Effectiveness of Empirical Mode Decomposition Based Features Compared to Kurtosis Based Features for Diagnosis of Pinion Crack Detection in a Helicopter

Canh Ly*[1], Kenneth Ranney[1], Kwok Tom[1], Hiralal Khatri[1], and Harry Decker[2]

[1]*U.S. Army Research Laboratory, Adelphi, Maryland, 20783, U.S.A*
**POC: canh.ly@us.army.mil*
kenneth.ranney@us.army.mil
kwok.tom@us.army.mil
hiralal.khatri@us.army.mil

[2]*VTD-Glenn U.S. Army Research Laboratory, Cleveland, Ohio, 44135, U.S.A*
harry.j.decker@us.army.mil

ABSTRACT

Features based on empirical mode decomposition (EMD) of measured vibration data were developed for a Bell OH-58 helicopter main rotor gearbox. A tooth on the input pinion of the gearbox was notched and run for an extended period at several over-torque conditions to induce a tooth fracture. Vibration data were recorded at regular intervals until a tooth fractured. The EMD features were found to be more sensitive to the gearbox condition than the kurtosis-based feature (FM4), and they diagnosed the onset of cracked tooth much earlier.*

1. INTRODUCTION

The National Aeronautics and Space Administration (NASA) Glenn Research Center has been working on improving diagnosis of gear damage since 1987. Experiments were conducted on the main rotor transmission of an OH-58 helicopter in NASA's 500HP Helicopter Transmission Test Stand. The objective of the experiments was to evaluate vibration-based diagnostic tools for detecting gear crack initiation. The transmission, set-up of the experiment, vibration measurements, and some of the results are described in (Decker *et al.,* 2003; Lewicki *et al.,* 1987). Herein, we give the results of processing this data using empirical mode decomposition (EMD) and generate a new feature that gives more timely warning of gear crack and fracture. Since the FM4 vibration diagnostic metric is one of the most popular metrics used (Stewart, R. M., 1977; Decker *et al.,* 2003; Lewicki *et al.,* 1987), these results are compared with those obtained with the FM4 metric. In addition, the kustosis-based technique does not follow a consistent trend (Shiroishi, *et al.,* 1999). On the other hand, the EMD shows that the trend increases as a function of the run time. We first summarize the EMD procedure that was described in (Huang *et al.,* 1998; Khatri *et al.,* 2008) for completeness, and then summarize the set-up of the experiment and measurements that were detailed in (Decker *et al.,* 2003; Lewicki *et al.,* 1987). Finally, we describe the measured data and the results.

2. EMPIRICAL MODE DECOMPOSITION

The use of the EMD to create intrinsic mode functions (IMFs) for nonlinear and non-stationary time series analysis is described in (Huang *et al.,* 1998). In this work, we also present a numerical procedure for performing the decomposition, and we illustrate the procedure using appropriate examples. Since EMD is based on the local characteristic time scale of the data, it can help generate new features for detection of faults in electromechanical systems (Khatri *et al.,* 2008). Briefly, local maxima and minima of the data, $x(t)$, are identified, and the maxima are connected by a cubic spline as the upper envelope. Similarly, a lower envelope is generated from the minima. The two envelopes should bracket all the data between them. Let $E_u(t)$ and $E_l(t)$ denote the values of upper and lower envelopes as a function of discrete time, *t*. The mean of the two envelopes is

$$M(t) = \frac{E_u(t) + E_l(t)}{2}. \quad (1)$$

The difference between the data and the mean of the envelopes gives the residue:

$$h_1(t) = x(t) - M(t), \quad (2)$$

where the subscript of $h(t)$ denotes level of iteration. This process is iterated with $h_1(t)$, representing a new time series whose upper and lower envelopes and mean envelope, $M_1(t)$, are computed to obtain the value for next iteration:

$$h_{i+1}(t) = h_i(t) - M_i(t), \text{ where } i = 1, 2, \ldots \quad (3)$$

Ideally, this process is continued until the residue, $h_i(t)$, satisfies the conditions specified for IMFs: (1) the number of extrema and the number of zero crossings in $h_i(t)$, must either be equal or differ by at most one; and (2) $M_i(t) = 0$ for all *t*. It is not always advisable or efficient to test for these conditions. Instead, the iteration process is terminated when the value of the standard deviation, *SD*,

