Remaining Useful Life Calculation of a Component using Hybrid Fatigue Crack Model

Eric Bechhoefer[1], Lei Xiao[2], and Xinghui Zhang[3]

[1]*GPMS Int. Inc., Cornwall, VT, 05753, USA*
eric@gpms-vt.com

[2]*Donghua University, Shanghai, China*
leixiao211@dhu.edu.cn

[3]*Intelligent Mobile Robot Research Institute (Zhongshan), Beijing, China*
dynamicbnt@163.com

ABSTRACT

High value asset vehicles, or vehicles where safety/operational readiness is important, benefit from an accurate remaining useful life (RUL) estimate. For these assets, RUL allows operators to realize increase revenue because of improved availability. This paper uses a hybrid algorithm based on two high cycle fracture mechanics models: a linear elastic fracture mechanics model, and the dislocation theory fracture mechanic model. Additionally, the hybrid model uses two separate Kalman filters to linearize the nonlinear component degradation process resulting in an improved RUL estimate. The hybrid model's performance is validated using prognosability, trendability and monotonicity against the two existing models using a real-world data set. The improved hybrid model allows a longer prognostic time horizon over which to marshal the resources needed for repair and give operations personnel an extended window to bring other assets to cover missions that would otherwise be unavailable.

1. THE OPERATIONAL NECESSITY FOR RUL

Preventive maintenance is designed to ensure the reliability and safe operation of equipment between servicing. The design goal of the maintenance plan is that there will be no failures caused from fatigue, neglect, or normal wear. For example, the time between overhaul on the M250 C47 engine turbine section is 2000 hours. Yet, inherent in this design, due to safety of flight, is the incorporation of two magnetic chip detectors. These sensors are installed in to detect ferrous debris in the oil indicative of component wear which could result in a failure. As a result of improper assembly, contamination, unanticipated loads, etc., even this highly reliable engine can fail prematurely. For a helicopter, the chip detector is in effect, a "on condition" indicator triggering an immediate and disruptive maintenance event.

An online condition monitoring system can extend the operational readiness, reliability, and safety of the system by giving a longer timeline prior to a potentially failure and an unplanned maintenance event. As such, extending the estimate as to when unscheduled maintenance should be performed requires an accurate prognostic capability. Calculating the remaining useful life (RUL) allows operation and maintenance personnel to better schedule assets and logisticians to order long lead-time parts. This preemptive action helps improve asset availability. In addition, higher asset availability allows revenue to be generated through missions/operations that would otherwise be lost because the asset was down for maintenance. While several RUL techniques have been reported, this proposed hybrid model improves upon exiting, model-based RUL algorithms. The term improvement implies that the estimate of RUL is more consistently accurate over time (prognosability, monotonicity and trendability). This improvement was achieved using two novel ideas.

As shown in "Contending Remaining Useful Life Algorithms" (Bechhoefer, Dube, 2020) three high cycle fracture mechanics theories were used to design RUL models. While these models were powerful, no single model fit the data over the entire 700 hours of damage propagation. Dr. Kai Goebel suggested that a combination of models might work better. Dr. Goebel's comment led to the hypothesis that fatigue damage, such as spalling of a bearing, gear tooth root bending, or shaft coupling failure, could result from a combination of degradation modes.

Fatigue crack growth can be characterized as a Mode 1 (opening mode) failure (Beer, 1992), where the crack surface is forced directly apart. The Linear Elastic fracture mechanism characterizes this type of failure model. Alternatively, in a Mode 2 failure model, the crack surfaces move normal to the crack and remain in the crack plane. Head's theory is one such model that used this failure mode. Finally, in a Mode 3 model, the crack surface moves parallel to the crack front and remains in the crack plane. This mode is characterized by dissociation theory (Beer, 1992).

The hypothesized improvement was to realize that complex components, such as gears/bearings, which have complex shapes, suffer from fatigue damage as a combination of failure modes. Hence, a better model for fatigue failure is a model that is a combination of Mode 1 and Mode 3 damage (as Mode 1 over estimated damage propagation while Mode 3 underestimated the propagation. Mode 2 was not considered, as its performance was worse than the Mode 3 model, see Bechhoefer, Dube, 2020). Clearly, an improved damage model allows for a better prediction of crack growth. This, in turn, allows for an improved estimate of the RUL.

The other improvement is an observation that while crack growth is nonlinear (e.g., the rate of change of a crack length grows faster as the crack length increases), the RUL is linear. That is, for any relatively constant load, the RUL rate of change (d_{RUL}/d_t) is approximately -1 when the model was converged (for both Mode 1 and Mode 3). For example, if the life of a component is given as 100 hours, after one hour of usage, the life should logically be 99 hours. This seemingly obvious observation (as measured in Bechhoefer and Dube 2020) can be used mathematically to improve the estimate of the RUL calculation as it is an extra observable when using state reconstruction algorithms.

