Proceedings of the 6th European Conference of the Prognostics and Health Management Society 2021 - ISBN – 978-1-936263-34-9

Age and Condition-Based Preventive Replacement Timing for Periodic Aircraft Maintenance Checks

Floris C. Freeman[1], Paul J. van Kessel[1], Wim J.C. Verhagen[2]

[1] *KLM Royal Dutch Airlines, Schiphol, 1117 ZL, The Netherlands*
floris.freeman@klm.com
paul-van.kessel@klm.com

[2] *RMIT University, Melbourne, 3053 VIC, Australia*
wim.verhagen@rmit.edu.au

Abstract

By anticipating impending failures and addressing those with preventive replacements, Condition-Based Maintenance (CBM) can provide several economic benefits for aircraft maintenance. While Prognostics and Health Monitoring (PHM) methods for CBM are widely available, scheduling of tasks originating from those methods is a relatively new challenge. To avoid incurring extra cost and ground-time, these maintenance tasks are typically scheduled during already existing (conventional) maintenance checks such as periodic checks. Following this strategy, an aircraft component would have to be replaced if a failure precursor is detected that is expected to result in failure between the next two checks. The decision following from that detection depends on the chosen decision threshold of the prognostic or diagnostic model, which can be represented by a point on the Receiver Operating Characteristic curve. Selecting the optimal operating point for each maintenance check is challenging, as - depending on the component in question - one or more of the following factors may be in play: age-dependent reliability of the component, the performance of the prognostic (diagnostic) model, the interval of the periodic check and the cost of (corrective and preventive) maintenance. This paper presents an innovative method for selecting optimal operating points for all periodic checks throughout the lifetime of an aircraft component. This is done by means of a numerical optimization model that finds operating points that minimize the component's total maintenance cost per flight hour. A case study on a compressor of a wide-body aircraft is presented, which shows that by using this method, additional economic value from existing PHM can be realized without the need for additional investments.

1. Introduction

Modern aircraft today make use of thousands of on-board recordable parameters for safe and efficient flight operations. The increasing availability of these parameters to aircraft operators and new methods to process sensor data effectively has made Prognostics and Health Monitoring (PHM) a promising enabler for Condition-Based Maintenance (CBM) (Hölzel, Schilling, & Gollnick, 2014). In aircraft maintenance, the most suitable maintenance policy for each aircraft part or subsystem depends on the probability, impact and detectability of each potential failure. Impact of failures that are critical to safety, economics or operations are adressed by (periodic) preventive maintenance tasks. Failures that are not critical to safety, economics or operations are adressed by corrective maintenance tasks. The full methodology to design a maintenance program is described in MSG-3 (ATA, 2013). For preventive maintenance tasks, PHM offers a potentially more efficient future alternative, by shifting from time-based to condition-based maintenance (IMRBPB, 2018). For failures that are adressed by corrective maintenance (the scope of this paper), PHM already offers a way to avoid operational disturbances and reduce maintenance cost (Sun, Zeng, Kang, & Pecht, 2010).

Maintenance tasks to mitigate critical failures are mandatory and have fixed due-dates. These tasks are often packaged into blocks of tasks and executed in periodic maintenance blocks, known as letter checks (e.g. A- and C-checks). PHM-driven tasks are not imposed by regulations, and are optional. These tasks are less common in aircraft operations, but they introduce a new challenge because the operator must decide if and when to schedule the task, taking into account the expected economical benefit under prognostic uncertainty (Engel, Gilmartin, Bongort, & Hess, 2000).

Scheduling of prognostic driven tasks can be approached in two ways. Either the task is executed in a dedicated maintenance check (the task is *pushed* into the airline schedule), or the tasks can be executed during an already existing slot such as a periodic letter check (the task is *pulled* into the airline schedule). Whether *pushing* or *pulling* is preferred is situation specific and may depend on several factors such as the availability of future maintenance opportunities, resource availability and competing demands from other aircraft. As multiple aircraft in the fleet compete for the same maintenance checks, pulling instead of pushing a PHM driven task is often more efficient. This is because it is relatively inexpensive to execute the task parallel to existing tasks (no extra ground time is required). In addition, the cost incurred by pulling a task into a schedule is easier to estimate and manage than the cost for pushing it into the schedule. The underlying reason is that airline schedules are highly dynamic and it is therefore difficult to predict which new maintenance opportunities are available in the future, and at what cost. As a consequence, a PHM practitioner may prefer to focus on pulling prognostic driven tasks into an existing maintenance check where possible, before considering assigning a new dedicated maintenance check. While some methods exist for *pushing* PHM driven tasks into an aircraft schedule (Vianna & Yoneyama, 2018), (Sandborn & Wilkinson, 2007), methodologies for *pulling* PHM driven tasks into existing checks are scarce in literature.

The decision whether or not to pull a prognostic-driven task into an existing maintenance check depends on the selected decision threshold of the PHM model, also known as operating point. Put simply, the component is replaced only if the output from the PHM model exceeds a pre-selected threshold. Selecting this point is a compromise between expected frequencies of true -and false positives, which can be visualized by the Receiver Operating Characteristic (ROC) curve of the PHM model. Generic considerations for selecting an operating point from the ROC curve has been proposed by Metz (1978), and illustrated for a use-case in aircraft engine maintenance by Koops (2018). Both studies provide useful insights but have two limitations for practical applicability. Firstly, only examples to minimize cost per (maintenance) event are given, whereas cost- minimization on aircraft lifecycle basis could be more relevant for an aircraft operator. As a consequence, the (opportunity) cost of wasting remaining useful life in case of a false-positive is not sufficiently accounted for. Secondly, while both authors show that the optimal operating point depends on the component's (age dependent) failure probability, they do not provide a way to dynamically select the optimal operating point given the component's time-varying failure rate. This last point touches upon a third, more fundamental problem. Age-dependent failure probability is typically known to the operator and is often used to time preventive maintenance actions (Letot, Equeter, Dutoit, & Dehombreux, 2017). However, the majority of PHM models found in recent literature rely only on on-board monitoring data, and do not use age as a prognostic parameter (Jia, Huang, Feng, Cai, & Lee, 2018), (Jardine, Lin, & Banjevic, 2006), (Lei et al., 2018). Arguably, age as a prognostic parameter only works if there is 1) a clear age-reliability relationship that can be established for the sample, and 2) individual components behave in accordance with the overall sample, while displaying a (relatively) small amount of variance with respect to the overall sample behaviour. If this is indeed the case, then the omission of age as a parameter leaves potential useful data for timing preventive replacements unexploited. It has to be noted that age is not the only factor at play in driving reliability behavior over time; the influence of operating and environmental conditions has been established in prior work (Thijssens & Verhagen, 2020), though its considered beyond the scope of this research.

