Maintenance, Engineering, and Operational Decision-Making Metrics Derived from Simple Maintenance and Aircraft Datasets

Paul Bordush[1,2]

[1] *Department of National Defence, Government of Canada*

[2] *The Paul Aerospace Company Ltd., Ottawa, Ontario, Canada*
paul@paulaerospace.ca

Abstract

Using simple in-service maintenance data, it is possible to predict and forecast the propulsion system contribution to aircraft and fleet level unavailability and identify sub-system degraders of overall engine reliability. While more complex means of assessing reliability exist, increased layers of complexity can lead to increasing difficulty when used to convince military commanders or higher management of the appropriate action to take. Furthermore, increased complexity increases the time required to produce, analyze, and assess the results of reliability assessments. In a time-critical situation, when faced with the need for an immediate maintenance or engineering decision, the best information is that which is the simplest and easiest to understand, quickest to produce, and fastest to apply. In this work, a minimum list of data requirements will be developed with an associated means of analyzing this data to produce meaningful indicators to predict and forecast unavailability and mission abort rates that can be used to plan for deployed or sustained operations. Further analysis can produce a prioritized listing of sub-system reliability degraders to drive engineering decisions for component improvement. The Royal Canadian Air Force's CT114 Tutor aircraft will be the basis for analysis demonstrating that sophisticated sensors and data systems are not required to produce meaningful metrics suitable for significant fleet level decisions. Statistical methods and appropriate data filtering were applied to the engine system to derive rates for overall mission aborts, aircraft unavailability and aircraft unreliability for the top sub-system degraders. Conclusions drawn include that this information, if calculated correctly, provides decision makers with the critical information required to make significant fleet level decisions. Recommendations and methodology are presented that should be applicable to any military or civil fleet at the sub-system, aircraft, and fleet level.

1. Introduction

1.1. Motivation

The motivation for this work is to provide a standardized method to conduct a reliability and availability assessment for fleets of equipment using datasets limited to simple maintenance data, which is likely to contain inconsistencies and errors. Large organizations, the Royal Canadian Air Force (RCAF) for example, operating multiple fleets of equipment of varying vintage tend to establish separate metrics and unique terminology which create unnecessary barriers to continuous improvement. Establishing simple standardized measures based on the available datasets can provide common terminology stimulating innovation and collaboration and leading to significant gains in cross-organizational capability. Effectively, the intent is to create a common framework, and then diligently promote it to establish a culture of continuous improvement. While these steps are an essential part of the desired process, it will not be further expanded upon as it is more closely related to organizational behaviour and management. This paper aims to demonstrate that from an existing maintenance dataset, that a standardized framework is possible and can be used to provide measures that have intrinsic value to engineers, operators, and managers within large organizations driving improved collaboration and decision making capability amongst these diverse groups. The ultimate goal is therefore finding the optimum balance between simplicity, accuracy, and informative value based on the availability and accuracy of data.

The issue with many traditional reliability assessments is that they are often too high level, thereby providing limited ability to identify insights into the required actions to improve system and fleet reliability. At other times, these high level assessments incorporate numerous assumptions, sometimes with differences in definitions of terminology, that render them potentially prone to misinterpretation by a non-reliability engineer or fleet manager or operator. This is not to say that high level fleet and system reliability assessments are flawed, because they do provide an overall indication of system and fleet health. The work presented here will attempt to bridge the gap between high level reliability analyses and make them usable and understandable by anyone involved in fleet level maintenance, operations, or in-service engineering. Furthermore, the means by which simple maintenance data sets can be used to generate meaningful system reliability data will be explored to assist the reader with the development of their own analysis capability.

1.2. Regulatory Requirements for Maintenance Data

Maintenance record keeping requirements ensure the airworthiness of aeronautical products. The Canadian Aviation Regulations, Standard 571 (TC, 2025a), states that each Designated Aeronautical Design Authority is required to establish and maintain a record keeping system for all the maintenance performed on any aircraft component. This record keeping system must include records of daily maintenance performed including but not limited to date, employee who performed the work along with records of inspection, component failures, rectification and other essential information. Similarly, the Federal Aviation Administration outlines requirements in the United States under FAR Chapter I Subchapter C Part 43 (FAA, 2025b) and in Europe under Regulation (EU) No 2018/1139 (EASA, 2018).

1.3. On-board Sensor Suites

Older fleets which do not have on-board data acquisition systems will require extensive effort and cost to retrofit them with a monitoring system. Depending on the expected in-service timeframe, it may not be cost effective to implement an on-board engine usage monitoring system after entry to service (NATO, 2000; Wang et al., 2023). However, as mandated by regulating bodies, maintenance data recording requirements provide a data set sufficient for predicting reliability without the need for on-board sensor suites and at a fraction of the implementation cost.

1.4. Demonstration Data Set

The J85-CAN-40 turbojet engine of Royal Canadian Air Force's fleet of CT114 Tutor aircraft will be used as a demonstration of the proposed approach, selected because of its vintage from the 1960's with no sophisticated on-board sensor suite to show that sophisticated and costly sensor systems are not always required for fleet management.

The methodology will be presented followed by its application to the CT114 Tutor maintenance dataset. Applications of this work in civil aviation will also be touched on and avenues for future work discussed.

2. METHODOLOGY AND MEASURES OF FLEET RELIABILITY

The most cited measure of fleet reliability is availability (A). Availability can be further redefined as operational availability (A_o), which may account for uptime and downtime in slightly differing ways based on available information. In its simplest form, availability is defined as the ratio of up time to total time which is equivalent to one minus the unavailability (U) according to Eq. (1) (Andresen & Williams, 2006; Hurst, 2006; Elsayed, 2021). Up time here refers to the period of time that the equipment is operational and performing its intended function, and down time refers to the period of time that the equipment is not operational or in a state of repair or maintenance.

$$A = \frac{\text{Up Time}}{\text{Total Time}} = \frac{\text{Total Time} - \text{Down Time}}{\text{Total Time}} = 1 - U \tag{1}$$

Because of the differing definitions of availability and the underlying measures of uptime and downtime, the term availability is inherently problematic and misleading without clearly defining its meaning. The author would argue that all definitions of availability are valid, but that all definitions of availability must be defined by different personnel for their specific requirement and availability of data. For this reason, the term availability will not be used further in this paper, and instead the terms reliability and unreliability will be used.

