RECEIVED

JUL 26 1993

OSTI

NUREG/CR-5993
BNL-NUREG-52362
Vol. 2

Methods for Dependency Estimation and System Unavailability Evaluation Based on Failure Data Statistics

Detailed Description and Applications

Prepared by
M. A. Azarm, F. Hsu, G. Martinez-Guridi/BNL
W. E. Vesely/SAIC

Brookhaven National Laboratory

Science Applications International Corporation

Prepared for
U.S. Nuclear Regulatory Commission

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

AVAILABILITY NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 2120 L Street, NW, Lower Level, Washington, DC 20555-0001
2. The Superintendent of Documents, U.S. Government Printing Office, Mail Stop SSOP, Washington, DC 20402-9328
3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Office of Information Resources Management, Distribution Section, U.S. Nuclear Regulatory Commission, Washington, DC 20555-0001.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

DISCLAIMER NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability of responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

Methods for Dependency Estimation and System Unavailability Evaluation Based on Failure Data Statistics

Detailed Description and Applications

Manuscript Completed: May 1993
Date Published: July 1993

Prepared by
M. A. Azarm, F. Hsu, G. Martinez-Guridi/Brookhaven National Laboratory
W. E. Vesely/Science Applications International Corporation

Brookhaven National Laboratory
Upton, NY 11973

DISCLAIMER

Subcontractor:
Science Applications International Corporation
655 Metro Place South
Suite 745
Dublin, OH 43017

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Prepared for
Division of Systems Research
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001
NRC FIN A3295

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

JP

ABSTRACT

This report introduces a new perspective on the basic concept of dependent failures where the definition of dependency is based on clustering in failure times of similar components. This perspective has two significant implications: firstly, it relaxes the conventional assumption that dependent failures must be simultaneous and result from a severe shock; secondly, it allows the analyst to use all the failures in a time continuum to estimate the potential for multiple failures in a window of time (e.g., a test interval), therefore arriving at a more accurate value for system unavailability. In addition, the models developed here provide a method for plant-specific analysis of dependency, reflecting the plant-specific maintenance practices that reduce or increase the contribution of dependent failures to system unavailability.

The proposed methodology can be used for screening analysis of failure data to estimate the fraction of dependent failures among the failures. In addition, the proposed method can evaluate the impact of the observed dependency on the system unavailability and plant risk. The formulations derived in this report have undergone various levels of validations through computer simulation studies and pilot applications. The pilot applications of these methodologies showed that the contribution of dependent failures of diesel generators in one plant was negligible, while in another plant, it was quite significant. It also showed that in the plant with significant contribution of dependency to Emergency Power System (EPS) unavailability, the contribution changed with time. Similar findings were reported for the Containment Fan Cooler breakers. Drawing such conclusions about system performance would not have been possible with any other reported dependency methodologies.

This report is written in two volumes; Volume 1 describes a summary of the study and Volume 2 contains detailed discussions of the methods, formulations, analyses, and applications. We believe these methods and the associated software can be useful to analyze the failure data to estimate the fraction of dependent failures, and to evaluate its impact on system unavailability. At the same time, further work is needed before these methods can be used in conventional PRAs to evaluate a plant-specific risk profile. We also showed that the methods can evaluate the safety significance of dependent failures, their plant-specific nature, and their dynamic variations with time. These methods appear useful to enhance early detection of risk significant dependencies and to take prompt corrective actions.

CONTENTS

Appendix	<u>Page</u>
A. DERIVATION OF MINIMAL SEQUENCE LENGTH DURATION FOR CONSTANT FAILURE RATE	A-1
B. DERIVATION OF SYSTEM UNAVAILABILITY FORMULA USING THE DISTRIBUTION OF THE MSC LENGTH DURATION	B-1
C. MONITORING INDICATORS OF DEPENDENCY AND SYSTEM UNAVAILABILITY	C-1
D. STATISTICAL TEST FOR DETECTING OF CLUSTERING AND DEPENDENCY	D-1
E. ESTIMATION PROCEDURE FOR DETERMINING AVERAGE SYSTEM UNAVAILABILITY	E-1
F. EVALUATION OF FAILURE DEPENDENCY THROUGH SIMULATION STUDIES	F-1
G. PILOT APPLICATIONS	G-1

ACKNOWLEDGMENTS

The authors wish to thank the following NRC staff, C.E. Johnson, L. Abramson, P. Baranowsky, J. Johnson, and D. Rasmuson for their constructive comments and feedback. The authors are also indebted to C.E. Johnson, the NRC Technical Monitor, who guided us and provided technical input through the various phases of this study. We want to acknowledge L. Abramson from NRC, for development of the statistical test method given in Appendix D1, and R. Travis from BNL for his contribution in screening of the field data used for the applications presented in Section 3.2 and in Appendix G. Last but not least, we are indebted to Ms. J. Danko for her effort in typing and organizing this report.

APPENDIX A. DERIVATION OF MINIMAL SEQUENCE LENGTH DURATION FOR CONSTANT FAILURE RATE

A.1. Introduction

An important issue in the clustering of failures (or dependent failures) is the timing of the failures of similar components. To verify if a significant clustering has taken place, we compare the observations of times of failures with those expected under the assumption of independence. The objective of this appendix is to formulate the distribution of the interval between the failures of similar components when the components have failed independently. These distributions are then used in Appendix D to formulate the statistical analyses for detecting clustering.

We shall discuss some of the physical mechanisms of clustering to determine what are the meaningful time intervals for which we need to construct the distributions.

Clustering occurs when a cause is shared by several components resulting in significantly higher failure rates than their normal failure rates. This cause is usually referred to as the triggering event. If the triggering event is very severe and affects k out of n similar components, the failure rate of those components will become very large and the components will fail shortly a part. When the component failures are detected, the components will be repaired and the effect of the shared cause will be eliminated; thereby the component failure rate returns to its normal value. Triggering events that are very severe are known as the common-cause/common-mode failures and they also are sometimes referred to as shock model. However, when the triggering events are not very severe, the component failure rate increases to a higher value, and the failure times of similar components become clustered. The effect of the triggering event is removed whenever a failed component is repaired. Furthermore, in plants with a good maintenance program, whenever two or more failures clustered, the other similar components would be required to undergo preventive maintenance, thereby removing the effect of the triggering event before actual observation of all the failures. Therefore, two plants with the same frequency of the triggering events and the triggering severity levels may manifest different observations mainly due to more effective preventive maintenance in one plant compared to the other plant. The preventive maintenance, root-cause analyses, and corrective actions are important feedback mechanisms in controlling the risk from clustered or dependent failures.

Since a microscopic look at all triggering events, their frequency, severity, scope (i.e., number of similar components affected), and the effect of the plant operational feedback is not practical, a more macroscopic data analysis is proposed. Our proposed approach attempts to reduce the observed failure data to a set of time intervals: for example, an observed interval associated with failures of two similar component, an observed time interval associated with failures of three similar component, etc. We more formally define patterns of failure associated with each of these intervals, which discussed in Section A.2. Finally, we develop the distribution of the interval (or duration) associated with each pattern under the assumption of independence in Section A.3.

A.2. Definition Of Clustering Patterns

We assign a rank (or code) to any sequence of failures. A sequence of failure which involves failure of three similar components is assigned a rank of 3. As an example, if A, B, C, and D, represent four similar components, then a sequence of failure, which chronologically can be shown by AABCBA,

will be assigned a rank or code of three since the sequence only contains the failures of three components, A,B,C, but not D. For assigning rank or code to a sequence of failure, the chronological ordering of the failures are of no concern. That is, the sequences AABCBA, BCBAAB, and BCAABA are all sequences of Code 3.

We also define Minimal Sequence Code (MSC). A minimal sequence code of code k, is a sequence of failures where code k has reached for the first time. In our earlier example, AABCBA is a sequence code of 3; however, the first four failures of this sequence, i.e., AABC is a minimal sequence code of 3, because it reaches Code 3 for the first time. Generally, a portion of the failure sequence, which reaches a Code k for the first time (with minimum number of failures), is defined to be an MSC-k. Similarly, an MSC-k is a sequence of failure, which, if the last failure of the sequence is eliminated, the sequence code changes to k-1. The period within which an MSC-k is observed is referred to as the length duration of MSC-k. The shorter the duration of an MSC-k compared to its expected value under independent assumption, the higher the probability that the failures are clustered or dependent. The remainder of this section formulates the length duration distribution of an MSC-k under the assumption of independence and constant failure rates. It is important to note that the length duration distribution for various MSCs can be empirically estimated from the data. If this distribution is denoted by $f_k(t)$, then the probability that k fails to be simultaneously detected after a standby period of T; $P_k(T)$ would be:

$$P_k(T) = \int_0^T f_k(t) dt \quad \text{A-1}$$

It is important to note that for estimating the length duration distribution; $f_k(t)$, all observed data on MSC-k have been used regardless of failures being clustered (in a short window of time) or not. The methodology differs from the conventional methodologies by looking at failure clusters in a continuum of time rather than discrete windows of time. The probability of clustering in a shorter period is predicted based on distributional extrapolation. There are certain practical functional forms for $f_k(t)$ under clustering which are discussed in Appendices B and D and are supported by simulation studies in Appendix F.

A.3 Length Distribution of MSCs

It is of interest to derive the distributions of the length of times for MSCs of various codes. These distributions are derived taking advantage of the following simplifying assumptions:

1. failure times for different components are independent,
2. time between failures are distributed exponentially for each of the components, and
3. the static (long-term average failure rate) is the same for all components.

The length period distributions are discussed for MSC-1, MSC-2, MSC-3, and MSC-K, where "K" is an arbitrary number from 1 through "n". For all cases, "n" is the total number of components in the population.

Case 1 - Length Distribution for MSC-1

If n similar components are under observation from time zero; the distribution of " t ", where " t " is the time of the first failure, is discussed here. A simple diagram of the observations is shown below:

where " x " stands for the failure of a component in the population. If the waiting time associated with the components in the populations is denoted by $W_A, W_B, \dots, W_N$, " t " can be written as:

$$t = \min(W_A, W_B, \dots, W_N) \quad A-2$$

where " $\min$ " stands for the minimum. The distribution of " t " denoted by $f_1(t)$ is then determined by:

$$f_1(t) = \sum_S f_s(t) \prod_{I \neq s} (1 - F_I(t)) \quad A-3$$

where f_s and F_s stand for the probability density function (pdf) and the Cumulative Distribution Function (CDF) of the failure times of the component " S ". In case of exponential with a parameter λ_s for component " S ", we have:

$$f_1(t) = \Lambda_1 e^{-\Lambda_1 t} \quad A-4$$

where

$$\Lambda_1 = \sum_S \lambda_s \quad A-5$$

If all components have the same failure rate, then:

$$\Lambda_1 = n\lambda \quad A-6$$

This is a well known theorem which states that the superposition of the exponentially generated failures times can be described by an exponential distribution with a failure rate equal to the summation of the individual failure rates.

Case 2 - Length Distribution for MSC-2

A MSC-2 is depicted in the following diagram:

where "D" represents a failure of a component which is not S. "W" is the incremental length period for MSC-2 compared to MSC-1. "W" can be written as:

$$W = \min(W_{A \neq s}, W_{B \neq s}, \dots) \quad A-7$$

where $W_{A \neq s}$ represents the time to next failure of component "A", which is not the same as component "s". The distribution of "W" denoted here as $g_2(W_s)$ then can be determined by:

$$g_2(W_s) = \sum_{I \neq s} f_I(W) \prod_{\substack{J \neq I \\ J \neq s}} (1 - F_J(W)) \quad A-8$$

In case of exponential failure times for individual components, $g_2(W_s)$ can be written as:

$$g_2(W_s) = \xi_2 e^{-\xi_2 W_s} \quad A-9$$

where

$$\xi_2 = \sum_{I \neq s} \lambda_I \quad A-10$$

note that ξ_2 depends on s and more appropriately shall be denoted by ξ_{2s} . If all the components have the same failure rate " λ ", then " ξ_2 " is given by:

$$\xi_2 = (n-1)\lambda \quad A-11$$

The above assumption, i.e., all components have the same failure rate is used from here on to simplify the formulations significantly.

