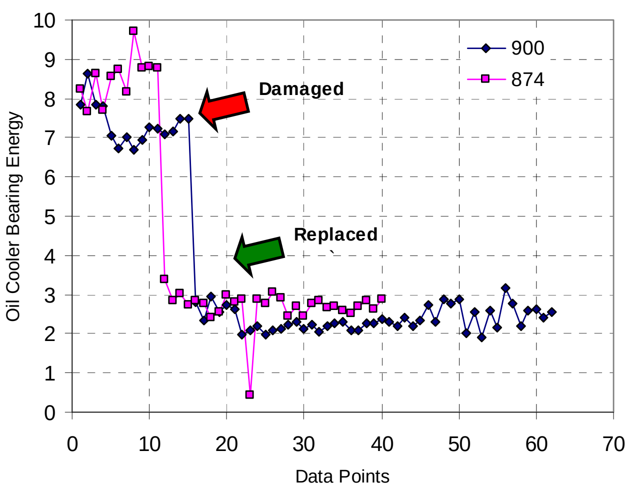

Figure 5: Bearing Energy Values for Oil Cooler Fan Bearings.

HELICOPTER DATA ANALYSIS

CI data were collected from an onboard commercial HUMS system, Vibration Management Enhancement Program (VMEP), on two Black Hawk helicopters when pitting occurred on the races, cages, and rolling elements of the oil cooler fan bearings (Ref. 4). A photo of the damage to the race and balls is shown in Figure 4. In addition to the two helicopters with faults, CI data from forty helicopters with healthy oil coolers were analyzed.

Bearing Energy values before and after bearing replacement are plotted in Figure 5 versus the number of data points collected over a given time period. For helicopter 874, data was collected from February 2004 until December 2006. For helicopter 900, data was collected from January 2004 until September 2007. The arrows indicate data collected when the bearing was damaged, then data collected after the bearing was replaced. CI values exceeded 6 during the time period when the components were damaged and dropped to below 4 when the bearings were replaced. Preliminary review of the bearing energy data shown in Figure 5 indicates the CI responded differently for damaged and healthy (replaced) bearings.

The forty additional helicopters that did not experience maintenance actions were analyzed to compare the CI values for these helicopters with undamaged bearings to the two helicopters with damaged bearings. Histograms were generated using BE values from the two helicopters with damaged bearings and for the helicopters with no oil cooler bearing damage. From these values two distributions, "no damage" and "damage," were plotted and are shown in Figure 6. If significant overlap occurred in these two plots, this CI would not be a good choice for differentiating between a healthy and faulted components and applying the ROC analysis would not be beneficial. The CDF for the no damage and damage dataset is also shown in Figure 6. Per the KS test, the distance statistic D of .9951 indicates the two data sets are not from the same distribution.