Based on the author's experience as a maintenance officer for a fighter squadron, an in-service support engineer, and a propulsion systems program manager, and derived from numerous maintenance and reliability handbooks, the most pertinent and useful measures defining fleet reliability have been:

1. Measures of Operational Effectiveness
   (a) Pre-Flight Abort Rate
   (b) Mission Abort Rate
2. Measure of Safety and Airworthiness
   (a) Accident and Hazard Occurrence Rate
3. Measures of Maintenance Burden and Sustainability
   (a) Post-Flight Unserviceability Rate
   (b) Component Removal Rate
   (c) Flight Line Maintenance Occurrence Rate
   (d) Maintenance Hours per Flight Hour

In most handbooks and papers, the prominent means of quantifying a component's reliability is its Mean Time Between Failure (MTBF - average time between failure events), and not its failure rate (average number of failures per unit time). MTBF is simply the inverse of failure rate, therefore the two quantities effectively represent the same thing. The author makes no claim that one is better than the other as they are equivalent, however the author has a clear preference for rates in conveying results to fleet managers and that the scatter in mean values of the rate is also pertinent in understanding failure processes.

All of these measures can and should be generated at a fleet level, as well as at the sub-component level to provide a means to effectively drill down to the root cause factors. In this way, the same measures provide operational and in-service managers with the information needed to perform their individual duties and with a common language for discourse. Having a common language between managers, engineers, operators, and technicians is a force multiplier in that less time is spent discussing the definition of a measure allowing the focus to be on the issues themselves and issue resolution. This point cannot be emphasized enough, to be truly effective everyone in the organization needs to speak the same language and agree on the underlying data.

We will define several general parameters which will then be expanded on in more detail and applied to the measures listed above. The first is failure rate (F_r), according to Eq. (2), for which the number of events, which could be failures or some other undesired occurrence, divided by flying hours over some specified averaging interval, the selection of which will be expanded on further in section 6.3. It is important to highlight that failure in this context is the undesired event, which does not have to refer only to a catastrophic failure of a component or aircraft. What constitutes a failure will change depending on the measure being derived which we will see later in the paper.

$$F_r = \frac{\Sigma \text{Failure Events}}{\Sigma \text{Flight Hours}} \tag{2}$$

The failure rate can be converted into a per sortie or per flight failure rate (F_s) based on average flight time (T_{avg}) according to Eq. (3).

$$F_s = F_r \, T_{avg} \tag{3}$$

Because a flight or sortie can either be successful or unsuccessful (failure), we can represent the per sortie failure rate as a failure probability (F) bound between 0 and 1:

$$F = F_s = F_r \, T_{avg} \text{ such that } 0 \leq F \leq 1 \tag{4}$$

In this way, we are then able to represent the probability of a successful flight or sortie (R) as:

$$R = 1 - F \tag{5}$$

The discrete number of expected events (N_F) can be calculated over a specified period of time, T, or a specified number of flights or sorties, N_s, according to Eqs. (6) or (7) respectively.

$$N_F = F_r \, T \tag{6}$$

$$N_F = F_s \, N_s \tag{7}$$

3. MEASURES OF OPERATIONAL EFFECTIVENESS

3.1. Pre-Flight Abort Rate

Perhaps of more interest in the military context, where at a specific point in time a specific number of aircraft are to be launched for a given mission, the pre-flight abort rate, $F_{r,pre}$, can be used as an indicator of how many total aircraft (primary and backup) are required to meet the mission requirement. The pre-flight abort rate is calculated using Eq. (2) over any given time period where the number of events is the number of pre-flight abort events.

Pre-flight abort rates can be used directly to estimate the number of pre-flight aborts over a specified number of fleet flying hours, however this is less useful to the maintenance organization responsible for providing the right number of aircraft to ensure mission success for a given launch. In this case, with the average flight time we can determine the average per sortie pre-flight abort rate, $F_{s,pre}$, with Eq. (3). The required number of mission ready aircraft for a launch consisting of a minimum number of required aircraft for the mission, $N_{ac,req}$, which is the sum of the number of aircraft required plus the number likely to fail pre-flight for which a backup aircraft is required, can then easily be found with Eq. (8).

$$\begin{aligned} N_{ac,req} &= N_s + N_{F,pre} \\ &= N_s + F_{s,pre} N_s = N_s(1 + F_{r,pre} \, T_{avg}) \end{aligned} \tag{8}$$

We can also consider pre-flight aborts in terms of pre-flight reliability if we use the pre-flight abort rate per sortie, first determining the unreliability and bounding it between 0 and 1:

$$F_{pre} = F_{s,pre} = F_{r,pre} T_{avg} \text{ such that } 0 \leq F_{pre} \leq 1 \tag{9}$$

From which the pre-flight reliability becomes:

$$R_{pre} = 1 - F_{pre} \tag{10}$$

3.2. Mission Abort Rate

Mission abort rates are a more general representation of operational effectiveness since they include pre-flight aborts as well as in-flight aborts. The mission abort rate cannot be used in the same way as pre-flight abort rates to determine a number of required aircraft because once the aircraft takes off, a component failure resulting in a mission abort does not afford the opportunity to switch to a backup aircraft as is possible before taxi and take off.

A better way to think of mission abort rates is how effective is the aircraft and its sub-systems in being able to achieve the mission. In this way, targeting the top offenders will result in improved mission readiness and mission capability.