$$SD = \frac{\sum_{t=0}^{T} \left| h_i(t) - h_{i-1}(t) \right|^2}{\sum_{t=0}^{T} h_{i-1}^2(t)} \quad (4)$$

is below a selected value, generally in the range of (0.2-0.3), as suggested in (Huang *et al.,* 1998). This value was defined by (Huang *et al.,* 2005) as the stoppage criterion. The results given here for IMFs were obtained when the stoppage criterion was less than 0.3 or the number of iterations in the IMF calculation was greater than 10000. It was a good practice to use the combination of these two criteria to prevent the infinite loop of the calculation. In addition, the stoppage criterion was predetermined based on the engineering judgment. Note that when $h_i(t) = h_{i-1}(t)$ for all *t*, no further reduction in residue is possible, so $h_i(t)$ is considered to have converged and *SD* = 0. The value of *SD* is used to gauge the level of convergence.

This gives us the first IMF, denoted as $c_1(t) = h_i(t)$. To obtain subsequent decompositions, $c_i(t)$, *i* = 2, 3…, treat the residue, $r_i(t)$,

$$r_i(t) = r_{i-1}(t) - c_i(t), \text{ with } r_0(t) = x(t),$$

as the signal to be decomposed, and repeat the previous procedure to obtain the IMF of the residue. This decomposition process can be stopped based on predetermined conditions. Note that if *n* IMF's are obtained, then the signal, $x(t)$, can be found:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \quad (5)$$

3. HELICOPTER TRANSMISSION EXPERIMENT

Experiments were conducted on the main rotor transmission of an OH-58 helicopter. Following is a brief description of the system and the experiments as they pertain to the results obtained with EMD.

The main rotor transmission is a two-stage reduction gearbox (Decker *et al.,* 2003). The first stage is a spiral bevel gear set with a 19-tooth pinion that meshes with a 71-tooth gear. The article reports on one of the experiments where a notch, shown in Figure 1, was machined into the fillet region of one tooth of the spiral-bevel pinion using electro-discharge machining. The dimensions of the notch were approximately 0.1 in wide by 0.005 in tall by 0.005 in deep. After a significant amount of run time at extreme torques, it was determined that the notch was not of a sufficient size to facilitate crack initiation. The notch was machined in the same area, and the dimensions of the notch were increased to 0.12 in wide by 0.01 in tall by 0.08 in deep. This second notch geometry was sufficient to initiate a crack. A suite of sensors consisting of a tachometer, a proximity probe, and five accelerometers were mounted to facilitate the detection of crack initiation and propagation.

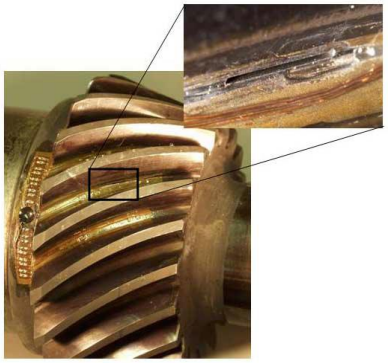


Figure 1: Photograph of notch

A tachometer signal (once per revolution) was used to synchronize measured acceleration to shaft position so as to compute average acceleration (time synchronous averaging) as a function of shaft position. A proximity probe was mounted inside the transmission on one of the support webs to detect the passing of the top land of the teeth.

The five accelerometers (A1-A5) were mounted on the gear box at selected locations shown in Figure 2. The locations were selected based on previous experience (Lewicki *et al.,* 1987). Accelerometer 1 is located on the input bevel gear housing immediately above where the input shaft connects to the pinion and is aligned to be most responsive in the vertical direction. Accelerometer 2 is at the same location and is aligned to the rotational axis of the input shaft. Accelerometers 3 and 4 are mounted around the circumference of the ring gear housing and are located 45° and 225° from the input pinion gear, respectively. Accelerometer 5 is mounted on an attachment bolt near accelerometer 4. Accelerometers 3, 4, and 5 are mounted in the axial-transverse plane, and have sensitivities in both directions. The accelerometers are linear to 20 kHz and have a resonance frequency of 90 kHz.

