ENHANCING GEAR PHYSICS-OF-FAILURE MODELS WITH SYSTEM LEVEL VIBRATION FEATURES

Gregory J. Kacprzynski  
Michael J. Roemer, Carl S. Byington,  
Girish A. Modgil and Andrea Palladino  
Impact Technologies, LLC  
125 Tech Park Drive  
Rochester, NY 14623  
mike.roemer@impact-tek.com

Kenneth P. Maynard  
Applied Research Laboratory  
The Pennsylvania State University  
State College, PA 16801

Abstract: To truly optimize the deployment of DoD assets, there exists a fundamental need for predictive tools that can reliably estimate the current and reasonably predict the future capacity of complex systems. Prognosis, as in all true predictions, has inherent uncertainty, which has been treated through probabilistic modeling approaches. The novelty in the current prognostic tool development is that predictions are made through the fusion of stochastic physics-of-failure models, relevant system or component level health monitoring data and various inspection results. Regardless of the fidelity of a prognostic model or the quantity and quality of the seeded fault or run-to-failure data, these models should be adaptable based on system health features such as vibration, temperature, and oil analysis. The inherent uncertainties and variability in material capacity and localized environmental conditions, as well as the realization that complex physics-of-failure understanding will always possess some uncertainty, all contribute to the stochastic nature of prognostic modeling. However, accuracy can be improved by creating a prognostic architecture instilled with the ability to account for unexpected damage events, fuse with diagnostic results, and statistically calibrate predictions based on inspection information and real-time system level features. In this paper, the aforementioned process is discussed and implemented first on controlled failures of single spur gear teeth and then on a helical gear contained within a drivetrain system. The stochastic, physics-of-failure models developed are validated with transitional run-to-failure data developed at Penn State ARL. Future work involves applying the advanced prognostics process to helicopter gearboxes.

Keywords: Prognostics, Material Capacity, State Awareness, Gears

Introduction: Military personnel currently have limited information to evaluate their assets’ operational capability or the next mission’s effect on that capability. Thus, they must make “go”/“no go” decisions that can result in the underutilization of assets with some remaining capability and the risk of overtaxing deployed assets that have insufficient capability to meet mission requirements. The assessment of current health state and the prediction of future capability for complex mechanical systems is a core enabling technology for such decision aiding.
The focus of this current effort was to demonstrate the methodology and develop tools to provide state awareness and capability prediction on a mechanical system. The approach follows an integrated mathematical (probabilistic) framework for utilizing ALL of the available data/information that can lead to better predictions (less uncertainty) on future capability of a component/machine. This includes material level fatigue models, system level feature models and raw health monitoring measurements that provide the data/interaction between system and component levels. This mathematical framework requires a consistent normalization and probabilistic updating process/algorithm that occurs at specified triggering points in the life of a component based on indication from features. These features, extracted from various system level measurements, can be used to understand the progression of a failure (potentially with competing failure modes). Considerable work and useful results have been identified in this technical area, primarily from the CBM and machinery health monitoring communities and the intent of this paper is not to reproduce those results. Rather, the focus of this paper is on maximizing the utility of features based on their ability to track material damage progression through particular failure modes and provide statistical correlation with actual component damage levels that can then be linked with the failure mode material damage models.

**Run-to-Failure/Seeded Fault Testing:** Developing validated models for predicting asset readiness requires normal, seeded, and transitional (run-to-failure) data that have well-documented ground truth information. In general, there is a lack of existing, high fidelity data sets pertaining to fault initiation and evolution. Thus the desire for accurate prognostic tools drives the accompanying need for realistic data sets to develop asset state prediction technology (e.g., failure models, diagnostic algorithms, predictive techniques, data fusion, reasoning).