2. Background on the Hybrid Model

For many vehicles (especially in aerospace), the manufacture determines the inspection and overhaul schedule. Installation of condition monitoring equipment (such as Health and Usage Monitoring Systems – HUMS) via a supplement type certificate (STC), will not change these maintenance intervals. The goal of HUMS in these applications is to reduce unscheduled maintenance while improving safety and availability.

In general, helicopters have inspections every 50 hours of flight time, with heavier maintenance conducted at 100 and 300 hours. Aircraft also have annual inspections. Typically, the number of hours flown per month is dependent on the operator's mission. It is not surprising to see fleets that average 300 to 500 hours per annum. Of course, for operators that fly inspection (inventorying power poles and examining power lines for encroachment) or other seasonal missions (firefighting, by dropping water, or delivering man/material to a fire), these aircraft can fly as much as 25 to 40 hours per week.

The importance of having an accurate RUL calculation is to support and supplement the already established maintenance practices. HUMS with an RUL capability allows the fleet operation manager to order parts, schedule the right personnel to perform maintenance. This turns an unscheduled maintenance action into scheduled maintenance. For example, as in figure 1, this M250 No. 5 bearing has been trending for 50 hours. The operations manager knows that in 70 hours, this engine bearing will need to be replaced. During this 70-hour interval, the aircraft and engine manufacture can be notified. The replacement engine can be ordered and at the convivence of the helicopter operator (e.g., while performing other scheduled maintenance), the repair can be performed opportunistically with other existing maintenance requirements. This prevents a potential unscheduled maintenance activity in the future, which improve availability, and hence, revenue.

For the light helicopter market or cases where the aircraft has no extended overwater flights, HUMS provides logistic support to improve availability (e.g., allowing the generation of revenue flights) rather than the safety of flight requirement. The aircraft, having been type certificated and adequately maintained, is inherently safe. The worst possible outcome for the bearing fault in figure 1 would be that the pilot sees a chips light (annunciator indicating metal debris in the gearbox) and is forced to land. As noted, a goal of HUMS is to generate a maintenance action before a chips light, so that this maintenance is done opportunistically while the aircraft is already down for some required scheduled inspection.

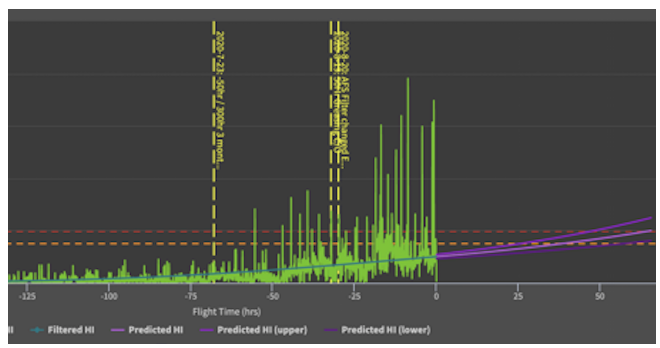

Figure 1 No. 5 Engine Bearing Propagating Fault

3. Prognostics Calculation of the RUL

A prognostic based on a fracture mechanics model, requires four inputs to calculate a RUL.

1. An estimate of the current component health.
2. An estimate of when it is appropriate to do maintenance, e.g., the threshold.
3. An estimate of the future component load.
4. A component degradation process model takes the current component health, the estimated future load and calculates the time/cycles to when it is appropriate to do maintenance.

HUMS typically acquires vibration data at an appropriate aircraft regime. An appropriate regime would be straight and level, at a benign flight envelope, for example, hover, 60 or 120 knots. Because the airframe is relatively flexible, it is not appropriate to acquire data while the aircraft is maneuvering, as those maneuvering loads are transferred to the drivetrain and can corrupt the acquisition measurements.

When in a benign regime, through signal processing techniques based on component configuration, the HUMS can calculate condition indicators (CI) representative of component health. Ideally, these condition indicators are also proportional to the extent of the component damage. This allows HUMS to provide, effectively, a virtual inspection of the drivetrain components. For a review of condition monitoring algorithm, Vercer (2005) is an excellent reference for many common condition indicators.

The estimate of when it is appropriate to do maintenance is a threshold-setting problem. A hypothesis testing approached has been adopted for this paper (Bechhoefer, 2005, 2011). In this paradigm, the measured set of condition indicators are used to provide evidence that the component is no longer good. That is, the CIs are used to reject the Null Hypothesis that the component is nominal. If the component is not nominal it is appropriate to perform maintenance.