In this paper, we aim to remove these limitations by minimizing total maintenance cost per flight hour, which is done by using component-reliability data in addition to maintenance costs and performance indicators of the considered PHM model. Component reliability data is used to estimate the life expectancy of a component as well as the age-dependent failure rate, while prognostic performance indicators are used to estimate the probability of PHM correctly detecting the failure precursor. Subsequently, total maintenance cost per flight hour are minimized by selecting the optimal operating point on the ROC curve for each periodic maintenance check throughout the lifetime of the aircraft component. With these operating points set, both component age and PHM output can be used to decide whether or not to preventively replace and service an aircraft component during periodic maintenance checks.

The next section in this paper discusses methods and importance of optimally timing preventive maintenance actions. The section ends with an overview of relevant PHM metrics for this research. Thereafter in Section 3, we explain our proposed method and derive a function for total maintenance cost per flight hour. Lastly in Section 4, the economic potential of the method is illustrated by means of an example from a compressor in a wide body aircraft, based on actual aircraft maintenance data and an arbitrary ROC curve. For this use-case, we have used a genetic algorithm to select optimal operating points for each periodic maintenance check. This case study shows an example were a reduction in maintenance cost of at least 6% can be achieved in comparison to traditional PHM maintenance policies, without the need for additional investments.

2. THEORETICAL CONTEXT

Preventive maintenance aims to perform maintenance before the failure of a system occurs. For non-critical failures (the scope of this study), preventive maintenance could deliver value when the cost of a preventive maintenance action before failure (C_p) is lower than the cost of a corrective maintenance action after failure (C_c). This could be the case when preventive maintenance results in less cascaded damage (cheaper repairs), better predictability of maintenance demand (lower logistic cost), faster troubleshooting (lower personnel cost), etc. Considering aircraft operations, additional benefits from preventive maintenance can be found in increased operational reliability and fleet availability. Regarding operational reliability, preventive maintenance can avoid an unexpected failure that would otherwise reduce immediate or future aircraft dispatch reliability. When these preventive tasks are scheduled in parallel, the airline may need fewer maintenance buffers and spare aircraft to mitigate the risk of operational disturbances. This could lead to an increase of the availability of the fleet.

2.1. Timing of preventive maintenance

For replaceable aircraft components, corrective maintenance often leads to longer operating times between replacements than preventive maintenance does. Hence when preventive maintenance is not correctly timed, the more frequent maintenance actions may diminish the benefits of cheaper repairs. This is especially the case when the cost for preventive repair is close to the cost of corrective repair. This issue stipulates the importance of accurately timing predictive maintenance tasks. As described by Ben-Daya, Kumar, and Murthy (2016), timing of preventive maintenance can be approached in several ways. Two of these ways (age-based and condition-based) will be illustrated with an example below.

2.1.1. Age-based preventive maintenance

In age-based approaches for non-safety critical systems, the probability of failure as a function of component's age may be derived from historical time-to-failure data. A distribution $f(t)$ can be fit on these data to describe the probability that a component fails at a certain age. Applied to aircraft operations, this age is usually expressed in Flight Hours (FHs) or Flight Cycles (FCs). From such a probability function, one can derive the survival function $S(t)$. The survival function (Eq. (1)) describes the probability of surviving past age t, or in other words, it describes the proportion of the original population that survives past age t. Finally, this survival function can be used to calculate the average life expectancy of a population till time t, see Eq. (2).

$$S(t) = 1 - \int_0^t f(s)\,ds \tag{1}$$

$$A_{exp} = \int_0^t S(s)\,ds \tag{2}$$

When a preventive replacement is planned at age t (which can be seen as an *economic hard-time* and is sometimes referred to as t_p in literature), a fraction of $[1-S(t)]$ of the population will have failed before t and is correctively maintained at cost C_c. A fraction of $S(t)$ is still operational at age t and is preventively maintained at cost C_p. Hence the total maintenance cost per flight hour (*CPFH*) is equal to the total expected cost divided by the expected life expectancy of the population, see Eq. (3).

$$CPFH = \frac{C_p S(t) + C_c \cdot [1 - S(t)]}{\int_0^t S(s)\,ds} \tag{3}$$

The optimal replacement time can be found by finding age t that minimizes *CPFH* (Eq. (3)). With t found, the aircraft operator can schedule a new maintenance check (*push*) or add the replacement-task to an existing maintenance check (*pull*) close to t. For further reading on optimal age replacement policy, see (Letot et al., 2017).

Although age-based preventive maintenance can reduce maintenance cost compared to corrective maintenance, it does not provide insights into which specific component will fail at what time. The introduction of on-board sensing technologies changes this. With continuous monitoring data on individual component level, the (future) condition of each component can be estimated, and preventive maintenance can be scheduled for specific components accordingly.

2.1.2. Condition-based preventive maintenance

In PHM enabled condition-based maintenance, a detection of a fault (diagnostics) triggers a maintenance action either directly, or after prediction of remaining useful life (prognostics) (Saxena et al., 2008). In both cases, the aircraft operator uses PHM output to decide whether and when to schedule (*push* or *pull*) a maintenance action. The decision to *pull* a PHM driven task into an existing maintenance check (the scope of this paper) can be challenging, as the optimal timing of these tasks depends not only on the PHM output and the cost-difference between C_p and C_c, but also on the underlying failure rate and the expected performance of the PHM model at use (Engel et al., 2000).

2.2. Performance indicators

PHM performance metrics inform a decision-maker of estimates of (un)certainty around the PHM output. A comprehensive summary of prognostic performance metrics can be found in literature (Saxena et al., 2008). For the scope of this study, the characteristic *trajectory* of the PHM output and the Receiver Operating Characteristic curve are of interest.

2.2.1. Trajectory

Consider a trajectory of PHM output over time. This output could be a health-index, an anomaly score, a remaining-useful-life estimate, etc. Let F be the time index for the time when the fault occurs. Depending on the sensitivity of the on-board sensing system, the time of detecting of the fault (D) will be somewhere after F. Furthermore, let EOL be the time index for the *end-of-life* (failure) of the aircraft component. The first prognostic metric we use in this study is the horizon h, which is the time-difference between fault detection D and the EOL. Note that h depends on the progression speed of the fault, and the sensitivity of the PHM technology, see Figure 1.