As illustrated in Figure 1, Run-To-Failure/Seeded Fault (RTF/SF) results are used to 1) provide realistic indications of state awareness, that is, health indicators (i.e. vibration features) and non-destructive evaluations related to actual material damage levels 2) identify sensitivities and uncertainties in the effects of material properties and manufacturing defects on component capacity and finally, 3) provide invaluable calibration of any range of prognostic model at various times in the life of a component so that it can be evaluated both in terms of long-term capability prediction (asset management) and more near-term damage minimization (fault accommodation).
Gear Prognostics: Under this program, two gear-based prognostic models were developed. The first was for a single spur gear tooth and validated Run-to-Failure results and deflection and acoustic emission data from a simple test rig at Penn State ARL. While quite removed from a realistic gearbox application, the purpose of the effort was to very accurately correlate a 2-D Finite Element fracture mechanics model and associated crack initiation and propagation lifing algorithms to tooth stiffness and acoustic emissions changes. This is achieved by updating/adapting material property distributions or choice of lifing algorithm during damage progression (used in the local strain stochastic fatigue models) based on measured/inferred conditions at the local failure location/region. The second prognostic model was developed for a helical gear in an industrial gearbox run to failure at ARL. Using a high fidelity model (build upon contact element and 3-D fracture mechanics FE models of the gear and a 6 DOF torsional model of the drivetrain) and well-proven vibration features, an attempt was made at predicting current and future material level damage (capability) as a function of torque from system level vibration. Thus, system level features and operational state awareness are used to ascertain current and future material capacity.

Single Spur Gear Tooth: Figure 2 shows a schematic of the gear tooth testing machine used to fail spur gear teeth.
In this study, the material was ASM-6265 and the teeth were case hardened to approximately 60-62 HRC to a depth of 0.036 – 0.044 inches. During the fatigue tests, tooth loading was along a line of contact at the highest point of single tooth contact and the gear tooth was cyclically loaded using a Haversign profile with alternating loads ranging from 300-800 min. to 7,200-11,000 max. Each test was faulted prior to complete tooth failure based on a tooth deflection limit. An ANSYS model of the tooth was used to calculate true strain at the tooth root for a number of different test cases. As shown in Figure 3, each test case was plotted on a strain-life curve to empirically determine the fatigue parameters n,c,K,and $\varepsilon_f$ used in the Eq. #1 with published data as a starting point [2].

The banding shown illustrates approximated +/- 3$\sigma$ bounds on the low, high, and combined S-N curves. With this information the lifing equation given in Eq. 1 could be employed to find the number of stress reversals to crack initiation [3].

\[ Nf_{1L} = \frac{1}{2} \left[ \sigma_{L(true)} - \sigma(mean) \right] \cdot K \cdot \frac{1}{\varepsilon} \cdot Ef \cdot \left( \frac{1}{n} \right) \]

$Nf_{L}$ = the LCF life for the gear (L)
\( \sigma_{L(true)} \) = localized true plastic stress amp. at a tooth root
\( n = \) cyclic strain hardening exponent
\( c = \) fatigue ductility exponent
\( K = \) cyclic strength coefficient
\( \varepsilon_f = \) fatigue ductility coefficient

This model was run as a Monte Carlo simulation whereby random samples were taken from each of the material property distributions and then randomly combined in the above equation. The results of this statistical crack initiation model for a mean loading of 4950 lbf. is shown in Figure 4. The actual crack initiation occurred around 3.32E8 total accumulated lbs. (as inferred from the acoustic emission level and tooth deflection shown in Figure 5) or approx. 70K cycles/140K reversals.

A fracture mechanics stochastic model and associated 2-D FE model was also created for this gear (Figure 6). While the cycles from “like new” condition to crack initiation is much less that the cycles from crack initiation to unstable crack growth, the purpose of this effort was to show that a life prediction can be updated based on state awareness (crack detected or core-to-case material transition).

The FE model automatically propagated a crack at fixed increments perpendicular to the maximum principle stress if an initial notch was created. This model provided the orthogonal threshold intensity ranges at the crack front for the Foreman model given in Eq. 2.
The Forman equation states that for Phase II and Phase III crack growth with mean stress effects [3]:

\[
\frac{\partial a}{\partial N} = \frac{C \Delta K^m}{(1 - R) K_{\text{th}} - \Delta K}
\]

(2)

where:

\( a \) = crack length
\( N \) = Number of cycles
\( \Delta K \) = Threshold intensity range
\( C, m \) = Phase II fracture characteristics
\( R \) = mean/alternating stress ratio

The threshold intensity range for a given crack length, coupled with gross estimates of mean “\( C \)” and “\( m \)” value of 6.86E-9 and 2.8 respectively, was used to obtain distributions on the number of stress cycles to the crack length at which the test was halted with a crack of 0.083” in 81,820 cycles. Thus it can be estimated that the crack propagated from microscopic size to 0.083 inches in approx. 24K reversals, which also falls within the model-based results for crack propagation. This distribution is shown in Figure 7 with the 29 failure cases represented by the underlying histogram.
In this model, the total probability of failure is defined as the Probability of crack initiation (Pi) * the Probability of the crack propagating to unstable crack growth (Pp). The observed features clearly marked the transition from crack initiation to crack growth and provided a critical piece of information for the prognostic model. Specifically, when the system features suggest a crack has initiated the probability of crack initiation can be estimated as close to unity, therefore reducing the uncertainty in the total probability of failure prediction significantly. In addition, the predicted stress intensity factor used in the model can be updated based on the relationship between the measured deflection and inferred crack length/location prediction. This update, based on the inferred knowledge on crack length/location, will allow the model to remain “informed” on the current state of damage so that improved predictions on remaining capability can be made.

