

Impact of Early Life Failures in Services of Engineering Asset Fleets

Arinan Dourado¹ and Felipe A. C. Viana²

^{1,2}*University of Central Florida, Orlando, FL, 12309, USA*

arinandourado@knights.ucf.edu

viana@ucf.edu

ABSTRACT

Services and warranties of large fleets of engineering assets is a very profitable business where original equipment manufacturers and independent service providers offer contracts designed to cover events in day-to-day service as well as major maintenance and repairs over the life of the asset. Accurate reliability modeling, as a way to understand how the complex stochastic interactions between operating conditions and component capability define useful life, is key for services profitability. The modeling task is daunting as factors such as aggressive mission mixes introduced by operators, exposure to harsh environment, inadequate maintenance, and problems with mass production (bad batch of materials) can lead to large discrepancies between designed and observed useful lives. This paper is focused on how to quantify the impact of infant mortality in fleets of industrial assets. A simple numerical experiment is used to address the fundamental question: how does number of observations and fleet size interact with each other in fleet management? The results demonstrate that material capability, penetration of bad batch of material in the fleet, and commissioning time can drastically influence fleet unreliability. Moreover, infant mortality due to manufacturing problems/material capability is a manifestation of an outlier problem. As a consequence, the propensity to observe first failures depend on the actual fleet size. Since failure observations are used to build/update the reliability models, small fleet operators have to deal with large uncertainties when quantifying infant mortality. This impacts their ability to make provisions for service and maintenance (inventory, labor, loss of productivity, etc.). Although the large number of failure observations causes a financial burden in large fleet operators, it also allows for reduced uncertainty in building/updating the reliability models. In turn, this improves their ability to forecast future failures and make provisions for service and maintenance.

Arinan Dourado. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. INTRODUCTION

Managing fleet reliability of industrial equipment is a very profitable business that focuses on services, maintenance and warranties. Contracts carefully designed to accommodate minor to major maintenance over the life of the asset are usually sold by both large original equipment manufacturers (the interested reader is referred to GE Aviation collaborators, 2017, and Siemens collaborators, 2017 for some examples), as well as independent service providers (the interested reader is referred to Gemini Energy Services collaborators, 2017, and Lufthansa Technik AG collaborators, 2017 for some examples). A key capability for effective operations and maintenance is reliability modeling, as it defines the ability to comprehend hardware degradation and predict remaining useful life. This gives to operators the chance to make decisions that directly impact their financial outcomes through asset performance and availability levels, operation safety, etc.

Modern approaches to reliability modeling of industrial equipment take full advantage of physics through the understanding of machine design, materials, and manufacturing, as well as high-fidelity computational models (Bogdanoff and Kozin, 1985; Johnson and Hillberry, 2004; Kapur and Pecht, 2014; Rao, 1992; Rausand and Hoyland, 2004; and Stephens et al. 2000). Despite of all the effort in design and quality control, when looking into a large fleet of assets (hundreds to thousands of units), it can happen that observed machine performance and hardware reliability deviate from design intent (Al-Dahidi et al. 2016 and Volponi, 2014). Such deviation is usually attributed to one or a combination of the following:

- Aggressive missions (duty cycles) and mission mixes introduced by operators.
- Problems related to mass production such as a bad batch of materials, poor quality control of a specific vendor, assembly, etc.
- Exposure to harsh environmental conditions (seasonality, corrosives, dust, etc.).

- Inadequate services and maintenance practices.

Infant mortality is a major concern among original equipment manufacturer and operators of industrial assets. It always increases cost of ownership (maintenance, warranty, services, etc.). It can reduce asset performance and availability. In addition, it can impose difficulties in meeting compliance and regulations standards (as hardware degradation can be a lead cause of safety standard infringements, elevation of noise and emission levels, etc.).

This work aims at presenting a probabilistic analysis for characterization of emerging fleet issues due to infant mortality. We will focus on answering the fundamental question: *how does fleet size and number of failures interact with each other when characterizing an infant mortality problem?* We answer this question using prognosis, uncertainty quantification, reliability and fleet management.

We use physics-based prognosis as a way to forecast remaining useful life through progression of hardware distress by fusing design, manufacturing, and services information. There is recent debate between data-driven and physics-based models for prognosis (not the focus of the current work). The interested reader is referred to Baraldi et al. (2013) and Dawn et al. (2015) for further discussion. When using physics-based approaches, one has to focus on quantifying uncertainty in model form, model parameters, and data. Jiang et al. (2013) discussed the issue of bias correction, with systematic error being corrected by using a statistical model for the bias term (e.g., a Gaussian process) calibrated with actual experimental data, or through high-fidelity simulations. Peherstorfer et al. (2017) reviewed strategies for handling multifidelity models when performing computationally intensive uncertainty quantification. Asher et al. (2017) and Coppe et al. (2011) discussed how to calibrate important parameters in fatigue crack growth applications (initial flaw size and crack growth parameters). In real world applications, it is also very common that models are updated while data is gathered for a particular instantiation (asset of interest). Li et al. (2016) discussed the use of dynamic Bayesian networks for model updating with observed data (including loads). The updated model was used in diagnosis and prognosis of an aircraft digital twin.

Accurate prognosis models are at the core of fleet management. For example, Pattabhiraman et al. (2012) discuss how models that are constantly updated using sensors installed in aircraft structures can aid in condition-based maintenance. Authors showed that scheduled interval-based maintenance can be safely avoided depending on model predictions, which directly impact cost of ownership. A similar outcome is also the target of the work by Ling et al. (2017), where information gain theory is used to evaluate the usefulness of aircraft component inspection (which helps

deciding whether inspection is worthwhile or not). A dynamic Bayesian network tracks and forecasts fatigue crack growth and the detection of a crack is modeled through probability of detection. Information gain per cost of inspection is used to identify the optimal option for the next inspection in the future. Haddad et al. (2011) and Haddad et al. (2012) discussed a cost-benefit-risk approach to manage the actions to be taken following a prognostic model. The discussion included important aspects such as overall maintenance (cost of unscheduled maintenance, collateral damage during repair, fault isolation), shortening of remaining useful life, spare parts management, etc.. Applications discussed included electronic systems in commercial aircraft and gearbox maintenance in wind farms.

The remaining of the paper is organized as follows. Section 2 describes the case study that will illustrate the issue of infant mortality in fleet reliability. Section 3 presents and discusses the numerical results. Finally, section 4 closes the paper recapitulating salient points and presenting concluding remarks and future work.

2. CASE STUDY: INFANT MORTALITY IN A FLEET OF ASSETS DUE TO BAD BATCH OF MATERIALS

We use a simple numerical experiment to study how fleet size and number of failures impact the characterization of infant mortality in fleets of assets. We consider a component made out of the Al 2024-T3 alloy* subjected to alternating loads and assume that initiation cycles dominate fatigue life. This hypothetical component can be **mission critical** (its failure does not affect directly safety of asset operation). We use readily available S-N curves commonly found in material handbooks (MMPDS collaborators, 2017) to model low cycle fatigue life at different average and alternating stress levels. Then, we mimic problems with manufacturing (bad batch of materials) by degrading the S-N curves. We designed two missions and two mission mixes to emulate variation due to customer behavior. Finally, we simulate different fleet sizes to understand how failure observations affect overall fleet reliability and detection of emerging issues.

2.1. Damage accumulation at the component level

We used the readily available S-N curves illustrated in Figure 1 and the following suggested fatigue life model as a function of equivalent stress (MMPDS collaborators, 2017):

$$\begin{aligned}
 N_f &\sim \text{logNormal}(\mu_{N_f}, \sigma_{N_f}) \\
 \mu_{N_f} &= \theta_1 \log(S_{eq} + \theta_2) + \theta_3 \\
 S_{eq} &= \left(S_{max} (1 - (2 S_{mean} - S_{max}))^{\theta_4} \right) \\
 \theta_1 &= -3.33, \theta_2 = -12.3, \theta_3 = 9.2, \theta_4 = 0.68 \\
 \theta_5 &= \sigma_{N_f} = 0.89
 \end{aligned} \tag{1}$$

* Al 2024-T3 is commonly used in aircraft fuselage, flaps, trim tabs, servo tabs and control surfaces.

where:

- N_f is the fatigue life,
- μ_{N_f} and σ_{N_f} are the parameters of the fatigue life lognormal distribution,
- S_{eq} is the equivalent stress of a given load cycle,
- S_{mean} and S_{max} are the mean and maximum stress of a given load cycle, and
- θ_1 to θ_5 are calibration parameters.

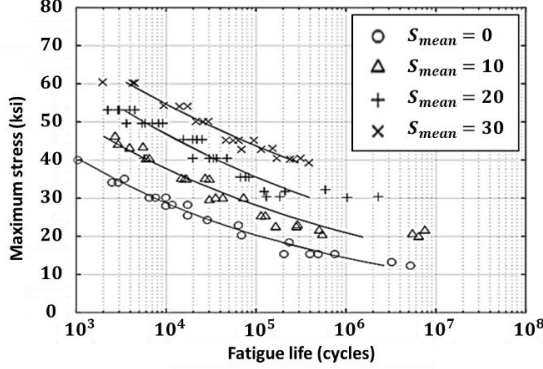


Figure 1. S-N curves for the Al 2024-T3 alloy.