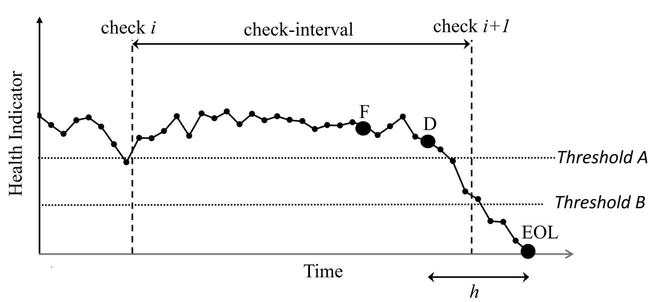


Figure 1. Generic trajectory of a health indicator. The horizon h is defined as the time difference between fault detection D and *end-of-life* (EOL). Thresholds A and B represent arbitrary decision thresholds.

2.2.2. Receiver operating characteristic curve

In this study, the goal is to optimally plan (*pull*) PHM driven tasks into existing maintenance checks such that the total maintenance cost per flight hour is minimized. With that goal in mind, the prognostic requirement is equivalent to the diagnostic problem of detecting a failure precursor just prior to the next maintenance check, which is expected to result in failure between the next two checks. Only if that failure precursor is found (with some confidence level), the component is replaced and maintained. For this type of prognostic tasks, similar metrics as in diagnostics may be used (Saxena et al., 2008). Hence, we use the True Positive Rate (TPR) as a metric of how sensitive the PHM model is in detecting the fault that will lead to failure before the end of the prognostic horizon h. Likewise, we can use the True Negative Rate (TNR), the False Negative Rate (FNR) and the False Positive Rate (FPR). Note that the FNR and TNR can be expressed as functions of TPR and FPR respectively, see Eqs. (4) and (5).

$$FNR = 1 - TPR \tag{4}$$

$$TNR = 1 - FPR \tag{5}$$

Only if the output of the PHM model exceeds a certain threshold just prior to the next maintenance check, the component will be replaced in that maintenance check. In Figure 1, two examples of a decision threshold are given. With threshold A, the anomaly just prior to maintenance check (i) would have resulted in a false positive for check (i). However, the same threshold would have resulted in a true positive for check $i + 1$. Likewise, if threshold *B* would have been selected, it would have resulted in a true negative for check (i), but also in a false negative for check ($i + 1$). Generally speaking, each possible threshold results in a unique combination of expected *FPR*s and *TPR*s. The set of all available combinations can be presented by a curve, known as the Receiver Operating Characteristic (ROC) curve (Metz, 1978), (see Figure 2). Selecting the operating point for the PHM model involves a trade-off between the *FPR* and the *TPR*. With a too low decision threshold, few failures will be missed but the many false alarms (*false positives*) will diminish the benefit from cheaper repairs. With a too high decision threshold, fewer false-alarms will be raised, but many failures will be missed (*false negatives*).

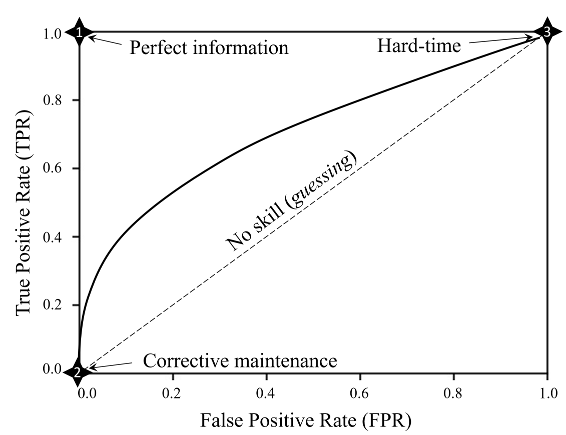


Figure 2. Arbitrary ROC curve, with operating points indicated for perfect information, corrective maintenance and hard-time.

A couple of operating points on the ROC curve are of special interest. As shown in the graph, an ideal PHM model would have perfect information about the component's condition, and would always provide true positives without false-positives, regardless of the decision threshold ($FPR = 0, TPR = 1$). A PHM algorithm with no skill (*guessing*) operates at the diagonal of the ROC curve. The origin of the ROC curve (at

$TPR = FPR = 0$) is equivalent to maintaining the component only correctively. The right upper corner (at $TPR = FPR = 1$) is equivalent to replacing the components preventively regardless of its condition (e.g. on *hard-time*). The actual curve of a feasible and potentially beneficial PHM models is a monotonically increasing line with decreasing gradient, that runs above the line of 'no skill', from original to the point where *TPR=FPR=1* (Metz, 1978).

According to Metz (1978), the ideal operating point comes from a business decision that depends on the underlying failure rate of the population of interest. When failure is rare, the decision maker should operate toward the lower left portion of the ROC curve. Conversely when failure is common, the best operating point is toward the upper right part of the ROC curve. Koops (2018) illustrates this point with a use-case on aircraft engine monitoring, but while recognizing that failure rate is age-dependent, the study assumes a fixed failure rate. As mentioned in the introduction, a method to use dynamic age-based reliability data and output from existing PHM to pull a preventive tasks into an existing maintenance check is currently lacking. The following Chapter presents an innovative approach for determining the most cost efficient maintenance strategy along the lifetime of the component.

3. METHODOLOGY

In this section, we first derive a cost function for the total maintenance cost per flight hour (*CPFH*) for an PHM-enabled replaceable aircraft component given selected operating points on the ROC curve. Finally, we minimize *CPFH* by optimally selecting operating points for each check. We make two assumptions for simplification purposes:

1. The performance of the PHM model is independent of the component's age and independent of the component's time to failure. In other words, we assume that the performance of the PHM model (in this case the ROC curve) is constant over time. This is based on an underlying assumption of similar behavior across components due to similar utilization characteristics (including operational and environmental exposure) across a fleet of aircraft as typically operated by commercial airliners. However, these assumptions can be challenged in at least two ways: i) in reality, the variability across component behavior may be sufficiently large to meaningfully change the ROC curve over time; ii) in reality, the constant addition of data over time (e.g. additional input data, additional failure occurrences) as time progresses may lead to increased model performance, leading to changes in the ROC curve over time.
2. False positives have no consequences on the life expectancy of the remaining population. This assumption holds when false-positives occur due to random anomalies in the PHM output. In other words, false positives do not occur due to imperfect estimation of the *timing* of the failure. For this assumption to hold, the interval of the check must be larger than the horizon h.

