Programmatic and Technical PHM Development Challenges in Forward Fit Applications

Carl S. Byington, P.E.,
Michael J. Roemer, Ph.D.,
Impact Technologies, LLC
200 Canal View Blvd.
Rochester, NY 14623
carl.byington@impact-tek.com
mike.roemer@impact-tek.com

Andrew J. Hess
Program Office
1213 Jefferson Davis Highway, Suite 600
Arlington, VA 22202-4304
andrew_hess@comcast.net

ABSTRACT

This paper describes the programmatic and technical challenges associated with developing Prognostics and Health Management (PHM) capabilities in an advanced new weapons platform. At the highest level, it presents specific strategies to tie PHM benefits and objectives to logistics support concepts in a measurable way. The authors also provide examples of the use of this approach to ensure that PHM elements have bought their way onboard the aircraft. At the next level, the authors evaluate the Verification and Validation (V&V) approach used within the PHM algorithm and sensor suite. Forward fit applications, where field and final system data is usually less available, present additional hurdles to the verification of diagnostic coverage, detection rates, and false alarm rates. Specific tools will be demonstrated by the authors to provide V&V use cases using a combination of test stand development data, similar component failure data, and, ultimately, fielded data. Metrics for fault detection, diagnosis, and prognosis functional elements will be presented. In addition, the authors discuss the use of simulation and real fault data, as well as a strategy developed to project the fielded performance of the system. The effects of signal noise, measurement uncertainty, and threshold settings are addressed. The concept of 'performance metric growth' with field data availability is also discussed and the specific application of these techniques and tools to the challenges in new aircraft deployment is offered.

INTRODUCTION AND MOTIVATION

With the U.S. Department of Defense (DoD) Condition-based Maintenance Plus (CBM+) initiatives, Acquisition Category (ACAT) program managers are required to "optimize operational readiness through affordable, integrated, embedded diagnostics and prognostics …automatic identification technology; and iterative technology refreshment." [1] It is also DoD policy that Condition-based Maintenance (CBM) be "implemented to improve maintenance agility and responsiveness, increase operational availability, and reduce life cycle total ownership costs." [2] *The goal of CBM is to perform maintenance only upon evidence of need. The primary tenets of CBM include: "designing systems that require minimum maintenance; need-driven maintenance; appropriate use of embedded diagnostics and prognostics through the application of reliability centered maintenance principles; improved maintenance analytical and production technologies; automated maintenance information generation; trend based reliability and process improvements; integrated information systems providing logistics system response based on equipment maintenance condition; and smaller maintenance and logistics footprints*."[1]

Condition-based Maintenance Plus (CBM+) expands on existing CBM principles and emphasizes their focus on "maintenance processes and capabilities derived, in large part, from real-time assessment of weapon system condition, obtained from embedded sensors and/or external tests and measurements."[1] This capability can only be achieved through appropriate diagnostic and prognostic technologies. As quoted directly from DoD Instruction 5000.2:

> ***Diagnostics:*** *Applicable and effective on-board monitoring/recording devices and software, e.g. built-in test (BIT), that provide enhanced capability for fault detection and isolation, thus optimizing the time to repair. Emphasis must also be on accuracy and minimization of false alarms.*
>
> ***Prognostics:*** *Applicable and effective on-board monitoring/recording devices and software, e.g. BIT, that monitor various components and indicate out of range conditions, imminent failure probability, and similar proactive maintenance optimization actions. [1]*

The Defense Acquisition Guidelines reiterate the need for the early teaming with appropriate systems engineering to clearly define and understand the operational envelope and range of conditions to properly design the Built-In-Test (BIT), Built-In-Self-Test (BIST), and diagnostic mechanisms, including false alarm mitigation. This DoD policy and the experience that underlies it clearly indicates the need for false alarm mitigation techniques, design principles, and analytical tools to assist developers and stakeholders.

The significance of this opportunity is also apparent considering the recent increase in Performance-based Logistics (PBL) business models [3], which are fast becoming the DoD's preferred approach for product support implementation. The essence of PBL is buying performance outcomes, not individual parts or repair actions. The motivation is to share the risk (and reward) over the life cycle of the asset. The contracts are balanced to meet the warfighter requirements and the support strategy is managed by the support of the Product Support Integrator. This is a major shift from the traditional approach to product support, which emphasizes buying set levels of spares, repairs, tools, and data. The new focus is on buying a predetermined level of availability to meet the warfighter's objectives.

This new paradigm is accomplished through Performance-based Agreements. Extensive business case analysis is conducted to assess the benefits for the DoD and contractor performing the integrator role. Earned Value Management (EVM) and associated program analysis tools are typically used in the implementation and negotiation of a PBL contract. Central to the business case analysis is the assessment of PHM system performance. If the health management system possesses insufficient coverage of the critical components/replaceable units, ineffective incipient fault detection capability, or high false alarm rates, then the risk for the contractor may be too great to make a cost effective business case. If such a situation occurs, neither the government nor the contractor benefits. Thus, verifying functionality, having sufficient PHM coverage, and mitigating false alarms are significant concerns in realizing these logistics and support paradigms.

Autonomic Logistics (AL) is the automation of this logistics environment such that little human intervention is needed to engage the logistics cycle. The idea for the AL system was drawn from the workings of the autonomic nervous system of the human body, whose functions occur autonomically: they are spontaneous and based on some internal stimuli. An AL system will be able to operate without the conscious intervention of a human. Examples of actions that will be automated within the JSF supportability concept include maintenance scheduling, flight scheduling, and ordering spare parts.

Prognostics and Health Management (PHM) is the key enabler for the Autonomic Logistics support concept. In fact, the cornerstone of Autonomic Logistics is an advanced diagnostic and PHM-based system. The PHM system provides the data, information, and knowledge to initiate the Autonomic Logistics chain of events.

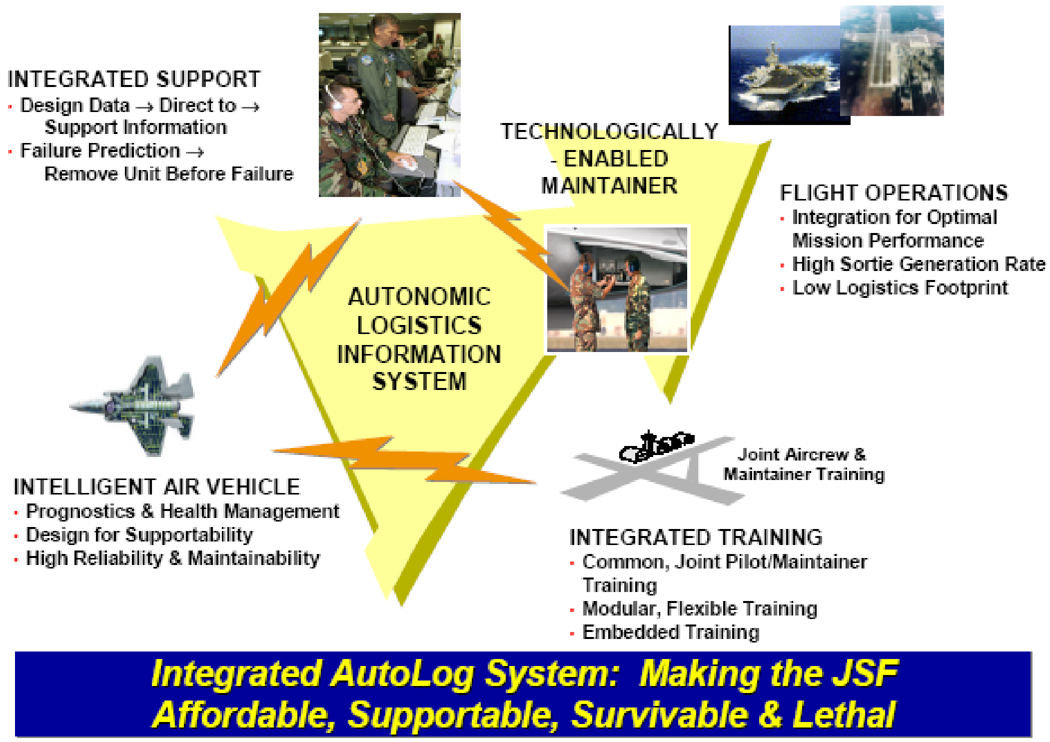

Figure 1: AL System Technical Solutions [11]

The AL system is based on five (5) key concepts:

1. Smart and Reliable Aircraft – An aircraft with reliability, maintainability, and an inherent, design-in PHM system enables the entire AL system concept.

