

Evaluating and Optimizing Analytic Signals

Shashvat Prakash¹, and Antoni Brzoska²

^{1,2}*Collins Aerospace, Windsor Locks, Connecticut, 06096, United States*

Shashvat.Prakash@collins.com

Antoni.Brzoska@collins.com

ABSTRACT

Condition-based maintenance is becoming a viable option for mitigating the high cost of unscheduled repairs. However, as data-driven approaches gain favor, there is a need to preserve the underlying physical degradation models in order to reasonably justify preventative maintenance. One solution is a class of models which augment physics with data-driven heuristics. The nature of the underlying degradation is explained with physics while detectability and decision nuances can be overcome with statistics and signal processing.

This paper describes a process for evaluating analytical models and using this evaluation for improving overall detection. The method involves optimizing a tunable filter to process signals such that the precursor signature preceding failure events approximates a known degradation behavior.

1. INTRODUCTION

Component failures in complex systems are often expensive. The loss of operation time is compounded by the costs of emergency repairs, excess labor, and compensation to aggrieved customers. Prognostic health management presents a viable mitigation when the failure onset is observable and the mitigation plan actionable. This means that the maintenance plan is clear and executable for a given, detectable failure mode.

Analytics for prognostic applications are, generally, degradation models. The model may estimate a remaining useful life (RUL) or, more generally, generate an alert for failure proximity. As more data have become available and as cloud computing capabilities grow, so does the potential for large scale model deployment. Regulatory bodies are considering prognostic monitoring regimes (Air Transport Assoc. of America, 2018) such as condition-based maintenance credits (Le, Ghoshal & Cuevas, 2011) as valid substitutes for certain routine maintenance inspections.

There is therefore a need to formalize methods which best evaluate the performance of prognostic models.

A representative prognostic for condition-based maintenance is shown in Figure 1. A suite of signals is continually providing information about the component health. It falls on the system to decipher the data and develop a decision architecture that drives a discrete maintenance.

Traditional maintenance schedules are based on a knowledge system using usage history and failure distributions (e.g. Weibull). More information on failure modes and usage profiles can drive a more efficient maintenance paradigm. Recent advances in sensors, electronics, and data have

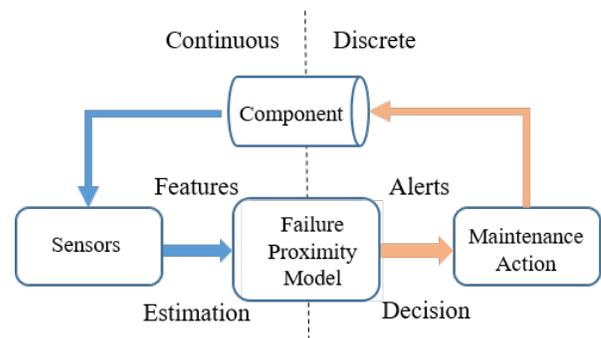


Figure 1. Prognostic Analytic Diagram

enabled this improvement, but they have also enabled more data-intensive approaches.

This paper will first review model types and evaluation constructs, then present an alternative evaluation method which can be used to optimize a discrete time filter. This method will amplify the known degradation elements of the signal while suppressing the attendant noise. An illustrative numerical example will conclude the paper.

2. MODEL TYPES

Model based and expert systems use, respectively, a model and a set of rules to infer when the degradation level requires attention (Zhang, Li & Yu, 2006). Models may be trained on

