P R E S S U R I Z E D   W A T E R   R E A C T O R   O W N E R S   G R O U P

PWROG-18042-NP
Revision 1

WESTINGHOUSE NON-PROPRIETARY CLASS 3

FLEX Equipment Data Collection and Analysis

Risk Management Committee

PA-RMSC-1651 R1

February 2022

Westinghouse

PWROG-18042-NP
Revision 1

FLEX Equipment Data Collection and Analysis

PA-RMSC-1651 R1

Matthew M. Degonish*
Risk Analysis

February 2022

Reviewer:	Kenneth L. Kiper* Risk Analysis
Approved:	Stacy A. Davis*, Manager Risk Analysis
Approved:	Damian S. Mirizio*, Program Director PWR Owners Group PMO

*Electronically approved records are authenticated in the electronic document management system.

Westinghouse Electric Company LLC
1000 Westinghouse Drive
Cranberry Township, PA 16066, USA

ACKNOWLEDGEMENTS

This report was developed and funded by the PWR Owners Group under the leadership of the participating utility representatives of the Risk Management Committee. The author would like to thank the following people and/or organizations for their valuable contributions to this report:

- Roy Linthicum – Chairman, Risk Management Committee
- Mike Powell – Arizona Public Service
- Allen Moldenhauer – Dominion Energy
- Thomas John – Dominion Energy
- Tracy St. Clair –Energy Harbor
- Ken Kiper – Westinghouse Electric Company
- R. Aaron Moreno – Westinghouse Electric Company
- Carroll A. Trull – Westinghouse Electric Company
- PWROG RMC members who participated in review and preparation of this analysis

LEGAL NOTICE

COPYRIGHT NOTICE

DISTRIBUTION NOTICE

PWR Owners Group
United States Member Participation* for PA-RMSC-1651 R1

Utility Member	Plant Site(s)	Participant Yes	Participant No
Ameren Missouri	Callaway (W)	X	
American Electric Power	D.C. Cook 1 & 2 (W)	X	
Arizona Public Service	Palo Verde Unit 1, 2, & 3 (CE)	X	
Dominion Energy	Millstone 2 (CE)	X	
	Millstone 3 (W)	X	
	North Anna 1 & 2 (W)	X	
	Surry 1 & 2 (W)	X	
	V.C. Summer (W)	X	
Duke Energy Carolinas	Catawba 1 & 2 (W)	X	
	McGuire 1 & 2 (W)	X	
	Oconee 1, 2, & 3 (B&W)	X	
Duke Energy Progress	Robinson 2 (W)	X	
	Shearon Harris (W)	X	
Entergy Palisades	Palisades (CE)	X	
Entergy Operations South	Arkansas 1 (B&W)	X	
	Arkansas 2 (CE)	X	
	Waterford 3 (CE)	X	
Evergy	Wolf Creek (W)	X	
Exelon Generation Co. LLC	Braidwood 1 & 2 (W)	X	
	Calvert Cliffs 1 & 2 (CE)	X	
	Ginna (W)	X	
Energy Harbor	Beaver Valley 1 & 2 (W)	X	
	Davis-Besse (B&W)	X	
Florida Power & Light \ NextEra	St. Lucie 1 & 2 (CE)	X	
	Turkey Point 3 & 4 (W)	X	
	Seabrook (W)	X	
	Pt. Beach 1 & 2 (W)	X	
Luminant Power	Comanche Peak 1 & 2 (W)	X	
Pacific Gas & Electric	Diablo Canyon 1 & 2 (W)	X	
PSEG – Nuclear	Salem 1 & 2 (W)	X	
So. Texas Project Nuclear Operating Co.	South Texas Project 1 & 2 (W)	X	

PWR Owners Group
United States Member Participation* for PA-RMSC-1651 R1

Utility Member	Plant Site(s)	Participant	
		Yes	No
Southern Nuclear Operating Co.	Farley 1 & 2 (W)	X	
	Vogtle 1 & 2 (W)	X	
	Vogtle 3 & 4 (W)	X	
Tennessee Valley Authority	Sequoyah 1 & 2 (W)	X	
	Watts Bar 1 & 2 (W)	X	
Xcel Energy	Prairie Island 1 & 2 (W)	X	

* Project participants as of the date the final deliverable was completed. On occasion, additional members will join a project. Please contact the PWR Owners Group Program Management Office to verify participation before sending this document to participants not listed above.

PWR Owners Group
International Member Participation* for PA-RMSC-1651 R1

Utility Member	Plant Site(s)	Participant	
		Yes	No
Asociación Nuclear Ascó-Vandellòs	Asco 1 & 2 (W)	X	
	Vandellos 2 (W)	X	
Centrales Nucleares Almaraz-Trillo	Almaraz 1 & 2 (W)	X	
CEZ**	Temelin		X
EDF Energy	Sizewell B (W)	X	
Electrabel	Doel 1, 2 & 4 (W)	X	
	Tihange 1 & 3 (W)	X	
Electricite de France	56 Units	X	
Elektriciteits Produktiemaatschappij Zuid-Nederland	Borssele 1 (Siemens)	X	
Eletronuclear-Eletrobras	Angra 1 (W)	X	
Emirates Nuclear Energy Corporation	Barakah 1 & 2	X	
Hokkaido	Tomari 1, 2 & 3 (MHI)	X	
Japan Atomic Power Company	Tsuruga 2 (MHI)	X	
Kansai Electric Co., LTD	Mihama 3 (W)	X	
	Ohi 3 & 4 (W & MHI)	X	
	Takahama 1, 2, 3 & 4 (W & MHI)	X	
Korea Hydro & Nuclear Power Corp.	Kori 1, 2, 3 & 4 (W)	X	
	Hanbit 1 & 2 (W)	X	
	Hanbit 3, 4, 5 & 6 (CE)	X	
	Hanul 3, 4, 5 & 6 (CE)	X	
Kyushu	Genkai 3 & 4 (MHI)	X	
	Sendai 1 & 2 (MHI)	X	
Nuklearna Electrarna KRSKO	Krsko (W)	X	
Ringhals AB	Ringhals 3 & 4 (W)	X	
Shikoku	Ikata 3 (MHI)	X	
Taiwan Power Co.	Maanshan 1 & 2 (W)	X	

* Project participants as of the date the final deliverable was completed. On occasion, additional members will join a project. Please contact the PWR Owners Group Program Management Office to verify participation before sending this document to participants not listed above.

** I&C Working Group Only

TABLE OF CONTENTS

List of Tables

List of Figures

Abbreviations

AC – Alternating Current

ANS – American Nuclear Society

ASME – American Standard

CAP – Corrective Action Program

CNID – Constrained Noninformative Prior Distribution

CR – Condition Report

CTG – Combustion Turbine Generator

DOF – Degrees of Freedom

DG – Diesel Generator

EB – Empirical Bayes

EDC – Engine-Driven Compressor

EDG – Emergency Diesel Generator

EDP – Engine-Driven Pump

EFW – Emergency Feedwater

EPRI – Electric Power Research Institute

FTR – Fail to Run

FTS – Fail to Start

INL – Idaho National Laboratory

JNI – Jeffreys Noninformative Prior

MDC – Motor-Driven Compressor

MDP – Motor-Driven Pump

MLE – Maximum Likelihood Estimate

NEI – Nuclear Energy Institute

NRC – U.S. Nuclear Regulatory Commission

PDDCP – Portable Diesel-Driven Centrifugal Pump

PDF – Probability Density Function

PM – Preventative Maintenance

PRA – Probabilistic Risk Assessment

PWR – Pressurized Water Reactor

PWROG – Pressurized Water Reactor Owners Group

RISC – Risk-Informed Steering Committee

RMSC - Risk Management Sub-Committee

SBY – Standby

Executive Summary

This report provides generic unreliability estimates (failure rates) for standard FLEX equipment in use in the probabilistic risk assessments (PRAs) for the United States (U.S.) nuclear industry. This report also documents the philosophy guiding the effort to update the inputs used for FLEX Equipment in PRA models.

In March 2012, the U.S. Nuclear Regulatory Commission (NRC) issued order EA-12-049 [Ref. 1] requiring nuclear power plants to develop mitigation strategies for beyond-design-basis external events (i.e., FLEX). The industry, through the Nuclear Energy Institute (NEI), developed a generic framework for response to this order that is documented in the NEI FLEX Implementation Guide (NEI 12-06, [Ref. 2]). The framework outlines an approach for adding diverse and flexible mitigation capabilities to increase defense-in-depth for beyond-design-basis scenarios to address an extended loss of AC power and loss of normal access to the ultimate heat sink at all units on a site [Ref. 3].

The U.S. nuclear sites (termed "industry" in this report) maintain PRA models that support day-to-day risk-informed decision making. Currently, when modeling FLEX equipment, plant-specific data (on a site-by-site basis) is used to estimate failure parameters for the components being credited. As the operating experience with FLEX equipment is still relatively small, this practice introduces a large amount of uncertainty into the derivation of failure parameters that are being used in the PRA models. Additionally, for consistency, it is desired to have one generic source of failure parameters which can be updated as necessary on a plant-specific basis.

In May 2017, the Risk-Informed Steering Committee (RISC) accepted the PWROG proposal to develop FLEX equipment failure rates for the industry. As a global organization, the PWR Owners Group provides a unique opportunity to share operating experience and to develop common solutions to operating issues confronting PWR members. The PWROG put together a project team to assist in defining the needs of utilities, as well as resolving some generic issues prior to collecting FLEX equipment data. Initial tasks performed by the project team were: (1) providing instructions for how to extract data from the EPRI Database[1], and (2) reaching consensus on certain modeling aspects of FLEX equipment to appropriately bound conditions for the data evaluation. As a result of this effort, component unreliability (UR) estimates for: portable diesel generators, portable combustion turbine generators, portable diesel-driven pumps, portable motor-driven positive displacement pumps, and portable air compressors are addressed in this report.

Initial drafts of this report were created and provided to the PWROG Risk Management Committee (RMC) members and the U.S. Nuclear Regulatory Commission (NRC) who provided comments and suggestions for improvement that greatly improved the fidelity of the analysis herein. The results presented in this report can be incorporated into plant-specific analyses as prior distributions in Bayesian updates using plant-specific data and may also have use for the NRC SPAR models.

[1] The EPRI FLEX Preventive Maintenance (PM) Database was originally developed to use operating experience to adjust FLEX PMs.

1 Introduction

This report provides generic unreliability estimates (failure rates) for standard FLEX equipment in use in the probabilistic risk assessments (PRAs) for the United States (U.S.) nuclear industry. This report also documents the philosophy guiding the effort to update the inputs used for FLEX Equipment in PRA models.

In March 2012, the U.S. Nuclear Regulatory Commission (NRC) issued order EA-12-049 [Ref. 1] requiring nuclear power plants to develop mitigation strategies for beyond-design-basis external events (i.e., FLEX). The industry, through the Nuclear Energy Institute (NEI), developed a generic framework for response to this order that is documented in the NEI FLEX Implementation Guide (NEI 12-06, [Ref. 2]). The framework outlines an approach for adding diverse and flexible mitigation capabilities to increase defense-in-depth for beyond-design-basis scenarios to address an extended loss of AC power and loss of normal access to the ultimate heat sink at all units on a site [Ref. 3].

The objective of FLEX is to provide a programmatic and controlled approach to transition to mobile equipment intended to mitigate a beyond-design-basis external event. Portable equipment that supplements installed systems will enable key safety functions to be maintained despite a postulated extended loss of normal AC power and loss of normal access to the ultimate heat sink. Protection, access and connections for the portable equipment must also be provided [Ref. 3].

The U.S. nuclear sites (termed “industry” in this report) maintain PRA models that support day-to-day risk-informed decision making. Currently, when modeling FLEX equipment, plant-specific data (on a site-by-site basis) is used to estimate failure parameters for the components being credited. As the operating experience with FLEX equipment is still relatively small, this practice introduces a large amount of uncertainty into the derivation of failure parameters that are being used in the PRA models. Additionally, for consistency, it is desired to have one generic source of failure parameters which can be updated as necessary on a plant-specific basis.

