Common Cause Failure Modeling: Aerospace vs. Nuclear

James E. Stott[a*], Paul T. Britton[b], Robert W. Ring[b], Frank Hark[b], and G. Spencer Hatfield[b]

[a] *NASA Marshall Space Flight Center*, AL, USA
[b]*Bastion Technologies* MSFC, AL, USA

Abstract: Aggregate nuclear plant failure data is used to produce generic common-cause factors that are specifically for use in the common-cause failure models of NUREG/CR-5485. Furthermore, the models presented in NUREG/CR-5485 are specifically designed to incorporate two significantly distinct assumptions about the methods of surveillance testing from whence this aggregate failure data came. What are the implications of using these NUREG generic factors to model the common-cause failures of aerospace systems? Herein, the implications of using the NUREG generic factors in the modeling of aerospace systems are investigated in detail and strong recommendations for modeling the common-cause failures of aerospace systems are given.

Keywords: Probabilistic Risk Assessment, PRA, Staggered Testing, Non-staggered Testing, Common Cause Failure, CCF, NASA

1. INTRODUCTION

Aggregate nuclear plant failure data is used to produce generic common cause factors that are specifically for use in the common-cause failure models of NUREG/CR-5485. Furthermore, the models presented in NUREG/CR-5485 are specifically designed to incorporate two significantly distinct assumptions about the methods of surveillance testing from whence this aggregate failure data came. What are the implications of using these NUREG generic factors to model the common-cause failures of aerospace systems? Herein, the implications of using the NUREG generic factors in the modeling of aerospace systems are investigated in detail and strong recommendations for modeling the common-cause failures of aerospace systems are given.

2. HISTORICAL BACKGROUND

2.1. Early developments

Since the series of letters between Pascal and Fermat over a gambling dispute started the mathematical formulation of probability theory in 1654 [1], quantitative risk analysis has been continuously evolving over various applications. The development of mass production and the issue of variability in these applications spawned the introduction of statistical quality control in the late 1920s [2]. After WWII, the United States became more dependent upon mass produced electronics. In the early 1950s, the U.S. Department of Defense initiated a study on how to increase the reliability of a ubiquitous piece of electronics prone to failure – the vacuum tube. From this study, came the birth of Reliability Engineering and quantitative reliability analysis [3].

2.2. Advent of PRA

The beginnings of Probabilistic Risk Assessment (PRA) centered around the aerospace industry in the 1960s with the development of Fault Tree Analysis (FTA) in 1961 by Bell Laboratories for the Minuteman Launch Control System [4]. FTA was used by the Boeing Company to study the Minuteman Missile System, and further used by Boeing in the design of commercial aircraft. After the Apollo 1 launch-pad fire in 1967, NASA contracted Boeing to perform a risk assessment, and a Fault Tree Analysis was performed for the entire Apollo system [5].

In the 1970s, with the number of nuclear power plants growing, the safety of reactors became an important policy issue. The nuclear industry borrowed the techniques used in the aerospace industry to perform a PRA, and significantly contributed to further developing these techniques during this time. With the publication of WASH-1400 in 1975, event trees become a part of the PRA as well as the concept of the "common mode" failure. However, the quantification that is provided in WASH-1400 for common mode failures was known to need improvement. As Joel Yellin of the Rand Corporation put it, "the validity of the particular procedure used to set numerical values of common mode failure rates and uncertainties is extremely weak." [6]

2.3. Common Cause Failure Analysis and the Return to NASA

From the release of WASH-1400 forward, many developments were made in the field of quantifying Common Cause Failures (CCF), some of which were the beta factor model (Fleming, 1975), Marshal-Olkin specializations (Vesely, 1977), and the MGL and alpha-factor methods (Mosleh, et. al., 1985), which culminated in the release of NUREG/CR-4780 in 1988 [7], a joint Nuclear Regulatory Commission (NRC) and Electric Power Research Institute (EPRI) prepared document. Most recently, the release of NUREG/CR-5485 in 1998 [8] updates NUREG/CR-4780.

Another significant event in the history of PRA was the return of the use of PRA techniques, improved by the nuclear industry, back to NASA following the Space Shuttle Challenger accident in 1986. PRA continued to be used on the Space Shuttle, the International Space Station, Orbital Space Plane [9], and the Constellation Program. The Constellation Program has established the Constellation Program Probabilistic Risk Assessment Methodology Document (CxP 70017) which requires the use of alpha factors to be used in CCF modeling: "the Alpha-Factor parameters should be calculated based on relevant failure history using the method documented in NUREG/CR-5485."

The use of the alpha factor method presents several challenges when applied to aerospace applications which will be addressed in this paper. In order to properly demonstrate the issue, an overview of parametric CCF models will now be given which focuses on the relationship of those models to the alpha factors and the impact that testing schemes have on the results.

3. COMMON CAUSE FAILURE ANALYSIS METHODOLOGY REVIEW

3.1. Parametric Models

Among the approaches for analyzing and quantifying the effects of dependent failures in a system-failure analysis, inter-component dependencies at the basic-event level that are not explicitly modeled in the fault tree are modeled using parametric methods. These parametric models avoid the need to explicitly identify and enumerate causes [10] and are classified as shock vs. non-shock and direct vs. indirect. We will present a brief overview of representative methods within each category, but develop the basic parameter model in greater detail. The reader should refer to source material for a complete exposition of the theoretical development and quantification of the various methods.

