Analytical Health Indices

Towards Reliability-Informed Deep Learning for PHM

Dersin, Pierre; Goglio, Dario; Bajarunas, Kristupas; Chao, Manuel Arias

DOI
10.36001/ijphm.2025.v16i2.4262

Publication date
2025

Document Version
Final published version

Published in
International Journal of Prognostics and Health Management

Citation (APA)
Dersin, P., Goglio, D., Bajarunas, K., & Chao, M. A. (2025). Analytical Health Indices: Towards Reliability-Informed Deep Learning for PHM. *International Journal of Prognostics and Health Management*, *16*(2), Article 4262. https://doi.org/10.36001/ijphm.2025.v16i2.4262

Analytical Health Indices: Towards Reliability-Informed Deep Learning for PHM

Pierre Dersin[1], Dario Goglio[2,3,4], Kristupas Bajarunas[2,3], and Manuel Arias Chao[2,3]

[1] *Luleå University of Technology, Luleå, 97187, Sweden*
pierre.dersin@ltu.se

[2] *Delft University of Technology, 292F+VC Delft, Netherlands*
d.s.goglio@tudelft.nl
k.v.b.bajarunas@tudelft.nl
m.a.c.ariaschao@tudelft.nl

[3] *Zurich University of Applied Sciences, 8400 Winterthur, Switzerland*
baja@zhaw.ch
aria@zhaw.ch

[4] *Swiss International Air Lines Ltd., 8302 Kloten, Switzerland*
dario.goglio@swiss.com

ABSTRACT

Deep learning has demonstrated significant potential for prognostics in complex systems. Recent advances in physics-informed machine learning have integrated physics-of-failure principles within data-driven models. Beyond physical laws, fleet-level time-to-failure (TTF) distributions provide valuable prior knowledge for individual asset life predictions.

In this paper we derive a probabilistic analytical health index (HI) model based on power-law degradation, enabling a probabilistic description that reconciles individual variability with fleet-wide trends. We show that, under Weibull, Gamma, and Pareto-distributed TTFs, the HI evolution follows an analytical form, allowing explicit characterization of time to reach intermediate degradation levels. Therefore, this work provides a theoretical foundation for integrating reliability principles with deep learning, advancing towards Reliability-Informed Deep Learning (RIDL).

The approach is validated on synthetic turbofan engine data and real-world battery degradation datasets. This work establishes a rigorous basis for embedding reliability engineering principles into deep learning, improving predictive maintenance and remaining useful life (RUL) estimation.

https://doi.org/10.36001/IJPHM.2025.v16i2.4262

1. INTRODUCTION

An important step in prognostics and health management of complex industrial systems is inferring their current health condition (Fink et al., 2020). To this end, a normalized health index is often defined as a metric that measures the degree of degradation of equipment. Conventionally, a value of 1 for the health index corresponds to perfect health, and a value of 0 to a failed state. An intermediate value characterizes a state where the item is still operating but less than perfectly. If the health index captures the physical condition of the asset correctly, the time evolution of the health index is an appropriate means for performing prognostics, i.e., predicting the evolution of a degradation, eventually up to a failure, and the time until that failure, or remaining useful life (RUL). Therefore, the health index for an asset constitutes a key tool for maintenance decision-making, as it enables health assessment (in particular, degradation severity) and prognostics.

The derivation of HIs has traditionally depended on extracting key features from condition monitoring (CM) data and integrating them with a physical understanding of the asset to create a health index (Atamuradov et al., 2020). This practice, while effective, is heavily reliant on domain-specific knowledge, presenting a significant barrier to scalability and adaptability across different systems. To address these limitations, diverse data-driven approaches have been proposed for estimating HI from condition monitoring data. For instance, supervised learning models have been applied when dealing

with datasets that contain labels of HIs (Roman, Saxena, Robu, Pecht, & Flynn, 2021). Similarly, residual techniques that identify deviations from a system's expected behavior (Ye & Yu, 2021; Hsu, Frusque, & Fink, 2023) offer another pathway, albeit contingent on the existence of a representative dataset of "healthy" state labels- an often challenging prerequisite in industrial settings due to difficulties in obtaining a representative data for complex systems.

Recently, unsupervised methods combining deep learning with traditional reliability engineering principles in the form of explicit, analytical representation of the health index have shown promise in inferring asset-specific HI (Bajarunas, Baptista, Goebel, & Arias Chao, 2023; Yang, Habibullah, & Shen, 2021; Qin, Yang, Zhou, Pu, & Mao, 2023). Therefore, these recent works highlight the potential of leveraging the extensive body of reliability engineering theory (Dersin, 2023), alongside deep learning algorithms, to model RUL dynamics effectively.

This paper builds on this foundation and previous work in (Dersin, Bajarunas, & Arias-Chao, 2024) to propose an analytical probabilistic framework for HI modeling that reflects both fleet-wide trends and asset-specific conditions. By doing this, we aim to enable the integration of reliability engineering models in machine learning algorithms by providing an analytical probabilistic description of the HI. Addressing this objective involves answering the following question: *How to find an analytical description for a time-dependent health index integrating random parameters to capture asset variability and align with observed times to reach various degradation severity levels including the time to failure?*

Hence, in this work, we assume the availability of time-to-failure (TTF) distributions for a fleet of assets. Given this assumption, we formulate the problem in a general context and provide an analytical solution when the TTF follows a Gamma distribution, a Weibull distribution or a Pareto distribution. In this scenario, with a health index defined by a power law featuring either an inverse-Gamma or a Fréchet-distributed or inverse-Pareto distributed coefficient, as the case may be, we demonstrate that the time to reach any intermediate degradation level follows a Gamma, Weibull or Pareto distribution, respectively, sharing the same shape parameter as the TTF. Moreover, the scale parameter explicitly depends on the degradation level. Also, a general formula for the expectation of the time to reach any intermediate level is given, which does not depend on the time-to-failure distribution. We also detail the procedure for estimating the power law exponent from field data through regression analysis and conduct a sensitivity analysis regarding this exponent. Further, we provide additional degrees of freedom to the model by making the exponent variable a function of time, operating conditions, or maintenance. A relationship is then discovered between the exponent and the shape parameter: if the shape parameter (which reflects the speed of degradation) varies, the exponent can be adjusted accordingly so that, for a given quantile the HI reaches a given degradation level at the same time.

To validate our methodology, we present case studies focusing on the N-CMAPPS turbofan (Arias Chao, Kulkarni, Goebel, & Fink, 2021) and randomized usage Li-ion batteries datasets (Bole, Kulkarni, & Daigle, 2014). The results confirm the proposed methodology and highlights its practical applications. Obtaining an explicit, analytical representation of the health index, including the random variability among assets, is a definite advance over the state of the art that offers a major advantage. The proposed approach enables maintenance decision-making with minimal computational demand.

The paper is organized as follows: Section 2 presents the methodology used in this work; we first formulate the problem in Section 2.1 and present a resolution method in Section 2.2. We then delve into specific cases involving Gamma (Section 2.3), Weibull (Section 2.4), and Pareto distribution (Section 2.5), followed by the general expected hitting time formula (Section 2.6), and a discussion on estimating the power law exponent controlling the shape of degradation for the analyzed distributions (Section 2.7). The case of incomplete degradation trajectories is addressed in Section 2.8. In Section 2.9, the concept of a generalized health index is introduced, to reflect the dependence on time or varying operating conditions or maintenance actions. Case studies from the N-CMAPSS and randomized battery usage datasets illustrate our approach (Section 3), with sensitivity analysis on the power law exponent (Section 3.2). In Section 3.3, the generalization of the health index to a variable exponent is studied, and illustrated with the N-CMAPPS data. The paper concludes with a summary of our findings and suggestions for future research in Section 5.

2. METHOD

In this section, we describe our methodology for deriving an analytical representation of the health index (HI) in prognostics. Our goal is to connect the evolution of the HI with fleet-level time-to-failure (TTF) distributions—specifically, when the TTF is modeled as a Gamma, Weibull, or Pareto distribution.

Figure 1 summarizes the methodological steps and the broader context of this work. The proposed workflow begins by modeling the HI as a monotonic, parametric function governed by interpretable degradation parameters ($h(b, p, t)$). Given a prior distribution for the TTF (Gamma, Weibull, or Pareto models), we derive closed-form analytical expressions for the probability distribution of the time required to reach any intermediate health level s (i.e., T_s or TTS). This framework enables population-level modeling of degradation while retaining interpretability and analytical tractability.

The broader context of this work is as follows. The methodology assumes the availability of asset-specific HI derived from CM data, for instance based on (Bajarunas, Baptista, Goebel, & Chao, 2024). In the future, we envision leveraging this framework for tasks such as remaining useful life (RUL) forecasting and incorporating variable exponents to capture operating context or maintenance effects.

