OPTIMIZED INTERPRETATION OF MAINTENANCE DATA FROM GEARBOXES SUBJECT TO TOOTH FAILURE

DAMING LIN, MURRAY WISEMAN

OMDEC Inc., Toronto, Ontario, Canada, http://www.omdec.com
E-mails: daming@omdec.com; murray@omdec.com

AND

DRAGAN BANJEVIC, ANDREW K. S. JARDINE

CBM Lab, Dept. of Mechanical and Industrial Engineering, University of Toronto,
5 King's College Road, Toronto, Ontario, Canada M5S 3G8
E-mails: banjev@mie.utoronto.ca; jardine@mie.utoronto.ca

CORRESPONDING AUTHOR

Prof. A.K.S. Jardine
Department of Mechanical and Industrial Engineering
University of Toronto
5 King's College Road
Toronto, Ontario
Canada, M5S 3G8

ABSTRACT

A Condition Based Maintenance (CBM) policy is a procedure used by maintenance personnel to interpret a set of measured machine condition indicators and decide whether or not to renew a physical asset at the current moment.

Traditionally CBM data interpretive policies have been obvious. For example, for process machinery, if a temperature, pressure, or vibration reading exceeds a pre-defined limit, maintenance should be carried out before functional failure. However our ability to collect large amounts of condition data has continually outpaced our ability to define policies for its interpretation. The condition indicators may sometimes contradict one another. Upward or downward trends are frequently obscured by randomness in the data. In many instances no clear set of limits or rules have been developed to indicate whether or not a failure process is underway and how much time is available before the physical asset is no longer able to perform one of its functions.

The EXAKT CBM optimizing methodology was applied to a set of experimental condition data on gearboxes in order to develop an optimal interpretation policy. An optimal policy is a procedure for data interpretation, which, if applied consistently in a CBM program, minimizes the cost of maintenance of a physical asset.

1. INTRODUCTION

Condition data from eleven gearboxes run to failure on the mechanical diagnostic test bed (MDTB) at Penn State University Applied Research Laboratory (ARL) was analyzed using the proportional hazards modeling (PHM) technique embedded in the EXAKT CBM optimizing program (see [1, 3, 4]). Several statistical and replacement decision models were built based upon the observed condition data and ensuing failure events.

The MDTB makes a large number of condition indicators available for analysis, for example, the conventional vibration features such as acceleration amplitudes at various gear and bearing frequencies. In particular the Fault Growth Parameter (FGP) was calculated from the residual error signal obtained by a signal processing algorithm (see Miller) developed at ARL. We call it, henceforward, the ARL algorithm. In this algorithm, a family of Wavelets is constructed to decompose the gear motion error signal and to extract the residual error signal for gear fault detection. In addition, we proposed a modified version of FGP, called FGP1, by weighting each point in the residual error signal spectrum proportional to its deviation from a reference baseline. Besides FGP and FGP1, other useful indicators were extracted from the residual error signal. It was found that the revised version of FGP (FGP1) is superior to the FGP and other condition indicators for building the EXAKT model and making the replacement decision.

The eleven test runs were designated Test Run Numbers: 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, and 15. All gearboxes were run in at 540 in-lbs torque for the first 96 hours of each test. Following the initial run-in period, the output torque was increased as follows:

Test Runs 05, 06, 12, and 13: by 300% from 540 to 1620 in-lbs
Test Runs 07, 08, 09, 10, and 11: by 200% from 540 to 1080 in-lbs
Test Runs 14 and 15: the load was alternated between 100% (540 in-lbs) and 300% (1620 in-lbs) at 30-minute intervals.

Vibration acceleration readings were taken at 8-hour intervals during the 96-hour run-in period and at 30-minute intervals during the high load 'operational' phase. Readings were of 10 seconds duration and sampled at a rate of 20 kHz. Accelerometers were located at various positions on the gearbox casing.

Test Runs 11, 13, and 15 ended in shaft failure. All other test runs ended with gear tooth failure. All gearboxes in the eleven tests consist of a gear and a pinion with gear ratio 1:1.533 for Test Runs 05 to 11 and 1:3.333 for Test Runs 12, 13, 14 and 15. We divide the test runs into two groups such that the gearboxes have the same gear geometry in the group. Group A is composed of Test Runs 5-11 and Group B consists of Test Runs 12-15. We found that the data from these two groups of physically different gearboxes have to be analyzed separately. “Gear tooth fracture” was the failure mode examined in this study. The three test runs where no gear tooth failure was observed were classified as suspensions. Only the gear in a gearbox was studied in this paper since there was no failure information about the pinion (pinion tooth fracture) in all eleven tests.

In this paper, we analyze the data set described above and apply the EXAKT CBM optimizing methodology to develop optimal maintenance policies for the gearboxes. The paper is organized as follows. In Section 2, a signal processing technique is used to extract useful information from the raw vibration signals. Based on the extracted information, we obtain the event and inspection data that are essential to applying EXAKT. Data cleaning and pre-processing are also included in this section. In Section 3, EXAKT software is used to analyze the data, build a PHM for the gearbox and develop an optimal maintenance policy. Finally in Section 4, the results are summarized and some concluding remarks are given.

2. DATA PRE-PROCESSING AND ANALYSIS

The original data provided by the ARL of Penn State University on a series of test runs of single reduction helical gearboxes contains vibration signatures captured by accelerometers mounted at different positions on the gearbox casing. As suggested by Miller [5], accelerometer A03, which is mounted in the axial direction, should be sensitive to the detection of helical gear tooth faults. Accordingly, data obtained from accelerometer A03 were used in our analysis. Since it is impossible to use directly the raw vibration data for CBM analysis, a signal processing and data pre-processing step is required in order to extract features that we may analyze.

2.1. VIBRATION SIGNAL PROCESSING

In each test, data were collected until the MDTB was shutdown as a result of two accelerometers exceeding a predetermined limit of 150% of RMS. The targeted failure mode, tooth failure, however, may occur at any time prior to shutdown. To detect the moment when a *potential* tooth failure occurs, a signal processing technique is required.

