

Continuous validation of the PHM function in aircraft industry

Jayant Sen Gupta

Airbus Group Innovations, Toulouse, 31300, France
Jayant.SenGupta@airbus.com

Christian Trinquier

Airbus Group Innovations, Suresnes, 92150, France
Christian.Trinquier@airbus.com

Kamal Medjaher, Noureddine Zerhouni

FEMTO-ST Institute, AS2M department
Univ. Bourgogne Franche-Comté, Univ. de Franche-Comté/CNRS/ENSMM
24 rue Alain Savary, F-25000 Besançon, France
kamal.medjaher@ens2m.fr, zerhouni@ens2m.fr

Abstract— In the literature, most of Verification and Validation (V&V) papers deal with the design of a prognostics function and the verification and validation process necessary to put this function into service. Nevertheless, from an industrial point of view, even if this V&V procedure is successfully performed during the design and testing phase, it only considers data and knowledge available at the time the V&V procedure is carried out. However, it is possible that the systems evolve during operations causing a drift from the initial assumptions which could lead to a loss of performance, especially as the systems age. This paper aims at taking into consideration the total lifecycle of a PHM function, from the design and testing phase to the operation phase (after its entry into service).

Keywords— Prognostics and Health Management, Verification & Validation, Remaining Useful Life, Prognostics metrics.

I. INTRODUCTION

During the last decade, PHM activity attracted significant industrial and research interest due to the need for health assessment, fault detection, fault prediction and decision models, which are important tasks for performing efficient condition based and predictive maintenances [1 – 4]. Most of the PHM works reported in the literature deal either with the design phase (how to integrate PHM functions within a system) or with the operation phase (how to assess and predict the health of a system in operation), but none concern both phases [5 – 10].

This paper aims at taking into consideration the total lifecycle of a PHM function, from the design and testing phase to the operation phase (after its entry into service). This work is motivated by the fact that most of Verification and Validation (V&V) papers deal with the design of prognostics functions and the verification and validation process necessary to put them into service. Nevertheless, from an industrial point of view, even if this V&V procedure is successfully performed before the entry into service, it only considers data and knowledge available at the time the V&V procedure is carried out. Indeed, it is possible that the information and the

assumptions made during design are no more valid during operations, especially as the systems age.

Let us consider a fleet of aircrafts. A set of systems on each aircraft is monitored by using PHM functions that were defined during the design of the aircraft or during its early operations. Different elements may evolve during the aircraft operations. For each system, the failure or degradation modes are selected during the design of the PHM function depending on their occurrence and operational or cost impact. Once in operation, these choices may not be valid anymore as unexpected failures occur. Moreover, during the design of the PHM function, assumptions were made on the operational scenarios. For instance, the first aircrafts that were delivered may have been used as reference for operational scenario. Once the aircrafts are delivered to other companies, the operations (missions, maintenance operations, etc.) may be different enough to make the performance of the PHM function not fit these new airlines requirements. Thus, the adaptation of PHM functions to the operations of a specific airline may be mandatory to maintain the level of performance. Moreover, as the fleet ages, the environment of each system may be different from what it was just after its entry into service. The effect of degradation on a related system may have an impact on the degradation of the studied system that was not taken into account during the set-up of the PHM function causing a drift in its performance.

All the previous examples demonstrate the need to monitor the performance of the PHM functions to ensure a satisfactory quality level of the predictions. But, when addressing the performance of a PHM function, one has to describe the performance indicators that are used. As for prognostics, literature provides interesting performance indicators [11, 12]. Nevertheless, we consider not only prognostics assessment output but also health assessment outputs. Moreover, during operations, runs-to-failure are not necessarily available as maintenance may prevent reaching the end-of-life. The prognostic performance indicators of the literature that are based on the availability of runs-to-failure are thus to be

adapted. Finally, if the performance can be computed continuously, once it has reached a non-quality threshold, the PHM function has to be updated to cope with the operation constraints.

A graphical representation of this could be the enhancement of the well-known V-cycle into a “square-root shape cycle” with a continuous line representing the operations after the entry into service, as shown in Figure 1.