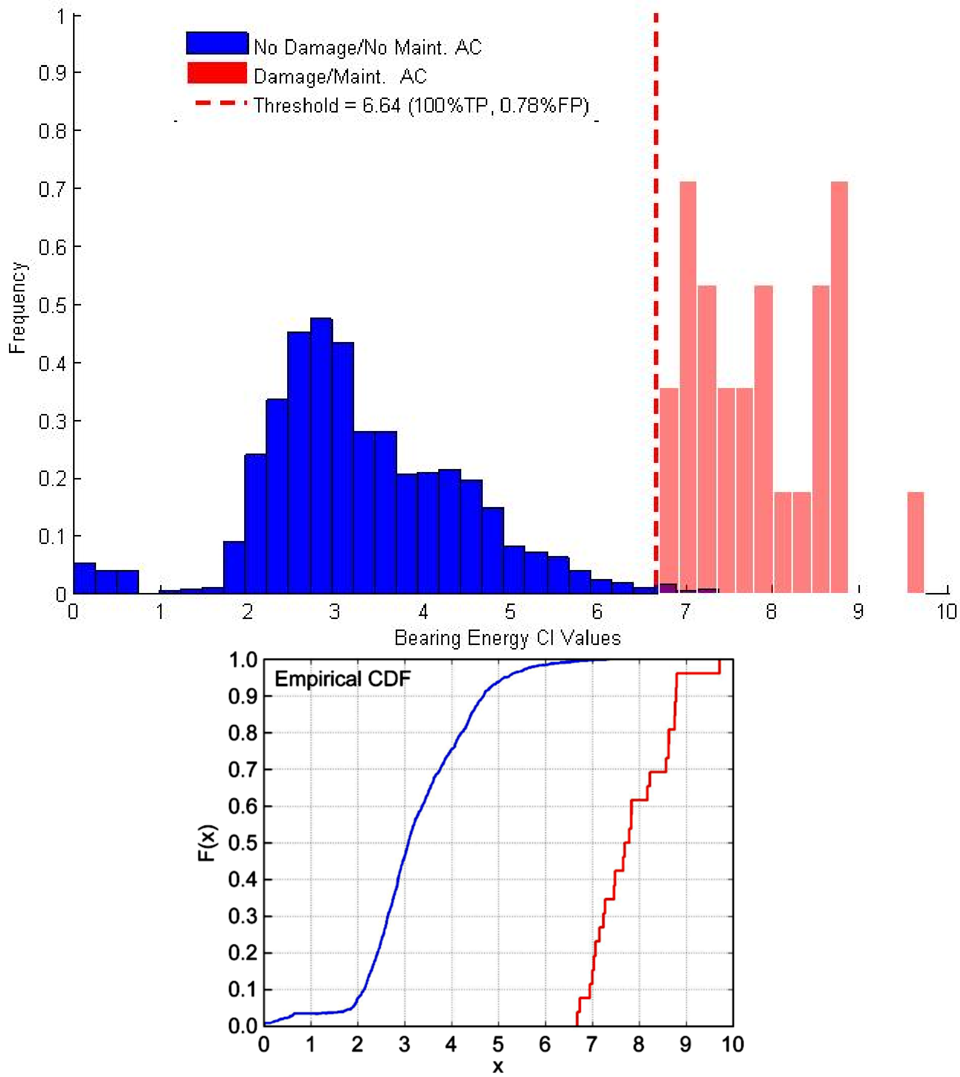

Figure 6: UH-60 Bearing Energy Probability Curves.

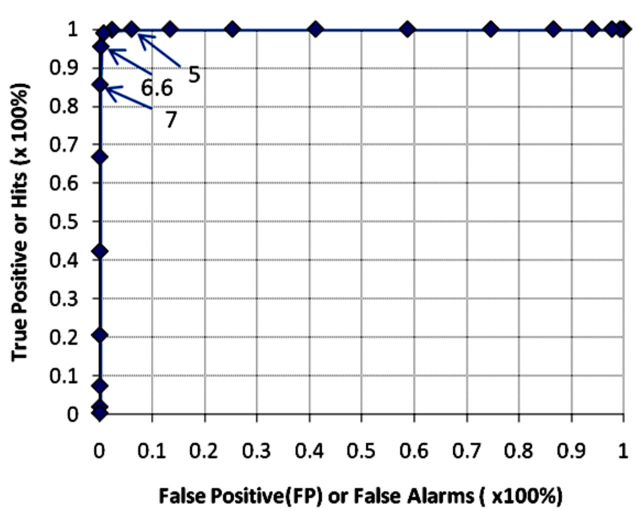

Figure 7: ROC Curves for UH-60 Bearing Energy Data.

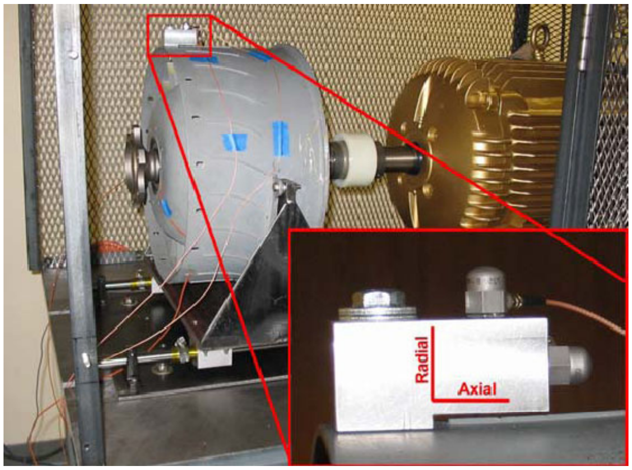

Figure 8: Oil Cooler Fan Bearing Test Stand.

