Improving the Estimates of International Space Station (ISS) Induced "K-factor" Failure Rates for On-Orbit Replacement Unit (ORU) Supportability Analyses

Leif F. Anderson[1]
Blackhawk Management Corporation, Houston, Texas, 77059

Sean P. Harrington[2]
The Boeing Company, Houston, Texas, 77059

Ojei Omeke, II[3]
Blackhawk Management Corporation, Houston, Texas 77059

and

Dr. Douglas G. Schwaab[4]
The Boeing Company, Houston, Texas 77059

This is a case study on revised estimates of induced failure for International Space Station (ISS) on-orbit replacement units (ORUs). We devise a heuristic to leverage operational experience data by aggregating ORU, associated function (vehicle sub-system), and vehicle 'effective' k-factors using actual failure experience. With this input, we determine a significant failure threshold and minimize the difference between the actual and predicted failure rates. We conclude with a discussion on both qualitative and quantitative improvements the heuristic methods and potential benefits to ISS supportability engineering analysis.

Nomenclature

A	=	observed set of induced failure
B	=	expected set of induced failure
DC	=	duty cycle
ε	=	failure rate error function
F_{id}	=	induced failure quantity
F_{ih}	=	inherent failure quantity
F_o	=	(other) failure quantity
F_t	=	total failure quantity
F_r	=	random failure quantity
F_w	=	wearout failure quantity
fr	=	failure rate
fr^*	=	selection failure rate
fr_a	=	actual failure rate
fr_b	=	baseline failure rate
fr_e	=	effective failure rate
fr_O	=	ORU effective failure rate
fr_f	=	function effective failure rate

---

[1] Analyst II, Boeing Product Support, Logistics & Maintenance, HB3-30, AIAA Member.
[2] System Engineer 3, Boeing Product Support, Logistics & Maintenance, HB3-30.
[3] Analyst I, Boeing Product Support, Logistics & Maintenance, HB3-30.
[4] System Engineer 5, Boeing Constellation Program, HB4-20, AIAA Associate Fellow.

fr_v	=	vehicle effective failure rate
fr_{id}	=	induced failure rate
fr_{ih}	=	inherent failure rate
fr_t	=	total failure rate
I	=	index (for Reciprocal Threshold)
kF	=	k-factor
kF_b	=	baseline k-factor
kF_e	=	expected k-factor
kF_f	=	function effective k-factor
kF_O	=	ORU effective k-factor
kF_v	=	vehicle effective k-factor
MTBF	=	Mean Time Between Failure
n	=	total failure instance
Ops	=	Operating Hours
PDF	=	Probability Density Function
Q	=	ORU quantity
r	=	induced failure instance
Γ	=	significant induced failure threshold

I. Introduction

This paper introduces heuristic methods for integrating on-orbit experience data for the International Space Station (ISS) with the original estimate of induced failures for sub-system hardware. The correlative estimate, labeled the 'k-factor' in the Fisher & Price 1990 study (Ref. 1) on external maintenance for the Space Station Freedom, is a product of maintenance experience from analog systems (i.e. C-5A and F-16 aircraft programs) to measure expected human, environmental, equipment, and false maintenance drivers, which were factored into aggregate rates for space hardware best described as 'one-off' systems. These k-factor estimates were applied to unique on-orbit replaceable units (ORUs) that make up the ISS vehicle in the effort to accurately predict resource demand attributable to induced failures.

Program knowledge of the actual on-orbit maintenance experience now exists from approximately ten years of ISS operations. There are some instances where the baseline k-factor adopted at the program's inception, have been seen to skew the supportability projections when compared with the actual experience. Some of the observed error results from the k-factor's multiplier adjustment to the ORU inherent failure rate to determine the ORU total failure rate. If the projected failures are greater than what becomes the 'true' number of failures, the program may risk expending limited resources on unnecessary spare hardware. On the other hand, if the projected failures are less than what becomes reality, the program risks loss of functional availability in critical systems. Refined estimates of the k-factor are necessary to avoid both scenarios.

While the Fisher & Price analog approach was valuable for determining an initial estimate of additional Station maintenance for induced failures, these estimates obviously need to be updated as sufficient ISS hardware failure experience becomes available. Because of the uniqueness of the ISS system and relatively limited experience data, techniques to estimate demand or error rate uncertainty in military maintenance or industry production systems were not directly applicable. New heuristic methods were researched and developed to update the k-factor that leverage the available ISS failure experience and that recognize that the influence of the k-factor on the total failure projection is sensitive to the specific ORU inherent failure rate. These update heuristic methods evaluate the actual induced failure experience and the ORU predicted failure rate to establish whether the baseline k-factor estimates warrant a revision. The heuristics help address the attributes in the ISS ORU experience data, where: 1) induced failure is not the primary failure mode, 2) some ORUs have significant operating time and no failure, 3) ORUs with more than one induced failure can have many inherent failures, 4) and conversely, a small set of ORUs have only induced failures.

The basis for updating heuristic methods is the valuation of the 'effective' k-factors assigned to the ORU with information granularity ranging from ORU-specific to the vehicle-wide level. The 'effective' k-factor is defined as the ratio of total actual failures divided by the number of failures that were not induced failures – the algebraic equivalent to the Fisher & Price definition in terms of failure mode. The effective k-factor is calculated at the ORU-specific, the function and the vehicle levels. These values for each ORU are the input to two algorithms that determine the 'effective' k-factor to select as revision value or default to the baseline value. The first algorithm

establishes an upper-bound revision value (this study used the vehicle-level 'effective' k-factor) for ORUs with failures that surpass a threshold test of induced failure experience data. Several approaches were investigated to define the threshold for the 'significant' number of induced failure, which may determine the criteria for constraining the ORU k-factor revision at the upper-bound value. The second algorithm seeks to minimize the actual failure rate and resulting predicted failure rate, given the candidate set of 'effective' k-factors.