While many classification problems have been solved using artificial neural networks (ANN) or deep learning techniques, this problem is better suited to this hypothesis testing process. Data-driven techniques such as ANN excel in the classification using symmetric datasets (e.g., there is a training set of known nominals and known faults to train against). This is not the case with many condition monitoring applications, particularly, helicopters. Because of the helicopters' high reliability, there are few, if any, training sets of know faults. Further, because of the large number of shafts, gears, and bearings within the helicopter drivetrain, it is highly unlikely that there will ever be a symmetric dataset for training.

Consider just the power turbine section of the M250 engine. It contains five shafts (with three failure modes for each shaft), six gears (with at least six different failure modes) and 13 bearings (with, at least four failure modes, each). For a comprehensive and symmetric dataset, one would require five × four (one extra for the nominal case) shaft example, six × seven gear sets and 13 × five bearings sets for a total of 127 different cases.

Hypothesis test functions in the domain of asymmetric data. The dataset is asymmetric as most if not all the training set are nominal data – that is, there is data from a health gearbox on a fielded aircraft. In this environment, one is not concerned with Type II errors (missed detection) as there are so few fault examples to test with. Instead, we employ hypothesis testing, using a specified by Type I error – the probability of fault alarm. This is the condition customarily found in condition monitoring, as the vast amount of data collected is nominal and can be used to establish the probability density function (PDF) of the measured CIs.

The difficult of asymmetric nature of helicopter data can be seen, for example in the Bell 407 drivetrain. This aircraft has 30 bearings, as noted with at least four fault modes. The aircraft has 22 shafts and 15 gears (at least six failure modes: chip, crack, pitting, micro pitting, scuffing, wear). The total number of training cases for this gearbox would be at least 343 conditions if approaching this as a classification problem.

To restate, due to the high quality of the gearbox, and regular maintenance, failures are rare. This suggests that training sets would require seeded fault testing – which is not practical due to the large number of shafts, gears and bearing in the helicopter drivetrain. As such, hypothesis testing is a good option. Typically, enough data is needed to estimate the condition indicator covariance matrix and PDF, in perhaps 50 to 100 data points. This can represent three hours of flight time for a helicopter, assuming 20 to 30 acquisitions per hour

3.1. Generation of Maintenance Appropriate Thresholds

In the context of hypothesis test, it is observed that all condition indicators (CIs) have a PDF. Any operation on the CI to define a health index (HI) is then a function of distributions. The HI function in the application is the weighted norm of *n* CIs (e.g., the normalized energy of *n* CIs), where the weights are determined by the Jacobian (the inverse covariance):

$$HI = {}^{0.35}/_{critical} \sqrt{\boldsymbol{Y}^T\boldsymbol{Y}} \tag{1}$$

where $\boldsymbol{Y}$ is the whitened, normalized array of CIs, and *critical*, is the critical value of the test. In a hypothesis test, the critical value is calculated from the inverse cumulative distribution function (ICDF) for a given probability of false alarm. For Eq. (1), the ICDF is the Nakagami where η is the number of CIs in the array and = n, and $\omega = \eta/(2-\pi/2)*2$, see Bechhoefer 2011 for the proof. A normalized HI > 0.35 for a component, indicates that the Null Hypothesis is rejected. That is, the component is no longer nominal. Note however that maintenance is not recommended until the HI > 1. These threshold values have been tested numerous helicopters, wind turbines, and seeded fault testing on 60+ gearboxes. The level of damage for an HI of 1.00 is typically moderate visible damage.

This function Eq (1) is valid if and only if the distribution (e.g., CIs) are independent and identical (e.g., IID). For example, for Gaussian distribution, subtracting the mean and dividing by the standard deviation will give identical *Z* distributions. The issue of ensuring independence is much more difficult. In general, the correlation between CIs is non-zero. This has been measured on numerous tests, see Bechhoefer 2011.

This correlation between CIs implies that for a given function of distributions to have a threshold that operationally meets the design probability of false alarm (PFA), the CIs must be whitened (e.g., de-correlated).

Whiting ensures that the realized false alarm rate of a fielded system will be the designed PFA. Consider a thought experiment where the HI is the sum of two CIs (*X* and *Y*) that are Gaussian with $\sigma = 1$. Then the standard deviation of $X+Y$ is:

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y} \tag{2}$$

Here ρ is the correction between *X* and *Y*. If the correlation is near 0, the standard deviation is sqrt(2). However, if the correlation is nonzero (or worst case near 1) the standard deviation is 2. Hence, the observed PFA (which is based on the

HI standard deviation) of the HUMS when the data is correlated with will be much higher than designed. For example, if independence is assumed with a PFA of 10^{-6}, and in fact the X and Y measurements are highly corrected, the observed PFA would be: 3.4^{-4}. The observed PFA would be or 388x greater than designed. This would result in needless maintenance and loss in system confidence.