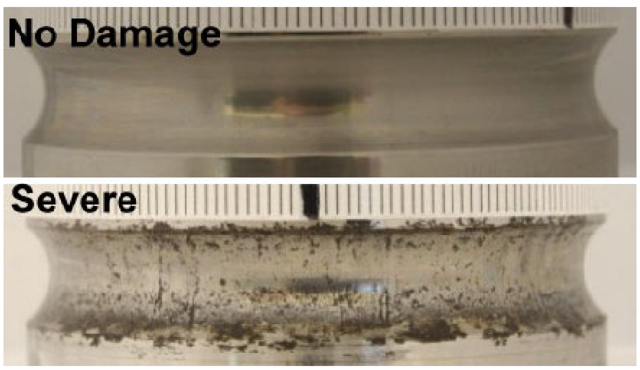

Figure 9: Bearing severity scale.

Threshold values used to generate ROC curves were identified by determining the highest fault hit rate (true positive-TP) with the smallest number of false alarms (false positive-FP). A threshold of 6.6 was selected providing 100% TP and 0.78% FP. This is shown as a dashed line in Figure 6. Using the hit rate (TP) and false alarm rate (FP) for thresholds ranging from 0 to 8, a ROC curve was plotted and is shown in Figure 7. The ROC curve shows the tradeoff between hit rates and false alarm rates for different thresholds. A threshold is then selected based on the level of risk (false alarms or missed detection) accepted for this specific component.

TEST STAND DATA ANALYSIS

CI data were also collected from oil cooler bearings in a test stand designed and constructed to test the entire Black Hawk oil cooler assembly. The oil coolers were provided by the U.S. Army from fielded helicopters (Ref. 6). All the oil cooler bearings tested were from different aircraft. The oil cooler fan assemblies were removed based on their Time Between Overhauls (TBO) service life limit. Figure 8 shows the oil cooler test stand and the location of the two accelerometers used for CI data. CI data were calculated from both the axial and radial mounted accelerometers. The vibration samples were taken every 5 minutes resulting in approximately seven independent samples per each 30 minute test run.

Two damage classes were defined: no damage and severe. Figure 9 provides a visual example of the severity rating. Based on the tear down and inspection of the used bearings, the dominant fault modes of the oil cooler bearings were corrosion and subsequent pitting of the races and balls. Six bearings were identified as “no damage” having no observable damage to the bearings or races. Four bearings, identified as “Severe”, had corrosion and pitting that covered the entire length of the races and substantial portions of the ball surfaces.

The CI calculated for the helicopters and discussed in the previous section, Bearing Energy (BE), was calculated for the oil cooler bearings when installed in the test stand. Two additional CIs were also calculated as the Root-Sum-Square (RSS) of the asynchronous vibration spectrum, but were filtered at different filtered frequency bands: BE (50 to 950 Hz), Shock Pulse Energy–SPE (2500 to 7500 Hz) and HP (10 K to 20 KHz). BE, SPE and HP values calculated from the radial and axial accelerometers for the bearings with no damage were compared to the bearings with severe damage in Figure 10. The x-axis identifies the number of readings collected over a time period for the six bearings tested. For the “no damage” data set, bearings were changed at data point 13, 25, 37, 49, 1nd 61. For the “damage” data set, bearings were changed at data point 13, 25 and 31. No overlap of the SPE and HP data was observed for the no damage and severe damage bearings. However, overlap was observed for the BE values for both the radial and axial accelerometers. Further analysis found that the test stand had resonant frequencies that were excited between the frequency bands (50 to 950 Hz), increasing the calculated BE values. This is an important consideration when developing condition indicators in a test stand to be used on a helicopter. In this case, the dynamics of the test stand decreased the performance of the BE for this component fault. Calculating SPE and HP values for the helicopter and comparing them to the test stand was not possible because the frequency data collected in the helicopter at the time of the fault were limited to 1170 Hz. SPE was later added to the VMEP system for detection of oil cooler bearing faults.

Histograms generated for the BE, SPE and HP data are shown in Figure 11. If significant overlap occurred in these two plots, this CI would not be a good choice for differentiating between a healthy and faulted component and applying the ROC analysis would not be beneficial. Significant overlap occurred for BE for both the axial and radial accelerometers. BE was not be a good choice for differentiating between healthy and damaged bearings in the test stand. However, no overlap occurred between the “no damage” and “severe damage” SPE and HP values indicating a threshold could be defined with 100% detection rate 0% false alarm rate at a value of 1.5. Both SPE and HP were good choices to differentiate between healthy and damaged bearings in this test stand.