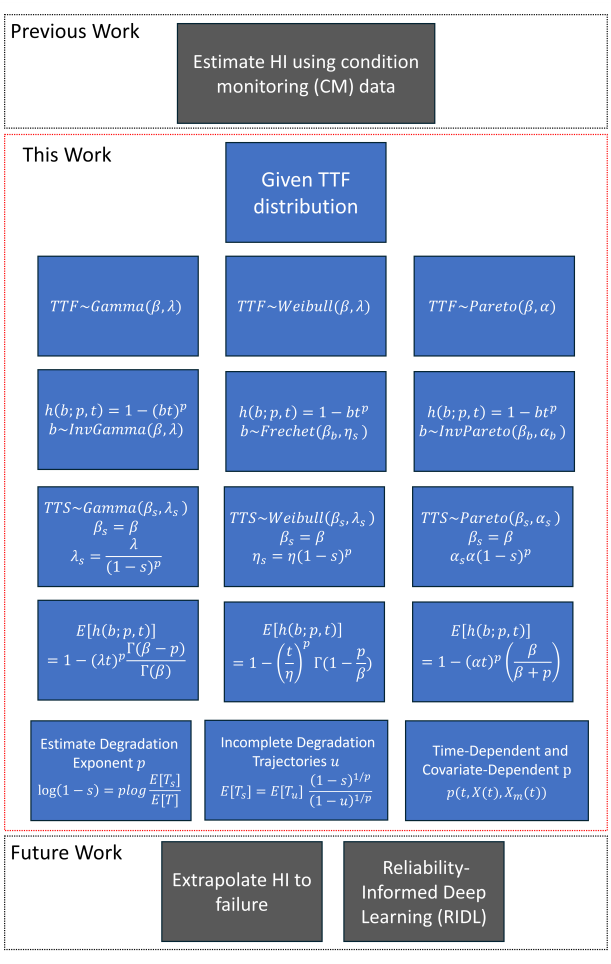


Figure 1. Flowchart illustrating the different steps of the proposed method along with the context of its applicability. The process of estimating begins with the assumption of a prior plausible Time to Failure (TTF) distribution (Gamma, Weibull or Pareto models) and a probabilistic parametric model of a Health Index ($h(b,p,t)$). Based on these assumptions, we derive the analytical forms of the distributions describing the time required for the HI to reach any specified intermediate health level (T_s), thus providing a comprehensive statistical framework to model a degradation process.

2.1. Problem Statement

A degradation process is characterized by a health index, $HI(t)$, that typically decreases monotonically with time from a value of 1 (perfect health) to 0 (failure). Let the time to failure be defined as:

$$T = \inf\{t : HI(t) = 0\} \tag{1}$$

Given that prior knowledge or data suggests that T follows a known probability distribution (e.g., Gamma, Weibull, or Pareto), our objective is to determine, for any intermediate health level $s \in (0,1)$, the distribution of the time

$$T_s = \inf\{t : HI(t) = s\} \tag{2}$$

2.2. Resolution Method: General Principle

Let $R(t)$ denote an assumed reliability function. Then a probabilistic model for $HI(t)$, as a non-increasing function of t, is selected, and the condition $P[T > t] = R(t)$ is imposed. Finally, Eq. (2) is applied to obtain the distribution of T_s:

$$R_s(t) = P[T_s > t] \tag{3}$$

Let us consider the following parametric model for the health index:

$$HI(t) = h(p_1, p_2, ...p_n; t) \tag{4}$$

with an assumed functional form h, where some of the parameters $p_1, p_2, \ldots, p_n$ are random variables.

Then, it should be noted that

$$h(p_1, p_2, ...p_n; t) > 0 \tag{5}$$

is equivalent to

$$T > t \tag{6}$$

therefore the following condition is imposed:

$$P[h(p_1, p_2, ...p_n; t) > 0] = R(t) \tag{7}$$

with the right-hand side of Eq. (7) known.

Similarly, the condition $T_s > t$ is equivalent to $HI(t) > s$ and hence from Eq. (7), one derives

$$P[h(p_1, p_2, ..., p_n; t) > s] = R_s(t) \tag{8}$$

for any value of s between 0 and 1. The method is quite general and can be applied to any TTF distribution.

In the next two subsections, the method is detailed and illustrated on three frequently encountered families of TTF distributions: Gamma, Weibull, and Pareto.

2.3. Gamma Case

Let us consider the case when the time to failure follows a Gamma distribution with shape parameter β and rate parameter λ. The Gamma reliability function for time T (Nachlas,

2017) can be expressed as:

$$R(t) = 1 - \frac{\gamma(\lambda t;\beta)}{\Gamma(\beta)} \tag{9}$$

where $\gamma(\lambda t;\beta)$ stands for the incomplete Euler gamma function.

A health index is sought, $HI(t)$, such that the time for the HI to reach the value 0 is Gamma-distributed.

We shall now show that a solution is provided by the following power law for the health index:

$$h(b;p;t) = 1 - (bt)^p \tag{10}$$

with a positive exponent p and a random variable b with an inverse-gamma distribution with shape parameter β and scale parameter λ (b has the dimension of a frequency, i.e., the inverse of a time, so does λ). The health index defined by Equation (10) decreases monotonically from 1 to 0 as the time or usage variable t increases from 0 to $\frac{1}{b}$. It is a convex function of t if $p < 1$ and a concave function if $p > 1$ (and linear in the limit case of $p = 1$). The property that b has an inverse-Gamma distribution is equivalent to $\frac{1}{b}$ having a Gamma distribution with parameters β (shape) and λ (rate).

Denoting by T the time to failure, there follows from the above health index definition that

$$P[T > t] = P[(bt)^p < 1] = P[bt < 1] = P[\frac{1}{b} > t] \tag{11}$$

Since $\frac{1}{b}$ is Gamma distributed, the right-hand side of Eq.(11) is the Gamma reliability function at time t, with shape and rate parameters respectively equal to β and λ. Therefore, it has been proved that the definition of Eq.(10) for the health index leads to a Gamma-distributed time to failure.

Now let us look at the distribution of the time for the health index to reach a level s, between 0 and 1.

Let us denote that first hitting time T_s.

$$P[T_s > t] = P[h(b;p;t) > s] = P[1 - (bt)^p > s] \tag{12}$$

Equation (12) is equivalent to:

$$P[T_s > t] = P[(bt)^p < 1 - s] = P\left[\frac{1}{b} > \frac{t}{(1-s)^{\frac{1}{p}}}\right] \tag{13}$$

Since $\frac{1}{b}$ is Gamma (β,λ) distributed, it follows from Eq.(9) that,

$$R_{T_s}(t) = P[T_s > t] = 1 - \frac{\gamma(\frac{\lambda t}{(1-s)^{\frac{1}{p}}};\beta)}{\Gamma(\beta)} \tag{14}$$

Therefore, it has been shown that T_s has a Gamma distribution with shape factor β, and rate parameter λ_s given by the following function of s and the exponent p:

$$\lambda_s = \frac{\lambda}{(1-s)^{\frac{1}{p}}} \tag{15}$$

The problem stated in the beginning has thus been solved in the case when the time to failure has a Gamma distribution. The mathematical expectations of T_s and that of the health index $HI(t)$ are then derived explicitly, as follows, from the properties of the gamma distribution and the inverse-gamma distribution (Llera & Beckmann, 2016):

$$E(T_s) = \frac{\beta}{\lambda_s} = \frac{\beta}{\lambda}(1-s)^{\frac{1}{p}} \tag{16}$$

which can also be written as :

$$E(T_s) = E(T)(1-s)^{\frac{1}{p}} \tag{17}$$

To derive the expectation of the health index $HI(t)$, we now use properties of the inverse-gamma distribution. If X has an inverse-gamma distribution with parameters β and λ, the nth-order moment of X is given (Llera & Beckmann, 2016) by:

$$E(X^n) = \lambda^n \frac{\Gamma(\beta - n)}{\Gamma(\beta)} \tag{18}$$

as long as

$$n < \beta$$

Therefore

$$E[HI(t)] = 1 - E(b^p)t^p = 1 - (\lambda t)^p \frac{\Gamma(\beta - p)}{\Gamma(\beta)} \tag{19}$$

assuming the exponent p to be smaller than the shape factor β.

2.4. Weibull Case

We shall now consider the case where the time to failure follows a 2-parameter Weibull distribution. Denoting β and η the shape and scale parameters, respectively, this corresponds to the well-known reliability function:

$$R(t) = e^{-(t/\eta)^\beta} \tag{20}$$

For the health index, let us take the following power law, slightly different from the one taken in the Gamma distribution case, for reasons which will become apparent:

$$h(b;p;t) = 1 - bt^p \tag{21}$$

where p is a positive exponent, and b is a random variable. It will be seen that, if b has a Fréchet distribution (Fréchet, 1927; Ramos, Louzada, Ramos, & Dey, 2020), then the time to failure is Weibull distributed.

Indeed, by definition of the Fréchet (also known as "inverse Weibull") distribution, if the random variable b is Fréchet-distributed with scale parameter λ_b and shape parameter β_b, then:

$$P[b > u] = 1 - exp[-(\frac{u}{\lambda_b})^{-\beta_b})] \tag{22}$$

The shape parameter β_b is dimensionless, and the scale parameter λ_b has the dimension of t to the power of $(-p)$, just as the coefficient b.

Then, by substituting

$$u = t^{-p} \tag{23}$$

in Eq.(22), the following is obtained :

$$P[HI(t) > 0] = P[b < t^{-p}] = exp[-(\lambda_b t^p)^{\beta_b}] \tag{24}$$

and this expression must be equated to $P[T > t]$, which is assumed to be the reliability function of a two-parameter Weibull variable (η ,β).

Therefore, the parameters of the Fréchet distribution for b are obtained as follows:

$$\lambda_b = 1/\eta^p \tag{25}$$

$$\beta_b = \beta/p \tag{26}$$

as it can be verified by substituting the right-hand sides of Eq.(25) and Eq.(26) respectively for λ_b and β_b in (24). Then the distribution of T_s, the first hitting time of level s, can be derived as well, for any value of s between 0 and 1.

$$P[Ts > t] = P[HI(t) > s] \tag{27}$$

$$= P[1 - bt^p > s] = P[b < (1-s)t^{-p}] \tag{28}$$

Therefore, by substituting for u in Eq.(22) the value $(1-s)t^{-p}$ and using Eq.(25) and Eq.(26),

$$P[Ts > t] = exp[-(t^p/\eta^p(1-s))^{\frac{\beta}{p}}] \tag{29}$$

or

$$P[Ts > t] = exp[-(1-s)^{-\frac{\beta}{p}}(\frac{t}{\eta})^{\beta}] \tag{30}$$

It is seen that Eq.(30) describes the reliability function of a Weibull random variable with: 1) the same shape factor β as the distribution of T; 2) A scale factor η_s expressed as follows as a function of s , the scale factor η of T and the exponent p:

$$\eta_s = \eta(1-s)^{\frac{1}{p}} \tag{31}$$

Thus, the problem stated in the beginning has also been solved in the Weibull distribution case.