Prakash et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

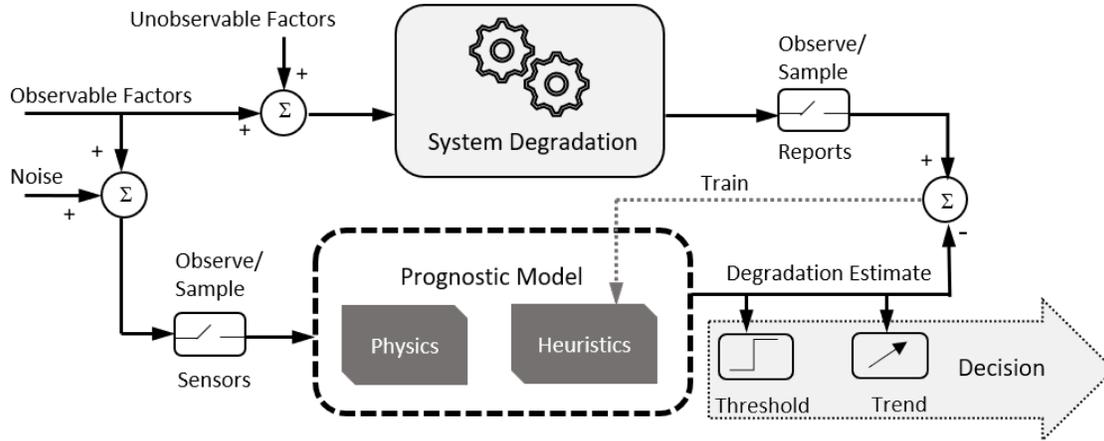


Figure 2. Model Structure with Physics and Data-Based Heuristic Components

a set of failure modes each with its own signature and associated probabilities in a Hidden Markov Model (HMM). Kwan, Zhang, Xu and Haynes (2003) and Zhang, Xu, Kwan, Liang, Xie, and Haynes (2005) developed and implemented an approach where principal components of the input signals were mapped to HMM degradation states. Capturing degradation modes in this manner may not always be scalable across multiple components of a complex system. Further, identifying and training against perceived discrete degradation states might not be needed if a simple signature can be processed out from the available data. The anomaly detection methods like HMM and multivariate Gaussian methods work best when there is some level of cleaned data. In other words we need to first separate, as best as possible, the signal from the noise in order to produce the best detection outcomes.

Further on the physical-data model spectrum are the purely data-driven approaches, where the outcome is driven by statistical inference (Luo, Namburu, Pattipati, Liu, Kawamoto & Chigusa, 2003, and Fornlof, Galar, Syberfeldt, & Almgren, 2016). While this allows for decisions outside the realm of human expertise, they are more difficult to productionize since the outcomes are less explainable.

Combining a physics model with a data-driven one was termed as a ‘physics – data hybrid’ by Sprong, Jiang, & Polinder (2019). Considering the expense involved in preventative maintenance, it makes sense to maintain the failure mode models grounded in physics but in a way that is enhanced by additional data.

One hybrid approach is shown schematically in Figure 2. The factors affecting system degradation are rarely known to a complete extent. Rather, only the observables inputs are captured at some time interval and even those signals are accompanied with noise. Similarly, reports on wear and failures are not real-time, but sampled versions are available. The available sensor signals and condition reports can inform

a physics model, with some data-based augmentation to compensate for the modeling and sensing deficiencies. The goal is to reduce estimation error while maintaining the intuition. The resulting approach captures the governing physical equations while relying on hyperparameters to compensate for incomplete or corrupted inputs.

3. BINARY CLASSIFIER CONSTRUCT

Common assessments of prognostic models have considered the overall goal of maximizing reliability (Zhang et.al., 2006) or the trade-off between part availability versus operational efficiency (Pipe, 2008). Such considerations are relevant when the model is already operational and logistic questions remain. This work will consider whether a given model can be tuned to optimize its key performance metrics, and what those metrics should be.

		Reality		
		No Event	Event	
Predicted	Alert +	False Positive FP	True Positive TP	Precision $\frac{TP}{TP+FP}$
	No Alert -	True Negative TN	False Negative FN	
			Recall $\frac{TP}{TP+FN}$	

Figure 3. Confusion Matrix of Binary Classifier

Standard assessments of prognostic model performance have often fallen under a broader category of binary classifiers (Fawcett, 2006). The capability to detect is quantified with ‘recall’ and the reliability of prediction is similarly quantified with ‘precision’. Recall and precision are based on a

confusion matrix, where one axis is detection and the other reality (Figure 3). If both reality and detection are in agreement, the model has scored a ‘true positive’ and if the model alerts without the event it is a ‘false positive’ while a ‘false negative’ means the model failed to adequately detect.

Recall is then the probability of alerting given the event will occur, while precision is the probability of the event occurring given an alert has been announced.

$$\text{Recall} = P(\text{Prognostic Alert} \mid \text{Impending Failure Event}) = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (1)$$

$$\text{Precision} = P(\text{Impeding Failure Event} \mid \text{Prognostic Alert}) = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

The ideal precision or recall values for a given problem can be evaluated by using derived values such as F1 score or receiver operating characteristic (ROC) (Fawcett, 2006). While the former attempts to balance recall and precision, the latter evaluates the tradeoff between the two quantities.