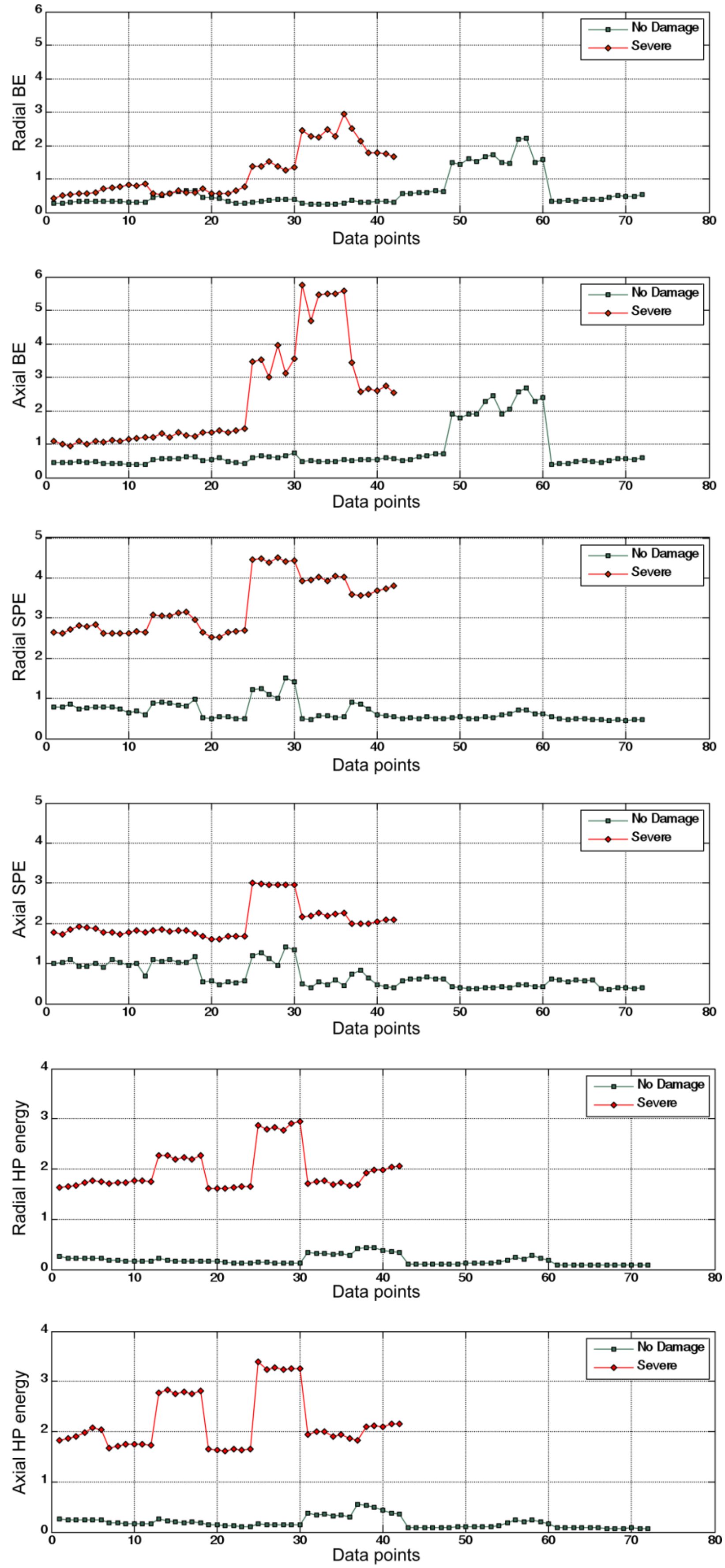

Figure 10: BE, SPE and HP for Test Stand Oil Cooler Fan Bearings.