In the literature, there are many signal processing tools for vibration data, such as power spectrum, time domain averaging, denoising, demodulation, time series, time-frequency distribution, wavelet, neural network, high order statistics, etc. In this paper, we are interested in using demodulation of a vibration signal. Wang and McFadden [8] proposed decomposition of motion error signal for gear fault diagnosis based on time domain average. But as mentioned by Miller [5], time domain averaging has some drawbacks such as the requirements of long signal length, precise shaft rate, synchronous signal, and a different signal for each gear. To get around these drawbacks, Miller [5] proposed to use wavelets for the decomposition of a vibration signal.

Wavelet is a powerful mathematical tool with wide applications [10]. Wavelet transformation has been applied in rotating machinery diagnostics (see, e.g., [2, 9, 11]). Miller's approach [5, 6, 7], however, is different in that a series of wavelets are used as a comb filter to decompose the motion error signal rather than transforming the signal. In this paper, we used Miller's approach, or the ARL algorithm, for the signal processing of the vibration data. The idea is briefly described as follows.

The gear motion error signal defined in [8] was used for the diagnostics of gear fault. It is a real signal described by an infinite cosine series with fundamental period f_r, which is the input shaft rate. Let N be the number of teeth for the pinion and M be the number of teeth for the gear. The composite gear motion error $s(t)$ can be written as a summation of three components:

$$\begin{aligned} s(t) &= \sum_{l=0}^{\infty} c_l^h e^{j(2\pi l(Nf_r)t+a_l)} + \sum_{n=0,\,\mathrm{mod}(n,N)\neq 0}^{\infty} c_n^p e^{j(2\pi n f_r t+\beta_n)} + \sum_{m=0,\,\mathrm{mod}(m,M)\neq 0}^{\infty} c_m^g e^{j(2\pi m(Nf_r/M)t+\gamma_m)} \\ &= s_{eh}(t) + s_{er,p}(t) + s_{er,g}(t), \end{aligned}$$

where $j=\sqrt{-1}$, c_l^h, c_n^p, c_m^g are amplitudes, $\alpha_l, \beta_n, \gamma_m$ are phases, $s_{eh}(t)$ is the harmonic error component, $s_{er,p}(t)$ is the residual error component due to the pinion and $s_{er,g}(t)$ is the residual error component due to the gear.

A series of Morlet wavelets having the form

$$\psi(t) = \sqrt{\pi}\sigma \exp(j2\pi f_0 t - (\pi\sigma t)^2),$$

where σ is the scale parameter and f_0 is the frequency parameter, were used as a comb filter to decompose the original signal to obtain the gear motion error signal, and the residual error signals for the gear and the pinion.

Based on the plot of the residual error signal, one can make a gear fault diagnosis. Typical plots of residual error signals are reproduced in Figure 1 from Miller [5] for illustration. In addition, a fault growth parameter (FGP) based on the residual error signal was proposed to track the gear tooth health condition over time. The FGP is defined as the part (percentage of points) of the residual error signal which exceeds three standard deviations from the baseline residual taken when the run began.

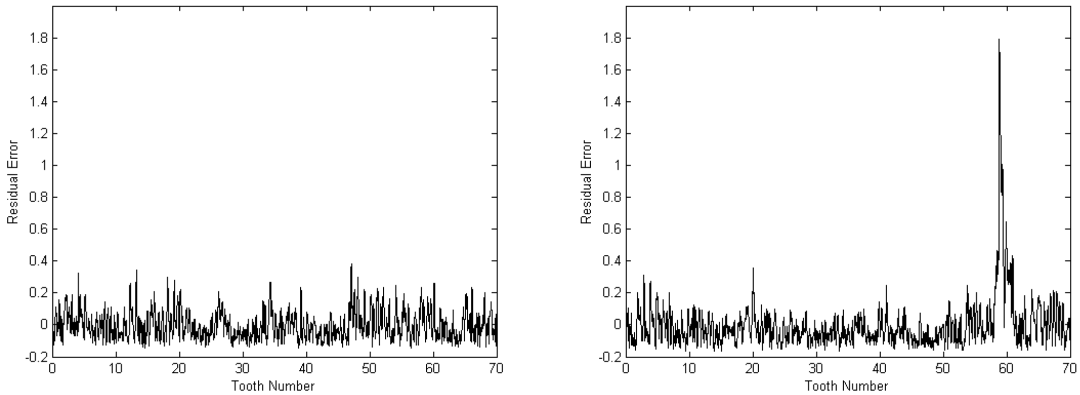

Figure 1: Residual error signals for a good gear at timestamp 200 (left) and for a gear with a broken tooth at timestamp 208 (right) in Test Run 12

The goal of signal processing in CBM is to filter out of the signal, as much operational and environmental data as possible, so that the magnitude of the remaining signal reflects the "ground truth" state of deterioration for the targeted failure mode. We modified the definition of FGP by assigning weights to the residual error signal points which exceed three standard deviations from the baseline residual. The weights were computed as proportional to the magnitude of their deviations. The modified version of FGP is called FGP1 in this paper. The signal processing technique described above enables us to prepare a table of inspection data related to the degradation process of the failure mode "tooth fractured" and a table of "event" data (installations, failures, suspensions, adjustments, etc). These tables are essential for applying the EXAKT procedure to create statistical models supporting predictive maintenance decision making.

2.2. EVENT DATA

To prepare the event data for applying the EXAKT CBM optimization technique, we are required to know when the gear was installed and when the gear tooth failure occurred. Also, other events, that would influence the measured variables or the failure mechanism, such as maintenance adjustments, operational changes, etc, should be included in the data. In this study, changes in torque loading as described in the introduction, were accounted for in the analysis process. The ARL algorithm was used as a method to detect the moment of tooth fracture providing the failure "event" data required for analysis. The event data for the gearboxes are presented in Table 1.

