Tackling Verification and Validation for Prognostics

Martin S. Feather[1]
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109

Kai Goebel[2]
NASA Ames Research Center, Moffett Field, CA, 94035

and

Matthew Daigle[3]
University of California, Santa Cruz, NASA Ames Research Center, Moffett Field, CA, 94035

Verification and validation (V&V) has been identified as a critical phase in fielding systems with Integrated Systems Health Management (ISHM) solutions to ensure that the results produced are robust, reliable, and can confidently inform about vehicle and system health status and to support operational and maintenance decisions. Prognostics is a key constituent within ISHM. It faces unique challenges for V&V since it informs about the *future* behavior of a component or subsystem. In this paper, we present a detailed review of identified barriers and solutions to prognostics V&V, and a novel methodological way for the organization and application of this knowledge. We discuss these issues within the context of a prognostics application for the ground support equipment of space vehicle propellant loading, and identify the significant barriers and adopted solution for this application.

Nomenclature

EOL	=	End Of Life
ISHM	=	Integrated Systems Health Management
PHM	=	Prognostics and Health Management
RUL	=	Remaining Useful Life
V&V	=	Verification and Validation

I. Introduction

INTEGRATED Systems Health Management (ISHM) will play a major role in future space missions to ensure safe operations while containing mission cost. One element of ISHM that is receiving increased attention is prognostics. The importance of the role of prognostics is acknowledged in the term "Prognostics and Health Management" (PHM) which is often used synonymously with ISHM. The task of prognostics is to quantify the health (or damage) of components or subsystems and – in case an abnormal condition has been detected – to estimate the remaining life of the component or subsystem. This functionality will provide advance notice of critical information to decision makers, and thus enable them to significantly improve operations.

This paper discusses a prognostics application for components in launch operations. Some of the underlying theoretical concepts are discussed and particular attention is given to verification and validation (V&V). Indeed, V&V will be one of the important considerations in gaining acceptance of systems health management in general and prognostics in particular. In preparation for this, we have conducted a literature survey pertinent to V&V of prognostics. Our aim is to use the information contained in those papers to guide how we will approach V&V for

---

[1] Principal, Quality Assurance Office, 4800 Oak Grove Dr, MS 125-233, Pasadena CA 91109.
[2] Senior Scientist, Prognostics Center of Excellence, NASA Ames Research Center, MS 269-4, Moffett Field, CA 94035. Member, AIAA.
[3] Associate Research Scientist, Intelligent Systems Division, NASA Ames Research Center, Moffett Field, CA 94035.

our particular application of prognostics. Towards this end, we extracted from those papers what we recognized as "barriers" to V&V of prognostics, what we recognized as "solutions" to overcome those barriers, and the connections between the two. While we do not make a claim for completeness of barriers and solutions found, we feel that a major contribution of this paper is the method to break down the V&V problem into manageable portions. It should also be noted that the large body of work for V&V outside the PHM domain has not been considered for this paper.

The paper is organized as follows. In Section II, we describe briefly our prognostics application, and the challenges to V&V that it introduces. In Section III, we present our method to determining barriers and solutions to prognostics V&V, and follow with specific examples of these in Section IV. The application of this method to our prognostics application is introduced in Section V, and specific examples follow in Section VI. Conclusions are found in Section VII, and listings of the titles of our barriers and solutions are in the Appendix.

II. Setting: Prognostics Application for Components in Launch Operations

Prognostics deals with forecasting the health state of components and systems, enabling condition-based maintenance. The goal of prognostics is to predict when a component can no longer fulfill a desired functional requirement. This time point is defined as the *end of life* (EOL), and the amount of time until that point is defined as the component's *remaining useful life* (RUL). Over time, components undergo wear and aging and may experience fault conditions due to normal usage and environmental factors but also due to abnormal usage stressors and environmental stressors. Prognostics is based on analysis of component failure modes and detection of early signs of wear, aging, and fault conditions, which are then correlated to damage propagation models to predict future damage and EOL.

We are developing a prognostics application for components used in launch operations of space vehicles. Specifically, we are interested in components that comprise the vehicle fueling system, which transfers cryogenic propellant from a storage area to the vehicle tank. In this system, we have identified various valves and pumps as components that would benefit most from the use of prognostics. We identified the components by focusing on the intersection of upfront analytical assignment of degree of fault impact, disruptions observed in past operation, and hardware configurations of the present system. Specifically, for the system at hand, we looked at the criticality of components, the number of components in operation, the frequency of fault, failure, and maintenance activities, and the available sensing capabilities.

Throughout the paper, we will describe the case of pneumatic valve prognostics as our example. These valves are actuated by gas, and can use different types of actuators. A normally-closed valve with a linear cylinder actuator is depicted in Fig. 1. The valve is opened by filling the chamber below the piston with gas up to the supply pressure, and evacuating the chamber above the piston down to atmospheric pressure. The valve is closed by filling the chamber above the piston, and evacuating the chamber below the piston. The return spring ensures that when pressure is lost, the valve will close due to the force exerted by the return spring. From looking at known failure modes and specific instances of failure modes that have actually occurred, we can determine which of these are important to capture. For example, leaks may occur at the valve ports or internally over the piston, the spring may degrade over time, and friction may increase as a result of sliding wear and lubrication breakdown. The only sensors available provide discrete valve opened and closed signals, from which only open and close times can be derived. This makes the task of prognostics difficult.

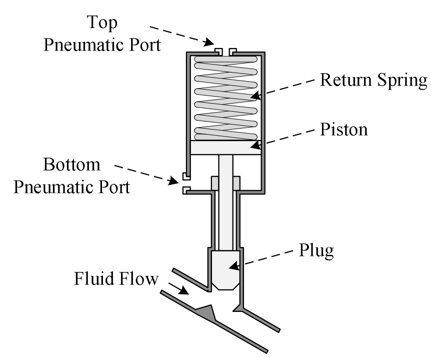

Figure 1. Pneumatic Valve.