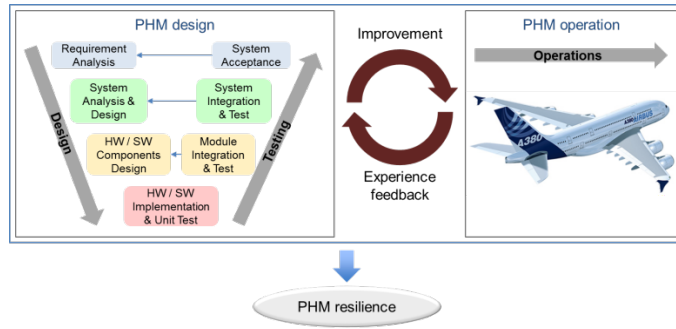


Figure 1: Enhanced V-cycle

Once put into service, the PHM function has to be continuously validated to detect potential drifts between initial assumptions and real life behavior. This point is essential to ensure the quality of the service provided by the PHM function throughout the life of the system it monitors.

The paper is organized as follows: after the introduction, section 2 describes the baseline organization that is considered in this work, section 3 details a list of possible reasons for which the once validated function may no longer be performing sufficiently well, section 4 presents the performance assessment that needs to be continuously done with real data, section 5 briefly describes the update principles and finally, section 6 concludes the paper.

II. BASELINE ORGANIZATION OF A MONITORING SERVICE ON A FLEET OF AIRCRAFTS

Before describing the different reasons of a loss of performance of a PHM function, we will briefly describe the setting we consider for the fleet.

This section presents the simple organization that is considered. A fleet of aircrafts is monitored by analyzing the data collected from each aircraft by PHM functions running on a ground station. These PHM functions monitor and predict the state of different systems and were developed and validated before the entry into service of the monitoring service. The monitored fleet is assumed to be civil aircraft and all predictions focus on operability and not safety.

The fleet of aircrafts may be owned by different operators, operating in different regions of the world with different mission profiles. The size of the monitored fleet usually increases as new aircrafts progressively enter into service.

This paper focuses on the computation of health and prognostic assessment but does not develop the different

usages that can be made of the information computed by the PHM algorithms.

III. POTENTIAL REASONS FOR A LOSS OF PERFORMANCE OF A PHM FUNCTION AFTER ENTRY INTO SERVICE

Now that the context has been described, the different reasons which could lead to a loss of performance of a PHM function between its initial validation process and the actual operating conditions will be developed. Note that this list is not exhaustive and may need to be completed or adapted to fit other contexts.

A. Arising of a failure mode or a degradation mode that was not foreseen by the engineering

Due to the potential lack of knowledge during the design of the PHM function, it is possible that a fault or degradation mode was not taken into account. For instance, it could be caused by an unexpected consequence of the fault of another component (leakage, overheat, etc.). If the fault or degradation mode is not taken into account, the prognostic or health assessment of this component will not be satisfactory and will not correspond to what was expected during the design phase.

B. Change in the missions of the fleet

The fleet composition is not constant in time. New aircrafts are produced and delivered constantly. So, the monitored fleet evolves in time, both in quantity and usage.

The evolution may also be within the same fleet operator. Indeed, it can operate new routes or change the repartition of the aircrafts among its existing routes.

For the aircrafts of the monitored fleet, it results in potential changes in operational (change in the operational cycles, loading of the system) and environmental (heat, humidity for instance) conditions that may question the technical choices (algorithms, learning database, etc.) of the initial PHM functions. The initial database on which algorithms were tuned may not be representative enough to ensure good performances of the PHM functions. It may also question the initial selection of failure and degradation modes that may be no more relevant for a part of the fleet.

C. Change in the maintenance operations of the fleet

The evolution of the fleet may result in differences in the maintenance operations, as each airline is able to negotiate with its own airworthiness authorities the implementation and adaptation of the maintenance program. The maintenance program may also evolve as the aircraft family gains maturity, following MSG-3 methodology Maintenance Review Board meetings. The potential evolution of the maintenance operations may have an impact on the validity of PHM functions. For instance, if the intervals between a lubrication task were increased, it could impact the performance of the prediction made by the PHM function. Another example could be the suppression of an inspection task by an airline where the operator made systematic measurements on the aircraft. In this case, if the PHM function needs these measurements, it cannot access them and the PHM function cannot be used.