It can be shown that a whitening solution can be implemented using Cholesky decomposition (Bechhoefer 2011). The Cholesky decomposition of Hermitian, positive definite matrix results in $\mathbf{A} = \boldsymbol{LL}^*$, where $\boldsymbol{L}$ is a lower triangular, and $\boldsymbol{L}^*$ is its conjugate transpose. Thus, by definition, the inverse covariance is positive definite Hermitian. It then follows that if:

$$\boldsymbol{LL}^* = \boldsymbol{\Sigma}^{-1}, \text{ then } \boldsymbol{Y} = \boldsymbol{L} \times \boldsymbol{CI}^{\mathrm{T}} \tag{3}$$

The vector CI is the correlated CIs processed due to data acquisition on the aircraft, which are used for the HI calculation. The transformed vector $\boldsymbol{Y}$ is 1 to n independent CIs with unit variance (one CI representing the trivial case). The Cholesky decomposition, in effect, creates the square root of the inverse covariance. This, in turn, is analogous to dividing the CI by its standard deviation (as in the case of one CI). It can be shown that $\boldsymbol{Y} = \boldsymbol{L} \times \boldsymbol{CI}^{\mathrm{T}}$ then creates the necessary independent and identical distributions required to calculate the critical values for a function of distributions.

The critical (*critical*, Eq. (1)) value is taken from the ICDF for the HI. The CIs used are assumed to have Rayleigh-like PDFs (e.g., heavily tailed). For magnitude-based CIs, it can be shown that for the nominal case, the CI probability distribution function (PDF) is Rayleigh (Bechhoefer, 2005, 2011). For Gear CIs and Bearing CIs (where magnitudes are biased by root mean square (RMS)), a transform is used to make the CI more Rayleigh. The transform "left shifts" the CI. For example, a shift such that the .05 CDF (cumulative distribution function) is assigned to 0.0.

Consequently, the HI function is based using the Rayleigh distribution. The PDF for the Rayleigh distribution uses a single parameter, β, defining the mean $\mu = \beta\,(\pi/2)^{0.5}$, and variance $\sigma^2 = (2 - \pi/2)\,\beta^2$. The PDF of the Rayleigh is: $x/\beta^2 \exp(x/2\beta^2)$. Note that when applying these operations to the whitening process, the value for β for each CI will then be: $\sigma^2 = 1$, such that:

$$1.526 = 1/\sqrt{2 - \pi/2} = \beta \tag{4}$$

For the HI equation in (1), the normalized energy of the CIs, it can be shown that the function defines a Nakagami PDF (2011). As note previously, the statistics for the Nakagami are $\eta = n$, and $\omega = 1/(2\text{-}\pi/2) \times 2 \times n$, where n is the number IID CIs used in the HI calculation.

4. The RUL Calculation

For this application, RUL is taken as the time when it is appropriate to do maintenance and *not* the time until the component fails. For aviation application, maintenance is a process to restore the equipment to the original design reliability. Worn or damaged parts have reduced reliability. Maintenance repairs those parts and restores the design reliability to the system. The concept that an HI exceeding 1 triggers a maintenance event is complementary to existing maintenance practices as it is design to restore the system's reliability to the manufactures design requirements.

For an example of a critical system, the design reliability is typically "six-nines," e.g., the probability of failure of the part under design loads is less than 10^{-6} per hour. For the damaged part, the reliability may be reduced to three-nines or a probability of failure of 10^{-3}. Thus, the appropriateness to repair the faulty component can be seen as an action to restore the designed reliability of the system. From a maintainer perspective, then:

- HI reflect the current components damage, where the probability of exceeding an HI of 0.35 is the PFA.
- A warning (yellow) alert is generated when the HI is greater than or equal to 0.75. Therefore, maintenance should be planned by estimating the RUL until the HI is 1.0.
- An alarm (red) alert is generated when the HI is greater than or equal to 1.0. Continued operations could cause collateral damage.
- This threshold setting model ensures that the probability of false alarm is exceedingly small when the HI reaches 1. However, from numerous installations and seeded fault tests in practice, a bearing at HI 1 has easily seen physical damage.

A component with a HI value does not define a probability of failure for the component nor that the component fails when the HI is 1.0. Instead, defining maintenance at an HI of 1 initiates a proactive policy to change operator behavior. The desire is to reduce cost and time associated with component failure by performing maintenance prior to the generations of collateral or cascading faults. As an example, by performing maintenance on a bearing before the bearing shedding extensive material, costly gearbox replacement can be avoided, and the reliability of the gearbox can be restored to its design requirements.