In the case study developed thus far, immediately prior to the detected crack initiation, the total probability of failure could only be calculated as Pi (15%) * Pp(99.43%) = 14.91% . As soon as the crack is detected, the Probability of crack initiation (Pi) is assigned a value of unity and the Probability of failure due to propagation (Pp) is reset to an alpha error level (0.05%). Thus the uncertainty in the current component capability is reduced by nearly an order of magnitude.
Gearbox application – Helical gear: Another prognostic model was developed for a helical gear in an industrial gearbox. Details on the test rig used to generate Run-to-Failure data may be found in [4]. The prognostic concept employed here is similar to the single gear tooth model already described in the sense that system level features are used to improve the accuracy of component level life models, but different in terms of level of modeling fidelity and relevance to real-world failures.

Figure 9 outlines the architecture of this prognostic model. Standalone physics-of-failure models predict current and future material damage levels as a function of speed and torque. If available, diagnostic results from system level features (i.e. crack detected) and NDE inspection results are used in a probabilistic fusion process to update/calibrate the model predictions.
The encapsulated prognostic models were built based on 3 different FE models and 2 stochastic lifing algorithms. The models were as follows:

1) 2-D FE model w/contact elements
   Purpose: Determine the torsional forcing function as a function of rotation and effective crack length.

2) 6 DOF Torsional model
   Purpose: Transform the torsional forcing function into simulated axial vibration.

3) 3-D FE of a cracked gear tooth
   Purpose: Obtain equivalent 2-D gear stiffness reduction as a function of crack length and stress intensity ranges at the crack front (Figure 10)

The modeling efforts assumed the following:

1) Planar, elliptical crack front (3D model)
2) Point load along helical contact line (3D model)
3) Elastic stresses only (All models)
4) Crack propagation path
   empirically determined
5) Cyclic symmetry / Bore Fixity (2D, 3D models)
6) Six DOF (Torsional model)
7) Rayleigh Damping (Torsional model)
8) Perfect Alignment (2D, 3D models)
While quite advanced, several assumptions were still made in the modeling effort. Principally, the 3-D crack was assumed planar which was deemed a reasonable assumption for phase II crack growth the experience with the failures. Secondly, the contact element analysis was 2-D and thus did not accurately reflect the forcing function created from a helical gear set in mesh but still provided an improvement in the calculated forcing function over basic sinusoidal forcing functions.

Figure 11 attempts to illustrate the interaction between the models both for the development of the encapsulated prognostic model and when it is functioning. It is interesting to note that when the model is in use, it produces estimates of material damage (life used / life remaining) for a given mission profile (speed, torque). Conceptually, these predictions are updated based on how the observed vibration response differs from the expected response and how that difference relates to run-to-failure experiences.
**Feature Fusion:** As previously noted, the model-based prognosis on time to crack initiation was “calibrated” with normalized, fused vibration features. As any given test progressed, a confidence number was generated based on the Dempster-Shafer fusion of 7 vibration features. Detailed information on the vibration features employed to detect gear tooth cracking can be found in [5] while discussion on feature fusion techniques are given in [6,7].

Figure 12 show a three fusion techniques (Dempster-Shafer, Bayesian, and weighted averaging) on seven vibration features out of 26 from a run-to-failure test. The fused features indicated a crack at 108 hrs (index 269 on the graph) with high confidence over a very short time. When the normalized feature confidence number exceeded 0.5 the probability of crack initiation was set to 1 as previously described. Thus, knowledge fusion allows the uncertainty in the prognostic model to be greatly reduced.

