# Gear Damage Detection Using Oil Debris Analysis

**Paula Dempsey**

**Abstract**

The purpose of this paper was to verify, when using an oil debris sensor, that accumulated mass predicts gear pitting damage and to identify a method to set threshold limits for damaged gears. Oil debris data was collected from eight experiments with no damage and eight with pitting damage in the NASA Glenn Research Center's spur gear fatigue rig. Oil debris feature analysis was performed on this data. Video images of damage progression were also collected from six of the experiments with pitting damage. During each test, data from an oil debris sensor was monitored and recorded for the occurrence of pitting damage. The data measured from the oil debris sensor during experiments with no damage was used to identify membership functions, which are required to build a simple fuzzy-logic model. Using fuzzy-logic techniques and the oil debris data, threshold limits were defined that discriminate between stages of pitting wear. Results indicate that accumulated mass combined with fuzzy-logic analysis techniques is a good predictor of pitting damage on spur gears.

**Introduction**

One of NASA's current goals, the National Aviation Safety Goal, is to reduce the aircraft accident rate by a factor of five within 10 years and a factor of 10 within 25 years. One of the leading factors in fatal aircraft accidents is loss of control in flight, which can occur due to flying in severe weather conditions, pilot error and vehicle/system failure. Focusing on helicopters' system failures, an investigation in 1989 found that 32% of helicopter accidents due to fatigue failures were caused by damaged engine and transmission components (Ref. 1).

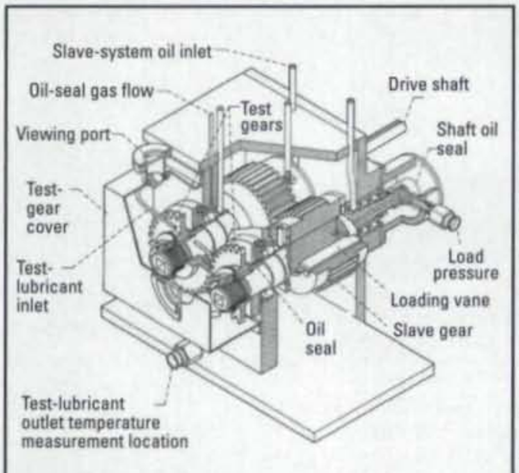


*Figure 1—Spur gear fatigue test rig.*

In more recent statistics, of the world total of 192 turbine helicopter accidents in 1999, 28 were directly due to mechanical failures with the most common failure in the drive trains of gearboxes (Ref. 11).

A study published in July 1998, in support of the National Aviation Safety Goal, recommended areas most likely to reduce rotorcraft fatalities in the next 10 years. The study of 1,168 fatal and nonfatal accidents that occurred from 1990–1996 found that, after human factor-related causes of accidents, the next most frequent cause of accidents was due to various system and structural failures (Ref. 2). Loss of power in flight caused 26% of this type of accident and loss of control in flight caused 18% of this type of accident. The technology area recommended by this study for helicopter accident reduction was helicopter health and usage monitoring systems (HUMS) capable of predicting imminent equipment failure for on-condition maintenance and more advanced systems capable of warning pilots of impending equipment failures.

Helicopter transmission diagnostics are an important part of a helicopter health monitoring system because helicopters depend on the powertrain for propulsion, lift and flight maneuvering. In order to predict transmission failures, the diagnostic tools used in the HUMS must provide real-time performance monitoring of aircraft operating parameters and must demonstrate a high level of reliability to minimize false alarms. Various tools exist for diagnosing damage in helicopter transmissions, the most common being vibration tools. Using vibration data collected from gearbox accelerometers, algorithms are developed to detect when gear damage has occurred (Refs. 16 and 20). Oil debris is also used to identify abnormal wear-related conditions of transmissions. Oil debris monitoring for gearboxes consists mainly of off-line oil analysis or plug-type chip detectors. And, although not commonly used for gear damage detection, many engines have on-line oil debris sensors for detecting the failure of rolling element bearings. These on-line, inductance-type sensors count the number of particles, measure their approximate sizes, then calculate an accumulated mass (Ref. 10).