The above set of parameters present a common metric to evaluate binary classifiers, particularly in medical applications, (Mallett, Royston, Waters, Dutton & Altman, 2010) however, key elements are lacking for evaluating a prognostic health model. First, the model output and reality are both discretized to, respectively, positive/negative and true/false. Details of the underlying correlation are obscured.

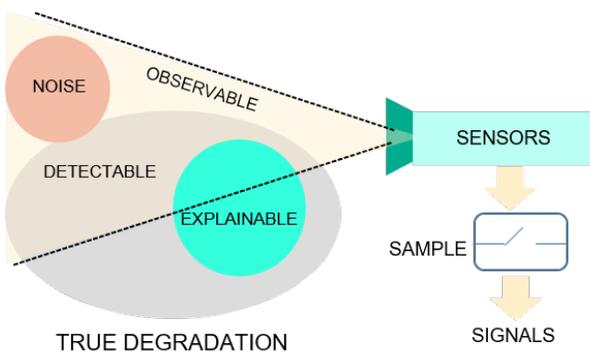


Figure 4. Composition of sensor signals

This becomes important if we wish to know margin to alert. Second, the time horizon is neglected. The time window to take action on an alert is limited; alerts without sufficient lead time cannot be acted upon and alerts too early will mean loss of useful life. Third, the full ROC analysis as presented by Fawcett (2006) has wide applicability in binary classifiers, but it becomes meaningless for prognostic signals where ‘true negative’ has no well-defined meaning. Once again, having no time basis means there is no widely applicable period where lack of both alerts and events qualifies as a successful prediction of a non-event.

One approach gaining favor is a simple precision-centric evaluation. The main idea is that induced downtime and removal of a functioning component must have a strong historical justification. Inherent in this logic is that the trust in the prognostic is tentative; only a record of minimal false positives can justify preventative measures. The as-yet unseen advantages of reducing unscheduled interruptions are not factored. Once again, without a time component, there is no clear demarcation between a timely true positive or a too-early positive that may be considered a nuisance false positive. Clearly, such a distinction exists but the rules governing it have not yet been considered.

4. SENSOR SIGNAL COMPONENTS

In a prognostic application, sensor signals intended to detect wear will also contain some amount of noise. In this case, noise is anything that is not the wear-out mode. It encompasses everything from random variations of the signal, to situations where the detection is intermittent or inconsistent. Hence, processing the raw sensor signal to maximize the wear-out precursors and minimize noise will provide an overall benefit to the detection before thresholds are applied.

The detectable portion of the degradation must be distilled from the noise, yet not all of it will be explainable, as shown in figure 4. Indeed, the explainable portion will be crucial in justifying the high-cost and high-risk actions. Explainability is enhanced with better physics models and physics-based intuition. It is not sufficient to have an explainable data model if the recommendations cannot be justified with known failure mechanisms. Finally, sample rate plays a critical role. Any dynamic elements with frequencies above half the sample rate are effectively inobservable, or, if not properly conditioned, aliased as noise (Ogata, 1995).

Developing the resulting signal into a degradation estimate requires then consideration of how well the signal correlates to historical failures, whether the captured events can be linked to explainable failure modes, and the value proposition of alerting on these failure modes in a manner that makes operational sense. It is important to first consider the resulting signals and features holistically, without thresholds applied. While the threshold represents a decision point, the

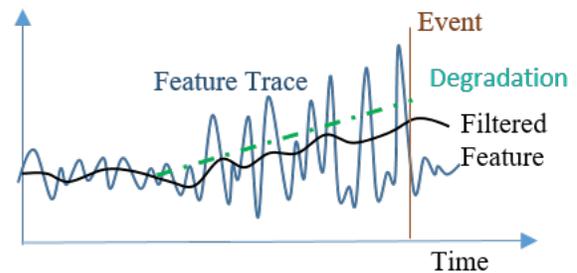


Figure 5. Noisy and Filtered Feature Trace with Event

underlying signal is intended to capture degradation, and its ability to do so will determine the overall performance of the analytic.

5. FEATURE SENSITIVITY

A prognostic model is only as effective as its ability to detect a set of given failure modes. This is a correlation problem. Evaluating a model only on the discrete outcomes (like true positives or false positives) misses the nuance of whether the underlying degradation is even well observed.