Hence, the RUL is defined as the time from the current HI until the HI is greater than or equal to 1.

4.1. The Linear Elastic Fracture Mechanics Model

For many materials, such as steel used in gears and bearings, which are subject to tensile loading cycle, the fatigue crack growth is Mode 1 and can be expressed as:

$$\frac{da}{dN} = D(\Delta K)^m \tag{5}$$

where

- da/dN is the rate of change in the half crack length per cycle
- D is a material constant
- m is the crack growth exponent for steal is 4.

Substituting in ΔK:

$$\frac{da}{dN} = D\left(2\sigma(\pi)^{1/2}\alpha\right)^m a^{m/2} \tag{6}$$

Inverting and integrating to get *N*, the number of cycles gives:

$$N = \int_{a_0}^{a_t} a^{-m/2} \Big/ D\left(2\sigma(\pi)^{1/2}\alpha\right)^m da \tag{7}$$

By taking *a* as a_0 to get the crack growth rate, the constants cancel out, leaving:

$$N = -dN/_{da}\; a_0 - a_f\left(a_0/a_f\right)^{m/2} \Big/ {m/_2 - 1} \tag{8}$$

Setting *m* to be 4, this gives:

$$N = -dN/_{da} \times a_0 \times ln\left(1/_{a_0}\right) \tag{9}$$

For constant rate machines, such as a helicopter gearbox, *N* is proportional to time.

We substitute the measured component health (the HI) for a_0, as it is proportional to damage. As our rule is to perform maintenance when the HI is 1, Eq. (9) then define the RUL estimate.

Eq (5), includes a term for strain. Strain is the cyclic load imparted on a component and cancels out eq (7). Inherent in this, is the assumes that the load at a_0 is constant. In general, load varies with time, but on average, similar missions have similar load.

For example, the torque, on average for a given mission, are very similar from one operation (start, mission, landing) to another mission. However, to account for a more aggressive mission, a correction factor as a percent change in mean load can be applied to the *-dN/da* term. For example, a mission that requires a sling load may be 10% more aggressive than a typical mission. The RUL for these more aggressive missions is then -*dN/da* × 1.10.

4.2. Dislocation Theory Fracture Mechanics Model

In the case where crack loading is in the anti-plane strain (e.g., Mode 3), the plastic zone at the crack tip can be represented as a continuously distributed array of small dislocations on the crack plane. It is assumed that crack growth occurs when the accumulated plastic strain distribution at the crack tip exceeds some critical value and continues as this value is exceeded at the crack tip. The rate at which the crack grows per stress cycle in terms of displacement leads to:

$$da = \frac{a^2\sigma_{max}^4}{DE\sigma^3} \tag{10}$$

This is similar to Eq. (6), with an exponent of 2. Inverting, integrating, and changing terms gives:

$$N = -dN/_{da} \times a_0 \times \left(2a_f - 2\sqrt{a_0}\right) \tag{11}$$

Again, *N* is proportional to time and the crack length is *a*.

4.3. The Combined Mode Fracture Mechanics Model Equations

The health paradigm is adopted where the RUL is the time from the current health to and health of 1. Then, by hypothesizing that fatigue failure is a combination of fracture modes, one can combine (9) and (11) to give a hybrid fracture mechanics model equation of:

$$N = -dN/_{dHI} \times HI \times \left(-2 + 2\sqrt{HI}\right) \tag{12}$$

Note that *N* is usually a cycle or count, but for many machines that operate at constant RPM (such as a helicopter), *N* is proportional to time. Note that in the mechanization of a solution, a_0 is the current health (HI), and *dN/da* is the inverse derivative of HI (e.g., *dHI/dt*). Hence, the calculation of the derivate is an essential part of solving for the RUL. By assuming the central limit theorem and its implied Gaussian noise, an unbiased estimate of the HI and *dHI/dt* can be calculated with a Kalman filter.