2. Technology Enabled and Supported Maintainer – A maintainer must be equipped with a leading edge technical support system that provides the training, information, tools and equipment to do his/her job.

3. Integrated Training Environment – A training environment that employs the latest technology and research on learning is needed to provide a comprehensive, integrated capability to mission–qualify pilots and maintainers, regardless of their locations.

4. Intelligent Information System – An information infrastructure that interfaces with the air vehicle, legacy support systems, supporting commercial enterprise systems, and the warfighter provides an effective portal to JSF information and create an intelligent system for maintaining and operating the JSF among multiple armed services and international partners.

5. Responsive Logistics Infrastructure – The system should be sufficiently responsive to support requirements within a timeframe that allows the JSF weapon system to generate the required number of effective sorties at the least cost.

To enable these maintenance and logistics concepts for our systems of the future, we need the development and implementation of advanced Prognostics and Health Management (PHM) technologies. The remainder of this paper will focus on the challenges involved in developing and implementing such incipient fault detection and true prognostic capabilities, as well as outlining the tools, techniques, and approaches undertaken to achieve these capabilities.

ANATOMY OF A FAILURE AND WHEN WE NEED TO KNOW

Prognostics and Health Management (PHM) is the name given to the capability being developed by the JSF to enable the vision of Autonomic Logistics and meet its overall affordability and supportability goals. In PHM, the term 'prognostics' includes the broader functions of fault/failure detection, fault/failure isolation, enhanced diagnostics, material condition, performance monitoring, and life tracking, rather than just prognostic functions alone. Envisioning an initial fault-to-failure progression timeline (shown in Figure 2) is one way of exploring the relationships between prognostic elements. This timeline starts with a new component in proper working order, indicates a time where an early incipient fault develops, then depicts how, under continuing usage, the component reaches a failure state and, eventually, a state of secondary system damage and complete catastrophic failure.

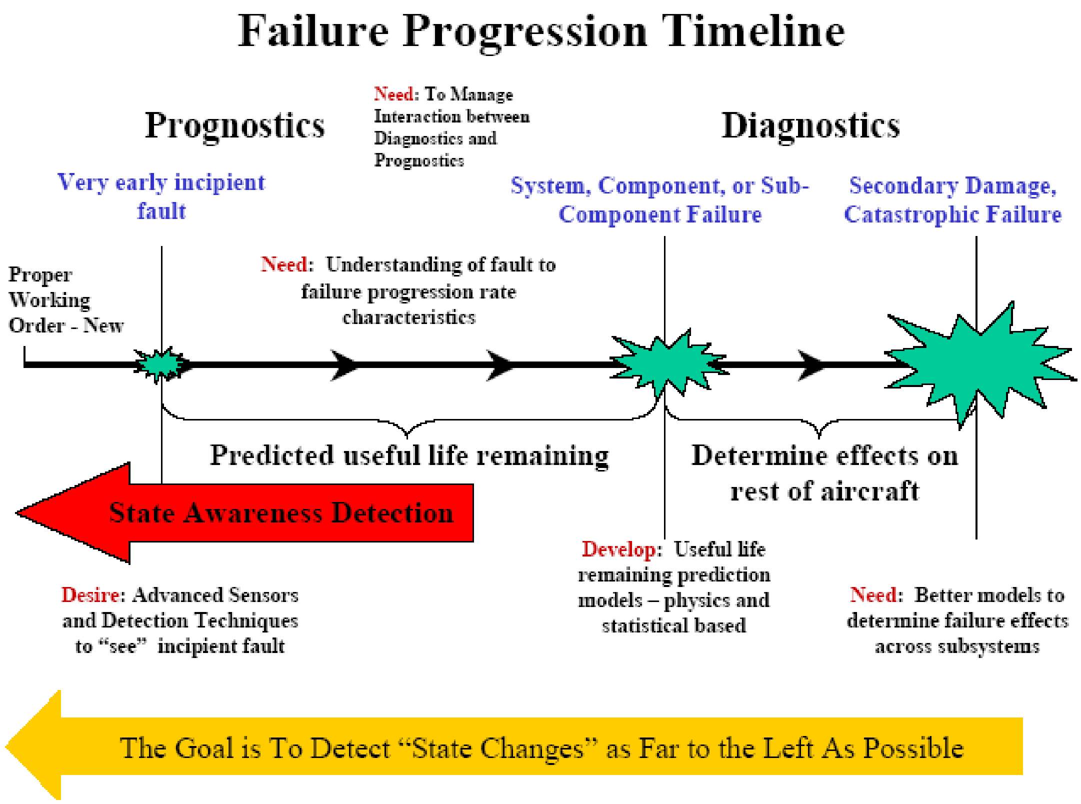

Figure 2: Failure Progression Timeline

Diagnostic capabilities have traditionally been applied at or between the initial detection of a system, component, or subcomponent failure and complete system catastrophic failure. More recent diagnostics technologies are enabling detections to be made much earlier at incipient fault stages. In order to maximize the benefits of continued operational life of a system or subsystem component, maintenance will often be delayed until the early incipient fault progresses to a more severe state (but before an actual failure event). This area between the very early detection of an incipient fault and its progression to system or component failure is the realm of true prognostics. If an operator has the will to continue to operate a system and/or component with a known incipient fault present, he will want to ensure that this can be done safely and will want to know how much useful life remains at any point along this particular failure progression timeline. This is the specific domain of real predictive prognostics, or "the Big P" – being able to accurately predict useful life remaining along a specific failure progression timeline for a particular system or component. To actually accomplish these accurate useful life remaining prediction capabilities requires many tools in your prognostic tool kit. Sometimes available sensors currently used for diagnostics provide adequate prognostic state awareness inputs; sometimes advanced sensors or additional incipient fault detection techniques are required. Other necessary prognostic tools include: a model or set of models that represent the understanding of a particular fault-to-failure progression rate; material physics of failure models; statistical and/or probabilistic based models; models to represent failure effects across interconnected subsystems; and models to account for and address future operational mission usage.