To get the distribution length of MSC-2, the value of "W" has to be added to the length of MSC-1. The distribution of the summation of two random variables becomes the convolution of the two distributions. That is:

$$f_2(t) = \int_0^t f_1(\tau)g_2(t-\tau)d\tau \quad A-12$$

where $f_2(t)$ is the length distribution of MSC-2, $f_1(t)$ is the length distribution of MSC-1, and $g_2(t)$ is the incremental length distribution of MSC-2 compared to MSC-1 as defined in Eqn. A-9

Substituting for f_1 and g_2 and performing the integration, $f_2(t)$ becomes:

$$f_2(t) = \frac{\Lambda_1 \xi_2}{\Lambda_1 - \xi_2} e^{-\xi_2 t} - \frac{\Lambda_1 \xi_2}{\Lambda_1 - \xi_2} e^{-\Lambda_1 t} \quad A-13$$

The expected duration of MSC-2 can then be determined by:

$$E(T_2) = \int_0^{\infty} t f_2(t) dt = \frac{1}{\xi_2} \left(\frac{\Lambda_1}{\Lambda_1 - \xi_2} \right) + \frac{1}{\Lambda_1} \left(\frac{\xi_2}{\xi_2 - \Lambda_1} \right) = \frac{1}{\Lambda_1} + \frac{1}{\xi_2} \quad A-14$$

Case 3 - Length Distribution for MSC-3

An MSC-3 is depicted in the following diagram:

where "W" is the incremental length of MSC-3 compared to MSC-2. The distribution of "W" can be derived similar to the previous case, that is:

$$W = \min(W_{A \neq s \neq D}, W_{B \neq s \neq D}, \dots) \quad A-15$$

$$g_3(W) = \sum_{\substack{I \neq s \\ I \neq D}} f_I(W) \prod_{\substack{J \neq I \\ J \neq s \\ J \neq D}} (1 - F_J(W)) \quad A-16$$

The definition of the terms are similar to those defined in the previous cases. In case of exponential failure times with a parameter of " λ " for individual components, we have:

$$g_2(W) = \xi_3 e^{-\xi_3 t} \quad A-17$$

where

$$\xi_3 = (n-2)\lambda \quad A-18$$

To get the total length distribution for MSC-3 (i.e., $f_3(t)$), the distribution of " g_3 " has to be convolved with " f_2 " as discussed previously, that is:

$$f_3(t) = \int_0^t f_2(\tau) g_3(t-\tau) d\tau \quad A-19$$

substituting for f_2 and g_3 from Eqns. 3 and 7, $f_3(t)$ can be expressed by:

$$f_3(t) = \frac{\Lambda_1 \xi_2 \xi_3}{(\xi_2 - \Lambda_1)(\xi_3 - \Lambda_1)} e^{-\Lambda_1 t} + \frac{\Lambda_1 \xi_2 \xi_3}{(\xi_3 - \xi_2)(\Lambda_1 - \xi_2)} e^{-\xi_2 t} + \frac{\Lambda_1 \xi_2 \xi_3}{(\Lambda_1 - \xi_3)(\xi_2 - \xi_3)} e^{-\xi_3 t} \quad A-20$$

The expected duration for MSC-3 can be determined by:

$$E(T_3) = \int_0^{\infty} t f_3(t) dt = \frac{\xi_2 \xi_3}{(\xi_2 - \Lambda_1)(\xi_3 - \Lambda_1)} \cdot \frac{1}{\Lambda_1} + \frac{\Lambda_1 \xi_3}{(\xi_3 - \xi_2)(\Lambda_1 - \xi_2)} \cdot \frac{1}{\xi_2} + \frac{\Lambda_1 \xi_2}{(\Lambda_1 - \xi_3)(\xi_2 - \xi_3)} \cdot \frac{1}{\xi_3} \quad A-21$$

Case 4 - Length Distribution for MSC-K

An MSC-K is depicted in the following diagram:

where " τ_{k-1} " is the length duration for MSC-(k-1), and "W" is the incremental length for MSC-k compared to MSC-(k-1). The distribution of "W", $g_k(W)$, can be determined using the same approach as previous cases for the exponential failure times and is given by:

$$g_u(W) = \xi_k e^{-\xi_k t} \quad A-22$$

where ξ_k in terms of " λ " for all the individual components becomes:

$$\xi_k = (n-k+1)\lambda \quad A-23$$

To get the total length distribution for MSC-k, i.e., $f_k(t)$, the distribution of " g_k " has to be convolved with the distribution of " f_{k-1} " as discussed previously, that is:

$$f_k(t) = \int_0^t f_{k-1}(\tau) g_k(t-\tau) d\tau \quad A-24$$

If we take a Laplace transform from both sides, we get:

$$f_k(s) = \int_0^t f_{k-1}(s) \cdot g_k(s) = \left(\frac{\xi_k}{s + \xi_k} \right) \cdot f_{k-1}(s) \quad A-25$$

Equation A-25 is a recursive formula. If we use the notation of " ξ_1 " for " Λ_1 ", $f_k(s)$ can be expressed by:

$$f_k(s) = \prod_{i=1}^k \frac{f_i}{(s + \xi_i)} \quad A-26$$

The inverse Laplace transform can be easily obtained and it is given below:

$$f_k(t) = \sum_{m=1}^k A_m e^{-\xi_m t} \quad A-27$$

where

$$A_m = \xi_m \prod_{\substack{j=1 \\ j \neq m}}^k \xi_j / (\xi_j - \xi_m) \quad A-28$$

The expected duration for MSC-k can then be determined by:

$$E(T_k) = \int_0^{\infty} t f_k(t) dt = \sum_{m=1}^k \frac{1}{\xi_m} \prod_{\substack{j=1 \\ j \neq m}}^k \left(\xi_j / (\xi_j - \xi_m) \right) \quad \text{A-29}$$

Substituting for ξ_k s from Eqn. A-23, the following equation results:

$$E(T_k) = \frac{n}{\lambda} \sum_{m=0}^{k-1} \frac{n}{n-m} = E(T_1) \cdot \sum_{m=0}^{k-1} \frac{n}{n-m} \quad \text{A-30}$$

Table 1 summarizes for various numbers of components the expected duration for the MSC of various codes in terms of multiples of the expected duration MSC-1 will be multiplied by. For example, for two components, the expected duration from failures of the first component to failure of the second component is triple the expected duration from the start of an observation to failure of the first component.

Table 1. Expected Duration of MSC* of Various Codes Divided by the Expected Duration of MSC-1 (this is defined as f_k)

MSC Code	n=1	n=2	n=3	n=4	n=5	n=6	n=10
1	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2		3.0	2.5	2.33	2.25	2.2	2.11
3			5.5	4.33	3.92	3.7	3.36
4				8.33	6.42	5.7	4.79
5					11.42	8.7	6.45
6						14.7	8.45
7							10.95
8							14.29
9							19.29
10							29.29

*Note: the expected duration of MSCs are calculated based on Eqn. (A-30).

A.4. Procedure For Data Reduction

The formulation of clustering distribution indicates two types of data reductions which are necessary to provide the needed observation on the model parameters. One is the observations on cycle duration, and the other is the observations on incremental cycle durations. The former provide direct observations or samples from the distribution of cycle distribution $f_k(t)$, where the latter provide observation on the parameters ξ_i of the $f_k(t)$. The following procedure for data reduction is suggested.

1. Identify a set of similar components that are expected to constitute a homogenous population and are expected to have the same failure rate.
2. Collect the failure data for a period. For each failure, record the component identification code and the time of failure (or an estimated time of failure based on the discovery time, cause information, and the test strategy and the interval).
3. Determine the non-overlapping MSCs of various codes and determine their length durations. These are observations on the $f_k(t)$. To clarify what is meant by a non-overlapping MSC, consider the following procedure for MSC-2.

To determine the length duration for non-overlapping MSC-2, start with the beginning of the observation period and continue until an MSC-2 is observed. The length duration is then the observation period within which the first MSC-2 was realized. To identify the second MSC-2, discard all the failures associated with the first MSC-2 and assume the beginning of the observation period is the time of the last failure in the first MSC-2. As an example, for a sequence of failures designated by ABDCADBA, the non-overlapping MSC-2s are: AB-DC-AD-BA.

4. Determine the running MSCs of various codes and calculate the incremental cycle durations. These are observations that are used to estimate the parameters ξ_i of the $f_k(t)$. For calculating these parameters the MSCs do not have to be non-overlapping, which is why they are referred to as the running MSCs. To clarify what is meant by running MSCs and to determine the length durations consider the following procedure for MSC-2.

To get the incremental length duration of MSC-2, start with the beginning of the observation period and continue until a MSC-2 is observed. The incremental length duration of the MSC-2 is the length duration of the MSC-2 minus the length duration of the first MSC-1 contained in the MSC-2. To identify the second running MSC-2, discard the first failure and assume the beginning of the observation period is the time of the first failure. Identify the second MSC-2. If the end point of the second MSC-2 is the same as the end point of the first MSC-2, discard the first failure of the second MSC-2 and repeat the process until the end point of MSC-2 is different. Calculate the incremental cycle duration as the second observation to be used for estimating the ξ_2 . The following sequence of failures can be used as an example to demonstrate the procedure.

Following the procedure described above, four distinct values for the incremental MSC-2 durations (denoted by W_1 , W_2 , W_3 , and W_4) can be identified and quantified that can be used to estimate or update the value ξ_2 .

The procedure for estimation is described in Appendix D.

APPENDIX B. DERIVATION OF SYSTEM UNAVAILABILITY FORMULA USING THE DISTRIBUTION OF THE MSC LENGTH DURATION

B.1. Introduction

To evaluate the safety implication of dependent failures in context of the clustering methodologies described in Appendix A, system unavailabilities are to be formulated in terms of the MSC length durations. It was shown in Appendix A that for a system composed of n redundant components, the probability that at least k out of n components to be found failed after a period " T " when all components were tested and made available at time zero is given by:

$$F_k(T) = \int_0^T f_k(\tau) d\tau \quad \text{B-1}$$

where $f_k(\tau)$ is the distribution of cycle duration of MSC- k , where under assumption of exponential distribution for incremental durations, can be expressed by:

$$f_k(\tau) = \sum_{m=1}^k A_m e^{-\xi_m \tau} \quad \text{B-2}$$

$$A_m = \xi_m \prod_{\substack{j=1 \\ j \neq m}}^k \xi_j / (\xi_j - \xi_m) \quad \text{B-3}$$

The incorporation of the above questions to evaluate the system unavailability are discussed in this appendix.

B.2 System Unavailability Under Sequential Testing Scheme

If a system is composed of n redundant components and the system is not operational if at least k out of n redundancy have failed, and under the sequential testing of all redundant components with an interval " T ", the system unavailability can be formally expressed by:

$$\bar{Q}_k = \frac{1}{T} \int_0^T f_k(\tau) \cdot (T - \tau) d\tau \quad \text{B-4}$$

where $\bar{Q}_k$ is the average system unavailability.

By performing the integration of Eqn. (B-4) and using Eqns. (B-2) and (B-3), the following formula can be obtained:

$$\bar{Q}_k = \sum_{j=1}^k (A_j | \xi_j) \cdot \frac{1}{\xi_j T} \left[\xi_j T - (1 - e^{-\xi_j T}) \right] \quad \text{B-5}$$

It is easy to show that Eqn. (B-5) is a general and accurate expression of system unavailability. Furthermore, for $\xi_j T \ll 1$ it was discussed in the same reference that under the independence assumption, the first order non-zero expansion terms of Eqn. (B-5) will yield the expression that is usually obtained using conventional methods for the case where the test interval is much shorter than mean time between failures.