Damage is accumulated following Palmgren-Miner's rule

$$D = \sum \Delta d_i = \sum \frac{n_i}{N_f^{(i)}}, \quad (2)$$

where:

- D is the damage accumulated throughout the life of the component
- Δd_i is the damage accumulated by running n_i cycles at the i th load level
- n_i is the number of cycles run at the i th load level (uniquely defined by mean and maximum stress).
- $N_f^{(i)}$ is the fatigue life at the i th load level, and
- the threshold for end of life is $d_{TH} = 1$.

Since fatigue life $N_f^{(i)}$ follows a lognormal distribution, D is a random variable with no closed-form expression for its probability density function. However, considering that all Δd_i have the same variance, then, D can be approximated by

$$D \sim \logNormal(\mu_D, \sigma_D),$$

$$\sigma_D^2 = \ln \left[\frac{e^{\sigma_{N_f}^2} - 1}{\left(\frac{\sum e^{2\mu_{\Delta d_i}}}{(\sum e^{\mu_{\Delta d_i}})^2} + 1 \right)} \right], \quad (3)$$

$$\mu_D = \ln \left(\sum e^{\mu_{\Delta d_i}} \right) + \frac{\sigma_{N_f}^2}{2} - \frac{\sigma_D^2}{2}, \text{ and}$$

$$\Delta d_i \sim \logNormal \left(-\mu_{N_f^{(i)}} + \ln(n_i), \sigma_{N_f} \right) \quad (4)$$

Since damage is accumulated after each mission, for a given component, the number of missions to failure (MTF) is a random variable with cumulative density function defined by $F_{MTF}(MTF = m) = \Pr[D^{(m)} \geq 1]$, where $D^{(m)}$ is the damage accumulated up to m missions. This implies that component reliability $R(m)$ and unreliability $Q(m)$ at mission m are given by

$$R(m) = 1 - \Pr[D^{(m)} \geq 1], \text{ and} \quad (5)$$

$$Q(m) = 1 - R(m) = \Pr[D^{(m)} \geq 1].$$

We designed the two missions shown in Figure 2 and two mission mixes detailed in Table 1. At any given mission, the load history can be modeled with the mission index M_{IDX} , which follows a Bernoulli distribution with probability θ_6 :

$$LH^{(m)} = \begin{cases} \text{Mission 1,} & \text{if } M_{IDX} = 1 \\ \text{Mission 2,} & \text{if } M_{IDX} = 2 \end{cases}$$

$$M_{IDX} \sim \text{Bernoulli}(k = M_{IDX}, p = \theta_6), \text{ and} \quad (6)$$

$$f_{M_{IDX}}(M_{IDX}, \theta_6) = \begin{cases} \theta_6, & \text{for } M_{IDX} = 1 \\ 1 - \theta_6, & \text{for } M_{IDX} = 2 \end{cases}$$

where:

- $LH^{(m)}$ is the load history for mission m ,
- M_{IDX} is index that defines which mission to assign,
- $f_{M_{IDX}}(M_{IDX}, \theta_6)$ is the probability mass function for the Bernoulli variable M_{IDX} with probability θ_6 , and
- θ_6 is a calibration parameter

This way, when variations due to both loads in the form of mission mixes, and material capability (spread in S-N curves) are considered, the distribution of fatigue life is illustrated in Figure 3.

We emulate debit in material capability by shifting S-N curves to the left (i.e., for the same stress level, the material has a shorter fatigue life as compared to the nominal material). This way, we model relatively the large deviation caused by problems during manufacturing (such as problems in surface treatment and/or microstructure). In other words

$$D_{Debit} \sim \logNormal((1 - \theta_7)\mu_D, \sigma_D) \quad (7)$$

where θ_7 is a calibration parameter (defining the debit in material capability).

Figure 4 illustrates the effects of considered material capability debit on the fatigue life distribution for the aggressive mission mix.

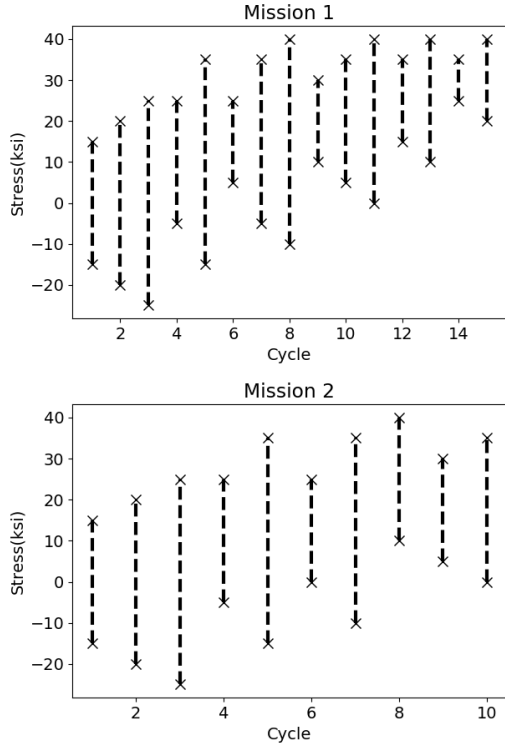


Figure 2. Alternating stress levels (S_{Min} to S_{Max}) for the two designed missions. At the end of each mission, the accumulated damage is distributed around 2.63×10^{-4} and around 6.55×10^{-5} for mission 1 and 2, respectively. The 50th percentile of fatigue life is approximately 3,800 and 15,260 missions for missions #1 and #2, respectively.

Table 1. Mission mix formulation. Every asset in the fleet is expected to operate at or between, or even alternating between, aggressive and mild mission mixes.

	Aggressive	Mild
Mission 1	50%	85%
Mission 2	50%	15%
50 th prctl N_f	9230	13360

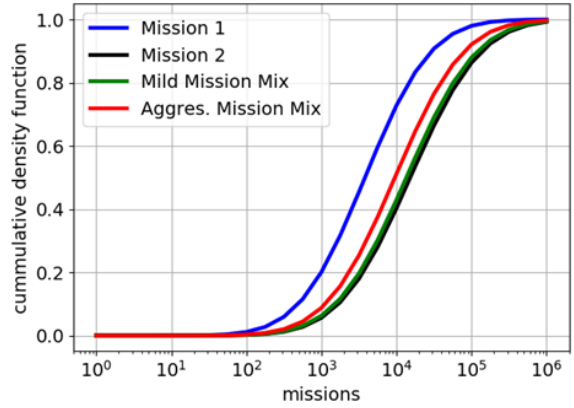


Figure 3. Fatigue life distribution in terms of missions to failure considering both load (mission mix) and material capability variations (spread in SN curve). Mission 1 (aggressive) and mission 2 (mild) bound the life distributions for the any mission mix.

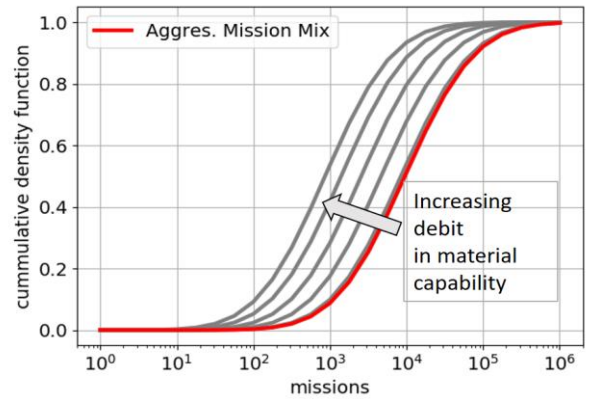


Figure 4. Fatigue life distribution for aggressive mission mix and different levels of material capability. At the highest debit considered (20%), the median of missions to failure can be reduced from 9230 to 860 missions.

2.2. Fleet commissioning, reliability, and failure observations

Large fleets of assets are usually commissioned over a period of time (as production follows a backlog of orders, commissioning ramps up for a while before it starts to decline). Commissioning schedule determines the number of units running (and as a consequence, it impacts the number of failure observations). In this study, we arbitrarily model commissioning time through a truncated Gaussian distribution (as illustrated by Figure 5). In real life, this distribution is first estimated based on market analysis and can be updated as units are sold and commissioned. Different commissioning time makes the units across the fleet to have different accumulated service lives (and damage, as a consequence).