TABLE 1

Events table for all eleven histories

Ident	Date	WorkingAge	Event
5	6/19/97 1:35:00 PM	0	B
5	6/24/97 4:16:00 PM	4297	EF
6	7/17/97 3:15:00 PM	0	B
6	7/21/97 6:31:00 PM	2389	EF
7	7/26/97 6:50:00 PM	0	B
7	7/31/97 11:31:00 PM	3471	EF
8	8/16/97 4:45:00 PM	0	B
8	8/21/97 1:27:00 AM	2563	EF
9	9/10/97 10:45:00 PM	0	B
9	9/20/97 3:01:00 AM	8057	EF
10	11/17/97 4:20:00 PM	0	B
10	11/23/97 7:31:00 PM	3904	EF
11	1/22/98 4:02:00 PM	0	B
11	1/30/98 1:06:00 AM	5748	ES
12	2/19/98 9:00:00 PM	0	B
12	2/24/98 5:01:00 AM	2775	EF
13	3/05/98 8:30:00 PM	0	B
13	3/10/98 12:01:00 PM	3280	ES
14	3/15/98 6:32:00 PM	0	B
14	3/20/98 8:13:00 AM	3007	EF
15	7/27/98 2:04:00 PM	0	B
15	8/01/98 7:48:00 PM	4010	ES

In the events table "Ident" refers to Test Run number. "Event" refers to events: "B" , that designates the installation of the gearbox, "EF", that designates the tooth failure event and "ES", that designates the end of the test due to shaft failure (considered as a suspension with respect to the "gear tooth fractured" failure mode). "Date" is the calendar date and time when the event occurred. "WorkingAge" refers to the working age as a measure of service usage. Since, in each test, the gearbox operated under varying loads with torques ranging from 540 in-lbs up to 1620 in-lbs, it would be inappropriate to use simple calendar "running time" as a service usage measure, which would ignore the different working conditions under which the gearbox operates. As a reasonable approach we used the integral of the product of actual running time and instantaneous torque as the working age reflecting the accumulated stress on the gear teeth. The unit for working age, as defined, is in-lb-day.

2.3. INSPECTION DATA

The signal processing technique is used to compile the table of inspection data related to the degradation associated with the targeted failure mode. Some tests ran for a period of time after failure occurred. Only inspections prior to detected tooth failure are included in the inspections table. For purposes of comparison, some 'conventional' vibration features were also included in the inspections table (not shown in Table 2). For example, the maximum amplitude of acceleration in a narrow frequency band around the gear mesh frequency and the sidebands were tested as potential covariates in a proportional hazards model. The proportional hazards modeling analysis revealed, however, these are not significant indicators of gear tooth failure.

Inspection data are summarized as shown in Table 2, which is a partial view of the entire Inspections table. "Ident", "Date", and "WorkingAge" have the same meaning as in the Events table. The other variables given in the Inspections table are extracted from the residual error signal, which are described as follows:

FGP - fault growth parameter,
FGP1 - revised FGP,
RFM - mean of the power spectrum of the residual error signal,
RFS - standard deviation of the power spectrum of the residual error signal,
RTM - mean of the residual error signal, and
RTS - standard deviation of the residual error signal.

TABLE 2

Inspections table

Ident	Date	WorkingAge	FGP	FGP1	RFM	RFS	RTM	RTS
5	6/19/97 2:00:01 PM	9.422793403	0	0	2.214	28.1485	0.083873	0.044863
5	6/19/97 10:00:01 PM	190.532660070984	0.66288	0.66288	1.9092	24.0698	0.071178	0.040783
5	6/20/97 6:00:00 AM	371.760867206249	0.47348	0.47348	1.6615	21.1355	0.068625	0.039672
5	6/20/97 2:00:00 PM	552.9594005409	0.37879	0.37879	1.9976	22.0851	0.071577	0.040055
5	6/20/97 10:00:00 PM	734.310333871595	0.94697	0.94697	1.6719	21.8022	0.077595	0.042951
5	6/21/97 6:00:01 AM	915.723966073461	0.47348	0.47348	1.9453	22.8788	0.07907	0.040879
5	6/21/97 2:00:00 PM	1097.21426410803	0.47348	0.47348	1.8623	24.118	0.091515	0.044289
5	6/21/97 10:00:01 PM	1278.74590040316	0.094697	0.094697	1.8029	21.2486	0.073868	0.037136
5	6/22/97 6:00:01 AM	1460.33453373781	0	0	1.7462	22.6369	0.072755	0.037574
5	6/22/97 2:00:00 PM	1641.98645949345	0.56818	0.56818	1.9661	25.6433	0.09066	0.043821
5	6/22/97 10:00:00 PM	1823.72942615747	0	0	1.9658	23.8469	0.081273	0.03821
5	6/23/97 6:00:00 AM	2005.39052615879	0	0	1.9189	24.0627	0.085294	0.043767
5	6/23/97 2:00:00 PM	2540.23665949602	3.5985	3.5985	2.9623	33.9679	0.12064	0.065243
5	6/23/97 2:30:01 PM	2573.7012364347	7.4811	7.4811	3.3846	35.8643	0.14462	0.081629
5	6/23/97 3:00:00 PM	2607.11836312754	4.072	4.072	2.4435	31.427	0.11078	0.062766
5	6/23/97 3:30:00 PM	2640.53895063143	2.0833	2.0833	2.8801	31.7443	0.0965	0.051916
5	6/23/97 4:00:00 PM	2673.96091104032	6.25	6.25	3.1761	31.0083	0.13418	0.075985
5	6/23/97 4:30:00 PM	2707.37974229421	3.4091	3.4091	2.986	34.0286	0.11848	0.063038
5	6/23/97 5:00:01 PM	2740.81539067039	5.0189	5.1088	3.0699	33.8125	0.13142	0.074357
5	6/23/97 5:30:01 PM	2774.23370109095	2.3674	2.552	3.1619	34.6763	0.11484	0.062556
5	6/23/97 6:00:00 PM	2807.63321465763	6.8182	7.0822	3.3282	36.9377	0.14324	0.083121
5	6/23/97 6:30:00 PM	2841.05906257819	3.5985	3.5985	2.8078	30.6903	0.12571	0.065146
5	6/23/97 7:00:00 PM	2874.47717715374	2.7462	2.7462	3.1056	35.9933	0.1229	0.063949
5	6/23/97 7:30:00 PM	2907.90208340763	4.6402	4.6402	3.1828	38.1417	0.14215	0.075681

In the data set, there are several outliers that were reported by the ARL investigators as invalid data. Two significant outliers, one in Test Run 10 and the other in Test Run 14, were corrected by interpolating between the preceding and next values. The last three inspections in Test Run 11 have very high values. These abnormal readings may have been caused by contamination from other vibration sources as a result of the shaft failing. These three inspections were removed from the table prior to analysis.

2.4. DATA PRE-PROCESSING

Before proceeding to the EXAKT modeling step, further investigation on the extracted features (called “covariates” in the analysis) was performed. First, FGP and FGP1 were compared over the test runs. FGP and FGP1 for a test run are plotted versus their timestamps on the same graph, as presented in Figure 2 and Figure 3.