By choosing a point along the ROC curve for a specific check, a two-alternative decision (replace or not replace) is made based on the PHM output. The expected cost resulting from that decision can be calculated as (Metz, 1978):

$$\begin{aligned} C =& C_o + C_{TP} \cdot p(TP) + C_{TN} \cdot p(TN) \\ &+ C_{FP} \cdot p(FP) + C_{FN} \cdot p(FN) \end{aligned} \tag{6}$$

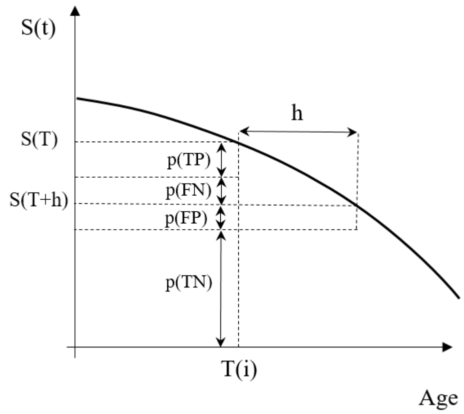
Here, C_0 is the overhead cost of operating the PHM technology, and C_{TP}, C_{TN}, C_{FP} and C_{FN} are the cost resulting from a true positive, true negative, false positive and false negative decision respectively. Probabilities p describe the probability of each of the four outcomes. For example, $p(TP)$ is the probability that a true positive decision is made. This probability is equal to the probability that a unit from the population will fail within a certain time-horizon, multiplied by the probability that an actually faulty component is identified as such by the PHM technology at use.

Let us consider a pool (*population*) of identical replaceable components with survival curve $S(t)$ for which PHM technology is available. Next consider an ROC curve, (Figure 2), describing the performance of the PHM technology. Finally, consider a maintenance policy with periodic maintenance checks of a certain interval. Let *i* be a counter that ranges from 1 for the first periodic check after the component was installed to *N* for the last letter check after failure of the last component in the population. A function $T(i)$ maps the sequence of letter checks *i* to the corresponding age *T* of the population. The probability that a unit from the population fails between age T and the end of the prediction horizon h is equal to $S(T) - S(T + h)$. The probability that an actually faulty component is identified as such by the PHM technology is equal to TPR. Similarly, all four probabilities in Eq. (6) can be defined as:

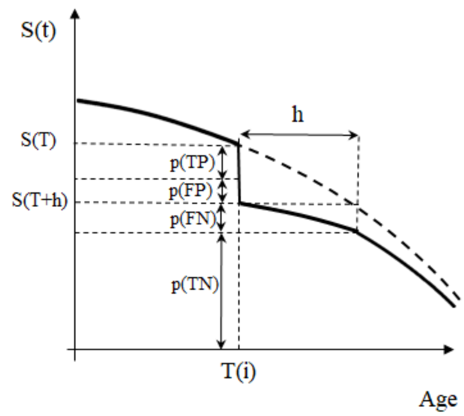
$$\begin{aligned} p(TP) &= [S(T) - S(T + h)] \cdot TPR \\ p(FN) &= [S(T) - S(T + h)] \cdot FNR \\ p(TN) &= S(T + h) \cdot TNR \\ p(FP) &= S(T + h) \cdot FPR \end{aligned} \tag{7}$$

These probabilities are graphically illustrated in Figure 3a.

As in similar work on this topic (Koops, 2018), true negative predictions are not associated with specific actions or cost, so we set C_{tn} equal to zero. Constant C_0 is irrelevant for cost



(a) Generic survival curve, with indicated impact of chosen operating point along ROC with horizon h.



(b) Generic survival curve (dashed), and altered survival curve (solid) due to a selected operating point.

Figure 3. Effect of choosing a point along the ROC on the survival curve due to TP, TN, FP and FN. Figure 3a displays the original survival curve and Figure 3b shows the adjusted surivival curve due to preventive replacements (TP & FP).

minimization and is also set to zero ($C_0 = 0$). Finally, C_{fn} is equal to the cost of maintaining the component after failure ($C_{fn} = C_c$). Hence, by substituting Eqs. (4,5,7) into Eq. (6), the expected cost of a replacement decision based on PHM output for a maintenance check (i) at time $T(i)$ becomes:

$$\begin{aligned} C_{t=T} &= C_{tp} \cdot [S(T) - S(T+h)] \cdot TPR \\ &+ C_{fp} \cdot S(T+h) \cdot FPR \\ &+ C_c \cdot [S(T) - S(T+h)] \cdot [1 - TPR] \end{aligned} \tag{8}$$

Based on the values chosen for TPR and FNR, the survival curve changes shape as a result of the replacements during the maintenance check. Both true positive and false positive rates cause a gap in the survival curve S at $t = T(i)$. Secondly, due to prognostic preventive replacements, less failures will take place between $T(i)$ and $T(i) + h$ as only the false negative part of the population remains.

The complete change in the survival curve during the decision made just prior a to maintenance check (i) at time $T(i)$ is subdivided in three regions:

1. Before the maintenance check: $t \leq T(i)$
2. After the maintenance check, before the end of the prognostic horizon: $T(i) \leq t \leq T(i) + h$
3. After the end of the prognostic horizon: $t \geq T(i) + h$

Below, the term S^* is used to indicate the adjusted version of the original survival curve S. The change in the survival curve is graphically illustrated in Figure 3b.The mathematical formulation for the adjusted survival curve can be formulated for the three regions as following:

$$S^*(t) = \begin{cases} S^*(t) & \text{for } t < T(i) \\ S^*(t) - \textit{p(TP)} - \textit{p(FP)} - p^*\textit{(FN)} & \text{for T(i)} \leq t \leq T(i) + h \\ S^*(t) - \textit{p(FP)} & \text{for t} > T(i) + h \end{cases} \tag{9}$$

As can be seen in Figure 3b the survival curve between $T(i)$ and $T(i) + h$ is flattened due to the preventive removal of faulty components (TP). We assume that the remaining corrective replacements are spread evenly between $T(i)$ and $T(i) + h$. It should be noted here that the preventive replacements are note taken into account to update the survival probability under a 'good-as-new' assumption. The reduction of the survival probability due to FNs between $T \leq t \leq T + h$ can be formulated as following:

$$p^*(FN) = (S^*(t) - S^*(T)) \cdot \textit{TPR} \tag{10}$$

By substituting Eq. (10) into Eq.(7) and subsequently Eq. (7) into Eq. (9), the change in the survival curve can be formulated in terms of a chosen point along the ROC curve. The equation can be further simplified by substituting Eq. (4) to eliminate the FNR term. The formulation for the adjustment in S^* then becomes as following:

$$S^*(t) = \begin{cases} S^*(t) & \text{for t} < T(i) \\ S^*(t) \cdot (1 - \textit{TPR}) + & \\ S^*(T(i) + h) \cdot \textit{(TPR-FPR)} & \text{for } T(i) \leq t \leq T(i) + h \\ S^*(t) - \textit{FPR} \cdot S(T(i) + h) & \text{for t} > T(i) + h \end{cases} \tag{11}$$

Maintenance decisions to replace a component are not taken only once; there is a decision to be made for each periodic maintenance check in the lifetime of the component. The component's age (hence its reliability) is different in each check. Therefore, different points along the ROC curve may be selected given the age of the component. For each periodic check (i), the selected operating point is represented by [FPR_i , TPR_i]. The altered survival curve S^* is obtained iteratively, by continuous adjustments after each periodic maintenance check using Eq. (11). For a graphical representation of establishing S^*, see Figure 4.