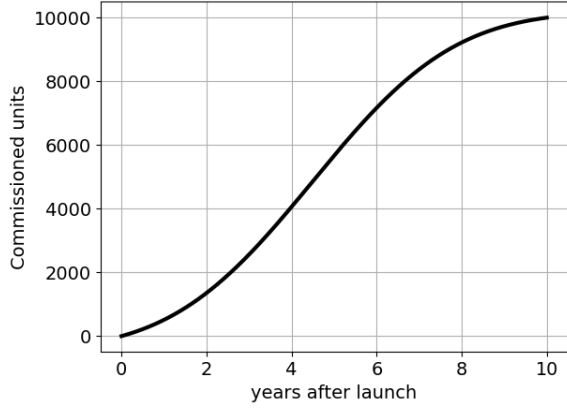


Figure 5. Commissioned units over time. In reality, commissioning is controlled by backlog of orders; here, we assume commissioning time after product launch $T_C \sim N(\mu = 4.5, \sigma = 2.625)$ and $0 \leq T_C \leq 7$ (which implies in fleet size of 10,000 units at year 10).

After commissioning, we assume that each unit runs one mission per day. Integration of asset unreliability, for pristine and material with debit in capability, up to fleet reliability is straightforward:

$$Q_{Pristine}^{(t)} = \frac{\sum_{i=1}^{N_{Fleet}} Q_i^{(t)} |_{Pristine}}{N_{Fleet}}, \text{ and} \quad (8)$$

$$Q_{Debit}^{(t)} = \frac{\sum_{i=1}^{N_{Fleet}} Q_i^{(t)} |_{Debit}}{N_{Fleet}},$$

where:

- $Q_{Pristine}^{(t)}$ and $Q_{Debit}^{(t)}$ are the fleet unreliability considering nominal material capability and material with certain capability debit, respectively, both at time t .
- $Q_i^{(t)} |_{Pristine}$ and $Q_i^{(t)} |_{Debit}$ are the unreliability of component i assuming it is made of pristine material and material with certain capability debit, respectively, both at time t . $Q_i^{(t)} |_{Pristine}$ and $Q_i^{(t)} |_{Debit}$ only start being computed if the unit is already commissioned (otherwise both are null).
- N_{Fleet} is the final fleet size.

As consequence, at time t , fleet unreliability $Q_{Fleet}^{(t)}$ is:

$$Q_{Fleet}^{(t)} = \theta_8 \times Q_{Debit}^{(t)} + (1 - \theta_8) \times Q_{Pristine}^{(t)}, \quad (9)$$

where θ_8 is a calibration parameter that defines the penetration of units in the fleet (in terms of fraction of the fleet) made of material with a certain debit level in capability.

With fleet unreliability we can predict number of failures at any year after product launch by using the binomial distribution to model number of failures:

$$f_{N_{Fail}}(n = N_{Fail}; p = Q_{Fleet}^{(t)}, N = N_{Fleet}) = \binom{N_{Fleet}}{N_{Fail}} \times Q_{Fleet}^{(t)^{N_{Fail}}} \times (1 - Q_{Fleet}^{(t)})^{N_{Fleet} - N_{Fail}}, \quad (10)$$

where:

- $f_{N_{Fail}}(\cdot)$ is the probability density function of the binomial distribution,
- N_{Fail} is the number of observed failures at year y ,
- $Q_{Fleet}^{(t)}$ is the fleet unreliability at year y , and
- N_{Fleet} is the fleet size at year y .

Conversely, given a number of observed failures, we can estimate the calibration parameters through the Bayes rule:

$$f_{Post}(\theta | N_{Fail}, N_{Fleet}) \propto f_{N_{Fail}}(N_{Fail} | Q_{Fleet}^{(t)}, N_{Fleet}) f_{Prior}(\theta), \quad (11)$$

where $Q_{Fleet}^{(t)}$ is a function of θ . We write the posterior in its proportional form as it is the way it is implemented in most numerical integration methods (such as Markov chain Monte Carlo). Also, computing the binomial coefficient $\binom{N_{Fleet}}{N_{Fail}}$ is not necessary when estimating fleet unreliability given a number of failures for a fleet. This is usually cumbersome and can cause numerical ill-conditioning depending on N_{Fleet} and N_{Fail} .

2.3. Fleet management

We build the fleet management model out of two Bayesian networks, one for the asset reliability and another one for the fleet unreliability. Figure 6 shows the asset-specific dynamic Bayesian network that relates material properties and loads with damage accumulation. LH stands for load history, S_{eq} is the equivalent stress of a given load cycle, μ_{Nf} and σ_{Nf} are parameters of the fatigue life lognormal distribution, $\Delta d_i^{(t)}$ is the damage accumulated after running through $LH^{(t)}$, and $D^{(t)}$ is the damage accumulated up to t . The full-blown model has seven calibration parameters θ_1 to θ_7 . Here, we freeze θ_1 to θ_5 (parameters defining material properties) to the values given by the MMPDS, as shown in Eq. (1). θ_6 (parameter defining the mission mix) will also be fixed at the values shown in Table 1. θ_7 , which defines the debit in material capability, will be calibrated with failure observations (as discussed in Section 3).

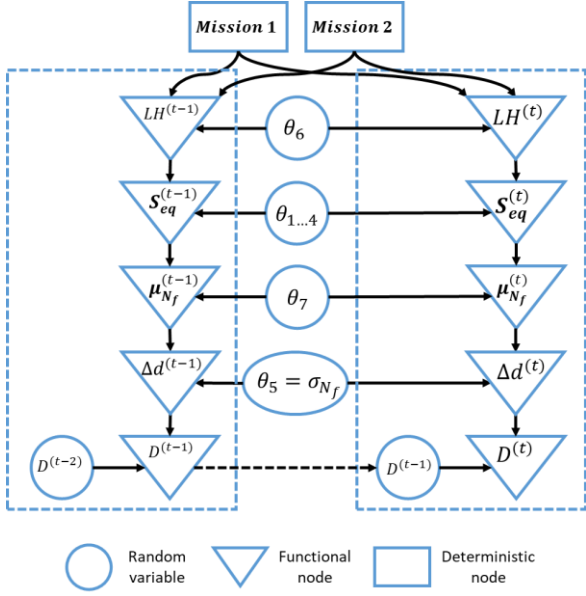


Figure 6. Asset-specific dynamic Bayesian network. Superscripts $(t-2)$, $(t-1)$, and (t) indicate the time stamps in which inference/estimation is performed. Load histories for missions 1 and 2 are defined in Figure 2.

Figure 7 shows the fleet dynamic Bayesian networks. $\mathbf{D}_{1...N_{Fleet}}^{(t)}|_{Debit}$ and $\mathbf{D}_{1...N_{Fleet}}^{(t)}|_{Prestine}$ are the vectors of damage accumulated up to t for the N_{Fleet} assets in the fleet with and without debit in material capabilities, respectively. $\mathbf{Q}_{1...N_{Fleet}}^{(t)}|_{Debit}$ and $\mathbf{Q}_{1...N_{Fleet}}^{(t)}|_{Prestine}$ are the vectors of unreliability at time t . $Q_{Debit}^{(t)}$ and $Q_{Prestine}^{(t)}$

represent fleet unreliability at time t . θ_8 is the fraction of the fleet made of material with a certain debit level in capability. One more calibration parameter is added to the list. θ_8 , penetration of material with low capability, will also be calibrated with failure observations. Both dynamic Bayesian network models are used to make inference about θ_7 and θ_8 , as well as estimate and forecast fleet unreliability $Q_{Fleet}^{(t)}$. With estimated/forecasted fleet unreliability $Q_{Fleet}^{(t)}$, one can model, estimate/forecast $N_{Fail}^{(t)}$, the number of failures at time t , through a binomial distribution

$$N_{Fail}^{(t)} \sim Bin(Q_{Fleet}^{(t)}, N_{Fleet}), \quad (12)$$

In this contribution, we study the effects of fleet sizes in the ability to forecast the number of failures and its implication to fleet management. From Eqs. (11) and (12), it is expected that inference performed with data from small fleets will result in large uncertainty about the calibration parameters. This is problematic as the calibration parameters are then used to estimate and forecast the number of future failures. Large uncertainty in number of future failures drive conservativeness in the way operators manage their fleets. On the other hand, large operators, large services and maintenance companies, and original equipment manufacturer tend to observe a large number of failures and should be able to benefit from it in terms of uncertainty quantification. Regardless of the fleet size, effective fleet management asks for continuous model update as new information is made available throughout service lives (including revisiting the assumptions about model form, failure modes, etc.).