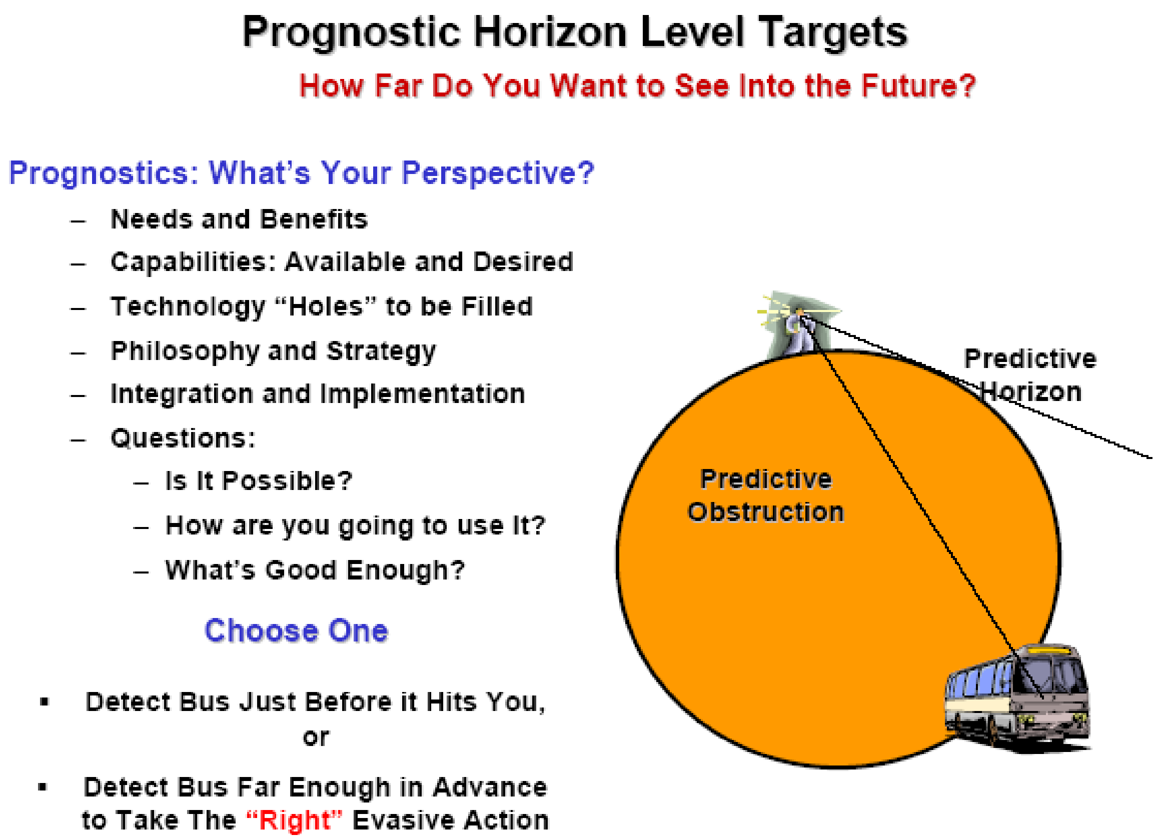

Figure 3: Prognostic Perspectives and Questions

As true prognostic capabilities evolve and are developed for particular applications, many difficult questions need to be addressed. Perhaps the first basic question to ask is: how far does your specific application want or need to "see" into the future? The answer to this question will very much depend on your application specific prognostic perspectives. From this perspective, this first question generates many additional considerations. Some of these include: specific needs vs. recognized benefits; what is possible and feasible; capabilities available vs. those desired or highly valued; technology shortfalls to be filled; integration, implementation, and usage strategies; and what is good enough. Figure 3 represents some of these questions from a prognostic perspective by using a bus traveling around a globe as an analogy. This depiction can be used to address the concept of the predictive prognostic horizon – in this case, how early you want to "see" or detect the bus in order to do something useful with the information.

Predictive prognostics is one of the fundamental factors that influence the decision to shift to a Condition-based Maintenance business approach. Original Equipment Manufactures (OEMs) and key Suppliers are incentivized to make capital investments in new, innovative prognostic technologies to improve logistic support capabilities, knowing that their involvement (and related profits) will extend well into the production and post-production phases of modern military aircraft. Additional long-term, performance-based contractual arrangements would be established between the prime contractors, their sub-suppliers, and government organic depots.

A performance-based approach is intended to:

- Reduce Total Ownership Costs.
- Increase Warfighter Confidence and Satisfaction.
- Facilitate Contractor-Government Integration and Communication.
- Reduce the Demand for Logistics.
- Incentivize Reliability Enhancements.
- Decrease the Resource Requirements for Support.
- Encourage Early and Continuing Emphasis on Diminishing Manufacturing Sources and Obsolescence Planning.
- Centralize Management.
- Create a Real Time Problem Response (24/7).
- Optimize the Technology Insertion.
- Utilize a consistent Life-Cycle Cost Analysis.
- Optimize Infrastructure Harmonization and Asset Utilization.

The vision is that legacy, renegotiable, price-based contracting will be abandoned over time as a gradual shift toward the new long-term Performance-based concept occurs. For example, cost-type contracts could be used during the Low Rate Initial Production (LRIP) phase. As the design matures, defined enablers are achieved and sufficient reliability and cost data are obtained. At this point, the contract would transition to a fixed-price type contract, with performance incentives linked to metrics identified during LRIP. Through joint cooperative forums, client (i.e., Government) and contractor would establish the criteria and performance metrics for the transition in accordance with substantiated data, maturity of process, pricing, and performance responsibility. Total transition to a fixed-price type contract would be tied to obtaining sufficient reliability and cost data. However, the actual transition point would depend on the system and air vehicle peculiarities and maturity. The final decisions for transition sequencing would be determined during LRIP using factors such as achieved and stable reliability, business case analysis impact, incentives, and the maturity of depot level repair.

In order to implement the Performance-based approach, the strategy will revolve around how core capabilities for dealing with OEMs/suppliers are established, refined, and executed in relationship to delivering performance to the user community. This is a challenge in itself – an additional challenge figuring out how PHM capabilities and their associated data product will assimilate into it.

Recognizing the need for the aforementioned contractual approach to ensure that system design is optimally balanced between total ownership cost and system/equipment performance requirements, it is mandatory to describe the development of a best value business arrangement. A best value business arrangement should contain affordability initiatives targeted at continuous cost reduction and technology refresh/insertion. The establishment of a business model is the framework for assessing the correlation of business elements and attributes, which are the functions and tasks required to perform the logistics operations and that will allow one to assess the best value solution. However, while it is clear that prognostics influences various attributes (i.e., support equipment, maintenance, supply chain, spares ownership/management, propulsion support, etc.), it is a challenge to develop valid and measurable metrics to quantify the impact of the various PHM technologies on individual model attributes and to include these in the business model so that the resultant analyses reflect the true value of prognostics to the proposed business case.

Key Technical Capabilities and Metrics

One of the major challenges to the designers of modern PHM systems is the need for developed diagnostic and prognostic methods that are truly capable of handling real-world uncertainties. Such real world uncertainties cause havoc with deterministic approaches leading to high false alarm rates, inaccurate predictions, incorrect decisions, and an overall PHM system that is not very robust. Some of the issues uncertainty presents to the designer are elaborated below, including issues associated with various steps in the predictive process, the estimate of current condition, the prediction of time-to-failure (or time remaining), and the choice of appropriate lead times (how far ahead to predict), as well as choices for an overall prognostic methodology.

Incipient Fault Detection: In order to predict remaining life or time-to-failure, you must first know where the current condition stands along the continuum of various possible health states, which ranges from normal to failing conditions. Assuming that the computed feature (or condition/health index) is sensitive to a change brought about by a damage or degradation mechanism, the system designer's goal is to maximize that sensitivity, thus separating the mean values while minimizing the variance to decrease the spread and resultant overlap of the distributions (see Figure 4). For successful diagnostic systems, threshold setting is most effectively accomplished by balancing the P(False Alarm) and P(Detection) using available data sets, simulations, and models, as appropriate. The P(Correct Rejection) is the complement of the P(False Alarm). Similarly, the P(Missed Detection) and P(Detection) are complementary. Note that P(False Alarm) and P(Detection) are related but are not complementary, as is sometimes incorrectly assumed.