There is an entire discipline devoted to signal filtering, but very often the application is data in the $>.01$ Hz range, with focus on higher processing throughput with advancing capability (Estrada & Starr, 2005). Nonetheless, all intuitions apply here. The general approach to filter design can be applied to data collected on a per flight cycle basis.

As depicted in Figure 2, the first problem in estimation is that the available data might not be complete. There are observable as well as unobservable system inputs that contribute to the degradation state. Second, even when the degradation is perfectly observable, the derived features will detect not only the degradation, but other confounding noise. Usually, there are a number of noise sources, each of which operate in a unique frequency spectrum.

In the case of aircraft components, often there is a strong seasonality effect, driven by ambient temperature variation. Second, the idiosyncrasies of flight schedules, flight patterns, and daily weather also result in high variability. The goal is then to find the appropriate filter that can best track the real degradation (Figure 5).

5.1 Filter Design

A base feature can be derived many ways, usually reflecting some amount of physical modeling. While a feature is sometimes a directly measured attribute like petal length, in PHM applications a feature can itself be an estimated quantity like an effective age or crack length. However, given limitations on sensing and observation, the computed feature will propagate these inaccuracies. Therefore, a signal processing step is required to improve the overall estimation.

A dynamic filter modifies the feature in a manner that amplifies certain spectral content and suppresses others. Filters are commonly applied for noise rejection, modeling, estimation, and data fusion. The generic discretized filter that produces filtered output Y from input U has the form:

$$a_0 Y_k = \sum_{i=0}^N b_i^* U_{k-i} + \sum_{j=1}^M -a_j^* Y_{k-j} \quad (3)$$

The filtered output at time instance k is Y_k and the output at preceding time samples is $Y_{k-1}, Y_{k-2}, \dots, Y_{k-M}$. Similarly, the input at the current k th instance is U_k and the preceding values are $U_{k-1}, U_{k-2}, \dots, U_{k-N}$. The filter output at the current time instance is therefore a weighted summation of current and previous inputs, and in some cases, previous outputs. The latter are termed infinite impulse response (IIR) due to their recursive nature, while filters which do not use previous output values are termed finite impulse response (FIR). The filter coefficients a_i and b_j are chosen to achieve a spectral objective: suppression and amplification of a specified frequency range within a phase shift tolerance. In this case, we can choose coefficients which best detect known events while suppressing noise.

5.2 Lead Time Aggregation

Designing the appropriate filter requires an evaluation construct. Since the main objective is detection, the ideal filter will produce a signal that deviates most from its standard values during time intervals preceding known events and will return to its standard values once events have

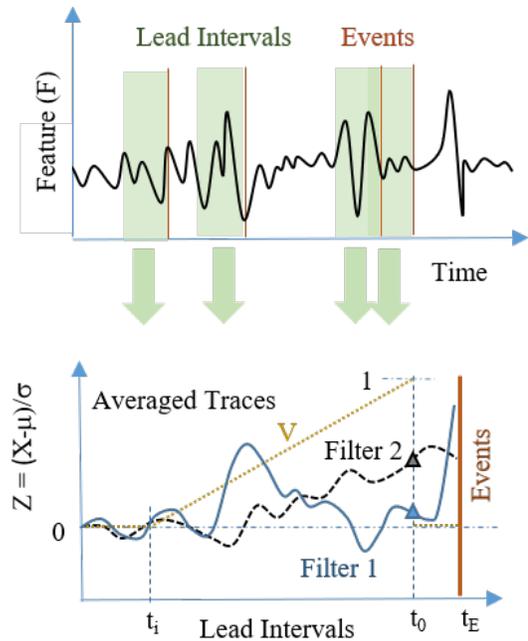


Figure 6. Aggregating the Lead Intervals into an Averaged Trace

transpired.

Capturing the filter behavior across all known events requires isolating the filter output for a set lead time interval before each event (Figure 6). Each lead interval data point is averaged on a time or cycle basis with all other lead time traces at the same relative distance from the event.

$$X_i = \sum_{j=1}^N F_{ij} / N \quad (4)$$

The lead interval value (X) at the i^{th} sample before the event is the average of all trace values (F) across N events, at the i^{th} value before each j event. The resulting signal represents typical behavior for the signal ahead of an event.

These aggregated averages X are then standardized using z -score normalization.

$$Z = (X - E[F]) / \sigma_F \quad (5)$$

The normalized value Z is the aggregated signal value X subtracted by the original signal mean $E[F]$ divided by the original signal standard deviation σ_F . The z -score normalization has the advantage of allowing comparisons across all signals with different base units. Further, normalization produces a signal in terms of its standard deviation value so that the larger values, either positive or negative, are more anomalous.