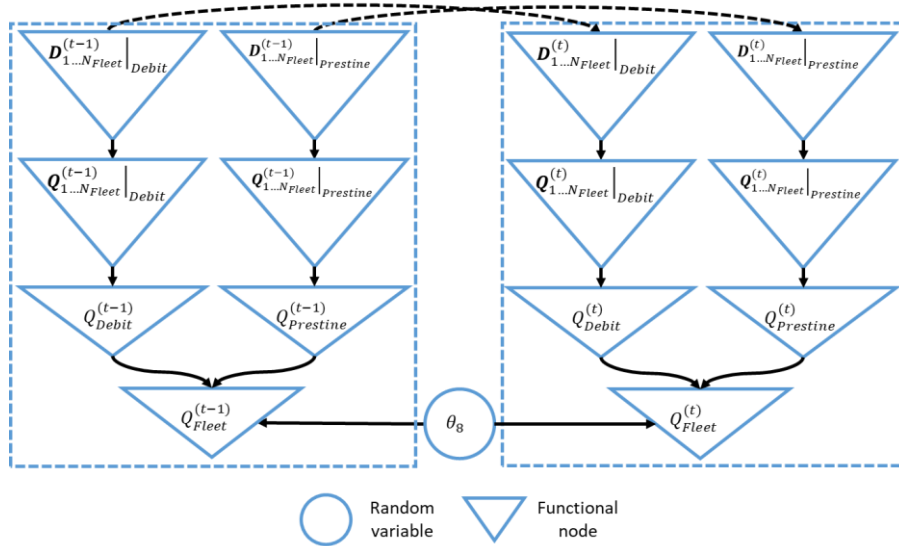


Figure 7. Fleet dynamic Bayesian networks. Superscripts $(t-1)$, and (t) indicate the time stamps in which inference/estimation is performed. Figure 6 illustrates the dynamic Bayesian network that models damage accumulation at the asset level.

The estimated number of failures can be used to build a risk metric associated with the forecast. One very straightforward measure of risk is the uncertainty about the forecasted number of failures (i.e., companies have to be prepared to absorb that variation from a financial perspective). There are a number of ways to quantify variation in number of failures. One can simply use the standard deviation, which might not be convenient given the asymmetric nature of the N_{Fail} estimator. Alternatively, risk can be defined as the difference between the 97.5 and 2.5 percentiles of the forecasted number of failures. Small operators can use this range to support the decision to either self-perform or buy a contractual service agreement from a third party company. Small operators tend to have difficulties in absorbing large variations in forecasted number of failures due to liability associated with it (both in terms of inventory, labor, etc., as well as in terms of loss of revenue). This can make small operators to be over zealous and perform excessive inspection and services in the hope to prevent costly maintenance or catch serious problems when units are still under manufacturer warranty (minimizing impact of unscheduled removals and cost of repairs/replacements). For large fleet operators, the problem shifts from unexpected downtime to excessive number of costly maintenance and contractual obligations regarding availability and reliability.

We compare results from a small and a large fleet to mimic the small operator vs large service provider dynamic. For this portion of the case study, we assume that the large service provider is more likely to provide an unbiased and accurate estimation of the number of failures. If that is the case, the difference in forecasted number of failures can help us judge whether self-performing is a good decision or not. Mathematically

$$RISK = N_{Fail@Small}^{Small} - N_{Fail@Small}^{Large} \quad (13)$$

where $N_{Fail@Small}^{Large}$ and $N_{Fail@Small}^{Small}$ are the estimated/forecasted number of failures on the small fleet coming from the large service provider and the small fleet operator models, respectively.

$RISK$ becomes an indicator of whether the small fleet operator is likely to save or lose money by self-performing services and maintenance:

- $RISK > 0$: small operator over predicts failures, which drives allocation of more resources than needed. In other words, the behavior is conservative and it translates in savings due to avoided unscheduled maintenance, reduced downtime, etc.
- $RISK < 0$: small operator under predicts failures, which drive allocation of less resources than needed (operator loses money due to unscheduled maintenance, downtime, etc).

Again, this assumes that $N_{Fail@Small}^{Large}$ is an accurate estimator of N_{Fail}^{True} . Although N_{Fail}^{True} can be obtained in this numerical example (through Eq. (12) since fleet reliability for known loads can be obtained at any point in time), we avoid using it as it is not available in real life though.

3. RESULTS AND DISCUSSIONS

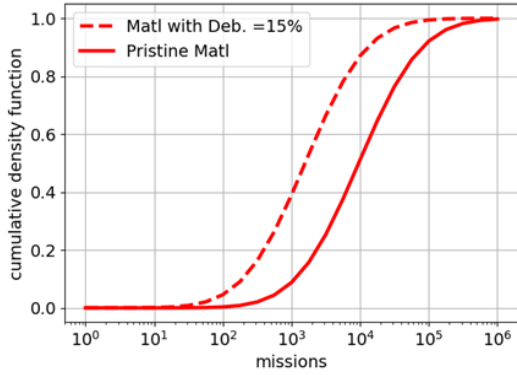
In order to evaluate the effect of fleet size in number of failure observations, we defined two distinct fleets:

- a large fleet of 10,000 units: emulating an original equipment manufacturer or a large service provider, and
- a small fleet of 1,000 units: emulating a small fleet operator. This units come from the larger 10,000 unit fleet, which also means that the large fleet operator has visibility into what happens with this small fleet.

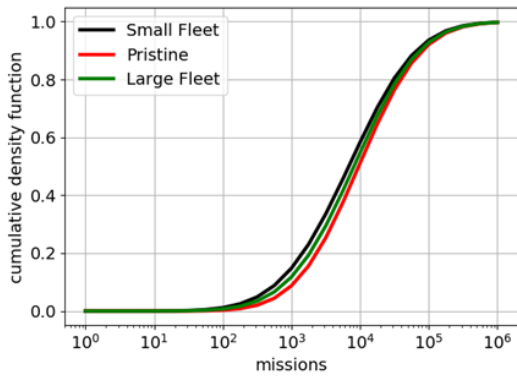
Both fleets are plagued with a material debit level of 15%. However, to make things more interesting, we distributed the failures across the fleet such that the small fleet operator has a penetration of 20% of units plagued with material of inferior capability (i.e., 200 out of 1,000 units are plagued), while the larger fleet has an overall 10% penetration (i.e., 1,000 out of 10,000 units are plagued). The implications in fatigue life distribution are shown in Figure 8.

As discussed in section 2.2, commissioning has an effect on fleet unreliability as the fleet grows bigger with asynchronous aging. As an illustration, Figure 9 shows a comparison between fleet unreliability over time with and without the effect of commissioning. The drastic reduction in unreliability values result in a delay in rising failures observations. Most industrial engineering assets (focus of this paper) are commissioned over a period of time. **In the reminder of this section, we will discuss the results following commissioning detailed in section 2.2.** The interested reader can find the case of simultaneous commissioning of the entire fleet in the appendix.

With the fleet unreliability over time, we can forecast the number of failure. Figure 10 highlights the contribution of each subpopulation by material type (pristine and with debit in capability) in the resulting failure observations. Besides the obvious penetration of material with poor capability (10% versus 20% for the large and small fleet, respectively), commissioning also affects the relative contribution of each material to the number of failures. Early on, most failures come from components plagued by material with poor capability. Over time, the unreliability for pristine material increases (see Figure 9), and the relative contribution of each population starts to change. Around the 3rd year after the product launch, at least for the large fleet, failures are dominantly coming from components made out of pristine material (although contribution from subpopulation with plagued material is still substantial).



(a) Impact of material debit in fatigue life when entire fleet is plagued.



(b) Pristine (expected) and actual (large and small) fleet unreliability at aggressive mission mix.

Figure 8. Fleet characterization of fatigue life distribution (without the effects of commissioning). Load history is exclusively coming from aggressive mission mix. There is considerable shift in fatigue life if the entire fleet is plagued with material of poor capability. Nevertheless, given the 10% and 20% penetration levels for the large and small fleet, respectively, the effect mostly manifested in the lower tail of fatigue life distribution.

At the 3rd year after product launch, we assume the following number of failure observations:

- Small fleet: 17 failures (lower tail of predicted number of failures).
- Large fleet: 127 failures (roughly 50th percentile of predicted number of failures). Obviously, the small fleet failures are contained in this set.

We use these failure observations and fleet unreliability (from known load histories) to calibrate:

- θ_7 : debit in material capability with uniform prior between 1% to 30%, and

θ_8 : penetration of units with poor material capability in the fleet with uniform prior between 0.01% and 20%.

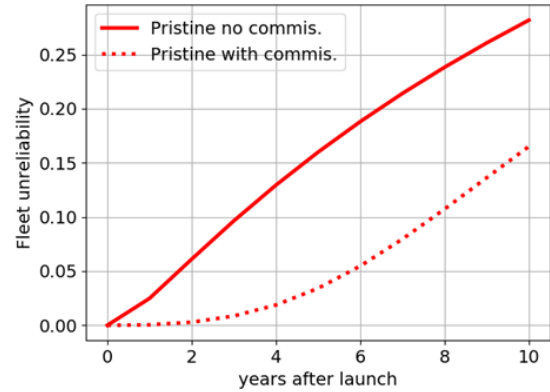
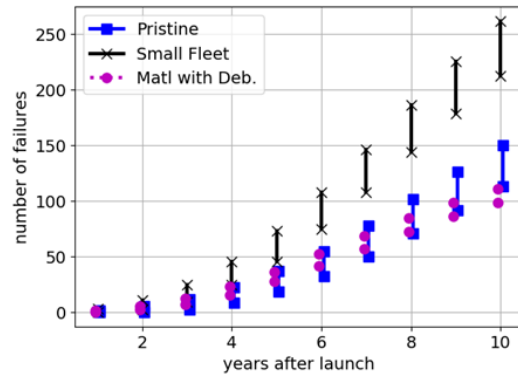
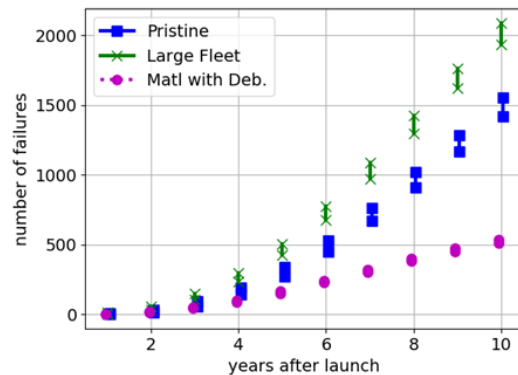


Figure 9. Commissioning effect in overall fleet unreliability. Asynchronous fleet aging is manifested in delayed increase in fleet unreliability, which delays failure observations.