The overall health management architecture is depicted in Fig. 2. Sensor data from the ground systems equipment are fed to the user interface and systems control module, which decides how to control the system based on user and automatic feedback mechanisms. It also displays results from the health management system, which may impact control decisions, for example, if a valve is predicted to reach EOL before loading operations are complete, a switch to a redundant path may be planned. The message bus transmits sensor data and commands to the health management interface, which packages the diagnosis and prognosis results from the individual modules. The prognostics module uses sensor data, commands to perform prognostics, and uses results from the diagnostic module to focus the efforts. For example, the prognostics component for the pneumatic valve would receive valve open and

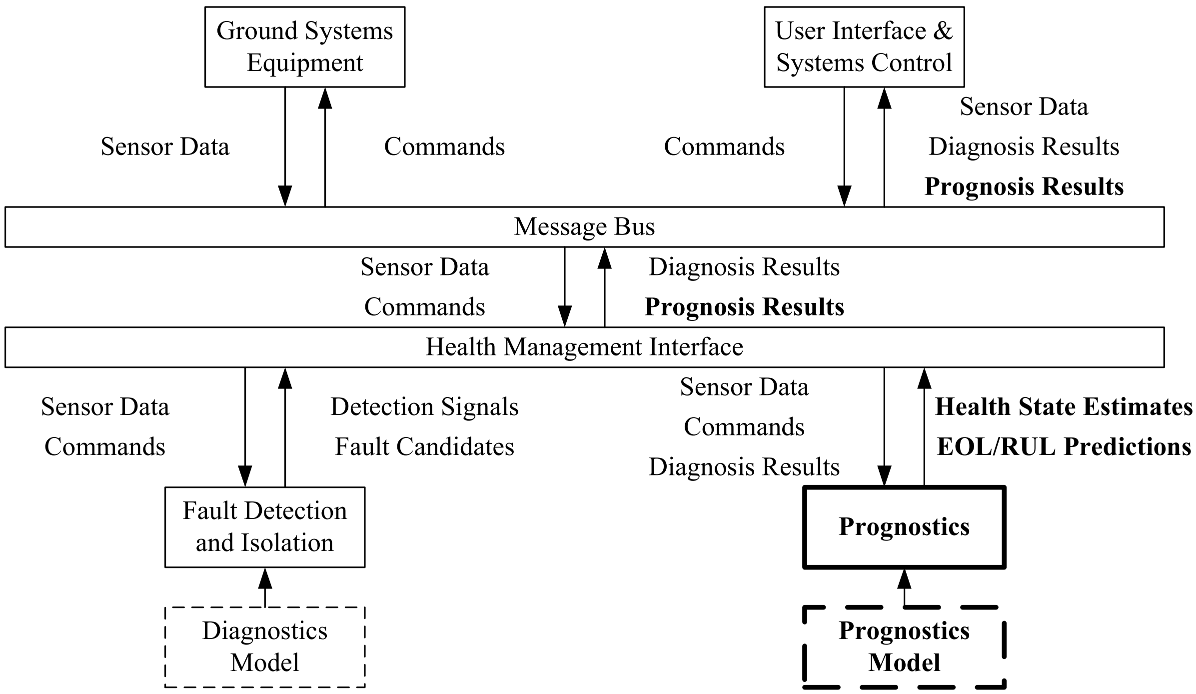

Figure 2. Health Management Architecture.

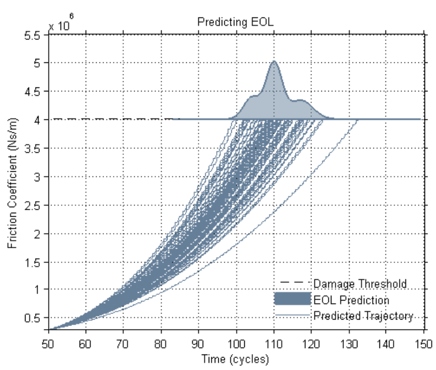

Figure 3. Sample trajectory output

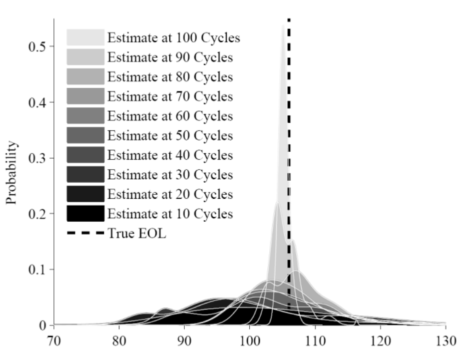

Figure 4. Progression of EOL predictions.

close commands, and receive measurements indicating when a valve is fully opened or closed, as well as the operating pressures of the fluid. Estimates of the health of components and component EOL and RUL predictions are computed in the prognostics component, such as the number of valve cycles until failure. Figs. 3 and 4 show sample outputs from the prognostics component. In Fig. 3, predicted trajectories are shown based on the current estimated health state distribution, based on an increase in the amount of valve friction. The predictions form a distribution that may inform the user as to the future component behavior and expected RUL. Figure 4 shows the progression of EOL predictions for the case of an internal leak. The probability distributions provide a measure of the prediction uncertainty. As time progresses and more data are collected, the results become more confident about the true EOL.