The goal of future HUMS is to increase reliability and decrease false alarms. HUMS are not yet capable of real-time, on-line health monitoring. Current data collected by HUMS is processed after the flight and is plagued with high false alarm rates and undetected faults. The current fault detection rate of commercially available HUMS through vibration analysis is 60%. False warning rates average one per 100 flight hours (Ref. 17). This is due to a variety of reasons. Vibration-based systems require extensive interpretation by trained diagnosticians. Operational effects can adversely impact the performance of vibration diagnostic parameters and result in false alarms (Refs. 5 and 3). Oil debris sensors also require expert analysis of data. False alarms of oil debris technologies are often caused by nonfailure debris. This debris can bridge the gap of plug-type chip detectors. Inductance-type oil debris sensors cannot differentiate between fault and no-fault sourced data (Ref. 8).

Several companies manufacture on-line, inductance-type oil debris sensors that measure debris size and count particles (Ref. 10). New oil debris sensors are also being developed that measure debris shape and size, and the shape is used to classify the failure mechanism (Ref. 8). The oil debris sensor used in this analysis was selected for several reasons. The first three reasons were sensor capabilities, availability and researcher experience with this sensor. Results from preliminary research indicate the debris mass measured by the oil debris sensor showed a significant increase when pitting damage began to occur (Ref. 4).

This sensor has also been used in aerospace applications for detecting bearing failures in aerospace turbine engines. From the manufacturers' experience with rolling element bearing failures, an equation was developed to set warning and alarm transmissions. A modified version of this

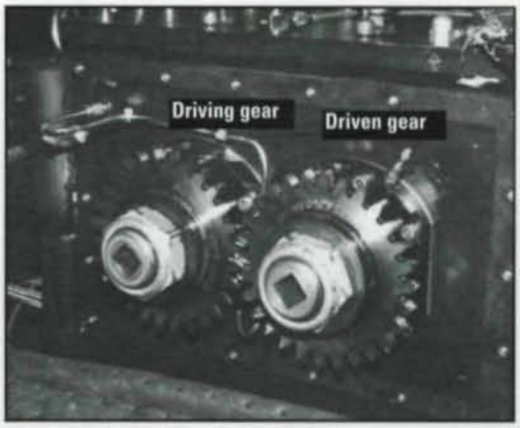


*Figure 2—Spur gear fatigue rig gearbox.*

**Table 1–Oil debris particle size ranges.**

| Bin | Bin range, μm | Average size, μm | Bin | Bin range, μm | Average size, μm |
|---|---|---|---|---|---|
| 1 | 125–175 | 150 | 9 | 525–575 | 550 |
| 2 | 175–225 | 200 | 10 | 575–625 | 600 |
| 3 | 225–275 | 250 | 11 | 625–675 | 650 |
| 4 | 275–325 | 300 | 12 | 675–725 | 700 |
| 5 | 325–375 | 350 | 13 | 725–775 | 750 |
| 6 | 375–425 | 400 | 14 | 775–825 | 800 |
| 7 | 425–475 | 450 | 15 | 825–900 | 862.5 |
| 8 | 475–525 | 500 | 16 | 900–1,016 | 958 |

sensor has been developed and installed in an engine's nose gearbox and is currently being evaluated for an operational AH-64 helicopter (Ref. 10), which is Boeing Co.'s Apache attack helicopter. Due to limited access to oil debris data collected by this type of sensor from gear failures, no such equation is available that defines oil debris threshold limits for damaged gears.

The objective of the work reported herein is to first identify the best feature for detecting gear pitting damage from a commercially available on-line oil debris sensor. Then, once the feature is defined, the objective is to identify a method to set threshold limits for different levels of pitting damage to gears. The oil debris data analysis was performed on gear damage data collected from an oil debris monitor in the NASA Glenn Research Center's spur gear fatigue rig.

### Test Procedure

Experimental data was recorded from tests performed in the NASA Glenn rig (Ref. 16). This rig is capable of loading gears, then running them until pitting failure is detected. A sketch of the test rig is shown in Figure 1. Torque is applied by a hydraulic loading mechanism that twists one slave gear relative to its shaft. The power required to drive the system is only enough to overcome friction losses in the system (Ref. 13). The test gears are standard spur gears having 28 teeth, 8.89 cm pitch diameters and 0.64 cm face widths. The test gears are run offset to provide a narrow


**Paula Dempsey, Ph.D.,**
*is an aerospace research engineer employed by the NASA Glenn Research Center in Cleveland, OH. She joined NASA in 1991, after performing experimental and analytical research work in the area of cryogenic fluid management systems for a government contractor from 1987–1991. She continued her cryogenic research until 1995, then served as lead project engineer for several aeronautics programs. She is currently performing research in the area of monitoring and diagnostics for helicopter drive systems and propulsion systems for transport aircraft.*

effective face width to maximize gear contact stress while maintaining an acceptable bending stress. Offset testing also allows four tests on one pair of gears. Two filters are located downstream of the oil debris monitor to capture the debris after the sensor measures it.