The other contribution to system unavailability stems from the component maintenance. In order to incorporate the maintenance contribution in the context of clustering methodologies, we have to differentiate between the two different types of maintenance. These are:

Corrective maintenance: This is the maintenance performed after a component failure. In this type of maintenance, failure of the component is equivalent to the realization of a MSC of one. This is also equivalent to assuming ξ_1 to be a delta function at initiation of maintenance with the remaining ξ_j s unchanged. If the duration of maintenance is expected to be θ , the unavailability from maintenance contribution would be:

$$\bar{Q}_{k1}^M = \xi_1 \theta \sum_{j=2}^k (A_j^* | \xi_j) \cdot \frac{1}{\xi_j \theta} \left[\xi_j \theta - (1 - e^{-\xi_j \theta}) \right] \quad \text{B-6}$$

where A_j^* is calculated based on Eqn. (B-3) but with ξ_1 set to infinity. ξ_1 in Eqn. (B-6) is the same as before and equals " $n\lambda$ ". In Eqn. (B-6), the expected number of maintenance for a single component is assumed to be:

$$E(\#M_1) = \int_0^T f_1(\tau) d\tau = \text{probability of a component failure at test} \approx \xi_1 T$$

If the plant technical specifications allow simultaneous maintenance of more than one component, then the unavailability contribution of " m " components under maintenance would be:

$$\bar{Q}_{km}^M = \frac{E(\#M_m)}{T} \cdot \theta \sum_{j=m}^k (A_j^* | \xi_j) \frac{1}{\xi_j \theta} \left[\xi_j \theta - (1 - e^{-\xi_j \theta}) \right] \quad \text{B-7}$$

where

$$E(\#M_m) = \int_0^T f_m(\tau) d\tau \quad \text{B-8}$$

and finally, the overall maintenance contribution from single maintenance, double maintenance, etc., can be evaluated by:

$$\bar{Q}_k^M = \sum_{j=1}^m \bar{Q}_{km}^M \quad \text{B-9}$$

Preventive Maintenance (PM): This is the maintenance performed not as a result of a component failure. In this type of maintenance, no failure is realized, however, one component would be unavailable due to maintenance work being performed. If the duration of maintenance is θ , the unavailability contribution from preventive maintenance is simply:

$$\bar{Q}_{km}^{PM} = f_m \theta \sum_{j=1}^{k-m} (A_j | \xi_j) \cdot \frac{1}{\xi_j \theta} \left[\xi_j \theta - (1 - e^{-\xi_j \theta}) \right] \quad \text{B-10}$$

where "m" is the number of components under PM, F_m is the rate for the m-PMs per unit time, and

$\bar{Q}_{km}^{PM}$ is the average unavailability contribution from m component being under PM at the same time. As before, m is the maximum number of components allowed to be simultaneously down for maintenance (m can be 1, 2, etc.) by the plant technical specification.

The total system unavailability under the sequential testing scheme then can be calculated accounting for both undetected downtime as well as maintenance contributions. As a final note, PMs have many positive effects on reliability improvement which are directly measured by observed data through estimation of ξ_i s.

B.3 System Unavailability Under Staggered Testing Scheme

The evaluation of system unavailability under the staggered testing scheme and the estimation of minimal cutset probabilities generated by conventional risk assessment in context of clustering formulation is of vital importance for future implementation of these methodologies. There is a certain commonality between evaluation of staggered testing and minimal cutsets. In staggered testing at the time of test, a specific component would be tested. Therefore, it is important to know the probability of occurrence of MSC-k, which involves a specific component. In a similar manner, evaluation of a minimal cutset requires the occurrence of one or more specific component failures. Generally speaking, as a first step to address both staggered testing and minimal cutsets, we need to estimate the probability of occurrence for an MSC-k involving "m" specific components.

If the clustering analyses are performed for n redundant component, the distribution for the duration of an MSC-k was denoted by $f_k(\tau)$ and was formulated for all possible combinations of k out of n components. The number of combinations of k out of n components, which includes one specific component, "A", $C(k,n,1)$, is:

$$C(k,n,1) = \binom{n-1}{k-1} \quad \text{B-11}$$

Therefore, the unavailability of k out of n components, including a specific component, $\bar{Q}_{k,n,1}$, is:

$$\bar{Q}_{k,n,1} = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} \bar{Q}_{k,n} \quad \text{B-12}$$

Similarly, the unavailability of k out of n components, including two specific components, $\bar{Q}_{k,n,2}$, is;

$$\bar{Q}_{k,n,2} = \frac{\binom{n-2}{k-2}}{\binom{n}{k}} \bar{Q}_{k,n} \quad \text{B-13}$$

Equations similar to (B-12) and (B-13) would allow the incorporation of clustering methodologies evaluation into conventional PRAs. The detail of these formulations has not yet been completed and it is recommended as potential future work in this area.

APPENDIX C. MONITORING INDICATORS OF DEPENDENCY AND SYSTEM UNAVAILABILITY

C.1 Introduction

Appendices A and B detailed an approach and formulations for evaluating system unavailability accounting for clustering of failure events. The objective of this appendix is to provide a road map of how these methodologies can be used for monitoring dependent failures and their effect on system unavailability. To describe this road map we also construct a logical bridge between Appendices A and B, and D through G. The following subjects are discussed in this appendix:

1. Capability - How to assure that the data are compatible with assumptions of clustering methodology described in Appendices A and B.
2. Screening Analysis - How to develop statistical screening indicators that can detect significant clustering.
3. Unavailability Evaluation - How to include dependency in system unavailability indicators that can monitor significant changes in system unavailability.

C.2 Compatibility

There are three main assumptions in the methodologies described in Appendices A and B:

1. During the observation period all components exhibit the same failure time distribution,
2. The failure time distribution shall not depend on the period of observation, and
3. During the observation period each component failure time distribution is approximated by an exponential with a constant failure rate.

Assumption 1 basically indicates that attention should be paid to selecting a set of similar components. Some of these considerations are discussed in Appendix D. There are statistical tests available to confirm the homogeneity of the failure data.

Assumption 2 deals with the stationary nature of the process. For long observation periods, it is not unusual to detect changes in the failure rate of a set of components. In these cases, the observation period is to be segmented such that the stationary process can be assumed within each segment. The first step in determining if the process is stationary is through graphical representation of data. For the case of clustering analyses, the cycle duration of various MSCs can be graphed vs time of observations. Visual examination of these graphs would provide important trend information. In addition, formal CUSUM analyses of the observations on MSC durations are discussed in Appendix D. These types of analyses will provide information on time trends. This information can be used for segmenting the observation period. An example of this type of analysis for Plant 2 diesel generator data is discussed in Volume 1.

Finally, the third assumption regarding the failure time distribution, even though necessary for formulations given in Appendices A and B, generally can be eliminated. Both in Appendices D and E models such as mixture exponential and Cox-interspacing model have been developed and discussed. These models generally result in more extensive (complex) formulation of system unavailability beyond what is discussed in Appendices A and B. Both the statistical tests and estimation routine for cases where the failure time distribution is non-exponential have been developed and reported in Appendices D and E with supporting simulation studies in Appendix F.

C.3 Screening Analysis

The objective of the screening indicator is to detect significant clustering shown by data. Significant clustering is assumed to be detected when the observed duration of an MSC of code k has significantly deviated from the expected duration when the failures occur randomly and independent. To do this we define a measure $\hat{\alpha}_k$ defined by:

$$\hat{\alpha}_k = \frac{\hat{T}_k}{E(T_k)} \quad \text{C-1}$$

where $\hat{T}_k$ is the observed duration of the most recent MSC-k duration and $E(T_k)$ is the expected duration of MSC-k under assumption of independence. Such a measure can be used as an indicator for screening the clustering events based on MSC duration. Measure $\hat{\alpha}_k$ is expected to have large variations. To smooth the indicator, one approach is to use the average of $\hat{\alpha}_k$ within each segment of observation. The statistical properties of this average measure are detailed in Section D.2 of Appendix D with supporting simulation studies reported in Appendix F. Additional simulation studies are needed to address a sequential indicator based on overlapping segments of observations (moving window). Based on the preliminary simulation studies in Appendix F, it appears that the three sigma bounds reported in Appendix D may be appropriate for use in moving window averaging as well.

C.4 Unavailability Evaluation

In addition to screening indicators of clustering, an indicator of system unavailability can be constructed that include the effect of the observed dependency. The basic formulations of system unavailability within a segment of observation period were discussed in Appendix B. Appendix E discusses the more comprehensive system unavailability estimation routine. Using these equations, the system unavailability can be calculated for three cases:

1. PRA type approximate unavailability,
2. Exact system unavailability based on independence, and
3. Exact system unavailability accounting for dependence.

These types of calculations were presented in Volume 1 of this report. In fact, the partial beta factors were back calculated to show the contribution of failure clustering. These types of analyses may be performed based on moving segments of observation period and may be used as a system unavailability indicator. The tolerance bound for these indicators can be established based on the similar

methodologies previously reported in NUREG/CR-5626¹. Preliminary studies on the characteristics of these indicators were studied through computer simulations and reported in Appendix F.

C.5 References

1. Azarm, M.A. and Vesely, W.E., "System Unavailability Indicators - Vol. 1. Summary Report, Vol. 2. Detailed Analysis and Appendices," NUREG/CR-5626, BNL-NUREG-52255, November 1991.

APPENDIX D. STATISTICAL TEST FOR DETECTING CLUSTERING AND DEPENDENCY

D.1 Introduction

The distribution of the length duration for MSC-k was derived in Appendix A and expressed by:

$$f_k(W) = \sum_{m=1}^k A_m e^{-\xi_m W} \quad D-1$$

where

$$A_m = \xi_m \prod_{\substack{j=1 \\ j \neq m}}^k \xi_j / (\xi_j - \xi_m) \quad D-2$$

and ξ_j is the exponential rate for incremental duration from MSC-(J-1) to MSC-J.

It was also shown that under the assumption of independence for a set of similar components, ξ_j can be expressed by:

$$\xi_j = (n-j+1)\lambda \quad D-3$$

where λ is the failure rate of a single component and n is the number of similar components.

The values of ξ_j s can also be directly estimated from the observed data without any assumption regarding dependency.

The purpose of the statistical tests described in this appendix is:

1. To verify whether or not the observed data are not consistent with the independent assumption (test for independence).
2. To verify if the exponential model is sufficient to describe the behavior of the observed data on the incremental durations (test for exponentiality).
3. To verify if the data are stationary and do not manifest a trend. If the existence of a trend is verified, the data shall be divided into smaller sets to minimize the effect of the trends observed (trend test).

The following describes the preliminary formulation on the three types of tests described above.

D.2 Test of Independence

Under the assumption of independence, the expectation for the duration length of MSC-k, i.e., $E(T_k)$ for "n" similar components can be expressed by:

$$E(T_k) = \frac{1}{n\lambda} + \frac{1}{(n-1)\lambda} + \dots + \frac{1}{(n-k+1)\lambda} \quad D-4$$

where λ , as defined before, is the failure rate for a single component.

The ratio of expectation of T_k over the expectation of T_1 therefore can be expressed by:

$$f_k = \frac{E(T_k)}{E(T_1)} = 1 + \frac{n}{n-1} + \dots + \frac{n}{n-k+1} = k + \frac{1}{n-1} + \frac{2}{n-2} + \dots + \frac{k-1}{n-k+1} \quad D-5$$

An estimator of f_k , denoted by $\hat{f}_k$, from observed data can be calculated by:

$$\hat{f}_k = \frac{\bar{T}_k}{\bar{T}_1} \quad D-6$$

where, for M_k observations of MSC-k from data, $\bar{T}_k$ can be expressed by:

$$\bar{T}_k = \frac{1}{M_k} \sum_{j=1}^{M_k} T_{kj} \quad D-7$$

$$\bar{T}_1 = \frac{1}{M_1} \sum_{j=1}^{M_1} T_{1j} \quad D-8$$

where T_{kj} is the length duration of the jth nonoverlapped MSC-k.

If m_1 is large; we approximate T_1 by $T_1 \approx 1/n\lambda$ and neglect the correlation between T_1 and T_k , then we have:

$$\hat{f}_k = n\lambda \bar{T}_k$$

and therefore,

$$E(\hat{f}_k) = n\lambda E(\bar{T}_k) = 1 + \sum_{j=1}^{k-1} n/(n-j) \quad D-9$$

If we define a test measure α_k by Eqn. (D-10)

$$\alpha_k = \hat{f}_k / E(\hat{f}_k) \quad D-10$$

α_k will have the following statistical properties.

The expectation (mean) of α_k obviously has to be one. That is:

$$E(\alpha_k) = E(\hat{f}_k) / E(\hat{f}_k) = 1 \quad D-11$$

The variance of α_k can be expressed by:

$$\text{Var}(\alpha_k) = \text{Var} \frac{(\hat{f}_k)}{[E(\hat{f}_k)]^2} = \frac{n^2 \lambda^2 \text{Var}(\bar{T}_k)}{n^2 \lambda^2 [E(\bar{T}_k)]^2} \quad D-12$$

and

$$\text{Var}(\bar{T}_k) = \frac{\text{Var}(T_k)}{M_k} \quad D-13$$

since T_k is the summation of W_i s, and each w_i is exponentially distributed, then the variance of each term is the square of its mean value. Therefore, the var (T_k) can be expressed by:

$$\text{Var}(T_k) = \left(\frac{1}{\lambda^2} \right) \sum_{j=0}^{k-1} \frac{1}{(n-j)^2} \quad D-14$$

Therefore, $\text{Var}(\alpha_k)$ is expressed by:

$$\text{Var}(\alpha_k) = \frac{1}{M_k} \cdot \left[\sum_{j=0}^{k-1} \frac{1}{(n-j)^2} / \left(\sum_{j=0}^{k-1} \frac{1}{n-j} \right)^2 \right] \quad \text{D-15}$$

The standard deviation for α_k , denoted by S_α is:

$$\frac{1}{\sqrt{M_k}} \cdot \frac{\left[\sum_{j=0}^{k-1} \left(\frac{1}{n-j} \right)^2 \right]^{1/2}}{\sum_{j=0}^{k-1} \frac{1}{n-j}} \quad \text{D-16}$$

For large values of k , Eqn. (D-16) approaches to:

$$S_\alpha = \frac{1}{\sqrt{k} \cdot \sqrt{M_k}} \quad \text{D-17}$$

Due to the correlation among the values of α_k for various k s, the exact significance bounds cannot be easily determined unless through a detailed simulation study. However, in our analyses of dependency, if the test measure α_k exceeds 1.5 standard deviations from mean, the assumption of independence is assumed to be rejected.