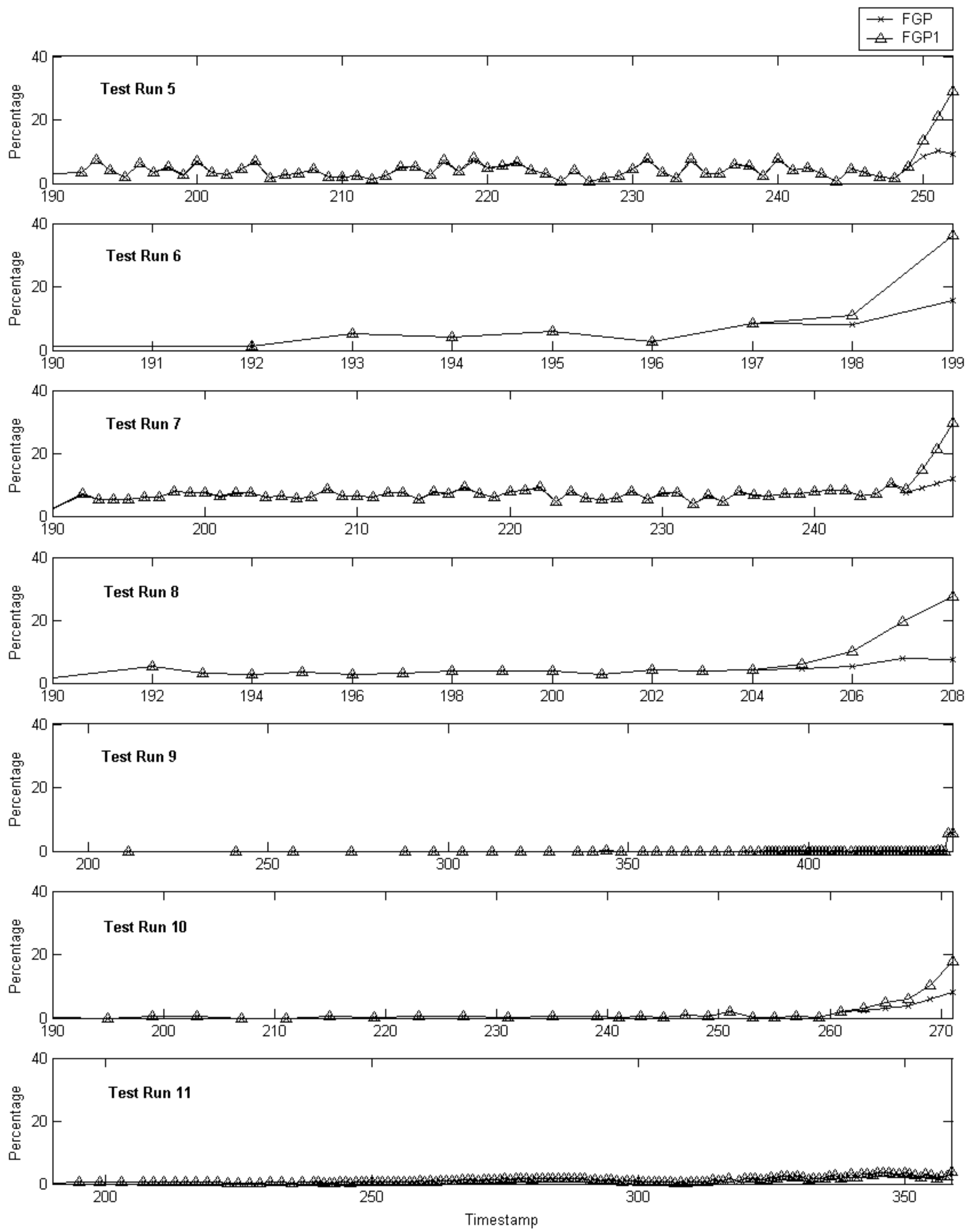

Figure 2:FGP and FGP1 vs Timestamp (Test Runs 5—11)

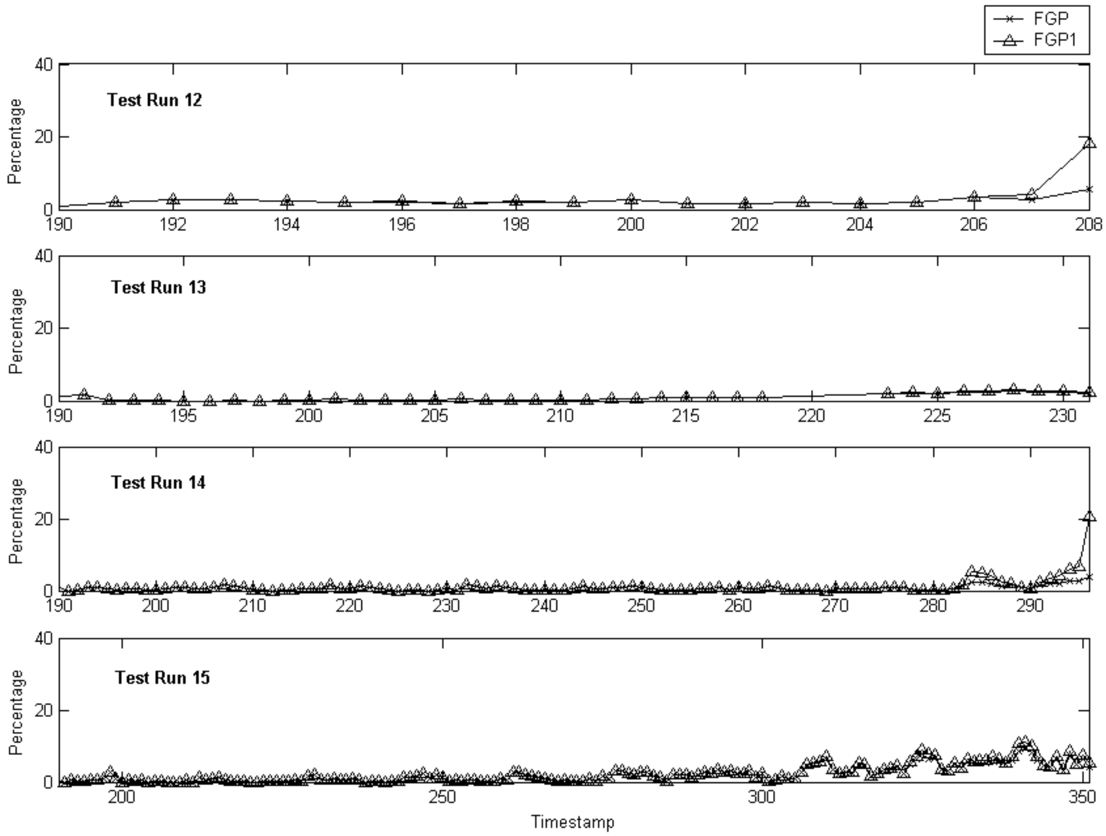

Figure 3: FGP and FGP1 vs Timestamp (Test Runs 12―15)