In some cases, there may be events which produce no detectable precursor, as may happen with a false negative. In that case, all filtered outputs X will be penalized equally, and the event will not play a role in filter selection. Conversely, there may be maxima in a signal that are not associated with an event. These false positives will raise the mean value $E(F)$ and result in a lower normalized Z value.

The aggregated and normalized pre-event trace Z of each filter can be considered at some fixed interval before the event for comparison (Figure 6). In the figure, the normalized and averaged traces of two candidate features are plotted, and the time axis has time of event (t_E), and minimum lead time to act before the event (t_0). Filter 1 in Figure 6 exhibits maxima both well ahead of t_0 as well as in the t_0 to t_E interval. The filter value between t_0 and t_E is irrelevant since there isn't enough lead time to mitigate the event. However, too much lead time reduces useful life. Filter 1's behavior is less desirable compared to Filter 2, which has a maximum just before t_0 , providing ample lead time ahead of the anticipated event without sacrificing much useful life.

5.3 Evaluation Methods

The Z value at the critical lead interval, $Z(t_0)$, can be a useful gauge of the relative performance of a given filter compared to others. This does not require any arbitrary rules or limits, only the process-defined, requisite minimum lead time.

In certain cases, the lead trace Z values can be better evaluated with a weighting function V (Figure 6) which rises

monotonically from 0 at some $t_i < t_0$ up to 1 at t_0 , then returning to 0 for the t_0 to t_E interval. For a given filter, each Z value can be combined with its weight V in a weighted sum:

$$S = \sum_i (Z_i V_i) / \sum_i (V_i) \quad (6)$$

The above equation ascribes a score S to the lead time averaged trace by weighting the i Z values with the weights V_i . High Z values near the event but before the critical actionable time (t_0) will have high V weights and increase the score while the other Z values will have less bearing on the score.

Table 1. Signal Evaluation Methods

<i>Method</i>	<i>Formula</i>	<i>Description</i>
Critical Z	$Z(t_0) = \frac{X(t_0) - E[F]}{\sigma_F}$	Aggregated normalized value at min lead time t_0
Score S	$S = \sum_i (Z_i V_i) / \sum_i (V_i)$	Correlation to value function V

The V function can be customized to suit the filter objectives. Any monotonically increasing function over t_i and t_0 will favor signals which show greatest anomaly immediately before the critical actionable time (t_0). A linearly increasing V value, either over cycles or flight hours, best captures the consumption of useful life on a cost basis. A sigmoid or step would acknowledge that any signal anomaly in the interval has comparable value, and loss of useful life is less important.

The score S (Equation 6) presents an objective function which not only quantifies the suitability of the filter, but it can also allow for the combination of multiple independent measurement signals. In the latter case, the score can be the weight of each independent signal in a weighted summation. The resulting aggregation will produce a signal with a higher score than any of its constituent parts.

$$F_{comb} = \sum_n S_n^b * F_n \quad (7)$$

Equation 7 describes a way to construct a new signal F_{comb} from n independent signals F_n by weighting these signals by their respective scores S_n raised to an arbitrary constant b , with $b \geq 1$.

6. FRAMEWORK FOR ESTIMATING DEGRADATION

The traditional precision, recall, and receiver operating characteristic apply to discrete classifiers, while in this case we have a continuous signal which can be correlated and optimally filtered. Therefore, the evaluation methods in

Table 1 are both a measure of signal correlation to a discrete event and the mechanism for obtaining the optimal signal.

For instance, if V is chosen to reflect the value of remaining useful life, the score reflects monetary benefit of the detector. Alternatively, V can represent any physical process that evolves over time, like crack growth or fatigue. Depending on the application elements of both operation and physics can inform the nature of the value function V .