Prognostics methods may typically be categorized as either data-driven or model-based. Data-driven approaches[1] use learning methods to identify trends and determine EOL and RUL. Such methods rely on large amounts of run-to-failure data that are used to train the algorithms. In our case, such information is not readily available. Instead, we are developing a model-based approach that exploits domain knowledge of the system, its components, and how they fail in order to provide EOL and RUL predictions[2-4]. The underlying physical phenomena are captured in a physics-based model that is derived from first principles. The particle filter[5] is a popular algorithm for model-based prognosis[3,4,6]. Particle filters approximate the posterior probability distribution of the system state as a set of discrete, weighted samples, called *particles*. The system state is augmented to include variables that characterize the health of the component and the amount of damage present. Although suboptimal, the advantage of particle filters is that they can be applied to systems which may be nonlinear and have non-Gaussian noise terms, where optimal solutions are unavailable or intractable. Further, they are very flexible and help to manage the various sources of uncertainty in prognostics through the use of explicit probability distributions.

Developing a model-based prognostics approach in the selected system domain raises many significant challenges to successful V&V. This entails the construction of both models of nominal operation, but also the progression of damage. Model-based approaches require accurate, reliable models to achieve useful predictions. However, constructing models in the extremes of cryogenic temperatures is difficult, as such factors may have a significant impact on how components age. Complex processes such as cavitation are difficult to model, and their effects cannot be ignored. Not only do models need to be validated in nominal operation, for which data are typically plentiful, but also in faulty operation, for which data are usually quite rare. For example, gas leaks appear often in the valves, but these are typically caught before the valves are used in a launch operation. Even if they occur during a launch operation, the evidence may be hidden within the sensor data. In Section V we will examine these and other challenges in further detail as they relate to pneumatic valve prognostics.

III. Gathering and Organizing V&V Information from the Literature

In preparation for planning V&V of our prognostics application, we began a literature survey to locate information that might be used as guidance for how to approach the V&V subject and to extract and record as much as possible the information useful to V&V planning. Generally, technical papers were found to be significantly more informative than presentation materials. The papers studied in this manner were Refs. 7-13.

Information was classified into one of two groupings and relationships between the groups were established. These groupings were:

- A group of “Barriers”, i.e., those things mentioned that could (depending on the particular situation) get in the way of successful, cost-effective V&V of prognostics. (“Barriers”, “Impediments”, “Problems”, “Obstacles”, and “Risks” could be used synonymously for naming this set of information; henceforth we will refer to this set of information as “Barriers”). Each instance of these that we found we added to our growing set of Barriers, giving it a short, pithy title (enough to convey the gist of what it involves), while recording a more extensive quotation taken verbatim from the source to serve as a more thorough explanation, and finally recording the reference to that source.
- A group of “Solutions”, i.e., those things mentioned that could be done to overcome the aforementioned Barrier. (“Solutions”, “Remedies”, “Approaches”, and “Options” could be used synonymously for naming this set of information; henceforth we will refer to this set of information as “Solutions”). Again, we added each one of these to our growing set of Solutions, giving it a pithy title, and recording a lengthier quotation and the reference to the source.
- Relationships between Barriers and Solutions: typically within the papers the discussion of a Barrier is followed shortly thereafter by a discussion of means to overcome that Barrier – in our terms, the Solution(s). We recorded this relationship between Barrier and Solution as simple linkage between the two. On occasion a prognostics paper describes how a Solution overcomes a Barrier, but its application induces some additional Barrier(s), this too is recorded as a linkage between the Solution and the Barrier it introduces (taking care to distinguish this kind of Barrier-inducing-linkage from the Barrier-overcoming linkages). Other than distinguishing between linkages of Solutions that overcome Barriers from Solutions that increase Barriers, we made no attempt to differentiate between magnitudes of the effect.

Error! Reference source not found. Fig. 5 shows a portion of the connectivity implied by the relationships we recorded between barriers and solutions. The row of small red circles represent barriers, and the row of small green circles solutions. A black line connecting solution to barrier indicates we recorded the solution as overcoming the barrier, and a red line indicates we recorded that solution as inducing (or exacerbating) the barrier. We used some of the capabilities of JPL-developed risk management software, DDP[14] to record the V&V information. This software offers several ways to view and utilize such relationships – e.g., for a specific

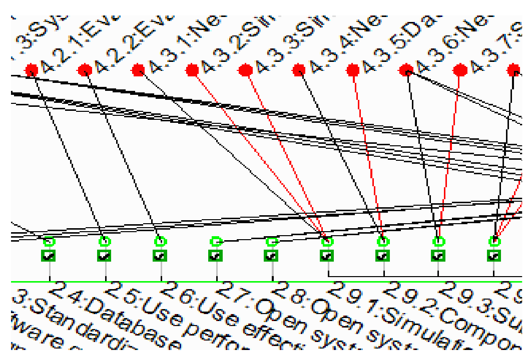

Figure 5. A view of a portion of the connections between barriers and solutions.

barrier, list the solution(s) available to help overcome it, keep track of the selected solutions and the additional barriers they induce, etc. This is a fragment of a screenshot taken of the software we used to record and help utilize the accumulation of V&V information.

When very similar information (Barriers and/or Solutions in our terms) was found to be repeated in multiple sources, no duplicates were created (Barrier or Solution). Instead, if the second source provided a helpful alternate phrasing of the same topic, it was recorded as additional descriptive information associated with the original instance (Barrier or Solution) and added a reference to that second source. In retrospect our feeling is it would have been useful to always have taken the time to add the second reference, even in those cases where we chose not to add anything to the original descriptive information – doing so would have given us at the end of this process a more thorough index into all the papers we processed.

In the course of this process, we observed that of all the papers we examined, we found only one discussion of *verification*, in Ref. 8, which stated:

> The process of determining that a PHM technology accurately represents the developer's conceptual description as a function of the design specifications. … Verification deals with the relationship between the conceptual system and the realized system and that Validation deals with the relationship between the experimental measurements and reality…Verification, while not simple, is much less involved than the more complex statistical nature of validation. Generally software quality assurance tools can be used to address programming errors and verify the correctness of the software.