Fatigue tests were run in a manner that allows damage to be correlated to the oil debris sensor data. For these tests, run speed was 10,000 rpm and applied torque was 72 N-m and 96 N-m. Prior to collecting test data, the gears were run-in for one hour at a torque of 14 N-m. The data measured during this run-in was stored, then the oil debris sensor was reset to zero at the start of the loaded test. Test gears were inspected periodically for damage either manually or using a micro-camera connected to a videocassette recorder and monitor. The video inspection did not require gearbox cover removal. When damage was found, it was documented and correlated to the test data based on a reading number. Reading numbers are equivalent to minutes and can also be interpreted as mesh cycles equal to reading numbers multiplied by $10^4$. In order to document tooth damage, reference marks were made on the driving and driven gears during installation to identify tooth 1. The mating teeth numbers on the driving and driven gears were then numbered from this reference. Figure 2 identifies the driving and driven gears with the gearbox cover removed.

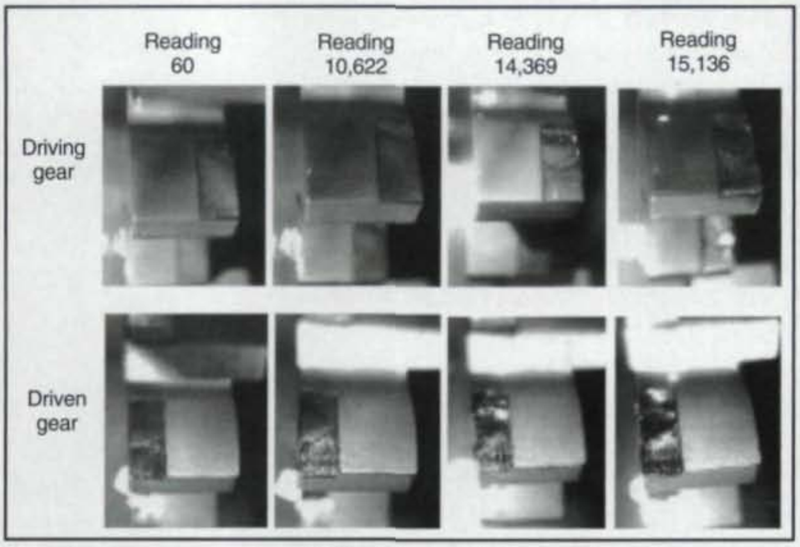


*Figure 3—Damage progression of driving/driven tooth 6 for experiment 1.*

Data was collected once per minute from oil debris, speed and pressure sensors installed on the test rig using the programs ALBERT, Ames-Lewis Basic Experimentation in Real Time, co-developed by NASA Glenn and NASA Ames Research Center. Oil debris data was collected using a commercially available oil debris sensor that measures the change in a magnetic field caused by passage of a metal particle where the amplitude of the sensor output signal is proportional to the particle mass. The sensor counts the number of particles, measures their approximate sizes (125–1,016 μm) and calculates an accumulated mass (Ref. 9). Shaft speed was measured by an optical sensor once per shaft revolution. Load pressure was measured using a capacitance pressure transducer.

The principal focus of this research is detection of pitting damage on spur gears. Pitting is a fatigue failure caused by exceeding the surface fatigue limit of the gear material. Pitting occurs when small pieces of material break off from the gear surface, producing pits on the contacting surfaces (Ref. 19). Gears are run until pitting occurs on several teeth. Pitting was detected by visual observation through periodic inspections on two of the experiments with pitting damage. Pitting was detected by a video inspection system on six of the experiments with pitting damage. Two levels of pitting were monitored, initial and destructive pitting. Initial pitting is defined as pits less than 0.04 cm in diameter and covering less than