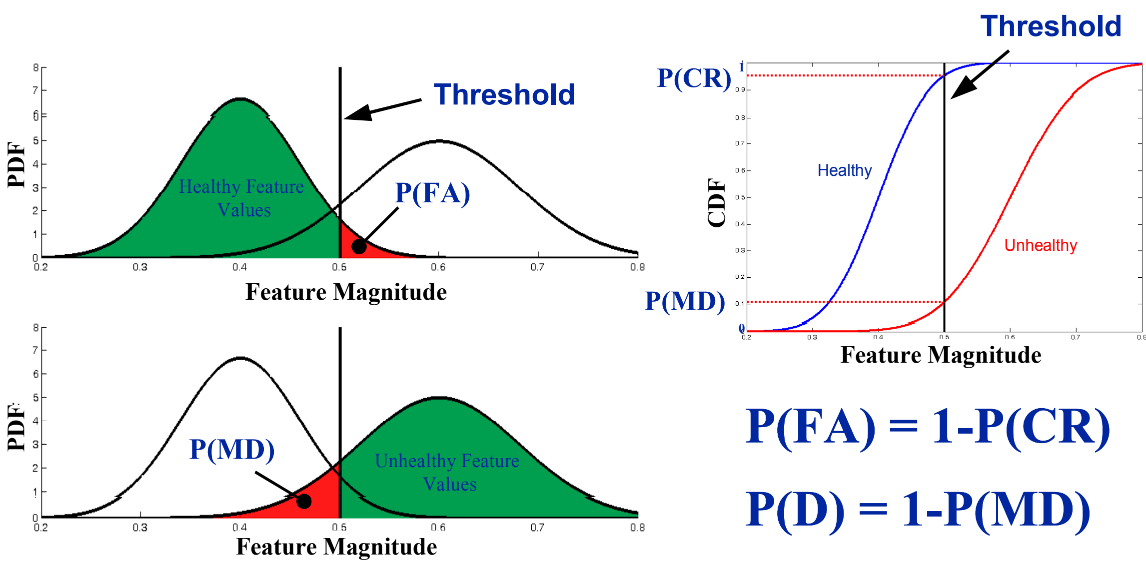

Figure 4: Relationship of Statistical Detection Metrics

The concept of improved performance with increased separability is also addressed with these curves. As seen in Figure 4, if we can increase the separability (that is, move the means further apart), then we will improve the P(FA) and P(D) performance. In addition, one can envision that, even if the mean values remain the same, increased separability will cause the distributions to become narrower and steeper with less probability in the tails, thus decreasing the P(FA) and increasing the P(D). Maximizing the separability and reducing the computed feature variances as much as possible is key to achieving good PHM system performance. *Fundamentally, good incipient fault detection and false alarm mitigation is handled using proper feature selection with (high separability/discriminability) and the use of intelligent decision fusion to effectively trade false alarm rates with missed detection rates.*

Unfortunately, for newly designed systems, there is often insufficient data to predict the faulty distribution. In this case, we need to extrapolate the data from incipient faults in similar systems to estimate where this distribution might reside. In many ways, this is analogous to estimating a failure rate on a new piece of equipment using legacy component reliability data. We will need to combine data and/or models from other similar degradation experiences with our engineering judgment to estimate the unhealthy or faulted distribution. We accept that this method has is uncertain and attempt to quantify the uncertainty as best as possible.

Figure 5 displays the probability distributions for the ImpactEnergy™ bearing fault detection algorithm, which provided a much higher confidence in fault detection than traditional methods. This figure clearly displays a lower variance in the spall progression data than typically achievable with traditional methods. This affects both the false alarm and missed detection rates. Also, note the decrease in probability percentages of both rates due to a greater separation between the respective means of the distributions. Table 1 provides a summary and comparison of the False Alarm and Missed Detection rates for three bearing fault detection methods tested under the seeded fault and spall conditions. From this example, it is apparent that feature selection and discriminability play major roles in false alarm mitigation and good incipient fault detection.

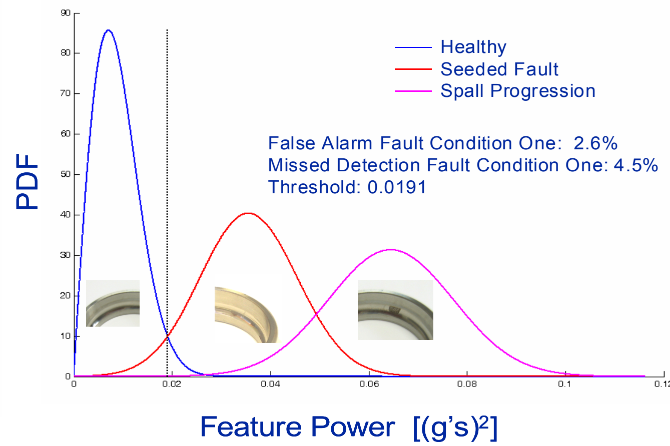

Figure 5: Feature Distributions with ImpactEnergy™ During Bearing Fault Progression

Many well-developed metrics for analysis of detection exist; however, a consolidated tool that incorporates raw data, algorithm interfaces, and metrics based analysis would prove valuable for development and long term support of specific applications. We have initiated work on such a tool and its current user interface is shown here. One of the primary objectives of the tool is to better automate and configuration control the analysis of this type to produce not only robust incipient detection algorithms, but also traceable performance characteristics for these algorithms.

Table 1 – Results of for Three Bearing Fault Detection Methods

Technique		Seeded Fault	Spall
ImpactEnergy™	F/A Rate	2.6%	0.1%
	Miss Rate	4.5%	0.2%
Conventional Frequency Processing	F/A Rate	15.4%	1.7%
	Miss Rate	5.5%	9.6%
Legacy Onboard	F/A Rate	28.2%	4.8%
	Miss Rate	29.3%	14.5%

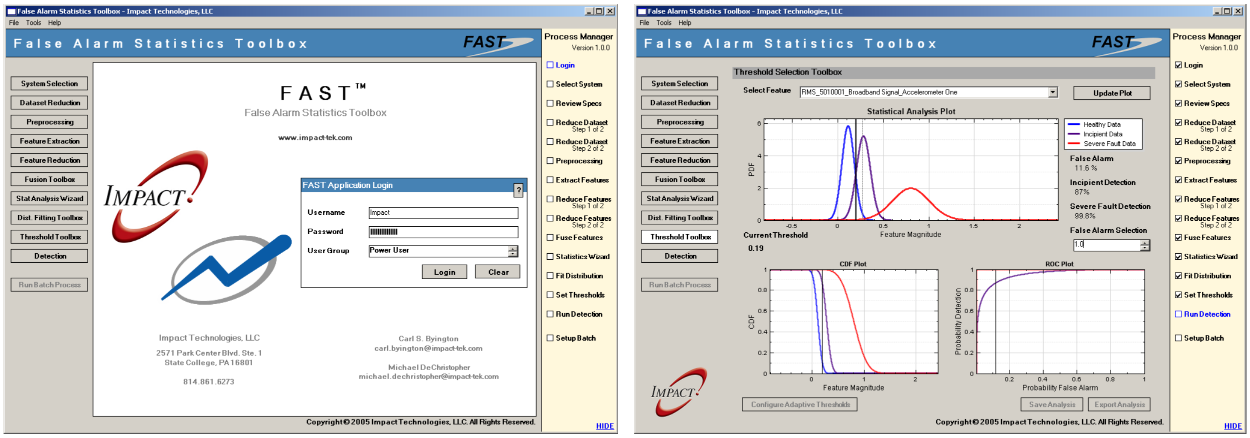

Figure 6: Two FAST PHM™ Tool Interfaces

Diagnosis: Perhaps better referred to as isolation, fault diagnosis is the next step in the formation of a PHM result. While the previous section dealt with the detection of the existence of a fault, diagnosis refers to the identification of the failed, or, more importantly, failing component. Table 2 presents three useful metrics for analysis of diagnostic effectiveness: the confusion matrix, the probability of isolation, and the Kappa Coefficient. Probability of isolation refers to the percentage of all component failures that the monitor is able to unambiguously isolate. The Kappa Coefficient represents how well an algorithm is able to correctly classify a fault (with a correction for chance agreement). The Kappa Coefficient is bounded by –1 and 1, with 0 representing chance agreement and negative numbers indicating that the algorithm performed worse than random guessing. Together, these metrics form a core analysis of diagnostic effectiveness. However, as these and related diagnostic metrics are treated in many texts on the matter, we now turn our attention to the less well known matter of prognostics.

Table 2 – Typical Diagnostic Metrics

Diagnostic	Confusion Matrix	*Table of Classification Results*
	Probability of Isolation (FIR)	$FIR = \frac{A_T}{A_T + C_T} * 100$
	Kappa Coefficient	$kappa = \frac{N(observed) - N(expected)}{N(total) - N(expected)}$

Prognostics: A prognostic vector typically indicates time-to-failure and an associated confidence in that prediction. However, the definition and implementation of prognostic performance metrics must consider multiple viewpoints and objectives. Two such viewpoints follow:

1. The maintainer's viewpoint: When should I perform maintenance? What is the most appropriate time (given such constraints as availability of spares, etc.) to maintain critical equipment?
2. The field commander's viewpoint: Are my assets available to perform the mission? What is the confidence level that critical assets will be available for the duration of the mission? Or, given a certain confidence level, what is the expected time-to-failure of a particular asset?