Accordingly, the mathematical expectation of the first hitting time T_s is obtained:

$$E(T_s) = \eta(1-s)^{\frac{1}{p}}\Gamma(1 + \frac{1}{\beta}) \tag{32}$$

Equation (32) can also be formulated as

$$E(T_s) = E(T)(1-s)^{\frac{1}{p}} \tag{33}$$

which is the same as in the Gamma-distribution case (Eq. (17)). Also, the expectation of the health index $HI(t)$ at time t can be derived from the expectation of the random coefficient b, assumed Fréchet distributed:

$$E(b) = \frac{1}{\eta^p}\Gamma(1 - \frac{p}{\beta}) \tag{34}$$

Therefore

$$E(HI(t)) = 1 - E(b)t^p = 1 - (\frac{t}{\eta})^p\Gamma(1 - \frac{p}{\beta}) \tag{35}$$

The quantiles of b can also be derived. The q-percent quantile is B_q, defined by

$$P[b > q] = B_q \tag{36}$$

Therefore,

$$B_q = \frac{1}{\eta^p(-lnq)^{\frac{p}{\beta}}} \tag{37}$$

In particular, the median (50-percent quantile) is given by:

$$B_{0.5} = \frac{1}{\eta^p(ln2)^{\frac{p}{\beta}}} \tag{38}$$

2.5. Pareto Case

After Weibull and Gamma, let us consider yet another distribution: the Pareto (type 1) distribution, defined as follows:

$$R(t) = P[T > t] = (\frac{\alpha}{t})^{\beta} \tag{39}$$

with scale parameter α and shape parameter β (both positive, and with β greater than 1). The Pareto distribution reflects the property that a nonnegligible number of items in the pop-

ulation has long life. Then, if the health index is defined by Eq.(21) just as in the Fréchet case,the probability distribution of b is derived as follows:

$$P[1 - bt^p] > 0] = (\frac{\alpha^p}{t^p})^{\frac{\beta}{p}} \tag{40}$$

Or equivalently,

$$P[\frac{1}{b} > t^p] = (\frac{\alpha^p}{t^p})^{\frac{\beta}{p}} \tag{41}$$

This is to say that the random variable $\frac{1}{b}$ has a Pareto distribution with shape parameter:

$$\beta_b = \frac{\beta}{p} \tag{42}$$

and scale parameter:

$$\alpha_b = \alpha^p \tag{43}$$

Equivalently, one can say that b has an 'inverse Pareto' distribution, also known as power distribution:

$$P[b > x] = 1 - (\alpha_b x)^{\beta_b} \tag{44}$$

From the properties of the Pareto and power distributions, there follow expressions for the mean time to failure and the expectation of the health index.

$$E(T) = \frac{\alpha\beta}{\beta - 1} \tag{45}$$

$$E[HI(t)] = 1 - \frac{\beta}{\beta + p}(\alpha t)^p \tag{46}$$

In the same way, as for the Gamma and Weibull distributions,the distribution of T_s,the time for the HI to reach level s,is determined :it is a Pareto distibution with shape parameter β and scale parameter α_s, with

$$\alpha_s = \alpha(1 - s)^{1/p} \tag{47}$$

Consequently, the expected time to reach level s is :

$$E(T_s] = \frac{\alpha\beta}{\beta - 1}(1 - s)^{1/p} \tag{48}$$

An expression for the quantiles can be derived as well. Denoting as previously by β_q the q-percent quantile, as defined in Eq.(36),

there follows, from Eq.(44),

$$B_q = \frac{1}{\alpha^p}(1 - q)^{\frac{p}{\beta}} \tag{49}$$

2.6. General formula for $E(T_s)$

The formula for $E(T_s)$ given for the Weibull, Gamma and Pareto distributions is in fact quite general. We now prove it in the general case i.e. regardless of the TTF distribution, for a health index described by a power law. Since T_s is nonnegative,

$$E(T_s) = \int_0^\infty P[T_s > t]\, dt \tag{50}$$

and, for any health index $HI(t)$, the condition $T_s > t$ is equivalent to $HI(t) > s$. There follows:

$$E(T_s) = \int_0^\infty P[HI(t) > s]\, dt \tag{51}$$

for s between 0 and 1. In particular, for $s = 0$,

$$E(T) = \int_0^\infty P[HI(t) > 0]\, dt \tag{52}$$

Therefore

$$\frac{E(T_s)}{E(T)} = \frac{\int_0^\infty P[HI(t) > s]\, dt}{\int_0^\infty P[HI(t) > 0]\, dt} \tag{53}$$

Let us now consider a health index defined by a power law such as in Eq. (21) and let $R(t)$ denote the complementary cdf (cumulative distribution function) of the random variable $\frac{1}{b}$.

$$R(x) = P[\frac{1}{b} > x] \tag{54}$$

Then

$$\begin{aligned} P[HI(t) > s] &= P[1 - bt^p > s] \\ &= P[\frac{1}{b} > \frac{t^p}{1 - s}] = R(\frac{t^p}{1 - s}) \end{aligned} \tag{55}$$

Therefore, from Eq.(53) ,

$$\frac{E(T_s)}{E(T)} = \frac{\int_0^\infty R(\frac{t^p}{1-s})\, dt}{\int_0^\infty R(t^p)\, dt} \tag{56}$$

Now consider the change of variable

$$u = \frac{t}{(1 - s)^{\frac{1}{p}}} \tag{57}$$

which implies

$$dt = (1 - s)^{\frac{1}{p}} du \tag{58}$$

There follows

$$\frac{E(T_s)}{E(T)} = (1 - s)^{\frac{1}{p}} \frac{\int_0^\infty R(u^p)\, du}{\int_0^\infty R(t^p)\, dt} = (1 - s)^{\frac{1}{p}} \tag{59}$$

This formula is true for any HI expressed by a power law as in Eq.(21), regardless of the particular TTF distribution. It can be shown in exactly the same way that the formula is

also valid for an HI of the type Eq. (10) used earlier for the Gamma distribution. An important remark is that Eq.(59) is valid, not only for constant p, but for a function $p(s)$. Indeed, Eq.(58) would remain valid.

2.7. Estimation of Exponent p from Data

From Eq.(59), there follows, by taking logarithms,

$$log(1-s) = p\log(\frac{E(T_s)}{E(T)}) \tag{60}$$

Therefore, in order to estimate p, it is equivalent to estimate $E(T_s)$ from the data samples corresponding to several values of s and then run the linear regression of $\log(1-s)$ with respect to $\log\left(\frac{E(T_s)}{E(T)}\right)$. The regression coefficient (slope) is the best estimate of p. The method applies in any distribution case since the dependence of E(T_s) on s is the same in all cases, as per the previous subsection (Eq.(59)). Furthermore, it can apply to a variable exponent p function of s.

2.8. Incomplete Degradation Trajectories

Our method for obtaining an analytical form of the HI does not require run-to-failure condition monitoring data [1]. Let us consider u as the smallest threshold of $HI(t)$ observed for all units in the fleet. Then in Eq. (33), instead of considering the expected TTF, $E(T)$, we would consider the expected time to hit the common threshold $E(T_u)$. The revised equation would be:

$$E(T_s) = E(T_u)\frac{(1-s)^{1/p}}{(1-u)^{1/p}} \tag{61}$$

Where $E(T_s)$ is the sample arithmetic mean for each value $s > u$. When $u = 0$, this is equivalent to Eq. (33). The exponent p can be estimated from linear regression in

$$\log(1-u)-\log(1-s) = p(\log(E(T_u))-\log(E(T_s))) \tag{62}$$

Once again, this method would be valid for an exponent p function of the threshold s instead of being constant.

2.9. Time-Dependent and Covariate-Dependent Degradation: Generalizing the Health Index Model

The classical degradation model assumes that the health index evolves according Eq. (21). While this formulation effectively models many degradation processes, it implicitly assumes that the degradation rate follows a fixed power-law behavior across all units. However, in practical applications, p does not necessarily remain constant but instead evolves dynamically based on operating conditions, maintenance history, and system wear mechanisms. This variability reflects the fact that some components degrade progressively, while others exhibit accelerated wear leading to failure.

To introduce a more flexible formulation, we allow p to vary over time and be influenced by external factors:

$$h(t) = 1 - bt^{p(t,\mathbf{X}(t),X_m(t))} \tag{63}$$

where $p(t, \mathbf{X}(t))$ is now a function of:

- **Time** t: Capturing how degradation dynamics change over the system's lifetime,
- **Operational conditions** $\mathbf{X}(t)$: External factors such as load, temperature, and humidity,
- **Maintenance interventions,** $X_m(t)$: Adjusting degradation based on past corrective or preventive actions.

In addition, p can be formulated recursively, where the current value depends on past values:

$$p(t) = \phi_0 + \phi_1 p(t-1) + f(\mathbf{X}(t), X_m(t)) + \epsilon_t \tag{64}$$

This recursive form allows for persistent effects and smoother modeling of degradation memory or explicitly making $p(t, s)$. Different functional forms of $p(t, \mathbf{X}(t), X_m(t))$ can be considered:

Linear growth in time:

$$p(t) = p_0 + \alpha t \tag{65}$$

where p_0 is the initial exponent and $\alpha > 0$ accounts for accelerating degradation.