**Table 2—Experiments with video inspection.**

| Experiment 1 | | Experiment 2 | | Experiment 3 | | Experiment 4 | | Experiment 5 | | Experiment 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Reading # | Mass, mg | Reading # | Mass, mg | Reading # | Mass, mg | Reading # | Mass, mg | Reading # | Mass, mg | Reading # | Mass, mg |
| 60 | 1.003 | 1,573 | 3.285 | 58 | 0 | 64 | 0 | 62 | 0 | 60 | 0 |
| 120 | 1.418 | 2,199 | 8.934 | 2,669 | 8.69 | 150 | 2.233 | 1,405 | 4.214 | 2,810 | 3.192 |
| 1,581 | 5.113 | 2,296 | 16.267 | 2,857 | 11.889 | 378 | 8.297 | 2,566 | 7.413 | 2,885 | 6.396 |
| 10,622 | 12.533 | 2,444 | 26.268 | 3,029 | 14.148 | 518 | 9.462 | 4,425 | 10.811 | 2,957 | 8.704 |
| 14,369 | 15.475 | | | | | 2,065 | 12.132 | | | 9,328 | 11.692 |
| 14,430 | 22.468 | | | | | 2,366 | 13.977 | | | 12,061 | 14.365 |
| 14,512 | 24.586 | | | | | 3,671 | 17.361 | | | 12,368 | 22.851 |
| 14,688 | 28.451 | | | | | 4,655 | 23.12 | | | | |
| 14,846 | 30.686 | | | | | 4,863 | 26.227 | | | | |
| 15,136 | 36.108 | | | | | | | | | | |

*Note: Highlighted cells identify reading and mass when destructive pitting was first observed.

approximately 25% of tooth contact area. Destructive pitting is more severe and is defined as pits greater than 0.04 cm in diameter and covering more than approximately 25% of tooth contact area. If not detected in time, destructive pitting can lead to catastrophic transmission failure if the gear teeth crack.

### Discussion of Results

The analysis discussed in this section is based on oil debris data collected during 16 experiments, 8 of which resulted in pitting damage. The oil debris sensor records counts of particles in bins set at particle size ranges measured in microns. The particle size ranges and average particle size are shown in Table 1. The average particle size for each bin is used to calculate the cumulative mass of debris for the experiment. The shape of the average particle is assumed to be a sphere with a density of approximately 7,922 kg/m$^3$.

Experiments 1–6 were performed with the video inspection system installed on the rig. Table 2 lists the reading numbers and the measured oil debris masses at those readings. The highlighted cells for each experiment identify the reading number and the mass measured when destructive pitting was first observed on one or more teeth. As this table shows, the amount of mass varied significantly for each experiment. A representative sample of the images obtained from the video damage progression system is shown in Figure 3. The damage progression of tooth 6 on the driving and driven gears for experiment 1 for selected readings is shown in this figure. The damage is shown on less than half of the tooth because the test gears are run offset to provide a narrow effective face width to maximize gear contact stress.

Experiments 7 and 8 were performed with visual inspection. Table 3 lists the reading numbers when inspection was performed and the measured oil debris masses at these readings. Only initial pitting occurred during experiment 7. During experiment 8, initial pitting was observed at reading 5,181 and destructive pitting at reading 5,314.

No gear damage occurred during experiments 9–16. Oil debris mass measured at test completion is listed in Table 4. At the completion of experiment 10, 5.453 mg of debris was measured, yet no damage occurred. This result is more than the debris measured during experiment 7 (3.381 mg) when initial pitting was observed. This result and observations made from the data collected during experiments when damage occurred made it obvious that simple linear correlations could not be used to obtain the features for damage levels from the oil debris data.

**Table 3—Experiments with visual inspection.**

| Experiment 7 | | Experiment 8 | | Pitting Damage |
|---|---|---|---|---|
| Reading # | Mass, mg | Reading # | Mass, mg | |
| 13,716 | 3.381 | 5,181 | 6.012 | Initial |
| | | 5,314 | 19.101 | Destructive |

**Table 4–Oil debris masses at completion of experiments with no damage.**

| Experiment | Reading # | Mass, mg | Experiment | Reading # | Mass, mg |
|---|---|---|---|---|---|
| 9 | 29,866 | 2.359 | 13 | 25,259 | 3.159 |
| 10 | 20,452 | 5.453 | 14 | 5,322 | 0 |
| 11 | 204 | 0.418 | 15 | 21,016 | 0.125 |
| 12 | 15,654 | 2.276 | 16 | 21,446 | 0.163 |