2 Historical Data Collection and Analysis Efforts

In May 2017, the Risk-Informed Steering Committee (RISC) accepted the PWROG proposal to develop FLEX failure rates for the industry. At the April 2017 RMSC Meeting, the PWROG put together a project team to assist in defining the needs of utilities, as well as resolving some generic issues prior to collecting FLEX equipment data. Initial tasks performed by the project team were: (1) providing instructions for how to extract data from the EPRI Database[1], and (2) reaching consensus on certain modeling aspects of FLEX equipment to appropriately bound conditions for the data evaluation. Note that the final component grouping was developed based on the recommendations from the NRC/INL audit of this report and is discussed in further detail in Section 4.1.

The components to be evaluated as part of this analysis are as follows:

- Portable Diesel Generators
- Portable Combustion Turbine Generators
- Portable Diesel-Driven Pumps
- Portable Motor-Driven Positive Displacement Pumps
- Portable Air Compressor

The structure laid out by the PWROG serves as the basis for this report and is documented herein.

[1] The EPRI FLEX Preventive Maintenance (PM) Database was originally developed to use operating experience to adjust FLEX PMs.

3 Data Compilation and Database Development Philosophy

The overall goal of this report is to provide failure rate parameters for FLEX equipment consistent with the process currently provided for permanently installed equipment in NUREG/CR-6928 [Ref. 4]. The goals of this report are the same as discussed in Section 3 of NUREG/CR-6928 [Ref. 4], but are summarized below:

1. Use data from comprehensive and consistently collected and interpreted sources (containing both failure and demand or run hour information) that are maintained and updated. For this report, the source information comes from the collection of condition reports (CRs) and preventative maintenance (PM) events related to FLEX equipment.
2. Characterize the current industry performance.
3. Structure the characterization of industry-average performance such that results can be updated periodically.

Additionally, the following goals were added for this report:

4. Gain insights on quality and effectiveness of FLEX equipment and PM tasks in ensuring long-term reliability of equipment.
5. Provide recommendations to ease the process for future data updates.

4 Data Collection

4.1 Component Boundaries

The available FLEX components will vary from site to site, but the list of component types considered in this analysis have been determined to be common amongst a group of utilities and thus are the focus of this analysis. Prior to requesting data from the respective utilities, the component boundary definitions for each piece of equipment considered in this analysis are required to be defined. Each component grouping consists of various operating range sizes (e.g., Portable Diesel Generator contains high, medium, and low voltage generators). The various operating range sizes were combined into one group because: (1) the data were not sufficient to further subdivide the component-type groups, and (2) the groups are considered to be generally homogenous with regard to the basic component type and having a common function in FLEX strategies. The following component boundaries will be used for this analysis.

Table 4-1: Component Types and Boundaries

Component Type	Description	Boundary
Portable Diesel Generator		
Portable Diesel Generator	The information collected for portable diesel generators includes various component capacities (i.e., high, medium, and low voltage)[1].	The diesel generator boundary includes the diesel engine with all components in the exhaust path, electrical generator, generator exciter, combustion air, lube oil systems, fuel oil systems and starting compressed air system, and local instrumentation and control circuitry. Additionally, starter batteries are included.
Portable Combustion Turbine Generator		
Portable Combustion Turbine Generator	The information collected for portable combustion turbine generators includes various component capacities (i.e., high and medium voltage)[1].	The combustion turbine generator boundary includes the gas turbine, generator, circuit breaker, local lubrication or cooling systems, and local instrumentation and control circuitry. Additionally, starter batteries are included.

Table 4-1: Component Types and Boundaries		
Component Type	**Description**	**Boundary**
Portable Diesel-Driven Pump		
Portable Diesel-Driven Pump	The information collected for portable diesel-driven pumps includes various component capacities (i.e., high, medium, and low pressure with combinations of high, medium, and low flow rate)[2]. Additionally, the data set includes both centrifugal and positive displacement pumps with diesel drivers.	The diesel-driven pump boundary includes the pump, diesel engine, local lubrication or cooling systems, and local instrumentation and control circuitry. Additionally, starter batteries are included.
Portable Motor-Driven Positive Displacement Pump		
Portable Motor-Driven Positive Displacement Pump	The information collected for portable motor-driven positive displacement pumps includes various component capacities (i.e., high, medium, and low pressure with combinations of high, medium, and low flow rate)[2].	The positive displacement pump boundary includes the pump, motor, local circuit breaker, local lubrication or cooling systems, and local instrumentation and control circuitry.
Portable Air Compressor		
Portable Air Compressor	The information collected for portable air compressors includes various drivers (i.e., motor-driven and diesel-driven).	The air compressor boundary includes the compressor, driver, local circuit breaker, local lubrication or cooling systems, and local instrumentation and control circuitry. Additionally, starter batteries are included.
Notes: 1. Capacity ranges for generators are as follows: High Voltage (4 kV) Medium Voltage (480 V) Low Voltage (120 V to 240 V) 2. Capacity ranges for Pumps are as follows: High Pressure (>1000 psig) and Low Flow Rate (< 200 gpm) Medium to Low Pressure (50 to 500 psig) and Medium to Low Flow Rate (50 to 500 gpm) Low Pressure (50 to 200 psig) and High Flow Rate (> 500 gpm) Low Pressure (50 to 200 psig) and High Flow Rate (> 500 gpm)		

4.2 Data Request and Use of EPRI FLEX PM Database

To support the data collection effort, a request was sent out to the industry for specific information (as specified below) regarding FLEX equipment including documented events, maintenance schedules, and run-hours. To support this effort, it was recognized that the EPRI FLEX PM Database may be used as a starting point to collect the requested information but was not required to be used. While the EPRI FLEX PM Database is recommended as a starting point for collection of FLEX equipment success and failure data, it is recognized that the database wasn't created with the purpose of supporting PRA analyses; therefore, a set of guidelines was provided to assist utility members in collecting the data in attempts to ensure that the provided inputs meet the needs of the analysis team for use in this report.

The following guidelines were provided to the utility members:

- Contact the appropriate personnel at each site who was designated to enter equipment information into the FLEX PM database.
- Consider use of the EPRI Database as a starting point for collecting FLEX equipment data; however, it is not necessary. If using the EPRI database, work with the appropriate contact at site to determine methodology for how PMs were entered into the database. Note that each site may have entered PM information differently (even within one utility) and understanding how the PMs were entered is critical for providing resolution of the information. Some important notes about information provided in the EPRI Database is provided below:
    - There is an "Event" field in the database which can classify the PM as "Test Failed"; this is designated for the FLEX strategy and may not be applicable as a PRA failure. These events should be reviewed to ensure that the event represents a PRA failure.
    - There is an "Equipment Operating Hrs." field in the database; this field provides an estimate of the equipment run time and is an estimate from the person responsible for entering the information. This information needs to be validated and updated if necessary.
    - An entry in the database may be applicable to any number of components (not just a single component necessarily). Resolution should be provided to identify the number of components involved with a PM event.
    - A large amount of equipment in the database is representative of FLEX support equipment. These components are often not considered in the component boundaries of the components of interest explicitly.
    - Each entry in the EPRI database provides a reference to the appropriate site document. If available, this information should be provided with the data.
    - The EPRI FLEX Equipment Maintenance Event Collaboration Site Event Entry Instructions serve as a help document for using the EPRI database. This document is available for those with access to the database and can be located using the "Help" menu.

- Provide the following information:
    - Component type (as classified in Section 4.1), make, model, and driver type
    - A count of the number of failures, demands, and run-hours for the component
    - The basis for classifying an event as a failure
    - The site reference document
    - The utility name and plant name
    - Any additional comments

4.3 Data Input

After submitting the request as discussed in Section 4.2, various responses were received from the U.S. sites. In general, the minimum set of information required from each plant consists of: number of FLEX components, PM activities for each component, number of demands and run hours for each component, and any adverse events identified for those components. In some situations, additional clarifications on the submitted data were requested from the applicable utility.

4.4 Data Compilation

There were 99 plants that reported data for inclusion in this analysis. Upon receipt of data from the survey participants, the data was reviewed for completeness and compiled in a database for further analysis. For each plant site, there are four (4) categories of information that need to be compiled: (1) the type, make and model of FLEX equipment at each plant site; (2) run-time on either a per component group or individual component basis; (3) Preventative Maintenance (PM) frequency for each component group reported; and (4) details of any adverse events related to the maintenance or operation of the FLEX equipment.

The compilation of data with regard to the type, make and model of FLEX equipment is relatively straightforward. Each component was assigned a unique component designator made up of initials representing the utility, site, component type (e.g., Portable Diesel-Driven Centrifugal Pump) operating range (e.g., High Pressure Low Flow) and a single digit identification number. The resulting number of component types used in this analysis is shown in Table 5-2. Note that although the final component grouping does not consider operating range, this level of detail is maintained in the database for potential use in future updates.

For example, NPP-S36 provided information on two portable diesel-driven centrifugal pumps (medium to low pressure/flow rate). These components were entered into the database (*tbl_Components*) to a unique key. After this information was entered, the database was updated using the custom name as a primary key to link the information together throughout multiple tables in the database.

Run time was generally provided by utilities in one of two basic ways: (1) on a per component basis or (2) as a summation for a component group. When run-hour figures were provided on a component basis, they were applied directly to the respective component. When run-hour data was provided for a component grouping, the run-hours were divided equally amongst the components in the group. In some instances, run-hours were not provided by the utility. In these cases, if the utility provided valuable information for the other categories, the data analysts made an estimate of the number of run-hours to make the best use of the data provided. Where run-hour estimates were made, these estimates were established based on the reported PM frequency, the date of the initial PM, and an assumed duration based on the type of PM. Based upon observation from the utilities that provided run-hour data, it appears that the PM related runs for FLEX equipment are relatively short in duration. This observation lead to assigning a value of 0.5 run-hours for PMs with a frequency of less than 1 year and a value of 1.0 run-hours for each PM with a frequency of greater than or equal to 1 year. The resulting total industry-level run-hours by component type are shown in Table 5-2.

The demand data was compiled based on the reported PM frequency, the start date of the PM, and the date in which the data from the respective utilities was received. Similar to the run time data, there was a varied range of completeness and usability across the responses that were provided. Attempting to utilize as much utility provided data as possible, some judgements were made with regards to compiling the number of demands. For instance, if a utility did not provide information related to the time in which the PM was established (date of initial PM) but had an identified event that occurred in 2016 then January 1, 2016 was assumed to be the start date of the PMs at the utility. The resulting total industry level demands by component type are shown in Table 5-2.

Finally, the identified events for each component were entered and a failure criterion was assigned (see Section 5.2 for more information on this process).

5 Data Analysis

5.1 Event Reporting

The 99 plants included in this analysis experienced a total of 794 events from the adoption of FLEX strategies at each site through (roughly) 2019[1]. The events in this analysis were binned into Failure Criteria based upon the brief description of the event from the respective utilities' Corrective Action Program (CAP). The level of detail describing each event varies greatly over a spectrum requiring some judgement by the analysts in regard to binning. It is recognized that in its current form, the data may contain infant-mortality type failures that do not represent long term component reliability. As future data collection efforts are performed, these failures may be pruned from the data used to calculate component failure rates.

The analysts made judgements regarding the reported events in two areas: (a) whether the event was indeed a failure and (b) for events determined to be failures, whether the failure mode was fail-to-start (FTS) or fail-to-run (FTR). The criteria used to judge failure events in provided in Section 5.2. In addition to the analysts' judgement, the classification of failure events was reviewed by utility representatives to confirm appropriate classification of the events.

Following the NRC Audit, the available events were reviewed to determine if any failures (specifically for the diesel generators) were representative of a fail-to-load failure mode. At most, one event was identified as a potential fail-to-load event. This review, as well as the current failure rate results, show that a fail-to-load failure mode would not be an important failure mode for this class of DGs given the current set of data. This review should be repeated with each update to ensure its applicability.