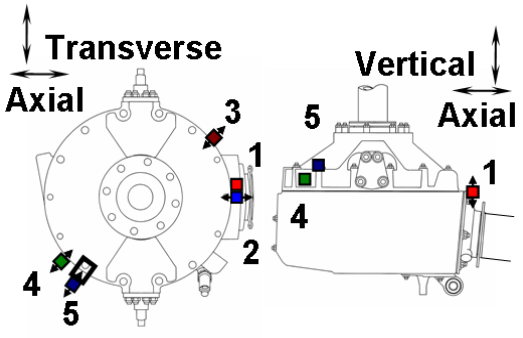


Figure 2: Accelerometer locations

The vibration, speed, and proximity probe signals were sampled through a low-pass, anti-aliasing filter with a 56 kHz cutoff frequency. These analog data were sampled at 150 kHz rate with 12-bit precision, and a record length of 1.5 s was taken every 15 s.

The pinion was run at the design speed of 6060 revolution-per-minute (rpm) (101 Hz) and the torque level was varied, with the run time in million-cycles and run time history in hours, as shown in Figure 3. Initially, the pinion was run at 80% of the maximum design torque, and the torque was gradually increased to try to speed up the propagation of the crack. Visual inspections using a 60X microscope were carried out at the periods shown in the figure by inverted triangles. At 80 h run time, the notch was deepened (solid square symbol) and the test was continued to failure, which occurred at approximately 101.6 h. This paper presents results of processing the accelerometer data acquired after the notch was deepened with torque settings at 120%, 140%, and 150% of the maximum design torque. There were 2523 records (each 1.5 s) collected at 120% torque, 1623 records collected at 140% torque, and 1054 records at 150% torque. Due to the intensively computational time in calculating IMFs for all data files, we only processed selected records (files) at each of the torque levels—namely, 23, 25, and 17 records at 120%, 140%, and 150%, respectively. These records covered the time zone of each torque level fairly uniformly.

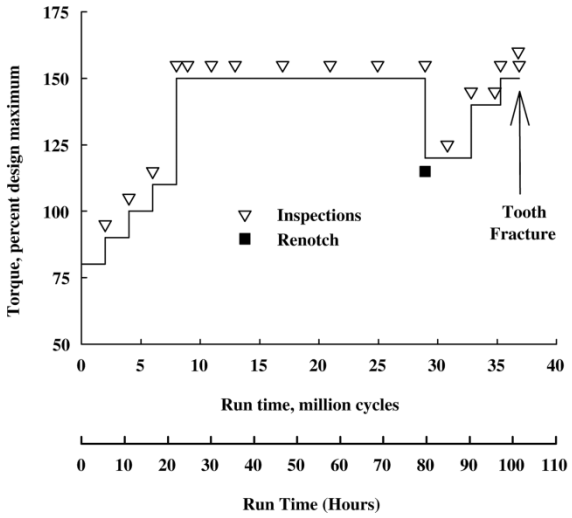


Figure 3: Loading history of the pinion

4. DATA PROCESSING AND RESULTS

First, we show samples of recorded data at the three torque levels in Figure 4—(a) 120 %, (b) 140 % and (c) 150%. For the demonstration of the feasible investigation of the EMD algorithm, we show the data for accelerometer #1 because it was the most responsive to the fault.

The nominal period of shaft rotation is10 ms, so we show the first 10 ms data from the 1.5 s record length. The data have been normalized by their respective root-mean-square (RMS) values. Figure 4(a), the top plot, shows the collected data soon after the notch was deepened at approximately 80 h of run time (Torque 120%). The data depicted in Figure 4(b), the middle plot, was collected at approximately 91 h of run time (Torque 140%), and the data depicted in Figure 4(c), the bottom plot, was collected at approximately 98 h of run time (Torque 150%). The figure shows the correlation between the vibration levels and the engagements of the 19 teeth of the pinion. We do not have an explanation for the low noise level in the data for the 150% torque case; however, one can speculate

that this may be because of higher torque or crack progression.