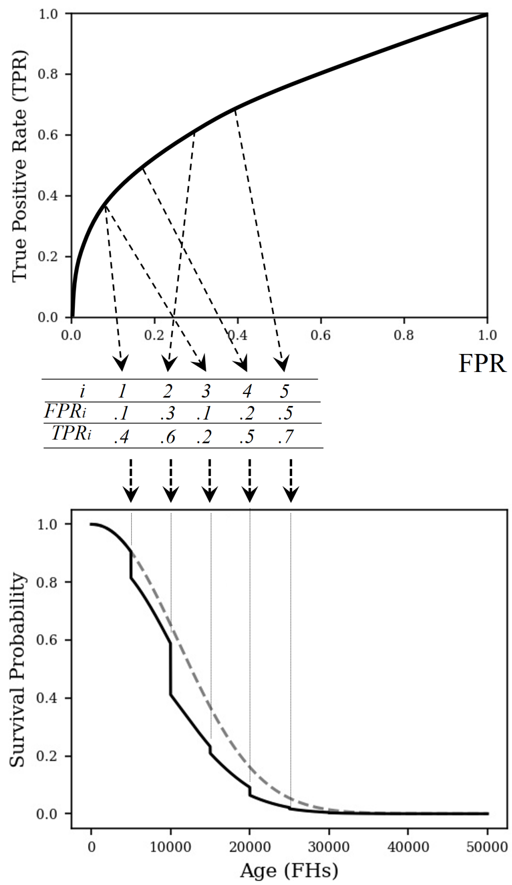


Figure 4. Graphical illustration of modifying the survival curve due to selected operating points during maintenance chekcs. Top graph: ROC curve. Middle Table: selected operating points from the ROC curve for each check (*i*). Bottom: component's altered survival curve.

As the overall goal is to minimize the component cost per flight hour, maintenance cost must be calculated over the entire lifespan of the population. Hence to calculate the maintenance cost over the expected lifetime of the component, we sum Eq. (8) over all periodic maintenance checks:

$$\begin{aligned} C_{total} = \sum_{i=1}^{N} & C_{tp} \cdot [S^*(T(i)) - S^*(T(i)+h)] \cdot TPR_i \\ & + C_{fp} \cdot S^*(T(i)+h) \cdot FPR_i \\ & + C_c \cdot [S^*(T(i)) - S^*(T(i)+h)] \cdot [1 - TPR_i] \\ & + C_c \cdot [S^*(T(i)+h) - S^*(T(i+1)] \end{aligned} \quad (12)$$

Note that Eq. (12) has one more term than Eq. (8), since corrective cost is also incurred by components that failed beyond the prediction horizon h, but before the next maintenance check $(i + 1)$.

As mentioned in the introduction of this paper, the relevant cost metric to be minimized is total maintenance cost per flight-hour ($CPFH$). To get there, Eq. (12) is divided by the expected life expectancy of the population (Eq. (2)):

$$CPFH = \frac{C_{total}}{\int_{t=0}^{\infty} S^*(t)\, ds} \quad (13)$$

As can be seen from Eqs. (11-13), the values for FPR_i and TPR_i determine the expected total maintenance cost as well as the life expectancy of the component. The optimization task is to find for each maintenance check (i), the PHM probability thresholds (represented by values for [FPR_i,TPR_i] on the ROC curve) that minimizes total maintenance cost per flight-hour (Eq. (13)). Selecting the type of solver to find optimal values for FPR_i and TPR_i is beyond the scope of this study. In the next section, we present an example with a genetic algorithm, but better alternatives may be available, depending on the size of the optimisation problem. With FPR_i and TPR_i found once for all checks (*i=1,2,... N*), the operator knows for each check which PHM decision threshold to select given the component's age.

4. CASE STUDY

The goal of this case-study is to demonstrate that additional savings from PHM can be realized when optimal operating points from the ROC curve are dynamically chosen for each periodic check throughout the lifetime of a replaceable aircraft component. For this case study, we consider a compressor in the environmental system of a modern wide body aircraft. Each aircraft has 4 compressors, which are all running (at equal power) during flight. Each compressor has 12 sensors installed that log parameters at a frequency of 1-4 Hz. A removal of the compressor takes up to 8 hours, and would therefore significantly impact the airline's flight schedule if not combined with existing maintenance checks. Failure data used in this case-study consists of removal and repair logs from multiple airlines.

The root cause of failure is known to be overheating of the compressor. When the internal heat of the compressor reaches a critical value, the air-bearing destabilizes, which results in damage that requires renewal of the compressor's rotor, bearings and housing. When overheating is detected prior to reaching the critical value, the subcomponents causing overheat can be removed preventively, without the need to renew the compressor's major parts. C_{fp} and C_{tp} in Eq. (12) represent preventive cost of maintaining the component before failure and are assumed to be equal ($C_{fp} = C_p$, $C_{tp} = C_p$). The costs for preventive and corrective repairs (C_p, C_c) are found by averaging recent repair cost for both preventive and corrective removals. These (normalised) averages were found to be $10k\$$ and $25k\$$ for C_p and C_c respectively. It must be noted that historical shop data contained variance in repair cost (standard deviation of $3k\$$ for preventive repairs and $9k\$$ for corrective repairs). Additional benefits from increased availability and operational reliability are expected, but are out of scope for this use-case.

Usage parameters between installation and failure of over 570 corrective removals were recorded, and about 910 compressors had not failed yet at the moment this case study was conducted. Correcting for these 910 right-censored samples, a maximum-likelihood estimation (Ferreira & Silva, 2017) was used to fit four parametric models; Weibull (*L=-6020*), lognormal (*L=-6160*), loglogistic(*L=-6070*), and exponential (*L=-6230*). See Figure 5 for QQ-plots for each distribution. A Weibull distribution was selected with shape parameter *2* (*se=0.1*) and a scale parameter of *15k FHs* (*se=200FH*), see Figure 6.