Answering the **maintainer's question** requires *an estimate of the remaining useful lifetime of a failing component and an assignment of uncertainty bounds to the trending curve that will provide the maintainer with the earliest and the latest times to perform maintenance and the associated risk factors when maintenance action is delayed.*

Answering the **commander's question** requires one of two things: *1) an estimate of the confidence level assigned to the completion of a mission considering mission duration, expected operating and environmental conditions, other contingencies, etc., or 2) a means to estimate the component's time-to-failure given a desired confidence level.* Thus, the prognostic curve (the output of prognostic routines) must be modulated by a time-dependent distribution (statistical/Weibull or possibilistic/fuzzy) that will provide uncertainty bounds. An appropriate integral of such a distribution at each point in time will result in the confidence level committed to completing the mission or, given a desired confidence level, the time-to-failure of a failing component.

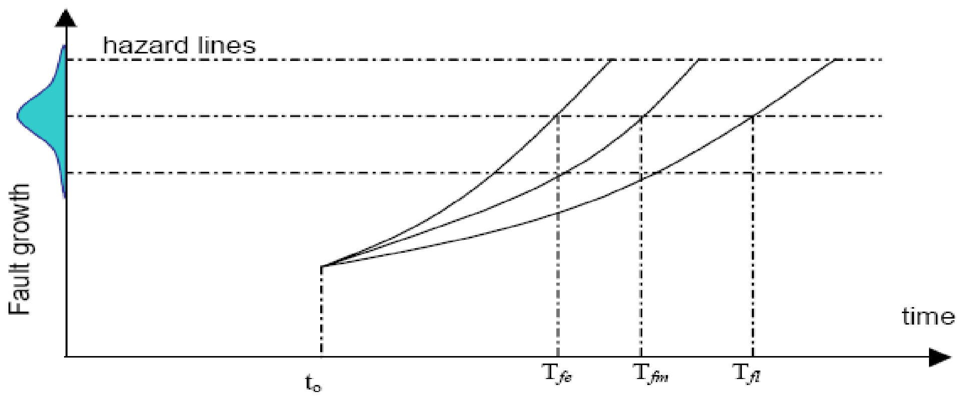

Figure 7: Confidence Limits and Uncertainty Bounds

Assume that a prognostic algorithm predicts the following progression or evolution of a fault with the associated uncertainty bounds: at current time (t_o), a fault has been detected and isolated and the prognostic routine is predicting the mean time-to-failure (T_{fm}), the earliest time to failure (T_{fe}), and the latest time to failure of (T_{ft}), as shown in Figure 7. The hazard line specifies the fault magnitude (dimension) at which the component ceases to be operational (failure).

Let us further assume that failure data is available to superimpose a distribution or distributions, as shown pictorially in Figure 8. Strictly speaking, these "distributions" are either possibilistic (fuzzy) functions or probability density functions.

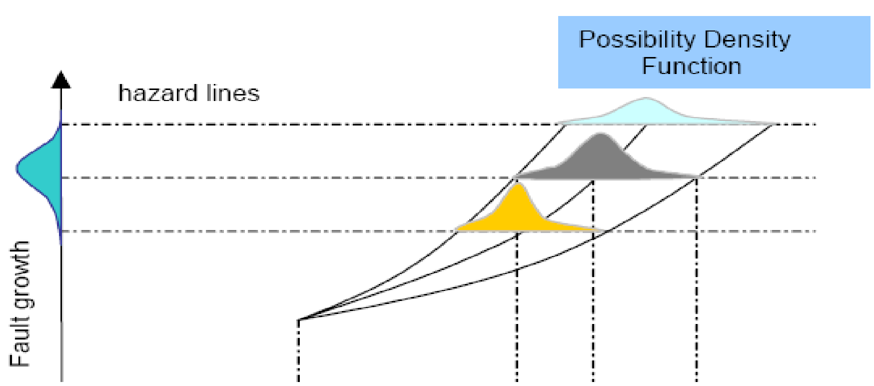

Figure 8: Possibility Density Functions for Confidence Bounds

Suppose further that the distribution at T_{fm} (crossing the hazard line) is as shown in Figure 9:

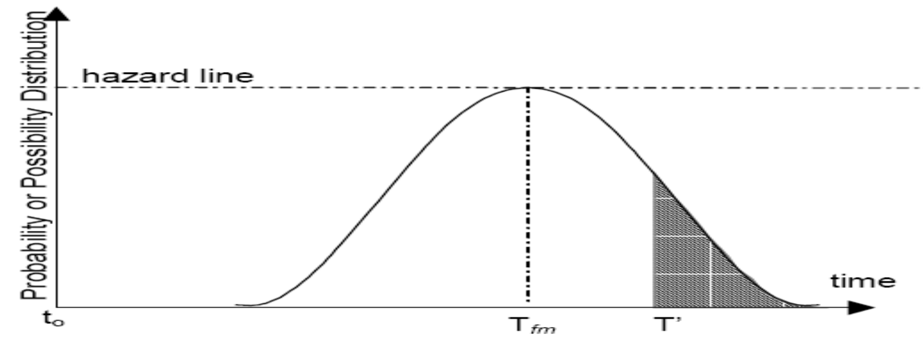

Figure 9: Distribution at the Hazard Line

As an example, assume that a pre-planned mission is estimated to require the availability of an asset under consideration for time (T'). The integral under the distribution curve from T' to infinity will give us an estimate of the confidence level in terms of probability (or possibility). In other words, it will tell us how confident we can be that the asset will not fail before the mission is completed. Now consider the same distribution for a second example, where we specify a certain confidence level, say 95%, and would like to find the time (T'') that the component will remain operational prior to complete failure. Integrating the distribution curve from T'' to infinity and setting the result equal to the confidence limit (95% in our example), we solve for T'', thus arriving at the length of time (starting at t_o) that the asset will be available within the desired confidence limit, as shown in Figure 10. We view this procedure as a dynamic evolution of our estimates. That is, as more data becomes available and time marches on, new confidence limits are derived and the uncertainty bounds shrink through appropriate learning routines. The procedure outlined above may eventually lead to specification of performance metrics for prognostic systems. Such additional metrics may refer to the reliability of confidence limits, risk assessments, safety specifications, etc.

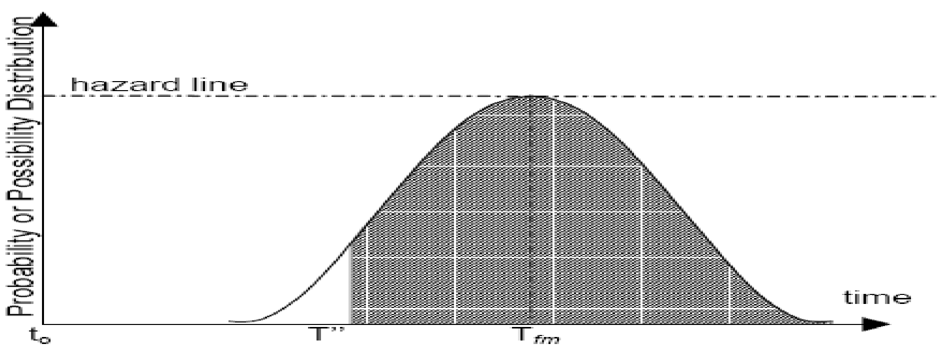

Figure 10: Distribution at Hazard Line

Similarity compares more than one predicted time series against the real series from a starting point to a certain point in the future. This measurement can be computed as:

$$similarity(x,y) = \sum_{i=1}^{n}\left(1 - \frac{|x_i - y_i|}{\max_i - \min_i}\right) \quad \textbf{(1)}$$

where x_i and y_i are two i_{th} elements in two different time series, and $\max_i$ and $\min_i$ are the maximum and minimum of all i_{th} elements.