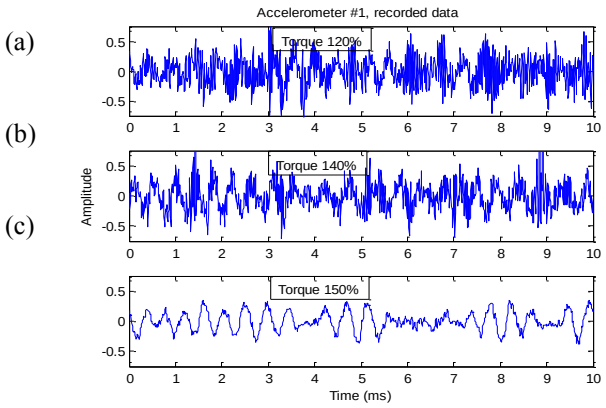


Figure 4: A sample of collected data for three different torques (a) 120%, (b) 140%, (c) 150% at near 80, 90, and 98 h (run time) after crack test was begun.

Figure 5 shows the power spectral densities of these three data sets. The power spectrum should peak at about 1900 Hz because the shaft rotates at 101 Hz (6060 RPM) and there are 19 teeth on the pinion. There is no significant change in the spectrum to reliably detect crack propagation.

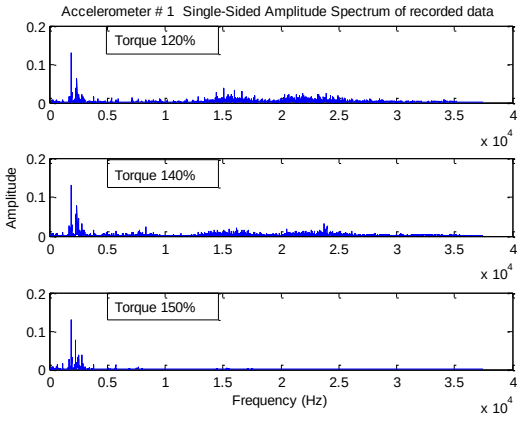


Figure 5: Amplitude spectrum of the three data samples given in Figure 4

The root-mean-square (RMS) values of the amplitudes measured by accelerometer # 1 are shown in Figure 6 for the selected 65 records. The RMS values are in the same range for 120% and 140% torques, but the data set at 150% torque has much lower RMS values. The exact cause of the difference in the RMS levels is unknown. One may reason that as the torque level increases, the gear has less room to spin. As a result, there will be less noise in the data. Consequently, the RMS values are much less than other torque levels.

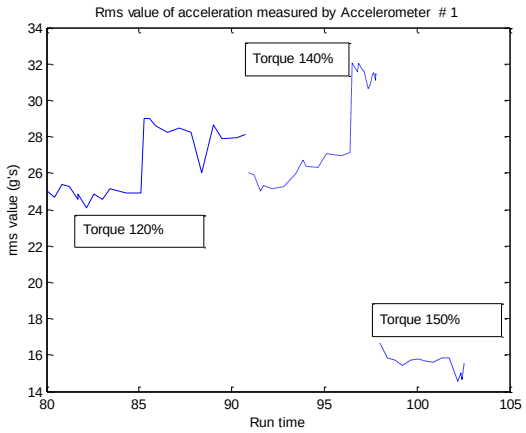


Figure 6: Root-Mean-Square values of the acceleration measured by accelerometer #1 at sampled run time.

4.1 Empirical Mode Decomposition of vibration data

We use cubic splines to generate the upper and lower envelopes of the measured vibration data. These splines introduce transients near the start and end points of the time series. As a result, the calculated IMFs have large values (errors) for a portion of samples in the vicinity of the start and the end of the time series. For this reason, we discard a fixed amount of data—approximately the initial and the final 5.6% of the total number of samples in any record. This corresponds to approximately 25,000 out of the 225,000 IMF samples available, and it leaves us with 200,000 samples for generating the results presented here.