With regard to the failure mode, analysts used the following logic to determine whether the failure was FTS or FTR:

Considering the short operating periods the FLEX equipment is typically run, the decision was made to distinguish between a run or start failure by determining if the component reached a functional level of steady-state operation rather than a set amount of time. For instance, if a component started but never reached a stable running state, it was classified as a failure to start rather than a failure to run, despite technically running for a short period of time. On the other hand, if a component had been started and reached a stable running state for any amount of time and then subsequently failed, it was classified as a failure to run event.

Additionally, when evaluating whether an event was a failure to start, consideration was given to repetitive attempts to start equipment. If there were multiple attempts to start equipment, and the equipment was successfully started within a short time period, without significant troubleshooting, the event was screened as a failure.

Finally, consideration was given as to whether an event was considered recoverable. Conversations were held with FLEX PM Coordinators on specific battery failures that could be corrected quickly (<15 minutes) and included actions that were proceduralized. Recovery could be explicitly credited in development of the failure rates, but for consistency with NUREG/CR-6928 [4] was not credited in this analysis.

---

[1] The start and end dates of the data from each utility differed somewhat due to differences in component in-service dates and dates of data reported to the PWROG. The specific start and end dates are used when it is necessary to estimate test demands. This reliability data generally represents the same time frame for all sites.

5.2 Failure Criteria and Identified Events

To systematically evaluate each identified event associated with FLEX equipment, a list of failure criteria was defined, as listed in Table 5-1 below:

Table 5-1: Failure Criteria Definitions

Failure Criteria ID	Component Failure Criteria	Notes
F1	Failure During Real Demand – Events where the component failed to function on demand (e.g., pump fails to start, valve fails to open) or failed to function over time (e.g., pump fail to run, valve fail to remain open), given a real demand. This includes start-failure events due to switch mis-positioning, even though the error might be easily identified and quickly restored.	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
F2	Failure During Test Demand – Events where the component failed to function on demand (e.g., pump fails to start, valve fails to open) or failed to function over time (e.g., pump fail to run, valve fail to remain open), given a test demand, where the component would have failed given a real demand. This includes start-failure events due to switch mis-positioning, even though the error might be easily identified and quickly restored.	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
F3	Degradation While Operating – Events where component degradation was identified while the component was successfully operating for real demand or for test but was manually shutdown because failure was imminent (e.g., pump bearing overheating). This also includes an event where continued operation would have created a condition that would challenge functionality (e.g., significant fuel oil leak in a DG skid with the potential to cause a fire).	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
F4	Degradation While in Standby – Events where component degradation was identified during tests or maintenance activities (as-found condition) or during other observations (e.g., operator walkdowns, engineering analysis) that would have prevented the component from functioning, given a real demand.	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
NF0	Duplicate Failure Reports – Failure reports that are repeats of previously reported failure events. These repeat failure reports often provide additional detail regarding the failure event but are not counted as a separate failure.	

Table 5-1: Failure Criteria Definitions		
Failure Criteria ID	**Component Failure Criteria**	**Notes**
NF1	Maintenance-Induced Failures – Events caused by maintenance activities where the component failure is identified before or just as the plant returns to an operational model where the component is required (e.g., EFW pump failure identified before return to Mode 1).	
NF2	Minor Degradation – Events caused by minor degradation in component performance, as long as the component would have been able to perform its mission with reasonable confidence. This include component performance slightly out of tolerance (e.g., DG start time slightly longer than 10 seconds, valve stroke time longer than design). It also includes minor maintenance issues even though the component may have been shut down to fix it, (e.g., pump shutdown for small oil leak, but the leak could be managed with the pump operating if it had been a real demand).	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
NF3	NOT USED.	NOT USED.
NF4	Not Applicable to Function – Events where the component failure is not applicable to its PRA function. This includes component failures that would have prevented component operation in the event of an emergency start demand (e.g., DG trip signals that are bypassed in emergency conditions), as long as the condition that caused the component trip is not indicative of an imminent failure. It also includes component inadvertent start, failure to trip on demand, or failure of other functions not applicable to its PRA function.	
NF5	Outside Component Boundary – Events where the component failures of a piece-part outside the component boundary (e.g., pump discharge valve, where the valve is not included in the pump component boundary).	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
NF6	Repetitive Problem Events – Subsequent failures within a short time interval from a single repetitive problem. The short time interval is taken to mean additional failures that occur within (roughly) one day of the previous failure. This does not apply to repetitive problems that occur over longer time intervals, even though it may be from the same cause.	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5]. This non-failure criterion is also based on supporting requirement DA-C5 of the ASME/ANS PRA Standard [Ref. 7].

Table 5-1: Failure Criteria Definitions		
Failure Criteria ID	**Component Failure Criteria**	**Notes**
NF7	Redundant Piece Part within Component Boundary – Failure of a piece part within the component boundary where a redundant part remains functional. For example, if a diesel generator has two redundant air start motors that are included in the diesel generator boundary definition, failure of one air start motor would not be counted as a failure of the diesel generator.	Similar criterion can be found in Section 5.2.3.1 of NUREG/CR-6823 [Ref. 5].
NF8	Event that occurred prior to the implementation of the site FLEX program and testing.	This failure criterion was added so that equipment events that occurred prior to the implementation of the FLEX program would not be counted as failures.

5.3 Calculation of Success Data

To appropriately estimate the failure parameters, the number of failures as well as the number of tests or run hours that have been performed are needed. The original request for information asked for the total number of demands and run-hours for each component provided. In general, run hours were provided as requested, but PM start dates and PM frequency were provided instead of number of demands. In most cases, the amount of time each piece of equipment was run during the PM was also provided. For demand data, the PM start dates, PM frequency, and date the data was received was used to calculate the number of demands. For run-hours, the input from the plant was used in most cases without performing any calculation. The exception to this are situations in which run-hours for each component were not provided, when this occurred, a calculation similar to that used for the demands was used.

The following equation was used to calculate the number of demands.

$$Total\ No.of\ Demands = \left\lfloor \frac{Recieved\ Date - PM\ Start\ Date}{PM\ Frequency} \right\rfloor + 1$$

The brackets in the equation are floor brackets and are used to represent the fact the term inside the brackets is rounded down to the nearest integer. Additionally, after the initial calculation an additional demand is added. This is done to ensure that the first test is counted in the number of demands. For example, given that a set of data was received on January 1, 2019, the PM for the component started on January 1, 2018, and the PM frequency was 6 months, as of January 1, 2019, it is assumed that 3 tests would have occurred (1/1/2018, 7/1/2018, and 1/1/2019).

Similarly, when calculating run hours, the following equation is used. Note that the expression is multiplied by the number of run hours per test to obtain a total number of run hours:

$$Total\ No.of\ Hours = \left(\left\lfloor \frac{Recieved\ Date - PM\ Start\ Date}{PM\ Frequency} \right\rfloor + 1\right) * No.Hours\ Per\ Test$$

Table 5-2 contains the results for the total number of demands and number of run-hours for each component type combined across the industry.

Note that in the calculation of success data, the shortest frequency tests (and associated duration) are used to estimate the total number of demands and run hours for a subset of equipment. It is standard industry practice to take credit for PMs to satisfy the shorter frequency tests. For example, a diesel generator may have a 6 Month and 36 Month PM. The 6 Month PM may be a short-duration run where the 36 Month PM may be an extended, full load test. When the 36 Month PM is performed, the 6 Month PM is considered to be satisfied. Therefore, the shorter frequency PM will provides a realistic estimate of the number of demands for each component, but may underestimate the number of run hours for each component. It is common that less frequent PMs are performed for longer durations (e.g., a 36 Month PM may run a piece of an equipment for 2 hours compared to a 6 Month PM where the equipment is only run for 30 minutes). Since the data is limited, excluding these additional run hours can result in conservative failure rate estimates. To provide more realistic run-hour success counts, the component run-hour count for each piece of equipment was increased by ten percent (see Assumption 9 in Section 7).

Table 5-2 contains a summary of the failure to start and failure to run by component group, as well as the respective number of demands and run hours by component group, across the industry.

Table 5-2: Industry Operating Experience

Component Type	# Components	# FTS Events	Demands	# FTR Events	Run Hours
Portable Diesel Generator					
Portable Diesel Generator (High Voltage)	9	2	68	0	22.0
Portable Diesel Generator (Medium Voltage)	157	64	1457	9	907.4
Portable Diesel Generator (Low Voltage)	14	1	24	0	7.3
Portable Diesel Generator (Total)	***180***	***67***	***1549***	***9***	***936.7***
Portable Combustion Turbine Generator					
Portable Combustion Turbine Generator (High Voltage)	7	1	110	1	60.5
Portable Combustion Turbine Generator (Medium Voltage)	13	6	116	1	73.7
Portable Combustion Turbine Generator (Total)	***20***	***7***	***226***	***2***	***134.2***
Portable Diesel-Driven Pump					
Portable Diesel-Driven Centrifugal Pump (High Pressure and Low Flow Rate)	21	11	134	2	51.7
Portable Diesel-Driven Centrifugal Pump (Medium to Low Pressure and Medium to Low Flow Rate)	121	20	1012	9	532.2
Portable Diesel-Driven Centrifugal Pump (Low Pressure and High Flow Rate)	94	42	945	5	427.3
Portable Diesel-Driven Centrifugal Pump (Subtotal)	*236*	*73*	*2091*	*16*	*1011.2*

Table 5-2: Industry Operating Experience

Component Type	# Components	# FTS Events	Demands	# FTR Events	Run Hours
Portable Diesel-Driven Positive Displacement Pump (High Pressure and Low Flow Rate)	9	0	67	0	37.1
Portable Diesel-Driven Positive Displacement Pump (Medium to Low Pressure and Medium to Low Flow Rate)	5	2	54	0	14.8
Portable Diesel-Driven Positive Displacement Pump (Low Pressure and High Flow Rate)	6	0	24	0	2.1
Portable Diesel-Driven Positive Displacement Pump (Subtotal)	*20*	*2*	*145*	0	*54.0*
Portable Diesel-Driven Pump (Total)	***256***	***75***	***2236***	***16***	***1065.2***
Portable Motor-Driven Positive Displacement Pump					
Portable Motor-Driven Positive Displacement Pump (High Pressure and Low Flow Rate)	44	1	353	2	153.0
Portable Motor-Driven Positive Displacement Pump (Medium to Low Pressure and Medium to Low Flow Rate)	29	2	65	1	39.3
Portable Motor-Driven Positive Displacement Pump (Low Pressure and High Flow Rate)	3	0	57	0	31.4
Portable Motor-Driven Positive Displacement Pump (Total)	***76***	***3***	***475***	***3***	***223.7***

Table 5-2: Industry Operating Experience

Component Type	# Components	# FTS Events	Demands	# FTR Events	Run Hours
Portable Air Compressor					
Portable Diesel-Driven Air Compressor	47	5	233	4	145.5
Portable Motor-Driven Air Compressor	7	1	30	0	14.3
Portable Air Compressor (Total)	***54***	***6***	***263***	***4***	***159.8***

.

5.4 Derivation of FLEX Equipment Unreliability Parameters

Three (3) approaches were defined to generate generic failure rate estimates, based on the amount of plant information available. These three methods are described below. The method in Section 5.4.2, Bayesian update using Jeffreys noninformative prior, was used as the default approach if the other methods could not be justified.

Prior to using any method, a pooling analysis is performed (see Appendix A). The pooling analysis involves a hypothesis test to determine if there is a statistically significant difference in the mean values between the sites being compared. This statistical test relies on the Chi-Squared "goodness-of-fit" test. A prerequisite of this test is that the expected number of events (for each defined category) needs to be greater than 1. In the cases where any of the categories have an expectation of 1 or less, the test will only provide approximate results, at best. As a second check, a Bayesian Chi-Squared test is used to determine how well a single failure rate predicts the observed data, if the Bayesian Chi-Squared shows that the failure rate reasonably predicts the data a JNI or CNID is used to develop failure rates for the set of equipment, otherwise the EB analysis is used. This validation technique is discussed in more detail in Section 5.5.