The mission abort rate, $F_{r,msn}$, is calculated using Eq. (2) over any given time period where the number of events is the number of mission abort events.

The probability of mission failure, per sortie, due to a system failure can be represented by:

$$F_{msn} = F_{s,msn} = F_{r,msn}\, T_{avg} \text{ such that } 0 \leq F_{msn} \leq 1 \tag{11}$$

Which allows for the determination of the probability of mission success, per sortie, as follows:

$$R_{msn} = 1 - F_{msn} \tag{12}$$

Care must be taken when the per sortie mission abort probability is calculated, because it can be significantly affected by the average mission duration, T_{avg}. Therefore if the average mission duration is highly variable then a tailored approach focusing on rates and current average mission duration would be needed.

4. MEASURES OF SAFETY AND AIRWORTHINESS

4.1. Accident and Hazard Occurrence Rate

In the United States, Chapter 7 Section 7 of the Federal Aviation Administration's Aeronautical Information Manual (AIM) (FAA, 2025a) outlines the voluntary and mandatory reporting requirements to the *National Transportation Safety Board* (NTSB). The NTSB website publishes this statistical data which can be used to infer safety rates for similar aircraft types where data is limited.

In Canada, Section 3 of Transport Canada's AIM (TC, 2025b) outlines the mandatory reportable incidents to the *Transportation Safety Board of Canada* (TSB).

For the Canadian Armed Forces, the Department of National Defence mandates accident and hazard occurrence reporting through its Flight Safety program (DND, n.d.). Additionally, each occurrence is categorized by phase of flight (ground or air), damage severity, and flight safety compromise level (how severe the outcome could have been) among other categories.

In Europe, Regulation (UE) No 376/2014 (EASA, 2014) outlines the mandatory and voluntary reporting requirements through their online reporting portal.

The main differences between these reporting requirements is that in general more pick-list style meta data is mandated through the Canadian Armed Forces Flight Safety program which allows for more specific data analysis to better identify and understand underlying safety trends and correlate them to maintenance records. For example, the Canadian Armed Forces Flight Safety reports include damage level (from negligible to catastrophic), an indication of the *potential* for a severe outcome, and references to maintenance work orders, making them more easily usable for data analysis in combination with maintenance records. While damage levels are included in equivalent civilian records, the others are either not required or are not presented in a manner that can be subjected to data analysis.

The accident and hazard occurrence rate, $F_{r,occ}$, more fully discussed in section 6.2, is calculated using Eq. (2) over any given time period where the number of events is the number of accident and hazard events.

While we could apply the same process as before to calculate general occurrence probabilities, without considering occurrence severity these results can become misleading. For risk assessments, the accident and hazard occurrence rate, combined with outcome severities can be very useful to determine the risk to the fleet and to develop appropriate mitigating actions. This is stated solely for illustrative purposes in this paper and is likely to form the subject of future work.

5. MEASURES OF MAINTENANCE BURDEN AND SUSTAINABILITY

5.1. Post-Flight Unserviceability Rate

The post-flight unserviceability rate, $F_{r,post}$, also sometimes referred to as the break rate, is a measure of the rate at which aircraft are unserviceable and in need of repair following their mission or flight. It is calculated using Eq. (2) over any given time period where the number of events is the number of events where the aircraft is found unserviceable post-flight.

This measure can be used to determine the number of aircraft following a launch that will be available for a subsequent launch without needing maintenance, to plan maintenance crew schedules for a given operational flying schedule, or the determine the number of aircraft required to meet a flying program.

The predicted number of aircraft to be found unserviceable upon return to base can be calculated according to Eq. (13) based on the number of sorties or Eq. (14) based on mission time.

$$N_{F,post} = N_s F_{s,post} = N_s F_{r,post} T_{avg} \tag{13}$$

$$N_{F,post} = N_s F_{r,post} T_{msn} \tag{14}$$

One thing to keep in mind is that the total number of aircraft requiring maintenance following a mission is the sum of the number of aircraft that broke pre-flight and post-flight (average flight time, T_{avg}, is shown, however mission time, T_{msn}, can also be used, depending on the variability in average flight time at a given moment in time):

$$\begin{aligned} N_{F,pre+post} &= N_{F,pre} + N_{F,post} \\ &= N_s T_{avg}(F_{r,pre} + F_{r,post}) \end{aligned} \tag{15}$$

Therefore the number of aircraft from a given launch that could be expected to be available post-flight for a subsequent launch is given by:

$$N_{ac,post} = N_{ac,req} - N_{F,pre+post} = N_s(1 - F_{r,post} T_{avg}) \tag{16}$$

Equation (16) can be combined with Eq. (8) and stacked multiple times to project aircraft requirements for multiple waves of missions with differing numbers of required aircraft per wave and even differing mission duration. This aggregation can be performed in a simplified manner but also lends itself well to a dynamic simulation.

Another means to consider post-flight unserviceability rates is through reliability or the probability that the aircraft will be available for a subsequent mission post-flight. To do this we use the post-flight unserviceability rate per sortie, and first determine the unreliability by bounding it between 0 and 1:

$$F_{post} = F_{s,post} = F_{r,post} T_{avg} \text{ such that } 0 \leq F_{post} \leq 1 \tag{17}$$

From which the post-flight reliability becomes:

$$R_{post} = 1 - F_{post} \tag{18}$$

5.2. Component Removal Rate

Component removal rates, other than removed for access, are an indicator of higher level maintenance requirements (intermediate or depot maintenance) and program cost. Reducing component removal rates is therefore the primary objective to achieve cost control and sustainability.

The component removal rate, $F_{r,c}$, is calculated using Eq. (2) over any given time period where the number of events is the number of component removal events (excluding removed for access).