Our impression is that verification challenges in prognosis are not very different from those in diagnosis, and verification challenges in general.

IV. Examples of Barriers, Solutions and Links between them

This section shows an illustrative example of identifying barriers and solutions, and the links between them to illustrate how the information was compiled.

The introduction of Ref. 8. begins by stating:

> In recent years, many Prognostics and Health Management (PHM) products have been released to market incorporating near real-time, automated monitoring, fault detection and isolation capabilities and advanced prognostic prediction capabilities. There is a growing need to develop standard methods for quantifying the performance and effectiveness of these tools.

This inspired the generation of a Barrier, given the pithy title of "Lack of standard methods for quantifying performance and effectiveness of PHM tools". The longer quotation was recorded to serve as its explanation; the reference to its source was also recorded.

The paper referenced above goes on to state:

> In order to assess the accuracy of the PHM technologies, Impact Technologies is developing a PHM Verification and Validation (V&V) Test Bench. The Test Bench is software application that contains fault pattern operational data and ground truth health information for a library of components to be provided to PHM technologies from the sensors through the reasoners.

This inspired the generation of a Solution, given the pithy title "V&V Test Bench". Again, the longer quotation and its reference were recorded. This Solution was linked to the aforementioned Barrier.

Another paper, Ref. 12, makes the following closely related statement about standard metrics:

> A key step in successful deployment of a PHM system is prognosis certification. Since prognostics is still considered relatively immature (as compared to diagnostics), more focus so far has been on developing prognostic methods rather than evaluating and comparing their performances. Tests are conducted based on specific requirements to declare the goodness of the algorithms but little or no effort is made to generalize the performance over a variety of other situations. Hence, there is no direct way of comparing different efforts if one needs to identify the most suitable algorithm among several. This calls for a set of general metrics that can be used in a standardized manner.

This inspired the generation of Barrier "Lack of a standard set of general metrics that can be used in a standardized manner (independent of application domain)". Interestingly, a Solution to this was found in the first paper Ref. 8, which in describing more about their Test Bench goes on to say:

> ...The PHM V&V Test Bench utilizes a standardized set of mathematical and business case metrics for evaluating the performance and effectiveness of PHM systems. The evaluation process is capable of assessing PHM technologies in terms of their ability to detect, diagnose and predict fault to failure progression of specific failure modes. Specific metrics for these capabilities include accuracy, reliability, sensitivity, stability, risk, economic cost/benefit, and robustness....

This inspired the addition of a Solution ("Standardized set of mathematical and business metrics"), linked to the metrics Barrier.

Shown above are examples where new entries to Barriers and Solutions were generated. The process of judging if another Barrier is part of an already existing one is subjective and is not shown here.

V. Applying the Gathered and Organized Information

The intent of gathering and organizing information from the literature in this Barriers-Solutions structure was to make it amenable to use when considering the V&V of a specific prognostics application. The approach we envisaged to using it was the following:

- Determine which of the listed Barriers are relevant to the prognostics application in question. The listed Barriers provide an organized and relatively well-populated structure to doing this. The pithy titles give a quick view into the information, the descriptive paragraphs (the key quotations extracted from the original papers) provide clarification, and finally the reference back to the sources allow quick access to those sources when there is the desire to read more of the context in the original.
- Consider the solutions listed as related to the relevant Barriers. Awareness of these Solutions may be helpful in suggesting approaches to overcoming the Barriers for the specific prognostics application.

We began an attempt to follow this approach by applying it to the prognostics application described in Section II. This was done by having the author (who had gathered and organized the V&V information into the Barriers-Solutions structure and who acted as the facilitator) step through the Barriers information and ask questions of the other two authors (who acted as the domain experts), taking notes on the discussions that ensued. The results of these discussions were then recorded as additional notes associated with the Barrier or Solution.

Overall, only some of the Barriers were relevant to the prognostics application in question, as might be expected given the information was gathered from multiple papers that as a whole spanned a wide range of prognostics applications. Making the determination of which were relevant to our applications was usually easy fairly straight forward.

VI. Examples of Specific Results

Some examples of following our approach are summarized in Table 1, which lists in each row:

- **Barrier**: the instance of a generic Barrier relevant to our prognostics application, taken from our previously assembled list of such (the number that precedes the name of the Barrier is its number in our list of all Barriers – see the Appendix).
- **Barrier Manifestation**: the specific manifestation of that Barrier in our application,
- **Solution**: the specific solution approach we are following in our application,
- **Solution Drawbacks**: the drawbacks, if any, of that solution approach.

For example, a common Barrier is managing different users' objectives, goals, and requirements. For our specific application, our specific use of PHM (by tracking the health of components and predicting how long they will last) aims to address the topmost metrics that motivate use of such predictions. In our case, that is indicating the need for proactive action to avoid an interrupted launch countdown due to failure of ground equipment, or, before the start of ground operations, recognizing which components will reliably last the duration of their corresponding ground operations, and initiating component replacements if warranted by confident prognostic predictions.

A number of the relevant Barriers were assessed to be 'solved problems' as far as this particular prognostics application was concerned. For example, for the Barrier (2) of "Different users of prognosis have different requirements; hence metrics should be tailored for each end user" we are able to take advantage of access to subject matter experts to help us understand what the user requirements are, so we can match offerings to those requirements. We have established that metrics that were previously developed[12,16] are sufficient to meet the user needs. For the valve, we have decided on using practical metrics, such as prognostic horizon (the time when the first confident prediction can be made), relative accuracy (a measure of prediction accuracy relative to true RUL), and the α-λ metric (a true or false indication of whether a fraction of β of the probability mass of the prediction within a bound of α from the true RUL, for a given time point λ). Figures 6 and 7 show two such α-λ plots for prediction of spring damage of the valve. The two plots illustrate the improvement in performance that can be achieved with the addition of particular sensors. Note that the percentages denote how much of the probability mass is contained within the α bounds.