We had generated 10 IMFs of the 65 selected data files, but the first seven respective IMFs show greater differences between one another, so we use only those seven IMFs in our condition monitoring algorithm. The plots of two IMFs, along with their single-sided amplitude spectrums, are shown in Figures 7-10, and they illustrate the manner in which the IMFs change as the damage to the gear tooth progresses. For example, in Figure 7, it is apparent that the energy in IMF 1 for the first two plots (120% Torque and 140% Torque) is greater than that in the third plot (150% Torque). The plots in Figure 10, however, lead us to conclude that the energy for 150% torque is greater than the energy for 140% torque, which in turn is greater than the energy for 120% torque. All of these observations suggest a new statistical, fault-detection feature suitable for a condition monitoring algorithm.

Note that we used non-time synchronous averaging (TSA) data to calculate IMFs. This is beneficial because collecting TSA data is not a trivial task, since it requires a tachometer sensor to capture synchronization data for all pulses and computational resources to

perform an average of many pulses for one data point. This is a major drawback of the FM4 calculation.

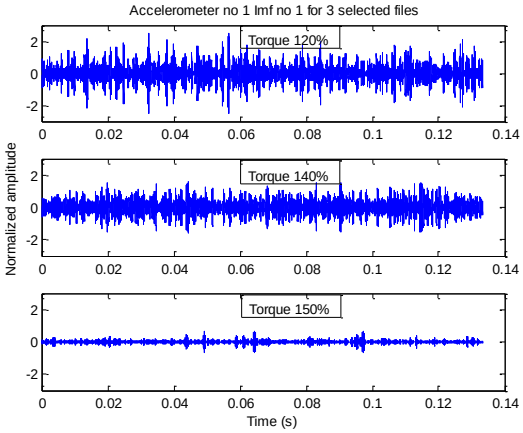


Figure 7: The IMF No. 1 for the three selected data sets given in Figure 4 near 80, 90, and 98 h (run time). Note the larger energy in the waveforms of the first two plots.

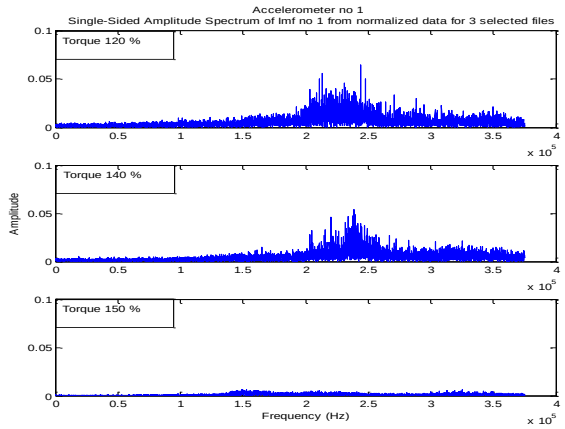


Figure 8: The amplitude spectrum of IMF No. 1 for the three selected data sets given in Figure 4 near 80, 90, and 98 h (run time). Again, notice the larger amount of energy in the waveforms of the first two plots.

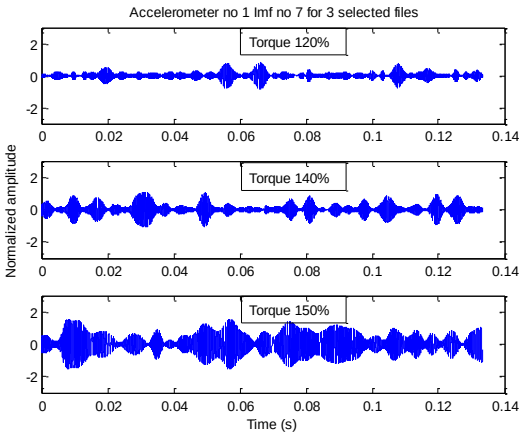


Figure 9: The IMF No. 7 for the three selected data sets given in Figure 4 near 80, 90, and 98 h (run time). Notice (in this case) the larger amount of energy for the case of 150% torque.