The forecasted number of component removals for which spare replacement components would be required during an interval of time where T flying hours are planned to be flown can be determined using Eq. (19). This approach forms the basis for asset management, and is capable of accounting for changing operational tempo. The usual method that asset managers typically plan for spare asset production is based on calendar averages, which only applies if the operational tempo is constant, and is inherently prone to gross under or over planning.

$$N_c = F_{r,c} T \tag{19}$$

The component removal rates need to be generated for all the components that are flight line removable if considering flight line maintenance requirements, but can be further broken down at the sub-component and sub-sub-component level if considering intermediate or depot level maintenance. For intermediate or depot level analysis, it may be preferred to calculate the per repair removal rates of sub-components and use the number of repairs instead of flying hours depending on the particular component because it is often desirable to set a specific number of intermediate or depot repairs in an effort to level load these repair organizations to maintain a steady flow of work.

5.3. Flight Line Maintenance Occurrence Rate

The sustainability of the maintenance organization is concerned with ensuring that the right number of qualified technicians are available to perform all the required maintenance to meet the flying program. While this leads to the planning of the size of the required maintenance organization, this paper will only go as far as determining the amount of maintenance to be performed.

The flight line maintenance occurrence rate, $F_{r,m}$, is calculated using Eq. (2) over any given time period where the number of events is the number of maintenance work orders generated (which is a combination of on-wing maintenance and component replacements). This can also be considered the sum of pre and post-flight occurrences.

It is then straight forward to determine the estimated number of flight line maintenance events over a period of time corresponding to T flying hours according to:

$$N_m = F_{r,m} T \tag{20}$$

This is only half of the asset management answer that we are seeking, which is why we require the average repair time for these maintenance events which will be discussed next.

5.4. Maintenance Hours per Flight Hour

A very important caveat to mention is that the number of maintenance hours for any repair likely includes at least a portion of concurrent hours. Therefore it is vital to understand the difference between average time to repair and the person hours to repair. The average time to repair is the time from the item becoming unserviceable to the time that it is repaired and serviceable, whereas the person hours to repair is the total number of hours that all technicians working on the repair performed to make the item serviceable. The latter allows us to size the maintenance organization, whereas the former enables an assessment of the speed at which the maintenance organization responds to repair requirements. We will focus solely on using person hours to repair for this paper, and incorporating mean time to repair assessments will be left for future work.

There are two approaches that can be taken to determine the average number of maintenance hour per flight hour, each with advantages and disadvantages stemming from the available data and its accuracy.

The first approach is to simply sum the number of maintenance person hours for all maintenance performed during a time period and divide it by the number of flying hours in the same time period according to Eq. (21).

$$T_{r,m} = \frac{\Sigma \text{First Line Maintenance Person Hours}}{\Sigma \text{Flight Hours}} \tag{21}$$

The total maintenance person hours over a specified period of time with T expected flying hours can then be determined by:

$$T_{mhrs} = T_{r,m} T \tag{22}$$

The second approach is to determine the average person hours to repair for each maintenance event using Eq. (23).

$$T_{m,avg} = \frac{\Sigma \text{First Line Maintenance Person Hours}}{\Sigma \text{Maintenance Events}} \tag{23}$$

With this approach, the total maintenance person hours over a specified period of time with T expected flying hours can then be determined by Eq. (24). The advantage of this approach is that the average repair times per sub-component can be used and then summed, as shown in Eq. (25). This is especially useful in cases where the sub-components exhibit cyclic repair rates that vary over different time intervals from component to component. This approach can also take into consideration inaccuracies in the number of maintenance person hours being recorded, where more appropriate repair person hours can be used specific to the component in question.

$$T_{mhrs} = T_{m,avg} N_m = T_{m,avg} F_{r,m} T \tag{24}$$

$$\begin{aligned} T_{mhrs} = \Sigma T_{m,avg} N_m \Big|_{\text{by component}} &= \\ \Sigma T_{m,avg} F_{r,m} T \Big|_{\text{by component}} & \end{aligned} \tag{25}$$

In the case of a predominantly condition-based maintenance program, where the majority of maintenance is unscheduled, the second approach is the only approach that should be used and it must be developed at the appropriate sub-component level for the analysis. This is the approach developed and used by the Royal Canadian Air Force (RCAF) for the modular F404-GE-400 engine powering the CF188 Hornet fighter aircraft, and being developed for other RCAF engine programs.

The approach taken should also consider the different classes of components in question and their individual failure modes or removal rate drivers. For example, some components are fly-to-fail, and are repaired or replaced on failure, while others may be lifed to a safe life or damage tolerant design philosophy with a life based on flying hours, operating hours, flights, component stress cycles (as is the case for many rotating gas turbine components).

6. MINIMUM DATA REQUIREMENTS

6.1. Maintenance Data

The ability to drill down to understand the sub-component contributions to the previously described measures is critical. Using the same methods at the aircraft level and applying them to the sub-component and sub-sub-component levels provides a common approach and language for all personnel in the organization. This leads to problem focussed collaboration between operators, maintainers, engineers, and managers instead of segregated silos of information with minimal interaction.

To calculate the measures described thus far, the following data must be available on every maintenance work order, ideally in a repeatable manner (i.e. a defined code or pick list) as opposed to free text, either explicitly or in a manner providing the means to extract this information:

1. Date of Maintenance Event
2. Root Cause Component Identifier

3. Phase of Operations:
   (a) Pre-Flight
   (b) Post-Flight
   (c) During Other Maintenance Activity
4. Component Replacement (including component serial numbers if applicable)
5. Total Person Hours to Repair

Additionally, the flight records are required with the following information:

1. Date of Flight
2. Flight Hours

6.2. Accident and Hazard Occurrence Data

Each accident or hazard occurrence report should include at least the following data, again in a repeatable manner as opposed to relying on free text:

1. Date of Occurrence
2. Aircraft and/or Component Identifier
3. Damage Level/Severity
4. Phase of Operations:
   (a) Air
   (b) Ground
5. Identifier to Link Occurrence to Maintenance Work Order

6.3. Averaging Intervals

The averaging intervals for all the measures are important. They should be long enough in duration to minimize large fluctuations in the measure but short enough to identify emerging trends. This is a delicate balance, and sensitivity studies should be performed to identify the optimal averaging period for a specific fleet of aircraft or specific aircraft sub-components. It is quite possible that the optimal aircraft level averaging interval is different than the optimal averaging intervals for some of the subcomponents, the engine for example. Each fleet of equipment is unique in terms of their operating environment, operational tempo, maintenance program, maintenance culture and operational culture. There are therefore too many permutations to consider in the scope of this paper. However, suffice to say that care must be taken to select the most appropriate interval that drives the right decisions at the right time and to recognize that the optimal interval can change over time in response to the changing operational environment. In this regard, evaluation of the variability in the rates should be useful.

The interval chosen should also account for seasonal effects if present, which may require an averaging interval accounting for between three to six months of operating data to be able to identify the seasonality of certain failures. Deployed operations in different environments can also have an effect, therefore this should also be considered when reviewing the results and making inferences from the results.

For the CT114 and J85-CAN-40 engine, an averaging interval of 24 months has been proven to be ideal in most cases based on internal analysis.

The averaging interval should also be long enough to account for a statistically significant number of flying hours, especially when measuring occurrences with a very low probability of occurrence. For example, if the predicted or target occurrence rate is $1 \times 10^{-5}/\mathrm{Flight\ Hour}$ then the averaging interval should account in general for at least $100\,000$ flying hours. If not, the measure would easily over state the rate if the interval was too short and under state the rate if the interval was too long.

While all the measures presented here use simple averages, additional fidelity can be achieved through generation of Weibull failure curves and interpreting their results. The preference should generally be to use a Weibull distribution and apply it to the measures developed in this paper, however there are instances where additional effects may prevent the direct use of a Weibull distribution. This can occur in cases where the failure distribution does not conform to flying hours, such as failures that are driven by take-off and landing cycles or some other usage parameter. The problem here may be that certain components may exhibit failure based on usage accumulation that is not tracked, and therefore simple averages may be the only means to assess reliability until tracking is mandated and sufficient time has accumulated to allow for further assessment. Attempting to generate a Weibull distribution can help determine if a Weibull model is appropriate and also if another time parameter is required to represent failure using a Weibull distribution. Confidence bounds on the failure rates could also be generated through the use of more sophisticated distributions, however this will be the subject of future work.

7. COMMERCIAL AVIATION APPLICATIONS

Thus far the focus has been on applications of a military context where often the interest is on the next launch with a minimum required number of aircraft that need to be part of that launch.

The measures developed here should be equally applicable to commercial aviation, assuming that the minimum data set requirements are met. However, their combined use is more pertinent when applied to planning aircraft requirements for either the entire fleet and its daily flying program, or considering an airline's operations at a single airport. Both applications are highly complex and depend on the timings of each flight which essentially becomes a dynamic rate of change problem. The goal then is to forecast the number of standby aircraft or the quantities of spare parts required at specific airports to prevent lost revenue service.

As a slight variation of the pre-flight abort rate, we can use the same approach to determine the number of flights likely to fail pre-flight in a given day as the sum of all the pre-flight failure probabilities for each individual planned flight, i, based on their individual planned flight times, T_{flight}:

$$N_{F,flights,pre} = \sum_{i=1}^{n} F_i = \sum_{i=1}^{n} F_{s,pre,i} = \sum_{i=1}^{n} F_{r,pre} T_{flight,i} \quad (26)$$

And the number of flights likely to result in the aircraft being unserviceable post-flight:

$$N_{F,flights,post} = \sum_{i=1}^{n} F_i = \sum_{i=1}^{n} F_{s,post,i} = \sum_{i=1}^{n} F_{r,post} T_{flight,i} \quad (27)$$

While these equations provide a macroscopic assessment, they can be combined and used together to model a daily flying program for a commercial airline to provide the information necessary to make an informed decision on the level of economic risk that the airline is willing to accept in terms of possible lost revenue flights.

8. J85-CAN-40 PROPULSION SYSTEM PERFORMANCE

The data and results presented here are normalized values for the J85-CAN-40 turbojet engine of the Royal Canadian Air Force's CT114 Snowbird air demonstration fleet, from 1996 to 2024. The focus of the results presented will be that of the J85-CAN-40 engine and its sub-systems, while the data has additionally been produced at the aircraft level internally to facilitate fleet level monitoring and continuous improvement initiatives.

8.1. Measures of Operational Effectiveness

The pre-flight and post-flight abort and overall mission abort rates are presented in Fig. 1. The top portion of the figure shows how the overall 24 month rolling average rates change over time, and the bottom portion provides the component drivers, which are identified by multi-letter codes assigned to specific components. The leading *B* identifies the engine system. The lower charts are also sorted by leading component offenders, so that the charts can be easily used to identify engineering priorities and effort with the potential for the largest gains.

Engine reliability is generally difficult to directly calculate, therefore the approach used, similar to that presented by Hurst (Hurst, 2006), is to calculate the failure probability (F) first and then derive the reliability as one minus the failure probability (as seen in Eqs. (10), (12), and (18). The reliability of the engine and its sub-systems as a whole has been relatively consistent over the last 30 years, with mission reliability averaging 99.8%, pre-flight reliability averaging 99.9%, and post-flight reliability averaging 99.6%. As expected, the reliability is seen to decrease from pre-flight, to overall mission, to post-flight.

8.2. Measure of Safety and Airworthiness

The engine and engine related accident and hazard occurrence rates are shown in Fig. 2, with the rolling 24 month average over time in the upper portion and the current (last 24 month) component drivers in the lower portion. Since there have been no air occurrences in the last 24 months, the lower portion of the figure only show ground occurrence rates. A very important caveat here is that the rates at which ground, and sometimes even air, occurrences are reported can vary significantly with shifts in organizational culture. It is most often seen with very low risk, low damage occurrences or where there is a perceived potential for an occurrence that significantly drives up reporting resulting in large spikes (for example between 2021 and 2024 for ground occurrences where additional visibility into maintenance issues not normally subject to reporting in the Canadian Forces Flight Safety Program). For this reason is it essential to be critical of measuring accident and hazard occurrence rates and not mis-interpreting these results without verifying the underlying data and individual occurrence reports. One means of data filtering to improve this issue is to ignore the low risk, low damage events and only consider the more severe events in the reported measure. This additional filtering was intentionally left out of Fig. 2 to demonstrate the potential effect that can easily lead to misinterpretation of the data.

8.3. Measures of Maintenance Burden and Sustainability

The post-flight unserviceability rates can be seen included in Fig. 1, with the component drivers shown in the lower right portion of the figure. Of particular interest here is that component *BEUE* is the top driver for pre-flight aborts, overall mission aborts, and post-flight unserviceability. Therefore these results show that investment in engineering effort into improving this component would have the largest effect on improving the overall reliability of the aircraft.

Flight line maintenance occurrence rates and component removal rates are both shown in Fig. 3. What stands out here

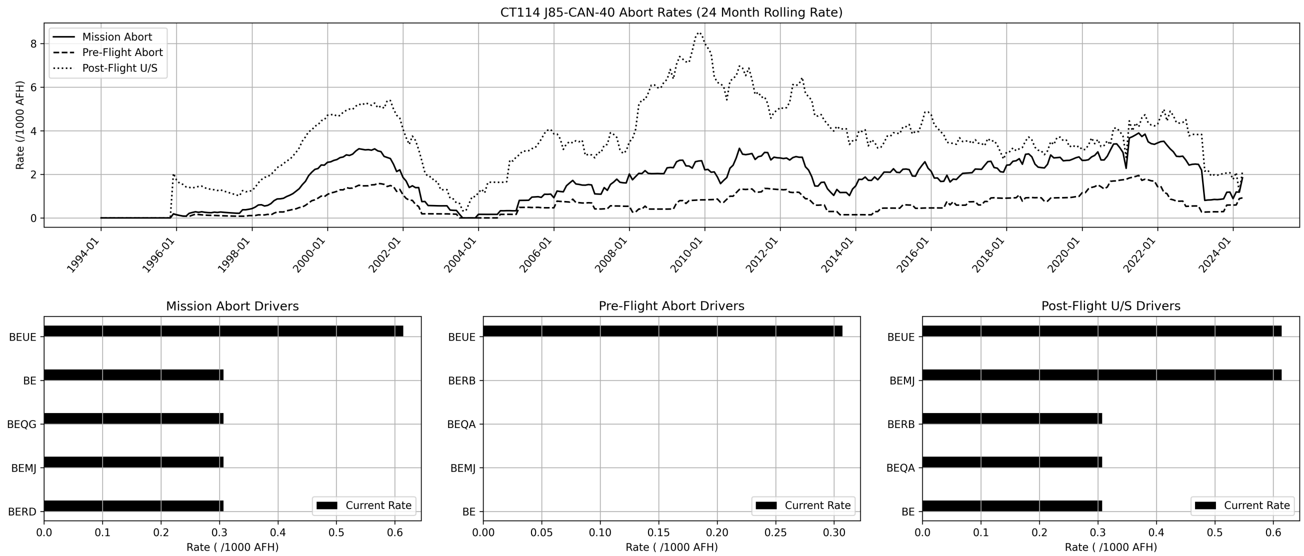

Figure 1. J85-CAN-40 Abort and Unserviceability Rates

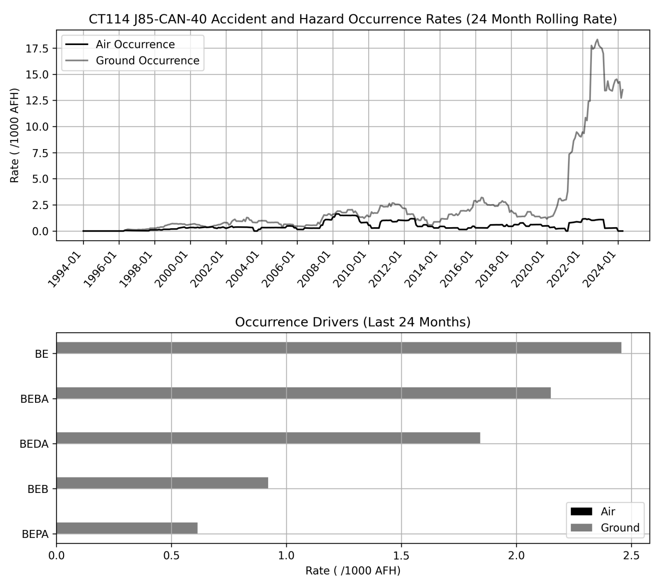

Figure 2. J85-CAN-40 Accident and Hazard Occurrence Rates

is that the amount of maintenance per flight hour has steadily increased, with the most notable jump occurring in 2018. Increased maintenance and inspection requirements introduced in this timeframe are one reason for this. Also of note, is that component *BEUE* does not appear as a first line maintenance driver, where it was the number one driver of pre-flight and mission aborts. This is not abnormal, and has been seen on other RCAF engine programs as well, where the maintenance drivers are not necessarily the same mission abort drivers due to the relative ease of replacing or repairing certain nuisance components. This introduces competing priority lists, and leads to the requirement to consider all of these drivers together to determine what corrective action can and should be taken.

The other important aspect of the data presented in Fig. 3 is that the current component driver rates are shown next to the all-time historical rate which provides an indication of improving or degrading trends. While the historical rates are not present on the other figures, the historical rates are produced and reviewed internally for all the measures presented in this paper.

The order of magnitude maintenance hours per flight hour for engine systems is provided in Fig. 4. The most interesting result is when we look at the component drivers for maintenance hours per flight hour. While components *BERC* and *BE* are the top two drivers for maintenance events, their order is reversed for maintenance hours per flight hour which is the true indicator of maintenance burden. Within the top five list, only *BEHA* additionally appears on both charts. Both figures present different information, and therefore both have complementary value in understanding the underlying performance of the system and its sub-systems. Reading both figures together allows an organization to identify nuisance offenders, which may appear to be of larger concern than they are in reality. In tandem they also provide the information necessary to justify taking action for a specific component over another and to effectively communicate that decision to all stakeholders in the organization, regardless of background.

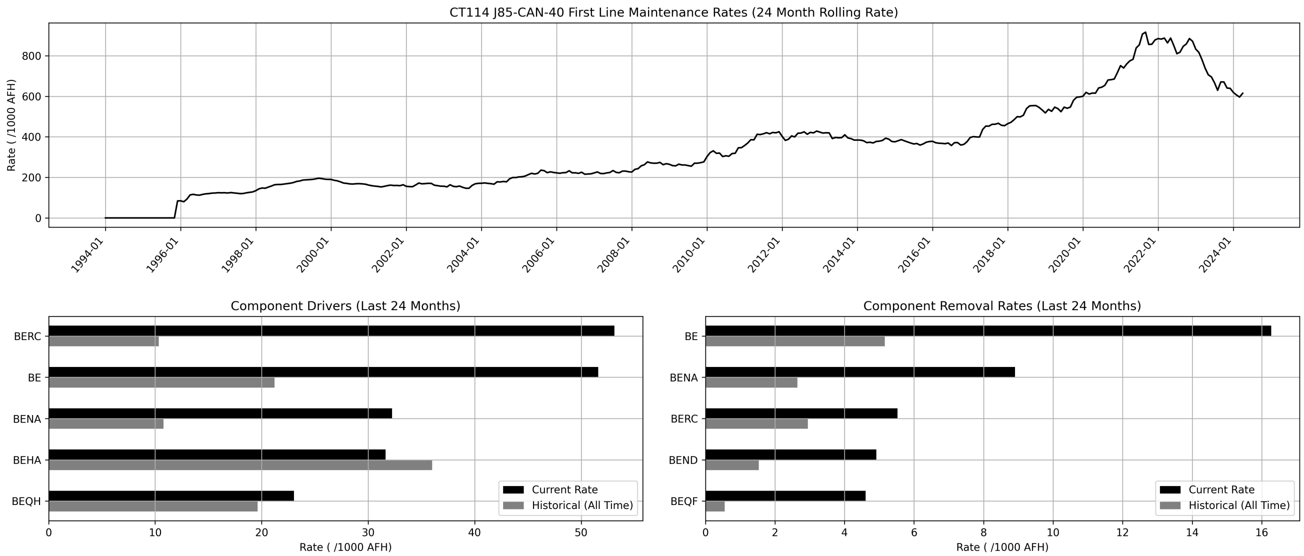

Figure 3. J85-CAN-40 Maintenance and Component Removal Rates

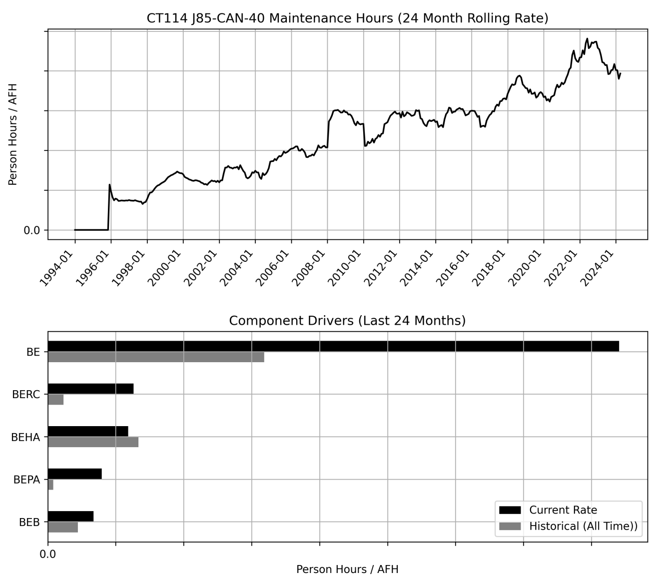

Figure 4. J85-CAN-40 Maintenance Hours per Flight Hour

9. DISCUSSION AND RECOMMENDATIONS

Development of key measures of fleet level performance are essential in understanding the fundamental drivers of unreliability of a fleet of aircraft and its sub-components. How this is developed and implemented is dependent on the data that is available, and the accuracy and variability of that data. Investing in developing common key, well-defined measures ensures that the entire organization has a foundational vocabulary with which to discuss technical and operational drivers of unreliability. Effort can then focus on issue resolution instead of issue definition.

Data fidelity, validation, and verification are also key aspects of performing any reliability assessment, which is not presented in this paper, but this requires inherent knowledge of the systems being assessed and the maintenance programs of these systems. Combining this knowledge, repeatable processes can be developed to properly filter records to ensure their accuracy and relevance. For example, records entered in error should be excluded, and multiple records referring to the same occurrence or maintenance action should be combined.

Identifying the link between maintenance records and accident and hazard occurrence records has been and still is one of the most difficult aspects of this work. Both record sets are separate and segregated, and for good reason, however this leads to difficulty in correlating the two sets of data together for data analysis. Within the RCAF, the accident and hazard occurrence reporting system is very good at recording most of the required information to be able to find the corresponding maintenance entry in an automated fashion, however there are instances where incorrect references to maintenance entries occurs. Although rare, there are also instances where there is no associated maintenance entry, mostly in cases where there is an aircraft crash or the aircraft or equipment is damaged beyond economical repair which can incorrectly skew the data.