5.4.1 Empirical Bayes Approach

If a statistically significant difference in mean values was determine for a given set of data based on the pooling analysis performed in Appendix A, generic failure rate estimates were derived using the empirical Bayes (EB) approach. This approach effectively pools component reliability data at the plant-level, rather than pooling data at the industry level. A lengthy discussion of the EB analysis is discussed in Section 8.2 of [Ref. 5], but, in general, the process uses the data pooled at the plant-level to estimate the parameters of a distribution and eliminates the disadvantages (e.g., underestimating plant-to-plant variability) of performing a typical Bayesian update by pooling the data at the industry level.

The U.S. NRC maintains an EB calculator on the NROD (Nuclear Reliability and Operating Experience Database) website. This calculator implements the methodology discussed in [Ref. 5] and provides a convenient user interface to provide consistent results.

5.4.2 Bayesian Update Using Jeffreys Noninformative Prior

In cases where a statistically significant difference in mean values was not found, a different approach is required for parameter estimation. A Jeffreys noninformative prior (JNI) is used where sufficient operating experience exists.

For demand failures (λ_d), the JNI is a beta distribution with parameters α= 0.5 and β = 0.5 [Page 6-37, Ref. 5]; similarly, for run-time failure rates (λ_h), the JNI is a gamma distribution with α = 0.5 and β = 0. By using the Bayesian update formulas [Page 6-36 and Page 6-61, Ref. 5], the posterior means for a beta and gamma distribution can be determined.

$$\lambda_d = \frac{\alpha_{prior} + n_f}{\alpha_{prior} + n_f + \beta_{prior} + n_d - n_f} = \frac{n_f + 0.5}{n_d + 1}$$

$$\lambda_h = \frac{\alpha_{prior} + n_f}{t + \beta_{prior}} = \frac{n_f + 0.5}{t}$$

Where n_f = number of failure events, n_d = number of demands, and t = number of run-hours.

5.4.3 Bayesian Update Using Constrained Noninformative Prior

The third method considered for this analysis is implemented by specifying a mean value, based on some prior belief, but ensuring that the dispersion corresponds to ignorance (in some objective sense). The distribution matching these characteristics is known as the Constrained Noninformative Prior Distribution (CNID).

Per [Ref. 10], when prior knowledge is vague, it is often not worth the effort of defending an assumed prior distribution. From the viewpoint of simplicity and defensibility, a noninformative prior is ideal. From the viewpoint of realism, however, a noninformative prior can be defective. In some situations the analysts truly have prior information, leading to a well-justified informative prior distribution. In other cases, however, the analysts have only a vague idea of the realistic values. A desirable prior distribution would be consistent with such understanding, but would be otherwise uninformative. One mathematical implementation of this is to specify the mean of the distribution, but to choose a dispersion that corresponds to ignorance in some objective sense (i.e., the CNID).

For situations where the exposure is less than 50 (demands or operating hours), a CNID would be calculated and updated with the industry specific failure information to obtain posterior distribution parameters, and in effect, posterior mean values. The final grouping of components resulted in no equipment meeting the criteria to implement the CNID. However, the discussion of the CNID is retained as it may be used in future updates.

A comparison of these values was performed against the parameters obtained using a JNI. The comparison showed a negligible difference between the results except in one case, which provided an estimate significantly less than the parameters estimated when using a Jefferys noninformative prior.

The CNID mean values calculated are dependent on the factor used to "adjust" the mean values of the permanently installed equipment. Comments from the NRC Audit recommended a more rigorous attempt at validating the adjustment factors used. One example given was to compare failures rates derived in this analysis as compared to the failure rates for the permanently installed equipment. This comparison would give some general idea on the appropriateness of an adjustment factor. When performing the comparison noted above, this method was used to develop adjustment factors to be used in the CNID calculations.

Generic failure rates for the permanently installed equipment were used from the latest electronic version (i.e., 2015 data set) of NUREG/CR-6928 [Ref. 8].

Per Page 6-15 and 6-38 of [Ref. 5], the CNID can be determined as follows:

Gamma Distribution

$$\alpha_{prior} = 0.5$$

$$\beta_{prior} \text{ satisfies } \alpha_{prior}/\beta_{prior} = prior\ mean$$

Beta Distribution

The CNID for a Beta distribution is more complex. Table C.8 of [Ref. 5] provides a table of values that can be interpolated to determine the parameters of a beta distribution given the prior mean. See Page 6-38 of [Ref. 5] for more information on this process.

5.5 Model Validation

Model validation is an important aspect of any data analysis and serves as a means to ensure that the selected models accurately replicate the observed data. For this analysis, the main tool used to validate the models is hypothesis testing via use of the Chi-Squared Test. Details of the Chi-Squared Test are provided in Section 6.2.3.1.2 of [Ref. 5] and it is used to study whether the failure rate parameters are the same or different between plants. The test assumes a constant failure rate and evaluates how the observed failures compare to the expected amount of failures. If the difference in the observed and expected failures is small, the analyst can reasonably expect a pooled data approach would reflect reasonable parameter estimates. If the differences in the observed and expected failures are large, the analyst should not expect a pooled data approach to reflect reasonable parameter estimates and should evaluate a different underlying model.

In the situation in which there is significant difference in the observed and expected failures, the empirical Bayes (EB) analysis is used (see Section 5.4.1 for more information on the EB approach). In short, the EB analysis is used to estimate parameters when there is sufficient variability in the observed data between plants.

Following the process above is generally straight forward when the number of failures and exposures are large. However, in the case where there is limited data, such as in this analysis, this process leads the analyst to reject any other models based on the limited data. More specifically, some specific criteria of the Chi-Squared Test used in this analysis are: (1) with equiprobable cells, the average expected frequency should be at least 1.0 when testing at the 0.05 significant level and (2) when the cells are not approximately equiprobable, the average expected frequency should be double (i.e., 2.0). As shown in Appendix A, no data sets are approximately equiprobable, and they also do not satisfy the second criterion. To mitigate this issue, a *Bayesian* Chi-Squared Test can be implemented. The main benefit of the Bayesian Chi-Squared Test is that there is no need to bin the associated data, and therefore the criteria stated above are irrelevant. Section 4.3.1 of [Ref. 11] discusses the Bayesian Chi-Squared Test in detail but key points are summarized here.

1. In the frequentist Chi-Squared Test, typically a *p-value* of 0.05 is used to determine whether or not the data is in conflict with the model.
2. In the Bayesian framework, a similar approach can be used, but to be more convenient, the *p-value* will be used to identify if the model appropriately replicates the observed data. In this method, the model with Bayesian *p-value* closest to 0.5, which is the value one would obtain if the distributions of the observed and replicated test statistics overlapped perfectly, is used.

For this analysis, if the Bayesian *p-value* is ≤ 0.15 or ≥ 0.85, the data will be processed through an EB analysis, as the Bayesian *p-value* at the extremes suggests that the proposed model (i.e., pooling the data) will not replicate the data. To perform the Bayesian Chi-Squared Test, the *winBUGS*® software is used. A tutorial of *winBUGS*® is not provided here, but [Ref. 11] provides a brief overview. Appendix A discusses the results of both the frequentist Chi-Squared Test and the Bayesian Chi-Squared Test.

6 Failure Rates for FLEX Equipment

The generic failure rates for FLEX equipment are summarized in Table 6-1 below.

Table 6-1: Generic Failures Rates for FLEX Equipment

FLEX Component	Failure Mode	Distribution	Mean	α	β	Method[1]
Portable Diesel Generator						
Portable Diesel Generator	Fail to Run	Gamma	1.03E-02	0.856	82.9	EB
	Fail to Start	Beta	4.35E-02	67.5	1482.5	JNI
Portable Combustion Turbine Generator						
Portable Combustion Turbine Generator	Fail to Run	Gamma	1.86E-02	2.5	134.2	JNI
	Fail to Start	Beta	3.30E-02	7.5	219.5	JNI
Portable Diesel-Driven Pump						
Portable Diesel-Driven Pump	Fail to Run	Gamma	1.55E-02	16.5	1065.2	JNI
	Fail to Start	Beta	3.38E-02	75.5	2161.5	JNI

Table 6-1: Generic Failures Rates for FLEX Equipment

FLEX Component	Failure Mode	Distribution	Mean	α	β	Method(1)
Portable Motor-Driven Positive Displacement Pump						
Portable Motor-Driven Positive Displacement Pump	Fail to Run	Gamma	1.56E-02	3.5	223.7	JNI
	Fail to Start	Beta	7.35E-03	3.5	472.5	JNI
Portable Air Compressor						
Portable Air Compressor	Fail to Run	Gamma	2.82E-02	4.5	159.8	JNI
	Fail to Start	Beta	2.46E-02	6.5	257.5	JNI

Notes:

(1) JNI – Bayesian update using Jeffreys Noninformative Prior, EB – Empirical Bayes Analysis, CNID – Bayesian update using Constrained Noninformative Prior Distribution.

7 Assumptions and Sources of Uncertainty

1. **Assumption:** It is assumed that PM activities are performed on the date predicted by the interval and start date of the PM. For example, if a PM for a component was started on 1/1/2018, and has a 6M frequency, it is assumed that the second test is performed on 7/1/2018. This assumption is used in calculation of success data.

   Justification: This assumption is realistic as FLEX equipment is required to be tested within a certain period of time consistent with the interval defined.

2. **Assumption:** When run hour data was provided for the entire component grouping from a site, the value provided for the group was divided equally amongst the components in the group.

   Justification: These components have been used for testing and maintenance purposes only. While these components may not run for the exact same amount of time each time they are run, they are tested at the same intervals via a common procedure. Given a total sum of run-hours for a group, it would be reasonable to assume that the run-hours were accumulated relatively evenly amongst the components within the group.

3. **Assumption:** In a few cases, run-hour data was not provided as part of the utility response. If the utility provided useable information for the other categories requested, estimates of run-hours were made based on the type and frequency of PMs performed.

   Justification: Based upon observations from the utilities reporting run-hours, it appears that the maintenance runs performed on FLEX equipment is relatively short in duration. Generally, for PMs with a frequency of less than 1 year, a value of 0.5 run hours was applied for each PM. For PMs with a frequency of 1 year or more, a value 1.0 run hours were applied for each PM. These estimates align with the PM durations specified in the EPRI FLEX PM database and are considered representative.

4. **Assumption:** In the absence of a clear start date for the establishment of PMs, the analysis assumed that the PM start date was January 1 of the first year in which an identified event occurred. If no events were identified, a conservative estimate was made on the start date of the PM.

   Justification: It is recognized that this assumption may have the effect of omitting some successful demands by beginning the count with the first year in which a failure was recorded. In the case of this data set, generating an artificially high failure rate due to undercounting is judged to be preferable and more defensible in applications as compared to generating an artificially low failure rate due to over counting. Additionally, the descriptions of the identified events make it appear that it was common that utilities experienced failures early in the life of their testing and maintenance runs. Given that insight, it seemed unlikely that a utility would have experienced a significant amount of demands and run hours without a failure and much more likely that they experienced a failure in their first year of testing and maintenance.

5. **Assumption:** The criterion for using a CNID as the prior distribution for failure rate estimates is where data sets have less than 50 demands or less than 50 hours of operating experience.

Justification: When prior distributions have little population basis, they can be difficult to construct. Typically, in these situations, a noninformative prior (such as the JNI) is used such that inferences are unaffected by information external to the available data. In the presence of weak data, the noninformative prior can be too influential on the failure rate estimate. To eliminate this impact, the cutoff for use of the JNI is selected. This cutoff is based on engineering judgement. It could be shown that a subset of additional information would still be considered weak (e.g., 100 demands, and 200 hours of operating experience) but an attempt to limit the use of a CNID is maintained by the current criteria.

6. **Assumption:** It is assumed that infrequent PMs are used to satisfy frequent PMs. For example, if a portable diesel generator has a 6M and 36M PM, when the 36M PM is performed, it is credited as satisfying the requirement of the 6M PM. In this scenario, only one demand is counted instead of two (and similarly for operating hours).

   Justification: This is standard practice for portable equipment. Furthermore, a sampling of utilities was performed to confirm this approach. It is recognized that from this sampling, some utilities may perform multiple PMs; however, the approach taken is conservative and consistent with the resolution of information provided.

7. **Assumption:** It is assumed that the distribution developed using the JNI is an adequate representation of the uncertainty associated with FLEX equipment.

   Justification: It is recognized that the uncertainty distributions associated with the failure rate estimates derived from using a JNI are generally narrow and may not adequately address the state-of-knowledge associated with FLEX equipment. However, the use of a JNI in situations where limited data are available is standard practice and no good alternative has been identified.

8. **Assumption:** When the Bayesian Chi-Squared test is used (See Section 5.5, and Appendix A), a Bayesian p-value ≤ 0.15 or ≥ 0.85 is assumed to represent a significantly deviated *p-value* such that the observed data are not adequately predicted through the use of a pooled data approach using the Jeffreys non-informative prior distribution.

   Justification: As discussed in Section 4.3.1 of [Ref. 11], Bayesian *p-values* near 0 or 1 are usually indicative of a problem with the assumed model. If a model with a single failure rate is adequate across the sites, the Bayesian *p-value* would be near 0.5. The values of 0.15 and 0.85 were selected to expand the range conservatively. More realistically, values much closer to 0.5 would be expected given that the assumed model could predict the observed data more reliably.

9. **Assumption:** Additional run-hour success data is added to each component run-hour estimate to account for the longer duration of less frequent PMs that are not used to calculate success data.

   Justification: 10% of each component run-hour estimate was added to the component run-hour success data to account for less frequent, and longer duration PM activities. Typically, less frequent PM activities are performed at longer durations than the more frequent PM activities. For example, 6M and 12M PM activities for a PDG may only be performed for 0.5 hours, as opposed to a 36M PM activity which may run the equipment for 2 hours. The algorithm used to calculate run-hour success data uses the shortest frequency PM, as well as the duration for that specific PM activity, to calculate estimates of success data. In general, this is determined to be an appropriate treatment because not all of the less frequent PM activities are performed for longer durations. A concern for under counting run-hour success data arises when the duration of less frequent PM activities differ from the more frequent PM activities. To determine a factor to include to each run-hour success data estimate, first, the minimum frequency of each PM activity was obtained. Second, duplicate PMs were found that were less frequent and had a different duration than the minimum PM activity. Comparing the additional run-hour success data from the less frequent PM activities to the estimates calculated using the minimum PM frequency showed that adding an additional 10% of the total run-hour estimates brought the run-hour estimates closer to the real accumulated operating experience.

8 References

1. EA-12-049, "Issuance of Order to Modify Licenses with Regard to Requirements for Mitigation Strategies for Beyond-Design-Basis External Events," U.S. Nuclear Regulatory Commission, March 12, 2012.
2. NEI 12-06, Revision 4, "Diverse and Flexible Coping Strategies (FLEX) Implementation Guide," Nuclear Energy Institute, December 2016.
3. NS-ES-0237, "FLEX Implementation," Westinghouse Electric Company, July 2012.
4. NUREG/CR-6928, "Industry-Average Performance for Components and Initiating Events at U.S. Commercial Nuclear Power Plants," U.S. Nuclear Regulatory Commission, February 2007.
5. NUREG/CR-6823, "Handbook of Parameter Estimation for Probabilistic Risk Assessment," U.S. Nuclear Regulatory Commission, September 2003.
6. PWROG-14003-P, Revision 0, "Implementation of FLEX Equipment in Plant-Specific PRA Models," March 2015.
7. ASME/ANS RA-Sa-2009, "Addenda to ASME/ANS RA-S-2008 Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications," February 2009.
8. U.S. Nuclear Regulatory Commission, "Component Reliability Data Sheets, 2015 Update," https://nrcoe.inl.gov/resultsdb/publicdocs/AvgPerf/CompnentUR2015.pdf/, February 2017.
9. NUREG/CR-1475, Revision 1, "Applying Statistics," U.S. Nuclear Regulatory Commission, March 2011.
10. Atwood, Corwin L., "Constrained noninformative priors in risk assessment." *Reliability Engineering and System Safety,* Vol. 53, 1996, Pages 37-46.
11. Kelly, D., and Smith, C, "Bayesian Inference for Probabilistic Risk Assessment – A Practitioner's Guidebook," ISBN 978-1-84996-186-8, 2011.

APPENDIX A – Pooling Analysis

Appendix A documents the result of the pooling analysis performed on each subset of data. For more information on the pooling analysis, see Section 5.5. Table A-1 summarizes the results of the detailed pooling analyses performed below. The conclusions of the pooling analysis are based upon the *p-values* calculated using the Bayesian Chi-Squared test rather than the frequentist Chi-Squared test as the Bayesian Chi-Squared test is a better tool for selecting a model based on limited data.

Table A-1: Pooling Analysis Summary

Component Type	**Failure Mode**	**Pooling Analysis Results**[1]	
		Frequentist Chi-Squared	**Bayesian Chi-Squared**
Portable Combustion Turbine Generator	Fail to Start	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared results in a p-value of 0.393. This value reasonably suggests that a pooled prior can accurately represent the observed data.
	Fail to Run	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared results in a p-value of 0.663. This value reasonably suggests that a pooled prior can accurately represent the observed data.
Portable Diesel Generator	Fail to Start	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared results in a p-value of 0.548. This value reasonably suggests that a pooled prior can accurately represent the observed data.

[1] The pooling analysis is the first step in the statistical process and determines if a component group should be processed through the empirical Bayes analysis. It is possible that the empirical Bayes analysis may fail, and in those situations, a different methodology may still be used to derive failure rates for the FLEX equipment.

Table A-1: Pooling Analysis Summary			
Component Type	Failure Mode	Pooling Analysis Results[1]	
		Frequentist Chi-Squared	Bayesian Chi-Squared
	Fail to Run	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 2.22E-04. This highly suggests that a single λ cannot model the variability in the observed data.
Portable Air Compressor	Fail to Start	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 0.301. This value reasonably suggests that a pooled prior can accurately represent the observed data.
	Fail to Run	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 0.444. This value reasonably suggests that a pooled prior can accurately represent the observed data.
Portable Diesel-Driven Pump	Fail to Start	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 0.144. This value suggests that there may be problems with the pooled prior; however, since it is borderline, a pooled data approach will continue to be used.

Table A-1: Pooling Analysis Summary			
Component Type	Failure Mode	Pooling Analysis Results[1]	
		Frequentist Chi-Squared	Bayesian Chi-Squared
	Fail to Run	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 7.78E-03. This highly suggests that a single λ cannot model the variability in the observed data.
Portable Motor-Driven Positive Displacement Pump	Fail to Start	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 0.343. This value reasonably suggests that a pooled prior can accurately represent the observed data.
	Fail to Run	The data do not provide enough evidence to reject the null hypothesis (i.e., poolability) – Pooled Data	The Bayesian Chi-Squared Test results in a p-value of 0.313. This value reasonably suggests that a pooled prior can accurately represent the observed data.

A.1 PORTABLE COMBUSTION TURBINE GENERATOR

A.1.1 Fail to Start

Site	Component Type	Failures	Demands
NPP-S35	PCTG FTS	3	36
NPP-S56	PCTG FTS	3	74
NPP-S36	PCTG FTS	0	24
NPP-S49	PCTG FTS	0	14
NPP-S37	PCTG FTS	0	48
NPP-S50	PCTG FTS	0	4
NPP-S6	PCTG FTS	1	26

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
6	3.10E-02	6.24	12.59	3.97E-01

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	3.93E-01	4.89E-01	4.81E-03

Summary:

The Chi-Squared test shows that there is not a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test shows a *p-value* of 0.393 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.1.2 Fail to Run

Site	Component Type	Failures	Exposure
NPP-S35	PCTG FTR	1	19.8
NPP-S56	PCTG FTR	1	40.7
NPP-S36	PCTG FTR	0	13.2
NPP-S49	PCTG FTR	0	15.4
NPP-S37	PCTG FTR	0	26.4
NPP-S50	PCTG FTR	0	4.4
NPP-S6	PCTG FTR	0	14.3

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
6	1.49E-02	3.04	12.59	8.04E-01

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	6.63E-01	4.73E-01	4.99E-03

Summary:

The Chi-Squared test shows that there is not a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test shows a *p-value* of 0.663 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.2 PORTABLE DIESEL GENERATOR

A.2.1 Fail to Start

Site	Component Type	Failures	Demands
NPP-S15	PDG FTS	0	16
NPP-S21	PDG FTS	1	13
NPP-S8	PDG FTS	1	112
NPP-S22	PDG FTS	8	28
NPP-S1	PDG FTS	0	8
NPP-S23	PDG FTS	1	28
NPP-S9	PDG FTS	2	51
NPP-S24	PDG FTS	1	16
NPP-S14	PDG FTS	4	67
NPP-S39	PDG FTS	0	104
NPP-S40	PDG FTS	2	20
NPP-S48	PDG FTS	1	67
NPP-S38	PDG FTS	2	40
NPP-S25	PDG FTS	0	14
NPP-S41	PDG FTS	2	28
NPP-S52	PDG FTS	1	12
NPP-S7	PDG FTS	1	26
NPP-S16	PDG FTS	0	14
NPP-S10	PDG FTS	1	18
NPP-S17	PDG FTS	2	22
NPP-S53	PDG FTS	1	35
NPP-S27	PDG FTS	2	7
NPP-S28	PDG FTS	0	22
NPP-S11	PDG FTS	0	63
NPP-S3	PDG FTS	0	18
NPP-S46	PDG FTS	2	30
NPP-S29	PDG FTS	1	24
NPP-S4	PDG FTS	3	18
NPP-S12	PDG FTS	0	60
NPP-S30	PDG FTS	1	10
NPP-S18	PDG FTS	0	18
NPP-S2	PDG FTS	3	112
NPP-S31	PDG FTS	2	13
NPP-S42	PDG FTS	2	10
NPP-S47	PDG FTS	0	27
NPP-S32	PDG FTS	0	5
NPP-S33	PDG FTS	0	12
NPP-S19	PDG FTS	1	63

Site	Component Type	Failures	Demands
NPP-S43	PDG FTS	2	21
NPP-S57	PDG FTS	0	24
NPP-S13	PDG FTS	0	34
NPP-S51	PDG FTS	0	29
NPP-S44	PDG FTS	5	27
NPP-S5	PDG FTS	0	14
NPP-S34	PDG FTS	0	4
NPP-S45	PDG FTS	3	35
NPP-S6	PDG FTS	3	47
NPP-S54	PDG FTS	1	11
NPP-S20	PDG FTS	2	14
NPP-S58	PDG FTS	1	36
NPP-S59	PDG FTS	2	2

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
50	4.33E-02	158.47	67.50	3.34E-13

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	5.48E-01	4.98E-01	5.06E-03

Summary:

The Chi-Squared test shows that there is a statistically significant difference in the mean values between plants (at the 5% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 2.0. where the cells are not approximately equiprobable. The Bayesian Chi-Squared Test shows a *p-value* of 0.548 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.2.2 Fail to Run

Site	Component Type	Failures	Exposure
NPP-S15	PDG FTR	0	8.8
NPP-S21	PDG FTR	0	3.575
NPP-S8	PDG FTR	0	92.4
NPP-S22	PDG FTR	0	7.7
NPP-S1	PDG FTR	0	4.4
NPP-S23	PDG FTR	0	7.7
NPP-S9	PDG FTR	0	28.05
NPP-S24	PDG FTR	0	4.4
NPP-S14	PDG FTR	1	145.97
NPP-S39	PDG FTR	0	28.6
NPP-S40	PDG FTR	0	11
NPP-S48	PDG FTR	0	18.425
NPP-S38	PDG FTR	0	22
NPP-S25	PDG FTR	0	3.85
NPP-S41	PDG FTR	0	15.4
NPP-S52	PDG FTR	0	6.6
NPP-S7	PDG FTR	0	14.3
NPP-S16	PDG FTR	0	3.85
NPP-S10	PDG FTR	0	9.9
NPP-S17	PDG FTR	0	12.1
NPP-S53	PDG FTR	0	19.25
NPP-S27	PDG FTR	0	1.925
NPP-S28	PDG FTR	0	7.7
NPP-S11	PDG FTR	0	34.65
NPP-S3	PDG FTR	0	9.9
NPP-S46	PDG FTR	1	33
NPP-S29	PDG FTR	0	6.6
NPP-S4	PDG FTR	0	9.9
NPP-S12	PDG FTR	0	33
NPP-S30	PDG FTR	0	2.75
NPP-S18	PDG FTR	0	9.9
NPP-S2	PDG FTR	1	61.6
NPP-S31	PDG FTR	0	3.575
NPP-S42	PDG FTR	0	11
NPP-S47	PDG FTR	0	44.55
NPP-S32	PDG FTR	0	1.375
NPP-S33	PDG FTR	1	3.3
NPP-S19	PDG FTR	1	17.325
NPP-S43	PDG FTR	0	11.55
NPP-S57	PDG FTR	0	9.9

Site	Component Type	Failures	Exposure
NPP-S13	PDG FTR	0	18.7
NPP-S51	PDG FTR	0	15.95
NPP-S44	PDG FTR	0	14.85
NPP-S5	PDG FTR	0	7.7
NPP-S34	PDG FTR	0	1.1
NPP-S45	PDG FTR	0	9.625
NPP-S6	PDG FTR	2	51.7
NPP-S54	PDG FTR	0	4.554
NPP-S20	PDG FTR	0	15.4
NPP-S58	PDG FTR	0	14.85
NPP-S59	PDG FTR	2	0.55

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
50	9.61E-03	799.13	67.50	0

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	2.22E-04	1.49E-02	1.55E-04

Summary:

The Chi-Squared test shows that there is a statistically significant difference in the mean values between plants (at the 95% confidence level).). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test results in a *p-value* of 2.22E-04 which highly suggests that a pooled data approach will not predict the observed parameters. Therefore, the data will be processed through an EB analysis (see Appendix C.1).