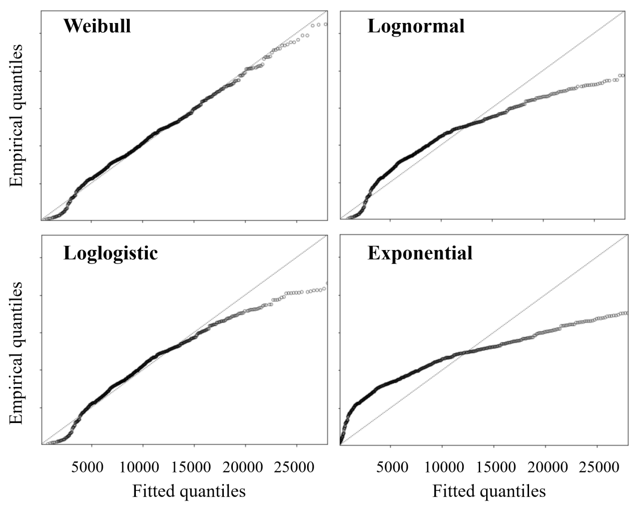


Figure 5. QQ-plots for fits of various failure distributions

For this use-case, one arbitrary PHM model with horizon h of 1000 Flight Hours (*FHs*) is chosen, based on experience with other prognostic models deployed at the airline. The corresponding ROC curve is visualized in Figure 7, with the data provided in the Appendix (Table 4). The considered periodic maintenance check is a 24 hours *A-check* with an interval of 1500 FHs. The 'population' of compressors is assumed to be extinct when less than 1% of the original population has survived. Doing so, this population is assumed to have $N = 26$ checks.

In the Genetic Algorithm (GA), initial values for TPR_i and FPR_i are populated by selecting randomly from points at the ROC curve for each of the 26 checks. For convenience, each value for TPR_i and FPR_i is represented by a (Nx2) matrix *M*, where the columns in *M* represent the TPR and FPR respectively, and the rows represent the checks in chronological order, see Table 1.

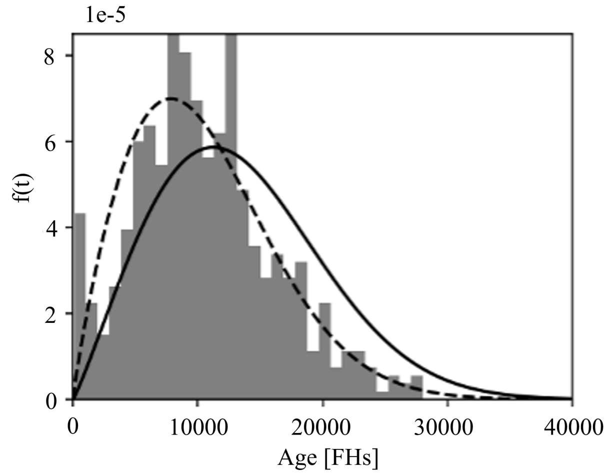


Figure 6. Histogram of time-to-failure data with fitted Weibull distributions. Dashed line is a Weibull fit on recorded time-to-failure data. Solid line accounts for censored data.

Table 1. Example of random M

i	FPR	TPR
1	0.15	0.68
2	0.35	0.86
...	...	...
26	0.85	0.985

The generation of random *Ms* is repeated until there are 500 initial versions of *M*. To limit the solution space of the GA, we make use of the fact that this component has a shape factor of > 1 and therefore has an increasing hazard rate. From the theoretical context, we know that the operating point on the ROC curve shifts to the right (higher $[FPR, TPR]$) when the underlying failure rate increases. Therefore, feasible solutions of *M* should be non-decreasing, so *M* is sorted in ascending order before being evaluated by the cost function from Eq. (13)

Table 2. Settings for genetic algorithm

Parameter	Value
Generation size	500
Number of iterations	40
Part of population mutating	50%
% of parents recombine	30%
% of M mutating	10%
Number of checks N	26
Weibull [scale, shape]	[1500, 2]
Periodic check interval	1500 FHs
Prediction horizon h	1000 FHs
Corrective cost C_c	25000$
Preventive cost C_p	10000$

The GA evaluates each version of *M* based on the maintenance cost per flight hour, as calculated by Eq. (13). The lower the total maintenance cost caused by a specific *M*, the likelier it is for that set of $FPRs$ and $TPRs$ in *M* to reproduce in the next generations of the genetic algorithm. In each iteration of the GA, the top 30% of the population recombines. Mutation occurs for 50% of all individual parents in each generation, by replacing 10% of the rows in *M* (selected randomly) by random rows from the ROC curve. We repeat the genetic algorithms for 40 iterations. Settings for the genetic algorithm (Table 2) have been found by grid-searching. Due to the coarse resolution of the ROC curve and the low number of maintenance checks in the lifetime of the compressor, the size of this optimisation problem is relatively small. For larger problems, different settings in the GA may needed.

After 40 iterations, we select the *M* that resulted in the lowest maintenance cost per flight hour. In Figure 7, the operating points for all checks are shown by indicating the check sequence number i at the corresponding locations of the ROC curve . The effect of the operating points on the population is shown in Figure 8 (solid line).

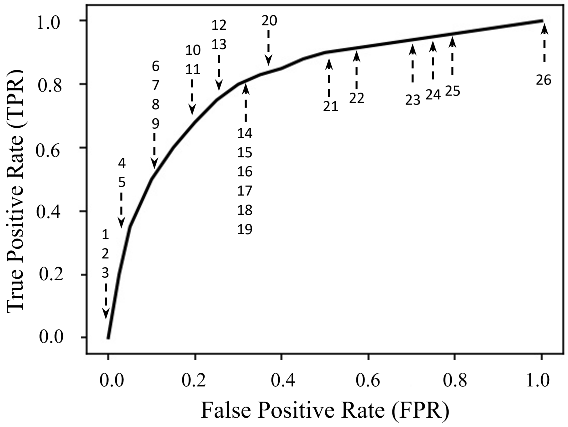


Figure 7. Operating points selected for each check. The numbers in the graph represent the check sequence numbers (*i*).

As can be seen in Figure 7 and Figure 8, the first 3 entries in *M* represent operating points at the origin of the ROC curve, where *FPR = TPR = 0*. This means that for the first 3 checks, one should not use the considered PHM model at all. After that, the operating points are at non-zero values of $[FPR_i, TPR_i]$, and moves to higher operating points on the ROC curve with increasing check sequence number i.