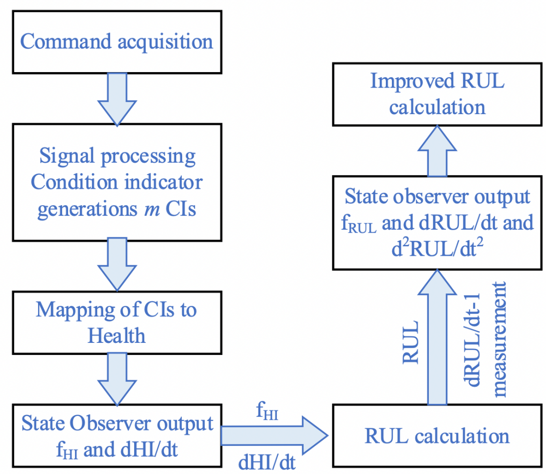

Figure 2 Process flow of RUL Estimation

4.4. Calculation of RUL with Two Inputs

Past mechanization of RUL has used state observer (for example, Kalman Filter, or Alpha-Beta-Gama filter) to give an unbiased estimate or RUL (e.g., *N* in Eq. (12)), and its first and second derivative. This hybrid model expands the measurement matrix to include inputs for the RUL and uses -1 as a measurement of the RUL derivative. As noted, this extra input, which has not been reported in earlier works, allows a second input to improve the observability of the state observer (figure 2).

Combining the improved RUL model using a combination (e.g., hybrid model) of fracture mechanics models and the use of a second RUL observable as -1 for dRUL/dt significantly improve the accuracy of the prediction early in the fault. Figure 3 shows a comparison of models, at time -450 hours. Note that on this fault, which is a high-speed bearing, both the Linear Elastic and Dislocation models overshot ideal RUL from -600 to -300 hours. The hybrid model is much closer to the ideal

RUL. This is due to the extra observable input of -1 for dRUL/dt. This can be seen the health indicator trend plot at approximately -520 hours (figure 4).

Note the performance of the models from -300 to 50 hours. Here it is seen that the hybrid model compares favorably with the existing models. We can quantify the accuracy by looking at the mean and standard deviation of the error between each model and the ideal RUL. The ideal RUL decreases by 1 for each hour of usage.

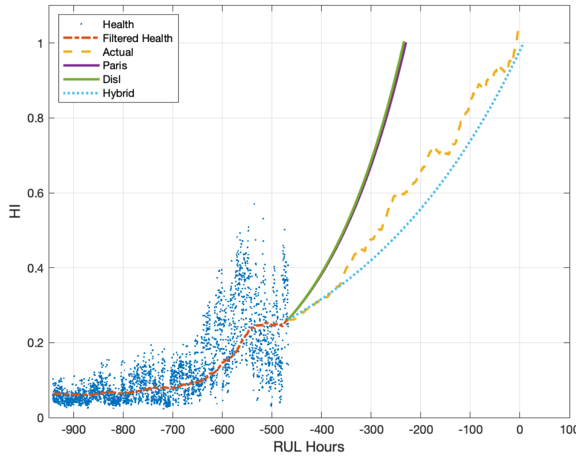

Figure 3 Trend Plot at -520 hours showing RUL Overshoot

The extra observable of *-1 dHI/dt* stabilizes and improves the RUL estimate, as can be seen in figure 4.

5. Quantifying RUL Performance

One measure of accuracy performance is to measure the mean error and standard deviation of error over the period. Here we see that the (Table I) that both the mean and standard error of the RUL. This is a rather gross measure of performance, although it does indicate that on average, the standard error is lowest for the hybrid model.

TABLE I. STATISTICAL COMPARISON OF THREE MODELS

Model	Mean Error	Std of Error
Linear Elastic Model	2.44 hours	18.22 hours
Dislocation Model	-11.68 hours	14.00 hours
Hybrid Model	-1.96 hours	10.72 hours

The mean error and standard error may not adequately capture the dynamics and complexity in RUL over time. Therefore, a more satisfactory metric was given in by Dr. Coble (2010), who introduces the concept of prognosability, monotonicity and trendability as RUL performance metrics.

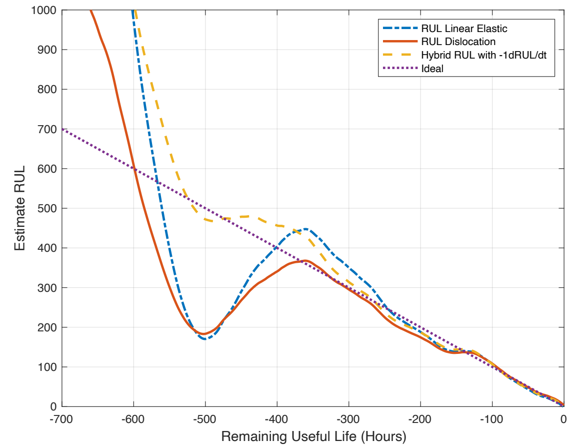

Figure 4 Linear Elastic, Dislocation and Hybrid RUL Models

Prognosability is usually defined as the deviation of the final failure values for each path divided by the mean range of the path. This would be exponentially weighted to give the zero to one scale. However, for this study, it was modified to capture the evolution of the RUL by taking the standard deviation of the RUL and dividing it by the ideal RUL value:

$$I_{prog} = exp\left(-{}^{std(Estimated\ RUL - RUL)}/_{RUL}\right) \quad (13)$$

In this model, if the estimated RUL is close to the Ideal RUL, the resulting exponent is small, returning a prognosability close to 1. The mean prognosability was calculated from -700 hours to 1 hour (Figure 5).