D.3 Test of Exponentiality

The cycle incremental duration distributions were assumed to be exponential for deriving the system unavailability equation. The assumption of exponentiality by no means prevents the models developed here to address dependency. The clustering/dependency exhibited by data will result in exponential rates somewhat different from those predicted by the independence assumption, thereby accounting for dependency contribution. However, it is important to design a test(s) to verify whether or not the exponential assumption is supported by the observed data. Two types of tests were designed:

- a. a test for validity of single exponential, and
- b. a test for the Cox interspacing model.

D.3.1 Test for Validity of Single Exponential

This test is designed to verify if the observed data on incremental cycle duration can be best described by a single exponential in comparison with a mixture of two exponentials. The concern is for short-cycle durations (comparable to two or three test intervals), which may occur more frequently than

those predicted by a single exponential fit. The test assumes a change point at the 2nd, 3rd, or 4th test interval, and it uses binomial distribution limits to verify acceptability of single exponential fit at 90% confidence. The following describes the test procedure.

First, all the data on incremental cycle durations of MSC-k are fit to an exponential by using a maximum likelihood estimate. Then the estimated exponential parameter $\hat{\lambda}$ is used to determine the probability of observing a data point less than the change point limit "T", i.e.,:

$$P = 1 - e^{-\hat{\lambda}T} \quad \text{D-18}$$

If there are n cycle durations of MSC-k, the probability of observing m out of n incremental cycle durations of MSC-k less than T is then given by a binomial distribution:

$$P_{m,n}(T) = \binom{n}{m} P^m (1-P)^{n-m} \quad \text{D-19}$$

The test rejects the assumption of single exponential if:

$$\sum_{j=m,n}^n \binom{n}{j} P^j (1-P)^{n-j} \leq 0.1 \quad \text{D-20}$$

where $P_{m,n}(T)$ is the probability of observing m or more cycle durations of MSC-k in less than T.

If the test rejects the assumption of single exponential, then the mixture parameters will be estimated based on the procedure discussed in Appendix F.

D.3.2 Test for Cox Interspacing Model

Cox's interspacing dependency model¹ is one approach that has been developed to analyze failure rates and conditional failure rates in terms of the interspacing times between failures.

Consider two types of events - A events and B events. B events initiate dependencies in A events.

B **B** **A** **B** **B** **B** **A** **B**
x **x** **x** **x** **x** **x** **x** **x**

If B events occur much more frequently than A events and they can be treated as realizations of an exponential distribution with a fixed (non-statistical) parameter Λ_1 , the occurrence rate for an A event may be expressed by:

$$\Lambda(u) = \Lambda_0 e^{\gamma u}$$

D-21

where

Λ_0 is a reference constant failure rate,

γ is the dependency parameter, and

u is the time interval since the last B occurrence.

The form of Eqn. (D-21) indicates that:

- a. for $\gamma=0$, the occurrence of A occurs with a constant Λ independence from the last occurrence of B. This shall not be misconstrued as independence as we will be discussing shortly.
- b. for $\gamma<0$, A occurs with a higher probability (higher failure rate) in shorter duration after the occurrence of B (shorter values of u). This is equivalent to the concept of common-cause failure (or failure time dependency) discussed in this report.
- c. for $\gamma>0$, A occurs with a higher probability (higher failure rates) for larger values of u . This is true of the case where a failure triggers a comprehensive preventive maintenance program thereby increasing the survival probability of the remainder redundant components (opposite to clustering).

The above discussion shows that the Cox interspacing model can address the issue of dependency. However, it is important to emphasize the following:

1. If γ is assessed to be zero, it does not mean that the data are not clustered. As we discussed in the previous section, the constant failure rate does not mean no dependency unless the failure rate estimated is comparable with those predicted from the independence model. One may visualize that when a dependency shock arrives, all component failure rates increase by a factor, however, the failure of component A does not depend on the time that other components have failed.
2. If γ is not equal to zero, it certainly means that the failure rate of a component depends on the times of failures of other components.

Given the above interpretation of the Cox interspacing model, we did not formulate the test procedure for verifying if γ equals zero. To do that we define:

v_i = the interval from an A occurrence to the nearest preceding B occurrence.

u_i = the interval between B occurrences.

Log likelihood "L" of the times of the A event, ignoring the additional constant is:

$$L = \gamma \sum_{i=1}^n v_i - n \ell n \sum_{i=1}^m \int_0^{u_i} e^{-\gamma x} dx \quad D-22$$

where m and n stands for number of events of type B and A, respectively.

The score statistic $dL/d\gamma = U(\gamma)$ is then:

$$U(\gamma) = \sum_{i=1}^n v_i - n \frac{\sum_{i=1}^m \int_0^{u_i} x e^{-\gamma x} dx}{\sum_{i=1}^m \int_0^{u_i} e^{-\gamma x} dx} \quad D-23$$

This reduces to:

$$U(\gamma) = \sum_{i=1}^n v_i - \frac{n}{\gamma} \frac{\sum_{i=1}^m (\gamma u_i e^{-\gamma u_i} - e^{-\gamma u_i} + 1)}{\sum_{i=1}^m (e^{-\gamma u_i} - 1)} \quad D-24$$

The information statistic $d^2L/d\gamma^2 = I(\gamma)$ is:

$$I(\gamma) = n \left\{ \frac{\sum_{i=1}^m \int_0^{u_i} x^2 e^{-\gamma x} dx}{\sum_{i=1}^m \int_0^{u_i} e^{-\gamma x} dx} - \left(\frac{\sum_{i=1}^m \int_0^{u_i} x e^{-\gamma x} dx}{\sum_{i=1}^m \int_0^{u_i} e^{-\gamma x} dx} \right)^2 \right\} \quad D-25$$

We can obtain an evaluated expression for $I(\gamma)$ by using the following substitutions.

Let:

$$A(\gamma) = \frac{\sum_{i=1}^m \int_0^{u_i} x e^{\gamma x} dx}{\sum_{i=1}^m \int_0^{u_i} e^{\gamma x} dx} \quad \text{D-26}$$

Then:

$$A(\gamma) = \frac{\frac{1}{\gamma} \sum_{i=1}^m (\gamma u_i e^{\gamma u_i} - e^{\gamma u_i} + 1)}{\sum_{i=1}^m (e^{\gamma u_i} - 1)} \quad \text{D-27}$$

Then the score statistic is given by:

$$U(\gamma) = \sum_{i=1}^N v_i - nA(\gamma) \quad \text{D-28}$$

Also, let:

$$B(\gamma) = \frac{\sum_{i=1}^m \int_0^{u_i} x^2 e^{\gamma x} dx}{\sum_{i=1}^m \int_0^{u_i} e^{\gamma x} dx} \quad \text{D-29}$$

Then:

$$B(\gamma) = \frac{\frac{1}{\gamma^2} \sum_{i=1}^m (\gamma^2 u_i^2 e^{\gamma u_i} - 2\gamma u_i e^{\gamma u_i} + 2e^{\gamma u_i} - 2)}{\sum_{i=1}^m (e^{\gamma u_i} - 1)} \quad \text{D-30}$$

The information statistic is then given by:

$$I(\gamma) = n(B(\gamma) - A^2(\gamma)) \quad \text{D-31}$$

The score function $U(\gamma)$ can be used to obtain the maximum likelihood estimate (mle) of γ by setting $U(\gamma)=0$. The inverse of the information statistic evaluated at the mle is an approximate estimate of the variance of the mle of γ .

A hypothesis test for $\gamma=\gamma_0$ also can be carried out by forming the ratio z , where:

$$z = \frac{U(\gamma_0)}{\sqrt{I(\gamma_0)}} \quad \text{D-32}$$

and comparing z to the percentiles of a standard normal distribution.

To specifically test the hypothesis that $\gamma=0$, the terms reduce to:

$$A(0) = \frac{\sum_{i=1}^m \int_0^{u_i} x \, dx}{\sum_{i=1}^m \int_0^{u_i} dx} = \frac{\frac{1}{2} \sum_{i=1}^m u_i^2}{T} \quad \text{D-33}$$

where $T = \sum_{i=1}^m u_i$ = the total observation interval

Also,

$$B(0) = \frac{\sum_{i=1}^m \int_0^{u_i} x^2 dx}{\sum_{i=1}^m \int_0^{u_i} dx} = \frac{\frac{1}{3} \sum_{i=1}^m u_i^3}{T} \quad \text{D-34}$$

Hence, for the hypothesis test of $\gamma=0$,

$$U(0) = \sum_{i=1}^n v_i - nA(0) \quad \text{D-35}$$

or

$$U(0) = \sum_{i=1}^n v_i - \frac{\frac{n}{2} \sum_{i=1}^m u_i^2}{T} \quad D-36$$

and

$$I(0) = n \left(\frac{\frac{1}{3} \sum_{i=1}^m u_i^3}{T} - \left(\frac{\frac{1}{2} \sum_{i=1}^m u_i^2}{T} \right)^2 \right) \quad D-37$$

Therefore, we calculate:

$$z = \frac{U(0)}{\sqrt{I(0)}} \quad D-38$$

and compare to the percentages of a standard normal.

The Cox interspacing model formula above can be used to determine dependency by assuming the failure of one of the redundant components as events of type A and the failure times of the remainder redundant components can be treated as events of type B. The procedure can be repeated for each of the redundant components.

D.4 Trend Test

As discussed previously, the trend test is designed to verify if the observations on cycle and the incremental cycle durations are stationary. If the data manifests a time varying behavior, a change point in time may be qualitatively identified and data may be divided into smaller segments to reduce the effect of the observed trend.

The CUSUM²⁻³ procedure was used in this study for detecting time trend.

If we consider the incremental cycle durations as the observations, Z_i , implementing the CUSUM scheme requires two cumulative sums, S_H and S_L , where at the i th observation:

$$S_{Hi} = \max \left[0, (Z_i - k) + S_{H(i-1)} \right] \quad D-39$$

$$S_{Li} = \max \left[0, (-Z_i - k) + S_{L(i-1)} \right] \quad D-40$$

If either S_H or S_L becomes greater than the decision-interval value h , the process is considered to be out of statistical control. A standard CUSUM has $S_{H0} = S_{L0} = 0$, while a fast initial response (FIR) CUSUM sets $S_{H0} = S_{L0}$ to a common non-zero value (S_0).

The properties of the CUSUM control scheme are determined by the value of h , k , and S_0 . The parameter k is determined by the shift in mean level, which the CUSUM control scheme is designed to detect. The values of h , k , and S_0 can be selected based on desired values of average run lengths (ARL). ARL is the expected number of observations for a given performance to alarm CUSUM. For example, when observations are generated from a normal distribution with a mean of zero and a standard deviation of one, the following two types of average run length may be of concern.

1. ARL when the observations are generated with no displacement in the mean of the normal distribution (for our example of normal distribution, the observed mean is zero). CUSUM shall be designed such that ARL for this case is as long as practical.
2. ARL when the observations are generated with some displacement in the mean of normal distribution. As an example, the ARL for the mean shift of 2 sigma is detected in a short period as practical.

For additional information the reader may refer to Reference 3 (pgs, 201 & 202, Tables 2&3) for detailed tabulations for values of S_H and S_L .

To apply the CUSUM technique to the clustering analyses, the following procedure was adopted.