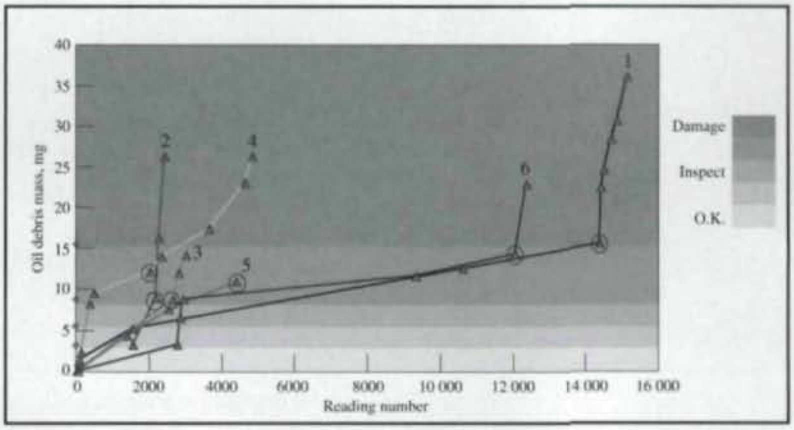


*Figure 4—Oil debris masses at different damage levels.*

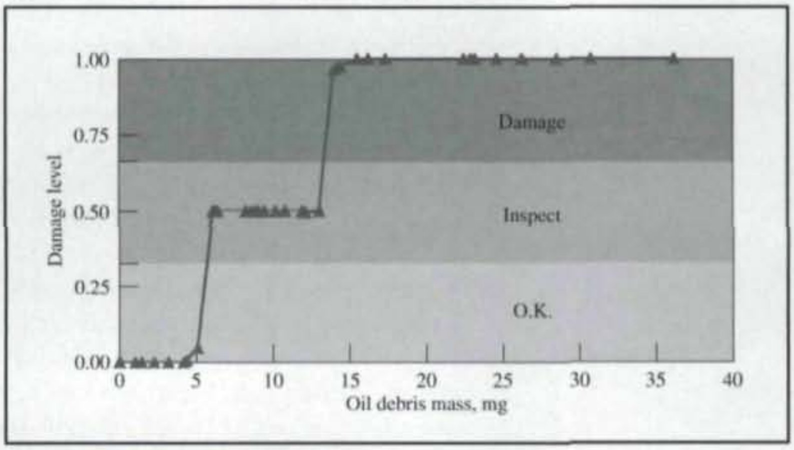


*Figure 5—Output of fuzzy-logic model.*

Prior to discussing methods for feature extraction, it may be beneficial for the reader to get a feel for the amount of debris measured by the oil debris sensor and the amount of damage to one tooth. Applying the definition of destructive pitting, 25% of tooth surface contact area for one tooth for these experiments is approximately 0.043 cm$^2$. A 0.04 cm diameter pit, assumed spherical in shape, is equivalent to 0.26 mg of oil debris mass. This mass is calculated based on the density used by the sensor software for calculating mass. If 0.04 cm diameter pits densely covered 25% of the surface area of one tooth, it would be equivalent to approximately 9 mg. Unfortunately, damage is not always densely distributed on 25% of a single tooth, but is distributed across many teeth, making accurate measures of material removed per tooth extremely difficult.

Several predictive analysis techniques were reviewed to obtain the best feature to predict damage levels from the oil debris sensor. One technique for detecting wear conditions in gear systems is by applying statistical distribution methods to particles collected from lubrication systems (Ref. 15). In this reference, mean particle size, variance, kurtosis and skewness distribution characteristics were calculated from oil debris data collected off-line. The wear activity was determined by the calculated size distribution characteristics. In order to apply this data to on-line debris data, calculations were made for each reading number for each bin. Mean particle size, relative kurtosis and relative skewness were calculated for each reading for six of the experiments with pitting damage. It was not possible, however, to extract a consistent feature that increased in value from the data for all experiments. This may be due to the random nonlinear distribution of the damage progression across all 56 teeth. For this reason, a more intelligent feature extraction system was analyzed and will be discussed in the following paragraphs.