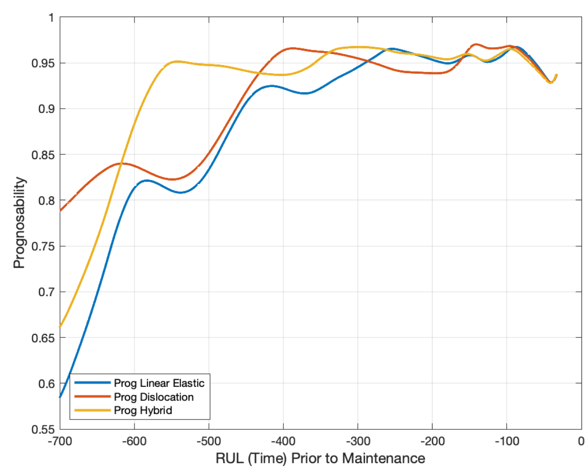

Figure 5 Prognosability of the Contending Models

Monotonicity is usually adopted to capture the underlying positive or negative trend of a series of health indicators. It is also used to evaluate the fitness of the extracted health indicator for RUL prediction. In [8], monotonicity can be measured by:

$$I_{\text{monotonicity}} = \text{mean}\left(\left|\frac{h^{+}_{\text{indicator}}}{n-1} - \frac{h^{-}_{\text{indicator}}}{n-1}\right|\right) \quad (14)$$

Here, n is the number of observations in a particular history. The monotonicity of a series of RUL is given by the average difference of the fraction of positive ($h^{+}_{indicator}$) and negative ($h^{-}_{indicator}$) derivatives for each RUL estimate.

Alternatively, trendability is defined to indicate the degree to which a series of RULs have the same underlying shape. In Coble 2010, trendability is given by the smallest linear correlation across a series RUL series as,

$$I_{\text{trendability}} = \min\left(\left|r_{\text{corrcoef}}\right|\right) \tag{15}$$

Initially, trendability was characterized by comparing the fraction of positive first and second derivatives in a series of health indicators. It should have a different functional form compared with Eq. (14). Even though Eq. (15) is an available measure of trendability, it cannot form the trend over the time horizon of the RUL, as the RUL prediction is a time-series regression problem. In consideration of the trendability definition, we propose a new measure for trendability as,

$$I_{\text{trendability}} = \frac{N^{-}_{H^{i*}_{tw} - H^{j*}_{tw}}}{N_{TW}} \tag{16}$$

Where TW means the size of a given time window. H^{i*}_{TW} is the average value of RUL in the i^{th} time window. H^{i*}_{TW} is then the corresponding values in the adjacent time window. N_{TW} indicates the total number of time windows. The numerator in Eq. (16) indicates the number of negative values from the comparison of average health indicators in each two adjacent time windows among N_{TW}.

The performance metrics (prognosability, monotonicity and trendability) compare the three different RUL extracted from the contending models. The results are listed in Table 2. Here we see that the prognosability, monotonicity and trendability from the hybrid method shows better performance than the linear elastic or dislocation methods. This supports the hypothesis that for complex shaped components the failure propagation is some combination of Mode 1 and 3 stresses. The result is that the RUL from the hybrid method is more monotonic and trendable, and hence, higher accuracy.

Table II. Comparison the RUL performance (prognosability, monotonicity, and trendability) of different models

MODLE	**HI from linear elastic method**	**HI from dislocation method**	**HI from hybrid method**
PROGNOSABILITY	0.8826	0.9095	**0.9219**
MONOTONICITY	0.5647	0.5528	**0.8495**
TRENDABILITY	0.7857	0.7810	**0.9286**

From Eq. (16), the trendability value is related to the time window. In Table II, the time window is set to 20. To stress the variation impacted by the time window, the time window is changed from 10 hours to 200 by 10 increments in each step; the results are shown in Figure 6. As can be seen, the hybrid method always shows a higher trendability value compared with the contending methods.