The observed cycle incremental duration of MSC- k , w'_k , were multiplied by $(n-k+1)/n$. The expectation of the resulting value, w'_k , will be the same as the observed expectation of time between failures denoted by w , which is directly calculated from all observed failure times. The following transformation was then applied to the w'_k to obtain a new variable, Z_k , which is approximately distributed by a unit normal.

$$Z_k = \frac{\text{Log}(w'_k) - \mu}{\sigma} \quad \text{D-41}$$

where μ and σ are the expectation and the variance of $\log(w)$.

The variable Z_k was then used in the CUSUM procedure and the parameters S_0 , h , and k are specified by the user.

References

1. "Stochastic Point Processes," Lewis, P.A.W., John Wiley & Sons, Inc., 1972, paper by D.R. Cox, "The Statistical Analysis of Dependencies in Point Processes," pp. 55-67.
2. "CUSUM vs Smoothed-Error Forecast Monitoring Schemes: Some Simulation Results," Everette S. Gardner Jr., Journal Operations Res., Vol. 36, pp. 43-7, 1985.
3. "Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Head Start," James M. Lucas and Ronald B. Crosier, Technometrics, Vol. 24, pp. 199-205, August 1982.

APPENDIX E. ESTIMATION PROCEDURE FOR DETERMINING AVERAGE SYSTEM UNAVAILABILITY

E.1 Introduction

The system unavailability formula accounting for clustering was described in Appendix B. This formula contains parameters ξ_j which are to be estimated from data. The observations on incremental cycle durations can be used to estimate these parameters. Based on the statistical approach discussed in Appendix D, the incremental cycle durations may be arrived from a:

1. single exponential with parameter ξ_i ,
2. mixture of two exponentials with parameters ξ_i , ξ_i^* , which are the parameters for the two exponentials, and p_i is the mixture ratio or probability, and
3. Cox-interspacing model, where parameter ξ_i is a function of u , the time since the last failure of other components, by:

$$\xi_i(u) = \xi_i e^{\gamma_i u} \quad \text{E-1}$$

This appendix will describe the approach currently used in this study for estimating these parameters. Three cases are described, single exponential, a mixture of two exponentials, and the Cox-interspacing model.

E.2 Estimation of a Single Exponential Parameter

In this case, the statistical tests described in Appendix D indicated that the incremental cycle duration of interest can be described by a single exponential. If the "m" observations on incremental cycle durations are denoted by W_j s, the maximum likelihood estimate for ξ is given:

$$\hat{\xi} = \left(\sum_{j=1}^m W_j \right) / m \quad \text{E-2}$$

The classical confidence bounds on ξ also can be determined using the conventional chi-squared bounds.

The Bayes approach also can be used to estimate the posterior distribution of ξ using the single-stage Bayes method. To arrive at a prior distribution for the Bayes process, we propose the following approach:

- The prior distribution for the incremental cycle duration of MSC-k shall be determined based on the posterior distribution of the incremental cycle duration of MSC-(k-1) by use of scaling factor " f_k ", defined by:

$$f_k = (n-k+2)/(n-k+1)$$

E-3

where n is the total number of similar components. Therefore, the mean of the prior distribution for the incremental cycle duration of MSC- k , denoted by $E_p(\tau_k)$ in terms of the mean of the posterior distribution for incremental cycle duration of MSC- $(k-1)$, denoted by $E^p(\tau_{k-1})$, is:

$$E_p(\tau_k) = f_k \cdot E^p(\tau_{k-1})$$

E-4

Since the mean of the incremental cycle duration is the reciprocal of the parameter ξ_k , the prior distribution for ξ_k can be directly determined.

Once the parameters of ξ_k are determined, the system unavailability can be estimated through the formulation given in Appendix B.

E.3 Estimation of Parameters for Mixture of Two Exponentials

In this case, the incremental cycle duration of MSC- k is assumed to be distributed by a mixture of two exponential distributions based on the test procedure given in Appendix D and the following formula:

$$f(\tau_k) = P_k \cdot \xi_k \exp(-\xi_k \tau_k) + (1-P_k) \xi_k^* \exp(-\xi_k^* \tau_k)$$

E-5

If we write the 0th, 1st, 2nd, 3rd moments about zero as:

$$1 = P_k + (1-P_k)$$

E-6

$$\bar{\tau}_k = (P_k/\xi_k) + (1-P_k)/\xi_k^*$$

E-7

$$\frac{m_2}{2} = P_k (1/\xi_k)^2 + (1-P_k) (1/\xi_k^*)^2$$

E-8

$$\frac{m_3}{6} = P_k (1/\xi_k)^3 + (1-P_k) (1/\xi_k^*)^3$$

E-9

and for convenience, we set:

$$1 = T_0, \quad \bar{\tau}_k = T_1, \quad \frac{m_2}{2} = T_2, \quad \frac{m_3}{6} = T_3$$

$$\mu_1 = 1/\xi_k, \quad \mu_2 = 1/\xi_k^*$$

If we multiply Eqn. E-6 by some constant α_2 , Eqn. E-7 by some constant α_1 , and Eqn. E-8 by (-1), and add them up, we get:

$$P_k(-\mu_1^2 + \alpha_1\mu_1 + \alpha_2) + (1-P_k)(-\mu_2^2 + \alpha_1\mu_2 + \alpha_2) = \alpha_2 T_0 + \alpha_1 T_1 - T_2 \quad \text{E-10}$$

Similarly, if we multiply Eqn. E-7 by α_2 , Eqn. E-8 by α_1 , and Eqn. E-9 by -1, and add, we get:

$$P_k(-\mu_1^3 + \alpha_1\mu_1^2 + \alpha_2\mu_1) + (1-P_k)(-\mu_2^3 + \alpha_1\mu_2^2 + \alpha_2\mu_2) = \alpha_2 T_1 + \alpha_1 T_2 - T_3 \quad \text{E-11}$$

Thus, if we choose α_1 and α_2 such that μ_1 and μ_2 are the root of

$$\mu^2 - \alpha_1 \mu - \alpha_2 = 0 \quad \text{E-12}$$

we get:

$$\alpha_2 T_0 + \alpha_1 T_1 = T_2 \quad \text{E-13}$$

$$\alpha_2 T_1 + \alpha_1 T_2 = T_3 \quad \text{E-14}$$

from which:

$$\begin{cases} \alpha_1 = (T_1 T_2 - T_0 T_3) / (T_1^2 - T_0 T_2) \\ \alpha_2 = (T_1 T_3 - T_2^2) / (T_1^2 - T_0 T_2) \end{cases} \quad \text{E-15}$$

Thus, we estimate the values of T_1 , T_2 , and T_3 from observed data on incremental cycle durations, which yields estimates for α_1 and α_2 , which can be used to estimate the parameters ξ_k and ξ_k^* from:

$$\xi_k, \xi_k^* = \left[\frac{\alpha_1}{2} \mp \frac{1}{2} (\alpha_1^2 + 4\alpha_2)^{1/2} \right]^{-1} \quad \text{E-16}$$

from ξ_k and ξ_k^* and Eqn. E-7, an estimate for P_k will be obtained. For parameters ξ_k and ξ_k^* to be positive and real, certain conditions must be satisfied for T_i s. These are discussed in detail in Reference 1.

The above procedure provides some estimates of the parameters involved; however, it does not yield any bounds. Bayesian approaches can be used for developing posterior distributions for the parameters of the mixture exponentials; however, the needed methodologies and the software have not been developed in this study.

When one or more incremental cycle duration is described by a mixture of two exponentials, the expression for system unavailability becomes more extensive. This is qualitatively discussed below:

If only one of the incremental cycle durations is expressed by a mixture of two exponentials representing one exponential for short incremental duration and one for long incremental duration, then the system unavailability can be expressed by:

$$Q_k = P_s Q_k^s + (1 - P_s) Q_k^L \quad \text{E-17}$$

where Q_k^s and Q_k^L is the same formulation as before when the short and long incremental duration distribution are used, respectively.

If two of the incremental cycle durations are expressed by a mixture of two exponentials, then Eqn. E-17 will have two additional terms accounting for short-long and long-short contributions. If the number of incremental cycle durations that are described by a mixture of two exponentials is " ℓ ", then we end up with $2^{\ell-1}$ additional terms, which will be very cumbersome to evaluate.

Due to the above complication, we use the following approximate bound for system unavailability calculations:

$$P_L Q_k^s + (1 - P_L) Q_k^L \leq Q_k \leq P_u Q_k^s + (1 - P_u) Q_k^L \quad \text{E-18}$$

where, for the case with ℓ mixture distributions, P_L and P_u are expressed by:

$$P_L = \prod_{j=1}^{\ell} P_j \quad \text{E-19}$$

$$P_u = 1 - \prod_{j=1}^{\ell} (1 - P_j) \quad \text{E-20}$$

Q_k^s and Q_k^L are the k out of n component unavailability when all " ℓ " incremental durations are assumed to come from short and long exponentials, respectively. Based on our preliminary analyses, the upper bound appears to be a better approximation of the system unavailability than the lower bound and may be used for most practical purposes.

E.4 Estimation of Parameters for Cox-Interspacing Model

The Cox-interspacing model was detailed in Appendix D (Section D.3.2). There are two parameters associated with the Cox-interspacing model, Λ_0 and γ . The classical estimate for parameter γ and its associated bounds was given in Eqn. D-32 of Appendix D and is duplicated here:

$$z = U \frac{(\gamma_0)}{\sqrt{I(\gamma_0)}} \quad \text{E-21}$$

where z is the standard normal variable. By varying γ_0 and comparing z to the percentiles of a standard normal distribution, we are able to arrive at the bounds for β (e.g., 90 percent confidence bounds). The parameter Λ_0 for the applicable purpose described in Appendix D would simply be the same as the combined failure rate for n similar components. Both the classical and Bayesian estimation processes for this case were discussed previously.

The system unavailability equation, based on Cox-interspacing, has not been developed in this study and is considered to be beyond the scope of the present program.

E.5 References

1. Everitt, B.S. and Hand, D.J., "Finite Mixture Distributions," Chapman and Hall Ltd., 1981, pp. 60-79.

APPENDIX F. EVALUATION OF FAILURE DEPENDENCY THROUGH SIMULATION STUDIES

As discussed in the main report, clustering of failures is caused by a conditioning/triggering event which increases the local failure rate among similar components. To study the effects of triggering events on the clustering of failures, and to evaluate the applicability of the methods developed in this study, a simulation approach was used. Therefore, the simulation studies also were focussed to evaluate the effect of failure clustering on the system unavailability on the plant risk.

F.1 Methods Used in the Simulation Study

One of the main objectives of the simulation studies was to generate times of component failures when the component failure rate changes locally as a result of a conditioning event. The simulated failure times were then used to evaluate the effect of clustering on system unavailabilities using the methodologies described here. More particularly, the Monte-Carlo simulation and computer modeling techniques were used to achieve the following fundamental objectives:

1. To simulate the conditioning event which results in a simultaneous change of the failure rates of a set of similar components.
2. To generate the times of failures before and after the conditioning events to evaluate and compare the impact of multiple changes of failure rate on the system unavailability.
3. To show that system unavailability of the conventional PRA type calculations perform poorly in approximating the actual system unavailabilities and that the resulting risk implications compare to the proposed approaches in this report.
4. To show how effective plant maintenance practices would be in the early termination of the local failure rate changes by initiating preventive maintenance based on clustering of times of failures.

A computer code, COMCUS, was developed using conventional Monte-Carlo techniques to generate times of failures corresponding to the local changes in failure rates. Simulations were performed for a set of 2 and 4 similar components to show how the effect of clustering is more pronounced as more components are affected. Throughout the simulation studies, the repair was assumed to take zero time (instantaneous repair).

In a system consisting of several similar components, if a single, unexpected shock event (or conditioning event) occurs, we should expect to see an immediate jump in the failure rates of all the components that were exposed to the conditioning event. After the jump in the failure rate, the time to component failure will become shorter. When a component fails, it is safe to assume that the repair will eliminate the effect of the conditioning event, thereby reducing the failure rate of the repaired component back to its value before the occurrence of the shock. When another component fails, the repair process again eliminates the effect of the conditioning event on that component. The impact of a shock, therefore, will be totally eliminated when every component in the set has at least been repaired once. This period is referred to as the impact duration of a conditioning event. The contribution of a conditioning event to a system unavailability is a strong function of this duration. There are various maintenance policies that can directly reduce the impact duration of a conditioning event, two of which are described here. In the first policy, the preventive maintenance or repair is assumed to be initiated

for all the components in the set, if two failures are observed less than a test interval apart. Therefore, the possibility exists that the effect of a shock be totally eliminated before at least one failure and repair of all the components in the set. In the second policy, the preventive maintenance or repair for all similar components may be initiated whenever there is failure of more than one component as a result of similar causes, but not necessarily the failure observed within a test interval. The effect of all of these policies in terms of reducing the clustering of failure times and system unavailability can be evaluated through sensitivity analyses of the impact duration of conditioning events. We performed several simulation runs for sensitivity studies, which will be discussed shortly. This by no means shall be considered as a complete set of simulation studies that can determine the statistical characteristics of the unavailability measures and clustering indicators discussed in Appendix C.

Finally, some triggering events may cause or accelerate the degradation of all the components in the set resulting in exponential time dependent increases in failure rates. This type of change in failure rate was to be modeled similarly to that described before except that the increase in failure rate during the local changes was modeled as $\lambda \cdot e^{at}$. Here t is not the running time, but is measured from the start of the change in local failure rate.

F.2 Summary Description of the COMCUS Code

The computer code (called the COMCUS Code) developed for this simulation study was written in Microsoft Fortran and designed for users in generating the random samples of dependent failure time data for a set of similar components. The COMCUS Code is written for PCs and can simulate various types of conditioning events, which result in the changes of the failure rates for multiple components. This provides the analysts with a tool for studying various types of failure dependencies as well as their impact on system unavailability and risk.

The program can generate a large number of samples based on specific information supplied by the user describing the desired failure time clustering or failure rate changes to be analyzed. These simulated data form the vectors of observations that are needed to calculate the clustering indicators and system unavailability monitors discussed in previous appendices.

The COMCUS Code simulates a variety of failure time-dependencies within a conventional Monte-Carlo method, and provides information on actual system unavailabilities that can be used for comparison with the indicators or other estimators.

The COMCUS Program allows the user to simulate the following three typical types of failure rate changes:

1. Multiple failure rate changes caused by a single conditioning event (i.e., one time failure rate increase),
2. Periodic failure rate changes by defining a non zero impact duration for the conditioning event, and
3. Failure rate changes with exponential time-dependent failure rates.

The input parameters required by the COMCUS Code are illustrated in Table 1, which shows an example of the simulation to generate 100 random samples of dependent failure times.

Table 1. Parameter Setup for Generating 100 Random Samples of Dependent Failure Times for 7 Years

1.	Total time period you wish to simulate? (# of months)	-	84
2.	Do you wish to see details for each component failure? Y/N	-	Y
3.	Enter a random seed? - preferably a 6 digit number	-	153642
4.	Please enter your test data file name?	-	JK1.Dat
5.	Input a file name for failure time history?	-	JK3.Dat
6.	Please specify the failure rate changes for Train 1: • for a piecewise increase failure rate ---- type P • for an exponential increase failure rate ---- type E • for a linear increase failure rate $\lambda(t)=2+\beta t$ --- type L	-	P
7.	Please specify the piecewise failure rate jump: • trigger event frequency? (i.e., once every # of months?) • time length impacted by each triggering event - # of months?)	- -	84 0
8.	What is your lambda for the normal time period (#/month?)	-	0.1
9.	What is your lambda for abnormal time period (#/month?)	-	0.0

F.3 Case Runs and Results

To evaluate the accuracy of the underlying dependent failure models, and particularly, to examine the system unavailabilities produced by various types of failure rate dependencies, we have selected six different case runs using the COMCUS Code. Even though the simulation studies are limited, the preliminary results obtained appear to demonstrate that the underlying failure dependent models are sound and the various related statistical test procedures are promising. The unavailability formula derived here provides a greater accuracy in estimating the average system unavailability under both independence and failure clustering conditions as compared to conventional PRA calculations.

The underlying simulation case studies using the COMCUS Code were carried out in the following six different case runs, in which Cases 1 through Case 3 were designed for a two train system and Cases 4 through Case 6 were configured for a four-train system.

Case 1 Total Independence and Randomness for a Two Component System

We can consider a two-component system operating under no impact of any conditioning events; therefore, the failure rate of each component was at a constant level with no jumps throughout the life of the component. An average failure rate of 0.1 per month per component was used in the COMCUS Code to generate samples of failure times for the two components for seven, ten, and fifteen years, respectively. The simulation process was repeated using different random seeds to obtain a database consisting up to 100 samples. Tables 2a through Table 2c each represent a single sample of the data on failure time generated by the simulation code.

Table 2a. Sample Data for Case 1 - Total Time - 7 Years

Failure Time (Month)	Component 1	Component 2
1.759	0	1*
4.411	1	0
6.695	1	0
7.230	1	0
9.886	0	1
20.803	0	1
30.039	1	0
35.022	1	0
37.441	0	1
40.718	0	1
44.833	1	0
54.160	0	1
57.279	1	0
69.032	0	1
70.468	0	1
79.330	0	1
84.071	1	0
*1 under component means the component is failed.		

Table 2b. Sample Data for Case 1 - Total Time - 10 Years

Failure Time (Month)	Component 1	Component 2
1.309	0	1
3.220	0	1
11.803	1	0
28.441	1	0
31.718	1	0
41.379	0	1
45.160	1	0
55.055	0	1
60.032	1	0
61.468	1	0
70.330	1	0
73.937	0	1
76.463	0	1
85.980	0	1
87.079	1	0
87.792	0	1
90.850	0	1
92.158	0	1
93.057	1	0
94.801	1	0
113.983	1	0
116.364	0	1
117.134	0	1
119.743	1	0

Table 2c. Sample Data for Case 1 - Total Time - 15 Years

Failure Time (Month)	Component 1	Component 2
5.022	0	1
14.833	0	1
16.544	1	0
19.892	1	0
20.303	1	0
27.279	0	1
34.120	1	0
54.071	0	1
55.759	0	1
63.886	0	1
74.803	0	1
77.671	1	0
79.126	1	0
84.571	1	0
86.938	1	0
88.767	1	0
90.005	1	0
91.441	0	1
93.499	1	0
94.718	0	1
95.382	1	0
96.434	1	0
107.348	1	0
108.160	0	1
123.032	0	1
124.468	0	1

Table 2c. Sample Data for Case 1 - Total Time Period 15 Years (Cont'd)

Failure Time (Month)	Component 1	Component 2
133.330	0	1
143.860	1	0
149.522	1	0
150.079	0	1
155.823	1	0
156.057	0	1
157.801	0	1
166.788	1	0
172.494	1	0
176.411	1	0
176.983	0	1
178.695	1	0
179.230	1	0

Using the simulated sample data, the underlying dependent failure model was used to estimate the system average unavailability, which was then compared with the conventional PRA estimates. These estimates refer to the PRA's estimation of system unavailability based on the average component unavailability and no common cause contribution (refer to Eqn. (23) of Volume 1 of this paper). The results clearly show that under total independence, the PRA estimates differ from the clustering formulation reported here by about 1.5, which results from the approximation usually made in PRAs (e.g., using $(1/4) \cdot (\lambda T)^2$ instead of $(1/3) \cdot (\lambda T)^2$). Table 3 shows the results of this simulation study.

The column headings of Table 3 are as follows:

Year: Period of Observation in Years.

Exp: The percentage of times where the single exponential was rejected (refer to Appendix D).

Cycle: The percentage of time where the cycle test rejected independence assumption (refer to Appendix D).

Cox: The percentage of time where the Cox-interspace model rather than simple exponential model was accepted (refer to Appendix D).

- Cusum: The percentage of the time Cusum indicated a failure trend (refer to Appendix D)
- Q_{PRA} : The expected average system unavailability calculated by PRA formula for 100 samples.
- Q_m : The expected average system unavailability calculated by system unavailability equations derived in this study based on 100 samples and not accounting for dependency (see Eqn. (25) and the associated discussion in Volume 1).
- Q_e : The expected average system unavailability calculated by system unavailability equations derived in this study based on 100 samples and accounting for dependency (see Eqn. (25) and the associated discussion in Volume 1).

The distribution of the unavailability measures accounting for sample to sample variability can be developed here, which can be used to calibrate the system unavailability indicators described in Appendix C.

Table 3. Case 1 - Total Independency and Randomness (Two Components)

Years	Exp.	Cycle	Cox	Cusum	Q_{PRA}	Q_m	Q_e
7	0.12	0.14	0.17	0.16	2E-3	3.0E-3	3.0E-3
10	0.15	0.34	0.1	0.25	2E-3	3.0E-3	3.0E-3
15	0.4	0.35	0.19	0.42	2E-3	2.9E-3	2.9E-3

Case 2 Two Components with Failure Dependency - a Sparse Clustering of Failure Times

In this case, an average failure rate of 0.1 per month per component was used in the simulation and a relatively sparse clustering of failure times was generated at 0.01 per month for 20 months, which then jumps to 0.6 per month for about 4 months. This type of jump in local failure rate was repeated periodically throughout the life of the component. Tables 4a through 4c give sample data for failure time generated for periods of seven, ten, and fifteen years. Table 5 shows the estimated results. The column headings are the same as those described for Table 3. For a two-train system with clustering, the estimated unavailabilities in PRA are 2 under the estimate, which is about a β factor of 0.05.

Table 4a. Sample Data for Case 2 - 7 Years

Failure Time (Month)	Component 1	Component 2
20.818	1	0
22.000	1	0
22.000	0	1
23.000	0	1
42.000	0	1
44.517	0	1
45.000	1	0
46.000	1	0
47.000	1	0
47.000	0	1
48.096	1	0
62.000	1	0
68.855	1	0
69.000	0	1
70.000	0	1
71.000	1	0
76.000	0	1

Table 4b. Sample Data for Case 2 - 10 Years

Failure Time (Month)	Component 1	Component 2
21.000	0	1
22.000	1	0
22.000	0	1
23.000	1	0
23.000	0	1
45.000	1	0
47.000	1	0
52.000	0	1
68.082	0	1
69.000	1	0
70.000	0	1
71.000	0	1
72.123	1	0
92.919	1	0
93.000	0	1
94.000	1	0
94.000	0	1
96.131	0	1
107.000	0	1
116.842	0	1
117.000	0	1
118.000	1	0
118.000	0	1
119.000	0	1

Table 4c. Sample Data for Case 2 - 15 Years

Failure Time (Month)	Component 1	Component 2
20.972	0	1
21.000	1	0
21.000	0	1
22.000	0	1
23.000	1	0
23.000	0	1
24.201	0	1
43.000	1	0
46.000	1	0
47.000	1	0
53.000	0	1
69.000	0	1
71.000	1	0
71.000	0	1
72.160	0	1
92.453	1	0
92.807	0	1
93.000	1	0
94.000	1	0
94.000	0	1
95.000	1	0
96.721	1	0
100.000	0	1
118.000	1	0
118.000	0	1
119.000	0	1

Table 4c. Sample Data for Case 2 - 15 Year Period (Cont'd)

Failure Time (Month)	Component 1	Component 2
140.459	0	1
141.000	0	1
142.000	1	0
143.000	1	0
164.273	1	0
165.000	0	1
167.000	0	1
168.470	1	0
188.850	1	0

Table 5. Case 2 - Two Components with Failure Dependency (Sparse Clustering Case)

Years	Exp.	Cycle	Cox	Cusum	Q_{PRA}	Q_m	Q_e
7	0.67	0.51	0.3	0.03	2.0E-3	2.8E-3	4.2E-3
10	0.88	0.63	0.23	0.23	2.0E-3	2.8E-3	3.9E-3
15	0.95	0.87	0.29	0.29	1.9E-3	2.4E-3	3.6E-3

Case 3 Two Components with Failure Dependency (a Frequent Clustering of Failure Times)

This trial was basically the same as described in Case 2 except a more frequent clustering of failure dependency (or more frequent of periodic failure rate jumps) was imposed on generating the failure time data, in which an average failure rate was still about 0.1 per month per component, but the failure rate jumps were at 0.02 per month for 10 months and then increases to 0.5 per month for 2 months. Again, the estimated results showed that the unavailabilities obtained in PRA were 2 under estimate compared with the system unavailability formulation accounting for clustering failures. Tables 6a through 6c show the sample data for each of the 3 periods of simulation. The simulation results similar to previous cases are give in Table 7.

Table 6a. Sample Data for Case 3 - 7 Years

Failure Time (Month)	Component 1	Component 2
1.000	0	1
10.508	1	0
11.000	1	0
22.731	1	0
24.014	1	0
24.448	0	1
35.000	1	0
36.206	1	0
41.000	0	1
47.000	0	1
48.843	0	1
59.000	1	0
60.103	1	0
71.000	0	1
72.016	1	0
80.000	1	0

Table 6b. Sample Data for Case 3 - 10 Years

Failure Time (Month)	Component 1	Component 2
1.000	1	0
6.000	0	1
10.736	1	0
12.174	0	1
23.000	0	1
24.216	0	1
35.000	1	0
44.000	1	0
46.727	0	1
47.000	1	0
48.777	1	0
57.000	0	1
65.000	0	1
77.000	1	0
82.846	0	1
83.000	1	0
84.080	0	1
96.086	0	1
99.000	0	1
106.652	1	0
107.000	1	0
108.596	0	1

Table 6c. Sample Data for Case 3 - 15 Years

Failure Time (Month)	Component 1	Component 2
5.000	0	1
10.679	1	0
11.000	0	1
23.000	1	0
35.000	1	0
35.000	0	1
44.000	0	1
47.000	1	0
47.000	0	1
59.000	1	0
70.432	1	0
71.000	0	1
82.359	0	1
82.660	1	0
83.000	1	0
83.000	0	1
92.000	0	1
94.852	0	1
95.000	0	1
108.027	1	0
119.000	1	0
120.821	0	1
126.000	1	0
131.000	1	0
131.000	0	1
132.454	1	0

Table 6c. Sample Data for Case 3 - 15 Year Period (Cont'd)

Failure Time (Month)	Component 1	Component 2
142.603	1	1
143.000	0	1
144.196	0	1
154.895	0	1
155.000	1	0
166.292	0	1
177.000	0	1

Table 7. Case 3 - Two Components with Failure Dependency (Frequent Clustering Case)

Years	Exp.	Cycle	Cox	Cusum	Q_{PRA}	Q_m	Q_e
7	0.37	0.31	0.07	0.07	2.3E-3	3.3E-3	4.1E-3
10	0.64	0.49	0.16	0.16	2.4E-3	3.5E-3	4.8E-3
15	0.86	0.62	0.25	0.25	2.2E-3	3.2E-3	4.1E-3

Case 4 Total Independence and Randomness for a Four Component System

This case is similar to Case 1 except that 4 components were used for the simulation. Table 8 shows a single sample data generated by the COMCUS Code. An average failure rate (with no changes of 0.1 per month per component) was imposed in the study, and the results (Table 9) show a significant difference between the estimates from conventional PRA calculations and the clustering formulation of system unavailability proposed here. This difference is mainly due to the approximate nature of the PRA formulations.

Table 8. A Single Sample Data Generated in Case 4

Failure Time (Month)	Component			
	1	2	3	4
1.309	0	0	0	1
3.220	0	0	0	1
6.522	0	1	0	0
11.803	0	0	1	0
12.823	0	1	0	0
16.544	1	0	0	0
19.892	1	0	0	0
20.303	1	0	0	0
23.788	0	1	0	0
28.441	0	0	1	0
29.494	0	1	0	0
31.718	0	0	1	0
33.411	0	1	0	0
34.120	1	0	0	0
35.695	0	1	0	0
36.230	0	1	0	0
41.379	0	0	0	1
45.160	0	0	1	0
55.055	0	0	0	1
59.039	0	1	0	0

Table 8. A Single Sample Data Generated in Case 4

Failure Time (Month)	Component			
	1	2	3	4
60.032	0	0	1	0
61.468	0	0	1	0
64.022	0	1	0	0
70.330	0	0	1	0
73.833	0	1	0	0
73.937	0	0	0	1
76.463	0	0	0	1
77.671	1	0	0	0
79.126	1	0	0	0
84.571	1	0	0	0
85.980	0	0	0	1
86.279	0	1	0	0
86.938	1	0	0	0
87.079	0	0	1	0
87.792	0	0	0	1
88.767	1	0	0	0
90.005	1	0	0	0
90.850	0	0	0	1
92.158	0	0	0	1
93.057	0	0	1	0
93.499	1	0	0	0
94.801	0	0	1	0
95.382	1	0	0	0
96.434	1	0	0	0
107.348	1	0	0	0
113.071	0	1	0	0

Table 8. A Single Sample Data Generated in Case 4

Failure Time (Month)	Component			
	1	2	3	4
113.983	0	0	1	0
114.759	0	1	0	0
116.364	0	0	0	1
117.134	0	0	0	1
119.743	0	0	1	0

Table 9. Four Components Analysis

10 Years	Exp.	Cycle	Cox	Cusum	Q_{PRA}	Q_m	Q_e
Case 4	0.24	0.35	0.14	0.26	3.3E-6	1.5E-5	2.3E-5
Case 5	0.64	0.65	0.85	0.22	3.4E-6	1.6E-5	3.5E-5
Case 6	0.32	0.68	0.14	0.39	2.7E-6	1.2E-5	1.1E-4

Case 5 Four Components Analysis Dependency (Periodic Clustering in Failure Times)

This case is similar to that of Case 3 except four components were designed in the system for the simulation. That was as average component failure of 0.1 per month was used as baseline failure rate, and it jumps up periodically in 10 months at 0.02 and two months at 0.5. The results (Table 9) has shown that the estimated unavailabilities in PRA are a factor of 10 under estimate, and this is about a single β factor of 1.3×10^{-3} or in terms of partial β factor, i.e.:

$$\begin{aligned}\beta_{2/1} &= 0.05 \\ \beta_{3/2} &= 0.1 \\ \beta_{4/3} &= 0.26\end{aligned}$$

Table 10 shows one of the samples of failure time data generated by the simulation code.

Table 10. Sample Data Generated in Case 5

Failure Time (Month)	Component			
	1	2	3	4
11.887	0	0	1	0
22.272	0	0	0	1
22.812	1	0	0	0
23.094	0	0	0	1
23.578	1	0	0	0
24.057	0	0	1	0
24.063	1	0	0	0
35.266	0	0	0	1
36.796	0	1	0	0
42.602	1	0	0	0
47.224	0	1	0	0
56.396	0	0	0	1
58.297	0	0	1	0
58.729	0	1	0	0
59.152	0	0	1	0
59.451	0	0	0	1
60.175	0	0	0	1
70.720	0	1	0	0

Table 10. Sample Data Generated in Case 5 (Cont'd)

Failure Time (Month)	Component			
	1	2	3	4
71.177	0	0	1	0
72.104	0	1	0	0
82.779	0	1	0	0
83.339	0	1	0	0
83.517	0	0	0	1
83.694	1	0	0	0
83.761	0	0	1	0
84.047	0	1	0	0
84.257	1	0	0	0
94.708	0	0	0	1
95.034	1	0	0	0
95.878	0	0	0	1
96.064	0	0	0	1
96.515	1	0	0	0
103.099	0	0	0	1
105.833	1	0	0	0
106.095	0	0	1	0
106.607	1	0	0	0
107.500	1	0	0	0
107.544	0	0	1	0
108.137	0	1	0	0
108.328	1	0	0	0
111.171	1	0	0	0

Case 6 Four Component Analysis on Failure Rate Changes with an Exponential Time Dependency

The simulation for this case is similar to that of Case 5 except that the increased failure rate during the failure rate changes was modeled as $\lambda \cdot e^{\lambda t}$. The average failure rate of the four components was still at 0.1 per month per component and it will be exponentially increasing starting at 0.01 per month (i.e., $\lambda_0 = 0.01/\text{month}$) ending with about 0.5 per month ($\lambda = 0.032$). The results show (Table 9) that for an exponential increase in failure rate for a period of 10 years, the PRA unavailability estimate for a four component system is under estimated by a factor of 50. This was the greatest deviation observed. Table 11 shows one of the sample data generated in the simulation.

Table 11. A Single Sample of Failure Time Data

Failure Time (Month)	Component			
	1	2	3	4
20.096	0	0	1	0
32.168	0	1	0	0
48.612	0	1	0	0
51.579	0	0	0	1
56.024	1	0	0	0
56.861	0	0	1	0
61.350	1	0	0	0
62.414	1	0	0	0
65.940	0	1	0	0
72.027	0	1	0	0
74.031	0	0	1	0
74.990	0	0	0	1
75.421	0	0	1	0
76.136	0	1	0	0
76.879	1	0	0	0
77.158	0	0	0	1
78.410	0	1	0	0
79.183	0	1	0	0
82.575	0	0	1	0

Table 11. A Single Sample of Failure Time Data (Cont'd)

Failure Time (Month)	Component			
	1	2	3	4
84.642	0	0	0	1
86.159	0	0	0	1
88.713	0	0	0	1
89.664	0	0	0	1
92.639	0	0	1	0
93.598	0	1	0	0
95.478	0	0	1	0
95.926	0	1	0	0
96.369	0	0	1	0
99.845	0	1	0	0
100.950	0	0	0	1
101.053	0	0	0	1
101.092	1	0	0	0
102.428	1	0	0	0
104.532	0	0	1	0
104.630	0	0	0	1
104.642	0	1	0	0
104.648	1	0	0	0
105.666	1	0	0	0
105.865	0	0	0	1
106.589	1	0	0	0
106.934	0	0	1	0
107.092	0	0	0	1
107.326	1	0	0	0
108.399	0	0	1	0
108.779	1	0	0	0

Table 11. A Single Sample of Failure Time Data (Cont'd)

Failure Time (Month)	Component			
	1	2	3	4
109.094	0	0	1	0
109.419	1	0	0	0
110.128	1	0	0	0
110.359	0	0	1	0
112.952	0	1	0	0
113.203	0	1	0	0
113.834	1	0	0	0
115.457	0	0	0	1
115.989	0	1	0	0
117.138	0	0	0	1
118.533	0	1	0	0
118.878	0	0	0	1
119.516	0	0	0	1
120.213	0	0	1	0

APPENDIX G. PILOT APPLICATIONS

This appendix presents pilot applications performed for fans and circuit breakers at six different nuclear power plants, Plants 3, 4, 5, 6, 8. The dampers associated with containment fans were studied in Plant 7.

The applications for Plants 3 and 4 show noticeable dependency, the applications conducted for Plants 5 and 6 show slight dependency, the study performed for Plant 7 shows very slight contribution from dependency and, finally, the circuit breakers at Plant 8 show positive dependency. The concept of positive dependency is introduced in this appendix. The order in which the results of the plants are presented is from high to low dependency, that is, first Plant 3 for which there is noticeable dependency, then Plant 5, which has slight dependency, and so on.

The raw data used in these applications were retrieved from the NPRDS. The data used for the applications are presented indicated by odd numbers (e.g., 61, 63, etc.), where the failure data are in chronological order. The headings of the tables show the code for the components under study. In the column corresponding to each component, a "1" signals failure, while a "0" signals "no failure." If 2 or more components failed in the same date, different rows are assigned to the failure of each component.

One way of detecting dependency is by comparing "Observed Minimal Cycle Duration (MCD)" against "Theoretical MCD," for a set of data. These two parameters are discussed in the tables in this appendix. The concept of "Minimal Cycle Duration" (MCD) was discussed in Section 2.3 of Volume 1. As discussed earlier, when there is a shock affecting several components, then the MCD obtained from the actual field failure data (Observed MCD) is shorter than the theoretical MCD.

Plant 3

Degree of Dependency: Noticeable

Type of Component: Circuit Breaker

Table G.1. Data for Circuit Breakers at Plant 3

Number of Failure	Date	FC A	FC B	FC C	FC D
1	05/18/86	1	0	0	0
2	01/08/87	1	0	0	0
3	10/28/87	0	0	0	1
4	12/09/87	0	0	1	0
5	12/10/87	0	0	0	1
6	12/12/87	0	1	0	0
7	07/19/88	1	0	0	0
8	10/02/88	0	0	0	1
9	05/02/89	0	0	1	0

The results for the data in Table G.1 are given in Table G.2 for a test interval of 30 days. The Exact Unavailability with and including Dependency in Table G.2 are calculated based on the unavailability formulation discussed in this report (Eqn. 25 in Volume 1). The PRA approximate unavailabilities are not shown in this table. However, they are given in Table 12 of Volume 1. Case 2 and Case 3 in Table 12 of Volume 1 correspond to Exact Unavailability without and including Dependency in Table G.2.

Table G.2. Results for the Circuit Breakers at Plant 3

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	126.44	126.44	1.10E-01	1.10E-01
2	252.86	295.04	6.35E-03	6.92E-03
3	398.40	547.93	1.88E-04	4.55E-04
4	566.33	1053.70	2.24E-06	1.60E-05

Table G.2 shows that there is dependency in the failures. The unavailability including dependency for 4 components is 1 magnitude larger than the unavailability calculation without considering dependency. The observed (real) MCDs are shorter, for all cycles of multiple components, than the theoretical ones.

Plant 4

Degree of Dependency: Noticeable

Type of Component: Fan Coolers

Table G.3. Data for Fan Coolers at Plant 4

Number of Failure	Date	FC 1A	FC 1B	FC 1C	FC 1D
1	03/21/83	1	0	0	0
2	04/02/83	0	1	0	0
3	04/27/83	1	0	0	0
4	10/06/83	0	0	0	1
5	03/19/84	0	0	0	1
6	03/24/84	1	0	0	0
7	04/26/84	0	1	0	0
8	09/15/84	1	0	0	0
9	11/02/84	0	0	0	1
10	12/13/84	0	0	0	1
11	01/01/85	0	1	0	0
12	02/10/85	0	0	1	0
13	04/01/85	0	0	1	0
14	06/20/85	1	0	0	0
15	10/20/85	1	0	0	0
16	11/20/85	1	0	0	0
17	01/15/86	0	1	0	0
18	02/10/86	0	0	1	0
19	06/12/86	1	0	0	0
20	08/15/86	0	1	0	0
21	12/01/86	1	0	0	0
22	09/01/87	1	0	0	0
23	12/03/87	0	0	1	0
24	10/01/89	0	1	0	0

Number of Failure	Date	FC 1A	FC 1B	FC 1C	FC 1D
25	04/09/90	0	1	0	0
26	09/03/91	0	1	0	0
27	11/01/91	0	1	0	0

The results for the data in Table G.3 are presented in Table G.4 for a test interval of 30 days.

Table G.4. Results for the Fan Coolers at Plant 4

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	118.63	118.63	1.16E-01	1.16E-01
2	204.35	276.80	7.17E-03	1.04E-02
3	364.42	514.06	2.26E-04	3.96E-04
4	513.50	988.58	2.88E-06	7.72E-06

Table G.4 shows that there is dependency in the failures. The unavailability calculation including dependency is about 2 larger than the unavailability calculation without considering dependency. The observed (real) MCDs are shorter, for all cycles of multiple components, than the theoretical ones.

Plant 5

Degree of Dependency: Slight

Type of Component: Circuit Breaker

Table G.5. Data for Circuit Breakers at Plant 5

Number of Failure	Date	1A	1B	1C	1D
1	03/12/84	0	0	0	1
2	03/24/85	0	0	0	1
3	05/20/85	0	0	0	1
4	02/13/86	0	0	0	1
5	03/08/87	0	0	1	0
6	03/15/87	0	0	0	1
7	12/09/87	0	0	0	1
8	12/25/87	0	1	0	0
9	04/05/88	0	0	0	1
10	04/05/88	0	0	1	0
11	04/11/88	1	0	0	0
12	04/11/88	1	0	0	0
13	04/17/88	0	1	0	0
14	04/30/88	0	0	0	1
15	10/14/89	0	0	0	1

The results for the data in Table G.5 are given in Table G.6 for a test interval of 30 days.

Table G.6. Results for the Circuit Breakers at Plant 5

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	139.93	139.93	9.99E-02	9.99E-02
2	261.87	326.51	5.2 E-03	5.18E-03
3	395.80	606.38	1.40E-04	4.41E-04
4	562.67	1166.11	1.51E-06	5.71E-05

Table G.6 shows that there is only a slight dependency in the failures. For cycle 2, the unavailability calculation with and without dependency is almost the same. For cycles 3 and 4, there is a factor of about 3 increase when the dependency is taken into account. For these cycles, the observed (real) MCDs are shorter than the theoretical ones.

Plant 6

Degree of Dependency: Slight
Type of Component: Fan Coolers

Table G.7. Data for Fan Coolers at Plant 6

Number of Failure	Date	HVH-1	HVH-2	HVH-3	HVH-4
1	05/17/82	0	0	1	0
2	04/10/83	0	0	1	0
3	09/07/83	0	1	0	0
4	09/26/83	0	1	0	0
5	09/28/83	0	1	0	0
6	10/12/83	0	0	1	0
7	11/03/83	1	0	0	0
8	11/03/83	0	0	1	0
9	11/04/83	1	0	0	0
10	01/31/86	1	0	0	0
11	01/01/87	0	0	0	1
12	01/01/87	0	1	0	0
13	01/01/87	1	0	0	0
14	01/01/87	0	0	1	0
15	05/23/87	0	1	0	0
16	05/23/87	0	0	1	0
17	05/24/87	1	0	0	0

The results for the data in Table G.7 are presented in Table G.8 for a test interval of 30 days.

Table G.8. Results for the Fan Coolers at Plant 6

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	111.24	111.24	1.24E-01	1.24E-01
2	212.92	259.55	8.09E-03	7.88E-03
3	508.44	482.02	2.72E-04	3.03E-04
4	1363.70	926.96	3.70E-06	4.76E-06

Table G.8 shows that there is only a slight dependency in the failures. For cycles 2 and 3, the unavailability calculation with and without dependency is almost the same. For cycle 4, the unavailability calculation including dependency is just above the calculation without dependency. For this cycle, the observed (real) MCD is shorter than the theoretical one.

Plant 7

Degree of Dependency: Very Slight

Type of Component: Fan Dampers

Table G.9. Data for Fan Dampers at Plant 7

Number of Failure	Date	1F	2F	3F	4F
1	01/05/78	0	1	0	0
2	03/06/80	0	1	0	0
3	02/05/81	0	1	0	0
4	10/09/81	0	1	0	0
5	06/03/82	0	0	1	0
6	11/13/82	1	0	0	0
7	03/09/83	0	0	1	0
8	03/09/83	0	1	0	0
9	10/01/85	0	0	0	1
10	01/24/89	0	0	1	0
11	09/01/89	0	1	0	0

The results for the data in Table G.9 are presented in Table G.10 for a test interval of 30 days.

Table G.10. Results for the Fan Dampers at Plant 7

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	392.18	392.18	3.73E-02	3.73E-02
2	996.57	915.09	7.07E-04	6.07E-04
3	1582.20	1699.45	6.76E-06	8.57E-06
4	2884.00	3268.18	2.59E-08	4.89E-08

From Table G.10, it can be observed that there is only a slight dependency in the failures. For cycle 2, the observed (real) MCD is larger than the theoretical one. The methodology developed is based on the concept of "shocks". These shocks are generally regarded as "negative" because they increase the unavailability of several components. However, in reality there can also be "positive" shocks which can *increase* the availability of a group of components. The main example of a positive shock is a good maintenance practice, such as initiating preventive maintenance and surveillance on all components when a single component experiences a potential dependent failure. Since the maintenance is performed adequately, it acts as a "positive shock" on the group of components. The accurate expression for unavailability described in this report takes into account this effect whenever the value in the column entitled "Exact Unavailability without Dependency" is larger than the value in the column entitled "Exact Unavailability Including Dependency".

The capability and flexibility of the methodology developed to take into account both *negative* and *positive shocks* is a very powerful tool in probabilistic safety assessments. It allows us to *monitor*, in risk terms, the performance of a group of components from the data on field failure. Both positive and negative *environments* can be identified and a more effective allocation of resources, with a corresponding gain both in safety and economic terms, may be carried out.

For cycles 3 and 4, the unavailability calculation including dependency is just above the calculation without dependency. For these cycles, the observed (real) MCDs are shorter than the theoretical ones.

Plant 8

Degree of Dependency: No Dependency

Type of Component: Circuit Breakers

Table G.11. Data for Circuit Breakers at Plant 8

Number of Failure	Date	1A4X	1A2X	1B4X	1B2X	1C4X	1C8X	1C6X
1	07/18/84	0	1	0	0	0	0	0
2	09/22/84	0	0	0	0	1	0	0
3	01/15/85	0	0	1	0	0	0	0
4	01/15/85	0	0	0	1	0	0	0
5	08/19/85	0	0	0	0	0	0	1
6	09/11/85	1	0	0	0	0	0	0
7	09/16/85	0	0	0	0	0	0	1
8	11/24/85	0	0	0	0	0	0	1
9	12/29/87	0	0	1	0	0	0	0
10	06/02/89	0	0	0	0	1	0	0
11	01/03/91	0	0	1	0	0	0	0
12	11/08/91	0	0	0	0	0	1	0

The results for the data in Table G.11 are given in Table G.12 for a test interval of 30 days.

Table G.12. Results for the Circuit Breakers at Plant 8

Cycle	Observed MCD	Theoretical MCD	Exact Unavailability without Dependency	Exact Unavailability Including Dependency
1	238.75	238.75	6.03E-02	6.03E-02
2	515.30	517.29	2.13E-03	2.22E-03
3	819.86	851.54	4.75E-05	4.37E-05
4	1008.67	1269.35	6.81E-07	6.32E-07
5	1303.24	1826.44	6.12E-09	5.10E-09
6	1641.50	2662.06	3.15E-11	1.93E-11
7	2865.0	4333.31	7.10E-14	3.23E-14

In this case, the failure data are very sparse in time. For example, Table G.11 shows that there is one failure in 1987 and one in 1989, but none in 1986, 1988, and 1990. From this framework, we concluded that these failure data show no dependency. This situation also is also by the methodology presented in this report. Table G.12 shows that the unavailability calculations with and without dependency are very close to each other for cycles 2, 3, and 4. For cycles 5, 6, and 7, the two unavailability calculations may also be regarded as being close to each other given both the very small unavailabilities (less than 10^{-8}) and the sparsity in time of the failure data.

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

2. TITLE AND SUBTITLE

Methods for Dependency Estimation and System Unavailability
Evaluation Based on Failure Data Statistics

Detailed Description and Applications

5. AUTHOR(S)

M. A. Azarm, F. Hsu, G. Martinez-Guridi/Brookhaven National Laboratory
W. E. Vesely/Science Applications International Corporation

1. REPORT NUMBER
(Assigned by NRC, Add Vol., Supp., Rev.,
and Addendum Numbers, If any.)

NUREG/CR-5993
BNL-NUREG-52362
Vol. 2

3. DATE REPORT PUBLISHED
MONTH YEAR

July 1993

4. FIN OR GRANT NUMBER

A3295

6. TYPE OF REPORT

Technical

7. PERIOD COVERED (Inclusive Dates)

8. PERFORMING ORGANIZATION -- NAME AND ADDRESS (If NRC, provide Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Brookhaven National Laboratory
Upton, NY 11973

Subcontractor:
Science Applications International Corporation
655 Metro Place South
Suite 745
Dublin, OH 43017

9. SPONSORING ORGANIZATION -- NAME AND ADDRESS (If NRC, type "Same as above"; if contractor, provide NRC Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address.)

Division of Systems Research
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001

10. SUPPLEMENTARY NOTES

11. ABSTRACT (200 words or less)

This report introduces a new perspective on the basic concept of dependent failures where the definition of dependency is based on clustering in failure times of similar components. This perspective has two significant implications: firstly, it relaxes the conventional assumption that dependent failures must be simultaneous and result from a severe shock; secondly, it allows the analyst to use all the failures in a time continuum to estimate the potential for multiple failures in a window of time (e.g., a test interval), therefore arriving at a more accurate value for system unavailability. In addition, the models developed here provide a method for plant-specific analysis of dependency, reflecting the plant-specific maintenance practices that reduce or increase the contribution of dependent failures to system unavailability.

The proposed methodology can be used for screening analysis of failure data to estimate the fraction of dependent failures among the failures. In addition, the proposed method can evaluate the impact of the observed dependency on system unavailability and plant risk. The formulations derived in this report have undergone various levels of validations through computer simulation studies and pilot applications. The pilot applications of these methodologies showed that the contribution of dependent failures of diesel generators in one plant was negligible, while in another plant, it was quite significant. It also showed that in the plant with significant contribution of dependency to Emergency Power System (EPS) unavailability, the contribution changed with time. Similar findings were reported for the Containment Fan Cooler breakers. Drawing such conclusions about system performance would not have been possible with any other reported dependency methodologies.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

Reactor components-availability, nuclear power plants-components;
nuclear power plants-risk assessment, reactor components-reliability
failure mode analysis, failures, mathematical models, probabilistic
estimation, system failure analysis, distribution functions failures-
data analysis

13. AVAILABILITY STATEMENT

Unlimited

14. SECURITY CLASSIFICATION

(This Page)

Unclassified

(This Report)

Unclassified

15. NUMBER OF PAGES

16. PRICE

END

DATE
FILMED
9 / 10 / 93