Environmental and operational influence:

$$p(t, \mathbf{X}(t)) = p_0 + f_e(X(t)) \tag{66}$$

where the effect of each operational condition $X(t)$.

Maintenance-modulated behavior:

$$p(t, X_m) = p_0 + f_m(X_m(t)) \tag{67}$$

where X_m represents maintenance.

This generalization enhances the predictive power of HI models by accommodating real-world variability in degradation behavior:

- **Heterogeneous Degradation Paths**: Each unit can exhibit unique degradation trajectories based on its operational environment and maintenance history.
- **Modeling Late-Stage Acceleration**: By allowing p to increase over time, the model can capture the rapid deterioration often observed before failure.
- **Dynamic Maintenance Integration**: Maintenance can now be incorporated as a factor that actively modifies the degradation curve rather than simply resetting $h_i(t)$.

[1] If no failures are observed the HI has a different meaning as it is normalized with respect to the most degraded unit in the fleet.

By introducing a time-varying exponent p, the proposed approach better captures the complexities of real-world degradation, improving both predictive accuracy and maintenance decision-making.

3. CASE STUDIES

To validate the proposed analytical health index framework, we apply our methodology to two distinct datasets: one from turbofan engine degradation and another from battery capacity degradation. In both cases, we demonstrate how the derived expressions for the time to reach intermediate degradation levels (T_s) can be used to characterize the degradation process and estimate key parameters such as the exponent p under constant and variable scenarios.

3.1. Constant Degradation Exponent (p)

3.1.1. Turbofan

The New Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS) dataset (Arias Chao et al., 2021) offers comprehensive degradation trajectories of turbofan engines until failure. Among the dataset's eight subsets, we focus on DS03, characterized by a failure mode impacting the efficiency and flows of both low-pressure and high-pressure turbines.

The N-CMAPSS dataset characterizes degradation at the component level across initial, normal, and abnormal degradation stages. Consequently, an HI is calculated through a nonlinear mapping of operational margins under reference conditions. System failure is determined when the HI reaches 0. The dataset also accounts for between-flight maintenance by allowing improvements in engine health parameters within specified limits. The ground truth HI is shown in Figure 2, and will be used to verify the findings of Section 2.3 and 2.4. Estimating the HI using condition monitoring data as highlighted in (Bajarunas et al., 2023) is also possible.

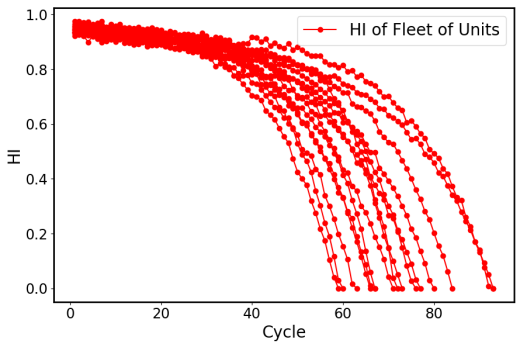


Figure 2. Observed HI in N-CMAPSS DS03 Dataset

The Akaike Information Criterion (AIC) (Akaike, 1974) was used to compare the goodness of fit with different probability distributions (Weibull, Gamma, Exponential), see Table 1. When a statistical model is used to represent the process that generated some data, some information is lost. The AIC, based on information theory, estimates the amount of information lost. It deals both with overfitting and underfitting by taking model simplicity into account as well as goodness of fit. The AIC is defined by

$$AIC = -2log(maxL) + 2P \tag{68}$$

where the term log(maxL) denotes the maximum value of the log-likelihood function, and P is the number of parameters in the model (for instance, for Weibull or Gamma, P is equal to 2). In our example, the best value of the AIC was obtained with the Gamma distribution for the time to failure as well as the time to reach level s for s ranging from 0 to 0.8. The AIC value for Weibull distribution is almost identical. In contrast, the AIC value for the exponential distribution is much higher.

Using the Maximum Likelihood Estimation technique, we estimated the best-fit Gamma parameters for various s thresholds. Figure 3 shows the estimated β_s and λ_s values for $s = [0, 0.1, 0.2, ..., 0.8]$. The results validate the conclusion presented in Section 2.3: the distribution of the first hitting time T_s shares the same shape factor $\beta = 52.83$ as the distribution of failure times T. Additionally, the rate parameter λ_s is a function of s and λ of T. We determined $p = 3.35$ following the description provided in Section 2.7. The wide confidence intervals of β_s and λ_s can be primarily attributed to the limited number of observations (15 run-to-failure curves), rather than to the choice of the Gamma distribution, which we have demonstrated to be the most suitable among the alternative distributions investigated.

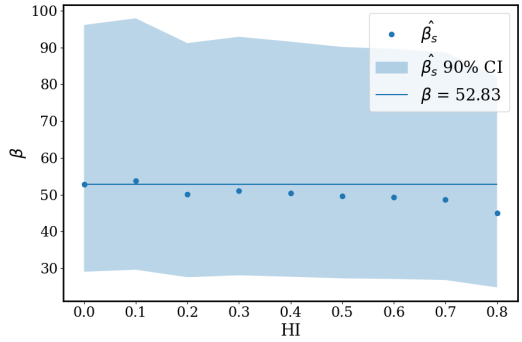
We then estimated the best-fit Weibull parameters for various s thresholds. In Figure 4, we estimated β_s and η_s using $s = [0, 0.1, 0.2, ..., 0.8]$. The results validate the conclusion presented in Section 2.4: the distribution of the first hitting time T_s shares the same shape factor $\beta = 7.32$ as the distribution of failure times T. Additionally, the scale parameter η_s is a function of s and η of T.

Figure 5 illustrates the mean, median, and 90% quantile of $HI(t)$, as described by Equations (35) and (37). Notably, we observe that the median closely aligns with the ground truth HI within the dataset.

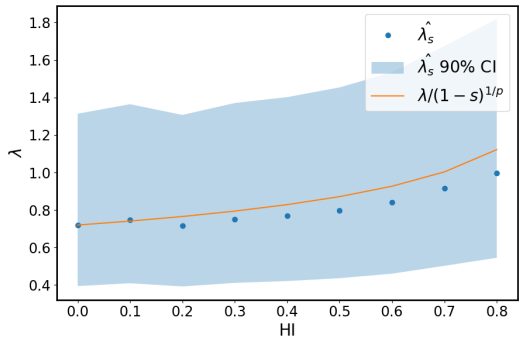
3.1.2. Battery

The methodology proposed in this study was further validated using a dataset obtained from the NASA Ames Prognostics Center of Excellence repository, specifically focusing on battery usage patterns (Bole et al., 2014). This dataset includes information collected from individual 18650 LCO cells undergoing various charging and discharging cycles following randomized protocols.

Batteries commonly exhibit several physical aging mecha-

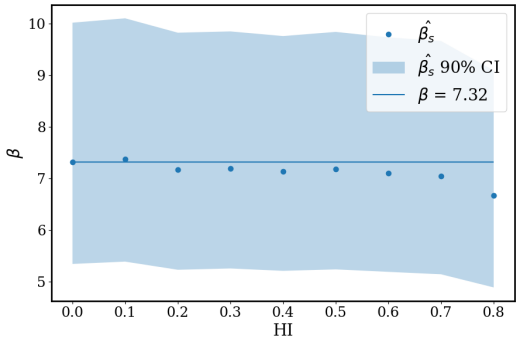


(a) β_s

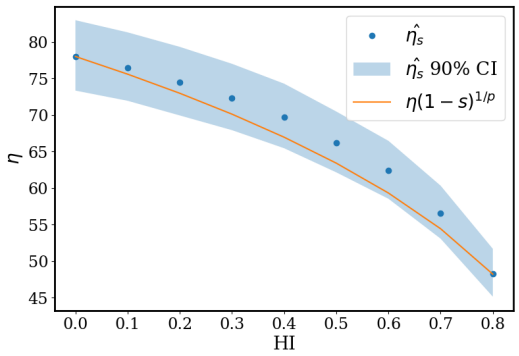


(b) λ_s

Figure 3. The Gamma distribution shape factor β_s and the rate parameter λ_s for various HI thresholds for N-CMAPSS dataset.



(a) β_s



(b) η_s

Figure 4. The Weibull distribution shape factor β_s and the scale factor η_s for various HI thresholds for N-CMAPSS dataset.

Table 1. AIC of distribution fits for CMAPSS turbofan case study.

s	AIC Gamma	AIC Weibull	AIC Exponential
0	118	120	161
0.1	117	120	161
0.2	117	120	160
0.3	116	119	159
0.4	115	118	158
0.5	114	116	156
0.6	112	114	154
0.7	109	112	151
0.8	106	108	147
0.9	116	113	127

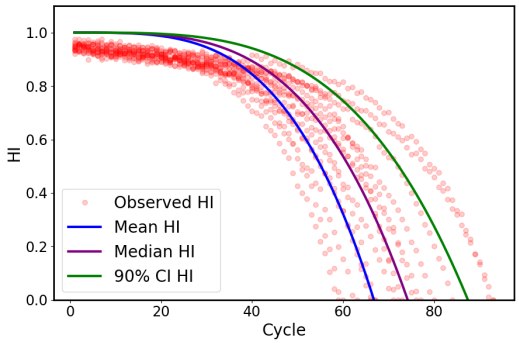


Figure 5. The mean, median, and 90% quantile of the health index obtained from Weibull distribution.

nisms such as graphite exfoliation, electrolyte loss, solid electrolyte interface layer formation, continuous thickening, and lithium plating, among others (Sui et al., 2021). These aging processes lead to two primary changes in battery behavior: capacity degradation and increased internal resistance. In this analysis, our focus will be on capacity degradation as the key health index for the batteries under investigation.