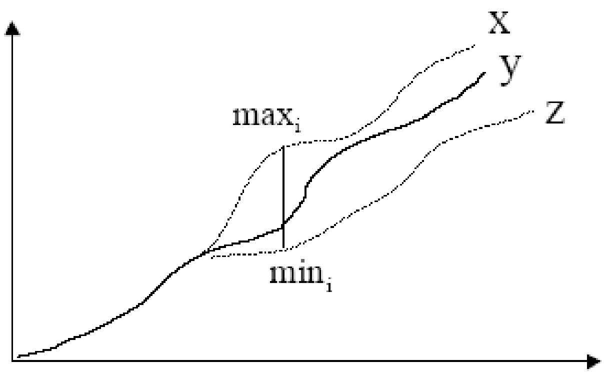

Figure 11: Determining Similarity

Sensitivity measures how sensitive the prognostic algorithm is to input changes or external disturbances. It is defined as:

$$SN = \frac{\sum_{i=1}^{N} \frac{\Delta_i^{Out}}{\Delta_i^{In}}}{N} \qquad \textbf{(2)}$$

where Δ^{out} equals the distance measure of two successive outputs and Δ^{in} equals the distance measure of two successive inputs.

The mean of the prediction-to-failure time is calculated as:

$$E\{t_{pf}\} = \frac{1}{N}\sum_{i=1}^{N} t_{pf}(i) \qquad \textbf{(3)}$$

where $t_{pf}(i)$ denotes the prediction-to-failure time for the i_{th} experiment and N is the number of experiments. Standard deviation of the prediction-to-failure time is calculated as:

$$S\{t_{pf}\} = [\frac{1}{N}\sum_{i=1}^{N}(t_{pf}(i) - E\{t_{pf}\})^2]^{1/2} \qquad \textbf{(4)}$$

Many other prognostic metrics exist and the interested reader is directed to the listed references.

Fundamentally, the same problem that exists with diagnostic techniques also exists with prognostics: validating and verifying their functionality and uncertainty drivers. For this reason, a Prognostics Test Bench is being designed to import "black box" physics or feature-based Remaining Useful Life (RUL) prognostic models and automatically interrogate them and report critical information, such as parametric sensitivities and certain prognostic metrics. A cornerstone of the program was to investigate and develop Statistical Influence Models (SIMs) for insertion into a Prognostics Testbench in the MATLAB® environment. SIMs are defined as independent models that ultimately generate parameters (either stochastic or deterministic) that influence a RUL distribution. These parameters include usage profiles, manufacturing defects, random damage events, build tolerances, material condition, inspection capabilities, and integration of state awareness and predictive prognostics. The overall concept envisioned for the Prognostic Testbench is shown in Figure 12.

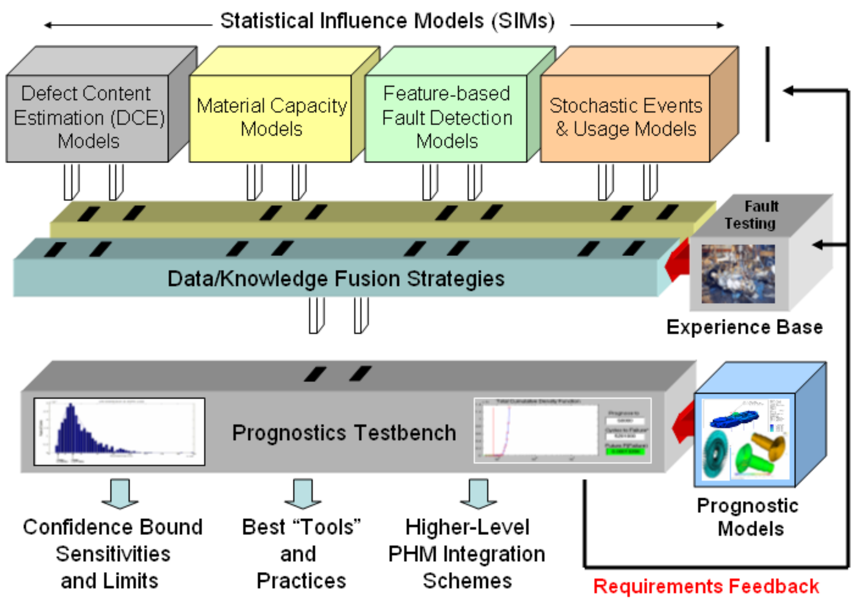

Figure 12: Overall Concept for Model and Test Bench Development

The Importance of Ground Truth Choice in Metric Assessment: Ground truth is a representation of the actual condition of the system and serves as a baseline for comparison between a feature/algorithm result and the actual condition. Validation at all three levels – detection, isolation, and prognosis – relies on ground truth data. Although the name itself implies an "absolutely true" representation, in most cases, ground truth is an engineering estimation of the actual condition that is derived from a measurable parameter that closely matches actual condition. Observable ground truth may be based on visible evidence, another feature or prediction result, or the damage severity curve. The ideal situation is for a feature/algorithm to possess perfect agreement with the ground truth in some measurable way. Metrics are used to assess and report the degree to which this agreement occurs. However, it is important to note that features and algorithms that score well against one ground truth estimate could score lower against a different ground truth estimate. Many factors can produce this effect. System overtraining is a large factor that can be controlled through careful advance planning. Acceptable control methods include robust system testing, use of multiple classifiers through knowledge fusion, and data reserve/hold-out techniques. The data hold-out technique, which uses a 70/30 development versus test data role, is also an effective preservation of ground truth integrity, given a substantive amount of data. Ground truth data for EFV will derive from test stand fault characterization efforts and from operational data.

Tools for Verification and Validation (V&V): As mentioned previously, robust tools for Verification and Validation (V&V) can provide useful insights that are applicable to system development as well as ongoing validation and improvement efforts. The ability to warehouse raw data, simulation models, and system configuration information, in combination with an interface that permits easy exercise of developed detection, isolation, and prognostic algorithms, would add significant value to many acquisition programs that seek to employ PHM technology. Such a tool could incorporate real data, simulation models, and noise augmentation techniques to analyze the effectiveness and overall robustness of existing and emerging technology. Impact Technologies, LLC, in cooperation with the Georgia Institute of Technology, is developing a web-based software application that will provide JSF (F-35) system suppliers with a comprehensive set of PHM Verification and Validation (V&V) resources. The application includes standards and definitions, V&V metrics for detection/diagnosis/prognosis, access to costly seeded fault data sets and example implementations, a collaborative user forum for the exchange of information, and an automated tool to impartially evaluate the performance and effectiveness of PHM technologies. This program is specifically focused on the development of a prototype software product that will illustrate the feasibility of the techniques, methodologies, and approaches needed to verify and validate PHM capabilities. The approach being pursued to assess overall PHM system accuracy, as illustrated in Figure 13, is to quantify the associated uncertainties at each individual subsystem level and combine them through the PHM architecture to generate a system level uncertainty estimate. A team of JSF system suppliers, including Pratt & Whitney, Northrop Grumman, and Honeywell, are being assembled to provide contributions, feedback, and recommendations regarding the product under development.