From the graphs, we observe that FGP and FGP1 are almost identical for Test Runs 11, 13, 15, which ended as suspensions; and that FGP1 has larger values than FGP when the timestamp is close to the end of the test for all the test runs (Test Runs 5, 6, 7, 9, 12, 14) that ended in gear tooth failure. This is not very clear for Test Run 9. The relatively low values of FGP and FGP1 (and brief warning period) prior to failure in Test Run 9 might be explained by high variation in the reference baseline of the residual error signal. Using FGP alone, it is difficult, or even impossible to distinguish between a "failed" history and a "suspended" history, (e.g., Test Run 14 and Test Run 15). Hence, we may expect that FGP1 is a better gear tooth failure indicator than FGP. It will be shown in the modeling phase that FGP1 is indeed a better indicator.

Next, correlation among the covariates was investigated for three cases: for data from Group A, for data from Group B and for the entire data set (Group A + Group B). Correlation analysis of the covariates is often useful to help in covariate selection in building a statistical (proportional-hazards) model. For each of the three cases, similar results were found, that is: FGP and FGP1 are highly correlated having a correlation coefficient of over 90%. Among the covariates RFM, RFS, RTM and RTS the correlation coefficients are over 90%.

However, the correlation between any two covariates, one from the grouping of covariates FGP and FGP1, and the other from the grouping of covariates RFM, RFS, RTM and RPS, was relatively low (correlation coefficient less than 50%). Then it may be expected that one representative from each grouping of covariates might be appropriate for inclusion as a covariate in the proportional hazards model.

3. MODELING AND MODEL ANALYSIS

Having prepared and validated the data as described, the next step is to build a proportional hazards model (PHM). The EXAKT software provides tools for selecting the covariates and building and analyzing the Weibull PHM (PHM with a Weibull baseline hazard).

The technique of PHM determines how the risk of failure, or hazard, depends on covariates. The influence of a covariate on the risk is expressed by the covariate parameters - covariate weights - which are the main outcome of the PHM analysis. The mathematical formula for the hazard at time *t* is:

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\left(\gamma_1 Z_1(t) + \cdots + \gamma_p Z_p(t)\right),$$

where t refers to working age, η is the scale parameter, β is the shape parameter, $Z_1(t), \cdots, Z_p(t)$ are the covariate values at time *t*, and $\gamma_1, \cdots, \gamma_p$ are the covariate parameters. The shape parameter reflects whether the hazard increases with the asset's working age ($\beta > 1$), decreases with the asset's working age ($\beta < 1$), or is independent of the asset's working age ($\beta = 1$).

The PHM is operating context specific. That is, if the physical asset's operating context or mechanical configuration changes, then a different failure risk model (different covariate weightings) may apply. In the following subsection, we investigate whether the PHM depends on gearbox geometry. If so, we would be inclined to build two separate PHMs, for each Group A and for Group B, rather than building a single PHM for both groups.

3.1. THE EFFECT OF GEARBOX GEOMETRY ON THE PHM

The physical configuration of the gearbox was changed in Test Runs 10-15. It is interesting to determine whether the PHM is unique to each gearbox geometry. An artificial covariate (a dummy variable) was created to denote the gearbox geometry. The value of the dummy variable is 0 if the test run is in Group A or 1 otherwise. The dummy variable and all other covariates were included in the initial PHM. Then the insignificant covariates were removed from the PHM, one at a time (as determined by their p-value), until only the significant covariates remained in the model. The dummy variable appeared to be a significant covariate, which tells us that the risk model is affected by different gear geometries. This means that we may build separate PHMs for Group A and Group B.

3.2. ANALYSIS OF GROUP A

Based on the correlation analysis discussed in Subsection 2.4, we may select either one covariate from the grouping of FGP, FGP1 or one covariate from the grouping of RFM, RFS, RTM, RTS, or one from each grouping (that is, two covariates in combination) to build a PHM. Since FGP1 outperforms FGP, we may only select FGP1 from the first grouping to build a PHM. PHMs for all three cases were examined. Other combinations were also attempted but they did not yield better results than the aforementioned three cases.

In the analyses of Group A gearboxes, six different PHMs were investigated. The results for the six models with significant covariates are presented in Table 3. Also the model with both covariates FGP and FGP1 was analyzed and, as anticipated, FGP appeared not to be significant in that combination, although it is significant on its own. This means, simply, that FGP1 provides a greater amount of useful information than FGP.

TABLE 3

PHMs built for Group A

Model	Hazard Function
FGP1	$h(t)=\frac{5.51844}{10319}\left(\frac{t}{10319}\right)^{4.51844}\exp(0.388431\cdot\mathrm{FGP1})$
FGP1, RFM	$h(t)=\frac{1}{2.79213}\exp(1.17955\cdot\mathrm{FGP1}-5.34302\cdot\mathrm{RFM})$
RFM	$h(t)=\frac{4.49062}{56160.6}\left(\frac{t}{56160.6}\right)^{3.49062}\exp(2.64606\cdot\mathrm{RFM})$
RFS	$h(t)=\frac{1}{1841840}\exp(0.113776\cdot\mathrm{RFS})$
RTM	$h(t)=\frac{1}{199259}\exp(22.8414\cdot\mathrm{RTM})$
RTS	$h(t)=\frac{9.32064}{14929.6}\left(\frac{t}{14929.6}\right)^{8.32064}\exp(69.3561\cdot\mathrm{RTS})$

The optimal decision policy is defined as one that minimizes the average cost per unit working age of replacements (preventive and reactive maintenance) (see [1]). The optimal policy was calculated for each of the six PHMs. We used an estimate of the costs of failure replacement and preventive replacement of $5000 and $1000 respectively. Alternatively, if maximum asset availability were the required optimization objective, one might apply a mean time to return to service (MTTR) of 1 week to 5 weeks respectively. A cost analysis of decision policies for the six models is presented in Table 4. The "Expected cost per in-lb-day" is the theoretical average cost per unit working age corresponding to the optimal policy. The "Average cost per in-lb-day applying the EXAKT decision policy" is the actual average cost that would have been obtained had the optimal decision policy been in force during gearbox operation.