By applying the proposed method (named *dynamic thresholds* for this use-case), the lowest total maintenance cost is found to be $\$1.62$ per flight hour. Looking at the results presented in Table 3, applying dynamic decision thresholds results in the lowest cost per flight hour of all available maintenance policies. If only a single operating point on the ROC curve had been selected for all checks (*fixed thresholds*), the total maintenance cost per FH would have been at least 6% higher, depending on which threshold is chosen. Operating components at 6% lower cost is significant, especially for components that fail frequently and are expensive to repair. The table also shows a comparison with using non-PHM policies. By using the dynamic thresholds method, the cost per flight hour is 13.8% lower than for the current practice of corrective maintenance. The method of applying dynamic decision thresholds also proves to be more cost efficient than a hard-time by a benefit of $\$0,18$ per flight hour. Looking at the cost breakdown provided in Table 3, it can be seen that the benefit of applying dynamic decision thresholds is achieved by a combination of low corrective maintenance cost per FH and a relatively high expected lifetime. The table also provides the minimum achievable cost if a perfect PHM system would have been available (*perfect information*). In such a scenario, corrective cost is only incurred for the periods beyond the prognostic horizon h. All other failures are preventively mitigated at cost C_p.

Table 3. Comparison in operating cost for different strategies

Method	**Cost per FH**	**Corrective cost per FH**	**Preventive cost per FH**	**Expected lifetime**
Dynamic thresholds	\$1.62	\$1.10	\$0.52	10639 FH
Fixed thresholds	\$1.73	\$1.25	\$0.48	10357 FH
Hard-time	\$1.80	\$1.26	\$0.54	9879 FH
Corrective maintenance	\$1.88	\$1.88	\$0.00	13310 FH
Perfect information	\$1.16	\$0.65	\$0.51	12979 FH

For verification of the results, the various replacement strategies can be evaluated by analysing the (adjusted) survival curves. In Figure 8, the effect of using a fixed operating point for all letter checks can be compared to the proposed dynamic thresholds approach. The figure also shows the effect of a hypothetical perfect PHM system with a limited prognostic horizon h. It is notable that at the first letter checks, the fixed approach is already 'decimating' the population whereas the dynamic approach isn't using the PHM output yet (*FPR = TPR = 0*). After 6000 FHs, the dynamic thresholds approach selects increasingly higher [*TPR, FPR*] points on the ROC curve, ultimately surpassing the fixed threshold approach in survival probability.

In Figure 9, the dynamic thresholds approach can be compared to non-prognostic maintenance policies, such as an optimal hard-time and corrective maintenance.

5. DISCUSSION & CONCLUSION

In this paper, a new method has been presented for selecting operating points on the ROC curve for each periodic maintenance check during the lifetime of a replaceable aircraft component. This method is applicable to existing PHM models, where component age is not used as a prognostic parameter.

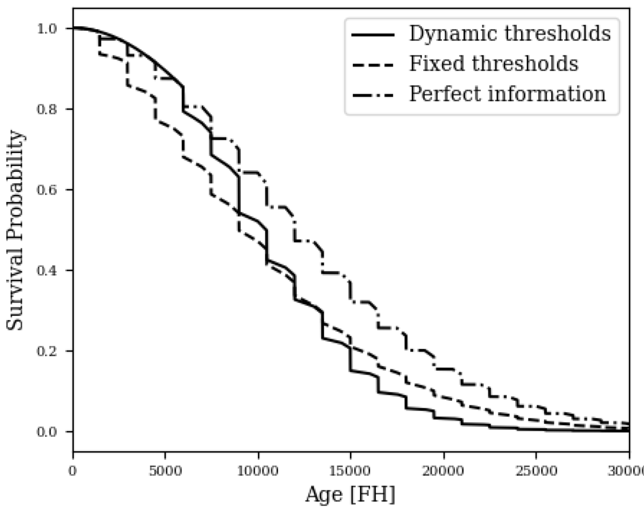


Figure 8. Adjusted survival curves for evaluated PHM based replacement strategies.

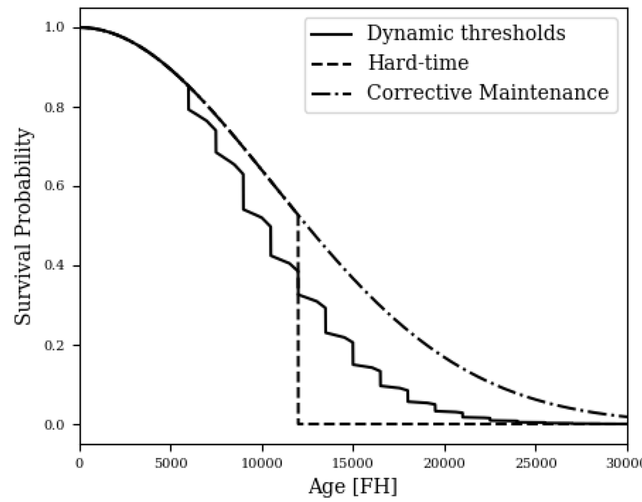


Figure 9. Adjusted survival curves for the proposed approach (dynamic thresholds) and for non-PHM based maintenance strategies.

Although the use-case shows that value can be realized compared to existing (PHM) methods, it must be noted that benefits from the case study cannot be generalized. The component's survival function $S(t)$, the airline's maintenance check interval, the cost ratio between corrective and planned maintenance ($C_c : C_p$), and the performance of the available PHM model (the ROC curve and horizon h) determine the ultimate cost saving. The type of solver and the resolution of the ROC curve is expected to have only a minor effect on expected benefits.

There are several reasons to support that the case study provides a conservative estimate of cost-savings compared to not using PHM at all. Firstly, benefits in operational reliability and fleet availability are not included in the calculation of (C_c) and (C_p). Secondly, it is expected that in reality, cost of a false-positive are smaller than the cost of a true-positive, because in the latter case, damage will have been progressed further.

This method comes with its limitations too. The maintenance opportunities in this paper are limited to periodic checks, which leaves the option to anticipate an upcoming failure between checks unused. This can be especially problematic for cases where the prediction horizon h is much smaller than the interval of the periodic maintenance checks, which could be the case in many other real-life applications than represented in this case-study. Furthermore, the scope of applicable PHM models is limited due to the assumptions made in Chapter 3. For example, prognostic models that combine monitoring data with usage data would violate the assumption of constant ROC curve. Setting optimal decision thresholds for those models is recommended as a topic of future research, also noting that the acceptability of dynamic thresholds from a company perspective may be a factor to take into account. Lastly, on the cost side, this case study does not take into account the different number of spare parts that may be needed due to changed maintenance policies (Fritzsche & Lasch, 2012), neither does it apply a discount factor for cost that is incurred later in the component's life.

This paper studied only the economic aspects of various maintenance policies. While this method is intended to prevent non safety-critical failures, there may still be some impact on safety, for example due to more frequent repairs. Hence, before implementation of the proposed method, a safety-risk assessment must be conducted by the airline. In a future with more condition-based maintenance and fewer (usage-driven) periodic tasks, long (letter) checks as we know today may cease to exist. Optimally scheduling prognostic driven tasks beyond periodic checks is therefore another recommended topic of future research.