The HI of a battery is defined as the ratio between its current capacity and the nominal capacity (Q/Q_{nominal}). The battery's capacity can be determined by reference discharge cycles conducted at a constant current (I). The current capacity is calculated as the integral of current over the entire reference discharge cycle, denoted as $\int_t I$.

In this work, the failure of a battery (HI = 0) is defined once the capacity ratio is less than 60%. The initial HI of the battery is equal to the initial capacity ratio. Figure 6 shows the estimated HI of the NASA battery dataset.

The AIC values of three different distribution fits are shown in Table 2. The best fit was obtained with Gamma distribution for the time to failure as well as the time to reach level s for s ranging from 0 to 0.9. The AIC value for Weibull dis-

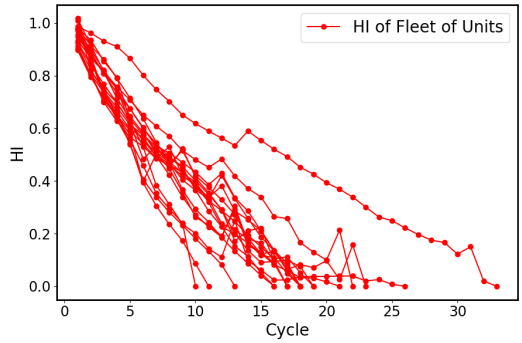


Figure 6. Observed HI in NASA battery dataset

tribution is almost identical,in contrast with the exponential distribution AIC, much higher.

We used the Maximum Likelihood Estimation (MLE) method to estimate the Gamma distribution parameters corresponding to different threshold levels s of the Health Index (HI). Specifically, we computed the shape (β_s) and rate (λ_s) parameters for thresholds $s = [0, 0.1, 0.2, \ldots, 0.8]$, where $s = 0$ denotes failure. For each value of s, the estimation was based on the first hitting times—i.e., the time at which each unit's HI trajectory crossed the threshold s. The results are shown in Figure 7, with the horizontal blue line indicating the value of β estimated at failure ($s = 0$), as also seen in Figures 3 and 4.

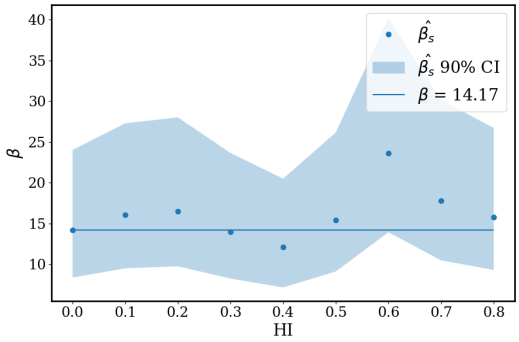
Confidence intervals were computed using the Fisher Information Matrix, leveraging the fact that the MLE is asymptotically normally distributed. These results confirm that the shape parameter β_s of the first hitting time remains close to the shape parameter β of the original TTF distribution across the range of s, supporting the validity of the analytical formulation. Additionally, using the previously estimated value $p = 0.94$, we show that the rate parameter λ_s varies systematically with both s and λ. The condition $p < 1$ results in a convex degradation curve for the HI, consistent with the observed behavior.

The best-fit Weibull parameters for various s thresholds are shown in Figure 8. Once again, we show that a reasonably good approximation for the shape parameter of the first hitting time β_s is the shape parameter β of the failure time T and that the scale parameter η_s varies with s and η as expected.

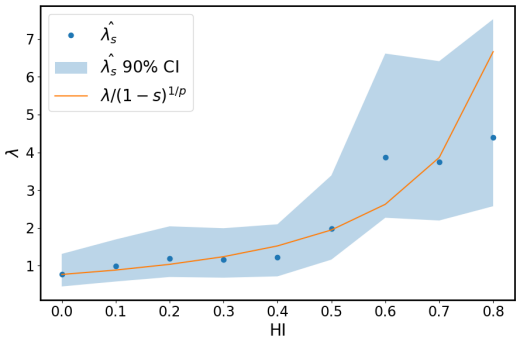
3.2. Sensitivity Analysis

Sensitivity analysis has been conducted on the N-CMAPSS dataset, to investigate the effect of the exponent p in the parametric model of the health index.

For the Gamma case, it is immediate from (15) that, for given s, λ_s is a decreasing function of p (for p greater than, or equal to 1). In the limit of p going to infinity, λ_s converges to λ. For the Weibull case, a similar conclusion is drawn, but in-



(a) β_s



(b) λ_s

Figure 7. The Gamma distribution shape factor β_s and the rate parameter λ_s for various HI thresholds for NASA battery dataset.

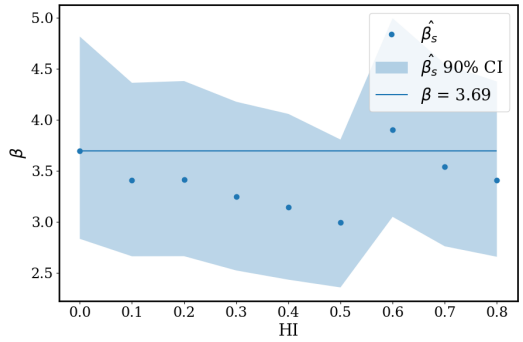
Table 2. AIC of other distribution fits for NASA battery case study.

s	AIC Gamma	AIC Weibull	AIC Exponential
0.0	118	121	151
0.1	111	118	146
0.2	104	112	140
0.3	102	108	135
0.4	97	102	128
0.5	84	93	118
0.6	67	75	109
0.7	62	70	99
0.8	54	60	89
0.9	44	50	71

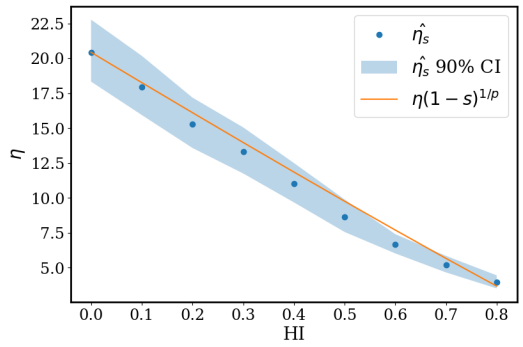
stead from (31) it follows that, for given s, η_s is an increasing function of p.

From Eq.(17) and Eq.(33) it follows that for both considered distributions the average time to reach threshold s, $E(T_s)$, is an increasing function of p, as illustrated in Figure 9.

For both distributions, when p increases, the average value of the HI is first higher than, and subsequently (for greater values of the time variable t), lower than, the HI corresponding to a lower value of p. Increasing p corresponds to delaying the decrease in HI, i.e., delaying the onset of the degradation; but, once the degradation occurs, it is more sudden. See Fig-



(a) β_s



(b) η_s

Figure 8. The Weibull distribution shape factor β_s and the scale factor η_s for various HI thresholds for NASA battery dataset.

ure 10 for an illustration.

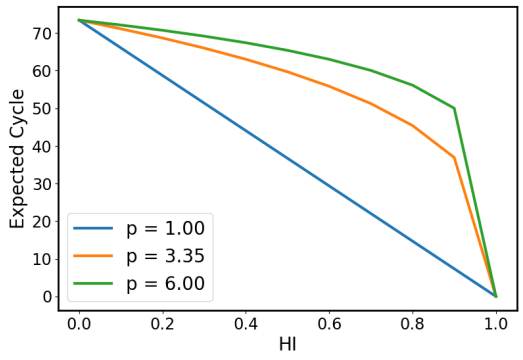
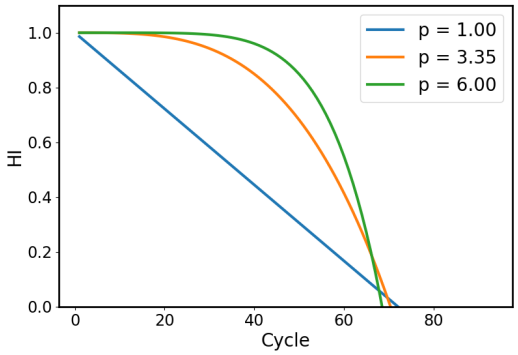


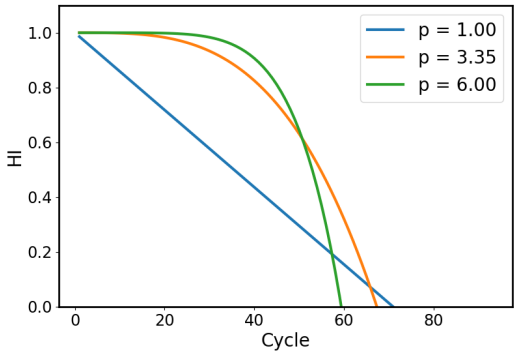
Figure 9. Gamma and Weibull distribution $E[T_s]$ as a function of s for three values of p. N-CMAPSS dataset.

3.2.1. Relationship Between shape factor and Exponent p

Let us focus on the Weibull case, although a similar reasoning could be carried out with other distributions. If we compare two data sets with different shape factors β , the corresponding health indices will have different exponents p. One can ask the question: what is the relationship between those two parameters? As seen in the previous subsection, a higher



(a) Gamma distribution $E[HI(t)]$ as a function of t for three values of p. N-CMAPSS dataset.



(b) Weibull distribution $E[HI(t)]$ as a function of t for three values of p.