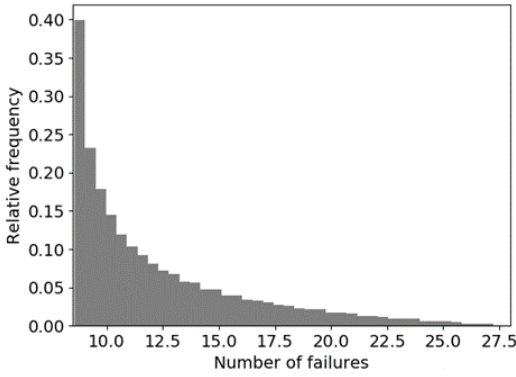
(a) Small fleet.



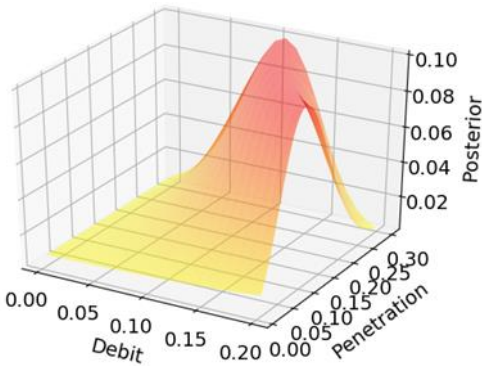
(b) Large fleet.

Figure 10. Estimated number of failure in each fleet. Error bars represent the 95% prediction intervals. Two factors contribute to the breakdown of failure observations. First, the asynchronous aging due to commissioning, and second the difference in penetration of material with poor capability. In the small fleet, 20% of units are plagued as opposed to only 10% in the large fleet.

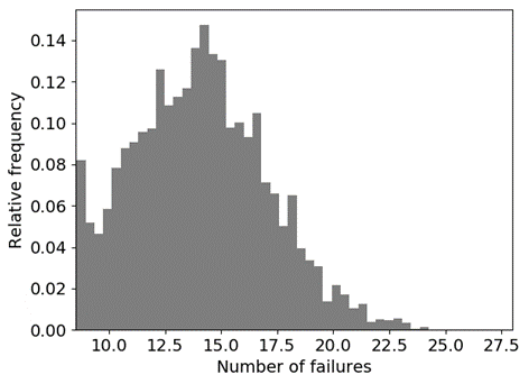
Figure 11 and Figure 12 detail the calibration results with regards to both calibration parameters and estimated number of failures for the small and large fleets, respectively. Even for small fleet operator, there is considerable uncertainty reduction and failure estimates are much improved as compared with non-informative priors.



(a) Prior distribution of number of failures.

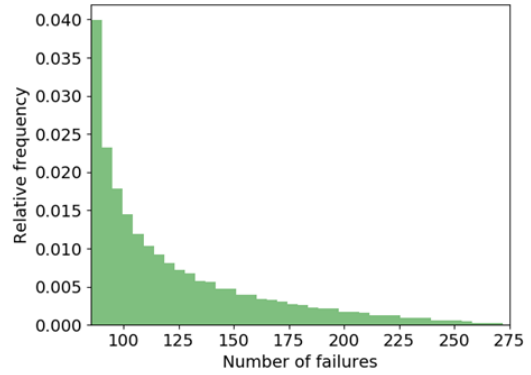


(b) Posterior distribution of calibration parameters.

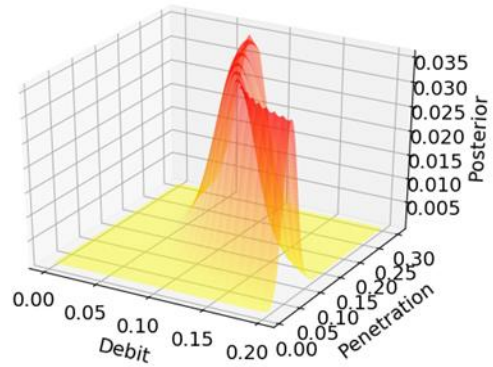


(c) Posterior distribution of number of failures.

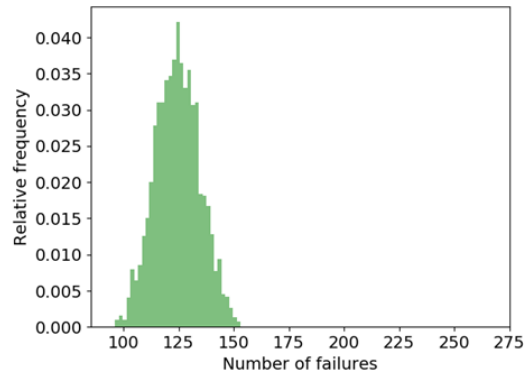
Figure 11. Calibration results for the small fleet. Failure observations at the 3rd year after product launch and uniform priors feed the Bayesian update.



(a) Prior distribution of number of failures.



(b) Posterior distribution of calibration parameters.



(c) Posterior distribution of number of failures.

Figure 12. Calibration results for the large fleet. Failure observations at the 3rd year after product launch and uniform priors feed the Bayesian update.

As we mentioned before, the large fleet operator has full visibility into what happens with the small fleet. In this numerical example, the relative number of failure with respect to the fleet size can be used to map the posterior distribution of number of failures at the large fleet into the small fleet, as illustrated by Figure 13. Table 2 summarizes the estimates regarding the number of failure.

The models with updated calibration parameters can be used to forecast the number of failures over time. Figure 14 shows how these forecasted values look like for the small fleet. The uncertainty in posterior distribution of calibration parameters for the small fleet model, Figure 11-(b), is larger than the one for the large fleet model, Figure 12-(b). The result is the larger uncertainty that the small fleet model exhibits when compared to the large fleet model.

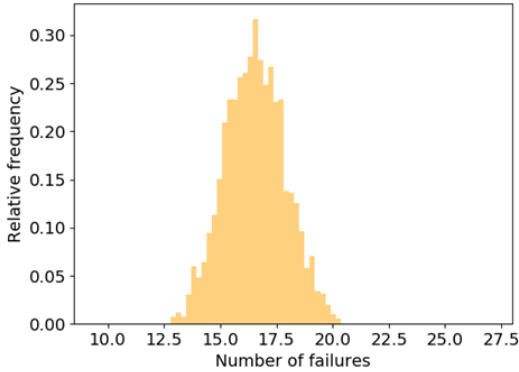


Figure 13. Posterior distribution of number of failures at the 3rd year after product launch for small fleet as estimated by the large fleet operator. When compared to Figure 11-(c), the uncertainty in Figure 13 is much smaller and clearly attributed to the much richer information available at the large fleet level.

Table 2. Estimates for number of failures at the 3rd year after product launch for small fleet coming from updated models. Small fleet observed 17 failures.

Model	Percentile		
	2.5	50	97.5
Small fleet	4	14	30
Large fleet	11	16	22

Once the infant mortality issue is quantified, operators undergo a number of risk mitigation actions to reduce costs associated with unscheduled maintenance, asset unavailability, etc. Since this numerical example explores failure mode due to a manufacturing problem, it is hard to identify the problem at an asset level purely by looking at operation (i.e., through sensors and performance). A massive fleet-wide inspection can be considered, but it can be costly due to fleet size and associated downtime. Another option is to recommission the fleet by changing the mission mix to a mild one. This can also be costly, as mild mission mixes are usually associated with some loss in performance or productivity. In this study, we show results for recommissioning and leave the investigation of inspection and maintenance for future work.