When defining an intelligent feature extraction system, the gear states that a person plans to predict must be defined. Due to the overlap of the accumulated mass features, three primary states of the gears were identified: OK (no gear damage), inspect (initial pitting) and damage (destructive pitting). The data from Table 2 was plotted in Figure 4. Each plot is labeled with experiment numbers 1–6. The triangles on each plot identify the inspection reading numbers. The triangles circled indicate the reading number when destructive pitting was first observed. The background color indicates the OK, inspect and damage states. The overlap between the states is also identified with a different background color. The changes in states for each color were defined based on data shown in Tables 2–4. The minimum and maximum debris masses measured during experiments 1–6 when destructive pitting was first observed were used to define the upper limit of the inspect scale and the lower limit of the damage scale, respectively. The maximum amount of debris measured when no damage occurred (experiment 10) was above the minimum amount of debris measured when initial pitting occurred (experiment 7). The former was used as the lower limit of the inspect state. The next largest mass measured when no damage occurred (experiment 13) was used as the upper limit of the OK scale.

Fuzzy logic was used to extract an intelligent feature from the accumulated mass measured by the oil debris sensor. Fuzzy logic was chosen based on the results of several studies to compare the capability of production rules, fuzzy logic and neural nets. One study found fuzzy logic the most robust when monitoring transitional failure data on a gearbox (Ref. 7). Another study comparing automated reasoning techniques for condition-based maintenance found fuzzy logic more flexible than standard logic because it made allowances for unanticipated behavior (Ref. 14). Fuzzy logic applies fuzzy set theory to data, where fuzzy set theory is a theory of classes with unsharp boundaries and the data belongs in a set based on its degree of membership (Ref. 20). The degree of membership can be any value between 0 and 1.

Defining the fuzzy logic model requires inputs (damage detection features), outputs (state of gear), and rules. Inputs are the levels of damage, and outputs are the states of the gears. Membership values were based on the accumulated mass and the amount of damage observed during inspection. Membership values are defined for the three levels of damage: damage low, damage medium and damage high. Using the mean-of-the-maximum (MOM), fuzzy-logic defuzzification method, the oil debris mass measured during the six experiments with pitting damage was entered into a simple fuzzy-logic model created using commercially available software (Ref. 6). The output of this model is shown on Figure 5. Threshold limits for the accumulated mass are identified for future tests in the spur gear fatigue test rig. Results indicate accumulated mass is a good predictor of pitting damage on spur gears and fuzzy logic is a good technique for setting threshold limits that discriminate between states of pitting wear.

### Conclusions

The purpose of this research was to first verify that accumulated mass predicts gear pitting damage when using an inductance-type, on-line oil debris sensor. Then, using accumulated mass as the damage feature, the purpose was to identify a method to set threshold limits for damaged gears that discriminate between different levels of pitting damage. In this process, the membership functions for each feature state were defined based on the level of damage. From this data, and a simple fuzzy-logic model, accumulated mass measured by an oil debris sensor combined with fuzzy-logic analysis techniques can be used to predict transmission health. Applying fuzzy logic incorporates decision making into the diagnostic process that improves fault detection and decreases false alarms.

This approach has several benefits compared with using the accumulated mass and an arbitrary threshold limit for determining if damage has occurred. One benefit is that it eliminates the need for an expert diagnostician to analyze and interpret the data since the output would be one of three states: OK, inspect and shutdown. Since benign debris may be introduced into the system due to periodic inspections, setting the lower limit above this debris level will minimize false alarms. In addition to these benefits, a more advanced system can be designed with logic built in to minimize these operational effects. Future tests are planned to collect data from gears with initial pitting to better define the inspect region of the model and the severity of gear damage. Tests are planned for gears of different sizes to determine if a relationship can be developed between damage levels and tooth surface contact area to minimize the need for extensive tests to develop the membership functions for the threshold levels.