Figure 10. Sensitivity to various p for the turbofan case study. N-CMAPSS dataset.

value of p provides a better fit to data that correspond to a faster degradation, therefore a higher shape factor. One can quantity that relationship. To do so, consider the q-percent quantile, B_q, which can be expressed in terms of the Weibull parameters (see Eq.(37)). Namely:

$$B_q = \frac{1}{\eta^p(-lnq)^{\frac{p}{\beta}}}$$

Therefore the health index trajectory corresponding to the q-percent quantile is described by:

$$HI_q(t) = 1 - B_q t^p = 1 - (\frac{1}{\eta^p(-lnq)^{\frac{p}{\beta}}})t^p \qquad (69)$$

This trajectory reaches 0 (failure) at time $t_{EOL}(q)$, where

$$t_{EOL}(q) = \eta(-ln(q))^{\frac{1}{\beta}} \qquad (70)$$

Clearly, $t_{EOL}(q)$ is independent of the exponent p. As expected, the failure time depends only on the parameters of the Weibull distribution, and on where in the population the particular sample is situated, i.e., which quantile it corresponds to. Those parameters determine when the asset will fail, but not how fast it degrades. That indication is given by the ex-

ponent p, as seen in Figure 11.

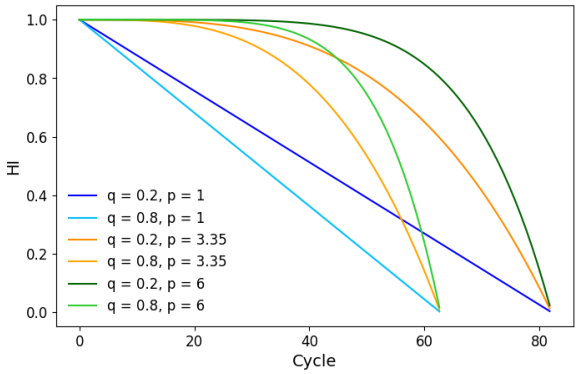


Figure 11. HI for various quantiles and p exponents

Indeed, to determine the time t_s at which the HI reaches the level s, one solves

$$HI_q(t) = s \tag{71}$$

for t, which provides t_s:

$$t_s = (1-s)^{\frac{1}{p}} \eta \, (-ln(q))^{\frac{1}{\beta}} \tag{72}$$

From this equation, one can answer the following question: for a different shape factor β, how must the exponent p in the health index be adapted so that the HI keeps reaching level s at the same time, i.e., t_s remains the same. This defines a relationship between p and β for a given threshold s:

$$\frac{1}{p(\beta)} = 1 - (\frac{1}{\beta} - 1)\frac{ln(-ln(q))}{ln(1-s)} \tag{73}$$

As an illustration, let us compare, for the 20-percent quantile, two health indices that cross the threshold 0.5 at the same point in time, corresponding respectively to, for the first, a shape factor of 2 and an exponent equal to 1.52, and, for the second, a shape factor of 4 and an exponent equal to 2.06. This case is illustrated in Figure 12. After crossing the threshold, it is seen that the curve corresponding to a higher shape factor (faster degradation) reaches level 0 earlier. The HI corresponding to a higher value of p lies above the other one before crossing the threshold and below it afterwards;

3.3. Variable Degradation Exponent (p)

The formulation as in Eq. (31) implies that, for any fixed degradation level s (i.e., horizontal cuts along the degradation curves), the parameter p influences the probability distribution of the crossing times t_s, which represents the moments when individual units reach the given degradation level. Furthermore, p affects the distribution of the random variable b, as seen e.g. in the Weibull case, the time t_s can be expressed as shown in Eq. (72), where q is the cumulative probability of the random variable b, and η and β are the Weibull scale and shape parameters, respectively. This relationship highlights

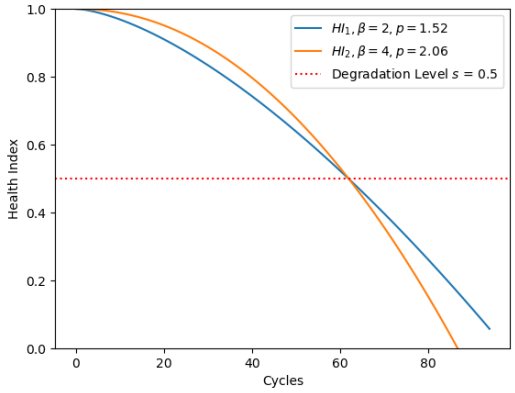


Figure 12. $HI_2(t)$ (orange), faster degradation, reaches failure earlier than $HI_1(t)$ (blue)

that p and q together influence the timing to cross a degradation level and hence the degradation trajectory.

The assumption of fixed values B_q per unit as in Eq. (37) implies that the degradation curves described by the stochastic health index $HI(t) = 1 - bt^p$ should not intersect. However, as illustrated in Figure 13, crossings are observed in the N-CMAPSS dataset between degradation curves of units 1 and 7, and units 3 and 8. Such crossings challenge the assumption of fixed B_q and p parameters for each unit. If p is truly fleet-wide and q (or equivalently B_q) is unit-specific and fixed, crossings should not occur.

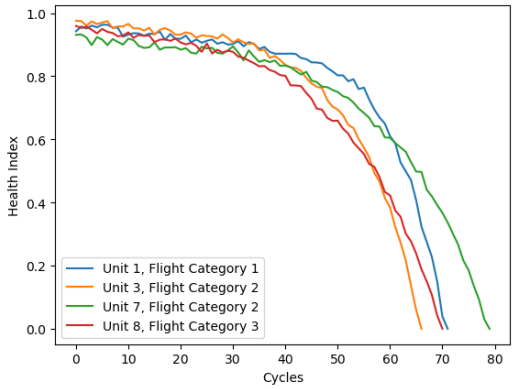


Figure 13. Crossing of degradation curves

To address this discrepancy, we propose following approach: introducing time variability in p and hence in B_q as well. If we assume that q is unit-specific and fix, we can calculate q from the end-of-life (EOL) time points t_{EOL}:

$$q_{EOL} = e^{\left(-\frac{t_{EOL}}{\eta}\right)^{\beta}} \tag{74}$$

The corresponding B_q values can be expressed as in Eq. (37). Inserting B_q into the health index formula t_s for specific degra-

dation levels, yields:

$$p(t_s, s) = \frac{log(1-s)}{log(t_s) - log(\eta) - \frac{1}{\beta} log(-log(q_{EOL}))} \quad (75)$$

By substituting the expression for q_{EOL} in Eq. (74) derived from the Weibull distribution, we obtain a simplified formulation that is independent of the Weibull parameters:

$$p(t_s, s) = \frac{log(1-s)}{log(\frac{t_s}{t_{EOL}})} \quad (76)$$

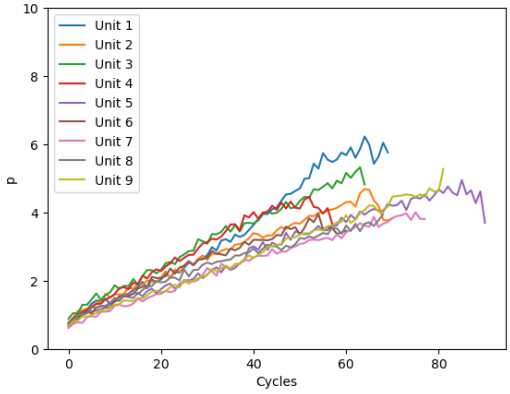


Figure 14. Degradation Exponent p varying in time

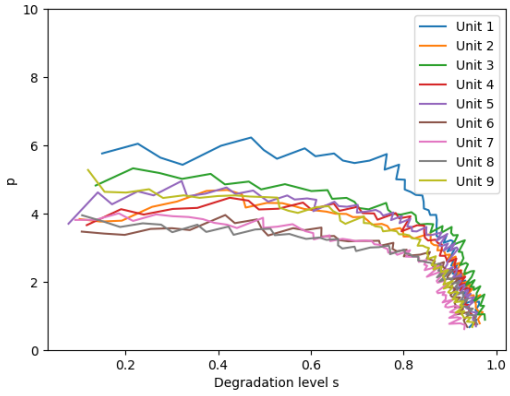


Figure 15. Degradation Exponent p varying per degradation level

Note that the last two samples of each unit is not plotted, as the p exponent goes to 0 when t approaches t_{EOL}, which is consistent with Eq. 76. This evolution of p can also be visualized as a function of the degradation level s, as shown in Figure 15. Using this approach, Figure 14 illustrates unit-specific p values computed from degradation data. These p values deviate from the initial fleet-wide p. The following is a heuristic approach; we assume linear growth as in Eq.(65) following observations from Figure 14 and also account for differences between units:

$$p(t) = p_0 + \alpha t + \epsilon(t), \quad (77)$$

where p_0 and α are fleet-wide parameters and $\epsilon(t)$ is an asset-specific parameter and can be a function of e.g. environmental or operational influence $X(t)$ or dependent on maintenance interventions X_m as elaborated in section 2.9. This individual contribution to the parameter p makes crossings possible. Fitting this formula to the N-CMAPSS dataset yields values for $p_0 = 0.913$ and $\alpha = 0.053$.

We assume that Eq. (59) still holds for variable p (it was shown in 2.8 that it does if p depends on t only through s). Furthermore, we replace, in the exponent, the value p, for constant p, with the conditional expectation of p given s. We estimate that expectation from ground truth degradation curves by the empirical mean, i.e. the arithmetic mean of the values taken by p on all the samples when the respective HIs cross different degradation levels $s = [0, 0.1, 0.2, ..., 0.8]$:

$$E[p \mid s] = \frac{1}{N_s} \sum_{i=1}^{N_s} p_i(t_i), \quad (78)$$

where $p_i(t_i)$ are the respective individual p values calculated with Eq.(76). To illustrate the improved fit of the varying p value compared to the constant p value, two Q-Q plots are presented. The theoretical quantiles are obtained from the 2P Fréchet distribution, with shape parameter β_b and scale parameter λ_b, as defined in Eq.(25) and Eq.(26), respectively. For the theoretical quantiles, the cumulative probabilities q are evenly spaced across the sample, excluding the boundary cases 0 and 1. Specifically, q_i is given by:

$$q_i = \frac{i}{n+1}, \quad (79)$$

where i is the index of the sample, and n is the total number of samples. The corresponding quantiles are then determined using Eq. (37).