The Binomial Failure Rate (BFR) model is a shock model specialized (Veseley, 1977) from a more general model developed by Marshall and Olkin (1967). It assumes individual similar components in a system configuration fail randomly at a constant rate (λ) with exponential mean time to failure, and common-cause failures result from common system shocks that occur randomly at an assumed constant rate (μ) according to a Poisson arrival process. Associated with the occurrence of a shock is the conditional probability of failure (p). The product of the shock rate times the conditional probability of failure given a shock produces the system failure frequency. The number of component failures realized given a failed component is modeled as a binomial distribution (hence the name). The shock rate is not directly available from the data because the shocks that do not result in a failure are not observable. It can be seen that three parameters (p, λ, μ) must be estimated regardless of the number of similar components in the system configuration [10].

Non-shock models do not model the process that generates the occurrence and number of common-cause failures. These are classified as direct and indirect. The basic parameter model is termed direct because it develops an expression for the total basic-event failure frequencies (Q_t) of each of the components in a common-cause group of m similar components directly from the minimum cutset representation of failure. Other models, such as the beta-factor model, the Multiple Greek Letter (MGL) model, and the alpha-factor model are termed indirect because they are re-parameterizations of the basic parameter model. The utility of the indirect methods is due to practical reasons arising from the demands imposed on the experience database due to the number of model parameters to be estimated and the parameter uncertainties resulting from estimates based on sparse or non-existent data. All of these methods require common-cause failure data, but among the indirect methods, the beta-factor model is the least demanding because it requires the estimation of only one common-cause parameter (β) in addition to the independent component failure rate (λ_I) to model the total component failure rate (λ_T).

Following the presentations in NUREG/CR-4780 and NUREG/CR-5485, the basic parameter model is best explained with an example using a two-out-of-three parallel configuration of similar components (A, B, and C). Without considering common-cause, the simple 2/3 parallel system configuration is shown below in Figure 1 as a fault tree diagram.

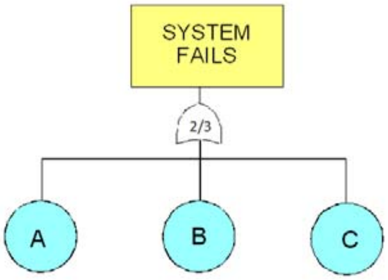

Figure 1 2/3 without CCF

Solving the above tree for the minimum cutsets we obtain the following Boolean expression (with notation defined below Eq. 2) for system failure:

$$S = A_I{*}B_I + A_I{*}C_I + B_I{*}C_I \qquad (1)$$

To incorporate common-cause into the tree, the basic events need to be converted to gates and expanded as sub trees to represent the ways in which each of the three components can fail with common-cause involvement. This is illustrated in the Figure 2 below.

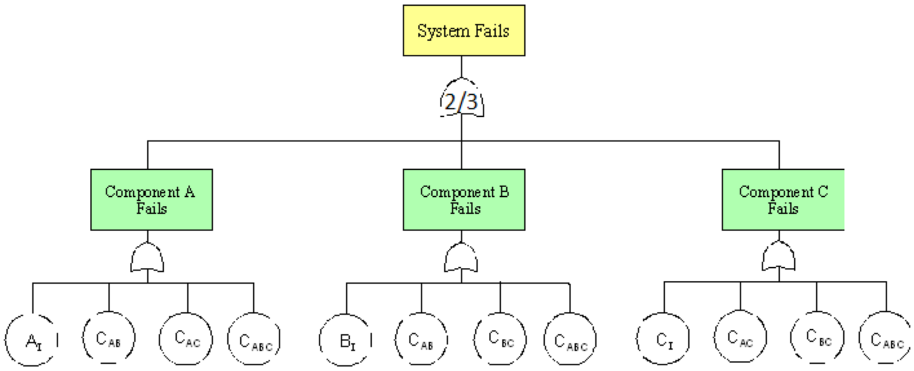

Figure 2 2/3 with CCF

The Boolean representation of the total failure frequency of component A is

$$A_T = A_I + C_{AB} + C_{AC} + C_{ABC} \tag{2}$$

Where
A_T = Total failure of components A.
A_I = Failure of component A from independent causes.
C_{AB} = Failure of components A and B (and not component C) from common causes.
C_{AC} = Failure of components A and C (and not component B) from common causes.
C_{ABC} = Failure of components A, B and C from common causes.

The reduced Boolean representation of the system failure in terms of these cutsets is

$$S = A_I*B_I + A_I*C_I + B_I*C_I + C_{AB} + C_{AC} + C_{BC} + C_{ABC} \tag{3}$$

For a one out of three success criteria, the Boolean expansion would have produced events such as $C_{AB}*C_{AC}$, which have questionable validity. One solution is to define these events as mutually exclusive, which implies $P(C_{AB}*C_{AC}) = 0$. Defined in this way, the Boolean representation of the total failure (A_T) of component A is a partition of the failure space of A into mutually exclusive event sets according to the explicit impact on the other components in the common-cause group.

Using the rare-event approximation, $P(A \cup B) \approx P(A)+P(B)$ and calculating probabilities yields

$$\begin{aligned} P(S) &= P(A_I*B_I) + P(A_I*C_I) + P(B_I*C_I) + P(C_{AB}) + P(C_{AC}) + P(C_{BC}) + P(C_{ABC}) \approx \\ &P(A_I)*P(B_I) + P(A_I)*P(C_I) + P(B_I)*P(C_I) + P(C_{AB}) + P(C_{AC}) + P(C_{BC}) + P(C_{ABC}) \end{aligned} \tag{4}$$

Since the components A, B, and C are identical, an exchangeability (symmetry) argument is made for the components A, B, and C, which allows for simplification and reduction of the above expression. Accordingly, we introduce the following variables:

$$\begin{aligned} &P(A_I) = P(B_I) = P(C_I) = Q_1 \\ &P(C_{AB}) = P(C_{AC}) = P(C_{BC}) = Q_2 \\ &P(C_{ABC}) = Q_3 \end{aligned} \tag{5}$$

Then substitute the variables in Eq. 5 into the expression in Eq. 4 to get the basic parameter model for the system failure frequency

$$P(S) \approx 3Q_1^2 + 3Q_2 + Q_3 \tag{6}$$

The expression for the total failure frequency of a component (Eq. 2) is easily generalized. For a configuration of m similar components in a redundant configuration, define the sequence of k probabilities

$Q_k^{(m)}$ – Probability of a common-cause basic event involving k <u>specific</u> components in a common-cause component group size m, $(1 \leq k \leq m)$

For a component-group size of m and for a given component (A), there are $\binom{m-1}{k-1}$ distinct subsets of k elements that include component A. Clearly, the total component-level failure frequency for each of the components in the configuration is given by

$$Q_t = \sum_{k=1}^{m} \binom{m-1}{k-1} Q_k^{(m)} \tag{7}$$

In the example given above for each of the components A, B, and C, the total failure frequency is

$$\mathrm{P(A_T) = P(B_T) = P(C_T) = Q_t = Q_1^{(3)} + 2Q_2^{(3)} + Q_3^{(3)}} \quad (8)$$

It's important to note that numerical estimates of $\mathrm{Q_k^{(m)}}$ derived from the generic common-cause database of component failure data is aggregated from multiple nuclear power-plant databases. There are uncertainties due to judgments inherent in the data-analysis process. Assigning failures to common-cause events also involves judgments. This is due in part because common-cause events occur in rough proximity over time and the classification is based on a heuristic method. Other sources of uncertainty are due to differences in nuclear power-plant design, operating procedures, maintenance procedures, hardware, and testing schemes. Consequently, the estimates of $\mathrm{Q_k^{(m)}}$ derived from the failure database are uncertain. Furthermore, the analysis process produces estimates of the numbers of common-cause failure events $(n_1, n_2, \dots, n_k)$. However, the number of demands is not recorded in the database; therefore, it must be estimated from power-plant records. Assumptions must be made about the testing scheme; since it directly impacts the number of challenges on a component group per surveillance test interval. The testing scheme affects the estimate of the number of demands because most of the demands on a redundant standby system are due to surveillance testing. There are two surveillance testing schemes to consider -- non-staggered and staggered. These are briefly explained below.

3.2 Staggered versus Non-staggered Testing

The power plant's technical specification requires that all m components in a component group be tested at least once in the surveillance interval, for example monthly. The actual interval may vary from one component type to another. Within a surveillance interval there may be several test episodes depending on the testing scheme used. In non-staggered testing, all m components are tested in a single test episode. Notice that in non-staggered testing; all $\binom{m}{k}$ distinct component subsets of k components (1 ≤ k ≤ m) are challenged in a single test episode. Except for classifying failure events that occur in consecutive episodes using impact vectors, if a demand failure occurs or is discovered during surveillance testing, the occurrence of a common-cause event $\mathrm{Q_k^{(m)}}$ will be discovered in a single test episode.

If staggered testing is used, there are multiple test episodes per surveillance interval. Typically, one component is tested per episode, but all m components are tested within the test interval. For example, suppose we test component A and it succeeds, to determine the number of k-component subsets challenged in the test episode, we infer that certain combinations of common-cause events have not occurred. For example, in addition to $\overline{\mathrm{A_I}}$ (not $\mathrm{A_I}$), we infer $\overline{\mathrm{C_{AB}}}$, $\overline{\mathrm{C_{AC}}}$, and $\overline{\mathrm{C_{ABC}}}$. So, in this case, we test component A, which by inference challenges $\binom{m-1}{k-1}$ additional distinct k-component subsets of the m component group for (1 ≤ k ≤ m) in a single test episode.

If we test component A and find that it failed, we cannot determine if other components have also failed due to a common-cause unless we test them. Therefore, all m-1 additional components in the group are also tested within that same test episode. Clearly, as for the non-staggered scheme, all $\binom{m}{k}$ distinct k-component subsets of the m-component group for (1 ≤ k ≤ m) are then challenged in a single test episode.