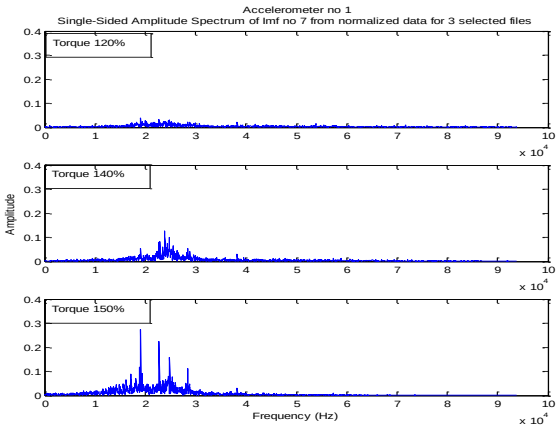


Figure 10: The amplitude spectrum of IMF No. 7 for the three selected data sets given in Figure 4 near 80, 90, and 98 h (run time). Notice (in this case) the largest amount of energy for the 150% torque, and the second-largest amount of energy for 140% torque.

4.2 Algorithms for monitoring the condition of the gearbox

As described above, certain relative, qualitative differences in the IMF series of the three selected data sets were evident for each of the seven modes. These differences are apparent in the amplitude spectrums shown in Figure 8 and Figure 10. Specifically, the peak values and the frequency distributions of spectrums of the IMFs are quite different.

We propose to monitor the condition of the gear-box with an algorithm based on the three sets of features: (1) the centroids of the amplitude spectrum of each of the seven IMFs, (2) the second moments of the spectrums about the centroids of each of the seven IMFs and (3) the energies of each of the seven IMFs,

referring to the amplitude square of each of the seven IMFs. Here, we refer to the center of mass of the amplitude spectrum as the *centroid*, and we calculate it according to the equation:

$$centroid = \frac{\sum_i freq_i \times s_i}{\sum_i s_i}, \qquad (6)$$

where $freq_i$ is i^{th} element of the frequency index and s_i is the i^{th} element of the amplitude spectrum.

Since the calculations of IMFs using EMD is extremely computationally intensive, we use a subset from the whole data set (covering run time 80-81.7 h) as the standard (baseline) and use the mean values of the amplitude spectrums, the second moments, and the energy of each of the seven IMFs from the subset. The energy is defined as the amplitude square of the frequency spectrum or FFT of each of the seven IMFs. Thus, we have a 21-dimensional mean vector of the features, denoted by mf_i. We also compute the variances of these features, vf_i, for normalizing distances of test cases. The proposed algorithm computes the distance from the mean as follows:

$$R = \sqrt{\sum_{i=1}^{21} \frac{[f_i - mf_i]^2}{vf_i}}, \qquad (7)$$

where f_i is the i^{th} feature vector. Normally, the distance is computed only with respect to the mean of the baseline. We have normalized this with the variance of the baseline feature values. We refer to this as the *modified radial basis distance*. The advantage of the normalization is to avoid the uncertainty in each data file.

In order to illustrate the effectiveness of the algorithm and the FM4, Figure 11 shows the amplitude of the proximity probe, defined in Section 3, versus Run Time history in hours. The onset of the fault occurred at around 101 h, which indicated when the broken tooth occurred.

The results of applying this algorithm to the processed data are shown in Figure 12(a). For comparison, Figure 12 (b) shows the corresponding values of the FM4 feature given in (Decker *et al.,* 2003). Although there are many techniques to compare with the proposed algorithm, we chose the FM4 method in this paper for the comparison purposes because FM4 is a well-established, kurtosis-based algorithm and its value is near 3 when the measured data is Gaussian. If there is a fault, the data is expected to be non-Gaussian and FM4 deviates substantially from 3. By considering the baseline data collected at the beginning of the test, we can identify a similar threshold for the test statistic, *R*. Based on this data (shown in Figure 12(a)), we conclude that a value of 30 would be a reasonable choice.

Since the range of values assumed by *R* is large, we display the results for run time up to 96 h in Figure 13(a) to better show the change in *R* before the pinion fails. Using the threshold identified above, the algorithm output warns us of a failure after 90 h, which is much earlier than the FM4 warning. Note that in Figure 13(b), while the value of FM4 rises to the amplitude of 4, it does not persistently increase after 85 h. In other words, the FM4 method does not provide a confident warning of early failure.