This paper is organized as follows. The paper provides a brief review of previous uses of a k-factor estimate in supportability engineering. The paper then introduces the upper-bound algorithm and presents several approaches for significance testing given relative uncertainty in current ORU experience data. Next, the paper discusses the heuristic algorithm to minimize the difference between effective k-factor failure rates to the actual failure rates. The paper shows some preliminary results for applying these heuristic methods to the ISS logistics and maintenance data and on-orbit experience. The paper addresses the main advantages and limitations identified for using this technique to revise the baseline k-factor estimate. Finally, the paper concludes with recommendations and extensions for adopting these heuristic methods for revising the station's ORU k-factor estimates in support of future space-exploration programs.

II. Past Use and k-factor Estimate Revision Purpose

The ISS program uses the k-factor definition from the Fisher and Price study. K-factor was a multiplier to the ORU inherent failure rate to cover the induced maintenance events, which was based on hardware categorization and experience from previous analog aerospace systems. For this study the inherent failures include both the random and wear-out failure mode, but ignores dormant failure mode that have been negligible in ISS experience. The inherent failures are considered failures caused by a source internal to the part, while induced failures are failures caused by an external cause. Failures for which the actual failure mode does not fit into the Fisher-Price k-factor categories are defined as other failures (F_o), and are assumed to be induced failures.

The equation for k-factor as provided in Ref. 1 is restated for our calculation of the 'effective' k-factors (kF_e).

$$
\begin{aligned}
kF_e &= \frac{induced}{inherent} + 1 = \frac{F_{id}}{F_r + F_w} + \frac{F_o}{F_r + F_w} + 1 \\
&= \frac{F_{id} + F_o}{F_r + F_w} + \frac{F_{ih} + F_w}{F_r + F_w} = \frac{F_{ih} + F_w + F_{id} + F_o}{F_r + F_w} \\
&= \frac{F_t}{F_t - \left(F_{id} + F_o\right)}
\end{aligned}
\tag{1}
$$

The concept of a k-factor has had various definitions in different industries for maintenance and supportability analysis. Extensive use of the k-factor estimate and consequently estimate revision methods long precedes the ISS program. Extensive research has been conducted throughout the military branches as reported in the Defense Acquisition University on-line library. For instance, maintenance on ground systems on USAF bases was studied for methods to improve operational planning (Ref. 2). While the ground maintenance estimate uses a multiplicative factor of induced maintenance upon predicted maintenance demand, the researchers relied on a questionnaire technique to derive the impact, proportion, and length of time spent during induced maintenance at different base locations. Likewise, the Fisher and Price study had adjusted the preliminary k-factor estimates from the analog aerospace maintenance systems with inputs from the astronaut office and other space logistics and maintenance experts.

There are somewhat different concepts and applications of the k-factor estimate in industry. When a k-factor estimate is used, greater importance is placed in the ability to measure secondary impacts to product availability and less emphasis on identifying specific failure modes. In industry applications, the k-factor becomes a measure of effectiveness of the quality regimes in Supply Chain Management. For supply chains feeding large scale projects, typically there is ample standardization and commonality among sub-components to reduce the risk of product loss (Ref. 4). This is a beneficial by-product of mass-reproduction. However for the ISS as a one-of-a-kind vehicle, the capability to perform on-orbit maintenance to restore critical system availability and redundancy closely depends on the accuracy of the failure rate projections. The induced failures increase the demands on Logistics and Maintenance planning, resources, and long-lead spares procurement and budget requirements. The k-factor can be viewed as a

logistics risk mitigation or management reserve factor for the unknowns and uncertainties for factors outside of the scope of the part reliability analysis and predicted inherent failure rate.

With this latter point in mind, the heuristic methods developed in this study also serve to identify 'significant' induced failure occurrences and flag potential risk to the overall ORU failure projection that requires further supportability assessment and strategic logistics replanning action. With few exceptions, there is not sufficient ORU experience data to provide enough samples for hypothesis testing or other data sampling techniques. This is an important point in our heuristic method study, since logistics planning is made against both 'epistemic' and 'aleatoric' uncertainty (Ref. 6, 8). 'Aleatoric' uncertainty "refers to phenomena occurring in a stochastic way" (Ref. 8). It is applicable to probabilistic assessment such as Monte Carlo simulation. Whereas, 'epistemic' uncertainty captures the analyst's confidence in the model validity by quantifying the degree of belief of the analysts on how well it represents the actual system (Ref. 8). Within the ISS program, engineering teams perform extensive analysis of the ORU failure events, and when possible perform tear-down evaluation of the failed ORUs after return to the ground to classify the actual failure mode and root cause. Therefore, the ORU on-orbit failure history data can help determine when the actual data demonstrates a reduction in the epistemic uncertainty in the ORU induced failure projection that may justify revision of k-factor parameter estimate.

The study considered several approaches to find a best fit for current ORU experience data and establish the 'significant' induced failure threshold criteria. The three most promising methods were:
 ⇒ the Failure Quantity (FQ) Threshold algorithm, a subjective decision approach;
 ⇒ the Reciprocal Threshold algorithm, a non-robust parametric model approach, and
 ⇒ the Binomial Set Test algorithm, a probabilistic model relying on set-logic approach.

The description and validity of these three methods in terms of their ability to reduce epistemic uncertainty is discussed next.