In contrast, for the observed quantiles, the B_q values are directly calculated from the N-CMAPSS dataset by rearranging Eq. (21):

$$B_q = \frac{1-s}{t_s^p}, \quad (80)$$

where p was either the constant p value or the $E[p \mid s]$, where at different degradation levels $s = [0, 0.1, 0.2, ..., 0.8]$, the mean of the respective p values according to Eq.(76) was calculated:
Comparing Figure 16, where the p exponent is a fixed value of 3.35 as described in section 3.1.1 and Figure 17 with varying p exponent as described in this section, the fit of the variable p seems better visually. When calculating the MSE to the ideal reference line, we observe $4.11e-5$ for fixed p and $4.80e-6$ for the varying p exponent. Figure 18 shows the analytical

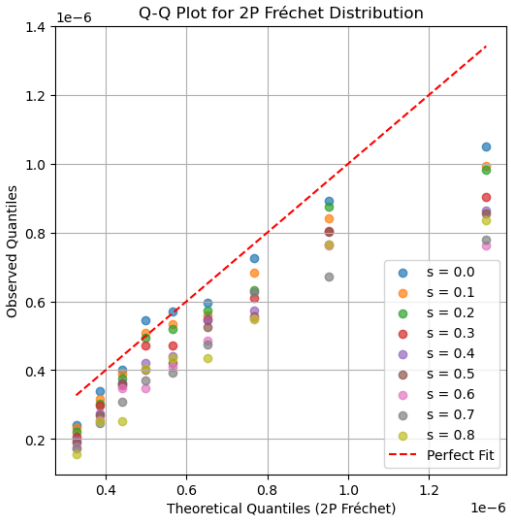


Figure 16. Q-Q plot with constant p value

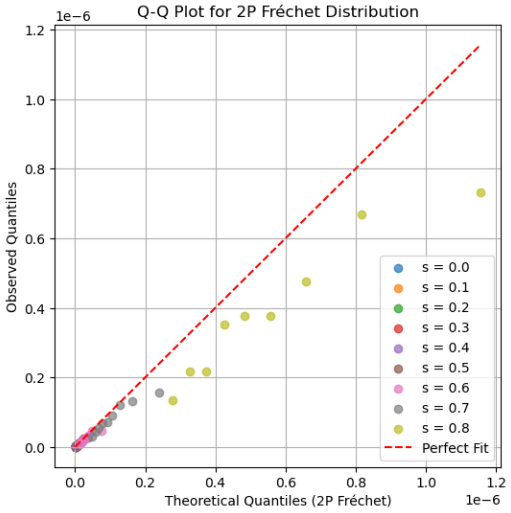


Figure 17. Q-Q plot with varying p value

curves with the varying p values according to Eq. (77), omitting the $\epsilon(t)$. Therefore, these curves can be interpreted as the behavior of the unit in case it would observe mean operational and external conditions as well as mean maintenance interventions:

4. DISCUSSION

It is fully realized that the case studies considered in this paper involve data sets which are either synthetic or laboratory data. Real-world data sets will be characterized by additional complexity; in particular, censored data, multiple failure modes, possibly interaction between failure modes. This added complexity will call for pre-processing methods such as data segmentation, outlier detection, and data imputation, all of which may affect parameter estimation quality.

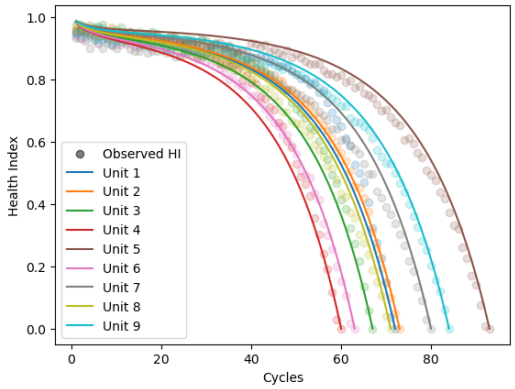


Figure 18. Ground truth health index curve and corresponding analytical health index curves

It should be noted that the proposed method can be applied at both system and component levels. Also, more complex probability distributions can be addressed.

In applying the maximum likelihood estimation method, it was noticed that the data corresponding to a HI value of 0.9 appeared a little as 'outliers', as shown for instance in Table 1 where the AIC value corresponding to s=0.9 does not follow the decreasing trend from 0 to 0.8. This may be explained by the fact that $s = 0.9$ is encountered after just a few cycles, i.e. corresponds to a small sample. Also, for the battery case (laboratory data) it was observed that the HI is not totally monotonic (Fig.6) and that fluctuations about the average (Fig.7 and Fig. 8) are more important than in the turbofan case. In general, the question of missing data points and outliers is the same as can be encountered in most data processing situations. Outlier detection methods, interpolation and imputation methods for missing points will help deal with such situations, much as in any statistical data processing situation.

We would like to point out that the variable degradation exponent model described in Section 3.3 provides several ways of handling increasing complexity. One way consists of incorportating more variability in the p exponent; another, of modifying q. As explained in Section 3.2, modifying p affects the speed of degradation, while modifying q impacts the end of life.

Real-world variability in the operating conditions could be incorporated into the $p(s)$ or q parameters by using Eq.(69):

$$HI(t) = 1 - \left(\frac{t}{\eta \left(-\log(q + q_{operational}) \right)^{\frac{1}{\beta}}} \right)^{p(s)} \tag{81}$$

or equivalently

$$HI(t) = 1 - \left(\frac{t}{\eta \left(-\log(q) \right)^{\frac{1}{\beta}}} \right)^{p(s)+p_{operational}} \tag{82}$$

Depending on the interpretation one can choose to incorporate the operational condition in the exponent (degradation shape) or the q value (change in population behaviour due to operational conditions) as shown in Figure 19. Both $p_{operational}$ and $q_{operational}$ are related via the following equation, which states the equality of $HI(t)$ as expressed by Eq. (81) or Eq.(82):

$$q_{operational} = e^{\left(-(-\log q)^{\frac{p_{operational}}{p}}+1\right)} - q \tag{83}$$

A third way of fitting the health index to a more-complex dynamic situation would of course be to complexify the power law by adding another term to the current HI definition.

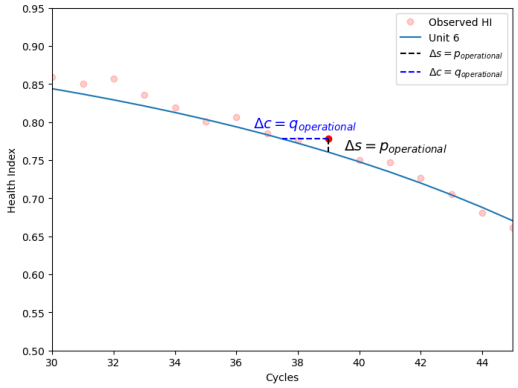


Figure 19. Residual of analytical curve and ground truth

Integration of Analytical Health Indices with Deep Learning for Hybrid Prognostics.

Recent advances in hybrid prognostics have highlighted the benefits of integrating reliability-based health modeling with deep learning architectures (Bajarunas et al., 2024). The probabilistic Health Index (HI) framework introduced in this work serves not only as an interpretable model of degradation but also as a structural prior that can be embedded into deep learning models, particularly autoencoders (AEs), to support constrained and meaningful latent representation learning.

In conventional unsupervised learning approaches, AE latent spaces are often optimized solely for reconstruction fidelity, leading to abstract embeddings that may not correlate with actual degradation. By contrast, our HI formulation, which models degradation trajectories using interpretable parameters such as initial condition, degradation rate, and reliability parameters of the failure time, introduces a health-aware constraint that can be directly incorporated into the AE training process. This can be implemented through latent anchoring strategies, where the AE is regularized to produce latent curves that resemble reliability-consistent HI progressions over time.

This integration not only increases the robustness and interpretability of the learned representations but also enables structured extrapolation for remaining useful life prediction. By calibrating the probabilistic HI parameters on inferred latent trajectories, future HI evolution can be analytically forecasted, offering uncertainty-aware RUL estimates even in data-sparse or label-scarce environments. Furthermore, since this analytical HI contains random coefficients, it can easily describe random variations from asset to asset in a fleet. Adaptation to different systems or different operating conditions is performed naturally through adaptation of the coefficients.

The feasibility and effectiveness of these hybrid strategies, integrating analytical health modeling with deep learning, have been evaluated in recent studies. Bajarunas et al. (Bajarunas et al., 2024) showed that integrating analytical HI structures with unsupervised learning yields improved interpretability and transferability across systems. Furthermore, (Goglio, Bajarunas, Dersin, Santos, & Arias Chao, 2024) explored the use of quantile regression built on HI trajectories to forecast RUL under uncertainty, providing an example of how these structured degradation representations can directly support predictive maintenance applications in aviation.

These results collectively support the argument that the proposed analytical HI modeling is not an isolated theoretical tool, but a foundation for robust hybrid prognostics systems. It enables deep learning models to retain flexibility while incorporating structural knowledge essential for reliability and decision-making in industrial contexts.