## Update

Due to the success of oil debris analysis in predicting damage on the spur gear fatigue rigs, an oil debris sensor was installed on the NASA spiral bevel gear test facility, and further tests were run. Details of that research are found in the report "Spiral Bevel Gear Damage Detection Using Decision Fusion Analysis," available at *www.grc.nasa.gov*. ○

**This article also appeared in the proceedings of the 14th International COMADEM (Condition Monitoring & Diagnostic Engineering Management) Congress, September 4–6, 2001 in Manchester, U.K.**

### References

1. Astridge, D.G. "Helicopter Transmissions—Design for Engineering Safety and Reliability." Proceedings of the Institution of Mechanical Engineers, Part G, Vol. 203, No. G2, pp. 123–38, 1989.
2. Aviation Safety and Security Program, the Helicopter Accident Analysis Team. "Final Report of the Helicopter Accident Analysis Team," summary of a cooperative analysis between the U.S. Dept. of Defense, NASA and the FAA, July 1998. See *http://safecopter.arc.nasa.gov/Pages/Columns/Accidents/ColAcc.html*
3. Campbell, R.L., C.S. Byington, and M.S. Lebold. "Generation of HUMS Diagnostic Estimates Using Transitional Data," Proceedings of the 13th International Congress on Condition Monitoring and Diagnostic Engineering Management, Houston, TX, December 2000.
4. Dempsey, P.J. "A Comparison of Vibration and Oil Debris Gear Damage Detection Methods Applied to Pitting Damage." NASA TM-210371, NASA Technical Memorandum, Glenn Research Center, Cleveland, OH. December 2000.
5. Dempsey, P.J. and J.J. Zakrajsek. "Minimizing Load Effects on NA4 Gear Vibrations Diagnostic Parameter," NASA TM-210671. February 2001. NASA Technical Memorandum Glenn Research Center, Cleveland OH.
6. "Fuzzy Logic Toolbox for use with METLAB," MathWorks Inc., Natick, MA, 1998.
7. Hall, D.L., A.K. Garga, and J. Stover, "Machinery Fault Classification: The Case for Hybrid Fuzzy Logic Approach." Proceedings of the 53rd Meeting of the Society for Machinery Failure Prevention Technology, Virginia Beach, VA, April 19-22 1999.
8. Howard, P.L. and J. Reintjes. "A Straw Man for the Integration of Vibration and Oil Debris Technologies," Presented at the Workshop on Helicopter Health and Usage Monitoring Systems, Melbourne, Australia, February 1999. Editor: Graham F. Forsyth. DSTO-GD-0197. Published by DSTO (Defense Science and Technology Organization) Australia, February 1999.
9. Howe, B. and D. Muir, "In line Oil Debris Monitor (ODM) For Helicopter Gearbox Condition Assessment." January 1, 1998. ADa347 503, Defense Technical Information Center, Ft. Belvoir, VA.
10. Hunt, T.M. "Handbook of Wear Debris Analysis and Particle Detection in Fluids," Elsevier Science Publishers Ltd., London, 1993.
11. Learmont, D. "Rotary Woes," Flight International No. 4725 Vol. 157, April 18–24, 2000.
12. Lynwander, P. Gear Drive Systems Design and Application. New York: Marcel Dekker Inc., 1983.
13. McGonigal, D.L. *A Comparison of Automated Reasoning Techniques for Condition Based Maintenance*. (Penn State University) Master's Thesis. August 1997.
14. Roylance, B.J. "Monitoring Gear Wear Using Debris Analysis-Prospects for Establishing a Prognostic Method," Proceedings of the 5th International Conference on Tribology, Vol 4. June 15, 1989.
15. Scibbe, H.W., D.P Townsend., and P.R. Aron,. "Effect of Lubricant Extreme Pressure Additives on Surface Fatigue Life of AISI 9310 Spur Gears." NASA TP-2408, NASA Lewis Research Center, Cleveland, OH, December 1984.
16. Stewart, R.M. "Some Useful Data Analysis Techniques for Gearbox Diagnostics," Report MHM/R/10/77, Machine Health Monitoring Group, Institute of Sound and Vibration Research, University of Southampton, U.K., July 1977.
17. Stewart, R.M. "Advanced HUM and Vehicle Management Systems Implemented through IMA Architecture," Proceedings from the 53rd American Helicopter Society Forum, Virginia Beach, VA, 1997.
18. Townsend, D.P. *Dudley's Gear Handbook*. 2nd Edition. New York: McGraw Hill Inc., September 1991.
19. Zadeh, Lofti. "Fuzzy Logic: Advanced Concepts and Structures." New Jersey: IEEE 1992.
20. Zakrajsek, J.J., D.P. Townsend and H.J. Decker. "An Analysis of Gear Fault Detection Methods as Applied to Pitting Fatigue Failure Data," NASA TM-105950, NASA Lewis Research Center, Cleveland, OH, April 1993.