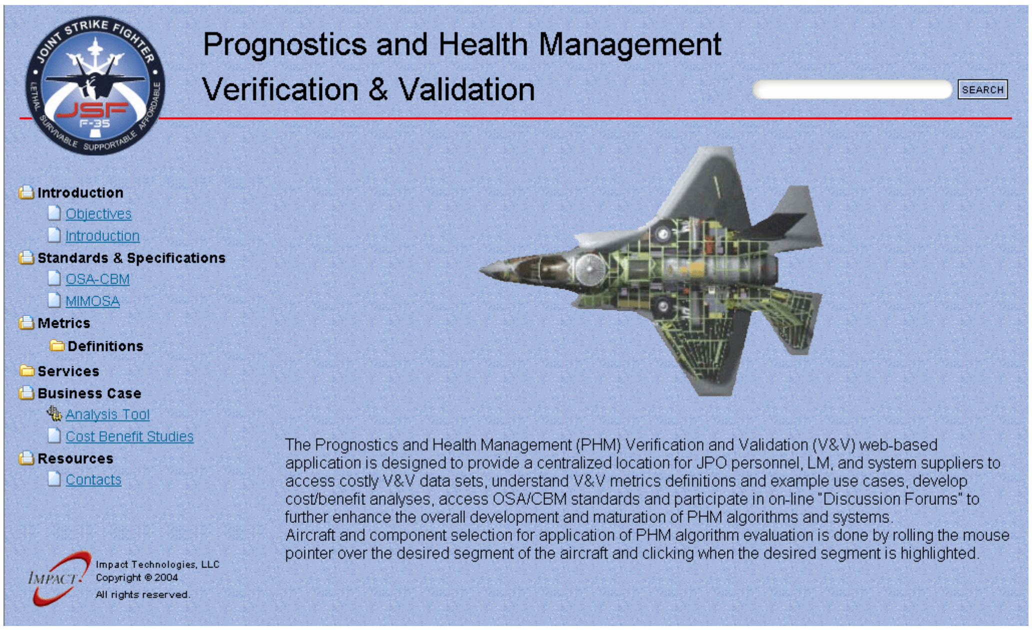

Figure 13: JSF PHM V&V by Impact and Georgia Tech

SOME LESSONS LEARNED… SO FAR

As various programs to develop comprehensive diagnostic, prognostic and health management (PHM) systems progress, many lessons are learned. Similarly, as specific prognostic-directed projects are undertaken, additional lessons are accrued. The following paragraphs discuss some of these accumulated lessons learned to date.

Prognostic capabilities can be hard to develop and often take time to mature, but they are feasible in many cases. Having said this, prognostics are certainly not feasible for all cases. There is a need to identify the clearly unachievable cases and then address them in other ways than through significant prognostic efforts. Rather, focus limited resources on systems, subsystems, or components that are both attainable and high value, such as the propulsion system, drivetrain, or, occasionally, electronics equipment (depending on your mission and system design).

Look to share developments (costs) and find people with prognostic experiences in other platforms. It is unlikely that any single platform or program can afford all of the resources required to support the development and validation of accurate useful life remaining prognostic capabilities on every desired system, subsystem, and component. This is true even for only the feasible and high value components of a complex platform like a modern fighter aircraft. To mitigate this high expenditure of resources on a single platform, a smart strategy needs to be implemented. This strategy should attempt to share prognostic capability development costs across similarly interested program platforms and across groups of similar component types. The intelligent application of specifically targeted science and technology efforts outside of their particular platform program may also offer additional resources.

Really good diagnostics and incipient fault detection techniques that detect faults before functional failure are important precursors to health-based prognostics. If we can detect a problem before it becomes a functional failure to be diagnosed, then we are achieving substantially more predictive capability than we have in legacy systems. This indication will also allow us to trigger the appropriate prognostic model. The corollary statement is also very true and real: if you have developed very capable and comprehensive diagnostic capabilities, then it necessarily follows that you will attempt to develop prognostics. There are a few exceptions to these statements. For one, system or subsystem overall performance degradation trending can be a very useful prognostic capability by itself.

Multiple types of models may be needed to achieve prognostic capabilities with accurate useful life remaining predictions. These types of models can include sensor calibration or correlation models, models of accumulated usage up to crack initiation, incipient fault or crack propagation models, statistical and/or probabilistic based models. All of these individual supporting models need to be developed for accuracy, validated, and integrated into a global prognostic model for the specific component and system application.

PHM is a multidisciplinary engineering problem. The successful development of global prognostics models and accurate useful life remaining predictive capabilities requires a development team with expertise in various disciplines. This development teams needs a mixture of material science experts; state awareness sensor experts; several types of modeling and data fusion experts; specific component of interest design specialists; statistical and probabilistic modeling specialists; and traditional diagnostic experts. Don't make the mistake of thinking, "Since it's largely software, I'll let the software design folks handle it." That will result in highly-efficient but largely worthless PHM software. The core of PHM, like any other engineered system, is applied engineering knowledge from multiple disciplines.

PHM a life cycle endeavor and maturity takes time and fielded experience. It takes significant data, experience, and maturation time to develop accurate prognostic and useful life remaining prediction capabilities. This is true in some cases more than others; since there are always "low hanging fruit" examples where degradation always occurs along easily understood trend lines. This statement is particularly true for accurate useful life remaining prediction capabilities and for specific components where physics of failure models are not well understood or where root cause failure mechanisms are very random. It is important to remember that sophisticated prognostic capabilities require lots of data to develop and ample time to mature, so be sure to include sufficient maturation time when planning your program.

Performance degradation capture and trending is "low hanging fruit" in many systems. Often straightforward system or subsystem performance degradation trending, without any accurate useful life remaining predictions, can be a very useful form of prognostics. This represents a prognostic capability in its simplest form, but it can be extremely useful to the fleet operator. This type of prognostic trending is best applied in combination with case-based reasoning or when physics of failure models for specific components are not available.

Although good Failure Modes and Effects Analyses (FMECA) won't give us prognostics by themselves, they are a necessary starting point. Thorough and up-to-date FMECAs identify failure modes and their related symptoms while prioritizing high occurrence failure modes and critical failure path components. These capabilities help to identify high value components for prognostic coverage. FMECAs can also be very useful in uncovering the interrelationships between the dependent elements of components, whether they are located within a single system or across interconnected subsystems.

Subsystem component seeded fault testing is extremely useful but can be prohibitively expensive. Plan for a few seeded fault tests and use them wisely. Don't jump into testing too soon just to try to make some progress. Instead, begin with a resourced testing strategy to maximize the capture of performance degradation AND incipient fault-to-failure data whenever possible in a piggybacked environment. This would include planning to capture this invaluable data during all possible system, subsystem, and component developmental testing, qualification testing, environmental tests, final acceptance tests, etc.

Be ready to justify why you are developing PHM … again! When the platform program funding cuts come (and they always will), prognostic capabilities will often be the first to be scrutinized for cost reduction. Prepare for this and develop a well designed strategy to counter these arguments and articulate the specific benefit justifications.

SUMMARY AND CONCLUSIONS

The fleet needs and benefit impacts of real predictive prognostics are evident, real, and substantial. There are several types of prognostic capabilities with varying levels of sophistication, from simple trending techniques to multiple integrated modeling approaches. The prognostic definitions used for these different approaches may also vary and are continuously evolving with the technology in this multidisciplinary field.

Real predictive prognostic capabilities are just one element among many interrelated and complementary functions in the field of PHM. Beginning with a solid foundation for monitoring and detection, a good PHM system deploys a wide array of techniques to achieve its ultimate result. Analysis of PHM effectiveness at each level provides not only the verification and validation required to measure overall performance, but also provides the legitimate feedback required for ongoing PHM system development and evolution.

This paper has explored the background, benefit, impact, and current playing field for predictive prognostics; highlighted some specific design challenges and issues; discussed the various degrees of prognostic capabilities and metrics; and drew heavily on lessons learned from previous and current prognostic development efforts. Techniques were discussed to mitigate the effects of system specialization and algorithm overtraining. The prognostic measures presented encompass measures of effectiveness from two different viewpoints (maintainer/commander). Finally, the need for ongoing PHM metrics analysis and tool development was established. The Joint Strike Fighter, along with many other current acquisition programs, will benefit from a more in depth look at verification and validation tool development from performance and requirements development perspectives.

The challenges involved in developing and implementing prognostic capabilities are many, and this paper has focused very much on those impacting and enabling benefits associated with new performance-based logistic support concepts, global sustainment strategies, and new business practices. The presentation of this paper will focus on these concepts and demonstrate some of the tools employed.

PRIMARY AUTHOR BIOGRAPHIES

Carl S. Byington is a Professional Engineer and the Director of Systems Engineering at Impact Technologies, LLC. He possesses over 18 years experience in the design, analysis, and testing of mechanical, thermal, and fluid power systems. He currently performs as the Principal Investigator for the development of innovative Prognostics and Health Management (PHM) technologies for military and commercial customers and he has successfully led more than 15 Phase I and II awards for the Army, Navy, Air Force, and DARPA. Several Phase III contracts and successful transitions to the warfighter have resulted from this work. He has also led health monitoring developments on multiple programs with Boeing, General Electric, Northrop Grumman, and NASA Glenn.