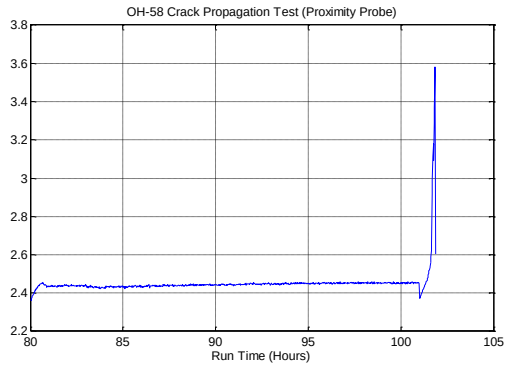


Figure 11: Proximity probe signal

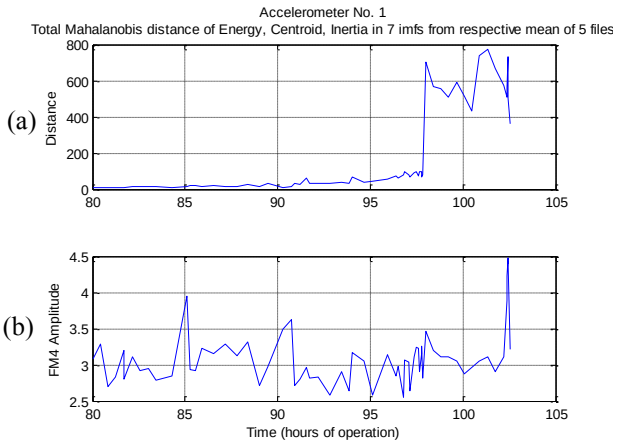


Figure 12: Comparison of outputs of the proposed algorithm and FM4

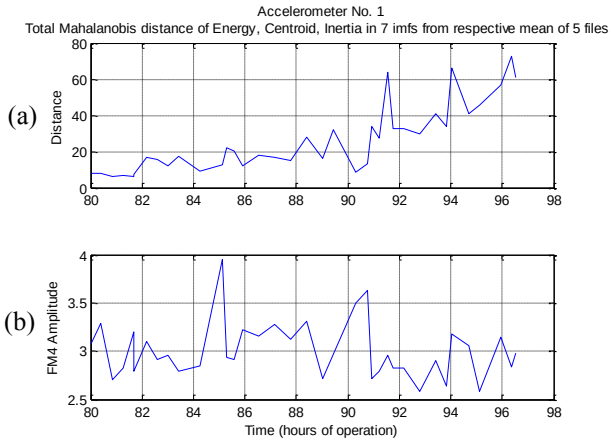


Figure 13: Comparison of outputs of the proposed algorithm and FM4 for 80-96 hours

5. CONCLUSION

EMD has been shown to be a useful tool for feature extraction of information related to the characteristic of the *modified radial basis distance*. The performance of EMD-based features compares favorably with the performance of a well-known feature FM4, based on the fourth order statistics (i.e., kurtosis). Although the EMD is not suitable for a real-time or on-board implementation because of its computational intensity, it certainly is an important concept that merits exploration with a larger, more extensive data set. Our results indicate that the algorithm output produce a failure warning after approximately 90 h of operation, which is much earlier than the FM4 warning after approximately 102.5 h of operation. This may be an advantage for developing the prognostic algorithm.

As we note that EMD outperformed the FM4 statistical technique for the gear box on selected sample data files, we also stress that further analysis is critical in order to determine the consistency of the EMD-based feature calculation across a wide variety of data sets. While the available data set was too limited in size and scope to draw statistically significant conclusions, it still provided a qualitative indication of the potential of the approach. We plan to extend this concept to the analysis of data obtained from ball-bearing defects due to corrosion, defective inner races, defective outer races, or defective cages. The results of this proposed analysis will be presented in another paper.

ACKNOWLEDGMENT

The authors would like to thank Dr. Hiralal Khatri (ARL, retired) for his initiation of this paper and his major contribution on the research of EMD in this important field. The authors also acknowledge David Lewicki for his expertise on collecting experimental data at the Vehicle Technology Directorate, NASA Glenn, ARL.