![Figure 12](image)

**Results:** Employing the same damage algorithms used for the single tooth module, the crack initiation and propagation distributions were calculated for cases where the gearbox was operated at 300% load and 1750 RPM. These tests were denoted as #5,6,12,14, and 20. These distributions are shown in Figure 13 and Figure 14. It is important to note that the variance of the initiation distribution is much greater than the propagation distribution. In fact, the COV is 63% less for the propagation distribution. *This means that, if a crack is reliably detected by the features, the uncertainty in time at which catastrophic failure is dramatically reduced.*
The table given in Figure 15 provides the load and RPM for each run-to-failure test. The “Approx. Initiation Time” and “Approx. Time to First Failure” are the time (in hours) to when the fused vibration features first indicated gear tooth cracking and the time (hours) to when the test was either halted or a borescope inspection indicated a single tooth failure respectively. Since the load and RPM are the same for each case the model predicts a consistent mean cycles to initiation as 6.96E6. The “Percentile” is the percent the actual number of cycles to initiation was with respect to the predicted cycles to initiation. The process is performed for the crack propagation. Run 14 and 20 are distinguished from the others because an alternating, 100% to 300% load was applied after a 90 hr break-in period. While the prognostic model calculated the damage as a function of load, there was no special provision to modify the damage assessment on the basis of an alternating load.

<table>
<thead>
<tr>
<th>RTF #</th>
<th>Load</th>
<th>RPM</th>
<th>Approx. Initiation Time</th>
<th>Approx. Time to First Failure</th>
<th># Cycles Initiation</th>
<th>Mean Initiation Prognosis</th>
<th>Percentile</th>
<th># Cycles Propagation</th>
<th>Mean Propagation Prognosis</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1600</td>
<td>1750</td>
<td>122</td>
<td>127</td>
<td>1.2E+07</td>
<td>6.98E+06</td>
<td>91.2453</td>
<td>5.20E+05</td>
<td>6.90E+05</td>
<td>24.8040</td>
</tr>
<tr>
<td>6</td>
<td>1600</td>
<td>1700</td>
<td>96</td>
<td>101</td>
<td>9.90E+05</td>
<td>6.90E+06</td>
<td>81.265</td>
<td>5.20E+05</td>
<td>6.50E+05</td>
<td>23.4550</td>
</tr>
<tr>
<td>12</td>
<td>1600</td>
<td>1700</td>
<td>100</td>
<td>104</td>
<td>1.10E+07</td>
<td>6.90E+06</td>
<td>86.2139</td>
<td>1.00E+05</td>
<td>6.50E+05</td>
<td>0.0038</td>
</tr>
<tr>
<td>14</td>
<td>1600</td>
<td>1750</td>
<td>102</td>
<td>106</td>
<td>1.11E+07</td>
<td>6.90E+06</td>
<td>86.216</td>
<td>1.00E+05</td>
<td>6.50E+05</td>
<td>1.9735</td>
</tr>
<tr>
<td>20</td>
<td>1600</td>
<td>1750</td>
<td>42</td>
<td>44</td>
<td>4.41E+06</td>
<td>6.90E+06</td>
<td>29.601</td>
<td>2.10E+05</td>
<td>6.50E+05</td>
<td>0.0214</td>
</tr>
</tbody>
</table>

The results showed that both the initiation and propagation models adequately modeled these limited run-to-failure cases. In general, the initiation prognosis was not conservative enough (with 4 of the tests being in the 80 or 90 percentile). This may be due in part to the some misalignment between the pinion and driven gears (yielding higher tooth root stresses) and/or the first tooth failure occurring before the “Time to failure” value indicated. The main outliers in the propagation model are runs 12 and 20. The fast failure in run 12 may indicate that the crack propagation model needs to be slightly more conservative or could be due to inaccurate estimates of when a crack actually formed.
Conclusions: The authors have demonstrated a methodology and prognostic tool development approach to provide accurate state awareness and improved capability prediction for machinery systems. Two examples of prognostic models have been shown based on a prognostic architecture that allows features to update/calibrate stochastic physics-of-failure models. Using system level observable features to orient and update the models tracking material condition provides a sound approach for improved awareness of current health state. The real-time calibrated model is then available to predict future performance over a range of potential loads and environmental conditions. The development of prognostic tools such as those described here can be accomplished through the application of material damage models to run-to-failure data sets. Deployment of calibrated prognostic tools within a tactical decision-aiding environment will provide intelligent asset management and lead to improved readiness and mission assurance.

Acknowledgments: We would like thank Dr. Leo Christodoulou of DARPA for funding this effort as part of a Phase I SBIR program as well as many fine engineers that have contributed to this effort behind the scenes including Avinash Sarlashkar of Impact Technologies and Rob Campbell of Penn State ARL.

References:


