DOT/FAA/TC-17/66

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

Analysis of Three Approaches for Estimation of Uncorrected Fleet Risk Due to Wear-out Failures

February 2018

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-17/66		

4. Title and Subtitle	5. Report Date
ANALYSIS OF THREE APPROACHES FOR ESTIMATION OF UNCORRECTED FLEET RISK DUE TO WEAR-OUT FAILURES	February 2018
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
M.H.C. Everdij, J.W. Smeltink, A.L.C. Roelen	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
NLR - Netherlands Aerospace Centre P.O. Box 90502 1006 BM Amsterdam The Netherlands	
	11. Contract or Grant No. 31090611
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
FAA Northwest Mountain Regional Office 1601 Lind Ave SW Renton, WA 98057	Final Report
	14. Sponsoring Agency Code AIR-677

15. Supplementary Notes

The FAA William J. Hughes Technical Center Aviation Research Division COR was Cristina Tan.

16. Abstract

The Transport Airplane Directorate of the FAA pursues Continued Operational Safety (COS) for transport category aircraft. Structured risk assessment methods are used to support consistent, objective decisions regarding concerns on continued operational safety of transport category aircraft. The Transport Airplane Risk Assessment Methodology (TARAM) provides guidance on the risk analysis methodology and acceptable risk guidelines to be used for Transport Category Airplanes. If the uncorrected risk exceeds the acceptable risk guidelines, mandatory corrective action (Airworthiness Directive) is the typical response.
In recent years a number of alternative methods have been developed in support of the TARAM approach. This report investigates several such methods and provides a comparative analysis.

17. Key Words	18. Distribution Statement
Fleet risk; Wear-out failures; Risk assessment; Analysis; Transport airplane risk assessment method; Weibull distribution	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	75	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACO	Aircraft Certification Offices (FAA)
cdf	Cumulative distribution function
CP	Conditional probability
DA	Defect airplanes
IR	Injury ratio
MED	Multiple element damage
MSAD	Monitor Safety/Analyze Data
MSD	Multiple site damage
ND	Nondetection
pdf	Probability density function
TARA	Transport airplane risk analysis
TARAM	Transport Airplane Risk Assessment Methodology
WFD	Widespread fatigue damage

EXECUTIVE SUMMARY

The Transport Airplane Directorate of the FAA pursues Continued Operational Safety (COS) for transport category aircraft. Structured risk assessment methods are used to support consistent, objective decisions regarding concerns on continued operational safety of transport category aircraft. The Transport Airplane Risk Assessment Methodology (TARAM) provides guidance on the risk analysis methodology and acceptable risk guidelines to be used for Transport Category Airplanes. If the uncorrected risk exceeds the acceptable risk guidelines, mandatory corrective action (Airworthiness Directive) is the typical response.

This study concerns uncorrected risk due to wear-out failures in which parts are more likely to fail at higher age.

In recent years, a number of alternative methods have been developed in support of the TARAM approach. Two specific alternative approaches are TARA (Transport Airplane Risk Analysis) and a modification of the TARAM approach in which an additional 'location parameter' is used in the estimation of the number of affected airplanes. The considered methods each write the uncorrected fleet risk as a product of four parameters:

$R_T = DA \times ND \times CP \times IR$ Where:

- DA is the predicted number of airplanes with a wear-out defect during the life of the affected fleet
- ND is the average probability that an occurrence of the defect is not detected before resulting in an unsafe outcome or condition throughout the life of the affected fleet
- CP is the conditional probability that the defect will lead to an unsafe outcome or condition
- IR is the average rate of fatality per person exposed to a specific airplane outcome or condition.

From this point onward, the methods deviate. This report describes how each method addresses the four parameters DA, ND, CP, and IR, and provides a detailed analysis of the differences. The two main differences are 1) TARAM determines DA by anchoring wear-out cracks to a particular size, whereas TARA does not; and 2) TARA provides easy-to-use flowcharts and spreadsheets to assist in the calculations of ND, CP, and IR, whereas TARAM does not. These differences have consequences in terms of uncertainties in the results.

The main conclusions are:

- Both TARAM and TARA are approaches to determine uncorrected fleet risk due to wear-out failures.
- Using the additional 'location parameter' (3-parameter Weibull) improves accuracy, although with slightly conservative results if there are no undiscovered cracks.
- The TARAM handbook contains a lot of information, though the guidance is not always very specific; an important part of the guidance is hidden in examples in the appendices.

TARA helps the user by providing easy-to-use flow charts and spreadsheets, aiming at improved consistency and repeatability.

- The outcome of the TARAM approach will include a certain level of uncertainty, but the uncertainty introduced in the TARA approach is estimated to be significantly larger. The main reason is that TARA adopts several major assumptions and simplifications, due to:

    - Accounting for crack growth in conditional probabilities (ND, CP) rather than in the probability distribution part of the analysis (DA).
    - Using for these conditional probabilities a set of flow charts and spreadsheets that are deterministic and that cover a limited and incomplete set of event sequences.
    - Making several mathematical errors.

The main recommendations are:

- For TARAM: The TARAM handbook could be improved. Important guidance currently hidden in examples and appendices needs to be moved to the main part of the document.
- For TARA: The approach needs to be updated to repair the mathematical errors. It is recommended that the approach anchors cracks to dangerous size. All assumptions adopted need to be made explicit to allow the user to assess the level of bias and uncertainty in the result.
- In the meantime, use TARAM with additional ‘location parameter’ (3-parameter Weibull).

1. INTRODUCTION

1.1 BACKGROUND

FAA Order 8110.107A [1] requires that potential unsafe conditions be assessed using an approved risk-analysis method. The Transport Airplane Risk Assessment Methodology (TARAM) provides guidance on the risk analysis methodology and acceptable risk guidelines to be used for Transport Category Airplanes that are administered by the Transport Airplane Directorate[1]. If the uncorrected risk exceeds the acceptable risk guidelines, mandatory corrective action (Airworthiness Directive) is the typical response.

The TARAM handbook [2] outlines a process for calculating risk associated with continued operational safety issues in the transport airplane fleet. The handbook is intended for use by Aviation Safety Engineers performing and overseeing transport airplane risk analysis as part of the Monitor Safety/Analyze Data (MSAD) process (FAA Order 8110.107A). MSAD is a safety-management process to promote continuing operational safety throughout the lifecycle of aviation products.

The TARAM handbook considers risks due to early failures (in which parts are more likely to fail early in their life, commonly referred to as infant mortality), random failures (in which the parts are equally likely to fail whatever their age, i.e., constant failure rate), and wear-out failures (in which parts are more likely to fail at higher age). For each type, there are measures of fleet risk (i.e., the number of weighted events or fatalities expected in a defined time period if no mandatory action is implemented to correct the identified, potentially unsafe condition) and individual risk (i.e., the probability of individual fatal injury per flight-hour).

Our study focuses on total uncorrected fleet risk due to wear-out failures (i.e., the number of planeloads of fatalities due to wear-out failures statistically expected in the remaining life of the affected fleet if no mandatory corrective action is taken).

1.2 PROBLEM DESCRIPTION

Because a TARAM risk analysis is used to support safety decisions and to determine how fast an issue must be corrected, it is important that the resulting risk value be a realistic and accurate estimate of the actual risk. Conservative estimates are inappropriate (except to quickly show that an issue has very low risk and further analysis is not needed). However, the desire for accurate risk values has to be balanced with the effort needed to obtain them, especially considering the inherent uncertainty in these risk analyses. For example, it would be inappropriate to spend significant time to improve the risk analysis accuracy by a small percentage when the inherent uncertainty of the risk analysis is fairly significant.

TARAM risk analysis must work across all engineering disciplines. Risk values for a flight control system safety issue must be directly comparable to risk values for a structure wear-out failure

[1] Airworthiness Directives administered by the Engine & Propeller Directorate use Advisory Circular AC39-8.

safety issue. The same acceptable risk guidelines must be used, regardless of the engineering discipline involved.

In recent years, a number of approaches have been developed that are alternatives for the TARAM approach. Two specific alternative approaches are TARA [3] and a modification of the TARAM approach in which a 3-parameter Weibull is used in the analysis rather than a 2-parameter Weibull. The FAA wishes to obtain an independent comparison of the original TARAM approach and these two alternative approaches.

1.3 ORGANIZATION OF REPORT

This report is organized as follows:

- Chapter 2 gives a description of the TARAM approach.
- Chapter 3 gives a description of the TARAM approach with 3-parameter Weibull.
- Chapter 4 gives a description of the TARA approach.
- Chapter 5 compares the approaches restricting to the Weibull part of the risk analysis.
- Chapter 6 compares the approaches looking at the remaining part of the risk analysis.
- Chapter 7 gives the conclusions of the study.
- Chapter 8 provides a list of references.
- Appendix A collects information on the Weibull distribution available from the literature.

2. TARAM APPROACH TO FLEET RISK DUE TO WEAR-OUT FAILURES

This chapter outlines the TARAM approach to estimate uncorrected fleet risk due to wear-out failures, as described in the TARAM handbook [2].

Uncorrected fleet risk R_T is defined as the number of planeloads of people who are fatally injured over the life of the airplane fleet, assuming that no mandatory corrective action is taken. So if $R_T = 1$, one accident due to wear-out failures will occur with all onboard fatally injured, or two accidents will occur with 50% of those onboard fatally injured, etc. For wear-out failures, it is a product of 4 parameters:

$$R_T = DA \times ND \times CP \times IR \tag{1}$$

where:

- DA (defect airplanes) is the predicted number of airplanes that would have the wear-out failure, if left undetected, during the remaining life of the fleet.
- ND (nondetection) is the probability that, during future operation and maintenance, a wear-out failure will not be discovered by any means before the cracked element fails.
- CP (conditional probability) is the conditional probability that the wear-out failure will lead to a dangerous event.
- IR (injury ratio) is the injury ratio or the proportion of people fatally injured because of a single dangerous event.

2.1 DETERMINING DA

DA is the predicted number of airplanes that would have a wear-out failure, if left undetected, during the remaining life of the fleet. DA is determined by:

$$DA = \sum_{i=1}^{N_S} \frac{F(t_i^R) - F(t_i^{age})}{1 - F(t_i^{age})} \tag{2}$$

where:

- F is the cumulative distribution function (cdf) of the 2-parameter Weibull distribution, i.e.,

$$F(t;\eta,\beta) = \begin{cases} 1-\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{3}$$

- N_S is the number of suspensions (i.e., the number of aircraft that do not have a wear-out failure). These are aircraft that are still subject to failure but that have survived so far.
- t_i^{age} is the age of aircraft i (e.g., in number of cycles flown).
- t_i^R is the retirement age of aircraft i (usually the same[2] for all i)

Parameters β (the shape parameter) and η (the characteristic life) are determined by a Weibull analysis, typically using fleet data obtained from Aviation Safety Information Analysis and Sharing Program related to a set of aircraft that have been tested positive for wear-out failures. If the number of aircraft with failures is small (i.e., < 20), the Weibayes method is used. In this method, β is assumed known (e.g., $\beta = 4$ for aluminum), and η is calculated analytically:

$$\eta = \left[\frac{1}{r_a}\sum_{i=1}^{N_{FS}} \left(t_i\right)^{\beta}\right]^{1/\beta} \tag{4}$$

where:

- r_a is the number of aircraft with failures.
- N_{FS} is the number of aircraft with failures plus number of suspensions ($N_{FS} = r_a + N_S$).
- t_i is the number of cycles flown by aircraft i at time of failure or suspension.

The difficulty here is that each aircraft may have a different-sized crack, and some aircraft may not have a crack yet. More importantly, typically for only a few aircraft, the cracking status is known: for those in which a crack has been detected (group A), either incidentally or after a follow-on inspection, and for those that have been inspected and have been concluded crack free (group B). Uninspected aircraft (group C) normally cannot be used in the Weibull analysis because it is not known if they are failures or suspensions. This means that the analysis needs to be based on

[2] Although a more realistic test can be used (e.g., 40,000 flights, or 60,000 flight-hours, or 35 years), whichever comes first, then converted to the units of t based on that airplane's utilization.

very few data, resulting in a significant overestimation of the actual risk, and failing to meet the objective of accurately estimating the actual risk.

To solve this, TARAM anchors the age of each aircraft to the age when the aircraft has a crack of dangerous size (i.e., an accident or incident size or obvious major malfunction). The cracks are normalized to accident/incident or obvious major malfunction size. As a result:

- t_i is the number of cycles of aircraft i at time of dangerous event (or alternatively, the number of cycles until which aircraft i is dangerous-event-free).

This age is taken equal to the current age of each aircraft, modified to compensate for the time to grow a current crack to a dangerous-sized crack.

A very important side effect of anchoring the calculations to dangerous-sized cracks is that the set of suspensions (i.e., aircraft without failures) now does not only include the aircraft in group B, but also the aircraft in group C: It is not known if the aircraft in group C have a crack, because they have not been inspected, but it is known that they do not have a dangerous-sized crack, because that would have been obvious. Therefore, N_S is the number of aircraft in groups B + C, and N_{FS} is the number of aircraft in groups A + B + C, which is the entire fleet under consideration.

The groups can be formulated more precisely as follows:

A. Aircraft for which a crack has been discovered, either initially (group A1 for less than dangerous-sized cracks; group A2 for dangerous-sized cracks); or during follow-on inspection (group A3, which includes cracks of any size).
B. Aircraft that have been inspected for cracks and that have been concluded crack free.
C. Aircraft that have not been inspected for cracks but are known to not have been in a dangerous event.

For group A1+A3, the aircraft has a crack, but it will take some time $t_{i,CRgrow}$ (measured in cycles) for the crack to grow to a dangerous-sized crack. Let t_i^{age} denote the number of cycles flown by aircraft i at the current moment. Then, for group A1+A3, take
$t_i = t_i^{age} + t_{i,CRgrow}$. Here, $t_{i,CRgrow}$ may be different for each aircraft i in the group because the detected cracks may have various sizes; it is determined by using extrapolation of crack growth curves, or other engineering estimates.

For group A2, t_i is the number of cycles the aircraft has flown at the time when the dangerous event occurred. The crack already reached dangerous size, and no additional time is needed to grow to it.

For group B, the aircraft has no cracks, but it could develop a detectable crack tomorrow, which will take some time $t_{i,grow}$ to grow into a dangerous-sized crack. This group will at least survive $t_{i,grow}$ cycles from now on. Therefore, for group B, take
$t_i = t_i^{age} + t_{i,grow}$.

For group C, it is not known if the aircraft already developed a crack, but it has no dangerous-sized crack, so it is known that this aircraft survived until now. Take $t_i = t_i^{age}$ (i.e., conservatively assume it is about to fail).

r_a is the number of aircraft in groups A1+A2+A3.

It is interesting to note that the estimate for η is not very sensitive to errors in the estimate of $t_{i,CRgrow}$.

2.2 DETERMINING ND

With the approach discussed in section 2.1, DA is the predicted number of airplanes that would have a dangerous-sized crack during the remaining life of the fleet if left undetected. However, through normal maintenance procedures, inspections, or pre-flight checks, a number of those DA cracks will be detected before they can actually lead to an accident, and the corresponding aircraft are no longer at risk. Only the aircraft with undetected cracks remain at risk. This is modeled by parameter ND.

Parameter ND is the probability that a crack is not detected before it grows to a dangerous size. ND is determined by consideration of factors such as the following [2]:

- How many cases of crack findings are there?
- How many crack lengths are found?
- What is the estimate of the dangerous event crack size?
- What is the estimate of time to grow from discovered crack size to dangerous-event crack size (review of crack growth curves if they are available; extrapolating a little bit past the critical crack length if the curve stops there)?
- How often is the area visible?
- How was the damage found?
- Are there other ways the damage may be found?

Historically, almost all wear-out failures are discovered before they lead to a dangerous event, which indicates that ND is typically much less than 1. ND would be close to 1 if the crack-growth time from a detectable crack to critical crack length is short, and a directed nondestructive inspection is needed to find the crack, but such inspection is not currently being performed. However, in normal situations, ND is a very small number.

2.3 DETERMINING CP

With the above, DA × ND is the number of aircraft with an undetected dangerous-sized crack, occurring during the lifetime of the fleet. However, depending on the type and location of the dangerous cracks, there may be a chance that some of those aircraft could still safely land and end up without major injuries. Parameter CP models the probability that the undetected dangerous-sized crack leads to an accident with injuries.

For conditions under study that lead directly to dangerous events (e.g., a wing crack growing to a size that results in wing failure), $CP = 1$. This is often the case for wear-out failures in principal structural elements. In other cases, a wear-out failure does not completely preclude the possibility of a safe landing, and a CP value of less than 1 would be appropriate (e.g., a crack in a flap fitting that grows to failure of the fitting, but the crew can still land the airplane an estimated 70% of the time ($CP = 0.30$)).

2.4 DETERMINING IR

By combining the results of sections 2.1, 2.2, and 2.3, it is found that DA × ND × CP is the number of aircraft that are in an accident due to an undetected dangerous-sized crack during the lifetime of the fleet. This number is multiplied by the IR to account for the chance that some people on board may survive. For example,
$IR = 1$ if the accident is fatal to all people on board; $IR = 0.5$ if the accident is fatal to 50% of those on board. This is determined from statistics on dangerous events.

3. TARAM APPROACH WITH 3-PARAMETER WEIBULL

The TARAM model described in Chapter 2 used a 2-parameter Weibull distribution to predict the number of airplanes that would have a failure due to wear-out, if left undetected, during the time period being analyzed. This chapter describes an approach in which a 3-parameter Weibull is used instead.

The difference between a 2-parameter Weibull and a 3-parameter Weibull is that in the latter, the cdf is shifted to the right or to the left along the time axis. More precisely: the 3-parameter Weibull distribution has the following cdf:

$$F(t;\eta,\beta,\gamma) = \begin{cases} 1-\exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] & t \geq \gamma \\ 0 & t < \gamma \end{cases} \tag{5}$$

γ is the location parameter. It has the effect of sliding the cdf to the right ($\gamma > 0$) or to the left ($\gamma < 0$) along the time axis. For $t \in [0,\gamma]$, there are no failures, which is why γ is also referred to as failure-free life. A negative γ may indicate that failures have occurred (e.g., prior to actual use). For $\gamma = 0$, the 3-parameter Weibull reduces to a 2-parameter Weibull.

As in the 2-parameter Weibull, the TARAM approach with 3-parameter Weibull anchors failures to dangerous-sized cracks. The location parameter accounts for the time to grow from the initial crack to dangerous-event-size damage. The advantage is improved accuracy. This is because typically, the shape parameter β for crack initiation and the shape parameter for crack initiation plus extensive growth are different. The shape parameter is currently based on a moderate-or detectable-size crack. The location parameter then accounts for the extensive growth part (i.e., $t_{i,grow}$).

The approach used to determine values for the parameters is generally the same as for the 2-parameter version, except that an estimate for the location parameter is now also required.

DA is determined by:

$$DA = \sum_{i=1}^{N_S} \frac{F(t_i^R) - F(t_i^{age})}{1 - F(t_i^{age})} \quad (6)$$

where:

- F is the cdf of the 3-parameter Weibull distribution, i.e.:

$$F(t;\eta,\beta,\gamma) = \begin{cases} 1-\exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] & t \geq \gamma \\ 0 & t < \gamma \end{cases} \quad (7)$$

- N_S is the number of suspensions, (i.e., the number of aircraft that are still subject to failure but that have survived so far).
- t_i^{age} is the age of aircraft i (e.g., in number of cycles flown).
- t_i^R is the retirement age of aircraft i (usually the same[3] for all i).

Parameters β and η are determined by following the procedure in Chapter 2, but with a changed formula for η. In this method, β is assumed known (e.g. $\beta = 4$ for aluminum), and η is calculated analytically:

$$\eta = \left[\frac{1}{r_a}\sum_{i=1}^{N_{FS}}\left(t_i - \gamma\right)^{\beta}\right]^{1/\beta} \quad (8)$$

However, in this formula, use $t_i - \gamma =: 0$ for those i for which $t_i - \gamma \leq 0$. Estimates for γ can come from deterministic-damage-tolerance analysis or from expert opinion estimates. The remainder of the approach is similar to the one for 2-parameter Weibull:

- N_{FS} is the number of aircraft with failures plus number of suspensions.
- r_a is the number of aircraft with failure.
- t_i is the age of aircraft i at time of dangerous event (or, alternatively, the age until which aircraft i is dangerous-event-free), in number of cycles. This age is equal to the current age of each aircraft, modified to compensate for the time to grow the current crack to a dangerous-size crack.

It is interesting to note that the estimate for η is not very sensitive to errors in the estimate of γ.

4. TARA APPROACH TO FLEET RISK DUE TO WEAR-OUT FAILURES

An alternative approach is used in some FAA offices. In this alternative approach, no attempt is made to adjust the crack findings to a common size. The detected cracks are the initial damage

[3] However, in actual practice, a more-realistic test can be used (e.g., 40,000 flights or 60,000 flight-hours or 35 years, whichever comes first, then converted to the units of *t* based on that airplane's utilization).

condition that is considered. Additionally, there are flow charts and spreadsheets to support the calculation of ND, CP, and IR. This approach is referred to as TARA and is described in [3].

In the TARA approach, each risk analysis is performed for a single type of damage (e.g., cracking detected in a specific structural component in a particular model of airplane). For a given defect or damage state, the uncorrected fleet risk (weighted events, in planeloads of fatalities) is calculated as the product:

$$R_T = DA \times ND \times CP \times IR \tag{8}$$

Where:

- DA is the predicted number of airplanes with the defect (damage condition) during the life of the affected fleet.
- ND is the average probability that an occurrence of the defect is not detected before resulting in an unsafe outcome or condition throughout the life of the affected fleet.
- CP is the conditional probability that the defect will lead to an unsafe outcome or condition.
- IR is the average rate of fatality per person exposed to a specific airplane outcome or condition.

4.1 DETERMINING DA

In [3], DA, the expected number of defect airplanes, is determined by:

$$DA = \sum_{i=1}^{N} \frac{F(t_i^R) - F(t_i)}{1 - F(t_i)} \tag{9}$$

where:

- F is the cdf of the 2-parameter Weibull distribution, i.e.

$$F(t;\eta,\beta) = \begin{cases} 1-\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{10}$$

- N is the active fleet of airplanes.
- t_i is the total accumulated number of flight cycles for aircraft i, at the time at which the damage was detected, or at the current time for airplanes in which damage has not been detected.
- t_i^R is the retirement age of aircraft i (usually the same for all i).

If the number of damaged aircraft is small (i.e., < 20), the Weibayes method is used to determine parameters β and η. In this method, β is assumed known, and η is calculated analytically:

$$\eta = \left[\frac{1}{r_a}\sum_{i=1}^{N}(t_i)^{\beta}\right]^{1/\beta} \tag{11}$$

Where r_a is the number of airplanes for which the damage condition has been detected, and N and t_i are as above. Reference [3] provides the following values for β presented in table 1:

Table 1. Shape parameters in Weibayes method, table 1 in [3], which refers to [4] and [5] for the first four values and to experience for the last value

Material	Shape parameter β
Aluminum	4
Titanium	3
Low-strength steel ($F_{tu} \leq 240$ ksi)	3
High-strength steel ($F_{tu} > 240$ ksi)	2.2
Stress corrosion cracking	2

It is noted that TARA does not anchor damage to dangerous-sized cracks, and uses the current age or the age of each aircraft at the time of the crack instead. Another main difference is that the sum for DA includes the entire active fleet (including the failures and the uninspected aircraft), rather than only the aircraft without damage. These differences are discussed in Chapters 5 and 6.

Reference [3] notes that aircraft that have been retired from active service may also be included in the dataset used to fit η. For aircraft that had the damage condition before retirement, take $t_i =$ age at the time the crack was detected, and include the aircraft in the count of r_a. If the retired airplane did not have the damage condition, then t_i = retirement age. However, the retired aircraft are not taken into account in the calculation because they are no longer at risk for unsafe outcomes.

4.2 DETERMINING ND

Reference [3] provides two connected flowcharts to assist in the determination of ND (i.e., the probability that the crack is not detected before resulting in an unsafe outcome). The starting point is one or more aircraft with damage (i.e., a crack of any size). The flowcharts take into account how easy it is to inspect the structure to find the damage, whether the design has redundant load paths, and whether the structure is susceptible to widespread fatigue damage (WFD), see figures 1 and 2.

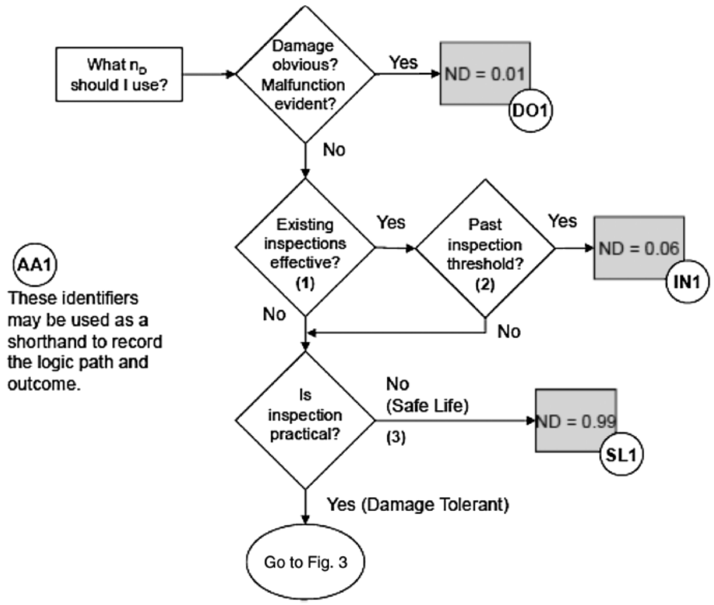


Figure 1. First flowchart used by TARA to determine ND

The inspection threshold refers to the time of first inspection in flights. Inspections are not considered effective if the damage occurs before the inspection threshold. Practicality of inspection is considered in the sense of Title 14 of the Code of Federal Regulations (14 CFR) 25.571. In a safe life policy, the part is removed and replaced at predetermined intervals, rather than when it shows signs of fatigue.

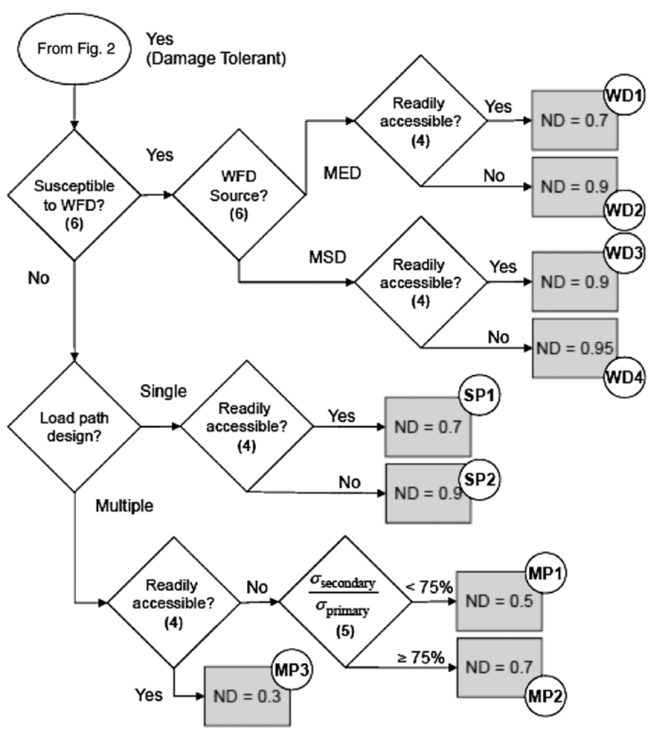


Figure 2. Second flowchart used by TARA to determine ND

WFD is widespread fatigue damage per 14 CFR 25.571. MED is multiple element damage (at least three elements). MSD is multiple site damage. “Readily accessible” means that structural damage may be found by activity not considered effective in item (1) in figure 1 when access is performed (system checks, walk around, etc.). A multiple-load path design is able to redistribute loads after the failure of a component; a single-load path design is not. The ratio at (5) is the operational stress of the secondary load path divided by the operational stress of the primary load path.

4.3 DETERMINING CP AND IR

CP is the overall conditional probability that an airplane with damage will experience an unsafe outcome. Each unsafe outcome has an IR representing the ratio of fatalities to people onboard.

A causal chain is introduced to trace the steps from the initial damage condition (the detected crack) to various unsafe outcomes:

$$CP = PA1 \times PA2 \times PA3 \times PA4 \tag{12}$$

where PAi is the conditional probability from each step in the causal chain (e.g. $PA3$ is the probability that condition $A3$ will occur, given that the airplane has condition $A2$). Condition $A4$ is referred to as unsafe outcome. Four possible unsafe outcomes $A4$ are considered, and their IRs are given (i.e., $IR = 1$ for $A4 =$ in-flight break-up, $IR = 0.98$ for $A4 =$ crash, $IR = 0.03$ for $A4 =$ runway departure and $IR = 0.001$ for $A4 =$ individual fatality). These numbers are based on historical data for transport airplane accidents and were developed in conjunction with the FAA’s Transport Airplane Directorate staff.

For a given initial condition (i.e., the detected crack), the total $CP \times IR$ is then equal to the sum $\sum_{k=1}^{K} CP_k \times IR_k$ over the various unsafe outcomes $k = 1, \ldots, K$ of the initial condition.

Microsoft® Excel® spreadsheets (there are different ones for damage in fuselage, in wing/pylon/empennage, and in landing gear) assist in computing $CP \times IR$. The user selects from pulldown menus conditions $A1$ and $A2$ that best correspond to the initial condition. Next, the user estimates the number of cycles required for the damage to progress from the initial condition to condition $A1$, as a percentage of the retirement age t_R (also in number of cycles). This leads to probability $PA1$ (i.e., $PA1 = 1$ if percentage between 0 and 10%; $PA1 = 0.75$ if percentage between 11 and 30%; $PA1 = 0.5$ if percentage between 31 and 50%; $PA1 = 0.1$ if percentage between 51 and 70%; $PA1 = 0.01$ if percentage between 71 and 90%; $PA1 = 0.005$ if percentage between 91 and 100%). The spreadsheet then automatically populates the possible conditions $A3$, the possible conditions $A4$ (unsafe outcomes), the conditional probabilities $PA2$, $PA3$, and $PA4$, and for each unsafe outcome the injury ratio IR and the product $\sum_{k=1}^{K} CP_k \times IR_k$. All calculations are deterministic.

Figure 3 is an example of the computation of the causal chain in the spreadsheet for “Fuselage.” The user has typed “Cracked stringer near crown” as the initial condition detected and has selected condition $A1$ “Stringer Failure” and condition $A2$ “Other skin failure” from pulldown menus as corresponding best to this initial condition. The user has estimated the time from initial condition to condition $A1$ to be 0 cycles, leading to $PA1 = 1$. The other results then follow automatically. For condition $A3$, there are four possible outcomes: decompression, total loss of control, reduction of control, and failure of emergency equipment, but the last one is not applicable to the selected condition $A2$ (it has been greyed out). In the same way, there are four possible outcomes $A4$ for each condition $A3$, but some have been greyed out. The sheet computes CP and $CP \times IR$ by summing over all applicable unsafe outcomes. The highest risk contributors are highlighted in red.

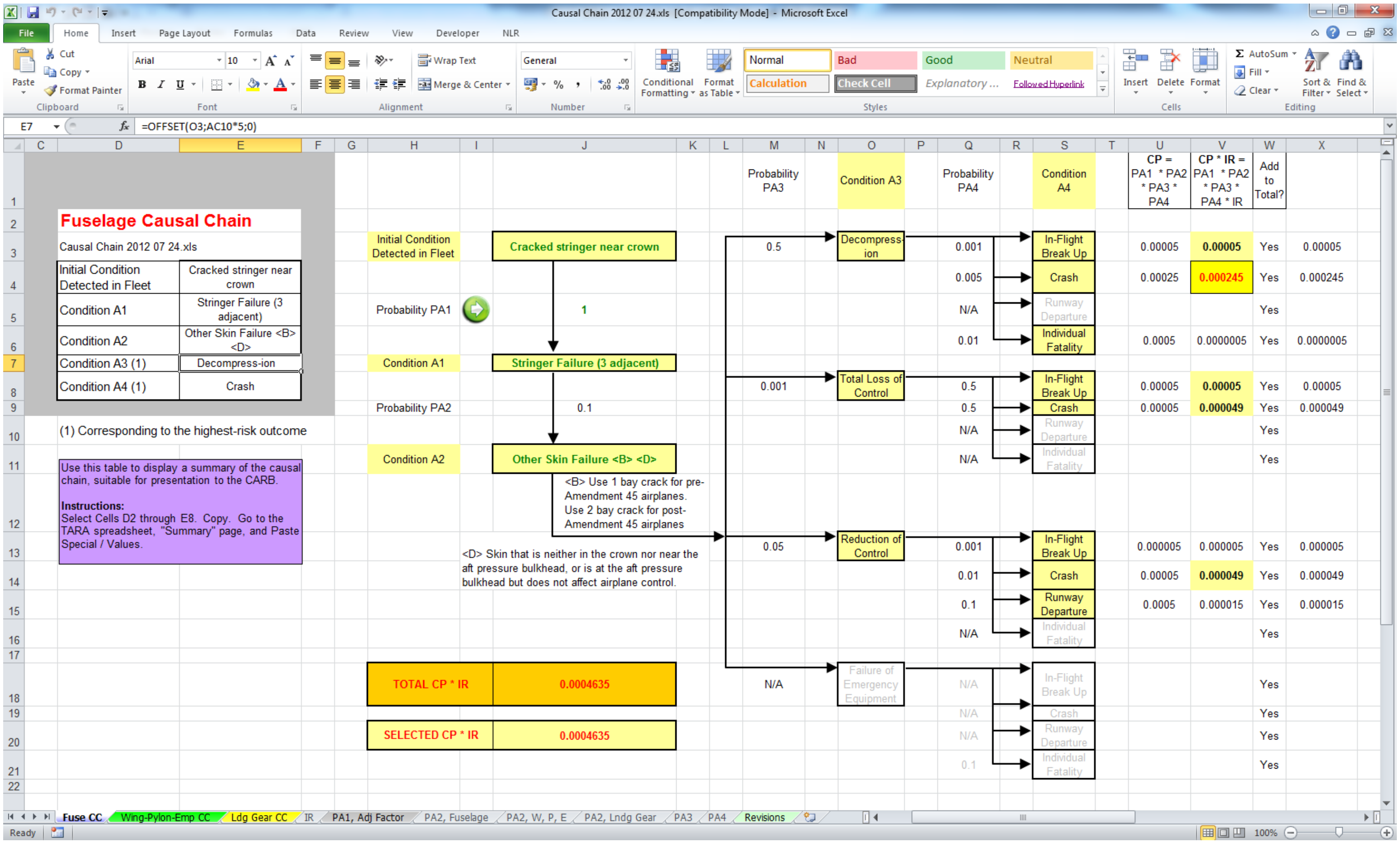


Figure 3. Worksheet for the fuselage causal chain

The values for $PA2$, $PA3$, and $PA4$ are tabulated values obtained from given tables. These tables were developed by the FAA's Seattle Aircraft Certification Offices (ACO) in consultation with the Los Angeles ACO and Boeing Commercial Airplanes based on expert judgment.

5. COMPARISON OF APPROACHES W.R.T. DA

To compare the Weibull part of the three approaches, Monte Carlo simulations have been performed by the FAA [6] to see which approach provides the best predictions of DA (i.e., the number of aircraft with cracking at the fleet retirement age, if no mandatory corrective action was taken). This chapter provides an interpretation and discussion of those results.

5.1 MONTE CARLO SIMULATIONS BY THE FAA

In the Monte Carlo simulations done by the FAA, a set of $N = 1000$ aircraft was considered. Before the actual simulations were started, the situation for the initial time was set up in three stages:

First stage: Each aircraft i was given a current age (counted in number of flights), denoted by t_i^{age}. In one set of runs, these ages t_i^{age} were uniformly distributed from 50 flights to $50{,}000$ flights, and in another they were uniformly distributed from $30{,}000$ to $50{,}000$ flights.

Second stage: This stage required several iterations of steps 1–3 below, with the aim to obtain a situation with a given number of aircraft $n_{moderate}$ having a moderate-sized crack at the initial time, a given number $n_{initial} \leq n_{moderate}$ of which would be initially discovered:

1. For each aircraft i, it was determined at which age t_i^{mod} (counted in number of flights) it will develop a moderate-sized crack. Each such t_i^{mod} was generated to satisfy the Weibull distribution $F(t; \eta_{mod}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{mod}}\right)^{\beta}\right]$. The formula to generate the failure times is $t = \eta_{mod} \cdot (-\ln(1 - \mathrm{Rnd}(\quad)))^{1/\beta}$ with shape parameter fixed at $\beta = 4$ (for aluminum) and using the initial choice (seed) for the input scale parameter η_{mod}; $\mathrm{Rnd}(\quad)$ generates a random number between 0 and 1.The first pass uses the seed η_{mod} and generates the failure times using the random number generator formula; the failure times are saved in an array. Subsequent passes adjust the value of η_{mod} until the desired number of moderate size cracks is produced (the failure times in the array are all factored up or down by the same amount as they are directly proportional to η_{mod}).
2. For each aircraft i, the generated moderate crack-age t_i^{mod} was compared with the current age t_i^{age} to determine whether the moderate crack would occur in the future or had occurred in the past. If the number of cracks of at least moderate size at the initial time was unequal to $n_{moderate}$, a new value for the input scale parameter η_{mod} was calculated using an iterative=solution adjustment scheme, and the procedure was repeated from step 1. The desired number of moderate-size cracks at the initial time can always be obtained given a reasonable seed η_{mod}; to avoid bias, the seed value was chosen so that, on average, no adjustment to η_{mod} was needed.
3. The time for a crack to grow from a detectable to obvious (dangerous) crack size was taken equal to 15 000 flights. Each crack was assumed to grow in a similar way, but no

assumptions were adopted on how the growth took place (e.g., no assumptions on crack growth being linear or otherwise). A moderate-sized crack was defined as the crack size after 7,500 flights. With this, the time for a crack to grow from a detectable to moderate crack size is equal to $t_{DM} = 7{,}500$ flights; the time for a crack to grow from a moderate to an obvious (dangerous) crack size is equal to $t_{MO} = 7{,}500$ flights. Using this, it was determined how many generated cracks had time to grow to obvious size at the initial time; these were automatically labelled "initially discovered." If this number was above $n_{initial}$, the iteration was discarded and labelled "excessive obvious" and the procedure was repeated from step 1 first pass. If the number of obvious size cracks was below $n_{initial}$, some aircraft with cracks of at least moderate size at the initial time were selected randomly, until the number $n_{initial}$ was reached.

The rest of the moderate-sized cracks at the initial time ($n_{undisc} = n_{moderate} - n_{initial}$) were labelled "initially undiscovered."

Third stage: A given fraction (P_{insp}^{vol}) of the fleet (of $N - n_{initial}$ aircraft) is inspected. During this inspection campaign, part of the (n_{undisc}) initially undiscovered moderate cracks may be found. In addition, if an aircraft has a crack of at least detectable but smaller-than-moderate size, then this crack will also be found if the aircraft is inspected. The total number of inspected aircraft, then, is $n_{initial} + P_{insp}^{vol} \cdot (N - n_{initial})$ and the total fraction inspected is $P_{insp} = \left(n_{initial} + P_{insp}^{vol} \cdot (N - n_{initial})\right)/N$. The set of discovered cracks may have various sizes, from detectable to obvious. The collection of parameters used to set up the Monte Carlo simulations is presented in table 2.

Table 2. Collection of parameters used to set up the Monte Carlo simulations

Parameter		Value
N	Number of aircraft in fleet	1,000
t_i^{age}	Age of aircraft i ($i = 1, \dots, N$)	In one set these ages are $0 + 50 \cdot i$ flights (hence ages are $50{,}100, \dots, 50{,}000$ flights). In another set they are $30{,}000 + \frac{20\,000}{N-1} \cdot (i-1)$ flights (hence ages are 30,000, 30,020, …, 50,000 flights)
t_i^R	Retirement age	70,000 flights, for all i
β	Shape parameter	4 (= aluminum)
η_{mod}	Input scale parameter that determines when moderate cracks are generated	This input parameter is adjusted (with trial and error) in such a way that a total of $n_{moderate}$ aircraft in the fleet has a crack of at least moderate size at the initial time
t_{DM}	Time for crack to grow from detectable to moderate size	7,500 flights
t_{MO}	Time for crack to grow from moderate to obvious size	7,500 flights
P_{insp}	Fraction of the fleet that is inspected for cracks (initially discovered plus inspected during inspection campaign)	Depends on scenario considered, from $0.003, 0.05, 0.1, 0.15, 0.2, 0.3, \dots, 0.9, 1.0$
$n_{initial}$	Number of cracks initially discovered; these are of at least moderate size	Depends on scenario considered, either 1 or 3
n_{undisc}	Number of moderate-sized cracks initially existing but undiscovered, except possibly during inspection campaign	Depends on scenario considered, 0, 1, 3, 6, 9, or 12.
$n_{moderate}$	$n_{moderate} = n_{undisc} + n_{initial}$.	

After the three-stage set-up, the Monte Carlo simulations could start. For each scenario (consisting of a combination of $(P_{insp}, n_{initial}, n_{undisc})$ and a distribution of initial ages), a simulation consisted of at least 10,000 trials.

For each scenario, a Weibayes analysis was performed to predict DA (i.e., the number of aircraft with cracking at the fleet retirement age, if no corrective action was taken). This was done by using the data to determine a Weibayes estimate for the output scale parameter, and next using this as input to the formula for DA.

By counting in the data, one can also determine the actual number of aircraft with cracks at retirement, based on which a correction factor was computed (i.e., the actual number of aircraft with cracks divided by the predicted number of aircraft with cracks). In addition, an error ratio can be computed as the predicted number of aircraft with cracks divided by the actual number of aircraft with cracks.

Correction Factor = Actual DA/Predicted DA

Error Ratio = Predicted DA/Actual DA = 1/Correction Factor

The correction factor and error ratio are calculated for each trial and are a measure of how well that trial predicted the outcome for a particular case (cases are described in the remainder of section 5.1). Each trial's correction factors are retained, and the average factor over all the trials (for a given scenario and case) is calculated and used as the statistical measure of how well that case performed for that scenario.

For some scenarios and cases, there may be a trial in which the actual DA is zero. This results in an infinite error ratio for that trial. For case/scenario combinations that had that issue, the error ratio was not calculated. The correction factor never has that issue, so it was used as a proxy to the error ratio to assess the relative performance of the different cases.

For any one trial, the error ratio = 1/correction factor, but this is not true of the average error ratio and average correction factor over a large number of trials (i.e., the average error ratio does not equal 1/average correction factor. Nevertheless, the correction factor is a good proxy for the error ratio.

The Weibayes analysis is done using six different approaches that are tested; these are referred to as cases.

For each case, the fleet of aircraft is split up into three groups A, B, and C, with group A sometimes being split up further into groups A1, A2, and A3:

A. Aircraft for which a crack has been discovered, either initially (group A1 for at least moderate but less than obvious cracks or group A2 for obvious cracks) or during follow-on inspection (group A3, which includes less-than-obvious-sized cracks).
B. Aircraft that have been inspected for cracks and have been deemed crack free.
C. Aircraft that have not been inspected for cracks but that are known to not having been in a dangerous event.

Groups A1 and A2 consist of aircraft with initially discovered cracks: therefore, the number of aircraft in these groups equals $|A1| + |A2| = n_{initial}$ (where '$|X|$' denotes 'number of aircraft in group X').

The numbers of aircraft in groups A3, B, and C depend on the value for P_{insp}, or, to be more specific, P_{insp}^{vol}. Some of the aircraft in group A3 have cracks of at least moderate size; these are the fraction of n_{undisc} that have been inspected. In addition, at the time of inspection, there may be a few, for example n_{small}, aircraft with cracks of detectable but less than moderate size. If these aircraft are inspected, these small cracks will be discovered, and will be included in group A3. Therefore, the average number of aircraft in group A3 is $|A3| = P_{insp}^{vol} \cdot (n_{undisc} + n_{small})$.

Further, on average $|B| = P_{insp} \cdot N - |A1 + A2 + A3| = P_{insp}^{vol} \cdot (N - n_{moderate} - n_{small})$ and $|C| = (1 - P_{insp}) \cdot N$.

This means the number of aircraft in groups A3 and B gets larger with increasing P_{insp}; the number of aircraft in group C gets smaller.

Case 1: The Weibayes analysis is anchored (normalized) to dangerous event/obvious size damage.

This case corresponds to the TARAM approach discussed in Chapter 2.

First, the data are used to estimate the scale parameter $\eta_{case1} = \left[\frac{1}{r_a}\sum_{i=1}^{N}(t_i)^\beta\right]^{1/\beta}$. For the aircraft in group A (crack discovered), the time that would be needed for the crack to grow from its size at time of observation (which can be anything from detectable to obvious) to obvious (damage) size was added to the time at observation and entered as a failure data point t_i. It is assumed that the researcher can accurately predict the time of obvious crack given the observed size of the crack; therefore, in the Monte Carlo simulations, this time is taken equal to $t_i = t_i^{mod} + t_{MO}$, where t_i^{mod} is the age at which the aircraft develops the moderate crack (which was sampled at the beginning of the simulation), and t_{MO} is the time to grow from moderate-sized crack to obvious damage. If an aircraft was inspected and found crack free (aircraft in group B), the time for the crack to grow from detectable to obvious size (i.e., $t_{DM} + t_{MO}$) was added to age at inspection and entered as a suspension data point; therefore, $t_i = t_i^{age} + t_{DM} + t_{MO}$ (it is known that the crack-free or less-than-detectable-size crack would not reach obvious size damage until then). If an airplane was not inspected at all (aircraft in group C), it was entered as a suspension data point at its current age, $t_i = t_i^{age}$ (all that can be said is that it currently does not have obvious size damage). Therefore:

$$\eta_{case1} = \left[\frac{1}{r_A}\sum_{group\ A}\left(t_i^{\text{mod}} + t_{MO}\right)^\beta + \frac{1}{r_A}\sum_{group\ B}\left(t_i^{age} + t_{DM} + t_{MO}\right)^\beta + \frac{1}{r_A}\sum_{group\ \text{C}}\left(t_i^{age}\right)^\beta\right]^{1/\beta} \qquad (13)$$

where r_A is the number of aircraft in group A.

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case1} = \sum_{group\ B+C}\frac{F_1(t_i^R) - F_1(t_i^{age})}{1 - F_1(t_i^{age})} \qquad (14)$$

where $F_1(t) = F(t; \eta_{case1}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case1}}\right)^\beta\right]$.

The Monte Carlo simulations also make a calculation of the predicted number of failures at the current moment. This is taken to be:

$$R_{case1}^{now} = \sum_{group\ A} 2 \cdot F_1\left(t_i^{\text{mod}} + t_{MO}\right) + \sum_{group\ B} F_1\left(t_i^{age} + t_{DM} + t_{MO}\right) + \frac{1}{r_A}\sum_{group\ \text{C}} F_1\left(t_i^{age}\right) \qquad (15)$$

Note the factor 2 included in the terms for group A. This factor is due to "Abernethy's reduced bias adjustment" (see ref [7] and section A.5 of appendix A).

All the data produced during the simulation can also be used to count the actual number of obvious cracks at the current moment, and (limited to aircraft that are suspensions now, i.e., groups B and C) the actual number of obvious cracks at retirement of the fleet. These can be compared with the predicted values above (correction factor = number of actual cracks divided by number of predicted cracks; or error ratio = number of predicted cracks divided by number of actual cracks).

Case 2: Weibayes analysis is not adjusted for crack size.

This case corresponds to the TARA approach outlined in Chapter 4, with one difference: the sum for DA does not include the aircraft in group A (detected failures); it was assumed that this was a mistake in the TARA approach documentation.

If an airplane was known cracked (group A), it was entered as a failure data point at the known crack age. For group A3 (discovery during follow-on inspection), this time is equal to the time of inspection (i.e., the current time, $t_i = t_i^{age}$). For group A1 (discovery initially of a less-than-obvious crack), the Monte Carlo simulations use a function "timeFoundSolver" that calculates a time $t_i^{modfound}$ of discovery of the crack from a particular probability distribution on interval $[t_i^{mod} - t_{DM}, t_i^{mod} + t_{MO}]$, where t_i^{mod} is the age at which the crack is of moderate size[4]. Then for group A1, $t_i = t_i^{modfound}$. For group A2 (discovery initially of an obvious crack), $t_i = t_i^{mod} + t_{MO}$. For groups B and C, take the current age of the aircraft, $t_i = t_i^{age}$. Therefore:

$$\eta_{case2} = \left[\frac{1}{r_A} \sum_{group\ A1} \left(t_i^{\mathrm{mod}\ found} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ A2} \left(t_i^{\mathrm{mod}} + t_{MO} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ A3+B+C} \left(t_i^{age} \right)^{\beta} \right]^{1/\beta} \tag{16}$$

where r_A is the number of aircraft in group A (including A1+A2+A3).

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case2} = \sum_{group\ B+C} \frac{F_2(t_i^R) - F_2(t_i^{age})}{1 - F_2(t_i^{age})} \tag{17}$$

where $F_2(t) = F(t; \eta_{case2}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case2}} \right)^{\beta} \right]$.

[4] Function timeFoundSolver () calculates the time of crack discovery for an initially discovered (pre-campaign) moderate-size crack. The cdf for crack discovery is: cdf $= 0.5 + 0.106103 * (0.25 * \sin(2\theta) + 2 * \sin(\theta) + 1.5 * \theta)$. The domain of θ varies from $-\pi$ to $+\pi$; $-\pi$ corresponds to a detectable size crack, 0 corresponds to a moderate-sized crack, and $+\pi$ corresponds to an obvious-sized crack.

The Monte Carlo simulations also make a calculation of the predicted number of failures at the current moment. This is taken to be:

$$R_{case2}^{now} = \sum_{group\ A1} 2 \cdot F_2\left(t_i^{\mathrm{mod}\ found}\right) + \sum_{group\ A2} 2 \cdot F_2\left(t_i^{\mathrm{mod}} + t_{MO}\right) + \sum_{group\ A3} 2 \cdot F_2\left(t_i^{age}\right) + \sum_{group\ B+C} F_2\left(t_i^{age}\right) \quad (18)$$

All the data produced during the simulation can also be used to count the actual number of cracks at the current moment, and (limited to the aircraft that are suspensions now, i.e., groups B and C) the actual number of moderate or larger cracks at retirement of the fleet. These can be compared with the predicted values above. Note in Case 1, only cracks are counted that had time to develop to obvious size before the retirement age ($t_i^{mod} + t_{MO} \leq t_i^R$). In Case 2, cracks of at least moderate size are counted (i.e., $t_i^{mod} \leq t_i^R$). The reason is that $F_2(t)$ is fitted on a scale parameter η_{case2} that considers any crack, whereas $F_1(t)$ is fitted on a scale parameter η_{case1} that considers obvious-sized cracks. Note that in Case 2, cracks of at least detectable but less-than-moderate size are not counted because the prediction was intended to be for moderate- or larger-sized cracks, though this could have been done by counting those (currently suspended) aircraft for which $t_i^{mod} - t_{DM} \leq t_i^R$.

Case 3: Weibayes analysis is not adjusted for crack size.

This is the same as for Case 2, the difference being that aircraft in group C (uninspected aircraft) are not included in the calculations for the scale parameter (the traditional approach). Therefore:

$$R_{case3} = \left[\frac{1}{r_A} \sum_{group\ A1} \left(t_i^{\mathrm{mod}\ found}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ A2} \left(t_i^{\mathrm{mod}} + t_{MO}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ A3+B} \left(t_i^{age}\right)^{\beta} \right]^{1/\beta} \quad (19)$$

where r_A is the number of aircraft in group A (including A1+A2+A3).

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case3} = \sum_{group\ B+C} \frac{F_3(t_i^R) - F_3(t_i^{age})}{1 - F_3(t_i^{age})} \quad (20)$$

where $F_3(t) = F(t; \eta_{case3}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case3}}\right)^{\beta}\right]$. Note this includes group C.

The predicted number of failures at the current moment is:

$$R_{case3}^{now} = \sum_{group\ A1} 2 \cdot F_3\left(t_i^{\mathrm{mod}\ found}\right) + \sum_{group\ A2} 2 \cdot F_3\left(t_i^{\mathrm{mod}} + t_{MO}\right) + \sum_{group\ A3} 2 \cdot F_3\left(t_i^{age}\right) + \sum_{group\ B+C} F_3\left(t_i^{age}\right) \quad (21)$$

Note that group C is included in the last term. Also, in the count of the number of actual cracks at retirement, Case 2 included only cracks of at least moderate size ($t_i^{mod} \leq t_i^R$), whereas Case 3 includes cracks of any size ($t_i^{mod} - t_{DM} \leq t_i^R$).

Case 4: Similar to Case 3, except the Weibayes analysis is adjusted to detectable crack size.

In this case, the Weibayes analysis is anchored to cracks of detectable size. Aircraft in group C are not used to estimate the scale parameter. For group B, no cracks have been detected, so any size of cracks is smaller than detectable; take $t_i = t_i^{age}$. For the aircraft in group A (including A1, A2, A3), the time that would be needed for the crack to grow from detectable size to its observed size is subtracted from the time at observation and entered as a failure data point. It is assumed that the researcher can accurately trace back the time of detectable crack given the current size of the crack; therefore, in the Monte Carlo simulations, this time is taken equal to $t_i = t_i^{mod} - t_{DM}$ (provided this is greater than zero; otherwise take zero), where t_i^{mod} is the age at which the aircraft develops the moderate crack. Therefore:

$$\eta_{case4} = \left[\frac{1}{r_A} \sum_{group\ A} \left(t_i^{\text{mod}} - t_{DM} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ B} \left(t_i^{age} \right)^{\beta} \right]^{1/\beta} \tag{22}$$

where r_A is the number of aircraft in group A.

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case4} = \sum_{group\ B+C} \frac{F_4(t_i^R) - F_4(t_i^{age})}{1 - F_4(t_i^{age})} \tag{23}$$

where $F_4(t) = F(t; \eta_{case4}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case4}}\right)^{\beta}\right]$.

The predicted number of failures at the current moment is:

$$R_{case4}^{now} = \sum_{group\ A} 2 \cdot F_4\left(t_i^{\text{mod}} - t_{DM} \right) + \sum_{group\ B+C} F_4\left(t_i^{age} \right) \tag{24}$$

In the formulas for η_{case4} and R_{case4}^{now}, $t_i^{mod} - t_{DM}$ is replaced by zero if it appears to be negative.

As in Case 3, the number of actual failures includes aircraft in groups B and C, with cracks of any size at retirement ($t_i^{mod} - t_{DM} \le t_i^R$).

Case 5: Similar to Case 1, except the inspected crack-free airplanes (group B) are not adjusted for the time to grow from detectable to dangerous-sized damage.

In this case, aircraft from group A and C are treated as in Case 1. Aircraft in group B are entered as a suspension data point at their current age, $t_i = t_i^{age}$ (i.e., similar to group C). So, for group B, there are no adjustments anymore for the crack to grow from detectable to dangerous-sized damage. Therefore:

$$\eta_{case5} = \left[\frac{1}{r_A} \sum_{group\ A} \left(t_i^{\mathrm{mod}} + t_{MO} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ B+C} \left(t_i^{age} \right)^{\beta} \right]^{1/\beta} \tag{25}$$

where r_A is the number of aircraft in group A.

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case5} = \sum_{group\ B+C} \frac{F_5(t_i^R) - F_5(t_i^{age})}{1 - F_5(t_i^{age})} \tag{26}$$

where $F_5(t) = F(t; \eta_{case5}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case5}} \right)^{\beta} \right]$.

The predicted number of failures at the current moment is:

$$R_{case5}^{now} = \sum_{group\ A} 2 \cdot F_5\left(t_i^{\mathrm{mod}} + t_{MO} \right) + \sum_{group\ B+C} F_5\left(t_i^{age} \right) \tag{27}$$

Case 6: Similar to Case 1, except a 3-parameter Weibayes is used.

In this case, the 3-parameter Weibayes is used (see Chapter 3), meaning that all failure times are shifted to the left with a fixed value named location parameter. The location parameter γ is taken to be the time to grow from a moderate-sized crack (what the fatigue shape parameter β is currently based on) to obvious size (i.e., $\gamma = t_{MO}$). Everything else is taken as in Case 1. Therefore:

$$\begin{aligned} &\left[\frac{1}{r_A} \sum_{group\ A} \left(t_i^{\mathrm{mod}} + t_{MO} - t_{MO} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ B} \left(t_i^{age} + t_{DM} + t_{MO} - t_{MO} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ C} \left(t_i^{age} - t_{MO} \right)^{\beta} \right]^{1/\beta} = \\ &= \left[\frac{1}{r_A} \sum_{group\ A} \left(t_i^{\mathrm{mod}} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ B} \left(t_i^{age} + t_{DM} \right)^{\beta} + \frac{1}{r_A} \sum_{group\ C} \left(t_i^{age} - t_{MO} \right)^{\beta} \right]^{1/\beta} \end{aligned} \tag{28}$$

where r_A is the number of aircraft in group A. In addition, if any of the terms turns out to be negative, it is taken to be zero.

Next, the predicted number of failures until retirement is calculated as:

$$DA_{case6} = \sum_{group\ B+C} \frac{F_6(t_i^R - t_{MO}) - F_6(t_i^{age} - t_{MO})}{1 - F_6(t_i^{age} - t_{MO})} \tag{29}$$

where $F_6(t) = F(t; \eta_{case6}, \beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{case6}} \right)^{\beta} \right]$ for $t \geq 0$ and $= 0$ for $t < 0$.

The predicted number of failures at the current moment is:

$$R_{case6}^{now} = \sum_{group\ A} 2 \cdot F_6\left(t_i^{\text{mod}}\right) + \sum_{group\ B} F_6\left(t_i^{age} + t_{DM}\right) + \sum_{group\ C} F_6\left(t_i^{age} - t_{MO}\right) \tag{30}$$

5.2 RESULTS OF MONTE CARLO SIMULATIONS

The result of the Monte Carlo simulation is a number of figures. Each such figure gives the results for one combination of $n_{initial}$ and n_{undisc} and for the whole range of P_{insp}. Each figure contains six curves, one for each of the six cases considered. The horizontal axis shows P_{insp} (i.e., the fraction of the fleet that is inspected in the scenario). The vertical axis shows either the correction factor (number of actual failures/number of predicted failures) or the error ratio (number of predicted failures/number of actual failures), ranging from 0.1 to 10 on a logarithmic scale.

As an example, see figure 4, with $n_{initial} = 1$ (i.e., one initial discovered moderate crack) and $n_{undisc} = 0$ (i.e., 0 initially undiscovered), and the correction factor on the vertical axis. Note that Cases 3 and 4 have nearly identical results; their curves are almost indistinguishable.

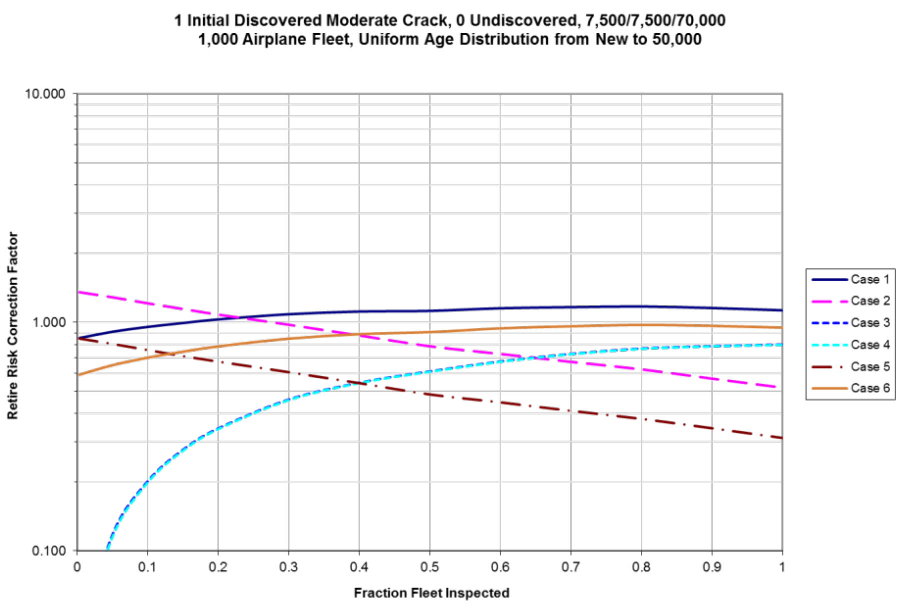


Figure 4. Retire risk-correction factor as a function of fraction inspected for a scenario with one initial discovered and zero initial undiscovered moderate cracks

A correction factor equal to 1 (1.000 in the figure) would mean a perfect match. If the correction factor is smaller than 1, there is an overestimation of risk (i.e., a conservative estimate); if the correction factor is greater than 1, there is an underestimation.

Figures for other combinations of $n_{initial}$ and n_{undisc} can be found in [6].

5.2.1 Which case provides the best results?

In this section, the Monte Carlo simulation results are used to determine which case provides the best results when looking at estimates for DA only (i.e., the Weibull prediction comparison).

Realistically, the fraction of aircraft inspected will be small. Therefore, the predicted number of failures should approximate the actual value, even if few aircraft are inspected. Based on these criteria, the results for Cases 3 and 4 are clearly unfavorable; therefore, they are not presently considered in more detail. Some retire-risk correction factor results for the other Cases are collected in tables 3–8. Green = correction factor closest to 1 (when comparing Cases 1, 2, 5, and 6 for a given parameter setting); Yellow = second best; Amber = third best; Red = result farthest from 1. Note that the correction factor = the actual/predicted number of failures; therefore, if the correction factor is smaller than 1, the predicted number of failures is an overestimation (i.e., a conservative estimate); if the correction factor is greater than 1, there is an underestimation.

Table 3 is for fraction of inspected aircraft equal to 0.3%, or $P_{insp} = 0.003$.

Table 3. Monte Carlo simulation-produced correction factors for various cases and various combinations of input parameters; $P_{insp} = 0.003$

$P_{insp} = 0.003$		Fleet ages uniform 50 … 50,000				Fleet ages uniform 30,000 … 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	0.721	1.145	0.721	0.502	0.688	1.151	0.688	0.484
3	1	0.942	1.482	0.942	0.655	0.934	1.532	0.934	0.656
3	3	1.384	2.151	1.384	0.963	1.422	2.289	1.422	0.999
3	6	2.018	3.092	2.018	1.403	2.136	3.380	2.136	1.501
3	9	2.644	3.997	2.644	1.838	2.823	4.428	2.823	1.983
3	12	3.203	4.786	3.203	2.227	3.572	5.536	3.572	2.510
1	0	0.854	1.356	0.850	0.591	0.850	1.411	0.846	0.597
1	1	1.579	2.472	1.571	1.091	1.609	2.580	1.601	1.129
1	3	2.911	4.502	2.896	2.012	3.066	4.832	3.051	2.151
1	6	4.773	7.288	4.749	3.298	5.149	8.000	5.125	3.613
1	9	6.498	9.813	6.465	4.489	7.339	11.304	7.305	5.149

Observations:

- If $P_{insp} = 0.003$, Case 6 scores best, except if there are very few undiscovered cracks n_{undisc}, in which case it provides a conservative estimate. Cases 1 and 5 have similar results; Case 2 scores worst, except if there are very few undiscovered cracks (see Section 6.2 for a possible argumentation of why Case 2 scores well if there are few undiscovered cracks).
- If $P_{insp} = 0.003$, all Case 2 results are greater than 1, even for small n_{undisc}, indicating underestimates of the actual number of cracks. All 4 cases underestimate the number of failures if the number of undiscovered cracks is more than 1.
- There is no significant difference between the scores for $n_{initial} = 3$ and those for $n_{initial} = 1$.

- Overall, for fleet ages uniform 30,000 ... 50,000, the scores are slightly further away from 1, compared to the situation with fleet ages uniform 50 ... 50,000. However, the differences are very small.

Table 4 shows the results for percentage of inspected aircraft equal to 5%, or $P_{insp} = 0.05$.

Table 4. Monte Carlo simulation-produced correction factors for various cases and various combinations of input parameters; $P_{insp} = 0.05$

$P_{insp} = 0.05$		Fleet ages uniform 50 ... 50,000				Fleet ages uniform 30,000 ... 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	0.761	1.079	0.679	0.547	0.726	1.098	0.654	0.527
3	1	0.970	1.360	0.865	0.696	0.950	1.414	0.856	0.689
3	3	1.367	1.898	1.220	0.981	1.390	2.023	1.253	1.009
3	6	1.887	2.585	1.687	1.355	1.972	2.829	1.778	1.431
3	9	2.310	3.121	2.067	1.660	2.446	3.476	2.207	1.775
3	12	2.665	3.567	2.389	1.916	2.883	4.046	2.602	2.093
1	0	0.910	1.288	0.805	0.650	0.875	1.386	0.828	0.624
1	1	1.582	2.203	1.400	1.129	1.604	2.327	1.437	1.162
1	3	2.641	3.637	2.339	1.885	2.806	3.999	2.515	2.033
1	6	3.873	5.262	3.435	2.765	4.180	5.884	3.749	3.029
1	9	4.563	6.165	4.049	3.261	5.250	7.352	4.708	3.806

Observations:

- Main difference with the $P_{insp} = 0.003$ situation is that for $P_{insp} = 0.05$; Case 5 scores slightly better than Case 1, whereas for $P_{insp} = 0.003$; Cases 1 and 5 score equally well.
- Overall, the scores for $P_{insp} = 0.05$ are better than the scores for $P_{insp} = 0.003$. This makes sense because, for $P_{insp} = 0.05$, more aircraft are inspected; therefore more data are available to support the analysis.

Table 5 shows the results for the percentage of inspected aircraft equal to 10%, or $P_{insp} = 0.1$.

Table 5. Monte Carlo simulation-produced correction factors for various cases and various combinations of input parameters; $P_{insp} = 0.1$

$P_{insp} = 0.1$		Fleet ages uniform 50 ... 50,000				Fleet ages uniform 30,000 ... 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	0.802	1.020	0.641	0.591	0.750	1.028	0.611	0.559
3	1	1.001	1.263	0.801	0.738	0.973	1.313	0.793	0.725
3	3	1.359	1.696	1.090	1.001	1.353	1.794	1.104	1.008
3	6	1.742	2.147	1.402	1.285	1.804	2.354	1.473	1.345
3	9	2.065	2.522	1.856	1.524	2.167	2.802	1.773	1.616
3	12	2.292	2.780	1.666	1.694	2.417	3.119	1.981	1.803
1	0	0.956	1.212	0.757	0.701	0.903	1.344	0.809	0.655
1	1	1.565	1.961	1.240	1.148	1.591	2.099	1.289	1.184
1	3	2.458	3.035	1.950	1.803	2.556	3.320	2.072	1.903
1	6	3.162	3.867	2.516	2.320	3.396	4.353	2.757	2.528
1	9	3.544	4.325	2.832	2.603	3.838	4.866	3.120	2.858

Observations:

- The overall observations are similar to those for the previous table.

Table 6 shows the results for percentage of inspected aircraft equal to 100%, or $P_{insp} = 1$, which means that all aircraft are being inspected.

Table 6. Monte Carlo simulation-produced correction factors for various cases and various combinations of input parameters; $P_{insp} = 1$

$P_{insp} = 1$		Fleet ages uniform 50 ... 50,000				Fleet ages uniform 30,000 ... 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	1.037	0.488	0.298	0.872	0.825	0.471	0.250	0.711
3	1	1.036	0.493	0.304	0.872	0.821	0.466	0.251	0.708
3	3	1.027	0.502	0.312	0.866	0.815	0.472	0.253	0.703
3	6	1.044	0.526	0.335	0.881	0.820	0.478	0.261	0.708
3	9	1.028	0.533	0.346	0.870	0.811	0.478	0.264	0.701
3	12	1.031	0.551	0.363	0.874	0.815	0.488	0.270	0.704
1	0	1.130	0.518	0.312	0.949	0.943	0.529	0.281	0.813
1	1	1.043	0.486	0.295	0.877	0.818	0.462	0.246	0.705
1	3	1.037	0.493	0.305	0.873	0.829	0.471	0.254	0.715
1	6	1.037	0.509	0.322	0.875	0.805	0.469	0.252	0.695
1	9	1.030	0.524	0.337	0.871	0.817	0.478	0.262	0.705

Observations:

- The overall picture is different from that in the previous tables. If $P_{insp} = 1$, Case 1 scores best, followed by Case 6. Cases 5 and 2 do not score well, although Case 2 scores slightly better than Case 5.
- With all aircraft inspected ($P_{insp} = 1$), Case 2 provides an overestimation of the number of failures of approximately a factor 2. For Case 5, this is a factor 3 to 4.
- Whereas for fleet ages uniform 50 ... 50,000, the results for $P_{insp} = 1$ and Case 1 slightly underestimate the number of failures (correction factor greater than 1); for fleet ages uniform 30,000 ... 50,000, these results for $P_{insp} = 1$ and Case 1 are slightly conservative.
- If $P_{insp} = 1$, the scoring of the cases is independent of the value of n_{undisc}; this makes sense because with all aircraft inspected, the initially undiscovered cracks will be discovered during inspection.

Overall observations:

- For low P_{insp}, Case 6 appears to provide the best predictions of the number of airplanes with failure due to wear-out failures, followed by Cases 1 and 5. Case 2 does not score very well, except if $n_{undisc} = 0$. Cases 3 and 4 score worst and are similar to each other.

Table 7 provides some results for the correction factor for the number of failures at the *current* moment, R_{case}^{now}. The fraction inspected is taken to be $P_{insp} = 0.05$.

Table 7. Monte Carlo simulation-produced correction factors for the number of failures at the current moment, for various cases and various combinations of input parameters; $P_{insp} = 0.05$

$P_{insp} = 0.05$		Fleet ages uniform 50 ... 50,000				Fleet ages uniform 30,000 ... 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	0.996	2.155	0.994	0.996	0.998	2.087	0.997	0.998
3	1	0.996	2.730	0.995	0.997	0.998	2.628	0.998	0.999
3	3	0.997	3.800	0.995	0.998	0.999	3.677	0.998	0.999
3	6	0.997	5.240	0.995	0.998	0.998	5.105	0.998	0.999
3	9	0.997	6.470	0.995	0.997	0.998	6.248	0.997	0.999
3	12	0.996	7.532	0.994	0.997	0.998	7.345	0.997	0.998
1	0	0.998	2.524	0.997	0.998	0.999	2.494	0.999	0.999
1	1	0.999	4.354	0.998	0.999	0.999	4.170	0.999	0.999
1	3	0.998	7.227	0.998	0.999	0.999	7.063	0.999	0.999
1	6	0.998	10.658	0.997	0.998	0.999	10.426	0.999	0.999
1	9	0.997	12.733	0.996	0.997	0.999	12.918	0.998	0.999

Some observations:

- Clearly, these predictions do not work out for Case 2. (See section 6.2 for a possible argumentation of why this might be the case.)
- For Cases 1, 5, and 6, the scores are independent of n_{undisc}.
- A look at the Monte Carlo simulation results reveals that for other values for $P_{insp,}$ the scores for Cases 1, 5, and 6 are also almost independent of P_{insp}. For Case 2, the scores get better with increasing P_{insp}, but they approximate the scores for the other cases only if $P_{insp} = 1$.

Another result produced by the Monte Carlo simulations were standard deviations for η, for the correction factor and for the error ratio. A large standard deviation indicates that the data points coming out of the Monte Carlo iterations can spread far from the mean, and a small standard deviation indicates that they are clustered closely around the mean. The typical number of Monte Carlo simulation iterations used was $M = 10\,000$ (although sometimes values up to $M = 100{,}000$ were used). If the correction factor resulting from the ith iteration is denoted by CF_i, then the standard deviation is calculated through:

$$SD_{GF} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(GF_i\right)^2 - \left(\frac{1}{M}\sum_{i=1}^{M}GF_i\right)^2} \tag{31}$$

Some results for $P_{insp} = 0.05$ are given in table 8.

Table 8. Monte Carlo simulation-produced standard deviations for correction factors for the number of failures at retirement, for various cases and various combinations of input parameters; $P_{insp} = 0.05$

$P_{insp} = 0.05$		Fleet ages uniform 50 ... 50,000				Fleet ages uniform 30,000 ... 50,000			
$n_{initial}$	n_{undisc}	Case 1	Case 2	Case 5	Case 6	Case 1	Case 2	Case 5	Case 6
3	0	0.400	0.552	0.356	0.287	0.409	0.598	0.369	0.297
3	1	0.449	0.614	0.400	0.322	0.451	0.657	0.406	0.327
3	3	0.523	0.711	0.466	0.375	0.551	0.796	0.496	0.400
3	6	0.640	0.857	0.569	0.458	0.689	0.978	0.619	0.499
3	9	0.731	0.963	0.650	0.523	0.803	1.130	0.722	0.582
3	12	0.837	1.076	0.743	0.598	0.922	1.294	0.827	0.668
1	0	0.732	1.017	0.647	0.523	0.739	1.135	0.699	0.527
1	1	1.017	1.406	0.900	0.726	1.020	1.483	0.913	0.739
1	3	1.369	1.871	1.210	0.976	1.477	2.111	1.322	1.070
1	6	1.963	2.641	1.736	1.397	2.120	2.967	1.897	1.534
1	9	2.436	3.244	2.152	1.733	2.843	3.946	2.539	2.059

Observations:

- If $P_{insp} = 0.05$, then for each combination of $n_{initial}$ and n_{undisc}, the standard deviation for the correction factor is the lowest for Case 6, followed by Cases 5 and 1; it is the highest for Case 2.
- A look in other Monte Carlo simulation results reveals that this situation changes for higher values of P_{insp}. For Cases 1 and 6, the standard deviation for the retire risk correction factor changes only slightly with increasing P_{insp} (for some combinations of $n_{initial}$ and n_{undisc} it increases, for other combinations it decreases), whereas for Cases 2 and 5, it decreases more significantly. For larger P_{insp}, the standard deviation is lowest for Cases 2 and 5, and highest for Cases 6 and 1.

An overall note is that all Monte Carlo simulation results presented in this chapter are dependent on the setting of the input parameter values in table 2. For other combinations of input parameters, other results may apply.

A final note is that these comparisons considered only the Weibull part of the risk analysis (i.e., predictions of parameter DA). Additional differences will be revealed when comparing the other factors that determine fleet risk (i.e., parameters ND, CP, and IR). This is done in Chapter 6.

5.2.2 Understanding the figures

The purpose of this section is to explain the behavior of the curves. As an example, consider figure 5, for a scenario with 1 initial discovered and 3 initial undiscovered moderate cracks:

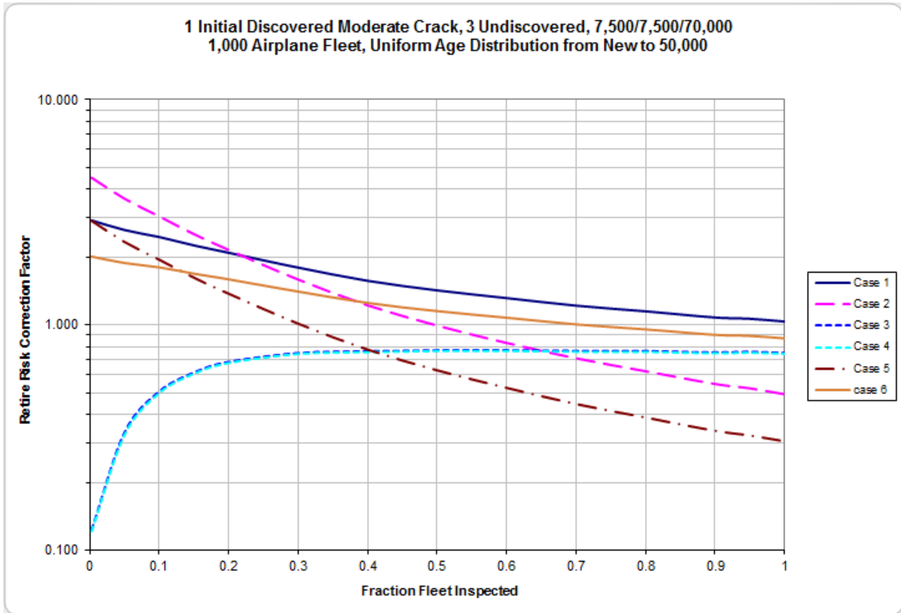


Figure 5. Retire risk-correction factor as a function of fraction inspected, for a scenario with 1 initial discovered and 3 initial undiscovered moderate cracks

Figure 5 shows the results for $n_{initial} = 1$ and $n_{undisc} = 3$. This gives the following (average) numbers of aircraft per group, where n_{small} is the initial number of less-than-moderate-sized cracks, and $P_{insp}^{vol} = (P_{insp} \cdot N - n_{initial})/(N - n_{initial}) \approx P_{insp}$ (this approximation is accurate for $P_{insp}^{vol} \geq 0.01$):

- $|A1| + |A2| = n_{initial} = 1$
- $|A3| = P_{insp}^{vol} \cdot (n_{undisc} + n_{small}) \approx P_{insp} \cdot (3 + n_{small})$
- $|B| = P_{insp} \cdot N - |A|$
- $|C| = (1 - P_{insp}) \cdot N$

The results of a Monte Carlo simulation of this situation show that $n_{small} = 5$ may be a good estimate; therefore:

- $|A1| + |A2| = 1$
- $|A3| \approx 8 \cdot P_{insp}$
- $|B| \approx 992 \cdot P_{insp} - 1$
- $|C| = (1 - P_{insp}) \cdot 1000$

This means the numbers of aircraft in groups A3 and B get larger with increasing P_{insp}; the number of aircraft in group C gets smaller. Depending on the case, this has an increasing or a decreasing effect on the value for η_{case}.

The predicted number of cracks is given by DA, which can be rewritten as:

$$DA=\sum_{i=1}^{N_S}\left(1-\exp\left[\left(\frac{t_i^{age}}{\eta}\right)^{\beta}-\left(\frac{t_i^{R}}{\eta}\right)^{\beta}\right]\right)=\sum_{i=1}^{N_S}\left(1-\exp\left[\frac{\left(t_i^{age}\right)^{\beta}-\left(t_i^{R}\right)^{\beta}}{(\eta)^{\beta}}\right]\right)=N_S-\sum_{i=1}^{N_S}\exp\left[\frac{\left(t_i^{age}\right)^{\beta}-\left(t_i^{R}\right)^{\beta}}{(\eta)^{\beta}}\right] \quad (32)$$

Therefore, because $\frac{(t_i^{age})^{\beta}-(t_i^R)^{\beta}}{(\eta)^{\beta}}$ is negative and $\exp(-x)$ is a decreasing function, if $(\eta)^{\beta}$ increases, $\left|\frac{(t_i^{age})^{\beta}-(t_i^R)^{\beta}}{(\eta)^{\beta}}\right|$ decreases, $\exp\left[\frac{(t_i^{age})^{\beta}-(t_i^R)^{\beta}}{(\eta)^{\beta}}\right]$ increases and DA decreases, provided N_S is constant. Figure 5 shows curves that represent the actual number of cracks divided by DA. The actual number of cracks is independent of P_{insp}; therefore, if DA is decreasing, the curves are increasing. Summarizing: if $(\eta)^{\beta}$ increases (and N_S is constant), the curves increase from left to right; if $(\eta)^{\beta}$ decreases, the curves decrease.

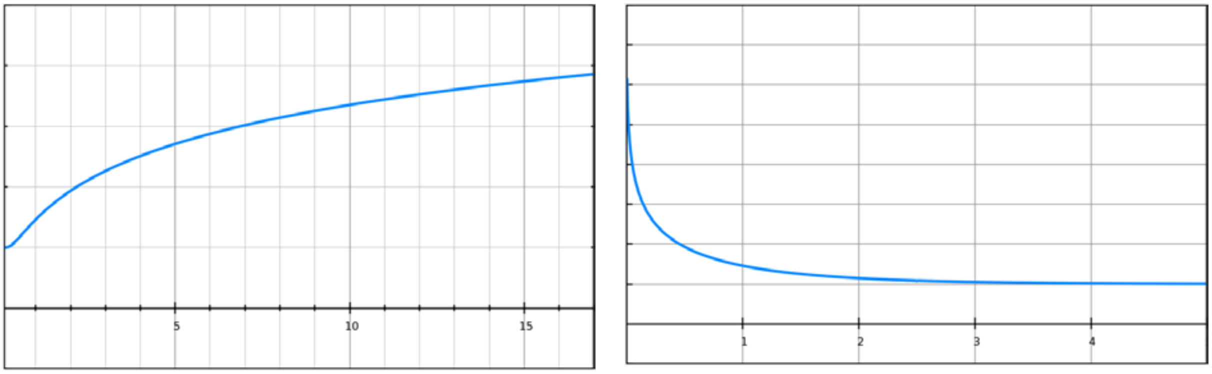
In $\left(t_i^{age}\right)^{\beta}-(t_i^R)^{\beta}$, note that $\left(t_i^{age}\right)^{\beta}$ is a value between $(50)^4=6.25\cdot 10^6$ and $(50\,000)^4=6.25\cdot 10^{18}$, while $(t_i^R)^{\beta}=70\,000^4=24.01\cdot 10^{18}$. Therefore, $\left(t_i^{age}\right)^{\beta}-(t_i^R)^{\beta}$ is between $-24.01\cdot 10^{18}$ and $-17.76\cdot 10^{18}$; on average, $\left(t_i^{age}\right)^{\beta}-(t_i^R)^{\beta}\approx -22.57\cdot 10^{18}$ hence $DA=N_S-\sum_{i=1}^{N_S}\exp\left[\frac{(t_i^{age})^{\beta}-(t_i^R)^{\beta}}{(\eta)^{\beta}}\right]\approx N_S-N_S\cdot\exp\left[\frac{-22.57\cdot 10^{18}}{(\eta)^{\beta}}\right]$. Denote the number of actual failures by AF, then the correction factor is $\approx AF/(N_S-N_S\cdot\exp\left[\frac{-22.57\cdot 10^{18}}{(\eta)^{\beta}}\right])$. On a logarithmic scale, the factor is $\log\left(\frac{AF}{N_S-N_S\cdot\exp\left[-22.57\cdot 10^{18}/(\eta)^{\beta}\right]}\right)=\log(AF/N_S)-\log\left(1-\exp\left[-22.57\cdot 10^{18}/(\eta)^{\beta}\right]\right)$.

Note that from the Monte Carlo simulation results of this scenario:

- For Cases 1, 5, and 6, $AF=51$ (which counts only cracks of obvious size).
- For Case 2, $AF=79$ (which counts cracks of at least moderate size).
- For Cases 3 and 4, $AF=113$ (which counts cracks of at least detectable size).

If $AF\approx 51$ and $N_S\approx 996$, $\log(AF/N_S)\approx -3$ (for log=ln; $^{10}\log(AF/N_S)=-1.3$). As an illustration, the following are plots of $f(x)=1-\ln(1-\exp(-1/x))$ and
$g(x)=1-\ln(1-\exp(-x))$.

These plots are very rough approximations of the behavior of the correction factor as a function of the scale parameter, but they show that on a logarithmic scale, if x is increasing (decreasing) in a linear way, then the correction factor also increases (decreases), but in a less-than-linear way (see figure 6).



$f(x) = 1 - \ln(1 - \exp\left(-\frac{1}{x}\right))$ $g(x) = 1 - \ln(1 - \exp(-x))$

Figure 6. Plot of functions *f(x)* = 1 – ln(1-exp(-1/*x*)) and *g(x)* = 1 – ln(1-exp(-*x*)), which (crudely) illustrate behavior of correction factor as a function of scale parameter

For **Case 1**:

$$\left(\eta_{case1}\right)^{\beta} = \frac{1}{r_A} \sum_{group\ A} \left(t_i^{\mathrm{mod}} + t_{MO}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ B} \left(t_i^{age} + t_{DM} + t_{MO}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ C} \left(t_i^{age}\right)^{\beta} \qquad (33)$$

The first term (group A) has $\approx 1 + 8 \cdot P_{insp}$ elements, which increases from approximately 1 to 9 as P_{insp} increases from 0.003 to 1. The second term (group B) has $\approx 992 \cdot P_{insp} - 1$ elements, which increases from approximately 2 to 991. The number of elements in group C is $(1 - P_{insp}) \cdot 1000$, which decreases from 997 to 0. Because the terms for group B, $(t_i^{age} + t_{DM} + t_{MO})^{\beta}$, include those for group C, $(t_i^{age})^{\beta}$, there is a (minor) net increase in $\sum_{group\ B}(t_i^{age} + t_{DM} + t_{MO})^{\beta} + \sum_{group\ C}(t_i^{age})^{\beta}$ as P_{insp} increases from 0.003 to 1, which is slightly dampened because of $|B| + |C|$ decreasing from 999 to 991. Conversely, r_A, the number of aircraft in group A, is also increasing more significantly from 1 to 9. Overall, $(\eta_{case1})^{\beta}$ decreases, resulting in a decrease of the correction factor and a decrease of the curve for Case 1.

For **Case 2**:

$$\left(\eta_{case2}\right)^{\beta} = \frac{1}{r_A} \sum_{group\ A1} \left(t_i^{\mathrm{mod}\ found}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ A2} \left(t_i^{\mathrm{mod}} + t_{MO}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ A3+B+C} \left(t_i^{age}\right)^{\beta} \qquad (34)$$

The terms for group A1 and A2 are independent of P_{insp}. The number of aircraft in group A3+B+C is constant; therefore, there is a net zero increase in $\sum_{group\ A3+B+C}(t_i^{age})^{\beta}$ as P_{insp} is increased. The only parameter that changes is r_A, which increases with $|A3|$; therefore, $(\eta_{case2})^{\beta} = \frac{1}{r_A} \cdot Const$, with $r_A = 1 + 8 \cdot P_{insp}$; therefore, $(\eta_{case2})^{\beta}$ is decreasing for increasing P_{insp}, which determines that the curve itself is also decreasing. Because groups A1 and A2 are much smaller than groups A3+B+C together, the following can be approximated:
$(\eta_{case2})^{\beta} \approx \frac{1}{r_A}\sum_{i=1}^{N}\left(t_i^{age}\right)^{\beta} = \frac{1}{r_A}\sum_{i=1}^{N} 50^{\beta} \cdot (i)^{\beta} = \frac{50^{\beta}}{r_A}\sum_{i=1}^{N}(i)^{\beta} = \frac{50^4}{r_A}\sum_{i=1}^{N}(i)^4$ with $\beta = 4$.

An equality known from the literature can be used: $\sum_{i=1}^{N}(i)^4 = \frac{1}{30}(6N^5 + 15N^5 + 10N^3 - N)$ with $N = 1000$ to find $(\eta_{case2})^\beta \approx \frac{50^4}{30r_A}(6N^5 + 15N^5 + 10N^3 - N) = 1.25 \cdot 10^{21}/r_A$.

Before, under the assumption that $\left(t_i^{age}\right)^\beta - (t_i^R)^\beta \approx -22.57 \cdot 10^{18}$, it was approximated that the correction factor $\approx AF/(N_S - N_S \cdot \exp\left[\frac{-22.57 \cdot 10^{18}}{(\eta)^\beta}\right])$. Using $(\eta_{case2})^\beta = 1.25 \cdot 10^{21}/r_A$, and using $N_S = N - r_A$, $r_A = 1 + 8 \cdot P_{insp}$ and $AF = 79$, the correction factor is:

$$\approx \frac{\left(\frac{AF}{N - r_A}\right)}{\left(1 - \exp\left[\frac{-22.57 \cdot 10^{18} \cdot r_A}{1.25 \cdot 10^{21}}\right]\right)} = \frac{\left(\frac{AF}{N - r_A}\right)}{\left(1 - \exp\left[-\frac{r_A}{55.4}\right]\right)} \approx \frac{\left(\frac{79}{999 - 8 \cdot P_{insp}}\right)}{\left(1 - \exp\left[\frac{1 + 8 \cdot P_{insp}}{55.4}\right]\right)} \tag{35}$$

The plot of function $h(x) = \ln\left(\frac{79}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$ is given in figure 7 ($x \in [0,1]$, $h(x) \in [-1,1]$):

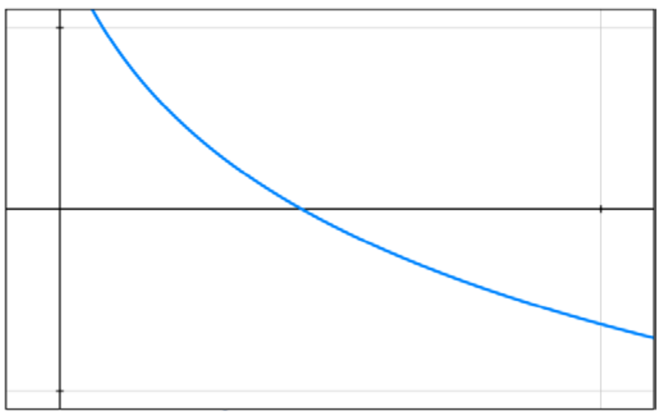


Figure 7. Plot of function $\boldsymbol{h(x) = \ln\left(\frac{79}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))}$, for $\boldsymbol{x \in [0,1]}$ **and** $\boldsymbol{h(x) \in [-1,1]}$

This shape is not too far off from the curve for Case 2 in figure 5.

It is also noted that Case 2 is based on the TARA approach, but with one difference: In the sum for DA, TARA uses the entire fleet of N aircraft, whereas Case 2 uses only groups B+C (i.e., excluding group A; this was assumed to be a mistaken omission in the TARA documentation).

Including group A in the sum for DA leads to the following approximation for the correction factor: $\approx AF/(N - N \cdot \exp\left[\frac{-22.57 \cdot 10^{18}}{(\eta)^\beta}\right])$. Using $(\eta_{case2})^\beta = 1.25 \cdot 10^{21}/r_A$, $r_A = 1 + 8 \cdot P_{insp}$, $N = 1000$ and $AF = 79$, the following correction factor is found:

$$\approx \frac{\left(\frac{AF}{N}\right)}{\left(1-\exp\left[\frac{-22.57\cdot 10^{18}\cdot r_A}{1.25\cdot 10^{21}}\right]\right)} = \frac{\left(\frac{AF}{N}\right)}{\left(1-\exp\left[-\frac{r_A}{55.4}\right]\right)} \approx \frac{\left(\frac{79}{1000}\right)}{\left(1-\exp\left[\frac{1+8\cdot P_{insp}}{55.4}\right]\right)} \tag{36}$$

The plot of function $h(x) = \ln\left(\frac{79}{1000}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$ is given in figure 8 ($x \in [0,1]$, $h(x) \in [-1,1]$) (blue curve), together with the curve of the previous plot (red curve). There is hardly any difference, likely because group A is very small.

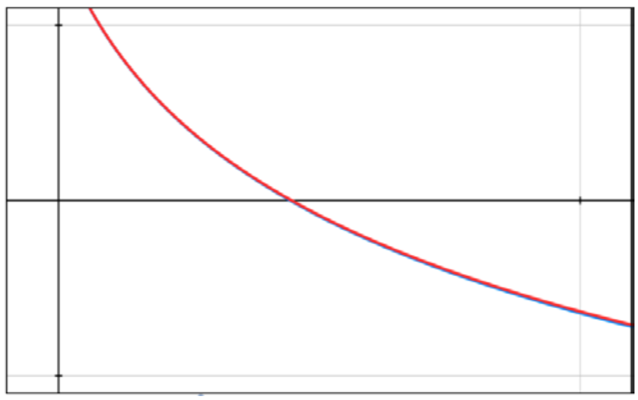


Figure 8. Blue curve: plot of function $h(x) = \ln\left(\frac{79}{1000}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$. Red curve: previous plot (figure 7)

Finally, recall that TARA considers cracks of any size, so it is uncertain that aircraft in group C are free of failures. Therefore, they should not officially be included in the sum for DA (this is further explained in section 6.2). Excluding group C and group A from the sum for DA leads to the following approximation for the correction factor:
$\approx AF/(|B| - |B| \cdot \exp\left[\frac{-22.57\cdot 10^{18}}{(\eta)^\beta}\right])$, where $|B|$ is the number of aircraft in group B, which is approximately equal to $992 \cdot P_{insp} - 1$. Using $(\eta_{case2})^\beta = 1.25 \cdot 10^{21}/r_A$, $r_A = 1 + 8 \cdot P_{insp}$ and $AF = 79$, the correction factor is:

$$\approx \frac{\left(\frac{AF}{|B|}\right)}{\left(1-\exp\left[\frac{-22.57\cdot 10^{18}\cdot r_A}{1.25\cdot 10^{21}}\right]\right)} = \frac{\left(\frac{AF}{992\cdot P_{insp}-1}\right)}{\left(1-\exp\left[-\frac{r_A}{55.4}\right]\right)} \approx \frac{\left(\frac{79}{992\cdot P_{insp}-1}\right)}{\left(1-\exp\left[\frac{1+8\cdot P_{insp}}{55.4}\right]\right)} \tag{37}$$

The plot of function $h(x) = \ln\left(\frac{79}{992x-1}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$ is shown in figure 9 ($x \in [0,1]$, $h(x) \in [-1,1]$) (blue curve), together with the curve of the original Case 2 plot including groups B+C (red curve).

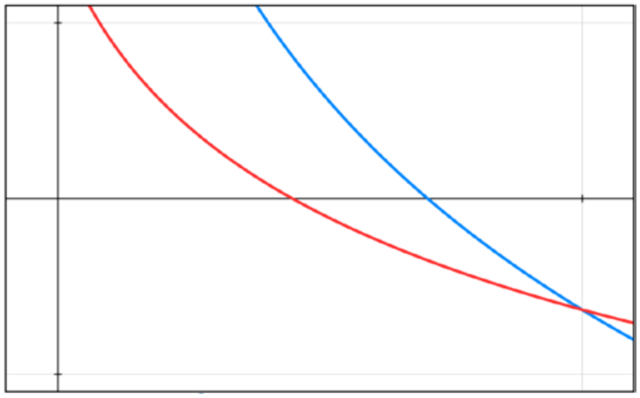


Figure 9. Blue curve: plot of function $h(x) = \ln\left(\frac{79}{992x-1}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$. Red curve: previous plot (figure 7)

This time, there is a major difference: the blue curve (excluding group C) is much steeper and higher. It seems that excluding group C is not a good idea, even though this group includes uninspected aircraft for which the cracking status is unknown.

For **Case 3**:

$$(\eta_{case3})^{\beta} = \frac{1}{r_A}\sum_{group\ A1}\left(t_i^{\text{mod}\ found}\right)^{\beta} + \frac{1}{r_A}\sum_{group\ A2}\left(t_i^{\text{mod}} + t_{MO}\right)^{\beta} + \frac{1}{r_A}\sum_{group\ A3+B}\left(t_i^{age}\right)^{\beta} \qquad (38)$$

The difference with Case 2 is that group C is no longer used in the calculations for scale parameter. For $P_{insp} \geq 0.1$, the number of elements in group A3+B is still much greater than the number of elements in group A1+A2, and this is equal to $1000 \cdot P_{insp} - 1$. Therefore, $(\eta_{case3})^{\beta}$ could be approximated by $(\eta_{case3})^{\beta} \approx \frac{1000}{r_A} \cdot P_{insp} \cdot (t_i^{age})^{\beta}$. It should be noted that this is similar to the approximation in Case 2, except for the factor P_{insp}. Therefore, in a similar way as for Case 2, the correction factor can be approximated as: $\approx \frac{\left(\frac{AF}{999-8\cdot P_{insp}}\right)}{\left(1-\exp\left[-\frac{1+8\cdot P_{insp}}{55.4\cdot P_{insp}}\right]\right)}$. Using $AF = 113$, the plot of function $h(x) = \ln\left(\frac{113}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4x}))$ is given in figure 10 ($x \in [0,1]$, $h(x) \in [-2,0]$). This is quite similar to the curve for Case 3 in figure 5.

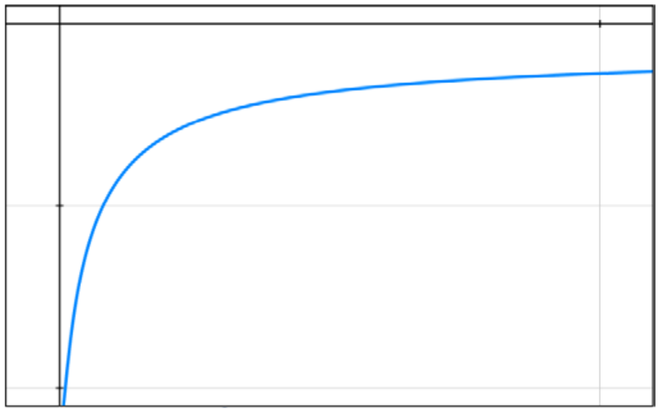


Figure 10. Plot of function $h(x) = \ln\left(\frac{113}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4x}))$, for $x \in [0,1]$ and $h(x) \in [-2,0]$

The version that excludes group C from the sum for DA as well can also be tried; therefore, the sum for DA includes only group B. Similar to Case 2, the correction factor is approximated by: $\approx \frac{AF/|B|}{1-\exp\left[\frac{-22.57\cdot 10^{18}}{(\eta)^{\beta}}\right]} \approx \frac{\left(\frac{113}{992\cdot P_{insp}-1}\right)}{\left(1-\exp\left[-\frac{1+8\cdot P_{insp}}{55.4\cdot P_{insp}}\right]\right)}$. The plot of function $h(x) = \ln\left(\frac{113}{992x-1}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4x}))$ is given in figure 11 ($x \in [0,1]$, $h(x) \in [-1,1]$) (blue curve), together with the curve for the original Case 3, including groups B+C (red curve).

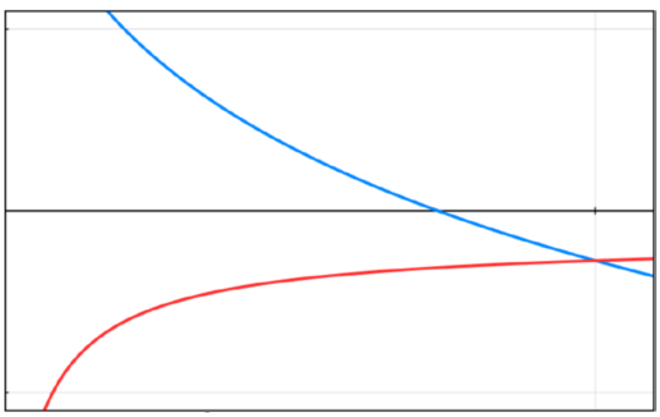


Figure 11. Blue curve: plot of function $h(x) = \ln\left(\frac{113}{992x-1}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4x}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$. Red curve: previous plot (figure 10)

The predictions now underestimate the number of failures, but there is no improvement otherwise. Also for this case, it seems that excluding group C does not help the predictions.

For **Case 4**:

$$(\eta_{case4})^{\beta} = \frac{1}{r_A}\sum_{group\ A}\left(t_i^{\mathrm{mod}} - t_{DM}\right)^{\beta} + \frac{1}{r_A}\sum_{group\ B}\left(t_i^{age}\right)^{\beta} \quad (39)$$

As in Case 3, the behavior of the scale parameter is largely determined by group B. Because the terms for group B in Case 4 are equal to the terms for group B in Case 3, and the contribution of the group A terms is small and close to those for group A in Case 3, the curves for Cases 3 and 4 appear to be almost indistinguishable.

For **Case 5**:

$$(\eta_{case5})^{\beta} = \frac{1}{r_A} \sum_{group\ A} \left(t_i^{\mathrm{mod}} - t_{MO}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ B+C} \left(t_i^{age}\right)^{\beta} \tag{40}$$

This can be compared with the expression for $(\eta_{case2})^{\beta}$ for Case 2: For groups B and C no difference is seen between the contributing terms. For group A2, there is also no difference, but for group A1, there is a minor difference $(t_i^{mod} + t_{MO})^{\beta}$ vs. $(t_i^{modfound})^{\beta}$, with the former (for Case 5) being larger than the latter (for Case 2). This leaves group A3. For Case 2, the contributory term for aircraft in group A3 is $(t_i^{age})^{\beta}$; for Case 5 the contributory term is $(t_i^{mod} + t_{MO})^{\beta}$. Since group A3 consists of aircraft for which a crack has been discovered during follow-on inspection, on average $t_i^{mod} \approx t_i^{age}$ for these aircraft. Therefore, typically, $(t_i^{mod} + t_{MO})^{\beta} > t_i^{age}$, and $(\eta_{case5})^{\beta} > (\eta_{case2})^{\beta}$ is expected; however, the differences seem to be very minor.

A more significant difference between Cases 5 and 2 seems to be the count of the actual number of failures. For Case 2, this is $AF = 79$ (counting cracks of at least moderate size), whereas for Case 5, it is $AF = 51$ (counting cracks of obvious size only). Curves for $h(x) = \ln\left(\frac{79}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$ (top curve) and $g(x) = \ln\left(\frac{51}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$ (bottom curve) are provided in figure 12. These appear similar to the ones for Cases 2 and 5 in figure 5.

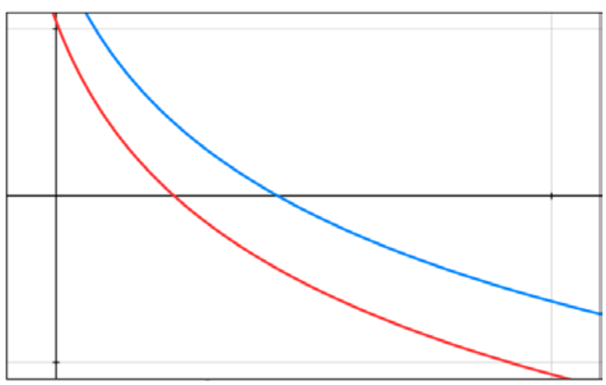


Figure 12. Blue curve: plot of function $h(x) = \ln\left(\frac{79}{999-8x}\right) - \ln\left(1 - \exp(-\frac{1+8x}{55.4})\right)$,, for $x \in [0, 1]$ and $h(x) \in [-1, 1]$. Red curve: plot of function $g(x) = \ln\left(\frac{51}{999-8x}\right) - \ln(1 - \exp(-\frac{1+8x}{55.4}))$

Note that in figure 5, the results for Cases 1 and 5 are equal if $P_{insp} = 0$. This can be explained by comparing the formulas for $(\eta_{case1})^{\beta}$ and $(\eta_{case5})^{\beta}$ and considering that $|B| = 0$ if $P_{insp} = 0$.

For **Case 6**:

$$\left(\eta_{case6}\right)^{\beta} = \frac{1}{r_A} \sum_{group\ A} \left(t_i^{\text{mod}}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ B} \left(t_i^{age} + t_{DM}\right)^{\beta} + \frac{1}{r_A} \sum_{group\ C} \left(t_i^{age} - t_{MO}\right)^{\beta} \qquad (41)$$

When compared to Case 1, each term in $(\eta_{case6})^{\beta}$ is smaller than the corresponding term in $(\eta_{case1})^{\beta}$; therefore, $(\eta_{case6})^{\beta} < (\eta_{case1})^{\beta}$. In addition, the expression for DA is different. The effects of these differences are not yet immediately clear, however. According to the figure 5, the curve for Case 6 is slightly lower than the curve for Case 1.

Note that the biggest difference between Cases 6 and 1 is not the correction factor but that Case 6 has significantly lower standard deviation of the correction factor (see table 8). The average correction factor and the standard deviation of the correction factor both need to be considered to assess how good the technique is in predicting future failures, which is the fundamental purpose.

Let us consider a different scenario, with $n_{initial} = 1$ (i.e., 1 initial discovered moderate crack) and $n_{undisc} = 0$ (i.e., 0 initially undiscovered).

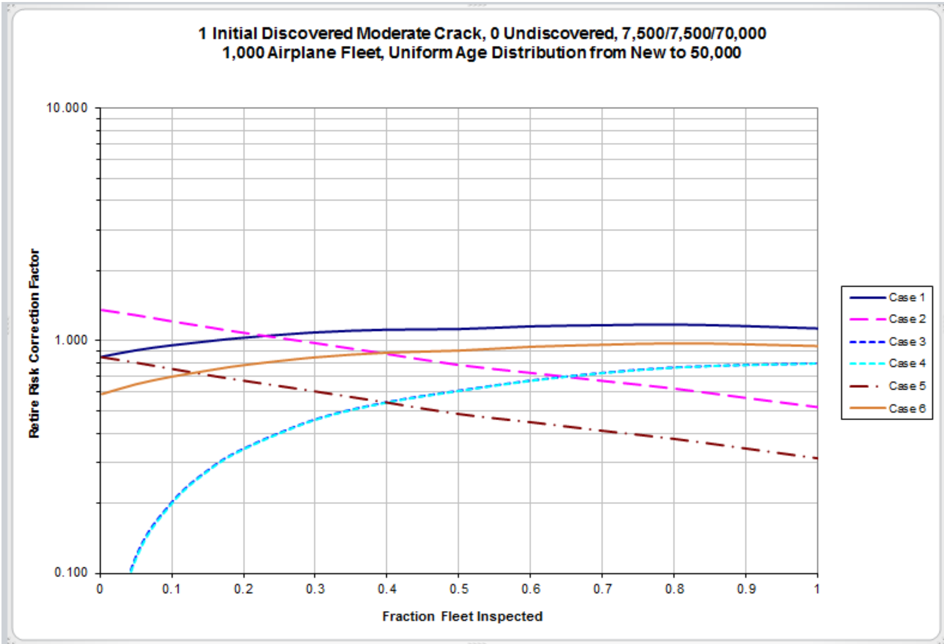


Figure 13. Retire risk correction factor as a function of fraction inspected, for a scenario with 1 initial discovered and 0 initial undiscovered moderate cracks

This gives the following (average) numbers of aircraft per group, where n_{small} is the initial number of less-than-moderate sized cracks, which from the Monte Carlo simulation results is found to be $n_{small} \approx 1.66$:

- $|A1| + |A2| = n_{initial} = 1$
- $|A3| = P_{insp}^{vol} \cdot (n_{undisc} + n_{small}) \approx P_{insp} \cdot (n_{undisc} + n_{small}) = P_{insp} \cdot n_{small} = 1.66\ P_{insp}$
- $|B| = P_{insp} \cdot N - |A| \approx P_{insp} \cdot (N - n_{small}) - 1 = P_{insp} \cdot (N - 1.66) - 1$
- $|C| = (1 - P_{insp}) \cdot N$

Then as P_{insp} increases from 0.1 to 1, the number of aircraft in group A3 increases from 0.166 to 1.66. The number of aircraft in group B increases from 99 to 997, and the number of aircraft in group C decreases from 900 to 0. Parameter r_A increases from approximately 1.17 to approximately 2.7 as P_{insp} increases from 0.1 to 1. In the first figure, r_A increased from 2 to 9.

For **Case 1**, there was a minor net increase in figure 5 due to the contributions from groups B and C, which was reduced by the more significant increase in r_A. In figure 13, there is still a minor net increase due to the contributions from groups B and C; however, the increase in r_A is less significant than in figure 5, leading to a curve for Case 1 that is almost constant.

For **Case 2**, the analysis below figure 5showed that $(\eta_{case2})^{\beta} = \frac{1}{r_A} \cdot Const$. The same situation applies for figure 13, but with r_A increasing less significantly as P_{insp} increases. As in the analysis

below figure 5, the correction factor can be approximated by $\frac{\left(\frac{AF}{N-r_A}\right)}{\left(1-\exp\left[\frac{-r_A}{55.4}\right]\right)} \approx \frac{\left(\frac{AF}{999-1.66 \cdot P_{insp}}\right)}{\left(1-\exp\left[-\frac{1+1.66 \cdot P_{insp}}{55.4}\right]\right)}$, using $r_A = 1 + 1.66\, P_{insp}$. Unfortunately, the initial Monte Carlo simulation results do not give a value for AF, which was equal to 79 in figure 5. However, because the number of initial plus undiscovered moderate cracks is a factor 4 lower in figure 13, AF will be much lower in this case. The plot of function
$h(x) = \ln\left(\frac{AF}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$ is given in figure 14 ($x \in [0,1]$, $h(x) \in [-1,1]$), for an assumed $AF = 25$.

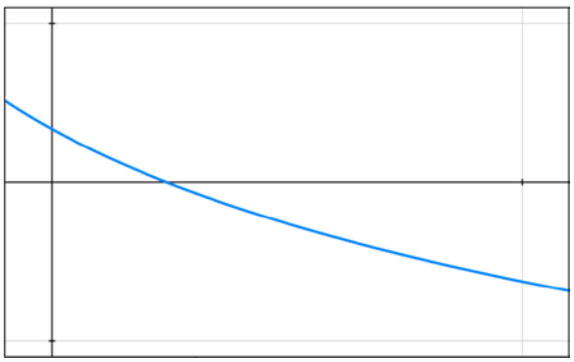


Figure 14. Plot of function $h(x) = \ln\left(\frac{AF}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$ and an assumed $AF = 25$

Including group A in the sum for DA (as is done by TARA) does not provide a noticeable difference, but excluding group C does, as is shown by the blue curve in figure 15:

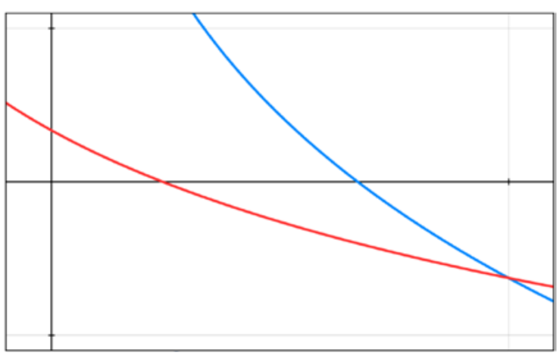


Figure 15. Blue curve: plot of function $h(x) = \ln\left(\frac{25}{998.34x-1}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$. Red curve: Previous plot (figure 14)

Again, excluding group C does not seem like a good idea.

Cases 3 and **4** can be reasoned in the same way, leading to an approximate correction factor of

$\approx \frac{\left(\frac{AF}{999-1.66}\right)}{\left(1-\exp\left[-\frac{1+1.66 \cdot P_{insp}}{55.4 \cdot P_{insp}}\right]\right)}$. Again, there is no value for AF, which was equal to 113 in the first figure. Using $AF = 40$, the plot of function $h(x) = \ln\left(\frac{AF}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4x}))$ is given in figure 16 ($x \in [0,1]$, $h(x) \in [-2,0]$).

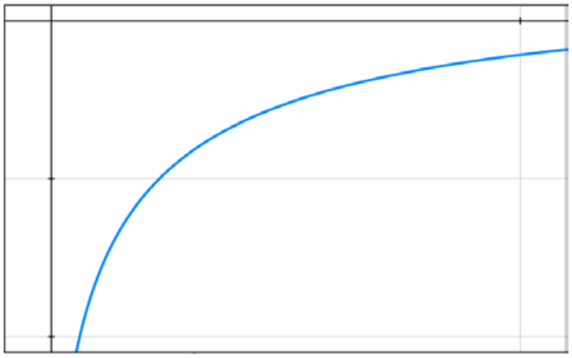


Figure 16. Plot of function $h(x) = \ln\left(\frac{AF}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4x}))$, for $x \in [0,1]$ and $h(x) \in [-2,0]$ and an assumed $AF = 40$

Case 5 reveals that the main difference with Case 2 seems to lie in the number of actual failures. Curves for $h(x) = \ln\left(\frac{25}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$ (top curve) and $g(x) = \ln\left(\frac{15}{999-8x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$ (bottom curve) are provided in figure 17. Indeed, these look similar to the ones for Cases 2 and 5 in figure 13.

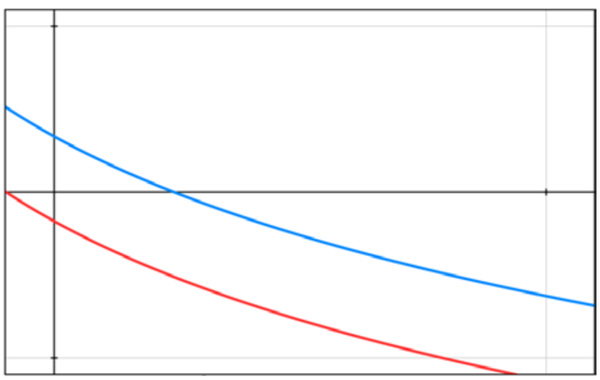


Figure 17. Blue curve: plot of function $h(x) = \ln\left(\frac{25}{999-1.66x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$, for $x \in [0,1]$ and $h(x) \in [-1,1]$. Red curve: plot of $g(x) = \ln\left(\frac{15}{999-8x}\right) - \ln(1 - \exp(-\frac{1+1.66x}{55.4}))$

The second figure for **Case 6** reveals that for Case 6, the situation in figure 13 is similar to the situation in figure 5. The curve is slightly lower than that for Case 1.

6. COMPARISON OF APPROACHES W.R.T. ND, CP, AND IR

This chapter discusses the TARAM and TARA approaches regarding the product $R_T = DA \times ND \times CP \times IR$, as well as regarding the parameters in this product.

6.1 DISCUSSION OF UNCORRECTED FLEET RISK

This section considers the expression for uncorrected fleet risk (i.e., $R_T = DA \times ND \times CP \times IR$) and discusses how the approaches of TARAM and TARA use it.

Given a certain fleet, R_T is the expected number of planeloads of people fatally injured as a result of wear-out failures during the remaining life of the fleet if no corrective action is taken. So if

$R_T = 1$, one accident due to wear-out failures will occur with all onboard fatally injured, or two accidents will occur with 50% of those onboard fatally injured, etc. The product $DA \times ND \times CP \times IR$ is composed of (see figure 18):

- The number of aircraft that gets a wear-out failure before retirement, during the lifetime of the fleet, if no corrective action is taken.
- The probability that the failure is not detected, given that the aircraft has a failure.
- The probability that the aircraft will be in an accident, given that it has an undetected failure.
- The proportion of people onboard fatally injured, given they are in an accident due to an undetected failure.

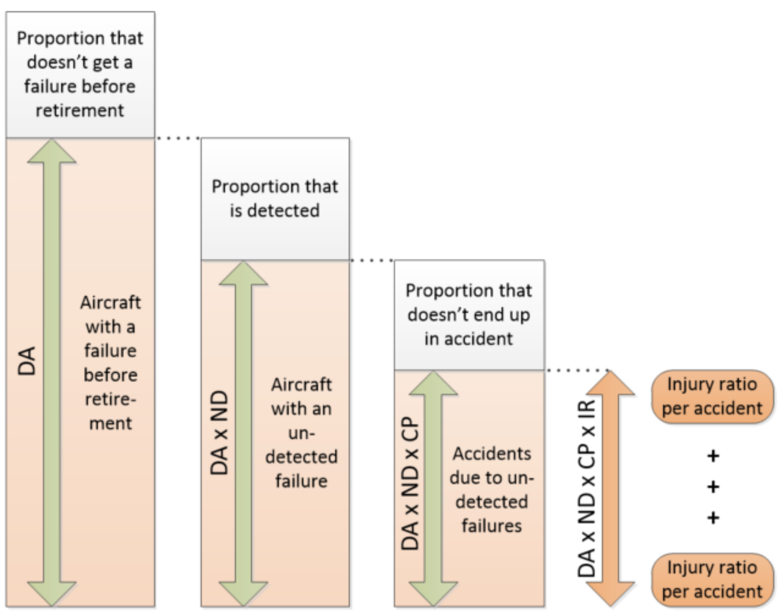


Figure 18. Structure of formula $DA \times ND \times CP \times IR$ for uncorrected fleet risk; the various proportions are not to scale

In TARAM, the term failure is defined to be a dangerous crack (accident or obvious malfunction size damage). The 4 parameters DA, ND, CP and IR are used and multiplied with the following interpretations:

- DA is the expected number of aircraft with a dangerous crack before retirement.
- ND is the probability of not detecting the dangerous crack, such that DA × ND is the expected number of aircraft with an undetected dangerous crack during the lifetime of the fleet.
- CP is the probability that the undetected dangerous crack leads to an accident, such that DA × ND × CP is the number of accidents involving aircraft with an undetected dangerous crack.
- IR is the injury ratio, such that DA × ND × CP × IR is the number of planeloads of people fatally injured in an accident due to an undetected dangerous crack.

In TARA, the term failure is defined to be a crack of any size. The formula DA × ND × CP × IR is used, but with a few differences:

- DA is the number of aircraft with a crack before retirement. The size of the crack is not specified; there is no anchoring to cracks of dangerous size.
- ND is the average probability that a given crack is not detected before resulting in an unsafe outcome or condition throughout the life of the affected fleet. Main difference with the definition in TARAM is that in TARA the crack concerns any sized crack, rather than a dangerous-sized crack. Therefore, DA × ND would be the number of aircraft with a crack of any size that remains undetected during the life of the fleet.
- CP is the probability that a given detected crack reaches an unsafe outcome. Therefore, this starts with a detected crack rather than with an undetected crack, as in figure 18 and TARAM. Therefore, it is not entirely clear what DA × ND × CP depicts. The growth of the crack, from any size to dangerous size, is addressed in CP rather than in DA.
- TARA explicitly considers multiple potential outcomes and provides the injury ratio IR for each potential outcome. In TARAM, this is implicit at most.

Therefore, even though the two approaches use the same formula for uncorrected fleet risk, there are several major differences in the interpretation and determination of the parameters.

- TARAM anchors DA, ND, and CP to dangerous-sized cracks, whereas TARA considers any sized crack.
- TARAM defines CP conditional on undetected cracks, whereas TARA defines CP conditional on detected cracks.
- In TARAM, the growth of the crack is addressed in DA, whereas in TARA, it is addressed in CP.
- TARAM follows the structure of the formula DA × ND × CP × IR as a sequence of conditional probabilities depicted in figure 18; TARA does not.

Sections 6.2–6.7 further analyze these and other differences, per component in the product.

6.2 DISCUSSION OF DA

This section discusses the differences between TARAM and TARA in their use of the expression for DA.

The origin of this expression is the formula for future risk taken from section 4.4.2 in reference [7]; see section A.5 in appendix A. This represents the number of failures at a future moment (e.g., Δ units of time from now), assuming that failed components are not replaced. Abernethy argues that this is equal to the sum over all aircraft in the fleet of the conditional probability that the aircraft will have a failure between now and Δ units of time from now, given that it had no failure until now:

$$\begin{aligned}\text{Number of future failures} &= \sum_{all\ aircraft} P\left(\text{Failure in } [Now, Now+\Delta] \middle| \text{No failure until } Now\right)\\ &= \sum_{\substack{Aircraft\ that\ currently\\ have\ a\ failure}} P\left(\text{Failure in } \left[Now, Now+\Delta\right] \middle| \text{No failure until } Now\right)\\ &+ \sum_{\substack{Aircraft\ that\ do\ not\\ currently\ have\ a\ failure}} P\left(\text{Failure in } \left[Now, Now+\Delta\right] \middle| \text{No failure until } Now\right)\end{aligned} \tag{42}$$

The first term to the right of the equal sign above, which considers aircraft that already have a failure, is zero, because the condition "No failure until *Now*" does not hold true for those aircraft. Therefore, the second term remains, leading to:

$$\begin{aligned}\text{Number of future failures} &= \sum_{\substack{Aircraft\ that\ do\ not\\ currently\ have\ a\ failure}} P\left(\text{Failure in } [Now, Now+\Delta] \middle| \text{No failure until } Now\right)\\ &= \sum_{\substack{Aircraft\ that\ do\ not\\ currently\ have\ a\ failure}} \frac{P\left(\text{Failure in } \left[Now, Now+\Delta\right] \text{AND No failure until } Now\right)}{P\left(\text{No failure until } Now\right)}\\ &= \sum_{\substack{Aircraft\ that\ do\ not\\ currently\ have\ a\ failure}} \frac{P\left(\text{Failure before } Now+\Delta\right) - P\left(\text{Failure before } Now\right)}{1-P\left(\text{Failure before } Now\right)}\\ &= \sum_{\substack{Aircraft\ that\ do\ not\\ currently\ have\ a\ failure}} \frac{F\left(Now+\Delta\right) - F\left(Now\right)}{1-F\left(Now\right)}\end{aligned} \tag{43}$$

where F is the cdf for failure occurrence.

Both TARAM and TARA use this formula to calculate DA, with $\Delta =$ time until retirement. However, TARAM defines failure as a dangerous-sized crack, and TARA defines failure as any sized crack. Since the number of aircraft with any crack will likely be higher than the number of aircraft with a dangerous-sized crack, TARA's DA will likely be higher than TARAM's. However, this difference does not need to be a problem because it may be compensated for in the calculation of ND, CP and IR.

However, there is another main difference, which considers the elements included in the sum. As before, divide the fleet of aircraft into three groups: Aircraft with failures (group A), aircraft that have been inspected and concluded free of failures (group B), and aircraft that have not been inspected (group C). TARAM's DA takes the sum over all aircraft in groups B+C. Aircraft in group B indeed have no failure. For aircraft in group C, TARAM argues that uninspected aircraft do not have dangerous-sized cracks, because that would be obvious. Therefore, with their definition of failure, the sum for DA can include the aircraft in group C as well. (Note that this is also a source of some confusion, though: for DA, it is assumed that dangerous cracks will all be obvious, whereas the nondetection probability ND is assumed non-zero.)

TARA's DA takes the sum over all aircraft in the active fleet (i.e., groups A+B+C). Their definition of failure includes any sized crack; therefore, aircraft in group B can certainly be included. For group C (uninspected aircraft), however, aircraft with failures are not excluded with certainty. Because the cracks can be of any size, their occurrence will not always be obvious. Therefore, respecting Abernethy's formula precludes using group C in the sum. For aircraft in group A, there cannot be discussion: these are aircraft with detected failures; therefore, they should not be included in the sum. As a result, the TARA's DA is likely too high (i.e., a conservative estimate of the number of cracks of any size).

In section 5.2.1, below table 3, Monte Carlo simulation results were compared for various cases, including Case 2, which is the TARA approach, with the difference that aircraft in group A are excluded from the sum for DA. For small fractions of inspected aircraft (P_{insp} small, and group C large), Case 2 provided conservative results for DA, except if the number of undiscovered cracks n_{undisc} was very small. If $n_{undisc} = 0$, Case 2 provided better results than Case 6 (TARAM with 3-parameter Weibull). Possibly, this can be explained now by the above reasoning: If there are no undiscovered cracks ($n_{undisc} = 0$), group C does not include aircraft with failures, leading to (accidental) correct use of the formula for DA. Note, however, that when an analyst is assessing an actual issue, the number of undiscovered moderate-sized cracks in the fleet is unknown; it could be zero, or it could be three or some other number. The approach used by the risk analyst must work for all situations faced by the analyst.

Another observation in section 5.2.1, below table 7, was made regarding the predictions of the number of failures at the current moment. These predictions did not work out at all for Case 2, whereas for Cases 1, 5, and 6, they were quite accurate. This may now be explained as follows: the formula for the number of current failures is taken from section 4.4.1 in reference [7], that is:

$$\text{Number of current failures} = \sum_{Failures} 2 \cdot F\left(t_i\right) + \sum_{Suspensions} F\left(t_i\right) \tag{44}$$

The factor 2 was introduced by Abernethy to reduce bias in earlier versions of the formula. Note that in Case 2, the failures are formed by the aircraft in group A, and TARA takes for the suspensions the aircraft in group B+C. However, in reality, some of the aircraft in group C are not suspensions in the sense that they may have a yet undetected crack. These aircraft should be moved to the first sum in the formula and, consequently, their contribution to the formula should be doubled by the factor 2. Since this is not done (the entire group C is kept in the second sum), the predicted current number of failures is too low, and the correction factors in table 7 are too high. This argumentation is corroborated by the fact that the predictions get worse for larger numbers of undetected failures n_{undisc}. It does not explain why the predictions are also bad if $n_{undisc} = 0$, however.

Another difference between the two approaches is the following: In TARAM, DA is the number of aircraft that have a dangerous-sized crack during the lifetime of the fleet. A dangerous-sized crack is counted only if it occurs between now and when the aircraft is retired: DA takes the sum over $P(\text{Failure in } [Now, Now + \Delta] \mid \text{No failure until } Now)$, where $[Now, Now + \Delta] = [t_i^{age}, t_i^R]$.

TARA counts all cracks of any size if they occur between now and when the aircraft is retired. Aircraft that would develop a crack after retirement are, correctly, not counted. However, it may happen that an aircraft has a crack, but retires before the crack has had the chance to grow to dangerous size, therefore not resulting in any injuries. Such cases are not filtered out in DA, so TARA needs to filter them out in one of the other parameters, presumably CP or ND. This is discussed in sections 6.3–6.7.

Note that the notion that TARAM anchors the Weibull analysis to dangerous cracks is not very apparent from the TARAM handbook [2]. It is not mentioned in the chapter that gives detailed guidance on Risk assessment for wear-out failures (i.e., Chapter 5), but it is hidden in an example in appendix C.1.5.2. In addition, the guidance Chapter 5 is not explicit on which groups of aircraft to include in the sum for DA. The text does indicate that the sum should include only those affected airplanes that have not failed, but the formula for DA is given as $DA = \sum_{fleet} \frac{F(t+t\Delta)-F(t)}{1-F(t)}$, which some readers could interpret as: the sum includes the (entire) fleet.

6.3 DISCUSSION OF ND

TARAM defines ND as the probability that, during future operation and maintenance, a wear-out failure will not be discovered by any means before the cracked element fails. The TARAM handbook gives high-level guidance for calculation of ND, in terms of some factors to consider, including:

- How many cases of crack findings are there?
- How many crack lengths are found?
- What is the estimate of the dangerous event crack size?
- What is the estimate of time to grow from discovered crack size to dangerous-event crack size (review of crack growth curves if they are available; extrapolating a little bit past the critical crack length if the curve stops there)?
- How often is the area visible?
- How was the damage found?
- Are there other ways the damage may be found?

Unfortunately, the TARAM handbook does not provide more practical guidance on how to address these factors to determine ND. This may be why the TARA developers decided to develop a tool to help the user. Reference [3] provides two flow charts that can be used to determine ND. The flow charts provide a sequence of questions; the yes or no answers provide values for ND (see figures 1 and 2 in Chapter 4). These flow charts are user-friendly, but there are a few issues.

The first is related to the structure of the formula for uncorrected fleet risk, $R_T = DA\ \times ND \times CP \times IR$. As shown in figure 18, DA delivers the number of aircraft that have a crack before retirement. The next parameter, ND, is designed to filter out those cracks that are detected during routine operation or maintenance, such that only the undetected cracks remain as contributors to fleet risk.

In TARAM, DA delivers the number of cracks of dangerous size. Many of those cracks will be readily detected during routine operation or maintenance before they result in an accident. Moreover, detection of cracks during the period of growth from detectable size until dangerous size will lead to risk-reducing maintenance activities, which therefore can also be credited in ND. Therefore, the nondetection probability in TARAM will typically be very small.

In TARA, DA delivers the number of cracks of any size. The TARA definition of ND is similar to the one for TARAM: "the average probability that an occurrence of the defect is not detected before resulting in an unsafe outcome or condition throughout the life of the affected fleet." However, when properly respecting the formula for uncorrected fleet risk, $R_T = DA\ \times ND \times CP \times IR$, as shown in figure 18, TARA's ND cannot take credit for detecting cracks that have grown beyond the size as considered in DA. It can take credit only for detection during the period of growth from detectable size until the size considered in DA, and not for the whole period of detection "before resulting in an unsafe outcome or condition." Moreover, since the size of the crack can be anything from detectable to obvious, there will be a lot of variation in the time period in which detection can take place, which TARA should take into account by means of some kind of uncertainty figure.

The confusion about the definition of ND, and the variation in the time period applicable to detection, will be sources of uncertainty regarding the determination of TARA's ND.

Next, consider the flow charts, figures 1 and 2 in Chapter 4. The starting point is one or more aircraft with damage (i.e., a crack of any size). The flow charts take into account how easy it is to inspect the structure to find the damage, whether the design has redundant load paths, and whether the structure is susceptible to WFD. The flow charts ask a sequence of questions, the yes or no answers to which provide values for ND. The values range from $ND = 0.01$ (in case the defect is obvious) to $ND = 0.99$ (in case inspection is not practical and a safe life philosophy is applied, i.e., the part is removed and replaced at predetermined intervals, rather than when it shows signs of fatigue). No uncertainty band is given for the ND values (e.g., accounting for other levels of visibility/accessibility of the crack besides the crack being either obvious or not obvious).

As an example, consider questions 1 (Existing inspections effective) and 2 (Past inspection threshold) in the first flow chart (figure 1), and which are answered conditional on the damage not being obvious. Reference [3] explains that "If inspections are in place, and the damage starts after the threshold for the inspections to begin, then the inspections are likely to find the damage (ND $=$ 0.06). If the damage occurs before the inspection threshold, then the inspections are not considered effective." This implies that according to TARA, for any nonobvious crack, if the airplane is being inspected, the crack will be found with probability $0.94 = 1 - 0.06$, irrespective of any other circumstances, such as the size and location of the crack at the time of inspection and the number of inspections, and irrespective of the remaining questions in the flow chart. A similar reasoning holds true for the other questions. No variation is considered, and the user is not notified of the possible variation.

Finally, consider completeness of the flow charts. The TARAM handbook provides factors to consider in determining ND, see beginning of this section. Some of these factors are considered in the flow charts, others are not:

- *How many cases of crack findings are there?* The second flow-chart does not consider the number of cases of crack findings amongst multiple aircraft, but does consider the occurrence of WFD within one aircraft. WFD can occur when adjacent structural details operate at similar stress levels and develop cracks simultaneously. It can grow from multiple element damage (MED) if it concerns structurally similar elements or from multiple site damage (MSD) if the elements are not similar. It appears difficult to detect these small cracks before they link up and cause catastrophic damage, as reflected in the relatively high probabilities of nondetection: $ND = 0.7$ or $ND = 0.9$ in case of MED, and $ND = 0.9$ or $ND = 0.95$ in case of MSD, depending on whether the area is readily accessible.
- The other factors in the TARAM handbook are not considered (i.e., *Crack lengths found. Estimate of the accident-critical crack size. Estimate of time to grow from discovered crack size to accident-critical crack size. How often is the area visible? How was the damage found? Are there other ways the damage may be found?).*

Also, no consideration is given to such factors as the time until retirement, the number of aircraft in the fleet, the number of aircraft with cracks, the number and type of inspections, the probability that a given aircraft in the fleet is inspected, cracking history, or the number of maintenance cycles during the period of crack growth. There is no evidence that the flow charts address the whole TARA definition of ND, including nondetection "before resulting in an unsafe outcome or condition" or nondetection "throughout the life of the affected fleet."

6.4 DISCUSSION OF CP

TARAM's parameter CP models the probability that an aircraft with an undetected crack of dangerous size in fact ends up in an accident. Depending on the type and location of the dangerous crack, there may be a chance that such aircraft could still safely land without major injuries; however, typically, $CP = 1$.

In TARA, the situation is very different. A causal chain traces the steps from the initial damage condition that is being detected in the fleet (e.g., cracking of a particular component) to an unsafe outcome (e.g., crash, in-flight breakup, runway departure, or individual fatality). Each step has a conditional probability of reaching the next step, and the overall conditional probability of reaching the unsafe outcome is the product of the conditional probabilities from each step in the chain: $CP = PA1 \times PA2 \times PA3 \times PA4$. Notice that the initial condition considers a detected crack, not an undetected crack.

The initial condition is not a dangerous-sized crack, but a crack of any size. Therefore, whereas TARAM addressed crack growth in its parameter DA, TARA needs to address crack growth to accident size in parameter CP.

The CP calculation is supported by an Excel spreadsheet, which requires the following user inputs:

- The user determines whether the damage considered occurs in fuselage, in wing/pylon/empennage, or in landing gear.
- The user selects from pull down menus conditions $A1$ and $A2$ that best correspond to the detected crack. For fuselage, there are 14 options for condition $A1$ and a maximum of 9 options for condition $A2$ (given a specific choice for $A1$, most of the options for $A2$ are not applicable); for wing/pylon/empennage, these numbers are 16 and 13; for landing gear they are 6 and 3.
- The user estimates the number of cycles required for the damage to progress from the initial condition to condition $A1$ as percentage of the retirement age t_R (also in number of cycles). This leads to probability $PA1$, i.e., $PA1 = 1$ if percentage between 0 and 10%; $PA1 = 0.75$, if percentage between 11 and 30%; $PA1 = 0.5$, if percentage between 31 and 50%; $PA1 = 0.1$, if percentage between 51 and 70%; $PA1 = 0.01$, if percentage between 71 and 90%; $PA1 = 0.005$, if percentage between 91 and 100%.

The spreadsheet then automatically populates the possible unsafe outcomes, the conditional probabilities $PA2$, $PA3$, and $PA4$, and—for each unsafe outcome—the injury ratio IR and the product $\sum_{k=1}^{K} CP_k \times IR_k$. In total, there are about 500 possible combinations for $(A1, A2, A3, A4)$. All calculations are deterministic, and no bounds of uncertainty are given on the values.

To evaluate this TARA approach for CP, first consider the notion that CP considers a detected crack as initial condition, rather than an undetected crack. As shown in figure 18, from a mathematical point of view, this is confusing at best. Apparently, TARA assumes that the value for $CP \times IR$ of a detected crack is representative for the value for $CP \times IR$ of an undetected crack. Therefore, this ignores any specific reasons why the crack is undetected. These reasons may have their effect on the probability of the crack leading to an accident. The assumption is never made explicit to the user, so the implications are not accounted for in the risk assessment.

A remark made in section 6.2 was as follows: DA counts the number of cracks of any size if they occur between now and when the aircraft is retired. Aircraft that retire before they develop a crack are correctly not counted. However, it may happen that the aircraft has a crack but retires before the crack has had the chance to grow to dangerous size. In TARA, such cases are not filtered out in DA, so they need to be filtered out in one of the other parameters.

After investigation of each of the components in the CP spreadsheet, the only likely candidate for this is the first conditional probability $PA1$. It is based on the number of cycles required for the damage to progress from the initial condition (the detected crack) to condition $A1$, as percentage of the retirement age t_R. If this percentage is small, the damage is considered to progress quickly, therefore increasing the probability of an accident. If the percentage is high, the damaged crack grows slowly, and the probability of an accident is lower.

Here, several observations can be made:

First, the numerator in the percentage is the time from initial condition to condition $A1$, which is not necessarily an accident-sized crack, and the denominator in the percentage is not the time from initial condition to retirement, but the retirement age (i.e., the total number of cycles an aircraft makes from delivery to retirement). This means that if the initial condition has only a few cycles to go before reaching condition $A1$, TARA estimates the damage to grow quickly and $PA1$ to be high. However, if the aircraft is close to retirement, the damage may still not have time to grow to accident size before the retirement age is reached. Such aircraft will not end up in an accident. However, this situation is not considered here; the time until retirement is not a parameter, and TARA does not filter out these cases.

Second, note that the options in the spreadsheet for condition $A1$ do not specify the size of the crack, so it is unclear how the user can estimate how many cycles it takes to grow the crack from the initial condition to condition $A1$.

Third, note that there is no account of the possibility that the initial condition may be a crack that has grown well beyond condition $A1$ (e.g., toward condition $A2$ or $A3$, which would probably lead to high $PA1 \times PA2 \times PA3$).

Fourth, note that this crack growth rate is nowhere a factor in the determination of ND, whereas a high growth rate would reduce the chances of the crack being detected.

A similar reasoning holds for the age distribution within the fleet. This is not a factor in the TARA approach.

Another issue is independence of the four conditional probabilities: As explained above, the first probability addresses quick or slow growth of the crack up to condition $A1$. The other three address the probability that condition $A1$ ends up in various other conditions, where condition $A4$ models the unsafe outcome. Mathematically (where $Init$ denotes the initial condition):

$$CP \times IR = \sum_{A4} P\left(A4 \middle| Init\right) \times IR\left(A4 \middle| Init\right) = \sum_{A1}\sum_{A2}\sum_{A3}\sum_{A4} P\left(A1 \wedge A2 \wedge A3 \wedge A4 \middle| Init\right) \times IR\left(A4 \middle| Init\right) = \tag{45}$$

$$= \sum_{A1}\sum_{A2}\sum_{A3}\sum_{A4} P\left(A1 \middle| Init\right) \times IR\left(A4 \middle| Init\right) \times P\left(A2 \middle| A1, Init\right) \times P\left(A3 \middle| A2, A1, Init\right) \times P\left(A4 \middle| A3, A2, A1, Init\right) \times IR\left(A4 \middle| Init\right) \tag{46}$$

Analysis of the spreadsheet shows that TARA writes this as:

$$CP \times IR = P\left(A1 \middle| Init\right) \times P\left(A2 \middle| A1, Init\right) \times \sum_{A3}\sum_{A4} P\left(A3 \middle| A2\right) \times P\left(A4 \middle| A3\right) \times IR\left(A4\right) \tag{47}$$

Therefore, TARA assumes that $P(A3 \mid A2, A1, Init) = P(A3 \mid A2)$ and that $P(A4 \mid A3, A2, A1, Init) = P(A4 \mid A3)$: the probability of condition $A3$ given condition $A2$ is independent of condition $A1$ and the initial condition, and the probability of condition $A4$ given condition $A3$ is independent of conditions $A1$ and $A2$ and the initial condition, and $\sum_{A1}\sum_{A2} P(A1|Init) \times P(A2 \mid A1, Init) = P(A1|Init) \times P(A2 \mid A1, Init)$.

This can be explained with an example. Examples of condition $A1$ are:

- Item of mass departing aircraft
- Failed wing/empennage ribs
- Floor beam/intercostal failure

- Stringer failure
- Gear-up landing.

Via various conditions $A2$, these condition $A1$ examples may all result in condition $A3$ reduction of control. Given a reduction of control, condition $A4$ may be a crash, an in-flight break-up, or a runway departure; however, the probability of condition $A4$ given reduction of control is assumed independent of condition $A1$ (i.e., independent of which aircraft part suffers a wear-out failure). This assumption is not always reasonable.

6.5 DISCUSSION OF IR

The IR is the average rate of fatality per person exposed to a specific airplane outcome or condition. The TARAM handbook refers to lists of historical injury ratios for a range of transport-airplane unsafe outcomes that have been compiled and are available on the FAA website.

In TARA, four possible unsafe outcomes are considered, and their IRs are given (i.e., $IR = 1$ for in-flight break-up, $IR = 0.98$ for crash, $IR = 0.03$ for runway departure, and $IR = 0.001$ for individual fatality). These numbers are based on historical data for transport airplane accidents and were developed in conjunction with the FAA's Transport Airplane Directorate staff. It is noted that in the TARA spreadsheet, these values are independent of the location of the crack, the age of the aircraft and fleet, or of the sequence of events leading up to the accident.

6.6 UNCERTAINTY INTRODUCED

In TARAM, the main uncertainty comes from the DA calculations and from the assessment of ND. In TARA, there is additional uncertainty in DA, ND, and CP. This section identifies the main sources.

Uncertainties in TARAM:

- Choice of probability distribution for the occurrence of wear-out failures. Typically the 2-parameter or 3-parameter Weibull distribution is used, but other distributions are available. Any choice will have an impact on the outcome.
- Assessment of shape parameter β. In case a Weibayes analysis is used, the value is assumed and depends on the material in which the crack grows. However, usually the shape parameter for crack initiation is different from the shape parameter for extensive growth, which creates some level of uncertainty in the result. More accuracy is obtained by using the 3-parameter Weibull rather than the 2-parameter version because, in the 3-parameter Weibull, the shape parameter is for the initiation and early growth phase, whereas the third parameter (the location parameter) accounts for the extensive growth phase.
- Assessment of the time it takes to grow the detected cracks to cracks of dangerous size. This is a source of uncertainty in the calculations for the scale parameter η. However, it is noted that the outcome appears not very sensitive to changes in this time to grow.
- Assessment of DA (i.e., number of aircraft with dangerous-sized cracks). There is a minor issue of confusion concerning the group of uninspected aircraft. Because TARAM anchors to dangerous-sized cracks, the uninspected aircraft are considered suspensions because

dangerous-sized cracks are considered obvious, even in uninspected aircraft. Nevertheless, ND is greater than zero; therefore, some of these obvious cracks are not detected anyway.

- Assessment of ND, nondetection probability. Typically, the value for ND is very small, but a difference of $ND = 10^{-3}$ or $ND = 10^{-4}$, both very small numbers, would still create a difference of a factor 10 in the fleet risk calculations, so there is a potential source of uncertainty.
- Assessment of CP, conditional probability. Typically, the value for CP is $CP = 1$. In some cases, the dangerous crack is in a location that does not preclude safe landing of the aircraft and CP will be smaller. However, overall, the level of uncertainty in CP will be less than, for instance, a factor 2.
- Assessment of IR, injury ratio. There may be variation in the number of fatal injuries, depending on the extent of the unsafe outcome.

Uncertainties in TARA:

- Choice of probability distribution for the occurrence of wear-out failures. The level of uncertainty is the same as for TARAM.
- Assessment of shape parameter β. The level of uncertainty is the same as for TARAM.
- Assessment of DA, number of aircraft with cracks. There are multiple issues:

    a. TARA includes failures in the sum for DA, which should not be included. This is a mathematical error, but, as is illustrated in section 5.2.2, the error made in the value for DA is probably small.
    b. TARA includes uninspected aircraft in the sum for DA that should only be included if they are free of failures, which cannot be guaranteed for nonobvious cracks.
    c. There may be aircraft that grow a crack but retire before the crack reaches dangerous size. These aircraft are included in the count for DA, and they are not filtered out in ND or CP.

- Assessment of ND, nondetection probability. There are multiple issues:

    a. The definition of ND suggests that credit can be taken for detection at any time until the crack reaches dangerous size, but TARA can take credit only for detection until the size as considered by DA.
    b. Because the crack can be anything from detectable to obvious, there will be variation in the time period during which detection can take place: the window of opportunity varies.
    c. The value for ND is based on a small set of yes/no questions that imply assumptions regarding the circumstances of the cracks involved. These assumptions create uncertainty.
    d. ND does not take account of the size of the cracks, the time until retirement of the aircraft, the number of aircraft in the fleet, the age distribution of the fleet, the number of aircraft with cracks, the size of the cracks, the number and type of inspections, the probability that a given aircraft in the fleet is inspected, cracking history, the growth rate of the cracks, the various ways in which the damage can be found, or the number of maintenance cycles during the period of crack growth.

e. There is no evidence that the flow charts address nondetection "before resulting in an unsafe outcome or condition" or nondetection "throughout the life of the affected fleet."

- Assessment of CP, conditional probability. There are multiple issues:

a. The initial condition is a detected crack rather than an undetected crack. Therefore, this ignores any specific reasons why the crack is undetected.
b. The four conditional probabilities in the sequence $CP = PA1 \times PA2 \times PA3 \times PA4$ are assumed independent, although this does not need to be the case.
c. CP does consider the rate of the crack to grow, but it does so only for the time of the initial condition (the detected crack) until condition $A1$. It does not account for the time of the crack to grow until dangerous size. It also does not account for cases in which the initial condition is beyond condition $A1$. In addition, the size of the crack or the number of cycles considered for condition $A1$ is not specified, so it is unclear how the user can estimate the number of cycles it takes for the crack to grow from the initial condition to condition $A1$.
d. CP does not filter out those aircraft that are retired before the crack has the chance to grow to dangerous size.
e. The value for CP is deterministic, given the choice for conditions $A1$ and $A2$ and the estimated time from the initial condition to condition $A1$.

- Assessment of IR, injury ratio. This is assumed independent of the location of the crack, or of conditions $A1$, $A2$, and $A3$.

From this reasoning, it may be concluded that both TARAM and TARA introduce uncertainty in the determination of uncorrected fleet risk. However, given the number of additional assumptions that need to be adopted because of the notion that cracks of any size are considered rather than cracks of accident size, and because of the notion that ND and CP are based on a fixed set of conditions that are incomplete and do not completely respect the mathematical background of the conditional probabilities underlying the model, it is expected that the level of uncertainty in TARA will be significantly larger than that in TARAM.

6.7 CONSISTENCY AND REPEATABILITY OF CALCULATIONS

According to [3], the TARA approach and tools were developed with the goal to provide a framework for the calculations and to make the risk assessment more consistent from one fleet of airplanes to another and more repeatable from one analyst to another. Therefore, it makes sense to provide a comparison in these terms as well.

TARAM comes with a handbook that is comprehensive. However, it provides factors to consider rather than practical guidance on what to do exactly in each step. As such, the user will need a thorough understanding of the material and additional guidance beyond the handbook. This may limit the level of consistency and repeatability of calculations to an extent. TARA takes the user by the hand and provides easy to use flow charts and spreadsheets to assist in the calculations, which should help consistency and repeatability of calculations. This effect will be much less if practitioners, when interpreting the results, take account of the levels of uncertainty introduced by the approach.

7. CONCLUDING REMARKS

7.1 CONCLUSIONS AND RECOMMENDATIONS

This report compares three approaches used by the FAA to estimate the uncorrected fleet risk R_T, which are TARAM (using 2-parameter Weibull), a modification of the TARAM approach (using 3-parameter Weibull), and TARA. TARAM uses crack data to fit a Weibull distribution that reflects the unsafe outcome due to wear-out failures. The number of cycles flown by each aircraft is adjusted to the time of a dangerous crack size. In TARA, the data are not adjusted, and the detected damage condition is used as a starting point. Next, flow charts and spreadsheets are used to determine the sequence of conditional probabilities towards the unsafe outcome.