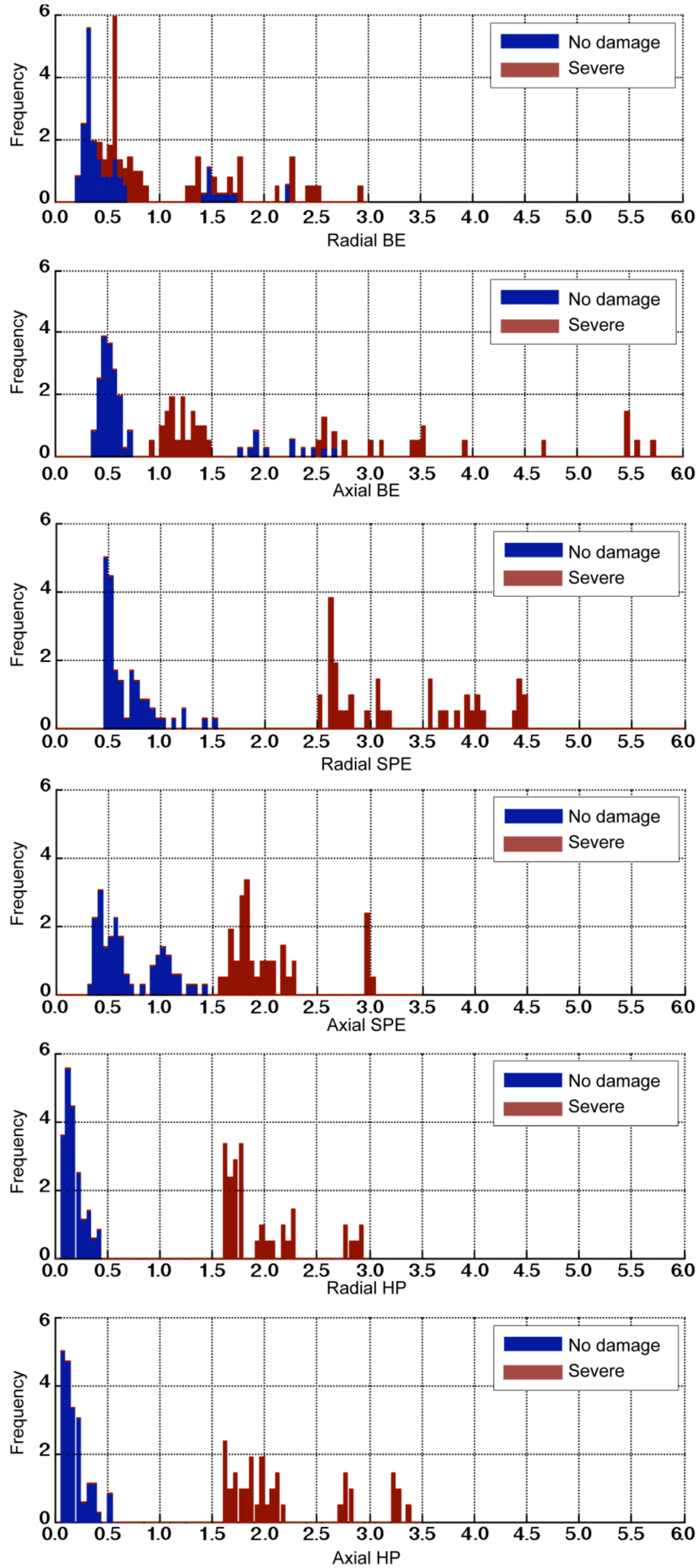

Figure 11: BE, SPE and HP Frequency Distributions for Test Stand Oil Cooler Fan Bearings.

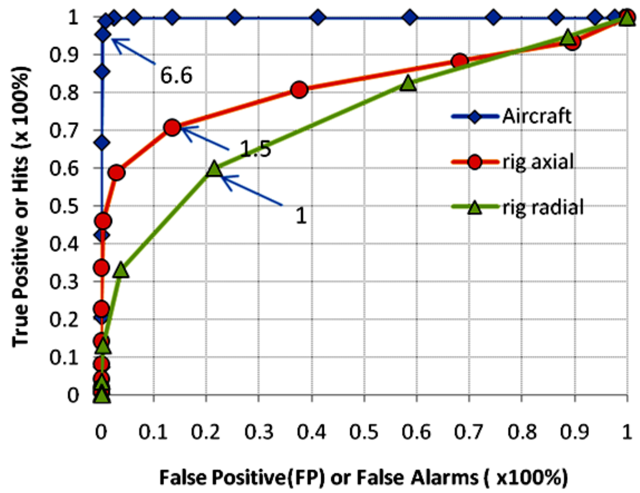

Figure 12: ROC Curves for UH-60 and Test Stand Bearing Energy Data.