In response to Barrier (4.1.1) we have developed a physics-based simulation of nominal and faulty valve behavior, including wear and aging processes[4]. Using this simulation, we may develop our prognostic algorithms and evaluate them under many different scenarios including different amounts of noise, different fault magnitudes, and different sensor suites[4,15].

Table 1. Summary of some specific examples of following our process.

Barrier	Barrier Manifestation	Solution	Solution Drawbacks
1. Confusion about terms and definitions	Require clear terms and definitions for communication of technical approach and results.	Adopt terms and definitions used in aerospace industry.	None identified.
2. Lack of standard methods and metrics for evaluation	Require evaluation of algorithm performance and comparison to other algorithms.	(Solutions 2.5, 3.2) Adopt prognostics performance metrics described in[12,16].	None identified.
4.1.1 Lack of ground truth information	Proper evaluation of prognostic results requires knowledge of actual ground truth information (true EOL and RUL values). This data is not present in actual launch operations due to regular maintenance and safety constraints that prevent running a component to actual failure.	(Solution 2.9.1) Develop simulation models based on detailed physics knowledge of component operation, fault modes, and damage progression processes. Evaluate algorithms using large sets of simulation data under various conditions. Also utilize inspection results of partly damaged components.	(Barriers 4.3.2, 4.3.3) Simulation models may be expensive to develop, require domain expertise, and they in turn require validation to be used reliably, which may be difficult due to the initial lack of ground truth information.
4.3 Incomplete mission data	Mission data covers nominal operation, but very few instances of fault conditions or failures occur during actual missions, with which prognosis algorithms may be validated.	(Solution 2.9.1) Develop system-level simulations that cover conditions of interest. Validate simulation against whatever data is available.	(Barriers 4.3.2, 4.3.3) System-level simulations are difficult and expensive to develop, require a great deal of domain expertise, and are themselves difficult to validate.
4.3.14 Design changes render existing databases unusable	Historical data may contain instances of failures, however since those times components have been modified or operating conditions changed to prevent reoccurrence of observed faults.	Determine extent to which design changes affect fault modes of interest, as some of these are still possible, but more unlikely.	New fault modes are not captured.
4.4.2 Difficulty to identify anomalous operation associated with incipient faults	Incipient faults are slowly developing and may be hidden within sensor noise before their effects are observable.	Utilize statistical methods to separate sensor noise from abnormal behavior using model-based algorithms.	An accurate model of the process is needed in order to reliably distinguish incipient faults from sensor noise.
4.4.5 Robustness to noisy data	Indications of faults can be hidden within noisy measurements, e.g., in the discharge flow sensor of a pump, where small changes in flow indicate onset of failure.	Utilize particle filter algorithm to handle sensor and process noise, and provide uncertainty estimates. Evaluate and tune algorithms under different noise conditions using simulation models or actual data with more noise superimposed.	Assumptions about mathematical form of noise sources may not be realistic.
4.4.9.3 Computational resources	Computational resources may be limited for certain equipment, or quality of prognostic results may not be reliable enough to justify the need for high computational resources.	Utilize particle filter algorithm, where the computational complexity may be tuned by limiting the number of particles used in the probability distribution approximation.	Reducing the number of particles used trades off estimation and prediction accuracy and precision.

Developing a simulation of the much larger cryogenic loading system that includes these component models is much more difficult, but the value of such a large-scale simulation would be to provide a more realistic idea of the conditions under which the component is being used, and to reveal unexpected side effects that might have been overlooked (e.g., a pressure shock wave, which may manifest only when the whole system is simulated). A system-level simulation is under development for the cryogenic loading scenario (Barrier 4.3).

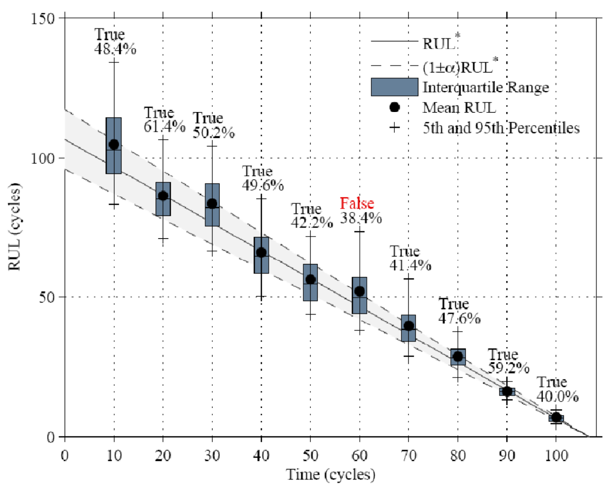

Figure 6. α-λ metric for spring damage prediction, where α=0.1, β=0.4, and continuous valve position and gas pressures are measured.

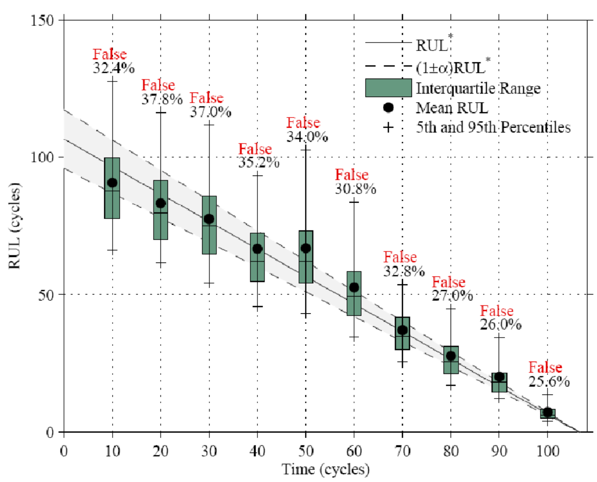

Figure 7. α-λ metric for spring damage prediction, where α=0.1, β=0.4, and only discrete valve position sensors are used.