.

A.3 PORTABLE AIR COMPRESSOR

A.3.1 Fail to Start

Site	Component Type	Failures	Demands
NPP-S1	PAC FTS	0	8
NPP-S40	PAC FTS	1	10
NPP-S48	PAC FTS	3	36
NPP-S52	PAC FTS	0	12
NPP-S7	PAC FTS	0	10
NPP-S49	PAC FTS	1	4
NPP-S17	PAC FTS	1	16
NPP-S3	PAC FTS	0	24
NPP-S29	PAC FTS	0	4
NPP-S4	PAC FTS	0	35
NPP-S18	PAC FTS	0	18
NPP-S37	PAC FTS	0	3
NPP-S33	PAC FTS	0	13
NPP-S5	PAC FTS	0	28
NPP-S6	PAC FTS	0	24
NPP-S54	PAC FTS	0	12
NPP-S59	PAC FTS	0	6

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
16	2.28E-02	23.04	26.30	1.13E-01

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	3.01E-01	4.59E-01	4.75E-03

Summary:

The Chi-Squared test shows that there is not a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-

Squared Test shows a *p-value* of 0.301 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.3.2 Fail to Run

Site	Component Type	Failures	Exposure
NPP-S1	PAC FTR	0	4.4
NPP-S40	PAC FTR	0	5.5
NPP-S48	PAC FTR	2	29.7
NPP-S52	PAC FTR	0	13.2
NPP-S7	PAC FTR	0	5.5
NPP-S49	PAC FTR	0	1.1
NPP-S17	PAC FTR	0	8.8
NPP-S3	PAC FTR	0	14.85
NPP-S29	PAC FTR	0	1.1
NPP-S4	PAC FTR	0	22
NPP-S18	PAC FTR	0	9.9
NPP-S37	PAC FTR	0	0.825
NPP-S33	PAC FTR	0	3.85
NPP-S5	PAC FTR	0	17.6
NPP-S6	PAC FTR	2	13.2
NPP-S54	PAC FTR	0	6.6
NPP-S59	PAC FTR	0	1.65

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
16	2.50E-02	13.48	26.30	6.37E-01

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	4.44E-01	4.97E-01	4.79E-03

Summary:

The Chi-Squared test shows that there is a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test shows a *p-value* of 0.444 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.4 PORTABLE DIESEL-DRIVEN PUMP

A.4.1 Fail to Start

Site	Component Type	Failures	Demands
NPP-S15	PDDP FTS	0	64
NPP-S35	PDDP FTS	2	36
NPP-S21	PDDP FTS	1	53
NPP-S56	PDDP FTS	6	142
NPP-S8	PDDP FTS	0	57
NPP-S22	PDDP FTS	0	42
NPP-S1	PDDP FTS	0	16
NPP-S23	PDDP FTS	0	58
NPP-S9	PDDP FTS	0	107
NPP-S24	PDDP FTS	1	16
NPP-S14	PDDP FTS	1	12
NPP-S39	PDDP FTS	0	78
NPP-S40	PDDP FTS	0	10
NPP-S36	PDDP FTS	1	20
NPP-S48	PDDP FTS	3	93
NPP-S38	PDDP FTS	1	42
NPP-S25	PDDP FTS	0	12
NPP-S41	PDDP FTS	0	27
NPP-S52	PDDP FTS	0	4
NPP-S7	PDDP FTS	0	48
NPP-S16	PDDP FTS	0	14
NPP-S10	PDDP FTS	0	104
NPP-S49	PDDP FTS	1	7
NPP-S17	PDDP FTS	3	92
NPP-S53	PDDP FTS	0	25
NPP-S27	PDDP FTS	0	5
NPP-S28	PDDP FTS	0	5
NPP-S11	PDDP FTS	0	120
NPP-S3	PDDP FTS	2	48
NPP-S46	PDDP FTS	1	50

Site	Component Type	Failures	Demands
NPP-S29	PDDP FTS	4	40
NPP-S4	PDDP FTS	3	63
NPP-S12	PDDP FTS	4	138
NPP-S30	PDDP FTS	2	10
NPP-S18	PDDP FTS	0	18
NPP-S2	PDDP FTS	1	40
NPP-S31	PDDP FTS	1	14
NPP-S42	PDDP FTS	4	20
NPP-S47	PDDP FTS	2	28
NPP-S32	PDDP FTS	0	2
NPP-S33	PDDP FTS	4	16
NPP-S19	PDDP FTS	1	27
NPP-S50	PDDP FTS	2	16
NPP-S43	PDDP FTS	2	12
NPP-S57	PDDP FTS	6	30
NPP-S13	PDDP FTS	0	33
NPP-S44	PDDP FTS	4	63
NPP-S5	PDDP FTS	1	74
NPP-S45	PDDP FTS	0	56
NPP-S6	PDDP FTS	0	63
NPP-S20	PDDP FTS	0	28
NPP-S58	PDDP FTS	11	30
NPP-S59	PDDP FTS	0	8

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
52	3.35E-02	230.98	69.83	0

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	1.44E-01	3.51E-01	3.80E-03

Summary:

The Chi-Squared test shows that there is a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 2.0, where the cells are not approximately equiprobable. The Bayesian Chi-Squared Test shows a *p-value* of 0.144 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.4.2 Fail to Run

Site	Component Type	Failures	Exposure
NPP-S15	PDDP FTR	1	35.2
NPP-S35	PDDP FTR	0	3.168
NPP-S21	PDDP FTR	0	14.575
NPP-S56	PDDP FTR	0	39.05
NPP-S8	PDDP FTR	0	31.35
NPP-S22	PDDP FTR	0	11.55
NPP-S1	PDDP FTR	0	8.8
NPP-S23	PDDP FTR	0	15.95
NPP-S9	PDDP FTR	0	58.85
NPP-S24	PDDP FTR	2	4.4
NPP-S14	PDDP FTR	0	14.828
NPP-S39	PDDP FTR	0	23.595
NPP-S40	PDDP FTR	1	5.5
NPP-S36	PDDP FTR	0	1.76
NPP-S48	PDDP FTR	1	25.575
NPP-S38	PDDP FTR	1	23.1
NPP-S25	PDDP FTR	0	6.6
NPP-S41	PDDP FTR	0	14.85
NPP-S52	PDDP FTR	0	2.2
NPP-S7	PDDP FTR	0	26.4
NPP-S16	PDDP FTR	0	3.85

Site	Component Type	Failures	Exposure
NPP-S10	PDDP FTR	0	57.2
NPP-S49	PDDP FTR	0	1.925
NPP-S17	PDDP FTR	0	50.6
NPP-S53	PDDP FTR	0	13.75
NPP-S27	PDDP FTR	0	1.375
NPP-S28	PDDP FTR	0	11
NPP-S11	PDDP FTR	0	66
NPP-S3	PDDP FTR	0	33
NPP-S46	PDDP FTR	0	55
NPP-S29	PDDP FTR	1	11
NPP-S4	PDDP FTR	1	34.65
NPP-S12	PDDP FTR	0	51.15
NPP-S30	PDDP FTR	1	2.75
NPP-S18	PDDP FTR	0	9.9
NPP-S2	PDDP FTR	0	22
NPP-S31	PDDP FTR	0	3.85
NPP-S42	PDDP FTR	0	22
NPP-S47	PDDP FTR	1	30.8
NPP-S32	PDDP FTR	0	0.55
NPP-S33	PDDP FTR	1	4.4
NPP-S19	PDDP FTR	0	7.425
NPP-S50	PDDP FTR	0	4.4
NPP-S43	PDDP FTR	0	4.4
NPP-S57	PDDP FTR	2	8.25
NPP-S13	PDDP FTR	0	18.15
NPP-S44	PDDP FTR	0	69.3
NPP-S5	PDDP FTR	0	40.7
NPP-S45	PDDP FTR	0	15.4

Site	Component Type	Failures	Exposure
NPP-S6	PDDP FTR	0	17.325
NPP-S20	PDDP FTR	0	15.4
NPP-S58	PDDP FTR	0	8.25
NPP-S59	PDDP FTR	3	2.2

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
52	1.50E-02	418.13	69.83	0

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	7.78E-03	8.79E-02	8.92E-04

Summary:

The Chi-Squared test shows that there is not a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test results in a *p-value* of 7.78E-03 which highly suggests that a pooled data approach will not predict the observed parameters. Therefore, the data will be processed through an EB analysis (see Appendix C.2).

A.5 PORTABLE MOTOR-DRIVEN POSITIVE DISPLACEMENT PUMP

A.5.1 Fail to Start

Site	Component Type	Failures	Demands
NPP-S35	PMDPDP FTS	0	4
NPP-S1	PMDPDP FTS	0	4
NPP-S9	PMDPDP FTS	0	57
NPP-S39	PMDPDP FTS	0	39
NPP-S36	PMDPDP FTS	0	2
NPP-S48	PMDPDP FTS	0	19

Site	Component Type	Failures	Demands
NPP-S10	PMDPDP FTS	0	52
NPP-S49	PMDPDP FTS	0	2
NPP-S53	PMDPDP FTS	0	18
NPP-S11	PMDPDP FTS	0	60
NPP-S12	PMDPDP FTS	1	60
NPP-S2	PMDPDP FTS	0	52
NPP-S33	PMDPDP FTS	0	3
NPP-S50	PMDPDP FTS	0	15
NPP-S13	PMDPDP FTS	0	34
NPP-S34	PMDPDP FTS	0	8
NPP-S54	PMDPDP FTS	2	40
NPP-S20	PMDPDP FTS	0	2
NPP-S59	PMDPDP FTS	0	4

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
18	6.32E-03	15.47	28.87	6.29E-01

Bayesian Chi-Squared Test Results

Parameter	Mean	Standard Deviation	MC Error
p-value	3.43E-01	4.75E-01	5.06E-03

Summary:

The Chi-Squared test shows that there is not a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test shows a *p-value* of 0.343 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.5.2 Fail to Run