Comparing the BE values from the helicopter and test stand in Figure 12, BE performed better in the helicopter differentiating between the no damage and damaged bearings, due to low frequency resonances in the test stand that caused false alarms for this algorithm These resonances appeared within the BE filtered bands for the test stand tests, causing poor performance of this CI. Due to the limited helicopter vibration data at the higher frequencies, the SPE and HP data sets could not be compared to the helicopter.

SUMMARY AND CONCLUSIONS

The focus of this paper was to compare the performance of HUMS condition indicators (CI) when detecting a bearing fault in a test stand or on a helicopter. This study compared oil cooler bearing CI data from two UH-60 Black Hawk helicopters and data collected from healthy and faulted oil cooler bearings that were removed from fielded helicopters and installed in a test stand. Receiver Operating Characteristic (ROC) curves were used to compare CI performance on the condition indicator bearing energy. Based on the results of comparing the condition indicators' performance in detecting a damaged bearing on a helicopter and a test stand, the following conclusions can be made for this application:

(1) *VMEP:* The response of BE to the damaged oil cooler installed in a test stand was different from the response in a helicopter equipped with a VMEP HUMS, due to resonances in the test stand that appeared within the BE filtered bands.

(2) *VMEP:* SPE and HP could not be calculated for the VMEP system due to limited spectrum data.

(3) *Test Stand*: The sensor mounting position in the test stand affected the CI value. The axial sensor was more sensitive to BE, while the radial sensor was more sensitive to SPE.

(4) The threshold values that provided the highest detection rate and the lowest false alarm rate were different for the aircraft and the test stand. BE values were higher in the helicopter when responding to the damaged bearings

(5) Both healthy and faulted data sets are required in tests stands and helicopters to evaluate CI performance and verify it can be maintained when damage occurs.

Many variables affect the ability of a CI to respond to a fault in a dynamic mechanical component. These factors can include sensor type, mounting, location, signal processing, structural dynamics, flight regimes, and history of the component. The relationship between all of these variables is complex, not well defined and attempts to develop physics based models of these relationships at the component level do not perform well when applied to complex mechanical systems. In this case, the same component and fault was monitored in a helicopter and in a test stand, but the response was significantly different due to the environment. Undefined limitations and constraints of a test stand can make it challenging to use a test stand to develop performance metrics for helicopter diagnostic tools. When defining a rig test to validate a CI for a specific fault for maintenance credit, it is critical to define environmental factors in the test stand and the field that can limit its performance. In this example, dynamics in the test stand at the frequencies that occur when the oil cooler bearing fails masked those signatures decreasing its performance. Additional research is required to develop seeded fault tests in test stands representative of fielded faults and to better understand the effect of helicopter flight regimes on CI performance.

ACKNOWLEDGMENTS

The authors would like to thank the U.S. Army Black Hawk PM, Apache PM, and AMCOM G-3 CBM office located at Redstone Arsenal, Mr. Bert Smith of the Army Aviation Applied Technology Directorate at Ft. Eustis and Traci Stadtmueller of the FAA Technical Center for their assistance and support of this work.

REFERENCES

1. Dempsey, P.J., Keller, J.A., Wade, D.R.: Signal Detection Theory Applied to Helicopter Transmission Diagnostic Thresholds. NASA/TM—2008-215262; AMRDEC PAO Control Number FN 3597.
2. Suggs, D.T. and Wade, D.R. Vibration Based Maintenance Credits for the UH-60 Oil Cooler Fan Assembly. Proceedings from the AHS CBM Specialist Meeting. Feb. 2007.
3. Federal Aviation Authority, Advisory Circular, "Airworthiness Approval of Rotorcraft Health Usage Monitoring Systems (HUMS)," PS-ASW100-1999-00063, July 15, 1999.
4. Keller, J.A., Branhof, R., Dunaway, D. and Grabill, P.: Examples of Condition Based Maintenance with the Vibration Management Program. AHS 61st Forum, Grapevine, TX, June 1-3, 2005.
5. Wickens, T.D.: Elementary Signal Detection Theory. Oxford University Press, 2002.
6. Hartin, J.R. Helicopter Drive Train Prognostics and Health Management. http://www.sentientscience.com/pdf/BlackhawkSummary_3.pdf accessed 1/23/09.