This solution itself presents additional barriers, since simulations may require a lot of resources to build, and the simulator must be validated as well so that predicted fault and failure simulations are trustworthy. In our case, individual models of the valves are relatively inexpensive to design, but the development of the system-level simulation is much more demanding. To cope with the model complexity we leverage component-based modeling strategies.

Assessing the valve model's fidelity is another key issue. We have specific information about the valve such as its dimensions, operating pressures, and properties of the pneumatic gas. This information allows particular parameters of the model to be preset, but the remaining parameters need to be determined from actual data. For the case under consideration, the only sensors available provide discrete valve-opened and valve-closed signals. One can derive only open and close times from this information and we have tuned the model parameters to match the real open and close times. However, in the absence of richer data, it is difficult to resolve whether the simulation model accurately captures all failure mechanisms. It is envisioned for this application to pursue a validation through time, i.e., field data (both sensor measurements and maintenance feedback) will be used to fine-tune the model parameters, if needed.

To manage model complexity, simplifying assumptions must also be made. Models of noise, which are often Gaussian, may be simplifications of the real world (Barrier 4.4.5). But, we can inspect the available data to see if such assumptions are valid. Fault and damage models are also typically mathematical simplifications of the true process. For example, a model of faults in pump bearings may model damage as a continuous exponential growth; in reality, damage is not completely continuous. Individual flakes (debris) are liberated at distinct points in time, and the sizes of those flakes vary. Thus damage actually progresses in a series of steps, to which we fit a line of continuous exponential growth. In this case, damage is a stochastic process, so prediction cannot be exact. Instead, we employ legitimate approximations, for which the inaccuracy is within the required bounds.

The use of the particle filter as the basis for our technical approach helps to manage some of the barriers. Particle filters are state observers, which automatically help to reduce the effects of sensor noise, process noise, and model uncertainty in determining a robust estimate of the state of the system. They deal directly with probability distributions (approximated discretely), so can readily incorporate various sources of uncertainty, including uncertainty in future system inputs, which may greatly affect the prognostics task. In our case, much of this uncertainty is naturally removed, as the propellant loading process is well-defined a priori. Because particle filters approximate probability distributions in a discrete, sample-based fashion, they may consume a lot of computational resources depending on the number of particles selected, which corresponds directly to accuracy and precision of estimation (Barrier 4.4.9.3). This number can be tuned depending on the available resources. In our application, this is not as much of an issue because the algorithms are operating in a ground-based scenario, where computational resources are not very limited, unlike for in-space scenarios.

VII. Conclusion

We believe our approach to gathering and organizing prognostics V&V information from relevant literature, and then applying it to our specific prognostics application, provides a novel methodological way to approach V&V. Conventional literature surveys have a different purpose – they attempt to distill general themes, offer comparisons and contrasts, etc. Our intention was to provide an aid that allows the application of knowledge gleaned from the literature. It also allows the user to break down the V&V process into smaller segments and to identify potential bottlenecks which then allow focusing the attention. The specific approach we took to organizing the information – into categories of "Barriers" and "Solutions", where each is accompanied by explanatory text from the original sources, coupled with references to the sources themselves – seems to have worked reasonably well overall. There are aspects that could be improved; the descriptive text recorded proved on occasion to be insufficient to serve as a standalone explanation of the item in question – fairly often we found the need to trace back to the original source and read more of the explanatory context; the reference to the source helped, but still meant a somewhat cumbersome process; we also feel it would have been better to have taken the time to record a reference back from an item (Barrier or Solution) to *all* the sources where that item, or its equivalent, were discussed; the overall hierarchical organization of these items (see Appendix for the listing of their titles) could probably be improved upon. What we found to be most useful was to use the Barriers as a series of talking points, taking notes as we went along as to our prognostics experts' understanding of whether, and if so how, that Barrier applied to our application. This process both serves as a means to capture the rationale that justifies faith in the prognostics application's approach, and to stimulate (and again capture) thoughts on areas of concern and possible approaches to addressing them.

Appendix

Tables 2 and 3 list the pithy titles we formed for the Barriers and Solutions identified from the literature. As explained in the body of the paper, we recorded additional information with each of these – more extensive quotations from source document to serve as explanatory text, and references to the source documents themselves. Such additional information is omitted from the list below. The hierarchical numbering of these items stems from our organization of these items into simple groupings. It should also be noted that the number of quotations for particular papers does not necessarily reflect the importance of such paper to the work in general but the order in which the papers were reviewed.

Table 2. Barriers (titles of).

1: Confusion about terms and definitions
- 1.1: Terms lack standard, uniform definitions[9]
- 1.2: Different users of prognosis have different requirements; hence metrics should be tailored for each end user[12]
- 1.3: V&V definitions and approaches differ by industry, application objective and scope[8]

2: Lack of standard methods and metrics for evaluation
- 2.1: Lack of standard methods for quantifying performance and effectiveness of PHM tools[8]
- 2.2: Lack of a standard set of general metrics that can be used in a standardized manner (independent of application domain)[12]

3: Verification challenges
- 3.1: Programming errors and software correctness[8]