TABLE 4

Optimal average maintenance costs for Group A

Model	Expected cost per in-lb-day	Average cost per in-lb-day applying the EXAKT decision policy
FGP1	$0.345661	$0.231
FGP1, RFM	$0.424178	$0.233
RFM	$0.346349	$0.408
RFS	$0.247331	$0.449
RTM	$0.562247	$0.494
RTS	$0.405053	$0.402

From Table 4, we see that the decision policy based on model “RFS” yields the lowest expected cost. The decision policy based on model FGP1 yields the second lowest expected cost. The question is, which model should we choose as the optimal CBM data interpretation policy. In principle, the best policy may be determined by applying all these models in practice, and then selecting one which gives the best results on average. This method, however, is not practical. There is a “cost comparison” function in EXAKT software that may be used to conveniently investigate the relative merits of alternative policies. The cost comparison in EXAKT generates the average cost per unit working age calculated when the policy is applied retroactively to the data used in the analysis. The results of the cost comparison are summarized also in Table 4. The Cost Comparison function may be considered as a final check of the statistical and decision model by reporting whether the decision model is useful, i.e., whether it improves current practice.

From the cost comparison it was found that models FGP1 and FGP1+RFM have close average cost (with average costs \$0.231 and \$0.233 respectively) and they are better than the other models. This could have been expected given the calculation methods and physical meanings of the variables. The difference between the theoretical and retroactively calculated costs (columns 2 and 3 of Table 4) may be explained by the small sample size that may affect the accuracies of parameter estimates and the average cost. We may, nonetheless, consider models FGP1 and FGP1+RFM as good models, useful for the interpretation of the gearbox vibration data. In the model FGP1+RFM, the working age appeared non-significant ($\beta = 1$). We may prefer to use FGP1 as the final model because it is simpler, having only one variable.

Model FGP1 was applied to all seven histories from Group A. The decision graphs for these tests are presented in Figure 4. If the point corresponding to the measurement (calculated as a composite covariate, that is a linear combination of covariates as obtained in the PH model) lies in the lower region of the graph, no maintenance action is recommended. In this case the expected remaining useful life (RUL), which is defined as the expected time to replacement due to either failure or preventive maintenance, is reported in the text box on the upper-right corner of the decision graph. If the point lies in the upper region, the policy recommends immediate renewal (or any appropriate action that would restore the gearbox to “as good as new” condition).