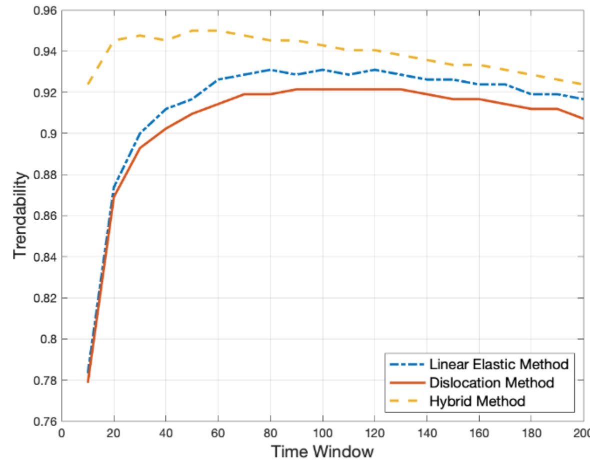

Figure 6 Trendability of Contending RUL Models

6. CONCLUSION

Presented is a physics-based model of a high cycle fatigue remaining useful life (RUL) algorithm. This model-based approach has the advantage of not needing extensive training from exemplars/fault data sets, as it is based on the relationship between processes noise of a nominal component and the rate of change of component over time. Three models were compared: the linear elastic model, a dislocation theory-based model, and a hybrid model, which combines both the linear elastics and dislocation model features. It was hypothesized that due to the complex shape of the component (e.g., a gears/bearing) that fractures due to high cycle fatigue have multiple propagation modes (Mode 1 and Mode 3). Further, it was hypothesized that the rate of change of the RUL should be -1 (e.g., the RUL decrements by one hour for each hour of life that is consumed). The second assumptions allow -1 to be used a as measurement the Kalman filter, increasing the state observability. This in improved the performance of the RUL calculation.

Using prognosability, monotonicity and trendability as accuracy metrics, the performance of these models' RUL was compared.

It was found that the hybrid model (combination of the linear elastic and dislocation theory models, with a secondary observable of -1 as a measurement) gave a much better estimate of RUL, especially during the initial fault propagation. Furthermore, as the RUL decreases, the extra observable reduces the standard deviation of the RUL estimate over other models and improves both the monotonicity and trendability of the RUL. This improved estimate and reduced variance allow maintainer and operators to schedule their assets better to improve availability and increase revenue.

REFERENCES

Bechhoefer, E., & Dube, M. (2020). Contending Remaining Useful Life Algorithms. Annual Conference of the PHM Society, 12(1), 9.

Beer, F., Johnston, P., Mechanics of Materials, McGraw-Hill, New York, 1992.

Vercer, P., Kreidl, M., Smit, R., "Condition Indicators for Gearbox Condition Monitoring Systems," Acta Polytechnica, vol. 45, no. 6, 2005.

Bechhoefer, E., Bernhard, A., "Use of Non-Gaussian Distribution for Analysis of Shaft Components," IEEE Aerospace Conference, 2005.

Bechhoefer, E., He, D., Dempsey, P., "Gear Health Threshold Setting Based On a Probability of False Alarm," Conference of the Prognostics and Health Management Society, 2011.

N. Frost, K. March, and L. Pook, Metal Fatigue, Dover Publication, Mineola, NY, 1999.

Bar-Shalom, Multitarget-Multisensor Tracking: Application and Advances, Artech House, 1992.

Coble, Jamie Baalis, "Merging Data Sources to Predict Remaining Useful Life – An Automated Method to Identify Prognostic Parameters" Ph.D. diss., University of Tennessee, 2010. http://trace.tennessee.edu/utk_graddiss/683

BIOGRAPHIES

Eric Bechhoefer received his B.S. in Biology from the University of Michigan, his M.S. in Operations Research from the Naval Postgraduate School, and a Ph.D. in General Engineering from Kennedy Western University. He is a former Naval Aviator who has worked extensively on condition-based maintenance, rotor track and balance, vibration analysis of rotating machinery and fault detection in electronic systems. Dr. Bechhoefer was a board member of the Prognostics Health Management Society and a senior member of the IEEE Reliability Society.

Lei Xiao is an assistant professor with School of Mechanical Engineering, Donghua University, China, since 2018. Her research interests include the key techniques in prognostics and health management (PHM), such as weak fault detection, remaining useful life prediction and the joint optimization of maintenance and production scheduling/spare part purchasing. She got her PhD degree from Chongqing University in 2016 and then worked as a postdoctoral researcher at Shanghai Jiao Tong University. She has published many papers on Reliability Engineering and System Safety, Journal of Sound and Vibration, Journal of Intelligent Manufacturing, among others. She also got several funding from national natural science foundation of China, China Postdoctoral Science Foundation, etc.

Xinghui Zhang received his B.S. in System Engineering, M.S. in Fault diagnosis, Ph.D. in PHM from the Mechanical Engineering College, Shijiazhuang. Now he is an associate research fellow in Intelligent Mobile Robot Research Institute, Beijing.