Site	Component Type	Failures	Exposure
NPP-S35	PMDPDP FTR	0	2.2

Site	Component Type	Failures	Exposure
NPP-S1	PMDPDP FTR	0	2.2
NPP-S9	PMDPDP FTR	0	31.35
NPP-S39	PMDPDP FTR	0	7.1643
NPP-S36	PMDPDP FTR	0	1.1
NPP-S48	PMDPDP FTR	0	5.225
NPP-S10	PMDPDP FTR	0	28.6
NPP-S49	PMDPDP FTR	0	2.2
NPP-S53	PMDPDP FTR	1	9.9
NPP-S11	PMDPDP FTR	0	33
NPP-S12	PMDPDP FTR	0	16.5
NPP-S2	PMDPDP FTR	0	28.6
NPP-S33	PMDPDP FTR	0	0.825
NPP-S50	PMDPDP FTR	0	11.55
NPP-S13	PMDPDP FTR	0	18.7
NPP-S34	PMDPDP FTR	0	0.88
NPP-S54	PMDPDP FTR	2	22
NPP-S20	PMDPDP FTR	0	0.55
NPP-S59	PMDPDP FTR	0	1.1

DOF	MLE	X^2 Statistic	X^2 Critical (α=0.05)	p-value
18	1.34E-02	18.08	28.87	4.50E-01

Bayesian Chi-Squared Test Results

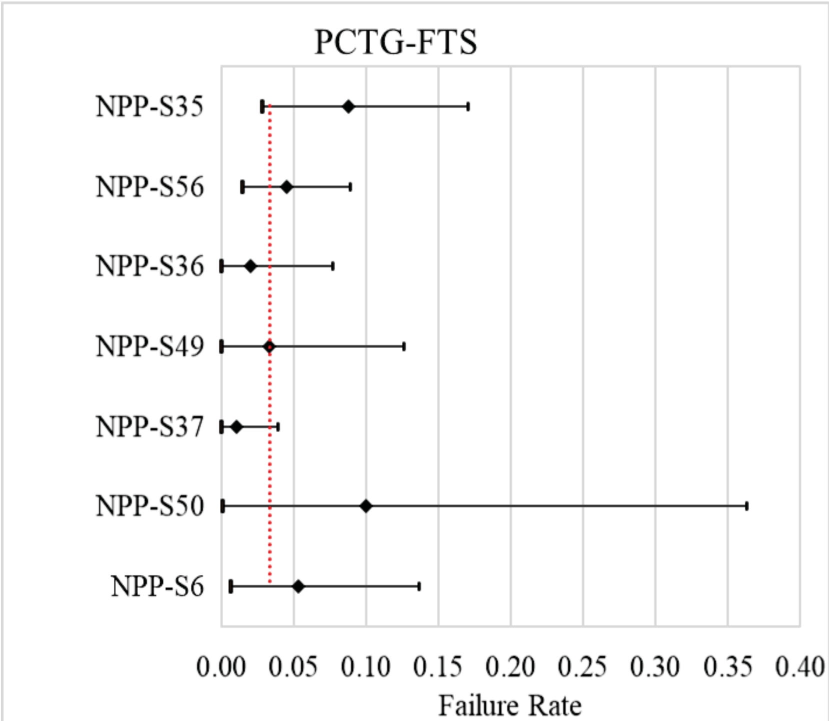
Parameter	Mean	Standard Deviation	MC Error
p-value	3.13E-01	4.64E-01	4.77E-03

Summary:

The Chi-Squared test shows that there is a statistically significant difference in the mean values between plants (at the 95% confidence level). The sample size of sites is such that the Chi-Squared approximation may not be appropriate as shown by the average expected frequency of less than 1.0. The Bayesian Chi-Squared Test shows a *p-value* of 0.313 which suggests that a pooled data approach reasonably predicts the observed data. Therefore, the data do not provide enough evidence to reject the null hypothesis (i.e., poolability), and the data will be pooled.

A.6 Posterior 90% Credible Intervals Plots

The following plots show the posterior 90% credible intervals (also known as caterpillar plots) for the failure rate for each site, component type, and failure mode, based on updating the JNI prior with the data from each source. The intent of these plots is to provide a visual comparison of each site as compared to a mean value developed from pooling the industry data (which is displayed as a dashed red line). These plots help assist in determining whether there may be extra variation amongst the sites such that a model with a single failure rate may not be adequate. The plots are used as additional confirmation to the Chi-Squared test results. The PDG-FTR and PDDP-FTR caterpillar plots show significant variability across each site.



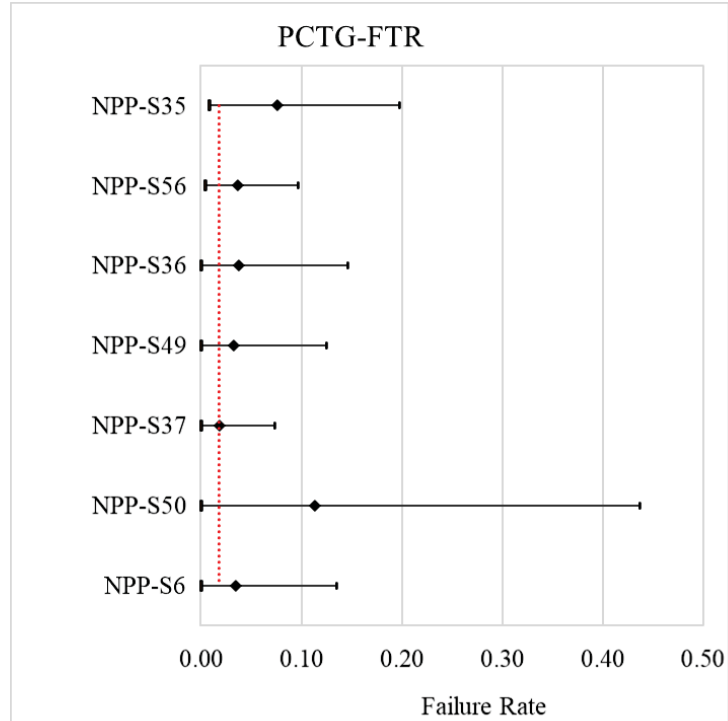
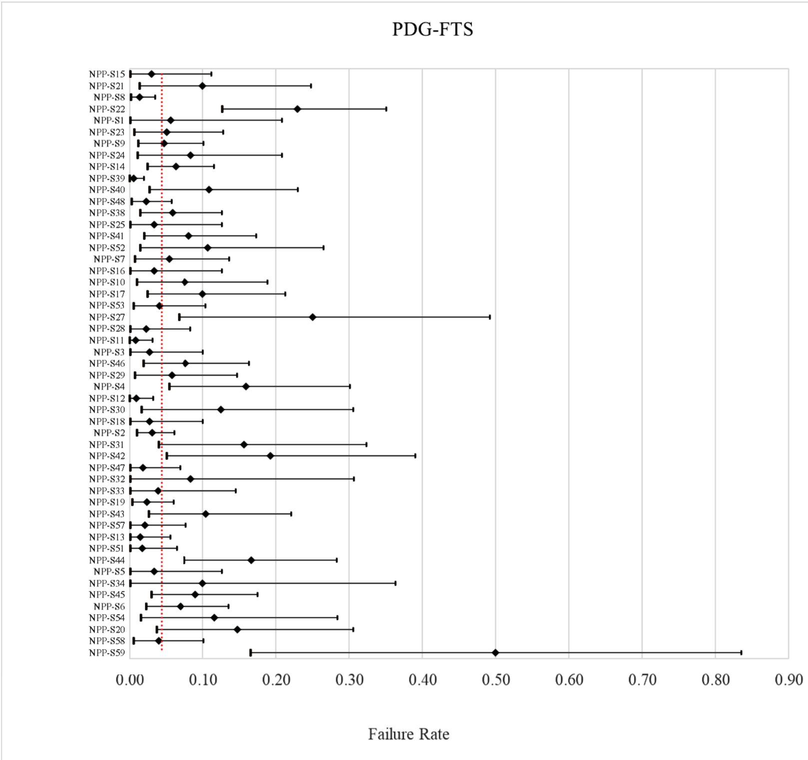


Figure A-1: PCTG 90% Credible Intervals



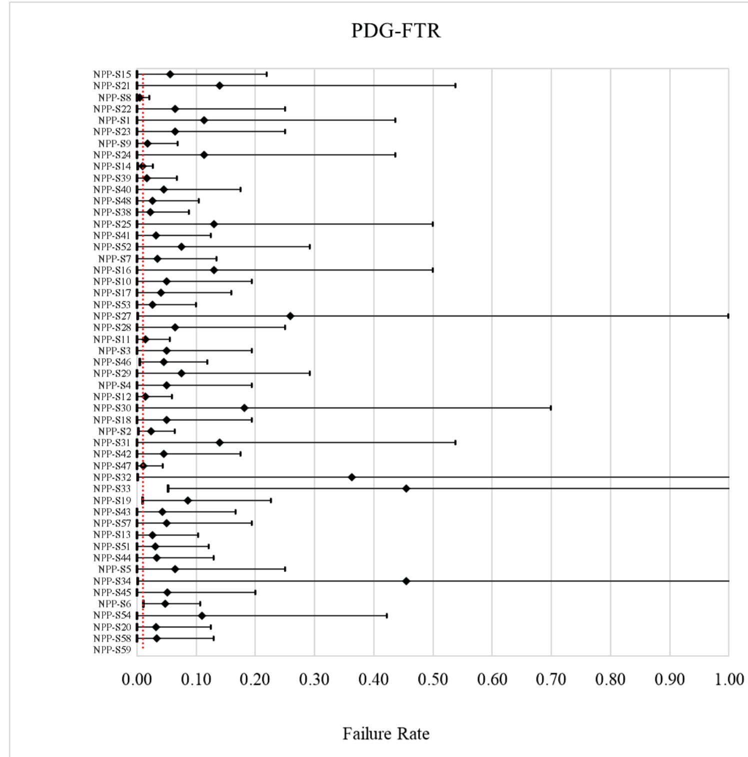
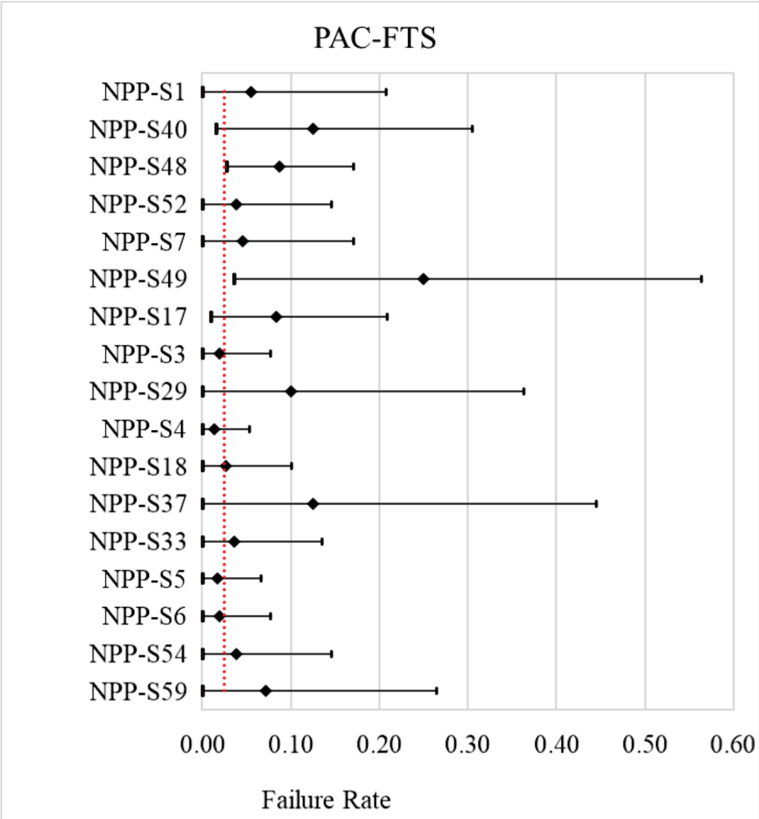