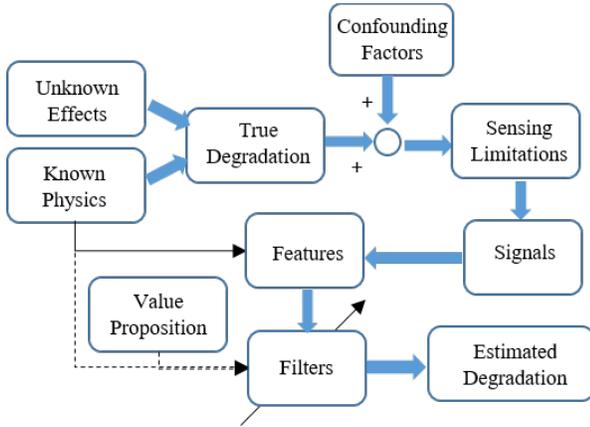


Figure 7. Filter Tuning based on physics and value

Conceptually, the filter tuning method is a way to model unmodeled elements in a catch-all filter, with the ultimate objective of producing a feature anomaly near failure events with sufficient lead time which in some way conform to known physical degradation signatures and/or reflect operational value.

This idea is shown in the model structure in Figure 7, which is a variation of the one in Figure 2. True degradation is driven by both known and unknown sources. Detection of the degradation is not perfect because confounding factors and sensing limitations respectively introduce noise and limit observability. The resulting signals are arranged into features using the known degradation mechanisms so that the features are physically explainable. The filter step at the end acknowledges the imperfections in the feature and attempts to compensate for them in order to estimate the degradation level.

7. NUMERICAL EXAMPLE

An example case has been constructed to demonstrate the method on synthesized data. First, we consider a component which has a lifespan distributed normally with mean 1250 cycles and standard deviation 250 cycles. Over the course of 10000 cycles, there are 7 failure events.

A degradation signature is modeled as a linearly increasing signal in the 500 cycles leading up to the event, and zero at all other points. This represents a form of physical process

where the degradation is evident only in the final stage of life and progresses at a constant rate until failure. In many practical scenarios, the degradation signal may not always be present before failure. To simulate this case, the degradation element has been removed for failure number 4. This is a false negative example.

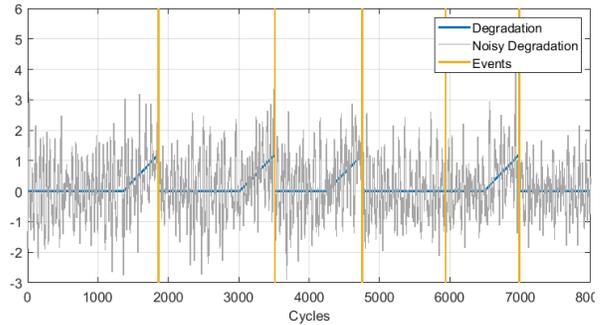


Figure 8. Synthetic data incorporating signal and noise – degradation is omitted for the fourth event

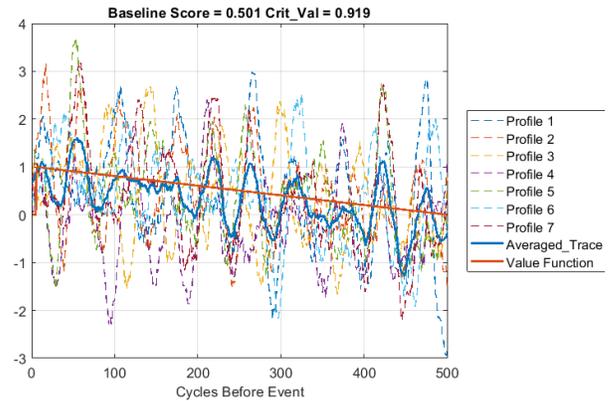


Figure 9. Baseline normalized traces

A noisy indicator will contain traces of the component degradation and noise from various sources. For this example, the noise is a set of 8 sinusoidal signals with random amplitudes up to .65 and frequencies spanning .01 to .03 Hz. Then, uniform random noise is added with zero mean and amplitude 2.5. The resulting noisy degradation signal is shown in Figure 8. This represents a raw sensor signal.

This raw signal now contains weak signatures prior to each event but one. Where the signatures exist, they are almost indistinguishable against the noise. A plot of lead time traces prior to each of the seven events is shown, mean-normalized, in Figure 9. The individual profiles are the signal value subtracted by the mean and then divided by the standard deviation, in the 500 cycles immediately before each event. Even though Profile 4 has no degradation component, it is

similar to all the other profiles which do contain the linear degradation.

Next, an optimal filter was developed in the form of equation 3. The coefficients a_i and b_i were chosen to maximize the objective function, which was simply S from Table 1. Running multivariate optimization yielded $a_0=0.8060$, $a_1=-0.7890$, $b_1=1.000$, $b_2=-0.9295$, $b_3=0.0011$.