In addition to economic benefits, there are several secondary benefits that can be expected. Firstly, the method can be used to estimate how well a PHM model should perform in order to reach target cost savings. This could help an operator prioritize its prognostic development options. Secondly, operating points from more than just one ROC curve can be selected for each check. This allows the operator to use different PHM models for different checks in the component's lifetime, which could be beneficial if both models outperform each other at different points in the ROC curve (e.g. the ROC curves cross each other). Thirdly, a wider range of PHM models becomes available to the airline, because PHM models that are disadvantageous when used with a fixed operating point can become valuable when different operating points (including *FPR=TRP=0*) are available for each maintenance check. Finally, the application scope of this paper can be expanded beyond PHM. For example, this study suggests that age-dependent, dynamic inspection limits for classical maintenance task may provide economic benefit over the fixed limits we know today.

6. ACKNOWLEDGEMENT

The work presented in this paper was financially supported by the European Union's Horizon 2020 research and innovation program ReMAP (Grant Agreement Number:769288). The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

BIOGRAPHIES

Floris Freeman works at KLM Engineering & Maintenance on Condition-Based Maintenance. After obtaining a degree in Aerospace Engineering at TU Delft, Floris worked as research engineer in well -and reservoir management for the oil & gas industry. In that role, Floris build experience in fibre-optics based distributed sensing technologies. His current focus is on predictive maintenance and optimisation for aircraft operations.

Paul van Kessel is a researcher within the Big Data team of KLM Engineering & Maintenance. His research is focused on planning and scheduling optimization within the aviation maintenance domain. This covers the aspects of long-term planning till day of operation decision making with disruptions taking place. Paul's work covers development of decision support tools for scheduling optimization to increase the MRO efficiency. Implementation of planning & scheduling optimization opens up the opportunity for predictive maintenance implementation.

Dr Wim Verhagen is a Senior Lecturer within the Aerospace Engineering and Aviation Discipline in the School of Engineering, RMIT University. Dr Verhagen's research covers predictive maintenance and decision support. With respect to predictive maintenance, Dr Verhagen's research efforts focus on the development, implementation and validation of data-driven and hybrid diagnostic and prognostic models for aircraft systems. With respect to decision support, his research covers a comprehensive and integrated set of models to streamline and optimize maintenance planning, execution and control on operational and tactical time horizons..

REFERENCES

ATA. (2013). *ATA MSG-3 Volume 1 (Fixed Wing Aircraft* (Tech. Rep. No. 2013.1). Airlines for America.

Ben-Daya, M., Kumar, U., & Murthy, D. (2016). Introduction to maintenance engineering: Modelling, optimization and management. *John Wiley & Sons: West Sussex, UK*.

Engel, S. J., Gilmartin, B. J., Bongort, K., & Hess, A. (2000). Prognostics, the real issues involved with predicting life remaining. *2000 IEEE Aerospace Conference. Proceedings (Cat. No.00TH8484)*, *6*, 457-469 vol.6.

Ferreira, L., & Silva, J. (2017, 03). Parameter estimation for weibull distribution with right censored data using em algorithm. *Eksploatacja i Niezawodnosc - Maintenance and Reliability*, *19*, 310-315. doi: 10.17531/ein.2017.2.20

Fritzsche, R., & Lasch, R. (2012). An integrated logistics model of spare parts maintenance planning within the aviation industry. *International Journal of Economics and Management Engineering*, *6*(8), 1958 - 1967.

Hölzel, N., Schilling, T., & Gollnick, V. (2014, 10). An aircraft lifecycle approach for the cost-benefit analysis of prognostics and condition-based maintenance based on discrete event simulation. *PHM 2014 - Proceedings of the Annual Conference of the Prognostics and Health Management Society 2014*.

IMRBPB. (2018). *Aircraft Health Monitoring (AHM) integration in MSG-3* (Tech. Rep. No. Issue Paper IP-180). International Maintenance Review Board Policy Board.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483-1510. doi: https://doi.org/10.1016/j.ymssp.2005.09.012

Jia, X., Huang, B., Feng, J., Cai, H., & Lee, J. (2018). A review of phm data competitions from 2008 to 2017: Methodologies and analytics. *Annual Conference of the PHM Society, 10(1)*. doi: https://doi.org/10.36001/phmconf.2018.v10i1.462

Koops, L. (2018). ROC-based business case analysis for predictive maintenance – applications in aircraft engine monitoring. *Proceedings of the European Conference of the Prognostic and Health Management Society European Conference, PHME18*.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, *104*, 799-834. doi: https://doi.org/10.1016/j.ymssp.2017.11.016

Letot, C., Equeter, L., Dutoit, C., & Dehombreux, P. (2017, 12). Updated operational reliability from degradation indicators and adaptive maintenance strategy.. doi: 10.5772/intechopen.69281

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283-298. doi: https://doi.org/10.1016/S0001-2998(78)80014-2

Sandborn, P. A., & Wilkinson, C. (2007). A maintenance planning and business case development model for the application of prognostics and health management (PHM) to electronic systems. *Microelectronics Reliability*, *47*(12), 1889-1901. (Electronic system prognostics and health management) doi: https://doi.org/10.1016/j.microrel.2007.02.016

Saxena, A., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S., & Schwabacher, M. (2008). Metrics for evaluating performance of prognostic techniques. In *2008 international conference on prognostics and health management* (p. 1-17). doi: 10.1109/PHM.2008.4711436

Sun, B., Zeng, S., Kang, R., & Pecht, M. (2010, 02). Benefits analysis of prognostics in systems. *2010 Prognostics and System Health Management Conference, PHM '10*, 1 - 8. doi: 10.1109/PHM.2010.5413503

Thijssens, O., & Verhagen, W. J. (2020). Application of extended cox regression model to time-on-wing data of aircraft repairables. *Reliability Engineering & System Safety*, *204*, 107136. doi: https://doi.org/10.1016/j.ress.2020.107136

Vianna, W. O. L., & Yoneyama, T. (2018). Predictive maintenance optimization for aircraft redundant systems subjected to multiple wear profiles. *IEEE Systems Journal*, *12*(2), 1170-1181. doi: 10.1109/JSYST.2017.2667232

APPENDIX

Table 4. ROC data used for case study

FPR	TPR
0	0
0.05	0.4
0.1	0.6
0.15	0.68
0.2	0.75
0.25	0.8
0.3	0.84
0.35	0.86
0.4	0.88
0.45	0.9
0.5	0.92
0.55	0.94
0.6	0.95
0.65	0.96
0.7	0.97
0.75	0.975
0.8	0.98
0.85	0.985
0.9	0.99
0.95	0.995
1	1