III. Model Structure

We take two sets of inputs for a model developed to select revised ORU k-factor estimates. These inputs include ORU reliability parameters for both inherent and induced failures, and the actual performance data gathered for the ORU. The latter is applied to our 'effective' k-factor calculation. The inputs feed two concurrent streams in our heuristic. As mentioned, we develop three methods that establish a threshold for 'significant' induced failure and thus identify those ORUs that warrant an increase from the baseline k-factor. Alternatively, various 'effective' k-factors associated with the ORU are input to an algorithm designed to minimize the difference between the failure rates determined by the 'effective' k-factor set, including the baseline k-factor, and actual performance data. We label this algorithm 'FRVA' (for Failure Rate Versus Actual). Since the ORU baseline k-factor is competing along with 'effective' k-factors in comparison to actual data, we will show that when an ORU does not have sufficient experience data, the FRVA algorithm defaults to the baseline k-factor. Otherwise, the FRVA allows for either an increase or decrease k-factor revision from the baseline. Our heuristic associates a pass/fail decision when adopting the k-factor revision determined by the 'significant' induced failure approaches. For any of the three investigated methods, if significant induced failures are not identified, the output from the FRVA algorithm applies, which recall, can again be either an increase or decrease revision. In both streams, the input 'effective' k-factors also serve as the output revision k-factor. Figure 1 provides an overview of the heuristic.
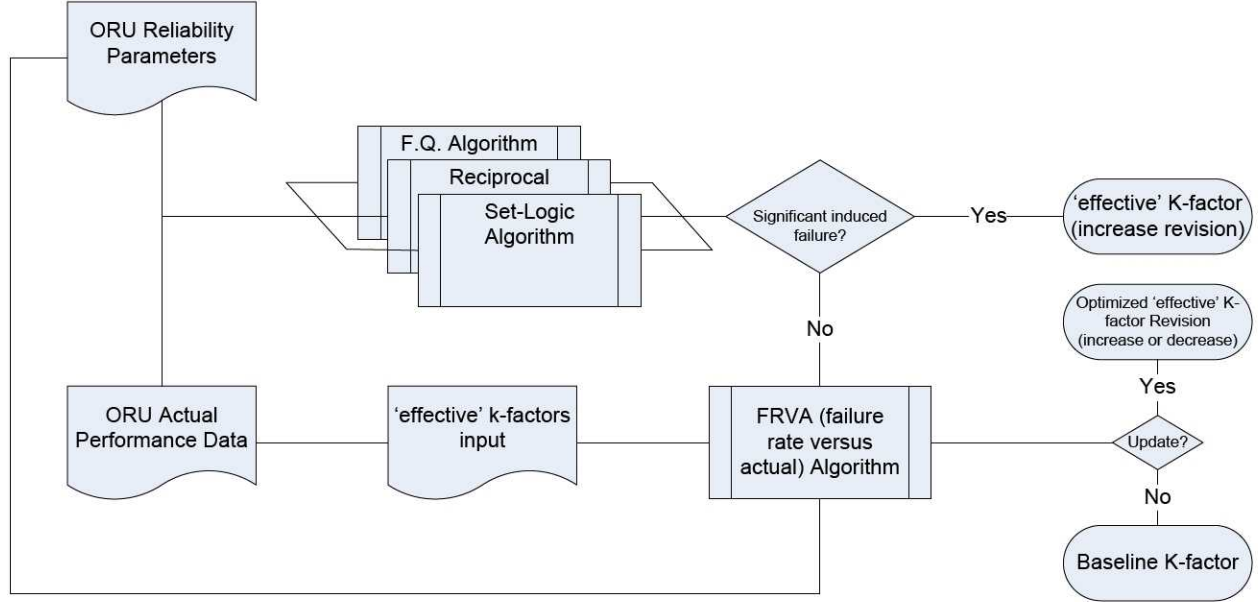
Figure 1. K-factor Revision Heuristic Flowchart.

A. Failure Quantity Threshold

The initial approach we consider for determining 'significant' induced failures, referred herein as FQ, is also most intuitive. For ORUs with two or more induced failures, the justification exists for increasing the baseline k-factor. The FQ threshold takes as significant the simple concept of a pattern. Its impact is measured by the ORU annualized failure rate equation:

$$fr = \frac{Q * DC * kF * 8760(\text{hrs}/\text{yr})}{\text{MTBF}} \quad (2)$$

In Eq. (2), we determine a generic failure rate for an ORU by multiplying the ORU operating quantity, Q, with a duty cycle, or portion of time activated, DC, and k-factor estimate, kF, divided by its Mean Time Between Failures, $MTBF$, which is measured in hours. Since the failure rate is typically annualized, we convert the hours to a year (using 8,760 hours in a year). The FQ threshold is expressed as the inequality between the actual failure quantities and the value represented by fr_e, the failure rate determined by the ORU 'effective' k-factor. The threshold is determined:

$$fr_e > \frac{Q * DC * \left(\dfrac{F_t}{F_t - 1}\right) * 8760(\text{hrs}/\text{yr})}{\text{MTBF}} \quad (3)$$

Where:

$$fr_e = \frac{Q * DC * \left(\dfrac{F_t}{F_{ih}}\right) * 8760(\text{hrs}/\text{yr})}{\text{MTBF}} \quad (4)$$

For ORUs satisfying this threshold criterion, FQ identifies the number of induced failures as significant and justifies an increase from the baseline k-factor. However, this FQ heuristic threshold (induced failure > 1) appears relaxed when we consider an ORU case with two failures, both of which are induced; and another ORU case with

5

twenty failures, out of which only two are induced. The FQ heuristic does not discriminate between these two cases and would adopt the same k-factor revision, even though in the second ORU case the number of induced failure mode is relatively minor effect compared to the total failure occurrences. In other words, the FQ heuristic method may unduly introduce conservatism into the k-factor updates.

B. Reciprocal Threshold

Therefore, we seek to constrain the threshold with the knowledge that failure uncertainty decreases as F_t increases. The simplest expression we find that meets this characteristic is by defining the threshold in terms of ORU induced failure by the following curve,

$$\Gamma(fr_a) = \frac{1}{F_{id}} \tag{5.1}$$

which is compared against an index value, or the proportion of induced failures to total failures per ORU.

$$I = \frac{F_{id}}{F_t} \tag{5.2}$$

Thus, the following expression determines significant induced failure:

$$\Gamma(fr_a) \leq I \tag{5.3}$$

This Reciprocal Threshold heuristic comparison equation provides a reasonable approximation of the epistemic uncertainty in the ORU experience data

Figure 1 graphically illustrates the Reciprocal Threshold criteria equation. The various potential ORU index lines cross the reciprocal threshold curve ($1/F_{id}$) with positive linear slopes of (F_{id}/F_t). The graph shows when the number of induced failures satisfies the threshold criteria. The threshold criteria will be met any time the number of induced failures equals or exceeds the square root of the number of total failures.
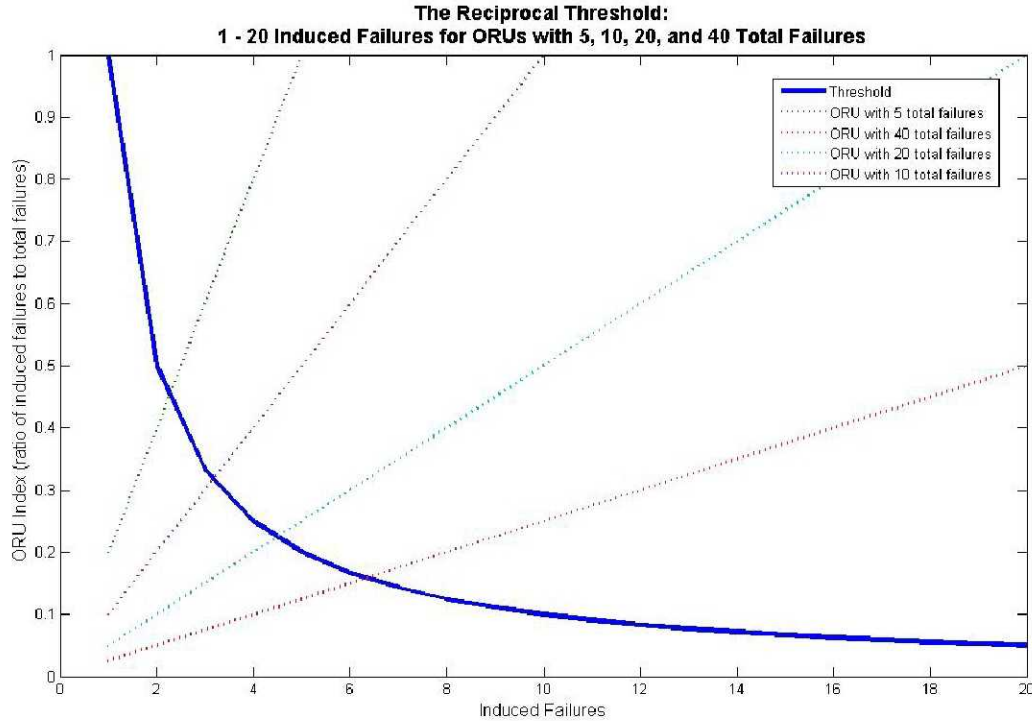
Figure 1. Various induced failure indices compared to the 'Reciprocal Threshold'. *Points on the lines depend on the induced proportion of total failure per ORU. Points above the threshold (blue curve) are considered significant induced failure quantity.*

In the previous example that compared two ORU cases: one case has two total failures, the other case has twenty, and both have two induced failures. Where the FQ approach identified both ORU cases as exhibiting significant induced failures; the 'Reciprocal' approach can discriminate between these two cases and only determines significance for the first ORU case with two total failures. The first case has an ORU index value of one, which is greater than the threshold, 0.50. The second ORU case with twenty total failures is illustrated by the light blue index line. It exhibits an insignificant number of induced failures given the total number of failure (threshold equal to 0.50, index value equal to 0.10). The Reciprocal heuristic provides a useful 'applicability check' on the amount of induced failure experience observed in the ORU actual performance data. However, the Reciprocal heuristic still depends on 'expert judgment' to establish the threshold criteria ratios in Equations 5.1 and 5.2, and is not based on a rigorous scientific analysis. The Binomial Set Test approach was investigated next as a more robust algorithm to define the threshold criteria for significant induced failure experience.

C. Binomial Set Test

Since a frequentist approach to hypothesis testing is unavailable due to the limited ISS experience data, a set-logic approach was considered to define induced failure significance. In this heuristic approach, a comparison is made between the number of induced failures expected with the baseline k-factor and the number of failures observed. As such, there is an inference component similar to the process used to update the ORU MTBF parameter. The ISS ORU MTBF estimate is updated annually using a Bayesian inference technique, with the actual ORU operating hours and the inherent random failures (Ref. 6). Revising the ORU k-factor is mutually exclusive from the MTBF Bayesian update. Bayesian inference was not considered a suitable technique to revise the k-factor estimate. Unlike the baseline MTBF estimate; there is no standard reliability and maintainability analysis method to determine the baseline k-factor values.

When constructing a set for expected and observed induced failure rates, an assumption is made on the underlying distribution of the ISS experience data. Following the example set forth by Wilson and Samaniego (Ref. 5), the induced failure is considered a 'success', r, in terms of a binomial distribution. This notion enables us to determine respective probabilities of the expected and observed sets. For the latter, the following statement defines a probability based on actual failures:

$$\Pr(F_{id}) = \frac{fr_t - fr_{ih}}{fr_t} = \frac{(kF_b - 1)}{kF_b} \tag{6}$$

For expected probability density function (membership function), we adopt, $\Pr(F_{id})$, derived from the baseline k-factor estimate, for a discrete binomial distribution given r induced failures out of n total failures per ORU:

$$P(r, \Pr(F_{id}), n) = \binom{n}{r} \left(\Pr(fr_{id})\right)^r \left(1 - \Pr(fr_{id})\right)^{(n-r)} \text{ for } r = 0, 1, 2, \ldots n \tag{7}$$

Alternatively, for the observed probability density function (PDF), we determine the observed probability of induced failure, $\Pr(F*)$, by substituting the expected induced probability, Eq. (6), with the observed ratio of induced failures instances out of n total failures.

$$\Pr(F*) = \frac{r}{n} \tag{8}$$

With the observed and expected probabilities in hand, we construct two sets for comparison, A and B respectively. Equation (1) serves as the membership function per the definition of a 'fuzzy set' established by Zadeh in 1965 (Ref. 7). The following statement determines our threshold for significant number of induced failure.

$$A \cap \overline{B} < A \cap B \tag{9}$$

This is read: if the area of the intersection between sets A and B is greater than the area of A not B, the ORU exhibits a significant number of induced failures. Figure 2 shows the relationship between A and B from a point where insignificant number of induced failures is observed with an arrow indicating movement towards meeting the threshold stipulated by $A \cap \overline{B}$. For an example of the threshold behavior, Fig. 2 illustrates the second ORU case (20 total failures with two induced failures) previously discussed, with a baseline k-factor of 1.40. Compare this scenario with Fig. 3 showing the same ORU, ten induced failures later.
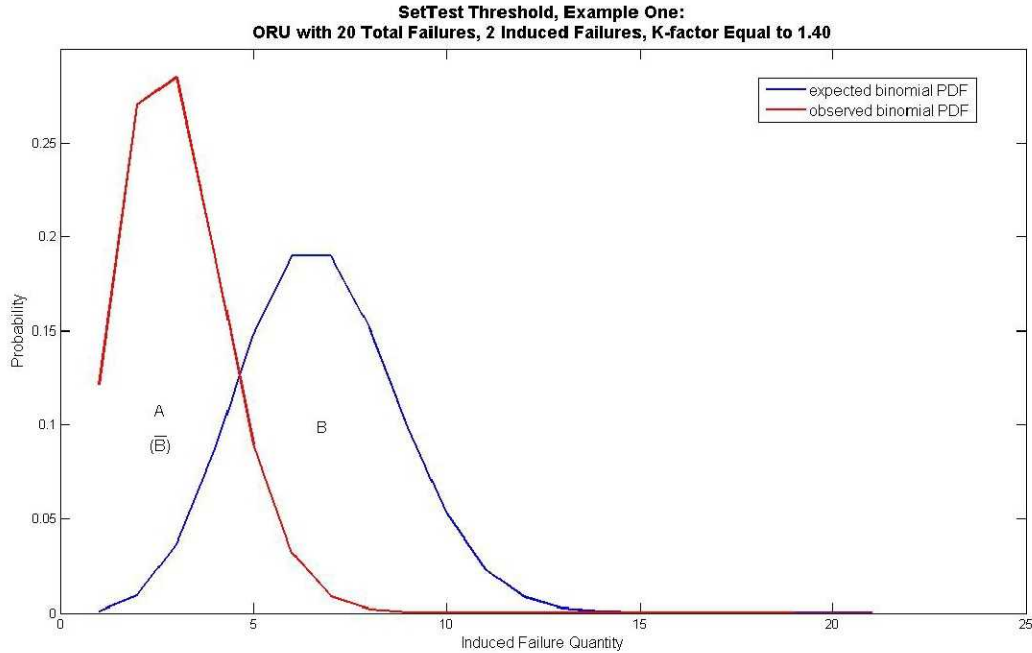
SetTest Threshold, Example One:
ORU with 20 Total Failures, 2 Induced Failures, K-factor Equal to 1.40

Figure 2. Example of Insignificant Induced failures under the 'Binomial Set Test' Threshold. *The observed set (red), A, has an area not intersecting with the expected set (blue), B, that is larger than the intersection area.*
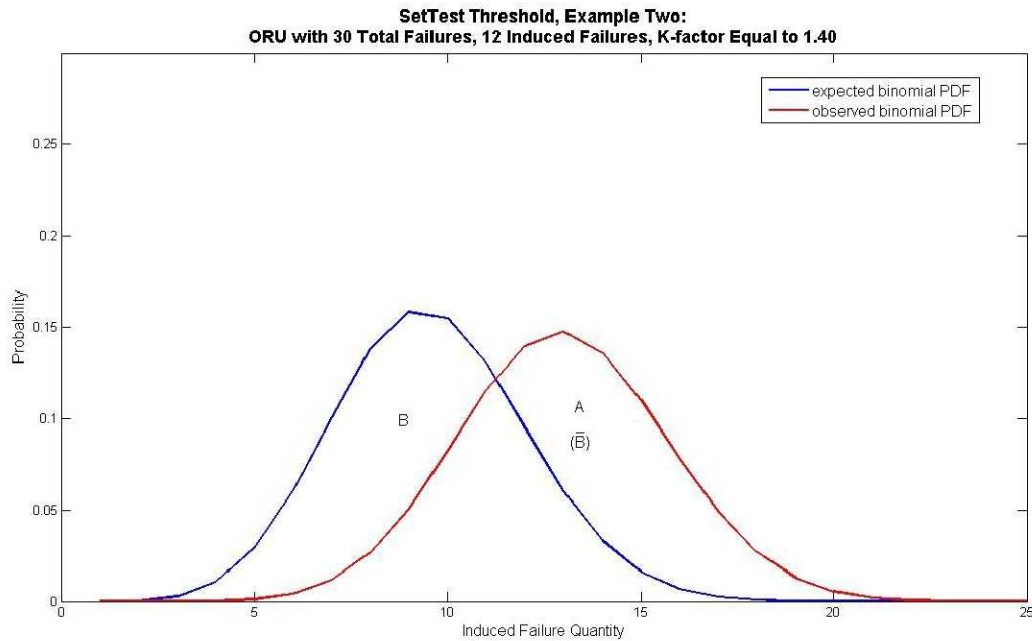

SetTest Threshold, Example Two:
ORU with 30 Total Failures, 12 Induced Failures, K-factor Equal to 1.40

Figure 3. Example of Significant Induced failure under the 'Binomial Set Logic' Threshold. *The set intersection is a larger area than that of the portion of set A (red) not part of set B (blue).*

In the first example, the area of A not B is approximately 0.73, whereas the area of the intersection is 0.27. Clearly, the ORUs observed performance data does not meet the expected number of induced failures to warrant 'significant' status. However, in our second example, the ORU has additional induced failures: 12 induced failures out of 30 total failures. The area of A not B here is 0.49, whereas the intersection has an area of 0.51; and since the intersection is greater, this ORU experience data now meets the threshold for having significant induced failures.

9
American Institute of Aeronautics and Astronautics

While the set-logic offers a more robust approach for defining significant induced failures, we noted an unexpected outcome when this heuristic was applied to ISS ORU experience data. ORUs with one total failure and no induced failure satisfied the set-logic threshold criteria to increase in k-factor baseline. This heuristic algorithm would need to be further modified to avoid this anomaly and prevent introducing even more conservatism into the k-factor revision than the FQ heuristic method.

The discussion of the Set Test method completes our investigation on heuristic methods to determine if the observed induced failure data satisfies the threshold criteria to justify an upper-bound revision to the baseline k-factor estimate. We detail how each approach is integrated into the broader k-factor revision heuristic process in the following section.

D. Calculating the 'Effective' k-factors

As defined in Eq. (1), the 'effective' k-factor is the ratio of total failures to failures not induced. Because of the limited ISS on-orbit experience and ORUs with insufficient failure data, the ORUs are aggregated into respective function-level and into a vehicle-level calculation of 'effective' k-factor. Thereby, every ORU baseline k-factor is subject to possible revision at some level. Appendix A maps the tree structure by which each ORU is assigned an effective k-factor. For each ORU, unique data, data relevant to its function, and the vehicle wide data is subject to the effective k-factor decision branches. These are characterized by the following four categories:

1) ORUs, functions, and the vehicle have only inherent failures
2) ORUs, functions, and the vehicle have only induced failures
3) ORUs, functions, and the vehicle have a combination of inherent and induced failures
4) ORUs, functions, and the vehicle do not have a failure instance

For ORU experience that falls in the third category, the associated function and the overall vehicle 'effective' k-factor will also fall into this category. Three of four categories (1, 2, 4) default to the baseline k-factor. This removes the instances where the lack of data or specific failure combinations result in 'not a number' (NaN) or a value of one. In the latter case, we default to the ORU baseline k-factor because the ORU still has intrinsic potential for induced failure. The aggregate quality of the 'effective' k-factor offers a practical means to reduce epistemic uncertainty. For a majority of the ISS ORU population, the current ORU failure instances can neither validate nor disprove the baseline k-factor estimate; so the heuristic process is allowed to choose between the baseline and the higher-level 'effective' k-factors. The next section discusses the FRVA heuristic method and explains how the ability to aggregate into higher-level 'effective' k-factor expands the usefulness of the heuristic's objective function.

E. Failure Rate Versus Actual (FRVA) - Error Minimization

With the 'effective' k-factor assignments per ORU as inputs, the FRVA algorithm selects the most appropriate revision to the baseline estimate. The algorithm chooses one of the input values of the k-factor estimate (baseline or 'effective') that minimizes the error between the predicted ORU failure rate and the observed ORU failure rate.

A more effective approach to determine whether an increase or decrease the baseline k-factor estimate is to calculate the failure rates Eq. (2) that are associated with the baseline, ORU, function, and vehicle level 'effective' k-factor values and the minimize the difference with the failure rate that is observed on-orbit as expressed by the objective function:

$$\min\left|\left\{fr_b,\ fr_O,\ fr_f,\ fr_v\right\} - fr_a\right| = \varepsilon \tag{10}$$

It is important to note that there are four subtraction operations in Eq. (10) applied to each ORU. The failure rate that results in the minimum difference is labeled the selection failure rate, fr^*. Therefore, if the difference between fr_O and fr_a results in the minimum error value, then fr_O is the selection failure rate. In practice, the absolute values for the four subtraction operators are hidden by the minimization operator. Therefore we use the value of the difference, or the error, ε, to provide the capability to solve for the selection failure rate using the ORU actual failure rate. Due to several constraints on the value of the selection failure rate in terms of the range of failure rates among the input parameters, the FRVA algorithm output (the 'optimized' k-factor) is determined under five unique scenarios:

Scenario:	Output Failure Rate Equals:

1	$fr_a = fr* \quad \therefore \varepsilon = 0$	fr_a
2	$fr_a > fr_b, fr_O, fr_f, fr_v$	$fr_a - \varepsilon$
3	$fr_a < fr_b, fr_O, fr_f, fr_v$	$fr_a + \varepsilon$
4	$fr_a \cap fr_b, fr_O, fr_f, fr_v$ $fr_a < fr* \quad \therefore fr_a + \varepsilon = fr_b, fr_O, fr_f, fr_v$	$fr_a + \varepsilon$
5	$fr_a > fr* \quad \therefore fr_a - \varepsilon = fr_b, fr_O, fr_f, fr_v$	$fr_a - \varepsilon$

The five scenario cases above are largely self-evident given the data. The selection failure rate can equal the actual failure rate, in which case the error is zero. The FRVA output failure rate is determined by the actual failure rate. Where the selection failure rate does not equal the actual failure rate, the selection failure rate is found given ε, where fr_a is greater than or less than all input failure rates. Under these circumstances, the output failure rate is found by subtracting ε from fr_a or adding fr_a and ε, respectively (Scenarios 2 and 3, respectively). Likewise, $fr*$ can also fall between the input failure rates. If fr_a is less than $fr*$, the output failure rate is the sum of fr_a and ε (Scenario 4), else if fr_a is greater than $fr*$, the output failure rate is the difference between fr_a and ε (Scenario 5). Given the output failure rate regardless of the five scenarios, the output k-factor estimate, baseline or revised, is determined by solving for k-factor in Eq. (2).

We can apply the ORU case examples to illustrate how the FRVA algorithm works. The second ORU case had two induced failures out of 20 total failures with a baseline k-factor of 1.40. The ORU has a MTBF of 20,000 hours and 100% duty cycle. Let us assume the example ORU is part of a function with observed twelve induced failures out of 40 total failures. Furthermore, we assume the function is part of a vehicle with 20 induced failure instances out of 100 total failures. The calculated ORU 'effective' k-factor is 1.11; function 'effective' k-factor is 1.42; and the vehicle effective k-factor is 1.25. We are evaluating this ORU k-factor after it has 300,000 operating hours. Thereby, it has an annualized actual failure rate of 0.5840 failures per year.

We compare this actual failure rate against the four determined by the various input k-factors. The baseline annual failure rate is 0.6132; ORU 'effective' failure rate is 0.4862; the function 'effective' failure rate is 0.6220; and the vehicle 'effective' failure rate is 0.5475. We then take the absolute value of the difference between the four failure rates and the actual and find the minimum value. This is our error, which is equal to 0.0292, which is determined by the baseline failure rate. Thereby, we consider the five data scenarios to determine the output failure rate from the FRVA algorithm. The error is not equal to zero. The actual failure rate is not greater than or less than all the comparison failure rates. It is not scenario one, two, or three. The actual failure rate plus error is 0.6132, while the actual failure rate minus the error is 0.5548. No failure rate matches the later value, so we find our data is under scenario four. When we solve for k-factor using Eq. (2), we do find indeed that the k-factor output from FRVA is 1.40, or the ORU baseline k-factor. In this ORU case example, applying the FRVA heuristic algorithm we find no revision is necessary, even though the experience data satisfies the FQ and Reciprocal threshold criteria. The preceding discussion should also show how the FRVA algorithm can produce possible revised k-factor values greater than the vehicle 'effective' k-factor.

IV. Model Application to Current ISS ORU Experience Data

At the time of writing, there are 275 unique internal ORU types in sixteen functions that make up the ISS vehicle population. The baseline k-factors range from 1.16 to 1.40, with a mean of 1.32. Approximately 90% of the ORU population has one or zero total failures, of which less than 9% are induced failures. Only 5% of the ORU population has two or more induced failures. There are six ORU types with ten or more failures. The mean ORU operating time is approximately 230,000 hours. However, we currently model 38 ORUs that have no on-orbit experience. (These ORUs have since been installed and activated, but have very limited on-orbit experience). Filtering out the ORUs that had not been activated, approximately 28% of remaining ORUs have operational experience less than a year. These are the basic statistics underlying our study for revising the ISS baseline k-factor estimate and characterize the heuristic input data.

F. ISS ORU K-factor Estimate Revision Assessment

The 'effective' k-factors are calculated from the current ISS ORU failure experience and have means that trend higher than the baseline values. Table 1 displays the summary comparison of the 'effective' and the baseline k-factor estimates. The ORU 'effective' k-factors range from 1.10 to 10.00 and have a mean of 1.38. The function 'effective' k-factors range from 1.08 to 10.00 and have a mean of 2.35. The vehicle 'effective' k-factor, the upper bound, is 1.54.

	Baseline k-factor	ORU 'effective' k-factor	Function 'effective' k-factor	Vehicle 'effective' k-factor
Mean	1.32	1.38	2.35	1.54
Min	1.16	1.10	1.08	1.54
Max	1.40	10.00	10.00	1.54
Revision (population %)	0%	5%	82%	100%

Table 1. Comparison of 'Effective' k-factors compared to baseline estimates. *The statistics are based on current ISS ORU experience data and reliability parameters. "Revision (population %)" measures the number of ORUs that result in an 'effective' k-factor estimate value not equal to the mean.*

Table 2 provides similar a comparison from our k-factor revision heuristic methods, showing the k-factor update results from the upper-bound heuristic approaches and failure rate minimization algorithm.

	F.Q.	Reciprocal	Set-Logic	FRVA
Mean	1.41	1.41	1.34	1.45
Min	1.08	1.08	1.08	1.08
Max	2.00	2.00	1.54	10.00
Revision (population %)	38%	38%	48%	38%

Table 2. Comparison of Upper-bound Algorithms and Failure Rate Minimization. *The statistics are based on current ISS ORU experience data and reliability parameters.*

For the set of ORUs that received a k-factor revision from the heuristic methods, the majority is characterized by having one or two failures and the k-factor revision value is not far from the baseline value. Dramatic increases from the baseline value are primarily to the ORUs more than four induced failures. In other words, we do not observe the heuristic methods producing widely divergent results. The FRVA algorithm is integrated in the overall heuristic revision process with the threshold approaches, when significant induced failure experience is not identified. In other words, for revised ORU k-factors, 91% were updated by FRVA to minimize the failure rate error after the FQ and Reciprocal heuristic methods were applied. For the portion of ORUs that received k-factor revisions from the Set-Logic heuristic, approximately 66% of the updates were from the FRVA algorithm. This is not to say the FRVA algorithm alone is entirely suitable for reducing epistemic uncertainty and neither is FRVA guaranteed to provide the 'best' revision estimates.
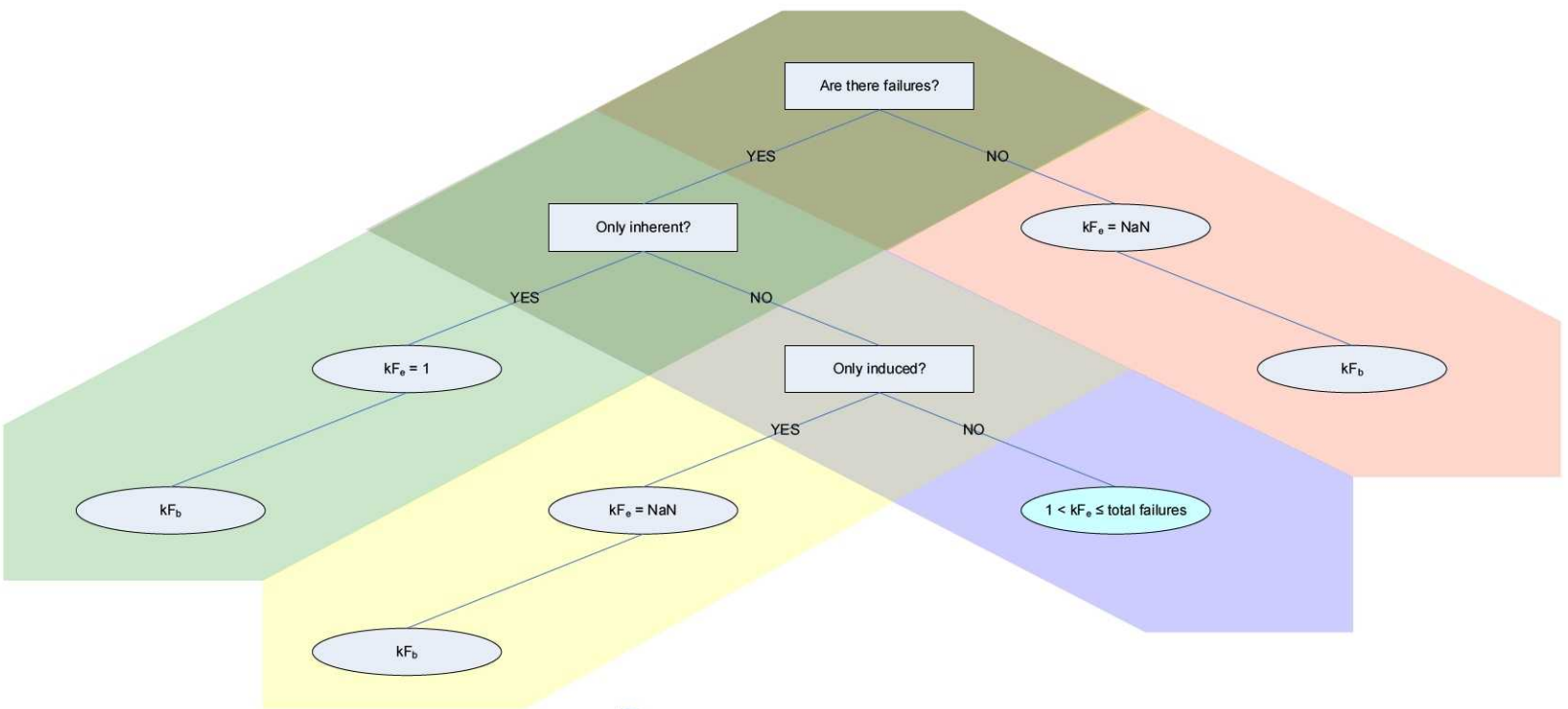
The overall revision heuristic process can be further improved with a subjective upper-limit revision value that would avoid revising the k-factor with gross over-estimate value, unless there really was sufficient ORU experience data to support an update value above the upper-limit. The upper-limit value could be set to the vehicle 'effective' k-factor or other values determined by logistics management or expert inputs. Likewise, the heuristic process can be improved with a lower-limit estimate for situations where FRVA reduced the k-factor estimate from the baseline values for ORUs with insufficient data. Due to the limitations in the ISS ORU on-orbit maintenance experience base and in the heuristic methods, a lower-limit for k-factor estimate value around 1.15 would be prudent to account for the uncertainty in predicting the induced failure rate. Again this update lower-limit could be over-ruled if there is sufficient ORU experience data to support a lower k-factor estimate value.