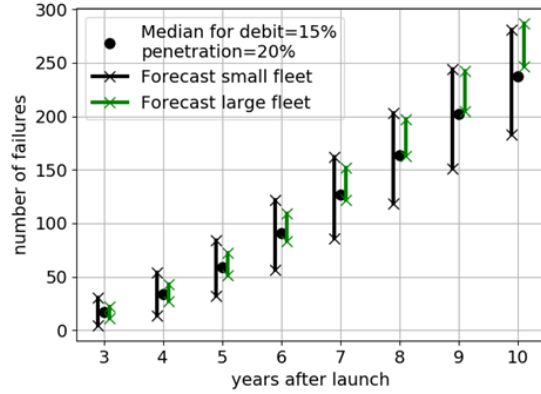
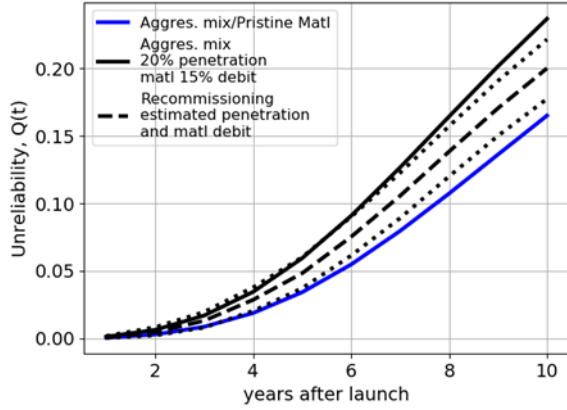
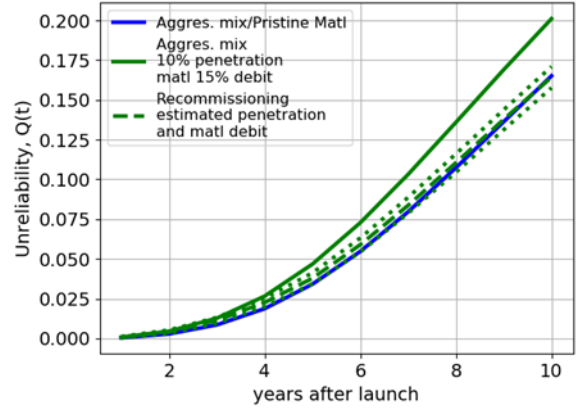


Figure 14. Forecasted number of failures for the small fleet as estimated by both small and large fleet models. Error bars represent the 95% prediction intervals. Expected observations represent the expected number of failures when debit and penetration assume actual (unknown) values of 15% and 20%. At the 3rd year, both models are unbiased. Over time, the small fleet model develop ever increasing uncertainties and the large fleet model becomes biased.

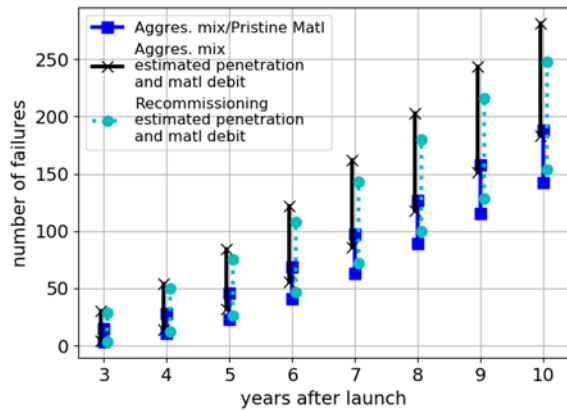
Figure 15 illustrates the estimated/forecasted fleet unreliability and forecasted number of failures over time for the small fleet. Figure 15-(a) shows that after the entire small fleet is recommissioned from the aggressive to the mild mission mix, the estimated fleet unreliability with the updated model falls between the estimates for the entirely pristine fleet and the actual fleet composition both operating at the aggressive mission mix. This means that although there is significant improvement in unreliability, the levels are still above design intent. Interestingly, the distribution in forecasted fleet unreliability might still be useful for estimating number of failures. Figure 15-(b) shows the forecasted number failures coming out of the unreliability estimates of Figure 15-(a). The aggressive mission mix and pristine material represents the design intent. The estimated penetration and material debit represent the forecasts if the fleet keeps operating at the aggressive mission mix. Visual comparison between the two cases makes it clear that the number of failures could be potentially much larger than what was intended. Recommissioning the fleet knocks down the number of failures and make the prediction interval overlap with the one from the design intent.



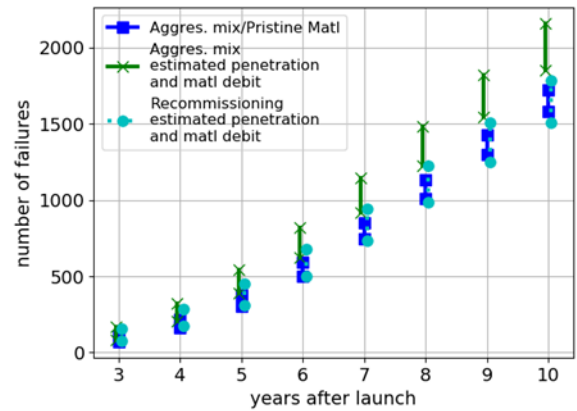
(a) Small fleet unreliability after recommissioning.



(a) Large fleet unreliability after recommissioning.



(b) Small fleet forecasted number of failures.



(b) Large fleet failure observations intervals.

Figure 15. Small fleet recommissioning. Recommissioning curves show the 50th percentile and the 95% prediction interval. Error bars represent the 95% prediction intervals.

Although uncertainty in fleet unreliability is large, recommissioning makes the forecasted number of failures to overlap with design intent (cyan versus blue error bars).

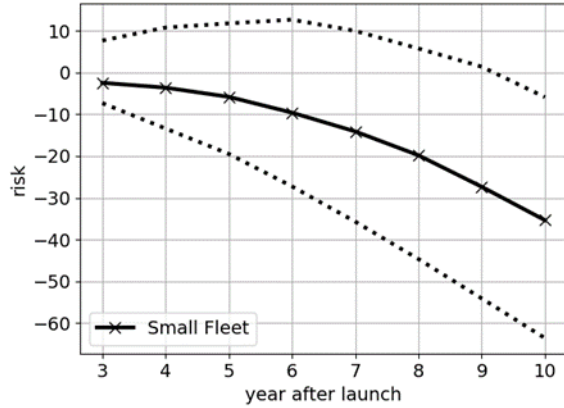
Figure 16. Large fleet recommissioning. Recommissioning curves show the 50th percentile and the 95% prediction interval. Error bars represent the 95% prediction intervals.

Small uncertainty in fleet unreliability implies in great agreement between forecasted number of failures and design intent (cyan versus blue error bars).

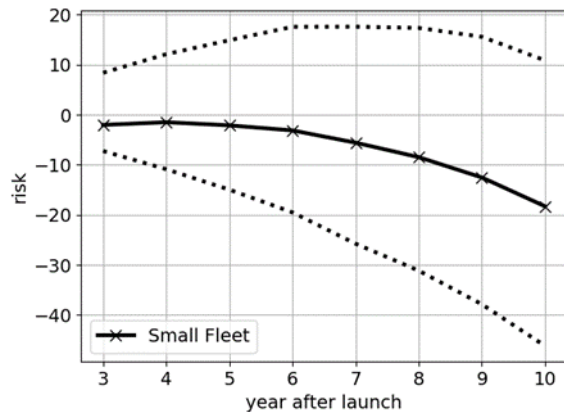
As expected, Figure 16 shows that recommissioning as a risk mitigation measure is much more effective at the large fleet level. Coincidentally, the fleet unreliability after recommissioning converges to design intent. Obviously, it comes at the cost of a mild mission mix (which, again, could imply in reduced performance). Figure 16-(a) shows uncertainty levels in fleet unreliability after recommissioning are much smaller than those shown in Figure 15-(b). This has direct implications in the forecasted number of failures, as illustrated Figure 16-(b), to the point that there is good overlap between estimated and intended error bars.

Besides recommissioning, the small fleet operator can also consider contracting out services and maintenance from a large service provider as a way to reduce financial exposure due to upcoming high number of failures. In real life, it is difficult to forecast the costs associated with such option, as the small operator does not know the outcomes of the large fleet operator model (and model form, assumptions, etc. also tend to be unknown). Nevertheless, we can study that in this synthetic example. Figure 17 shows the forecasted risk for the small fleet, as defined by Eq. (13). The small operator is likely to lose money by self-performing services and maintenance when the risk is negative (since unbiased predictions from large operator tend to be larger than the ones from the small operator). Conversely, the operator saves money by self-performing when risk is positive. Figure 17-

(a) and (b) show the forecasted risk before and after fleet recommissioning, respectively. Although the median risk is relatively small up to the 5th or 8th year, depending on recommissioning, the uncertainty about it tends to be large and continuously increasing.



(a) Before recommissioning.



(b) After recommissioning.

Figure 17. Risk associated with self-performing maintenance (as opposed to buying a contract from large fleet operator) for small fleet. Continuous and dotted lines represent the median and 95% prediction intervals, respectively. Risk is defined by Eq. (13). When risk is positive, the small operator is likely to save money by self-performing maintenance. When risk is negative, operator is likely to lose money by self-performing maintenance.