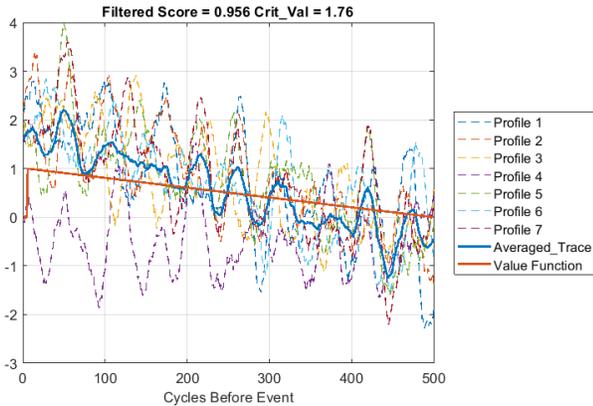


Figure 10. Filtered signal traces

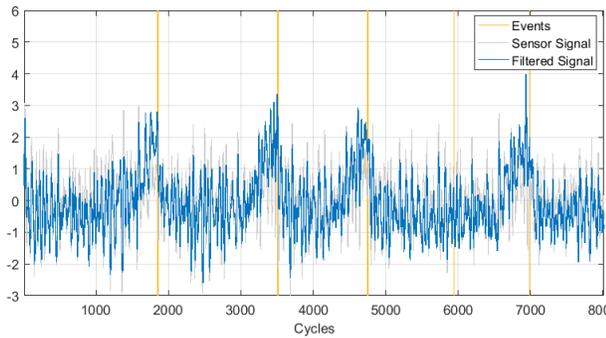


Figure 11. Resulting filtered signal

The normalized filtered signal traces leading up to each event are shown in Figure 10, with the resulting time series shown in Figure 11. Now there is a clear separation between the profiles containing the degradation and profile 4, the one without degradation. This indicates that the filter is amplifying the modeled degradation and suppressing the attendant noise.

Note that even with the filter, some residual noise remains. This left over noise is in some part related to the omission of event 4’s degradation, in effect pushing the filter to retain some noise in attempt to find the signal.

Table 2 summarizes the performance of the filter across each event and for each of the two evaluation metrics from Table

1. Although the filter was optimized for the score metric, it also improves on the critical Z metric for the cases where degradation was present. Event 4 did not have a degradation signal so the metrics decrease after the filter suppresses noise.

Table 2. Evaluation Results from Example

Event	Life	Degraded	Score S		Critical Z	
			Base	Filter	Base	Filter
1	1858	Yes	0.634	1.18	1.17	2.38
2	1656	Yes	0.692	1.30	1.60	2.79
3	1242	Yes	0.596	1.18	0.422	1.52
4	1187	No	-0.207	-0.490	0.988	0.402
5	1045	Yes	0.581	1.19	1.08	1.70
6	1634	Yes	0.577	1.08	0.332	1.68
7	848	Yes	0.633	1.25	0.846	1.86
AVG	1353		0.501	0.956	0.919	1.76

8. CONCLUSIONS

The framework presented here is an evaluation method for a prognostic analytic. While the analytic may be based on a fundamental understanding of physical degradation, unknown effects, confounding factors, and signal limitations will present estimation challenges. An appropriate filter can lead to a better estimation of true degradation.

The discrete classifier evaluation methods (Fawcett, 2009) require arbitrary boundaries between true positives and false positives, and a threshold value. The receiver operator characteristic (ROC) curve attempts to disambiguate the threshold, but the classification of positives and negatives remains unclear when alerts occur over time, eventually leading to a discrete event. All alerts preceding a particular failure, going back to infinite time, could be interpreted as a singular true positive. Conversely, all those alerts could be false positives until the very recent alert even if the degradation was evident for some time.

The presented framework resolves this ambiguity by assigning an objective value to a signal at a specific lead time or time intervals. Signal processing and heuristics can then be tuned to maximize this value. When the value is tied to actual costs, like remaining useful life, then the benefit becomes more justifiable.

Replacing the binary classifier means replacing the discretized true/false reality and positive/negative prediction with meaningful metrics that capture the quality of estimation at the key actionable intervals. Indeed, the single normalized trace will capture when the signal is anomalous, so missed or false detection instances (i.e. false negatives and false positives) will penalize the evaluation appropriately.