REFERENCES

Decker, Harry J. & Lewicki, David G. (June 2003), Spiral Bevel Pinion Crack Detection in a Helicopter Gearbox, *NASA TM-2003-212327, ARL-TR-2958*.

Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q., Yen; N.C.,Tung C.C.; & Liu, H.H. (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc., R. Soc. Lond. Ser. A 454*, pp. 903-995.

Huang and Shen, eds., (2008) Huang, N.E., Shen, S.S.P., eds. (2008). Hilbert-Huang Transform and Its Applications, New Jersey: World Scientific.

Khatri, Hiralal; Ranney, Kenneth; Tom, Kwok; del Rosario Romeo (April 2008), New Features for Diagnosis and Prognosis of Systems Based on Empirical Mode Decomposition, *International Conference on Prognostics and Health Management*, PHM 2008.

Lewicki, David G. & Coy, John J. (April 1987), "Vibration Characteristics of OH-58A Helicopter Main Rotor Transmission," NASA TP-2705, AVSCOM TR 86-C_42.

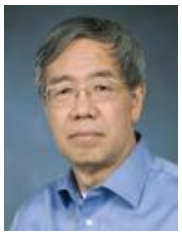
Shiroishi, J.; Li, Y.; Liang, S.; Danyluk, S; Kurfess, T. (1999), Vibration Analysis for Bearing Outer Race Condition Diagnostics, Jornal of the Brazillian Society of Mechanical Science, Vol. 21, No. 3, Rio de Janerio

Stewart, R. M. (July 1977), 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.

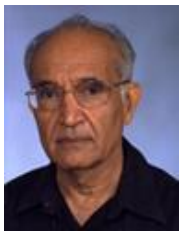
Dr. Canh Ly received his B.S. in Electrical Engineering from the University of Maryland, College Park, MD, an M.S. in Electrical Engineering from Johns Hopkins University, and a Ph.D. in Information Technology from George Mason University. Dr. Ly has been Member of Technical Staff at the U.S. Army Research Laboratory (ARL). He had three patents related to frequency estimation technique, patch array antennas, and published papers in IEEE Radar Conference Proceedings, Asilomar Conference, and SPIE Conference. His research interests include developing algorithms for fault diagnosis and failure prognosis, antenna design, superresolution algorithms

for millimeter wave (MMW) radars, statistical signal processing, advanced HEMP techniques, and radar array processing.

Mr. Kenneth Ranney received the B.S. degree in electrical engineering and computer science from the Johns Hopkins University in Baltimore, MD, and the M.S. degree in electrical engineering from the University of Maryland in College Park, MD. He currently works as an electronics engineer at ARL, focusing primarily on problems related to signal processing and automatic target detection.



Mr. Kwok Tom, is employed by the Army Research Laboratory. He has B.S. and M.S. degrees in Electrical Engineering from George Washington University. His past responsibilities have been as a team leader for the signal processing team for Ultra Wide Band radar in the detection and identification of underground objects, the development of a Ka Band, Monopulse, fully polarimetric, instrumentation grade radar system and fuze development for the PATRIOT PAC-2 missile. Since 2004, he has been an integral part of the Prognostic & Diagnostics (P&D) program and is currently the team leader for the P&D for Condition Based Maintenance program.



Dr. Hiralal C. Khatri received the Ph.D. degree from Purdue University, West Lafayette in 1966. In 1971 he joined the staff of the ARL in Adelphi, MD, where he worked until his retirement in 2009. While at ARL, his research interests included statistical processing and nonlinear signal processing, radar systems, optimization, prognostics and diagnostics of electro-mechanical systems.

Mr. Harry J. Decker is a Mechanical Engineer employed by the Vehicle Technology Directorate of ARL at the NASA Glenn Research Center in Cleveland, OH. He holds a B.S. in Aerospace Engineering Technology from Kent State University, and a B.M.E. and an M.S. in Mechanical Engineering from Cleveland State University. He has been performing analytical and experimental research in transmission, gearing and bearing areas for rotorcraft since 1988. He has performed various transmission and component experiments such as vibration, noise, temperature, and efficiency studies. His main area of concentration is in the area of Health and Usage Monitoring.