Let's summarize how non-staggered versus staggered testing impacts the estimate of the number of demands. For the class of similar redundant components in an m-component group, let $\mathrm{N_{TS}}$ be the total number of test intervals derived from the test specification represented in the database from all power plants, let $\mathrm{N_{TE}}$ be the total number of test episodes, and N_D be the total number of component demands. Assuming all plants use non-staggered testing, $\mathrm{N_{TE}} = \mathrm{N_{TS}}$ and $N_D = m\,\mathrm{N_{TE}}$. Now define N_k to be the total number of k-component subsets challenged. Since all m components in the group are demanded in a single episode, every k-component subset is also challenged. Therefore,

$$\mathrm{N}_k = \binom{m}{k}\mathrm{N}_{\mathrm{TE}} = \binom{m}{k}\mathrm{N}_{\mathrm{TS}} \tag{9}$$

Notice that $\sum_{k=1}^{m}\binom{m}{k} = 2^m - 1$. This is the total number of non-empty subsets contained in a set of m elements. The basic parameter model explicitly assumes the testing scheme is non-staggered. The maximum likelihood estimator (MLE) for $\mathrm{Q}_k^{(m)}$ is

$$\widehat{\mathrm{Q}}_k^{(m)} = {}^{n_k}\!/_{\mathrm{N}_k} \tag{10}$$

where n_k is the number of common-cause basic events involving k components derived from the experience database. Substituting

$$\widehat{\mathrm{Q}}_{\mathrm{k}}^{(\mathrm{m})} = {}^{n_k}\!\Big/\binom{m}{k}\mathrm{N}_{\mathrm{TE}} \tag{11}$$

Then, substituting $\widehat{\mathrm{Q}}_{\mathrm{k}}^{(\mathrm{m})}$ into the basic parameter model equation we get an estimate for Q_t.

$$\widehat{\mathrm{Q}}_{\mathrm{t}} = \sum_{k=1}^{m}\binom{m-1}{k-1}\widehat{Q}_k^{(m)} = \sum_{k=1}^{m}\frac{\binom{m-1}{k-1}n_k}{\binom{m}{k}\mathrm{N}_{\mathrm{TE}}} \tag{12}$$

Substituting the identity

$$\binom{m}{k} = \frac{m}{k}\binom{m-1}{k-1} \tag{13}$$

for the term in the denominator of Eq. 12 yields

$$\widehat{\mathrm{Q}}_{\mathrm{t}} = \sum_{k=1}^{m}\frac{\binom{m-1}{k-1}n_k}{\frac{m}{k}\binom{m-1}{k-1}\mathrm{N}_{\mathrm{TE}}} = \sum_{\mathrm{k}=1}^{\mathrm{m}}\frac{kn_k}{m\mathrm{N}_{\mathrm{TE}}} = \frac{\mathrm{N}_{\mathrm{T}}}{\mathrm{N}_{\mathrm{D}}} \tag{14}$$

where $\mathrm{N}_{\mathrm{D}} = m\mathrm{N}_{\mathrm{TE}}$ and the total number of component failures is given by N_T

$$\mathrm{N}_{\mathrm{T}} = \sum_{k=1}^{m} kn_k \tag{15}$$

For the staggered testing scheme, the NUREGs proceed to develop an expression for the total failure frequency (Q_{t}) as follows. In each test episode, only one component is tested unless that component is found to have failed in which case the remaining m-1 components are tested. To calculate the total number of challenges of k components and the total number of demands it is necessary to examine two cases.

Case 1: The component tested in an episode is found to be good. The total number of successful demands represented in the database is $\mathrm{N}_{\mathrm{D}}' = \mathrm{N}_{\mathrm{TE}} - \mathrm{N}_{\mathrm{CCBE}}$ where the total number of common-cause basic events is

$$\mathrm{N}_{\mathrm{D}}'' = \mathrm{N}_{\mathrm{CCBE}} = \sum_{i=1}^{m} n_i \tag{16}$$

This leads to the following expression for the number of k-component subsets challenged in the group of m components in one test episode.

$$\mathrm{N}_{\mathrm{k}}' = \binom{m-1}{k-1} \tag{17}$$

Recall, each time a component is found to be good following a test or operational demand, none of the common-cause events containing that component can occur, and there are N_k' such events.

Case 2: The component fails in a test episode. If an episode test or operational demand reveals a failed component, of which there are N_CCBE instances represented in the database, then all of the m-1 remaining components in the group are tested. Assume these additional tests are done in the same single test episode. Notice that the number of challenges for this test episode is the same as a non-staggered episode. Hence, the expression for the number of challenges on each k-subset in the group of m components in one test interval is

$$\mathrm{N}_\mathrm{k}'' = \binom{m-1}{k-1} \tag{18}$$

Eq. 18 is not $\mathrm{N}_\mathrm{k}'' = \binom{m}{k}$, since it is assumed that in the case of k+1 failures occurring in a single test episode, the failures would be mapped to a single common-cause and not a combination of an independent cause and a common-cause impacting the remaining k components. Therefore, combining Case 1 and Case 2 the total number of challenges of k-component subsets is given by

$$\mathrm{N}_\mathrm{k} = \mathrm{N}_\mathrm{D}'\mathrm{N}_\mathrm{k}' + \mathrm{N}_\mathrm{D}''\mathrm{N}_\mathrm{k}'' = (\mathrm{N}_\mathrm{TE} - \mathrm{N}_\mathrm{CCBE})\binom{\mathrm{m}-1}{\mathrm{k}-1} + \mathrm{N}_\mathrm{CCBE}\binom{\mathrm{m}-1}{\mathrm{k}-1}$$

$$= \binom{\mathrm{m}-1}{\mathrm{k}-1}[(\mathrm{N}_\mathrm{TE} - \mathrm{N}_\mathrm{CCBE}) + \mathrm{N}_\mathrm{CCBE}] = \binom{\mathrm{m}-1}{\mathrm{k}-1}\mathrm{N}_\mathrm{TE} \tag{19}$$

Then, substituting $\mathrm{N}_\mathrm{TS} = \mathrm{m}\mathrm{N}_\mathrm{TE}$ into the above result

$$\mathrm{N}_k = \mathrm{m}\binom{m-1}{k-1}\mathrm{N}_\mathrm{TS} = \mathrm{k}\binom{m}{k}\mathrm{N}_\mathrm{TS} \tag{20}$$

Therefore, the staggered estimate for the basic parameter $\mathrm{Q}_k^{(m)}$ is

$$\hat{\mathrm{Q}}_\mathrm{k}^{(\mathrm{m})} = {n_k}\Big/{\mathrm{k}\binom{m}{k}\mathrm{N}_\mathrm{TE}} \tag{21}$$

To calculate the approximate number of demands in the case of staggered testing, combine Case 1 and Case 2 and recognize that $\mathrm{N}_\mathrm{TE} \gg (\mathrm{m}-1)\mathrm{N}_\mathrm{CCBE}$ to get

$$\mathrm{N}_\mathrm{D} = \mathrm{N}_\mathrm{D}' + \mathrm{m}\mathrm{N}_\mathrm{D}'' = (\mathrm{N}_\mathrm{TE} - \mathrm{N}_\mathrm{CCBE}) + \mathrm{m}\mathrm{N}_\mathrm{CCBE} = \mathrm{N}_\mathrm{TE} + (\mathrm{m}-1)\mathrm{N}_\mathrm{CCBE} \approx \mathrm{N}_\mathrm{TE} \tag{22}$$

Substituting $\hat{\mathrm{Q}}_\mathrm{k}^{(\mathrm{m})}$ into the basic parameter model equation we get an estimate for Q_t:

$$\hat{\mathrm{Q}}_\mathrm{t} = \sum_{k=1}^{m} \frac{\binom{m-1}{k-1}n_k}{\binom{m-1}{k-1}\mathrm{N}_\mathrm{TE}} = \frac{\sum_{k=1}^{m} n_k}{\mathrm{N}_\mathrm{TE}} = \frac{\sum_{k=1}^{m} n_k}{\mathrm{N}_\mathrm{D}} \tag{23}$$

Notice, the estimate for the total failure frequency in Eq. 23, is equal to the total number of common-cause basic events (N_CCBE) in the database divided by the approximate total number of demands. Rather than this being the failure rate of a component it is a basic event frequency. Note the contrast between Eq. 23 and the non-staggered estimate of $\hat{\mathrm{Q}}_\mathrm{t}$ in Eq. 14.

3.2. Alpha Factor Model

The alpha-factor model for a component group of size m is a sequence of m conditional frequencies ($\alpha_1, \alpha_2, \ldots \alpha_m$), such that given the occurrence a of common-cause basic event, α_k is the probability that the event consists of k failures. The definition of the α_k in terms of the parameters of the basic parameter model is

$$\alpha_k = \frac{\binom{m}{k}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})}}{\sum_{k=1}^{m}\binom{m}{k}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})}} \tag{24}$$

Applying the definition for α_k in Eq. 24 along with the definition of Q_t in Eq. 7, we can express the basic parameters $\mathrm{Q}_\mathrm{k}^{(\mathrm{m})}$ in terms of α_k and Q_t. First, transpose the denominator in Eq. 24 to the other side of the equation, then substitute the identity, Eq. 13, to get

$$\mathrm{k}\alpha_k \sum_{k=1}^{m}\binom{m}{k}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} = \mathrm{m}\binom{m-1}{k-1}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} \tag{25}$$

Sum both sides over k ,($1 \le \mathrm{k} \le \mathrm{m}$)

$$\sum_{\mathrm{k}=1}^{\mathrm{m}} \mathrm{k}\alpha_\mathrm{k} \sum_{k=1}^{m}\binom{m}{k}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} = \mathrm{m}\sum_{k=1}^{m}\binom{m-1}{k-1}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} = \mathrm{m}\mathrm{Q}_\mathrm{t} \tag{26}$$

Letting $\alpha_\mathrm{T} = \sum_{\mathrm{k}=1}^{\mathrm{m}} \mathrm{k}\alpha_\mathrm{k}$ and rearranging we have

$$\sum_{k=1}^{m}\binom{m}{k}\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} = \frac{\mathrm{m}\mathrm{Q}_\mathrm{t}}{\alpha_\mathrm{T}} \tag{27}$$

Substituting the right side of the equation into the definition of $\alpha_k$and using the identity, Eq. 13, yields an expression for $\mathrm{Q}_\mathrm{k}^{(\mathrm{m})}$, ($1 \le \mathrm{k} \le \mathrm{m}$)

$$\mathrm{Q}_\mathrm{k}^{(\mathrm{m})} = \frac{\mathrm{k}}{\binom{m-1}{k-1}}\frac{\alpha_k}{\alpha_t}\mathrm{Q}_\mathrm{t} \tag{28}$$

Notice this equation is valid for both staggered and non-staggered testing. This re-parameterization of the basic parameter model is done in part to simplify data analysis; i.e., the alpha factors can be estimated from the common-cause frequency data without needing to know the number of demands as follows

$$\hat{\alpha}_k = \frac{n_k}{\sum_{k=1}^{m} n_k} \tag{29}$$

The derivation for Eq. 29 can be found in NUREG/CR-4780 [7].

4. ISSUES WITH THE STAGGERED METHODOLOGY

We have now shown the development of the basic parameter model and the alpha-factor method for both the non-staggered and staggered testing schemes. Recall, the contrast between Eq. 14 and Eq. 23; and between Eq. 11 and Eq. 21.

Table 1 Basic parameter estimates

Non-Staggered	Staggered
$\hat{\mathrm{Q}}_\mathrm{t} = \frac{\mathrm{N}_\mathrm{T}}{\mathrm{N}_\mathrm{D}}$	$\hat{\mathrm{Q}}_\mathrm{t} = \frac{\sum_{k=1}^{m} n_k}{\mathrm{N}_\mathrm{D}}$
$\hat{\mathrm{Q}}_\mathrm{k}^{(\mathrm{m})} = {n_k}\big/{\binom{m}{k}\mathrm{N}_\mathrm{TE}}$	$\hat{\mathrm{Q}}_\mathrm{k}^{(\mathrm{m})} = {n_k}\big/{\mathrm{k}\binom{m}{k}\mathrm{N}_\mathrm{TE}}$

The staggered estimates are not consistent with Eq. 28 and Eq. 29. Here, we will investigate in detail the cause of the anomalous result for the staggered case and will illustrate the problems that arise from using the staggered methodology.

4.1. Issue Statement

The issues with the staggered methodology can be summarized as follows: The staggered estimate for the basic parameters is based on a particular counting argument. This counting argument yields a greater estimate in the number of challenges on common-cause groups within a test interval which alters the frequencies of failure of the various common-cause groups. This reveals a subtly in the definition of the alpha-factor model. This subtly is not captured in the representation of the MLE of $\alpha_k, \widehat{\alpha}_k$. It has been, until now, incorrectly concluded that $\widehat{\alpha}_k$ is independent of the testing scheme assumption. The correct conclusion is that $\widehat{\alpha}_k$ is dependent upon the frequency of failure of a k-component group per system demand. And, this frequency is dependent upon both the testing scheme assumption and the database of failures, $\{n_k: k = 1, \dots, m\}$.

In order to clarify the issue, recall that the definition of a test interval is a period of time in which each component of a system is tested at least once. In addition, the definition of a system demand is a test that challenges each combination of failures exactly once. The distinction between these two is important and will become apparent in the following development.

In the context of a probabilistic model, we can consider a system demand to be an experiment that consists of a series of tests, precisely one for each of component, in one or more episodes along with an analytical step that maps component test results to a single outcome. The sample space for system demand on a three component system is given below:

$$\Omega^3 = \{\bar{F}, A_I, B_I, C_I, A_I * B_I, A_I * C_I, B_I * C_I, C_{AB}, C_{AC}, C_{BC}, A_I * C_{BC}, B_I * C_{AC}, C_I * C_{AB}, \\ A_I * B_I * C_I, C_{ABC}\} \tag{30}$$

where $\bar{F}$ represents that no failures occurred. Note that, in this model, the sample space of outcomes for a system demand is independent of any testing scheme assumptions. This is because the sample space for a system demand is a well-defined characteristic of the basic parameter model. To be consistent with the definition of an experimental outcome, only one outcome is possible for each test interval.

We now decompose the outcomes of a system demand into the parts that are used to estimate the model parameters, $Q_k^{(3)}$. Each outcome from a system demand is expressed as a combination of basic outcomes from the sample spaces shown below.

$$\Omega_1^3 = \{\bar{F}_1, A_I, B_I, C_I\}, \qquad \Omega_2^3 = \{\bar{F}_2, C_{AB}, C_{AC}, C_{BC}\}, \qquad \Omega_3^3 = \{\bar{F}_3, C_{ABC}\} \tag{31}$$

where $\bar{F}_i, i = 1,2,3$ represents no failure of a common-cause group of size i.

From these sample spaces, we can see that from a system demand, there are three opportunities for an individual component to fail independently; three opportunities for a common-cause group of size two to fail, and so forth. The method of enumerating the challenges to the members of $\Omega_i^3, (i = 1,2,3)$ for a system demand is straightforward and illustrated below in the case of a non-staggered test interval.

Table 2 Non-staggered using Counting Method 1

Non-staggered Test Interval										
Test Episode	A	B	C	$\mathrm{A_I}$	$\mathrm{B_I}$	$\mathrm{C_I}$	C_{AB}	C_{AC}	C_{BC}	C_{ABC}
	Components			Events Challenged						
Counting Method 1				$\mathrm{N_1} = \binom{3}{1} = 3$			$\mathrm{N_2} = \binom{3}{2} = 3$			$\mathrm{N_3} = \binom{3}{3} = 1$

It is clear that each outcome in $\Omega_i^3, (i = 1,2,3)$ is challenged exactly once per test interval. Notice, that a non-staggered test interval satisfies the definition of a system demand. Consequently, the estimator, $\widehat{\mathrm{Q}}_\mathrm{k}^{(\mathrm{m})}$, from Eq. 10, quantified by Counting Method 1, has an intuitive interpretation that agrees with the definition of the basic parameter, $\mathrm{Q}_\mathrm{k}^{(\mathrm{m})}$.

However, for the staggered case, the NUREG's adopt a different approach to enumerating the number of challenges to common-cause groups within a single test interval. Recall, that within each test interval, all components are tested exactly once unless a failure is observed. If a failure is observed in a test episode, then the remaining m-1 components are also tested in that episode. The result is additional challenges to certain common-cause groups for each test episode. The additional challenges that are taken into account are illustrated in the table below:

Table 3 Staggered using Counting Method 2

Staggered Test Interval										
Test Episode 1	A			$\mathrm{A_I}$			C_{AB}	C_{AC}		C_{ABC}
Test Episode 2		B			$\mathrm{B_I}$		C_{AB}		C_{BC}	C_{ABC}
Test Episode 3			C			$\mathrm{C_I}$		C_{AC}	C_{BC}	C_{ABC}
	Components			Events Challenged						
Counting Method 2				$N_1 = 1\binom{3}{1} = 3$			$N_2 = 2\binom{3}{2} = 6$			$N_3 = 3\binom{3}{3} = 3$

The table above illustrates that the challenges to k-component groups are counted k times each. We now come to the fundamental issue with the staggered model. It must be understood that for each system demand, each k-component group is given exactly one opportunity to fail. Therefore, using Counting Method 1, the frequency of failure of a k-component group per test interval equals the frequency of failure of a k-component group per system demand. However, using Counting Method 2, the frequency of failure of a k-component group per test interval equals k times the frequency of failure of a k-component group per system demand. For example, in Table 4, note the different representations of the frequency of failure of a k-component group using both counting methods,

Table 4 Frequency of Failure Representations

Frequency of Failure	Per Challenge	Per Test Interval	Per System Demand
Non-Staggered - Counting Method 1	$\mathrm{Q_k^{(m)}} = \dfrac{\mathrm{n_k}}{\binom{m}{k} N_{TS}}$	$\binom{m}{k} \cdot \mathrm{Q_k^{(m)}} = \dfrac{\mathrm{n_k}}{N_{TS}}$	$\binom{m}{k} \cdot \mathrm{Q_k^{(m)}} = \dfrac{\mathrm{n_k}}{N_{TS}}$
Staggered - Counting Method 2	$\mathrm{Q_k^{(m)}} = \dfrac{\mathrm{n_k}}{k\binom{m}{k} N_{TS}}$	$\mathrm{k} \cdot \binom{m}{k} \cdot \mathrm{Q_k^{(m)}} = \dfrac{\mathrm{n_k}}{N_{TS}}$	$\binom{m}{k} \cdot \mathrm{Q_k^{(m)}} = \dfrac{(\mathrm{n_k}/\mathrm{k})}{N_{TS}}$

The number of challenges per interval depends upon the testing scheme assumption. In Table 4, the two frequencies of failure per system demand are the correct quantities to use in the respective non-staggered and staggered estimates of $\widehat{\alpha}_k$. Therefore, the formula for $\widehat{\alpha}_k$ under the staggered assumption is $\widehat{\alpha}_k = \frac{(\mathrm{n_k/k})}{\sum_{k=1}^{m}(\mathrm{n_k/k})}$ and the formulae in Table 5, below, are consistent with Eq. 28.

Table 5 Basic parameter and alpha factor estimates

Non-Staggered	Staggered
$\widehat{\mathrm{Q}}_{\mathrm{t}} = \frac{\mathrm{N_T}}{\mathrm{N_D}}$	$\widehat{\mathrm{Q}}_{\mathrm{t}} = \frac{\sum_{k=1}^{m} n_k}{\mathrm{N_D}}$
$\widehat{\mathrm{Q}}_{\mathrm{k}}^{(\mathrm{m})} = {n_k}\Big/{\binom{m}{k}\mathrm{N_{TE}}}$	$\widehat{\mathrm{Q}}_{\mathrm{k}}^{(\mathrm{m})} = {n_k}\Big/{\mathrm{k}\binom{m}{k}\mathrm{N_{TE}}}$
$\widehat{\alpha}_k = \frac{n_k}{\sum_{k=1}^{m} n_k}$	$\widehat{\alpha}_k = \frac{(\mathrm{n_k/k})}{\sum_{k=1}^{m}(\mathrm{n_k/k})}$

Therefore, as seen in Table 5, the alpha-factors estimates are not necessarily only dependent upon the total number of failures of the various common-cause groups (as they are in the non-staggered case). Since, the alpha-factors represent, given the occurrence of an event from a system demand, the conditional frequency of failures of the various common-cause groups per system demand; and, because the failures are detected during test intervals, it is necessary to know the relative number of demands per common-cause group size per test interval. This has not been stated clearly until now.

5. NASA USE OF ALPHA FACTOR MODEL

The Constellation Program implements the NASA PRA guidance document [11] through the Constellation Program PRA Methodology Document CxP 70017, which specifies that common-cause be modeled using the Alpha Factor Model. Discussions among the Program's PRA analysts regarding the most appropriate testing scheme assumptions to use in the Constellation PRA were useful; but, ultimately failed to reach a consensus. Consequently, the Constellation PRA allowed each of the elements to select either non-staggered or staggered testing assumptions according to the method that best represented how their element hardware is to be tested.

However, it is improper to infer two distinct interpretations of the generic data within a system PRA. The reasons for this are clear. Using the staggered method, results in a lower incidence of common-cause. But, these benefits are an artifact of the model's assumptions. The difference that is captured by the mixed use of non-staggered and staggered methods is merely the difference that results from two interpretations of how generic data were collected less the effects of the error in the staggered use of the non-staggered $\widehat{\alpha}_k$.

PRA trade studies used to support Risk Informed Decision Making (RIDM) are being performed on the Constellation and Space Shuttle Programs. One of the main issues we see in this area regarding common-cause modeling is the use of both non-staggered and staggered testing assumptions to model common-cause within a single trade study. The trade study alternative that assumes staggered testing takes credit for a lower incidence of common-cause and therefore, a lower risk result than the alternative that assumes non-staggered testing. This practice biases the risk results by assuming the risk benefit going in and then taking credit for it in the results.

Aside from the significant amount of additional data that can be obtained with the relatively small number of additional tests that stem from the implementation of a staggered testing schedule, there may be real risk benefits to staggered testing because common-cause coupling factors induced by testing are reduced when similar components are tested in separate episodes. For example, the tests may be conducted by different operators using staggered testing as opposed to non-staggered testing. Common-cause failures due to same-operator would be mitigated using a staggered testing scheme.

These benefits, however, are not reflected in the basic parameter model. In order to quantify them, a cause level analysis of coupling factors should be performed.

6. CONCLUSION AND RECOMMENDATIONS

A thorough review of common-cause modeling methods should be conducted within NASA programs and projects of PRA. In light of the issues with the staggered methodology, the non-staggered method is the recommended approach. Instances in which models are found to be using the staggered method should be converted to the non-staggered method. This should be done regardless of the actual testing scheme used in surveillance testing or in pre-flight checkout.

The use of the staggered method results in an artificial reduction in model risk estimates and it biases trade-study results in favor of the trade alternative that uses the staggered method. The use of the staggered method results in a low risk estimate relative to the non-staggered method. A mixed use of the non-staggered and staggered methods within a trade study is particularly misleading because of the artificially induced bias introduced. Therefore, the decision-maker's ability to make correct decisions is undermined.

Future revisions of NASA documentation should incorporate a discussion of parametric methods with revisions that reflect the correct development of models. The use of the terms non-staggered and staggered with reference to parametric models should be eliminated because the parametric models do not explicitly model the root causes that lead to the difference in risk due to testing schemes. The terms should be reserved just for describing the testing schemes themselves.

References

[1] T.M. Apostol, "*Calculus, Volume II,*", 2nd edition, John Wiley & Sons, 1969, New York.

[2] A.J. Duncan, "*Quality control and industrial statistics*,", 4th edition, Richard D. Irwin, 1974, Illinois.

[3] A. Coppola, "*Reliability engineering of electronic equipment: an historical perspective*", IEEE Transactions Reliability, R-33(1), pp. 29–35, 1984.

[4] H.A. Watson, "*Launch Control Safety Study*,", Section VII Vol. 1, Bell Labs, Murray Hill, 1961, New Jersey.

[5] C.A. Ericson, "*Fault Tree Analysis – A History*", 17th International System Safety Conference, 1999, Orlando, FL.

[6] J. Yellin, "*The Nuclear Regulatory Commission's Reactor Safety Study*", The Bell Journal of Economics, Vol. 7, No. 1, pp. 317-339, 1976.

[7] A. Mosleh, et. al., "*Procedures for Treating Common Cause Failure in Safety and Reliability Studies (NUREG/CR-4780)"*, 1988.

[8] A. Mosleh, et. al., "*Procedures Guidelines in Modeling Common Cause Failures in Probabilistic Risk Assessment (NUREG/CR-5485)"*, 1998.

[9] J. Stott and Y. Lo, "*Application of Probabilistic Risk Assessment (PRA) During Conceptual Design for the NASA Orbital Space Plane (OSP)*", PSAM7, Berlin, Germany, 2004.

[10] NRC, "*PRA Procedures Guide (NUREG/CR-2300) Vol. 1"*, 1983.

[11] M. Stamatelatos, et. al. "Probabilistic Risk Assessment Procedures Guide for NASA Mangers and Practitioners", Aug 2002.