4: Validation challenges
- 4.1: Validation relies on ground truth information
  - 4.1.1: Lack of ground truth information sources[8]
  - 4.1.2: Inappropriate selection of ground truth[8]
  - 4.1.3: System overtraining[7]
- 4.2: Evaluation appropriate to level: component or system
  - 4.2.1: Evaluation of technical performance and PHM algorithms at component/ subsystem level[8]
  - 4.2.2: Evaluation of system level capabilities - operational goals, economic cost/benefit[8]
- 4.3: Incomplete mission data
  - 4.3.1: Need for signals, noise and fault signatures[8]
  - 4.3.2: Simulation models expensive to develop[8]
  - 4.3.3: Simulated data may not always be realistic[8]
  - 4.3.4: Need inexpensive way of producing fault and fault progression data[8,10]
  - 4.3.5: Data collected on component or subscale test rigs may have limited applicability for actual system fault observability[8]
  - 4.3.6: Need subsystem/system fault characterization data[8]
  - 4.3.7: Seeded faults may not be entirely realistic of natural fault[8]
  - 4.3.8: Need vehicle/mission data[8]
  - 4.3.9: Data from dedicated missions does not cover full range of potential conditions[8]
  - 4.3.10: Data from dedicated missions has no critical faults or progression[8]
  - 4.3.11: Need data from full range of mission, operational, and environmental conditions[8]
  - 4.3.12: Data from a technology maturation field program limited to opportunistic fault occurrence[8]
  - 4.3.13: A technology maturation field program delays use and implementation in field [actual mission] [8]
  - 4.3.14: Design changes render existing database of faults unsuitable for V&V purposes[10]
- 4.4: Assessment needs
  - 4.4.1: Detection confidence level[8]
  - 4.4.2: Ability to identify anomalous operation associated with incipient faults[8]
  - 4.4.3: Stability of output fault confidence level[8]
  - 4.4.4: Range of operating (duty) conditions over which PHM system will detect anomalies[8]
  - 4.4.5: Robustness to noisy data[8]
  - 4.4.6: Accuracy of prediction[8]
  - 4.4.7: Precision of prediction[8]
  - 4.4.8: Confidence of prediction[8]
  - 4.4.9: Cost Assessment
    - 4.4.9.1: Acquisition and implementation costs[8]
    - 4.4.9.2: Operation and maintenance costs[8]
    - 4.4.9.3: Computational resources[8]
    - 4.4.9.4: Susceptibility to unexpected behavior due to unforeseen events[8]
    - 4.4.9.5: Net value of a PHM technology[8]
  - 4.4.10: Different users' objectives/goals/requirements
    - 4.4.10.1: Program Manager's goals - assess economic viability[12]
    - 4.4.10.2: Field commander's / Plant Manager's goals - resource allocation and mission planning[11,12]
    - 4.4.10.3: Operator's goals - action and re-planning during mission[12]
    - 4.4.10.4: Maintainer's goals - know when to perform maintenance[11,12]
    - 4.4.10.5: Designer's goals - implement/improve the prognostic system[12]
- 4.5: Lack of benchmark datasets or models to evaluate prognostics systems[12]
- 4.6: Lack of standardized methodology for performance evaluation[12]

5: Other V&V challenges
- 5.1: Difficulty of developing information analysis tools[8]
- 5.2: Need to protect proprietary algorithms and software design approaches[8]

6: PHM design & maintenance challenges
- 6.1: Cannot identify the most suitable prognostic algorithm among alternatives
  - 6.1.1: Lack of generalized results for performance of prognostic algorithms[12]
  - 6.1.2: Confusion between need for general metrics and need for tailored metrics[12]
- 6.2: Difficult to specify crisp and unambiguous requirements[12]
- 6.3: Some applications lack an explicit declaration of fault detection[12]
- 6.4: Some algorithms cannot start predicting as soon as a fault is detected[12]
- 6.5: Overall PHM system performance depends on detection and diagnosis[7]
- 6.6: High cost of correcting an erroneous PHM system once fielded[8]
- 6.7: Limitations of the overall PHM approach
  - 6.7.1: Physics-of-failure (POF) models computationally prohibitive to apply at the system level[13]
  - 6.7.2: Reliability methods do not handle idiosyncrasies of specific systems[13]

Table 3. Solutions (titles of).

- 1: Definitions
  - 1.1: Verification[8]
  - 1.2: Validation[8]
  - 1.3: Conceptual System[8]
  - 1.4: Realized System[8]
  - 1.5: Prognostics[12]
  - 1.6: Confidence vs. performance[12]
- 2: Methods/approaches
  - 2.1: V&V prognostics system prior to deployment[8]
  - 2.2: Software quality assurance tools address programming errors and verify correctness[8]
  - 2.3: Standardized set of mathematical and business metrics[8]
  - 2.4: Database of testing, simulation or in-service fault data[8]
  - 2.5: Use performance metrics to evaluate technical performance and accuracy of the PHM algorithms at component or subsystem level[8]
  - 2.6: Use effectiveness measures to evaluate system level capabilities in terms of achieving overall operational goals and economic cost/benefit[8]
  - 2.7: Open systems architecture (OSA) standards[8]
  - 2.8: Open systems architecture (OSA) architecture[8]
  - 2.9: Obtaining data for use in PHM validation
    - 2.9.1: Simulation Models and Fault Generation[8]
    - 2.9.2: Component and LRU Fault and Failure Tests[8]
    - 2.9.3: Subsystem/System Fault Characterization Tests[8]
    - 2.9.4: Dedicated mission for normal and off-normal test and evaluation[8]
    - 2.9.5: Technology Maturation Field Program[8]
    - 2.9.6: Adaptive Learning[7]
    - 2.9.7: Highly Accelerated Stress Simulation[7]
    - 2.9.8: Seeded Fault Testing[11]
    - 2.9.9: Accelerated Mission Testing[11]
    - 2.9.10: Run components/systems until failure[10]
  - 2.10: Ground Truth Sources[8]
  - 2.11: Historical data akin to current process[12]
  - 2.12: Data sets with associated ground truth information[8]
  - 2.13: Benchmark datasets or models[12]
  - 2.14: Advance planning to avoid system overtraining
    - 2.14.1: Robust system testing[7]
    - 2.14.2: Use of multiple classifiers through knowledge fusion[7]
    - 2.14.3: Data reserve/hold-out techniques[7]
- 2.15: V&V Test Bench[8]
- 3. Metrics
  - 3.1: Metrics for detecting and classifying system faults[8]
    - 3.1.1: Accuracy metric[8]
    - 3.1.2: Detection Threshold Metric
- 3.2: Prognostics metrics
  - 3.2.1: Performance metrics
    - 3.2.1.1: Stability Metric[8]
    - 3.2.1.2: Detection Duty Sensitivity Metric[8]
    - 3.2.1.3: Noise Sensitivity Metric[8]
    - 3.2.1.4: Accuracy of Prediction[8]
    - 3.2.1.5: Precision of prediction[8]
    - 3.2.1.6: Confidence[8]
  - 3.2.2: Effectiveness metrics
    - 3.2.2.1: Implementation Cost Metric[8]
    - 3.2.2.2: Operation & Maintenance Cost Metric[8]
    - 3.2.2.3: Computational performance metrics
      - 3.2.2.3.1: Computer Resource Metric[12]
      - 3.2.2.3.2: Comp. Sci. Big O notation for describing computational complexity of algorithms[12]
      - 3.2.2.3.3: CPU time / elapsed time[12]
      - 3.2.2.3.4: Samples per unit time capacity[12]
      - 3.2.2.3.5: Memory space utilization[12]
    - 3.2.2.4: System Complexity Metric[8]
    - 3.2.2.5: Technical Value[8]
    - 3.2.2.6: Total Value / Life cycle cost[8,12]
    - 3.2.2.7: Return on investment (ROI)[12]
- 4: Design
  - 4.1: Good PHM system design techniques[7]
  - 4.2: Prognosis carried out on a decay process[12]
  - 4.3: Continuous data collection[12]

Acknowledgments

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology; NASA Ames Research Center; and UC Santa Cruz, under a contract with the National Aeronautics and Space Administration. This work was funded in part by the NASA Exploration Technology Development Program (ETDP). We thank Barbara Brown for her guidance of this work.

References

[1]Schwabacher, M. "A survey of data-driven prognostics," *Proceedings of the AIAA Infotech@Aerospace Conference*, Reston, VA, 2005, AIAA.

[2]Roemer, M., Byington, C., Kacprzynski, G., and G. Vachtsevanos, G., "An overview of selected prognostic technologies with reference to an integrated PHM architecture," *Proceedings of the First International Forum on Integrated System Health Engineering and Management in Aerospace*, Napa, CA, Nov 2005.

[3]Saha B., and Goebel, K., "Modeling Li-ion battery capacity depletion in a particle filtering framework," *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2009*, San Diego, CA, Sept. 2009.

[4]Daigle, M., and Goebel, K., "Model-based prognostics with fixed-lag particle filters," *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2009*, San Diego, CA, Sept. 2009.

[5]Arulampalam, M.S., Maskell, S., Gordon, N., and T. Clapp, T., "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[6]Orchard, M., Kacprzynski, G., Goebel, K., Saha, B., and Vachtsevanos, G., "Advances in uncertainty representation and management for particle filtering applied to prognostics," *Proceedings of International Conference on Prognostics and Health Management*, Denver, CO, Oct. 2008.

[7]Byington, C.S., Roemer, M.J., Kalgren, P.W., and Vachtsevanos, G., "Verification and Validation of Diagnostic/Prognostic Algorithms," *Machinery Failure Prevention Technology Conference (MFPT 59),* Virginia Beach, VA, 2005.

[8]Dzakowic, J.E., and Valentine, G.S., "Advanced Techniques for the Verification and Validation of Prognostics & Health Management Capabilities," *Machinery Failure Prevention Technologies (MFPT 60),* Virginia Beach, VA, 2006.

[9]Kalgren, P.W., Byington, C.S., Roemer, M.J., and Watson, M.J., "Defining PHM, A Lexical Evolution of Maintenance and Logistics," *Autotestcon 2006* Anaheim, CA, 2006, pp. 353-358.

[10]Aguilar, R., Luu, C., Santi, L.M., and Sowers, T.S., "Real-Time Simulation for Verification and Validation of Diagnostic and Prognostic Algorithms," *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, Tucson, Arizona,10-13 July 2005.

[11]Roemer, M.J., Dzakowic, J., Orsagh, R.F., Byington, C.S., and Vachtsevanos, G., "Validation and Verification of Prognostic and Health Management Technologies," *IEEE Aerospace Conference*, Big Sky, MT, 2005 pp. 3941-3947.

[12]Saxena, J.C., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S., and Schwabacher, M., "Metrics for Evaluating Performance of Prognostic Techniques," *International Conference on Prognostics and Health Management*, Denver, CO, 2008.

[13]Sheppard, J.W., Kaufman, M.A., and Wilmering, T.J., "IEEE Standards for Prognostics and Health Management," *IEEE AUTOTESTCON 2008*, Salt Lake City, UT, 8-11 September 2008, pp. 97-103.

[14]Cornford, S.L., Feather, M.S., and Hicks, K.A., "DDP – A tool for life-cycle risk management," *IEEE Aerospace Conference*, Big Sky, MT, 2001, pp. 441-451.

[15]Daigle, M., and Goebel, K., "Model-based prognostics under limited sensing," *Proceedings of the 2010 IEEE Aerospace Conference*, Big Sky, MT, to appear.

[16]Saxena, A., Celaya, J., Saha, B., Saha, S., and Goebel, K., "On applying the prognostic performance metrics," *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2009*, Sept. 2009.