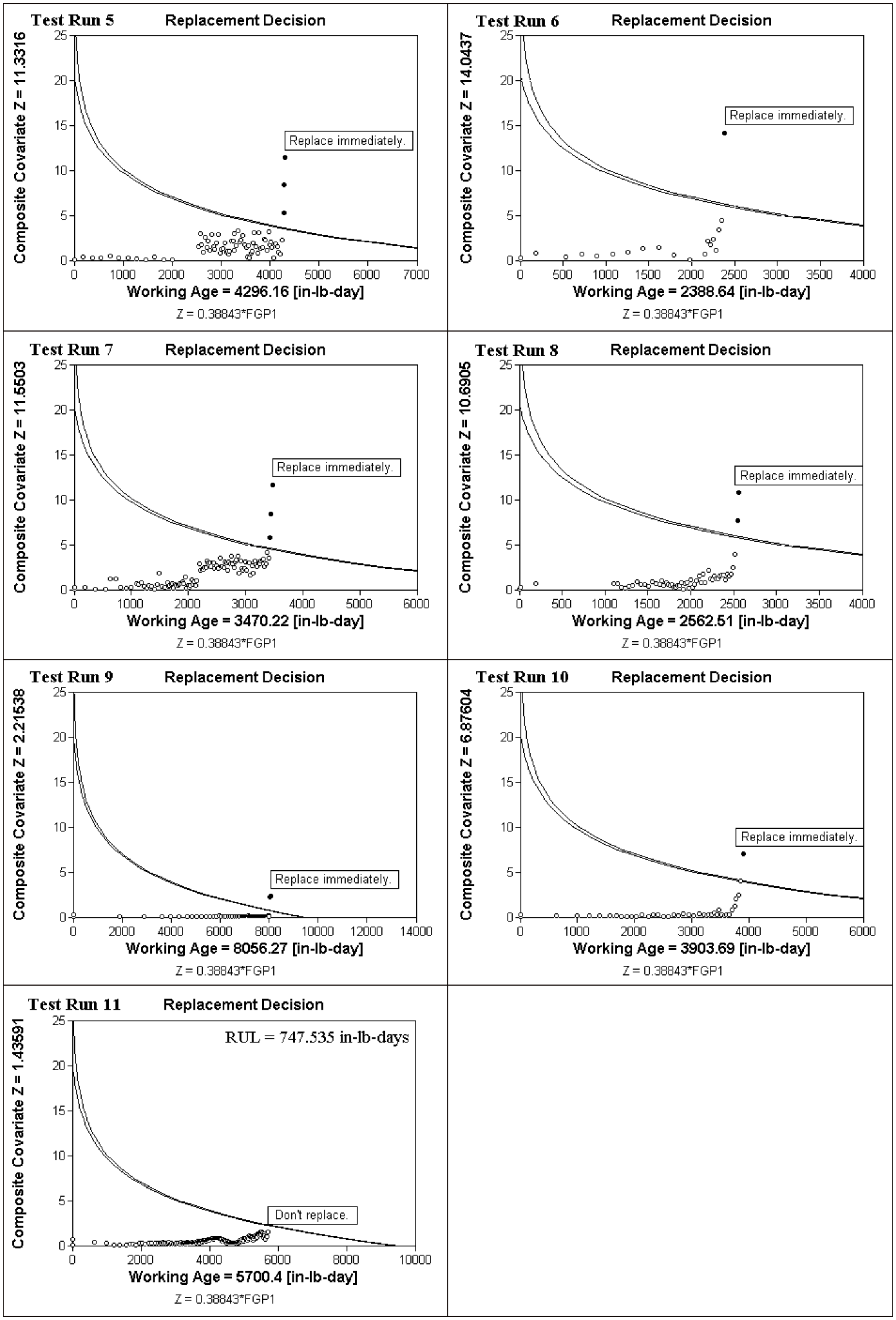

Figure 4: Decision graphs for Test Runs 5-11

From Figure 4, we observe that the application of model FGP1 would have resulted in a recommendation to renew the gearboxes (which actually failed) prior to their failure. Furthermore, we observe that no recommendation would have been made to unnecessarily remove the unfailed gearboxes. This shows that the policy appears to be realistic and good one, at least when applied to this data. Someone may argue that the policy was both estimated from, and applied to, the same data. Obviously, it is not the best method to check the policy. But it is the only one available to us with such a small sample size. If new data are available, one can apply the policy to the new data to validate the policy. Or if the sample size is much larger, part of the data may be used for the estimation of the model, and the other part for the validation of the model.

3.3. ANALYSIS OF GROUP B

Similar to the analysis of Group A, we analyzed Group B and obtained 5 tentative PHMs for Test Runs 12-15, as presented in Table 5. We may observe that in all models the working age is not significant ($\beta = 1$). The optimal expected costs per unit working age and the results of the cost comparison for all models are summarized in Table 6.

TABLE 5

PHMs built for Group B

Model	Hazard Function
FGP1	$h(t) = \frac{1}{13078400} \exp(0.819564 \cdot \text{FGP1})$
RFM	$h(t) = \frac{1}{5286380000} \exp(3.23274 \cdot \text{RFM})$
RFS	$h(t) = \frac{1}{62148900} \exp(0.229462 \cdot \text{RFS})$
RTM	$h(t) = \frac{1}{271342000} \exp(85.5958 \cdot \text{RTM})$
RTS	$h(t) = \frac{1}{927285} \exp(48.2463 \cdot \text{RTS})$

TABLE 6

Optimal average maintenance costs for Group B

Model	Expected cost per in-lb-day	Average cost per in-lb-day applying the EXAKT decision policy
FGP1	$0.181969	$0.311
RFM	$0.422995	$0.315
RFS	$0.448288	$0.700
RTM	$1.17534	$1.104
RTS	$0.274129	$0.306

From Table 6, we see that model FGP1 has the lowest expected cost per in-lb-day. The model with the second lowest expected cost is RTS. The cost comparison, however, shows that model RTS and FGP1 are close in average cost per in-lb-day when the EXAKT decision policy was applied retroactively (with average costs $0.306 and $0.311). Summarizing, we conclude that model FGP1 may be deployed as the final model. The decision policy based on model FGP1 is applied to Test Runs 12-15 and the decision graphs are presented in Figure 5. The values of FGP1 in Test Run 15 fluctuate in the later part of the test much more than in the other tests, in response to the ramping up and down of load in the test. To remove the fluctuation of FGP1 and to improve the model, we may smooth FGP1 by using some smoothing procedure such as moving average. EXAKT provides a few smoothing functions. The model was rebuilt based on the smoothed version of FGP1 and the corresponding decision graphs are presented in Figure 6. From Figure 6, we observe again that the application of the model would have resulted in a recommendation to renew the gearboxes (which actually failed) prior to their failures, and in no recommendation to unnecessarily remove the unfailed gearboxes. The same comment given at the end of the analysis of Group A applies here.