Another way of looking at the risk associated with self-performing maintenance is through the probability of reward and loss. With risk defined by Eq.(13), there are three things to keep in mind: (1) when risk is positive, the small operator saves money by self-performing services and maintenance, (2) conversely, when risk is negative loss of money is more likely; and finally, (3) unbiased number of failure estimates imply that the expected value of risk is zero.

With that in mind, Figure 18-(a) shows that self-performing is reasonable in the short term. For how long it is a reasonable option really depends on the operator attitude towards risk. If a threshold of $Prob[RISK \geq 0] \geq 0.4$ is imposed, then the small operator could sustain the aggressive mission mix until almost the end of the 4th year (without having to buy a services and maintenance contract). If the operator decides to recommission the fleet at the third year; then, with the $Prob[RISK \geq 0] \geq 0.4$ threshold, self-performing is reasonable until the 7th year.

Now, let us assume that the small operator is willing to accept the risk of under estimating failure of 10 units. Then, Figure 18-(b) shows the probability that the operator will have to pay for extra 10 units (unplanned failures). If a threshold of $Prob[RISK \leq -10] \leq 0.2$ is imposed, then the small operator could sustain the aggressive mission mix until the middle of the 4th year (without having to buy a services and maintenance contract). Switching to a mild mission mix early on, extends that window to the middle of the 5th year.

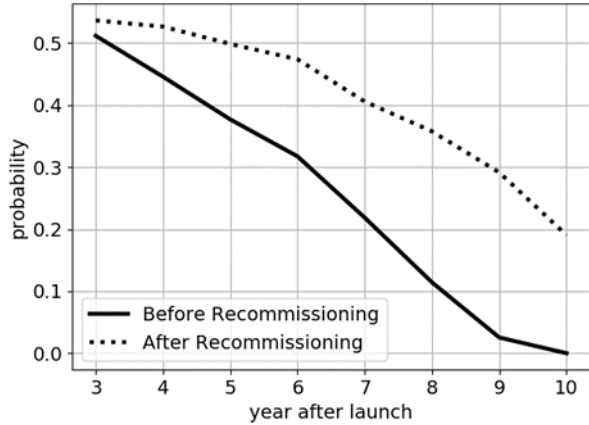
4. CONCLUSIONS AND FUTURE WORK

In this work, we studied early life failures as applied to fleet management. Depending on the scale of the problem, early failures can have significant impact in safety, availability, and operational profit of industrial equipment. We designed a simple numerical experiment where:

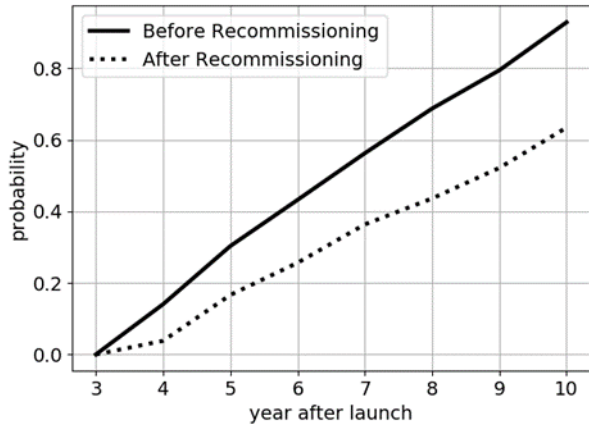
- debit in material capability is used to characterize infant mortality, and
- fleet commissioning is a function of time.

We have studied:

- The effect of debit in material capability: we learned that it can dramatically impact fleet unreliability.
- Fleet commissioning distributed over time: we learned how it can retard the overall increase in fleet unreliability.
- The role of fleet size in the number of observed failures: we verified that observing early life failures in smaller fleets is hard (due to the actual number of potentially affected units). We also found that characterizing the extent of poor material quality is challenging (even in larger fleets).



(a) $Prob[RISK \geq 0]$.



(b) $Prob[RISK \leq -10]$.

Figure 18. Self-performing reward and loss probabilities for small fleet. When risk is negative, the small operator is likely to lose money; otherwise, it saves money by self-performing services and maintenance. With unbiased estimators, there is ~50% chance that risk defined by Eq. (13) will be positive. That value goes down with time. Assuming a tolerance of maximum underestimation of 10 failures, the probability of exceedance starts with 0% and increases over time.

The results obtained so far are promising, and in order to have a better understanding of the impact that early failure in fleets of assets, we want to extend the study and include, among other factors:

- Improved physics of failure models: not only by separating the cycles spent in initiation and propagation, but also improving the stress models to account for geometry and boundary conditions.
- Other sources of uncertainty, such as

- field inspection (damage detection and characterization) and
- service level (repair versus replace failed units).
- Strategies for services: repair and replacement as seen by operator and service provider.

ACKNOWLEDGEMENT

This work was supported by the University of Central Florida (UCF). Nevertheless, any view, opinion, findings and conclusions or recommendations expressed in this material are those of the authors alone. Therefore, UCF does not accept any liability in regard thereto.

REFERENCES

Al-Dahidi, S., Di Maio, F., Baraldi, P. and Zio, E., (2016). Remaining useful life estimation in heterogeneous fleets working under variable operating conditions,” *Reliability Engineering & System Safety*, Vol. 156, pp.109-124. doi:10.1016/j.res.2016.07.019.

Asher, I.,Khan, K.M.K.G, Wang, L., Ling, Y., and Viana, F.A.C., (2017). Calibrating an equivalent initial flaw size distribution from da/dN and S-N data, *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Grapevine, USA, January 9-13, AIAA-2017-1299.

Baraldi, P., Cadini, F., Mangili, F., and Zio, E., (2013). Model-based and data-driven prognostics under different available information, *Probabilistic Engineering Mechanics*, Vol. 32, pp. 66-79.

Bogdanoff, J. L. and Kozin, F., (1985). *Probabilistic models of cumulative damage*, Wiley-Interscience, New York.

Coppe, A., Haftka, R.T., Kim, N.H., (2011). Uncertainty identification of damage growth parameters using nonlinear regression, *AIAA Journal*, Vol. 49, No. 12, pp. 2818–2821.

Dawn A., Kim, N.H, and Choi, J-H, (2015). Practical options for selecting data-driven or physics-based prognostics algorithms with reviews, *Reliability Engineering & System Safety*, Vol. 133, pp. 223-236. doi: 10.1016/j.res.2014.09.014.

GE Aviation collaborators, (2017). TrueChoice Commercial Services, <https://www.geaviation.com/commercial/truechoice-commercial-services> [retrieved 22 October 2017].

Gemini Energy Services collaborators, (2017). Gemini Energy Services: Wind Turbine Services, <http://www.geminienergyservices.com> [retrieved 22 October 2017].

Haddad, G., Sandborn, P., and Pecht, M., (2011). Using real options to manage condition-based maintenance enabled by PHM, *2011 IEEE Conference on Prognostics and Health Management*, Montreal, QC, Canada, doi: 10.1109/ICPHM.2011.6024318.

- Haddad, G., Sandborn, P. A., and Pecht, M. G., (2012). An options approach for decision support of systems with prognostic capabilities, *IEEE Transactions on Reliability*, Vol. 61, No. 4, pp. 872-883. doi: 10.1109/TR.2012.2220699.
- Jiang, Z., Chen, W., Fu, Y., and Yang, R., (2013). Reliability-based design optimization with model bias and data uncertainty, *SAE International Journal of Materials and Manufacturing*, Vol. 6, No. 3, doi:10.4271/2013-01-1384.
- Johnson, W. S., and Hillberry, B. M., (2004). *Probabilistic aspects of life prediction*, ASTM International.
- Kapur, K.C. and Pecht, M., (2014). *Reliability engineering*, John Wiley & Sons.
- Li, C., Mahadevan, S., Ling, Y., Choze, S., and Wang, L., (2016). Dynamic Bayesian network for aircraft wing health monitoring digital twin, *AIAA Journal*, Vol. 55, No. 3, pp 930-941.
- Ling, Y., Asher, I., Wang, L., Viana, F.A.C., and Khan, K.M.K.G, (2017). Information gain-based inspection scheduling for fatigued aircraft components, *19th AIAA Non-Deterministic Approaches Conference*, Grapevine, USA, January 9-13, AIAA-2017-1565.
- Lufthansa Technik AG collaborators, (2017). Engine Services - Lufthansa Technik AG, <https://www.lufthansa-technik.com/engine> [retrieved 22 October 2017].
- MMPDS collaborators, (2017). MMPDS-12: Metallic Materials Properties Development and Standardization, Federal Aviation Administration; Battelle Memorial Institute.
- Pattabhiraman, S., Gogu, C., Kim, N.H., Haftka, R.T., and Bes, C., (2012). Skipping unnecessary structural airframe maintenance using on-board structural health monitoring system, *Journal of Risk and Reliability*, Vol. 226, No. 5, pp. 549-560.
- Peherstorfer, B., Willcox, K. and Gunzburger, M., (2017). Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *SIAM Review*, to appear.
- Rao, S.S., (1992). *Reliability-based design*, McGraw-Hill.
- Rausand, M. and Hoyland, Å., (2004). *System reliability theory: models, statistical methods, and applications*, John Wiley & Sons.
- Siemens collaborators, (2017). Siemens Service Programs and Agreements, <http://www.industry.usa.siemens.com/services/us/en/industry-services/services-glance/service-programs-agreements/pages/service-programs-agreements.aspx> [retrieved 22 October 2017].
- Stephens, R. I., Fatemi, A., Stephens, R.R., and Fuchs, H. O., (2000). *Metal fatigue in engineering*, John Wiley & Sons.
- Volponi, A. J., (2014). Gas turbine engine health management: past, present, and future trends, *Journal of*

Engineering for Gas Turbines and Power, Vol. 136, no. 5. doi: 10.1115/1.4026126.