Applying a subjective upper-limit of the vehicle 'effective' k-factor (1.54) and a lower-limit of 1.15 to the initial case study heuristic results would further constrain the revised k-factor estimates for 23 ORUs. These upper and lower-limit adjustments created negligible difference from the initial heuristic revisions for supportability projections for ISS annual corrective maintenance actions, crewtime, and upmass for internal ORUs.

V. Conclusion

We have presented an investigation of heuristic methods to update ORU k-factor estimates for induced failure rates, which evolved over the course of many months and leveraged the available ISS on-orbit experience data. The ISS baseline k-factors were based on the Fisher and Price model for predicting induced maintenance. Our study focused on refining the k-factor estimates and removing some of the epistemic uncertainty in ISS supportability analyses and strategic logistics planning. We developed and evaluated several conceptual heuristic methods to revise k-factor estimates with limited ORU failure data. With the range of the observed ISS on-orbit data, only a few internal ORUs have experienced multiple failures (overall the station ORUs have been highly reliable), while most ORUs have insufficient induced failure data. So we first sought to assess the ORUs with induced failure data using heuristic methods that define threshold criteria for an upper-bound revision increase to the baseline k-factor. Contrary to expectations, we did not find that these heuristic methods, which seek to define significance in the induced failure data, produce the most accurate failure rate forecast. Also of interest, the Reciprocal approach yielded identical results to FQ approach on the data set we used. We had expected that for ORU experience data with a large quantity of induced and inherent failure instances, the Reciprocal heuristic would discriminate better and produce different revision estimates; but for current experience data, this is not the case. The Set-Logic heuristic proved effective, but exhibits a weakness in setting the threshold bar too low for 'significant' induced failures in the special situation with one total failure and no induced failures. Yet the majority of revision estimates for the ORU population is generated by FRVA algorithm, which minimize the difference between the selected 'effective' k-factor and actual failure rates. In this sense, the FRVA algorithm achieves its objective function. However, the overall heuristic process appears limited with respect to the remaining 62% of internal ORUs that do not warrant a revision to their baseline k-factor estimate.

This case study exemplifies the difficulty in truly resolving epistemic uncertainty. Ideally, supportability analyses are conducted to inform program management on the predicted state of the ISS through the end of life. However, past experience shows that in flying a vehicle as complex as ISS, the hardest decisions are often made on issues which can be very difficult to predict. The heuristic methods developed in this study provide a practical means to refine the baseline k-factor estimates and improve the ISS supportability projections and assessments accordingly.

Are there failures?

YES   NO

kF_e = NaN

Only inherent?

YES   NO

kF_e = 1

Only induced?

kF_b

YES   NO

kF_b

kF_e = NaN

$1 < kF_e \leq$ total failures

kF_b

Route Combinations			
failures?	inherent? (A)	induced? (B)	Progr... r...
yes →	no →	no	
yes →	yes →	(no)	
yes →	no →	yes	
no →	(yes →)	(yes)	

A("no" = 0, "yes" = 1), B = 2(A)

Acknowledgments

References

[1]Fisher, William F., and Price, Charles R., "Space Station Freedom External Maintenance Task Team, Final Report," NASA-TN-111430, 1990

[2]Bivens, Gretchen A., Kranz, Frank M., and Liberati, Gino L., "An Inside View of Air Force Ground Electronic Equipment Maintenance," *IEEE Annual Reliability and Maintainability Symposium,* 0149 144X/88/000-0262, IEEE, 1988, pp. 262-266

[3]Pecht, J. and Pecht M., *Long-term non-operating reliability of electronic products*. CRC Press, Boca Raton, FL, 1995

[4]Chopra, Sunil, and Meindl, Peter, *Supply Chain Management: Strategy, Planning, and Operations*, 2nd Ed., Pearson Prentice Hall, 2003, pp. 322, 394-395

[5]Wilson, Simon, P., and Samaniego, Francisco, J., "Nonparametric Analysis of the Order-Statistic Model in Software Reliability," *IEEE Transactions on Software Engineering,* Vol. 33, no. 3, 2007, pp. 198-208

[6]Parsons, Vickie S., Volovoi, Vitali, and Womack, James, "New Method for Updating Mean Time Between Failure for ISS Orbital Replaceable Units," NASA RP-05-131, v. 1.0, 2005

[7]Zadeh, L. A., "Fuzzy Sets," Information and Control, No. 8, 1965, pp. 338-353

[8]Zio, E., and Pedroni, N., "Building Confidence in the Reliability Assessment of Thermal-Hydraulic Passive Systems," Reliability Engineering and System Safety, No. 94, 2009, pp. 269-270