The approaches presented here address the shortcoming of using discrete classifier methods on prognostic health models. A discrete decision is the result of a set of continuous features which track an observable degradation, and the performance is therefore a function of correlation, with continuous signal anomaly preceding a discrete failure event. The nature of communication around prognostics should adopt this continuous-discrete mindset and evaluation criteria.

REFERENCES

- Air Transport Assoc. of America, International Maintenance Review Board Policy Board. (2018). Aircraft Health Monitoring (AHM) Integration in MSG 3, Issue Paper 180. MSG-3: Operator/Manufacturer scheduled Maintenance Development, Vol. 1- Fixed Wing Aircraft Rev. 2018.1, <https://www.easa.europa.eu/document-library/imrbpb-issue-papers>.
- Le D.D., Ghoshal A., and Cuevas E. (2011). Condition-based maintenance plus and maintenance credit validation. Army Research Laboratory, Dept. of Defense, Aberdeen Proving Ground, MD 21005. <https://apps.dtic.mil/sti/pdfs/ADA589574.pdf>
- Zhang L., Li X., Yu J. (2006). A review of fault prognostics in condition based maintenance. Sixth International Symposium on Instrumentation and Control Technology: Signal Analysis, Measurement Theory, Photo-Electronic Technology, and Artificial Intelligence, (6 November 2006); doi: 10.1117/12.717514
- Kwan C., Zhang X., Xu R., and Haynes L. (2003). A novel approach to fault diagnostics and prognostics, Proc. of the 2003 IEEE International Conf. on Robotics & Automation, pp. 604-609, (14-19 Sept.) doi: 10.1109/ROBOT.2003.1241660
- Zhang X., Xu R., Kwan C., Liang S. Y., Xie Q., and Haynes L., (2005). An integrated approach to bearing fault diagnostics and prognostics, Proceedings of the 2005 American Control Conference, pp. 2750-2755 vol. 4, (8-10 Jun.) doi: 10.1109/ACC.2005.1470385.
- Luo J. H., Namburu M., Pattipati K., Liu Q., Kawamoto M., Chigusa S., (2003). Model-based prognostic techniques [maintenance applications], Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference., pp. 330-340, (22-25 Sept.) doi: 10.1109/AUTEST.2003.1243596.
- Fornlof V., Galar D., Syberfeldt A. and Almgren T. (2016). Maintenance, prognostics and diagnostic approaches for aircraft engines, 2016 *IEEE Metrology for Aerospace (MetroAeroSpace)*, pp. 403-407, doi: 10.1109/MetroAeroSpace.2016.7573249.
- Sprong, J. P., Jiang, X., & Polinder, H. (2019). A Deployment of Prognostics to Optimize Aircraft Maintenance: A Literature Review. Annual Conference of the PHM Society, 11(1). doi: 10.36001/phmconf.2019.v11i1.776.
- Pipe K. (2008). Practical prognostics for condition based maintenance, *International Conference on Prognostics and Health Management*.
- Fawcett T. (2006). An Introduction to ROC Analysis, *Pattern Recognition Letters*. 27 (8): 861–874.
- Mallett, S., Royston, P., Waters, R. Dutton S., and Altman D. (2010). Reporting performance of prognostic models in cancer: A review. *BMC Med* 8, 21. doi: 10.1186/1741-7015-8-21
- Ogata, K. (1995) *Discrete-Time Control Systems*. Ann Arbor: Prentice Hall.
- Estrada R. F. and Starr E. A. (2005). 50 Years of Acoustical Signal Processing for Detection: Coping with the Digital Revolution. *IEEE Annals of the History of Computing* 65-78.
- Shashvat Prakash** received a B.S. in Mechanical Engineering from University of Illinois, Urbana (1997), an M.S. in Mechanical Engineering from Carnegie Mellon University (1999), and a Ph.D. in Mechanical Engineering from the Georgia Institute of Technology (2007). He has worked on satellite attitude and orbit control at NASA Goddard, combustion and control of aviation turbines at General Electric, and prognostics and health management (PHM) of aircraft components at Raytheon Technologies-Collins Aerospace. Currently a Principal Engineer, he has over 13 years of experience in the commercial aviation industry
- Antoni Brzoska** received a Ph.D. in Mathematics from the University of Connecticut in 2017. He has been with Collins Aerospace for one year.
- Prior to Collins, he was a data scientist at a neuroscience start-up