Prior to joining Impact, Carl worked at NASA Langley Research Center, where he performed air breathing propulsion research, and, more recently, at the Penn State Applied Research Laboratory (PSU ARL) as the Department Head in Condition-based Maintenance (CBM). He has served as the PI on a University Research Initiative for Integrated Predictive Diagnostics and subsequently led several programs related to Joint Strike Fighter subsystem prognostics. Carl is also a past Technical Program Chairman of the Machinery Failure Prevention Technology (MFPT) Society and a current member, instructor, and keynote speaker for the Society of Tribologists and Lubrication Engineers (STLE). Carl is also the current Chairman of the Machinery Diagnostics & Prognostics Committee within the ASME Tribology Division. He has BS and MS degrees in Mechanical and Aeronautical Engineering, respectively. Carl is a lead and co-author on 2 patents related to PHM technologies and he has published over 70 papers, book chapters, and magazine and journal articles related to signal processing, data fusion, statistical analysis, model-based prognostics, and predictive health management. Carl can be reached at carl.byington@impact-tek.com.

Andrew Hess is a graduate of both the University of Virginia and the U. S. Navy Test Pilot School. Andy is world renowned for his work in fixed and rotary wing health monitoring and is recognized as the father of naval aviation propulsion diagnostics. Beginning with the A-7E engine health monitoring program in the early 1970's, Andy has been a leading advocate for health monitoring in the Navy and has been instrumental in the development of every Navy aircraft application, from the A-7E through the F/A-18 and V-22. His efforts have largely led to the successful transition of a development program known as Helicopter Integrated Diagnostic System (IMD HUMS) into production for all H-60, H-53, and AH-1 aircraft. More recently, Andy has been a strong advocate of prognostic capabilities and has been involved in many efforts to advance the development of these predictive capabilities for a sundry of systems and their component elements. Currently, Andy is hard at work leading the development and integration of the Prognostic and Health Management (PHM) system for the Joint Strike Fighter program and is a frequent participant in the international technical community. He can be reached at andrew_hess@comcast.net.

Dr. Michael J. Roemer is the co-founder and Director of Engineering at Impact Technologies. He has over 19 years experience in the development of real-time monitoring, diagnostic, and prognostic systems for a wide range of military and commercial applications. He has extensive working knowledge in the areas of artificial intelligence, vibration analysis, aero-thermal performance monitoring, non-destructive structural evaluation and monitoring, and probabilistic risk assessment. Dr. Roemer is Chairman of the Machinery Failure Prevention Technology (MFPT) Society, Prognostic Lead for the SAE E-32 Engine Condition Monitoring Committee, a member of the ASME Controls and Diagnostics Committee, and an Adjunct Professor of Mechanical Engineering at the Rochester Institute of Technology. He is the co-author of a recent book entitled, "Intelligent Fault Diagnosis and Prognosis for Engineering Systems," and has written or co-authored more than 75 technical papers related to integrated systems health management. He can be reached at mike.roemer@impact-tek.com.

ACKNOWLEDGMENTS

The authors wish to thank their colleagues who have contributed ideas and opinions to this work in current and past collaborations. *Major Giulio Calvello*, an Italian Air Force Corps of Engineers Officer, graduated in 1994 from both the Italian Air Force Academy and the University of Naples "Federico II" in Aeronautical Engineering. He has served as Autonomic Logistics leader for Prognostic and Health Management (PHM) in the Joint Strike Fighter Program Office. *Dr Peter Frith* is a principal research scientist from the Australian Defence Science and Technology Organization (DSTO). He recently completed an assignment with the JSF Program Office as a Collaborative Program Personnel in the Air System PHM Team. *Dr. George Vachtsevanos* is a professor at Georgia Tech University in Atlanta, GA who has collaborated with the authors on numerous difficult prognostic problems. *Mr. Michael Begin* of the Naval Air Warfare Center in Patuxent River, MD, works on JSF propulsion system PHM, and has shepherded many PHM innovations. *Mr. Doug Gass* has recently retired from the Naval Air Warfare Center, but not before contributing his many ideas and experiences on false alarm mitigation and the fielding of effective PHM. *Dr. Steve Engel* is a researcher at Northrop Grumman with great interest in the art of prognosis and its application in new and legacy systems. Lastly, the authors would like to recognize the many project engineers and managers at Impact Technologies, LLC, who have contributed their hard work and dedication toward the development of these technologies and tools.

References

1. DoD Instruction 5000.2, http://akss.dau.mil/dag/DoD5002/DoD5002-3.9.2.asp#39241
2. DUSD(LMR) CBM+Memorandum, November 2002, Available at Defense Acquisition Guidebook, http://akss.dau.mil/dag/TOC_GuideBook.asp?sNode=R5-2-1-2&Exp=Y
3. DUSD(LMR) Memorandum, March 2003, Implementing the Future Logistics Enterprise End-to-End Customer Support, http://akss.dau.mil/dag/Guidebook/IG_c5.3.2.asp
4. Kay, Steven, "Fundamentals of Statistical Signal Processing," Volume II, Prentice Hall, ISBN 0-13-504135-X, 1998.
5. Kalgren, P., Byington, C., Amin, S., Kallappa, P., "Bearing Incipient Fault Detection: Technical Approach, Experience, and Issues," MFPT 2005.
6. Byington, C., Safa-Bakhsh, R., Watson, M. Kalgren, P., "Metrics Evaluation and Tool Development for Health and Usage Monitoring System Technology," AHS Forum 59, Phoenix, Arizona, May 6-8, 2003.
7. Orsagh R.F., Roemer M.J., et al, "Development of Metrics for Mechanical Diagnostic Technique Qualification and Validation," COMADEM Conference, Houston TX, December 2000.
8. Roemer, M. Dzakowic, J., Orsagh, R., Byington, C., Vachtsevanos, G., "Validation and Verification of Prognostic and Health Management Technologies," IEEE Aerospace, 2005.
9. Barlas, I., Zhang, G., Propes, N., and Vachtsevanos, G., "Confidence Metrics and Uncertainty Management in Prognosis," MARCON, Knoxville, May, 2003.
10. Hess A. and Fila L., "The Joint Strike Fighter (JSF) PHM Concept: Potential Impact on Aging Aircraft Problems," Paper # 403, 2001 IEEE Conference, March 2001.
11. Hess A. and Fila L., "Prognostics, from the Need to Reality – from the Fleet Users and PHM System Designer / Developers Perspectives," Paper #116, 2002 IEEE Conference, March 2002.
12. Calvello G., Dabney T., Hess A., "PHM a Key Enabler for the JSF Autonomic Logistics Support Concept," Paper #1601, 2004 IEEE Conference, March 2004.
13. Engel, S., Gilmartin, B., et al, "Prognostics, the Real Issues With Predicting Life Remaining," Proceedings of the 2000 IEEE Aerospace Conference; 0-7803-5846-5.
14. D. Muench, G. Kacprzynski, A. Liberson, A. Sarlashkar and M. Roemer, "Model and Sensor Fusion for Prognosis, Example: Kalman Filtering as Applied to Corrosion-Fatigue and FE Models," *SIPS Quarterly review presentation*, 2004.
15. Hall, D. L., Mathematical Techniques in Multisensor Data Fusion, Artech House, 1992.
16. M. J. Roemer, G. J. Kacprzynski, and R. F. Orsagh, "Assessment of Data and Knowledge Fusion Strategies for Prognostics and Health Management," IEEE Aerospace Conference, Big Sky, MT, March 2001.
17. Byington, C. and Garga, Data Fusion for Developing Predictive Diagnostics for Electromechanical Systems, Chapter 23 of CRC Press Handbook of Multi-sensor Data Fusion, edited by Hall and Linas, CRC Press, Spring 2001.