5. CONCLUSION AND PERSPECTIVES

This study has established an analytical framework for modeling health indices in cases where the time-to-failure follows a known probability distribution such as a Gamma, Weibull, or Pareto. By leveraging observed health index trajectories over time and specifically the failure times, we have derived an analytical form for the health index that is consistent with these observations. Additionally, we provided an analytical expression for the distribution of the time to reach any intermediate degradation level.

The key contributions of this work are summarized as follows:

- A systematic methodology to derive analytical, probabilistic HIs from TTF and degradation data.
- Closed-form derivations of the distribution of times required to reach intermediate degradation thresholds.
- Detailed analysis for widely-used TTF models, specifically Gamma, Weibull, and Pareto distributions.
- Demonstration of the methodology on public datasets, including turbofan engines and battery cells.
- Introduction of a heuristic extension to model heterogeneous degradation paths, late-stage acceleration, or maintenance events.

The availability of closed-form HI expressions is highly beneficial for predictive maintenance. In particular, it facilitates accurate remaining useful life estimation and enhances model explainability compared to black-box approaches that derive RUL estimates directly from sensor measurements. Moreover, once an analytical HI is established for a given asset, it can be adapted to similar assets or to the same asset operating under different conditions, reducing the need to re-learn the HI from scratch for each new dataset.

Looking ahead, several avenues for future work are promising. First, the current approach can be extended to incorporate other TTF distributions and alternative HI formulations by applying the general methodology outlined in Section 2.2. Second, integrating quantile regression techniques and extrapolating HIs from individual degradation trajectories (as discussed in Section 3.3) could further enhance prognostic accuracy. Finally, the analytical health index framework opens new opportunities for merging machine learning — particularly deep survival methods — with traditional reliability engineering, thereby advancing the state of the art in survival analysis and prognostics.

MAIN CONCEPTS

Symbol / Term	Definition
HI	Health Index
TTF	Time to Failure
MTTF	Mean Time to Failure
T	Random variable representing the TTF
$F(t)$	Cumulative Distribution Function: $P[T \leq t]$
$R(t)$	Reliability Function: $P[T > t]$
T_s	First hitting time of health level s
$h(t)$	Health Index at time t
B_q	q-percent Quantile
CI	Confidence Interval
MLE	Maximum Likelihood Estimation
AIC	Akaike Information Criterion

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723.

Arias Chao, M., Kulkarni, C., Goebel, K., & Fink, O. (2021). Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics. *Data*, *6*(1), 5.

Atamuradov, V., Medjaher, K., Camci, F., Zerhouni, N., Dersin, P., & Lamoureux, B. (2020). Machine health indicator construction framework for failure diagnostics and prognostics. *Journal of signal processing systems*, *92*, 591–609.

Bajarunas, K., Baptista, M., Goebel, K., & Arias Chao, M. (2023). Unsupervised physics-informed health indicator estimation for complex systems. In *Annual conference of the phm society* (Vol. 15).

Bajarunas, K., Baptista, M. L., Goebel, K., & Chao, M. A. (2024). Health index estimation through integration of general knowledge with unsupervised learning. *Reliability Engineering & System Safety*, *251*, 110352.

Bole, B., Kulkarni, C. S., & Daigle, M. (2014). Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use. In *Annual conference of the phm society* (Vol. 6).

Dersin, P. (2023). *Modeling remaining useful life dynamics in reliability engineering*. CRC Press,Taylor and Francis.

Dersin, P., Bajarunas, K., & Arias-Chao, M. (2024). Analytical modelling of health indices for prognostics and health management. In *Vol. 8 no. 1 (2024): Proceedings of the european conference of the phm society.*

Fink, O., Wang, Q., Svensén, M., Dersin, P., Lee, W.-J., & Ducoffe, M. (2020). Potential, challenges and future directions for deep learning in prognostics and health management applications. *Eng. Appl. Artif. Intell.*, *92*, 103678.

Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. de la Soc. Polonaise de Math.*.

Goglio, D., Bajarunas, K., Dersin, P., Santos, B. F., & Arias Chao, M. (2024). Enhancing predictive maintenance in aviation: A quantile regression approach for health index-based remaining useful life estimation. In *Proceedings of the 2nd international conference for condition-based maintenance in aerospace (iccbma).* Paris, France. (Presented at ICCBMA, 11–13 September 2024)

Hsu, C.-C., Frusque, G., & Fink, O. (2023). A comparison of residual-based methods on fault detection [Conference Paper]. In C. S. Kulkarni & I. Roychoudhury (Eds.), *Proceedings of the annual conference of the phm society 2023* (Vol. 15). s.l.: PHM Society. (15th Annual Conference of the Prognostics and Health Management Society (PHM 2023); Conference Location: Salt Lake City, UT, USA; Conference Date: October 28 - November 2, 2023) doi: 10.3929/ethz-b-000636893

Llera, A., & Beckmann, C. F. (2016). Estimating an inverse gamma distribution. *arXiv:1605.01019v2*.

Nachlas, J. (2017). *Reliability engineering- probabilistic models and maintenance methods, 2d edition*. CRC Press,Taylor and Francis.

Qin, Y., Yang, J., Zhou, J., Pu, H., & Mao, Y. (2023). A new supervised multi-head self-attention autoencoder for health indicator construction and similarity-based machinery rul prediction. *Advanced Engineering Informatics*, *56*, 101973. doi: https://doi.org/10.1016/j.aei.2023.101973

Ramos, P., Louzada, F., Ramos, E., & Dey, S. (2020). The fréchet distribution: Estimation and application-an overview. *Journal of Statistics and Management Sys-*

tems, *23*(3), 549-578.

Roman, D., Saxena, S., Robu, V., Pecht, M., & Flynn, D. (2021). Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence*, *3*(5), 447–456.

Sui, X., He, S., Vilsen, S. B., Meng, J., Teodorescu, R., & Stroe, D.-I. (2021). A review of non-probabilistic machine learning-based state of health estimation techniques for lithium-ion battery. *Applied Energy*, *300*, 117346.

Yang, F., Habibullah, M. S., & Shen, Y. (2021). Remaining useful life prediction of induction motors using nonlinear degradation of health index. *Mechanical Systems and Signal Processing*, *148*, 107183.

Ye, Z., & Yu, J. (2021). Health condition monitoring of machines based on long short-term memory convolutional autoencoder. *Applied Soft Computing*, *107*, 107379.

BIOGRAPHIES

Dr.Pierre Dersin: Dr. Dersin is currently Adjunct Professor at Luleå University of Technology (Sweden) in the Operation and Maintenance Engineering Division. He is also the president and founder of Eumetry sas, Louveciennes,France, a consulting firm in the fields of RAMS, PHM and AI. He holds a Ph.D. in Electrical Engineering from the Massachusetts Institute of Technology (MIT),as well as a Master's degree in Operations Research, also from MIT, and math and E.E. degrees from Université Libre de Bruxelles (Belgium). From 1990 to 2021, he was with Alstom Transport, France, where he occupied several technical and managerial positions, including RAM (Reliability-Availability-Maintainability) Director and RAM Master Expert,and founded the "RAM Center of Excellence".In 2015, he contributed to the launch of the predictive maintenance activity at ALSTOM and became PHM (Prognostics and Health Management) Director of ALSTOM Digital Mobility, St-Ouen, France. From 2014 to 2018, he was also the co-director of the joint Alstom-Inria Research Lab on Digital Mobility, and supervised several Ph.D. theses. Prior to joining Alstom, he was employed in the USA ,first as a research scientist in the US DOE large-scale system effectiveness program; and subsequently,with Belgian engineering firm Fabricom's US subsidiary, involved with fault detection and diagnostics in industrial systems. He has contributed a number of conference and journal papers in the fields of RAMS, PHM, automatic control, electric power systems, and AI. He was the keynote speaker at the 2014 European Conference of the PHM Society . Dr.Dersin is the author of the book " Modeling Remaining Useful Life Dynamics in Reliability Engineering" , CRC Press, Taylor and Francis, June 2023. His current interests focus on the confluence between RAMS and PHM as well as complex systems resilience and asset management.

Dario Goglio: Dario Goglio is a PhD candidate at Delft University of Technology (Netherlands) and Zurich University of Applied Sciences (Switzerland). He holds an MSc in Mechanical Engineering from ETH Zurich and is working for Swiss International Air Lines as an Aircraft Data Analytics Engineer, where he gains valuable domain knowledge and the opportunity to apply his research in real-world scenarios. His research focuses on developing hybrid models for prognostics using multimodal domain knowledge available to airlines.

Kristupas Bajarunas: Kristupas Bajarunas is a Ph.D. candidate at Delft University of Technology (Netherlands) and Zurich University of Applied Sciences (Switzerland). He holds an M.Sc. in Machine Learning from the Royal Institute of Technology (Stockholm). His current research is concerned with developing generic hybrid models for prognostics.

Dr. Manuel Arias Chao: Dr. Arias Chao is a Senior Lecturer at Zurich University of Applied Sciences and an Assistant Professor at the Operations and Environment Chair of the Delft University of Technology in the Netherlands. He has a PhD in Physics-informed Machine Learning for Prognostics and Health Management from ETH Zurich, a Master's degree in Thermal Power from Cranfield University, and a Bachelor's degree in Aeronautical Engineering from the Technical University of Madrid. Manuel has gained valuable industrial and research experience as a visiting researcher at the Diagnostics & Prognostics Group at NASA Ames, Thermodynamics & Performance Lead Engineer at General Electric and ALSTOM Power, and Aero Engine Maintenance Engineer at ITP. In his current role, Manuel also co-leads the Expert Group Smart Maintenance from the Swiss Alliance for Data-Intensive Services. He focuses on teaching & research for a broad range of application fields, including power generation, marine and aircraft propulsion, and manufacturing equipment.