D. Evolution of the aircraft

Another factor of potential drifts in PHM function performances is the evolution of the aircrafts. Regularly, the aircraft manufacturer proposes evolutions on the aircraft or its operations to take into account return of experience on the global fleet through "Service Bulletins". Moreover, with the huge development of computers and wireless communications in the aircraft, updates are made much easier than it was before, especially on software. The evolution of the software may have an impact on how the systems are controlled and operated and this could also have an impact on the performance of the PHM functions.

The aircraft configurations that may have been well mastered at the entry into service of the PHM function may quickly become much more difficult to track on the fleet. The different evolutions on each aircraft may have an impact on the performance of the PHM functions, changing the selection of most important failure and degradation modes, for instance.

E. Ageing of the fleet

In parallel to the evolution of the configuration of the fleet, each aircraft ages and this ageing may have an impact on its behavior. The degradation behavior of each system may evolve with time due to ageing effects on parameters of the system or due to the ageing of other systems in interaction. In this case, the models that used to be valid at the entry into service may not be as valid as they were before and this could lead to a degradation of the PHM function performance. A solution would be to use test-benches to study the effect of ageing, but it is highly unlikely that such studies could be made before the entry into service of the PHM function because of their induced cost.

F. Potential source of improvement due to new available parameters or new available data

If the PHM function is designed at the same time as the system of the aircraft, it is possible that not enough relevant data is available to use a data-driven model. The alternative is to use engineering models (based on the physical behavior of the system for instance) to characterize the nominal behavior of the systems. Once data is collected after the progressive entry into service of the fleet, PHM functions using data-driven models may prove to be more efficient than engineering models and the choice of modeling techniques that was best before the entry into service could change during the operation.

Another possibility which is similar to the previous one is the fact that new parameters are available thanks to an update of the system. Once again, the choice of the best technique for the PHM function may change.

IV. PERFORMANCE ASSESSMENT OF A PHM FUNCTION

Now that the reasons of a potential drift in the PHM function performance have been presented in the context of the monitoring of an aircraft fleet, the assessment of the performance of a PHM function will be addressed. Indeed, this performance assessment is necessary to detect performance

drifts induced by the different evolutions presented in the previous section.

A. Limitation of classical prognostic algorithms performance metrics

The literature on performance metrics for diagnostics and prognostics provides different metrics, for instance in the survey presented in [12]. In practice, most of the metrics described in the survey for diagnostics are computable. Nevertheless, in the case of prognostics, the computation of the performance metrics may become trickier. Saxena *et al.* [11] proposed different metrics for offline evaluation of prognostics performance. All of them rely on the knowledge of the end-of-life of the monitored system. In the context presented in this paper, during the operations of the fleet, the monitoring service aims at avoiding reaching the end-of-life. So, for systems that have a strong impact on operational reliability, the real end-of-life time will seldom be available. In the following, an adaptation of the definition of the end-of-life is proposed to cope with this operational constraint in the case where a health indicator is available.

B. Proposed alternative when a health indicator can be defined

First, the performance analysis is done *a posteriori* on a complete run-to-maintenance history. In the context of the work presented in this paper, the maintenance is done before reaching the end-of-life. Figure 2 is representative of the data available to compute the performance metrics *a posteriori*. In this figure, h is the health indicator, \bar{h} is the value of the failure threshold, h_m the value of health when maintenance is done and t_m is the maintenance time.

Using classical performance metrics is difficult because the prognostic algorithm will predict the time before reaching the failure threshold which is unknown in the case presented in Figure 2. The only solution to be able to use partially degraded histories in order to assess the performance of prognostic algorithms is to be able to set a common reference between prognostic predictions and each runs-to-maintenance.