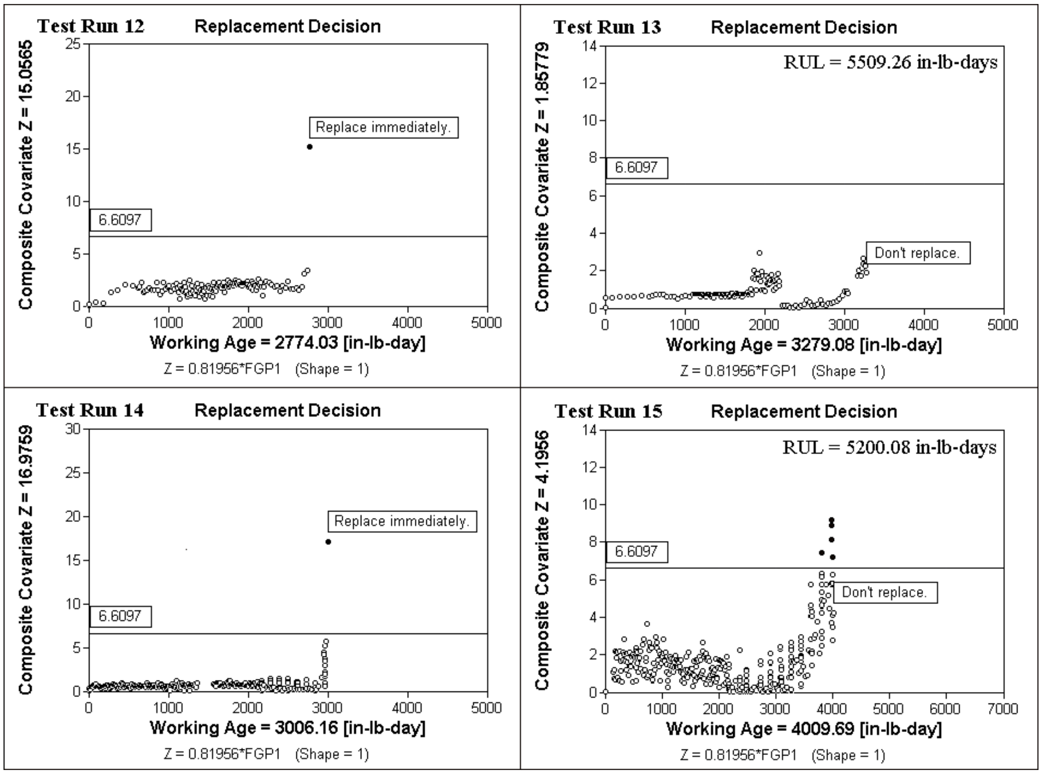

Figure 5: Decision graphs for Test Runs 12-15

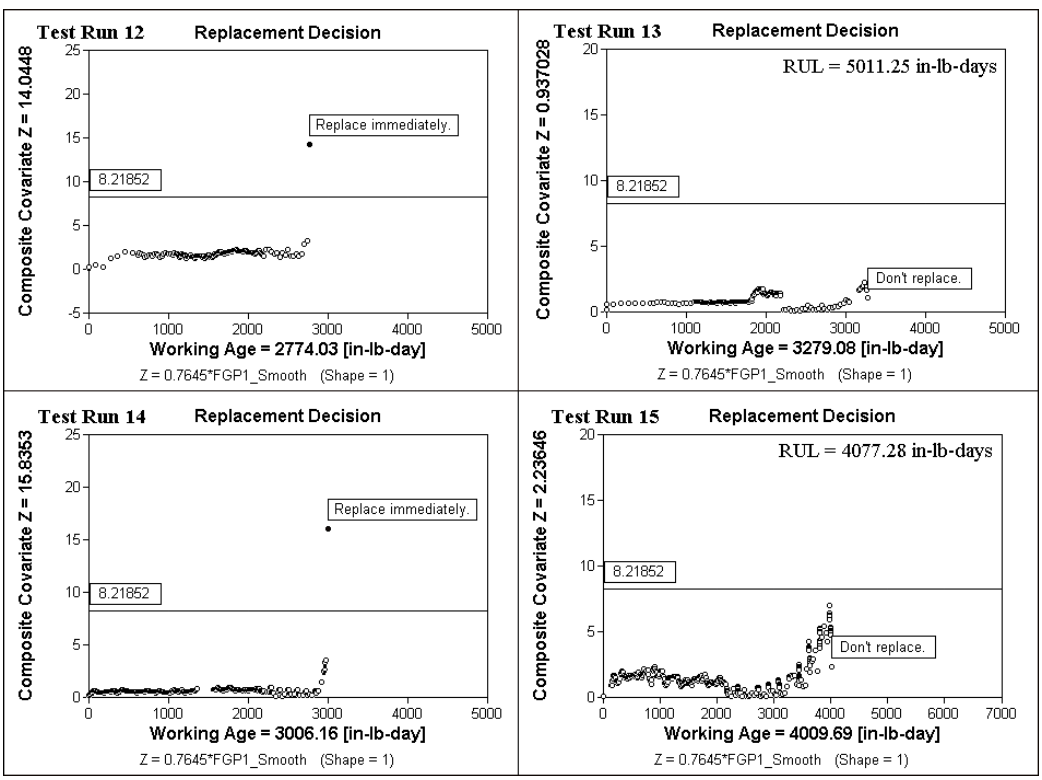

Figure 6: Decision graphs for Test Runs 12-15 using smoothed FGP1

4. CONCLUSION

Both the fault growth parameter (FGP) and its revised version (FGP1) provide a clear way to track the development of cracks (or spalls) in gear teeth. However, the point at which a replacement action should be taken to respect some sort of availability or cost criterion is not obvious. Using FGP or FGP1 in an optimization decision model results in setting a policy limit which responds to some stated economic objective for that physical asset. The objective may be, for example, to minimize the total cost of failure and maintenance, to maximize physical asset availability, to attain a certain level of reliability, or to achieve a particular performance measure such as a target ratio of planned to breakdown maintenance. It has been shown that FGP1 is superior to FGP in fault detection and decision modeling. Maintenance managers can use the methods described here as a practical way to improve the return on investment in their existing CBM programs. The sample size of the data (number of histories, not number of inspections) analyzed in this paper is relatively low. Although larger sample size would provide greater confidence, the MDTB test data was found to be adequate for demonstrating the usefulness of PHM and decision policy methodology described in this paper for predicting and preventing gearbox failures.

ACKNOWLEDGEMENTS

We are most grateful to the Applied Research Laboratory at Penn State University and the Department of the Navy, Office of the Chief of Naval Research (ONR) for providing the data used to develop this work. We also thank Bob Luby at IBM Consulting for his help to obtain the data. Finally the authors wish to thank the Natural Science and Engineering Research Council (NSERC) of Canada, Material and Manufacturing Ontario (MMO) of Canada, and the CBM Consortium members for their financial support.

REFERENCES

1. A. K. S. JARDINE, D. BANJEVIC, M. WISEMAN, S. BUCK and T. JOSEPH 2001 *Journal of Quality in Maintenance Engineering* **7**, 286-301. Optimizing a mine haul truck wheel motor's condition monitoring program: use of proportional hazards modeling.
2. J. LIN and L. QU 2000 *Journal of Sound and Vibration* **234**, 135-148. Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis.
3. V. MAKIS and A. K. S. JARDINE 1992 *INFOR* **30**, 172-183. Optimal replacement in the proportional hazards model.
4. D. BANJEVIC, A. K. S. JARDINE, V. MAKIS and M. ENNIS 2001 *INFOR* **39**, 32-50. A control-limit policy and software for condition-based maintenance optimization.
5. A. J. MILLER 1999 *A New Wavelet Basis for the Decomposition of Gear Motion Error Signals and Its Application to Gearbox Diagnostics*. M.Sc. Thesis, The Pennsylvania State University.
6. A. J. MILLER and K. M. REICHARD 1999 *Proceedings of Internoise 99, International Institute of Noise Control Engineering*. A new wavelet basis for automated fault diagnostics of gear teeth. Fort Lauderdale, FL.
7. K. M. REICHARD and A. J. MILLER 2000 *54th Meeting of the Society for Machinery Failure Prevention Technology*, 365-374. Wavelet-based filter design for gear tooth fault diagnostics and prognostics, improving productivity through applications of condition monitoring. Virginia Beach, VA.
8. W. J. WANG and P. D. MCFADDEN 1995 *Journal of Vibration and Acoustics* **117**, 363-369. Decomposition of gear motion signals and its application to gearbox diagnostics.
9. W. J. WANG and P. D. MCFADDEN 1996 *Journal of Sound and Vibration*, **192**, 927-939. Application of wavelets to gearbox vibration signals for fault detection.
10. R. YOUNG 1993 *Wavelet Theory and Its Applications*. Kluwer Academic Publishers.
11. H. ZHENG, Z. LI and X. CHEN 2002 *Mechanical Systems and Signal Processing* **16**, 447-457. Gear fault diagnosis based on continuous wavelet transform.