To make this information readily available in the organization is the most important recommendation. When everyone is aware of what is being measured, it instils a shared sense of responsibility for continuous improvement and becomes part of that process. Keeping this information isolated in a data silo, only to be reviewed by reliability engineers, has the opposite effect on the intended outcome and has caused division between operators, maintenance organizations, and engineering support. Generating the clearly understood reports on a quarterly basis and disseminating them throughout the organization ensures that priorities are aligned and aids in contributing to organizational objectives while avoiding information overload. Adjusting these reports periodically in response to changing and evolving requirements for information within the organization ensures that the reports remain relevant towards driving continuous improvement.

10. FUTURE WORK

The measures and methods presented in this paper is the foundational work to develop a framework for a holistic reliability assessment of a fleet. Accident occurrence rate estimation by occurrence severity is planned to better understand the specific failure mechanisms driving risk to the fleet. Incorporating Weibull failure distribution assessments is also planned which will be compared to the method of using simple averages as presented in this paper, and which will further be used to determine confidence bounds on the measures for improved sensitivity analysis and engagement of the diverse and essential stakeholders.

11. SUMMARY AND CONCLUSION

Simple maintenance data, with the right minimal data requirements, was used to develop pragmatic and effective measures of system level performance and reliability at a fraction of the cost of a sophisticated on-board sensor suite. Sophisticated sensor systems are therefore not always required to develop a fundamental understanding of the technical issues of an aircraft or engine system. The RCAF's CT114 fleet data was used to identify the top fleet degraders and forecast their contribution to future fleet unreliability. These methods, if developed appropriately, can perform as well as complex on-board sensor systems. Data fidelity, validation, and verification is required to ensure repeatability and confidence in the results. The advantage of these methods is that they are easy to understand, which makes the results accessible to all personnel within the organization, contributing to organizational effectiveness in ways that results from complex systems, that can be difficult to understand, cannot.

As mandated by regulatory bodies within the aviation industry, existing maintenance data recording requirements provide a data set that is sufficient for pragmatically developing component, aircraft, and fleet level asset management capabilities.

ACKNOWLEDGEMENT

The author would like to acknowledge the support and assistance of Jeff Bird and Satpreet Singh Sidhu in preparing this paper, and the Royal Canadian Air Force for the provision of data for this paper.

NOMENCLATURE

Variables and symbols:

A	Availability
F	Unreliability / Failure Probability
F_r	Hourly Failure Rate
F_s	Per Sortie Failure Rate
$MTBF$	Mean Time Between Failure
N_{ac}	Number of Aircraft
N_c	Number of Components
N_s	Number of Sorties
N_F	Number of Failure Events
N_R	Number of Successful Events
R	Reliability / Probability of Success
R_r	Hourly Success Rate
R_s	Per Sortie Success Rate
T	Time
U	Unavailability

Superscripts and subscripts:

avg	Average
c	Component
m	First Line Maintenance
$mhrs$	Maintenance Hours
msn	Mission
occ	Accident and Hazard Occurrence
pre	Pre-Flight
$post$	Post-Flight
r	Rate
req	Required

REFERENCES

Andresen, & Williams. (2006, October). 144-07 Metrics, Key Performance Indicators, and Modeling of Long Range Aircraft Availability and Readiness. In *RTO-MP-AVT-144 Enhanced Aircraft Platform Availability Through Advanced Maintenance Concepts and Technologies.*

DND. (n.d.). *A-GA-135-001/AA-001 Flight Safety for the Canadian Armed Forces*. Department of National Defence (Canada).

EASA. (2014). *Regulation (EU) No 376/2014 of the European Parliament and of the Council*. Official Journal of the European Union.

EASA. (2018). *Regulation (EU) No 2018/1139 of the European Parliament and of the Council*. Official Journal of the European Union.

Elsayed, E. (2021). *Reliability Engineering* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

FAA. (2025a). *Aeronautical Information Manual, Chap. 7 Sec. 7.* Federal Aviation Administration.

FAA. (2025b). *Federal Aviation Regulations, Title 14, Chapter I, Subchapter C, Part 43 Maintenance, Preventive Maintenance, Rebuilding, and Alteration*. Federal Aviation Administration.

Hurst, D. (2006, October). 144-31 Operational Availability Modeling for Risk and Impact Analysis. In *RTO-MP-AVT-144 Enhanced Aircraft Platform Availability Through Advanced Maintenance Concepts and Technologies.*

NATO. (2000). *RTO-TR-28 AC/323(AVT)TP/22 Recommended Practices for Monitoring Gas Turbine Engine Life Consumption*. North Atlantic Treaty Organisation (NATO), Research and Technology Organization.

TC. (2025a). *Acts and Regulations, Standard 571 – Maintenance – Canadian Aviation Regulations (SOR/96-433)*. Transport Canada.

TC. (2025b). *Aeronautical Information Manual, Sec. 3*. Transport Canada.

Wang, C., Puigh, Lei, Guo, Yuan, & Mazarek. (2023, September). Lessons Learned from Aircraft Component Failure Prediction using Full Flight Sensor Data. In *4th Asia Pacific Conference of the Prognostics and Health Management.*

BIOGRAPHIES

Paul Bordush completed his BEng in Aeronautical Engineering from the Royal Military College of Canada, in Kingston, Ontario, Canada in 2009, and his MSc in Thermal Power from Cranfield University, Cranfield, UK, in 2014. He is currently working towards a PhD in Mechanical Engineering at the Royal Military College of Canada. He served with the Royal Canadian Air Force as an Aerospace Engineering Officer for 15 years and is currently working part-time with the Department of National Defence as a Propulsion System Expert and Program Manager. Since 2019, he has been a lecturer at the Royal Military College of Canada in Mechanical and Aerospace Engineering. He recently founded The Paul Aerospace Company Ltd. He holds a Professional Engineering licence in the province of Ontario since 2017.