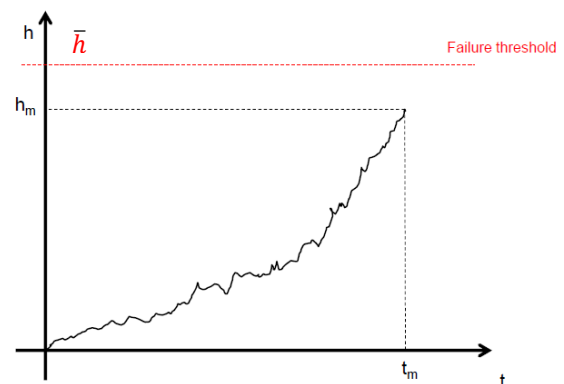


Figure 2: Available data after maintenance

The only reference time available for each run-to-maintenance is the time of maintenance. In the case where a health indicator is available, instead of predicting when the health indicator will exceed a failure threshold characterized by a limit value \bar{h} , we should rather predict the time before reaching the h_m threshold and compare it to our reference time t_m . For different runs-to-maintenance, these references may be different, as shown in Figure 3.

If maintenance was only decided using the health or degradation indicator, it could be expected that the states of a component when the maintenance is performed are relatively similar. Unfortunately, maintenance operations are more complex than that and the component may be removed using other decision criteria than the health state. Some examples of decision criteria are given below:

- the maintenance of the component may be decided because of a mandatory scheduled maintenance task;
- the maintenance of the component may be decided because of a failure which is not taken into account by the PHM function;
- the maintenance of the component may be decided to take advantage of another maintenance action in the neighborhood of the component (opportunistic maintenance).

The final state of the component may thus be quite different between different runs-to-maintenance.

This simple consideration implies an important change in the validation process. Instead of collecting Remaining Useful Life (RUL) predictions during the life of the component under use and compare them to the perfect prediction that will be known when the component reaches its end-of-life, a complete *a posteriori* process has to be set.

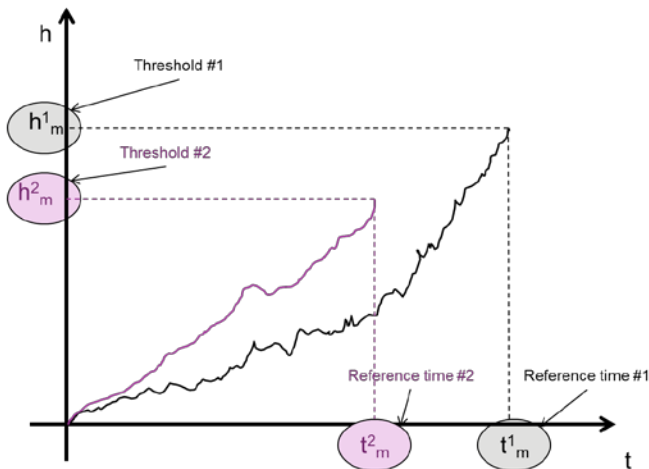


Figure 3: Different references for different runs-to-maintenance

C. Formalization of the adapted procedure

Let us consider a run-to-maintenance characterized by a set of couples: a health indicator and the associated time index, as shown in Equation 1.

$$\mathcal{H} = (h_i, t_i)_{i=0..m} \quad (1)$$

$\forall k = 1..m$, we can define a partial history of this run-to-maintenance $\mathcal{H}_k = (h_i, t_i)_{i=0..k}$ which can be used to predict the adapted RUL at time k, noted $aRUL_k$, and defined in Equation 2.

$$aRUL_k = \min\{t > t_k \mid h(t) > h_m\} \quad (2)$$

Now that this new definition of the Remaining Useful Life is set with a known reference time, it is possible to use all the classical performance metrics defined, for instance, in Saxena *et al.* [11].

D. Illustration for the prognostic horizon performance metric

The computation of performance metrics such as prognostic horizon, α - λ performance, relative accuracy or convergence is rather straightforward as it is very similar to the initial definition, differing only by the definition of the Remaining Useful Life (usage of the adapted remaining useful life) and of the reference time (usage of the time of maintenance instead of the end-of-life time). The following paragraph explains how to compute the equivalent of prognostic horizon. The interpretation of such a metric will be discussed in the next subsection.

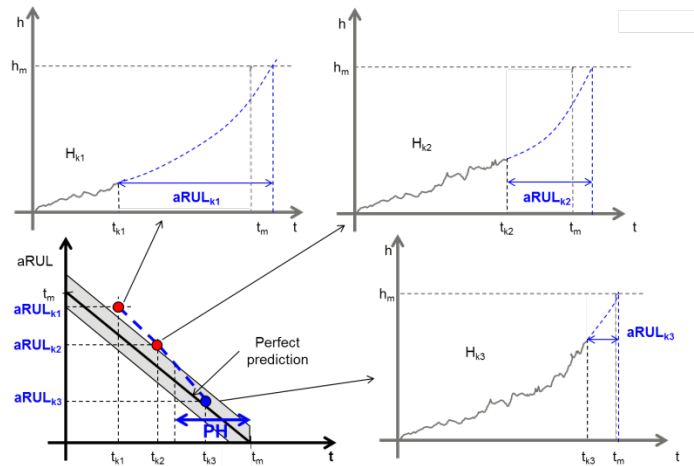


Figure 4: Prognostic Horizon Performance Metrics Illustration

Figure 4 shows how the $aRUL$ versus time plot is created. For each considered time step t_{k_i} (this could be done for each time step where data was collected or just a subset), the partial history \mathcal{H}_{k_i} is used to predict $aRUL_{k_i}$. In Figure 4, the Remaining Useful Life is predicted in a deterministic way to simplify the explanation in the graph but the prediction could

take into account the uncertainty and $aRUL_{k_i}$ could be a random variable.

This simple illustration shows that there is no difficulty in computing prognostic performance metrics similar to the one proposed in the literature with the adaptation proposed to cope with the absence of end-of-life references. Nevertheless, the question of the interpretation of the metric in terms of validation purpose is still open.

E. Interpretation of the adapted metrics

There is a real conceptual difficulty in interpreting the adapted metrics. Indeed, whereas the initial metrics were defined with a “common” reference in terms of state of the component (end-of-life), the adaptation proposed in this paper, although giving a way to use runs-to-maintenance for performance assessment purpose, uses time references that are different in terms of states of the component, as can be seen in Figure 3 and explained in section B. For simplicity reason, we will limit the explanation to the prognostic horizon metric.

Let $\Pi = \{\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(n)}\}$ be a set of runs-to-maintenance on which the prognostic horizons $(PH^{(1)}, \dots, PH^{(n)})$ are computed. In the case where the run-to-maintenance is in fact a run-to-failure which was not predicted by the PHM function, it should be discarded as it will be taken into account by another performance metric such as the missed estimation rate (number of missed detections over the total number of runs).

To ensure the validity of the prediction, we want the prognostic horizon to be greater than a limit value \overline{PH} . This limit is set by maintenance operations constraints (time to detect, analyze and schedule the maintenance action). This limit has a meaning when the maintenance is triggered by the health state but no meaning at all when maintenance is triggered in an intermediate health state by other triggers. Different mathematical solutions exist to filter the prognostic horizons that matter the most. Two of these solutions are given below.

- Consider the runs-to-maintenance for which the final health state is sufficiently close to the failure threshold by defining a subset of runs-to-maintenance Π_α ($0 < \alpha < 1$), as expressed by Equation 3.

$$\mathcal{H}^{(k)} = (h_i^{(k)}, t_i^{(k)})_{i=0..m^{(k)}} \in \Pi_\alpha \Leftrightarrow h_m^{(k)} > \alpha \bar{h} \quad (3)$$

- Build a weighted indicator which puts more importance into runs-to-maintenance that are closer to runs-to-failure which is noted \overline{PH} as in Equation 4.

$$\overline{PH} = \sum_{k=1}^n \left(\frac{h_m^{(k)}}{\sum_{i=1}^n h_m^{(i)}} \right) PH^{(k)} \quad (4)$$

In both propositions, the indicators are more comparable to the end-of-life ones and thus allow triggering an update process when the values are under the threshold value.

V. PRINCIPLES FOR THE UPDATE PROCESS OF A PHM FUNCTION

In the previous section, a focus has been made on prognostic performance metrics. These metrics allow detecting a potential drift in the performances of the PHM function. Other metrics could also be taken into account as the missed estimation rate that was already mentioned in the previous section for instance. The relevant metrics of interest mainly depend on the usage that is made of PHM outputs, so we will not develop them further.

Before entry into service, the PHM function has been designed as best as possible, with the available data and knowledge. At the entry into service, the PHM function is validated so the different performance metrics used meet the requirements. In parallel, a monitoring of the performance of the PHM function is also achieved. After each run-to-maintenance of the components monitored in the aircrafts, the different metrics are computed and eventually, a drift in terms of performance of the PHM function is detected. These changes will lead to a redesign as shown in Figure 5.

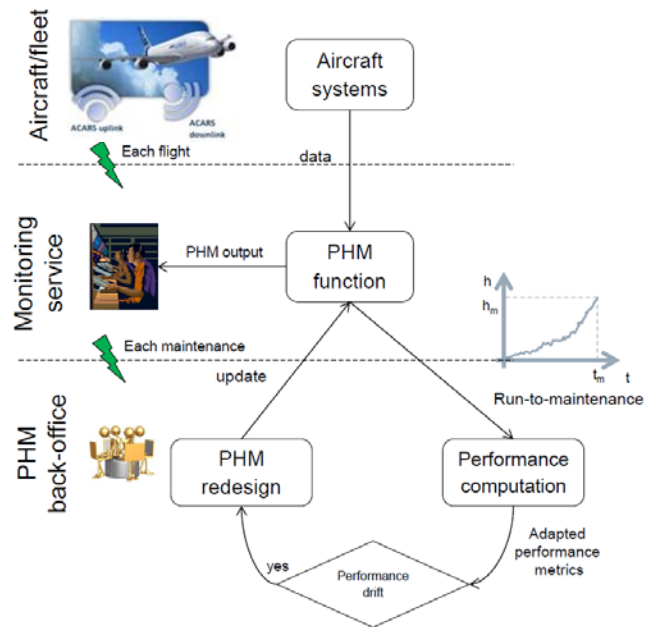


Figure 5: Performance monitoring of the PHM function

Once a drift is detected in the performance of the PHM function, a redesign of the function will be launched.

There are many challenges in the redesign of the PHM function. This paper will not solve them but, at least, it lists a series of issues to address in order to make the redesign as efficiently as possible. Of course, there is always the option of starting a completely new design process of the PHM function from scratch, discarding everything that was done in the

beginning, but this option is not conceivable in an industrial context.

The first issue is to be able to characterize on which part of the fleet the redesign is necessary. It is possible that the global performance of the PHM function decreases but only because the prediction is very bad on a part of the fleet and still good on another part. In this case, two main options could be tested:

- each aircraft could be defined by a set of parameters (operator, age, list of installed modifications, regions where the aircraft is operated, etc.) and a classification technique could help categorizing the aircrafts for which the PHM function is not good anymore;
- the list of potential evolutions described in Section III could be tracked and each evolution could be isolated to study its consequence on the performance of the PHM function.

Once the part of the fleet concerned by the redesign of the PHM function is defined, the second issue is to understand what can be kept from the former PHM function and what shall be discarded. The simplest way is to redo completely the design process of the PHM function, challenging each choice, each assumption that was made in the light of the configuration, operational and environmental conditions and age of the subset of the considered fleet. Is it possible to be smarter and understand *a priori* which part of the modeling process has to be redone, leaving the rest unchanged? In a way, a troubleshooting procedure has to be designed to understand which part of the modeling should be redone.

Some choices or assumptions could have a strong impact on the amount of work of the redesign. These choices concern:

- selection of failure and degradation modes;
- definition and validity of the health state.

Another key question concerns the health and prognostic assessment technique used. In practice, it could be very costly to develop a completely new technique because the one chosen in the design was not capable of dealing with the evolution of the fleet. This concern should be taken into account for the initial choice of methods. Moreover, the adaptability of the method to a different context could be a major selecting criterion for the technique to be used for health and prognostic assessment in order to avoid their complete new development.

Ideally, the evolution of the fleet should just require redoing a fine tuning of the algorithms adapted to the subset of the fleet for which the initial tuning is not good.

VI. CONCLUSION

This paper presented the continuous validation necessary to perform in order to ensure the sustainability of the performance of the PHM function. In the case of a monitoring service on a fleet of aircraft, the list of potential causes of evolution were listed, from the evolution of the composition of the fleet itself, introducing new operational and environmental

conditions or adapted maintenance operations, up to the evolution of the configuration of the newest aircraft. In order to measure the effects of the evolution of the fleet, a reference to literature was made to list a series of relevant performance metrics. A focus was made on prognostic algorithms performance metrics which usually use runs-to-failure when, in operations, the monitoring service is designed to avoid the failures. An adaptation of the metrics was proposed to take into account runs-to-maintenance instead of runs-to-failure. The principles of the update process were finally briefly presented, opening questions on the redesign of the PHM function.

This paper is really an overview of the questions underlying the sustainability of the PHM function performance. In a way, it stresses out the difficulties to make the diagnostics and health assessment of the PHM function itself:

- when the PHM function is not performing sufficiently well, what shall be changed to recover the required performance? (the questions is similar to a troubleshooting procedure);
- what “health indicator” allows measuring the “health” of the PHM function?

Moreover, the update process raises some open questions too:

- do the available metrics are enough to capture potential drifts?
- on which part of the fleet the redesign should be applied?
- how to ensure that the modeling and PHM techniques chosen during the initial design of the PHM function will be robust to potential evolutions of the fleet?

The answers to these questions will allow building relevant and valid PHM functions and maintaining the quality of the monitoring service. Nevertheless, most of these questions are still open.

REFERENCES

- [1] Andrew. K.S. Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7): 1483 – 1510, 2006.
- [2] A. Heng, S. Zhang, Andy C.C. Tan, and J. Mathew. Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23(3): 724 – 739, 2009.
- [3] K. Medjaher, D.A. Tobon-Mejia, and N Zerhouni. Remaining useful life estimation of critical components with application to bearings. *IEEE Transactions on Reliability*, pages 292–302, 2012.
- [4] J.Z. Sikorska, M Hodkiewicz, and L Ma. Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*, 25(5): 1803–1836, 2011.
- [5] Scott Kramer and Irem Y. Tumer. Towards StateCharts Based Failure Propagation Analysis for Designing Embedded PHM Systems, Annual Conference of the Prognostics and Health Management Society, San Diego, CA, USA, 2009.
- [6] Taoufik Jazouli and Peter Sandborn. A Design for Availability Approach for Use with PHM, Annual Conference of the Prognostics and Health Management Society, Portland, Oregon, USA, 2010.
- [7] Christian Modest and Frank Thielecke. A Design Methodology of Optimized Diagnosis Functions for High Lift Actuation Systems,

- Annual Conference of the Prognostics and Health Management Society, Minneapolis, Minnesota, USA, 2012.
- [8] Brian Bole, Christopher Teubert, Cuong Chi Quach, Edward Hogge, Sixto Vazquez, Kai Goebel and George Vachtsevanos. SIL/HIL Replication of Electric Aircraft Powertrain Dynamics and Inner-Loop Control for V&V of System Health Management Routines, Annual Conference of the Prognostics and Health Management Society, New Orleans, USA, 2013.
- [9] N. Scott Clements and David S. Bodden. Prognostic Algorithm Verification, Annual Conference of the Prognostics and Health Management Society, New Orleans, USA, 2013.
- [10] Johann Schumann, Vanesa Gomez-Gonzalez, Nagabhushan Mahadevan, Michael Lowry, Peter Robinson and Gabor Karsai. A Tool Chain for the V&V of NASA Cryogenic Fuel Loading Health Management, Annual Conference of the Prognostics and Health Management Society, Fort Worth, TX, USA, 2014.
- [11] Saxena, A., Celaya, J., Saha, B., Saha, S., Goebel, K., (2010) Metrics for Offline Evaluation of Prognostics Performance, International Journal of Prognostics and Health Management (IJPHM), vol. 1(1), 2010.
- [12] Wheeler, K., Kurtoglu, T., Poll, S., A Survey of Health Management User Objectives Related Diagnostic and Prognostic Metrics, International Journal of Prognostics and Health Management (IJPHM), vol. 1(1), 2010.