The main conclusions are:

- Both TARAM and TARA are approaches to determine uncorrected fleet risk due to wear-out failures.
- Compared to the 2-parameter Weibull, the 3-parameter Weibull improves accuracy, though it is slightly conservative if there are no undiscovered cracks.
- The TARAM handbook contains a lot of information, though the guidance is not always very specific; an important part of the guidance is hidden in examples in the appendices. TARA helps the user by providing easy-to-use flow charts and spreadsheets, aiming at improved consistency and repeatability.
- The outcome of the TARAM approach will include a certain level of uncertainty, but the uncertainty introduced in the TARA approach is estimated to be significantly larger. The main reason is that TARA adopts several major assumptions and simplifications, due to:

    - Accounting for crack growth in conditional probabilities rather than in the Weibull part of the analysis.
    - Using for these conditional probabilities a set of flow charts and spreadsheets that are deterministic and that cover a limited and incomplete set of event sequences.
    - Making several mathematical errors.

The main recommendations are:

- For TARAM: The TARAM handbook could be improved. Important guidance currently hidden in examples and appendices needs to be moved to the main part of the document.
- For TARA: The approach needs to be updated to repair the mathematical errors. It is recommended that the approach anchors cracks to dangerous size. All assumptions adopted need to be made explicit to allow the user to assess the level of bias and uncertainty in the result.
- In the meantime, use TARAM with 3-parameter Weibull.

7.2 OPTIONS FOR FURTHER WORK

A list of ideas and considerations for further study has been identified as follows (in arbitrary order).

Exploring alternative Weibull approaches. This option could include one or more of the following:

- One could opt to not use a Weibayes analysis and an assumed value for β, and instead use a Weibull analysis, by using detected failure data generated by the Monte Carlo simulation to determine an estimate for γ and β and techniques such as those described in section A.4 of appendix A. Note that in practice, this approach is rarely taken because of a lack of accident data.
- As explained in appendix A.4, the Weibayes method also works if there are no failure data. However, in that case, in the formula $\eta = \left[\frac{1}{r_a}\sum_{i=1}^{N}(t_i)^{\beta}\right]^{1/\beta}$, the number of detected failures cannot be taken equal to $r_a = 0$ because this leads to division by zero. A conservative engineering assumption is to take r_a equal to one (i.e., the first failure is imminent), but in the literature, e.g., [8], other values for r_a have been proposed that lead to less conservative estimates for η (see appendix A.4). It may be worth investigating whether these studies could help improve the Weibayes method used, in case of small r_a.
- Some alternative approaches may be investigated. These may include alternatives to the Weibull distribution, such as Lognormal distribution, Crow AMSAA, or Supersmith Weibull; see, for example, reference [7].

Sensitivity analysis. This option could include one or more of the following:

- At the start of the Monte Carlo simulations (during the three-staged setup), moderate cracks (inputs) were generated at aircraft ages t_i^{mod} sampled from a Weibull distribution $F(t;\eta_{mod},\beta) = 1 - \exp\left[-\left(\frac{t}{\eta_{mod}}\right)^{\beta}\right]$. The shape parameter was fixed at $\beta = 4$ (for aluminum). At the end of the simulation, the outputs were fitted using a Weibayes with assumed shape parameter $\beta = 4$. It would be interesting to determine to what extent the conclusions are dependent on the correlation between input and output distribution and shape parameter.
- Another option is to generate the Monte Carlo simulation failure data from a 3-parameter Weibull. It will be interesting to see whether Case 6 will do better then.
- Investigate why Case 6 does not score particularly well (compared to the other cases) if there are few undiscovered cracks and $P_{insp} = 0.003$.
- There are numerous other options for sensitivity analysis, in which Monte Carlo simulations are run, such as with other values for β, other values for crack-growth rates (e.g., including a crack growth model; also note that typically $t_{DM} \gg t_{MO}$), other fleet age distributions, and more.

Improvements to TARAM handbook:

- This option would include identifying improvements to the TARAM handbook aimed at more practical guidance for the user.

- As part of this, guidance could be developed for determining the location parameter in the 3-parameter Weibull analysis.

Implement a TARA-like approach to determine the location parameter:

- For the TARAM approach it was observed that the prediction of DA is relatively insensitive to the value used for the time to grow from a moderate-sized to a dangerous-sized crack (which is also the shape parameter in a 3-parameter Weibull). TARA uses an expert-system-like question/answer process to determine parameters to calculate CP. This same approach could be used to determine the location parameter in the TARAM approach. Errors in CP are directly proportional to risk, whereas risk is relatively insensitive to errors in the location parameter. This appears to be a win-win: a simple, repeatable, standardized process that the TARA authors sought while maintaining the accuracy and statistical/mathematical integrity sought by the TARAM authors.

8. REFERENCES

1. FAA Order 8110.107A, Monitor Safety / Analyze Data (2012). https://www.faa.gov/documentLibrary/media/Order/Order%208110.107A.pdf

2. FAA. (2011). *Transport Airplane Risk Assessment Methodology (TARAM) Handbook,* FAA Transport Airplane Directorate ANM/100.

3. Violette, M.G., Safarian, P., Han, N., Atmur, R. (2015). Transport Airplane Risk Analysis. *Journal of Aircraft*, *52*(2), 395–402.

4. AFML-TR-69-65: Air Force Materials Lab. A reliability analysis approach to fatigue life variability of aircraft structures.

5. AFML-TR-72-236: Air Force Materials Lab. Development of titanium and steel fatigue variability model for application of reliability analysis approach to aircraft structures.

6. Craycraft, J., Monte Carlo study on fatigue cracking safety, Draft paper, August 2015

7. Abernethy, R. B. (2010). *The New Weibull Handbook*, (5^{th} ed.). North Palm Beach, FL: Dr. Robert. Abernethy.

8. Meeker, W. Q., Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. Wiley Series in Probability and Statistics,

APPENDIX A—THE WEIBULL DISTRIBUTION

The Weibull distribution is one of the most widely used lifetime distributions in reliability engineering. There are three versions: 1-parameter, 2-parameter, and 3-parameter.

A.1 2-PARAMETER WEIBULL

Best known is the 2-parameter Weibull distribution, which has a probability density function (pdf) of:

$$f(t;\eta,\beta) = \begin{cases} \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{A-1}$$

Failure data plotted as a bar chart against time will have the same shape as the pdf.

The cumulative distribution function (cdf) or probability of failure up to time t is:

$$F(t;\eta,\beta) = \begin{cases} 1-\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{A-2}$$

The complement of the cdf, called reliability, is the probability that failure will not occur up to time t:

$$R(t;\eta,\beta) = \begin{cases} \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{A-3}$$

The failure rate h (sometimes referred to as hazard function) or the frequency of occurrence of a failure per unit of time is $f(t;\eta,\beta)/R(t;\eta,\beta) =$

$$h(t;\eta,\beta) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \tag{A-4}$$

β is the shape (or slope) parameter, $\beta > 0$; it determines the shape of the pdf. For $\beta \leq 1$, the pdf is decreasing; for larger β, the pdf looks like a skewed bell-shaped curve. When looking at the term $\beta - 1$ in the exponent in the failure rate function h, it can be easily derived that:

- $\beta < 1$ indicates an early failure condition: the failure rate decreases with time.
- $\beta = 1$ indicates random failure distribution: constant failure rate.
- $\beta > 1$ indicates wear-out failure distribution: the failure rate increases with time.

At $\beta = 1$, the Weibull becomes an exponential distribution. A high shape parameter β is an indication of failures that are predictable; all components fail around a predictable age.

If the cdf is plotted against time using “Q-Q plot,” it shows as a straight line. The slope of the line is equal to β, which explains that β is sometimes referred to as the slope. In this Q-Q plot, the axes are in logarithmic scales of $\ln(-\ln(1 - F(t)))$ versus $\ln(t)$, using that $F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$;

therefore, $\ln\left(-\ln\left(1-F(t)\right)\right) = \beta\ln(t) - \beta\ln(\eta)$. Figure A.1 provides an example of a Q-Q plot.

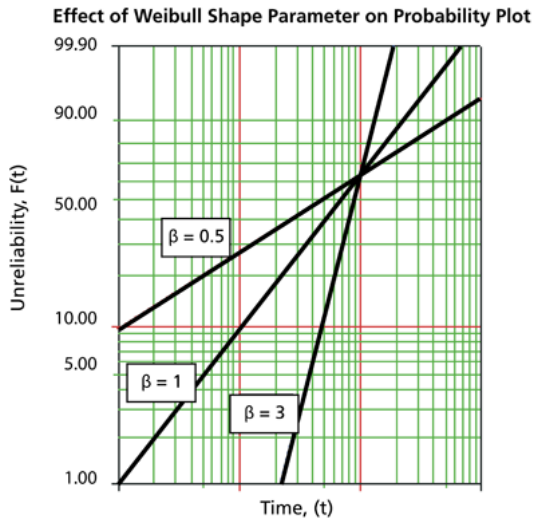


Figure A-1. Example Q-Q plot, taken from http://reliawiki.org/index.php/The_Weibull_Distribution

η is the scale parameter, $\eta > 0$; it scales the time parameter (this can be observed from the term t/η in the cdf). For $\eta > 1$, the cdf increases more slowly, and the pdf is stretched out; for $\eta < 1$, the cdf increases more quickly, and the pdf shape becomes narrower. At time $t = \eta$, approximately 63% of the population is expected to have the failure, independent of the value for β. This can be observed from the cdf at $t = \eta$:

$$F(\eta;\eta,\beta) = 1 - \exp\left[-\left(\frac{\eta}{\eta}\right)^{\beta}\right] = 1 - e^{-1} = 0.6321 \qquad \text{(A-5)}$$

For this reason, η is also called characteristic life.

At $\beta = 2$ and $\eta = \sqrt{2}\beta$, the Weibull becomes a Rayleigh distribution.

If $\beta = 1$, η is equal to the mean-time-to-failure (mttf). If $\beta < 1$, $\eta <$ mttf; if $\beta > 1$, $\eta >$ mttf. Generally, mttf $= \eta\,\Gamma(1+\frac{1}{\beta})$, where Γ is the gamma function $\Gamma(n) = \int_0^{\infty} e^{-x}x^{n-1}dx$. The mean and variance of the Weibull are: $E(T) = \frac{\eta}{\beta}\,\Gamma\left(\frac{1}{\beta}\right) = \eta\,\Gamma\left(1+\frac{1}{\beta}\right)$ and $\mathrm{Var}(T) = \eta^2\left[2\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma(1+\frac{1}{\beta})^2\right]$.

If U has the standard uniform distribution, then $T = \eta[-\ln(U)]^{1/\beta}$ has the Weibull distribution.

A.2 1-PARAMETER WEIBULL

The 1-parameter Weibull distribution is equal to the 2-parameter Weibull, but with the β parameter assumed a known constant.

A.3 3-PARAMETER WEIBULL

The 3-parameter Weibull distribution has a pdf of:

$$f(t;\eta,\beta,\gamma) = \begin{cases} \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] & t \geq \gamma \\ 0 & t < \gamma \end{cases} \quad \text{(A-6)}$$

The cdf is:

$$F(t;\eta,\beta,\gamma) = \begin{cases} 1-\exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right] & t \geq \gamma \\ 0 & t < \gamma \end{cases} \quad \text{(A-7)}$$

γ is the location parameter. It has the effect of sliding the pdf to the right ($\gamma > 0$) or to the left ($\gamma < 0$) along the time axis. For $t \in [0, \gamma]$, there are no failures, which is why γ is also referred to as failure free life. A negative γ may indicate that failures have occurred (e.g., prior to actual use. The mttf $= \gamma + \eta\,\Gamma(1 + \frac{1}{\beta})$). For $\gamma = 0$, the 3-parameter Weibull reduces to a 2-parameter Weibull. If $\gamma \neq 0$, the characteristic life equals $\eta + \gamma$: 63.2% of all failures fall below the characteristic life regardless of the value of the shape parameter. The mean and variance of the 3-parameter Weibull are: $E(T) = \gamma + \frac{\eta}{\beta}\,\Gamma(\frac{1}{\beta})$ and $\mathrm{Var}(T) = \frac{\eta^2}{\beta}\left[2\Gamma\left(\frac{2}{\beta}\right) - \frac{1}{\beta}\Gamma(\frac{1}{\beta})^2\right]$.

A.4 ESTIMATING WEIBULL PARAMETERS

There are several methods to estimate the parameters of a Weibull distribution, given a set of failure data. These include graphical methods via probability plotting paper (i.e., Q-Q plots) or analytical methods using either least squares (rank regression) or maximum likelihood estimation, [reference 7 in the main document], [A-1]. Assume $t_1, t_2, \ldots$ denote failure data in a particular chosen time unit (hours, flights, cycles). Sometimes, the list $t_1, t_2, \ldots$ also includes suspension times, which are times at which a component has been inspected and is considered failure free.

A.4.1 MAXIMUM LIKELIHOOD ESTIMATION

For 2-parameter Weibull, the maximum likelihood estimator for β is as shown in Appendix C.4 in reference [7] of the main document.

$$\hat{\beta}^{-1} = \frac{\sum_{i=1}^{n} t_i^{\hat{\beta}} \ln t_i}{\sum_{i=1}^{n} t_i^{\hat{\beta}}} - \frac{1}{r}\sum_{i=1}^{r} \ln t_i \quad \text{(A-8)}$$

Where r is the number of detected failures and n is the number of failures plus suspensions. This needs to be solved using iterative procedures. Appendix C.5 in reference [7] of the main document provides an example in which subsequent estimates for $\hat{\beta}$ are inserted in the equation above, until equality is obtained.

The maximum likelihood estimator of η given $\hat{\beta}$ is:

$$\hat{\eta} = \left[\frac{1}{r} \sum_{i=1}^{N} (t_i)^{\hat{\beta}} \right]^{1/\hat{\beta}} \quad \text{(A-9)}$$

Reference [7] in the main document explains that maximum likelihood estimators for the 3-parameter Weibull also exist, but are not presented in the book as being too complex.

A.4.2 WEIBAYES METHOD

The Weibayes method makes use of the maximum likelihood estimation. It is reported to be particularly useful in cases of few failure data[5]. In this method, β is not estimated but is assumed known (table A-1); therefore, this corresponds to a 1-parameter Weibull.

Table A-1. Shape parameters in Weibayes method (table 1 in reference [3] of the main document), which refers to references [4, 5] in the main document for the first four values and to experience for the last value

Material	Shape parameter β
Aluminum	4
Titanium	3
Low-strength steel ($F_{tu} \leq 240$ ksi)	3
High-strength steel ($F_{tu} > 240$ ksi)	2.2
Stress corrosion cracking	2

Next, η is calculated analytically (maximum likelihood estimate given β) (see Chapter 6 and Appendix E in reference [7] of the main document):

$$\eta = \left[\frac{1}{r_a} \sum_{i=1}^{N} (t_i)^{\beta} \right]^{1/\beta} \quad \text{(A-10)}$$

$N =$ total number of failures plus suspensions.

$r_a =$ number of aircraft with detected failure.

$t_i =$ number of cycles flown by aircraft i at time of failure or suspension.

[5] Some references (e.g., reference [7] in the main document) recommend using Weibayes in case of fewer than 20 failure data. In practice, Weibull is still used in case of few failure data.

For a 3-parameter Weibull, the same expression is used, with t_i replaced by $t_i - \gamma$. If $t_i - \gamma < 0$ then use $t_i - \gamma = 0$.

If no failures have occurred, a conservative engineering assumption is to take r_a equal to one (i.e., the first failure is imminent). In the literature, this is referred to as Weibayes method. However, assuming that the first failure is imminent is often very conservative. A less conservative approach is to select the 50% lower confidence bound with respect to the true Weibull failure distribution, the so-called Weibest method. Reference [8] in the main document explains that if no failures have occurred, a conservative $100(1-\alpha)\%$ lower confidence bound for η is $\eta = \left[\frac{1}{-\ln(\alpha)}\sum_{i=1}^{N}(t_i)^{\beta}\right]^{1/\beta}$. For $100(1-\alpha)\% = 95\%$, $\alpha = 0.05$ and $-\ln(\alpha) \approx 3$. Taking $r_a = 1$ corresponds to $100(1-\alpha)\% = 63\%$ (see [A-2], section 12.8).

A.4.3 PROBABILITY PLOTTING

In this method, failure data are used to plot the cdf, from which the parameters β and η are estimated. The first step is to take the times to failure (i.e., suspensions are not included) and to rank them in ascending order $t_1, t_2, \ldots$. Next, their median rank plotting positions are obtained. This can be done in various ways; the easiest one is by $\hat{F}_i \approx \frac{i-0.3}{n+0.4} \cdot 100$, where i denotes the index of the times to failure and n is the total sample size. Next, using Q-Q plot (i.e., special logarithmic scale), plot the $(t_i, \hat{F}_i)$. Draw the best possible straight line through the data points. The slope of this line determines β. The point on the time axis where the line equals 63.2% on the cdf axis determines η. Special Weibull plotting paper exists that assists in estimating the slope of the line. In figure A-1, all three lines correspond to the same estimate for η.

The failure data may better fit a 3-parameter Weibull if the failure data $(t_i, \hat{F}_i)$ on the Q-Q plot are not on a straight line.

- If the curve for $(t_i, \hat{F}_i)$ is concave down (∩) and the curve for $(t_i - t_1, \hat{F}_i)$ is concave up (∪), then γ has a positive value in $[0, t_1]$.
- If the curves for $(t_i, \hat{F}_i)$ and for $(t_i - t_1, \hat{F}_i)$ are both concave up (∪), then γ has a negative value.
- If neither case prevails, reject the Weibull or use a mixed version.

Parameter γ is constructed by trial and error: pick a value for γ (positive if $(t_i, \hat{F}_i)$ is concave down and negative if $(t_i, \hat{F}_i)$ is concave up). Next, plot $(t_i - \gamma, \hat{F}_i)$. Repeat until the line is acceptably straight. After this, determine β and η as described above, but by using the line $(t_i - \gamma, \hat{F}_i)$.

A.4.4 RANK REGRESSION

This method is the same as the one for probability plotting, except that the straight line through the data points is not drawn by looking at it, but by using mathematical fits. Least squares estimators are:

$$\hat{\beta} = \left[\sum_{i=1}^{N} x_i y_i - \frac{1}{N} \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i \right] \Bigg/ \left[\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} x_i \right)^2 \right] \tag{A-11}$$

and

$$\hat{\eta} = \exp\left(\frac{\hat{\beta}}{N} \sum_{i=1}^{N} x_i - \frac{1}{N} \sum_{i=1}^{N} y_i \right) \tag{A-12}$$

where $y_i = \ln(-\ln(1 - F(t_i)))$ and $x_i = \ln(t_i)$.

Other fitting approaches include maximum likelihood estimators, but these are reported to provide biased slope parameter estimations unless the number of observations is very large.

A.5 ESTIMATES FOR NUMBER OF FAILURES

Reference [7] in the main document presents formulas for the predicted number of failures, given a certain dataset. Consider a set of N components, r of which are shown to have failed, and N_S are suspensions. The ages of the components at the time of failure or suspension, are $t_1, t_2, ...$

A.5.1 CURRENT RISK

Section 4.4.1 in reference [7] of the main document presents a formula for the number of failures at the present moment. In earlier editions, this was taken to be the probability of failure by time t_i, summed over the number of units, including failures and suspensions (i.e., NowRisk $= \sum_{i=1}^{N} F(t_i)$). Later research showed that this estimate was biased. A better estimate was shown to be:

Now Risk =

$$\sum_{Failures} 2 \cdot F(t_i) + \sum_{Suspensions} F(t_i) \tag{A-13}$$

where F is the cdf of the Weibull distribution.

A.5.2 FUTURE RISK

Section 4.4.2 in reference [7] of the main document presents a formula for the number of failures at a future moment, for instance Δ units of time from now, assuming that failed units are not replaced:

Future Risk = Sum over P(Failure in [Now, Now $+\Delta$] | No failure until Now).

For already failed components, this conditional probability is zero. For components that have not yet failed (i.e., suspensions), this probability is equal to:

Sum over P(Failure in [Now, Now $+\Delta$] and No failure until Now)$/P$(No failure until Now) $=$

Sum over $(P(\text{Failure before Now } + \Delta) - P(\text{Failure before Now}))/(1 - P(\text{Failure before Now})) =$

$$\sum_{Suspenses} \frac{F(t_i^{age} + \Delta) - F(t_i^{age})}{1 - F(t_i^{age})} \quad \text{(A-14)}$$

A.6 REFERENCES

A-1. ReliaWiki, The Weibull Distribution, (Accessed November 19, 2015), Retrieved from http://reliawiki.org/index.php/The_Weibull_Distribution

A-2. IEC 61649, Ed. 2.0: Weibull Analysis, (2008-08-13). Retrieved from http://lyle.smu.edu/~jerrells/courses/emis7370sum08/56_1269e_FDIS.pdf