BIOGRAPHIES

Arinan Dourado is a graduate research assistant at UCF with Bachelor and Master's degrees in chemical engineering from the Federal University of Uberlandia. He also has a PhD degree in mechanical engineering from the Federal University of Uberlandia in the field of solids mechanics and vibration. He is currently working towards a second PhD at UCF in the field of probabilistic mechanics. His research topics involves design and optimization under uncertainty, probabilistic analysis of engineering systems and systems robustness and reliability analyses.

Felipe A. C. Viana is an assistant professor at the University of Central Florida. The vast majority of Dr. Viana's work has been applied to new designs and improvement of fielded products with focus on aircraft propulsion, power generation, and oil and gas systems. Before joining UCF, Dr. Viana was a Sr. Scientist at GE Renewable Energy, where he led the development of state-of-the art computational methods for improving wind energy asset performance and reliability. Prior to moving to that role at GE, he spent five years at GE Global Research, where he lead and conducted research on design and optimization under uncertainty, probabilistic analysis of engineering systems, and services engineering.

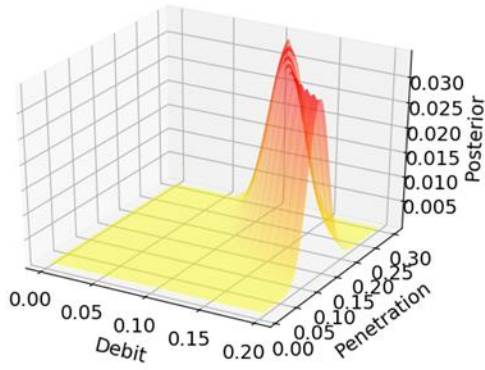
APPENDIX: SIMULTANEOUS COMMISSIONING

As discussed in section 3, most industrial engineering assets (airplanes, jet engines, gas turbines, etc.) are commissioned over a period of time. The results presented in the main body of the manuscript focused on that type of fleet. Nevertheless, infant mortality is also a problem for equipment and consumer goods that experiment virtually synchronous commissioning (when time to deployment of entire fleet is much smaller than asset lives). A potential example comes from the automotive industry. A model of a car for a given year is mostly sold within 3 to 9 months after the product is launched. Cars sold within that year can run for 20 years or more. To simplify the study, we present the simultaneous commissioning of large fleets, as a way to approximate what happens in such cases.

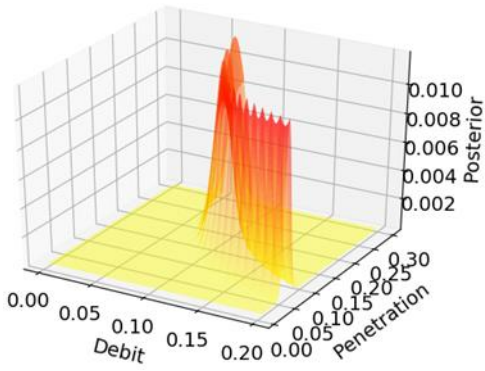
At the 3rd year after product launch, we assume the following number of failure observations:

- Small fleet: 160 failures.
- Large fleet: 1280 failures (again, failures from small fleet are contained here).

Figure 19 details the calibration results with regards to the calibration parameters for the small and large fleets. The uncertainties are significantly smaller when compared to Figure 11-(b) and Figure 12-(b). This is a direct result of the relatively higher number of failure observations.



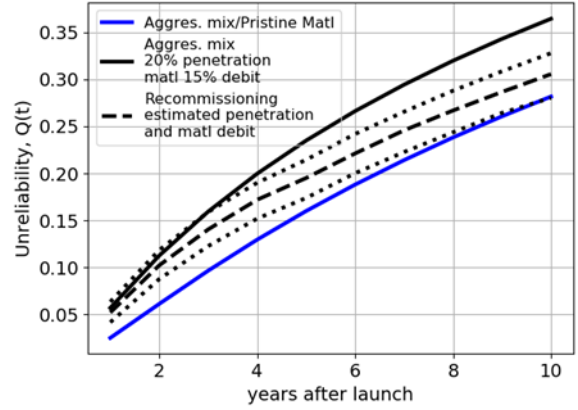
(a) Small fleet model.



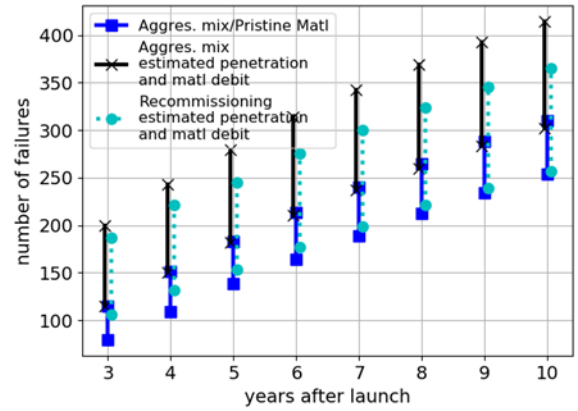
(b) Large fleet model.

Figure 19. Posterior distribution of calibration parameters when fleet initially deployed simultaneously.

The reduced uncertainty in calibration parameters is translated into fleet unreliability forecast. Figure 20 illustrates the estimated/forecasted fleet unreliability and forecasted number of failures over time for the small fleet. Here the comparison with Figure 15 is inconclusive. The fleet unreliability curves have different shapes overall (although recommissioned curves are still within the curves obtained from running the fleet at the aggressive mission mix).



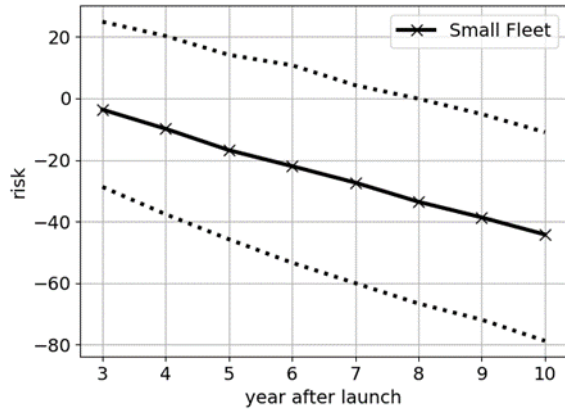
(a) Small fleet unreliability after recommissioning.



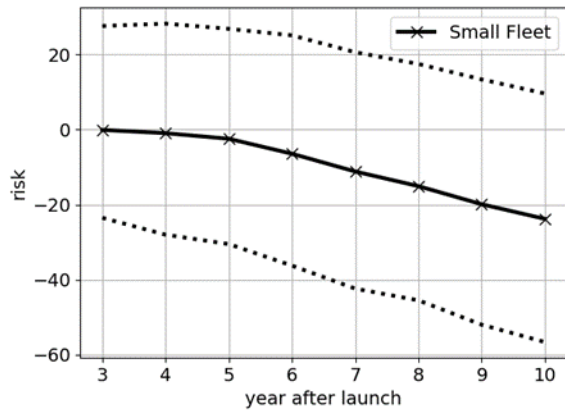
(b) Small fleet forecasted number of failures.

Figure 20. Small fleet recommissioning when fleet initially deployed simultaneously. Recommissioning curves show the 50th percentile and the 95% prediction interval. Error bars represent the 95% prediction intervals. Although uncertainty in fleet unreliability is large, recommissioning makes the forecasted number of failures to overlap with design intent (cyan versus blue error bars).

Figure 21 shows the forecasted risk for the small fleet, as defined by Eq. (13). Once again, the comparison with Figure 17 makes it clear the effect of higher unreliability levels in the number of failure observations. The analysis of self-performing reward and loss probabilities for small fleet was also performed. We reached similar conclusions as previously discussed in section 3.



(a) Before recommissioning.



(b) After recommissioning.

Figure 21. Risk associated with self-performing maintenance (as opposed to buying a contract from large fleet operator) for small fleet when fleet initially deployed simultaneously. Continuous and dotted lines represent the median and 95% prediction intervals, respectively. Risk is defined by Eq. (13). When risk is negative, the small operator is likely to lose money; otherwise, it saves money by self-performing services and maintenance.