Figure A-2: PDG 90% Credible Intervals



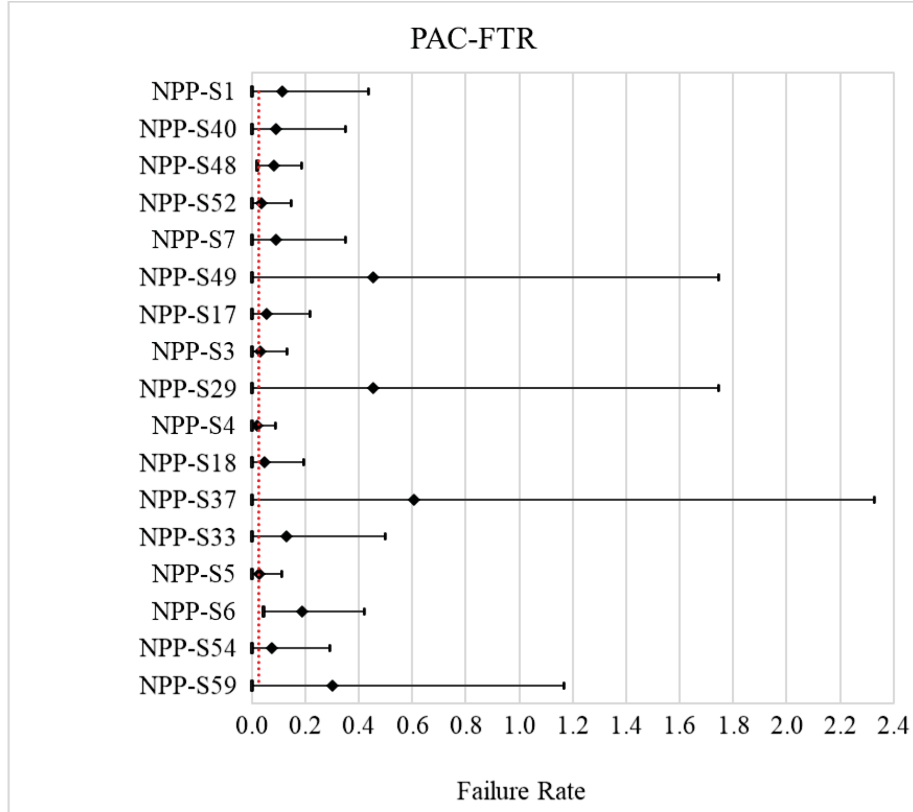
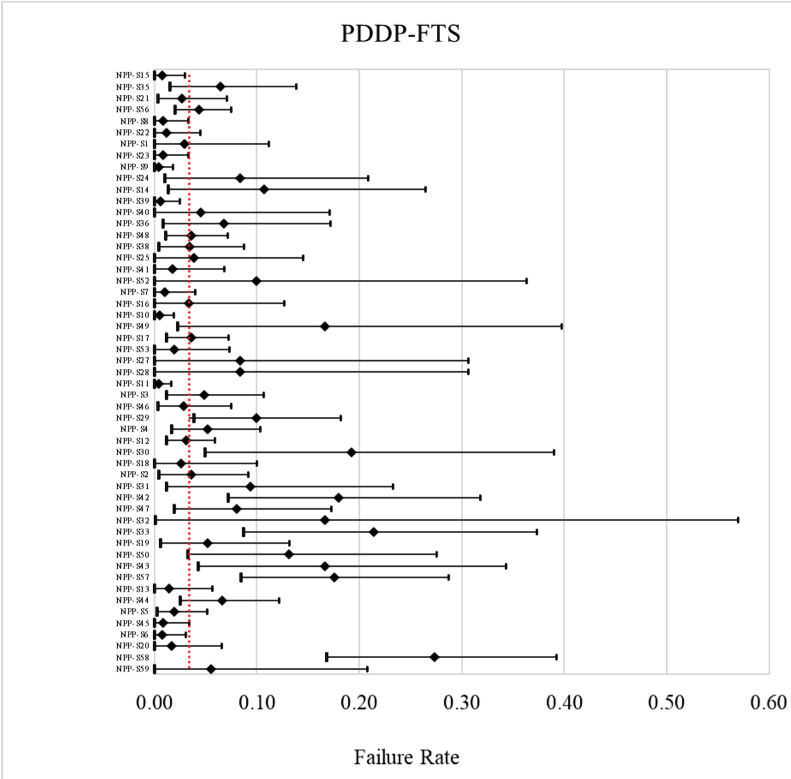


Figure A-3: PAC 90% Credible Intervals



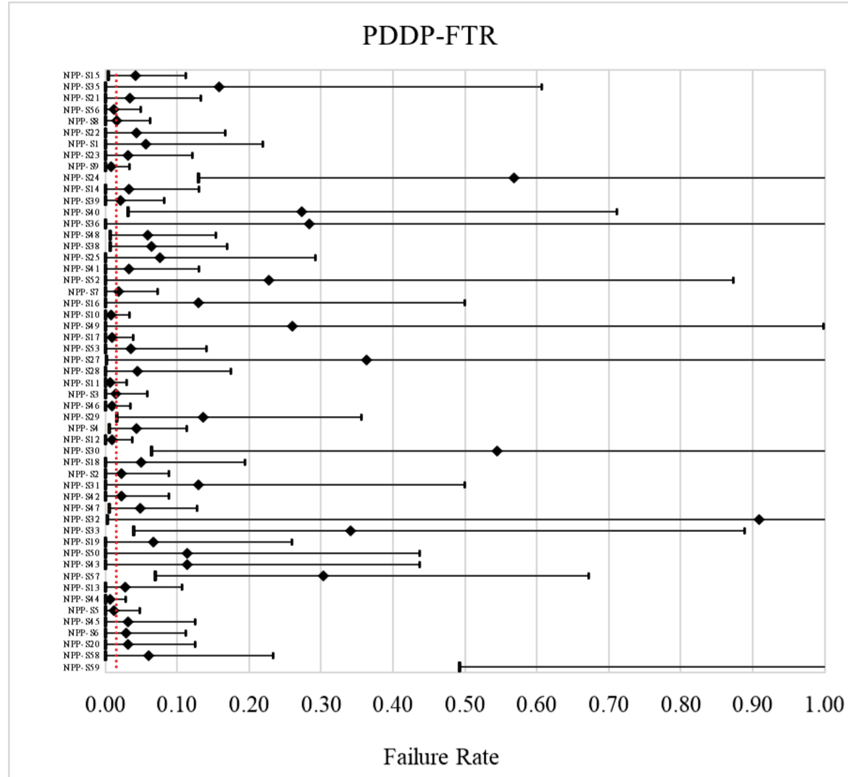
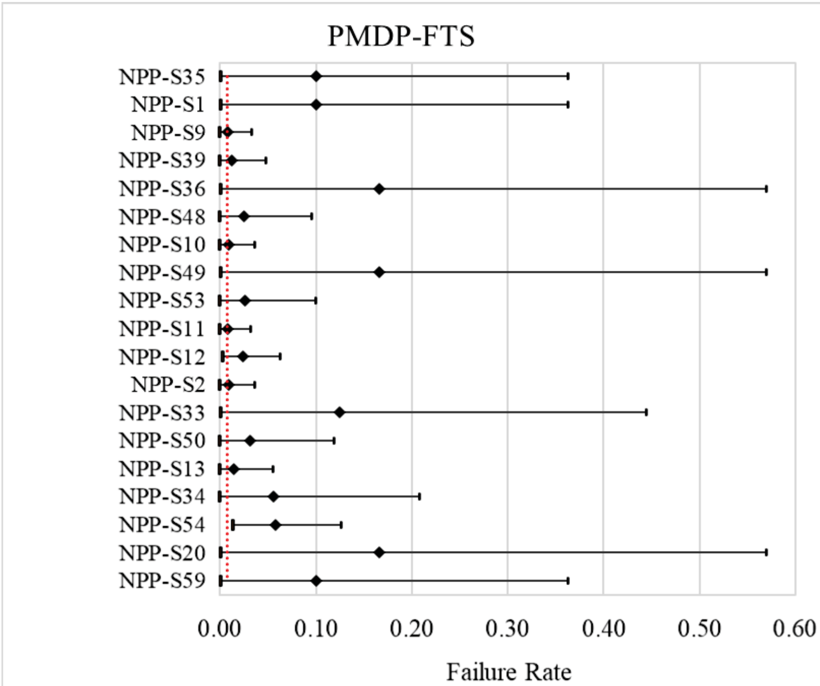


Figure A-4: PDDP 90% Credible Intervals



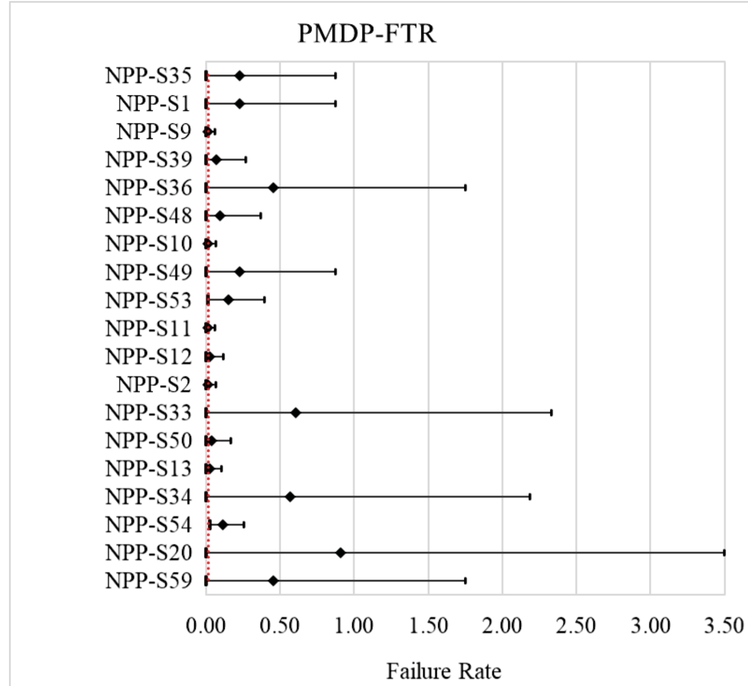


Figure A-5: PMDPDP 90% Credible Intervals

APPENDIX B – Investigation in Data Outliers

B.1 GENERAL OUTLIER IDENTIFICATION PROCESS

The following section investigates the impact of outliers on the data analysis. This process is performed after the data analysis as a means to validate the resulting failure rate estimates. As defined in Section 26.2 of [Ref. 9], an outlier is as an apparently erroneous observation that has been identified by some statistical procedure as due to error or some cause rather than randomness in the data. An observation can deviate from the remaining observations in a sample either because of random fluctuation or because it does not really belong in the sample. If a deviant observation is a legitimate manifestation of randomness but is removed from the sample, the reduced sample is no longer an unbiased sample of the population from which it was drawn. On the other hand, if an observation that does not belong in the sample is retained, the consequent analysis would be tainted. Once an observation has been identified as an outlier, it is important to examine it to try to identify the reason why it was included in the sample. A procedure for identifying outliers is developed by calculating quantiles of the observed data, and plotting this information using a box plot. In [Ref. 9], this is known as the box plot procedure for outlier identification.

The five (5) fundamental quantities that determine a box plot are obtained as follows:

Minimum value – the smallest MLE value in the dataset.

Maximum value – the largest MLE value in the dataset.

Median value – Also known as the second quartile, is the middle MLE value after the values are arranged in ascending order of magnitude. Note that the median is the middle value if the number of values is odd, and the average of the two middle values if the number of values is even.

Lower quartile (LQ) – is the median of the group containing the MLE values below the dataset's median. In general, the lower quartile may only approximate the 25th percentile of the data set.

Upper quartile (UQ) – is the median of the group containing the MLE values above the dataset's median. In general, the upper quantile may only approximate the 75th percentile of the data set.

Additionally, the following measures are defined:

Interquartile Range (IQR) – is the difference between the upper and lower quartiles.

Lower fence (LF) – The lower fence is equal to the lower quartile less 3 times the IQR (e.g., LQ – 3(IQR)). Note the procedure discussed in [Ref. 9] considers a 1.5 factor multiplier to the IQR; however, varying texts classify this as a weak outlier. For this reason, a factor of 3 is used to multiply the IQR which will denote strong outliers.

Upper fence (UF) – the upper fence is equal to the upper quartile summed with 3 times the IQR (e.g., UQ + 3(IQR)). Note the procedure discussed in [Ref. 9] considers a 1.5 factor multiplier to the IQR; however, varying texts classify this as a weak outlier. For this reason, a factor of 3 is used to multiply the IQR which will denote strong outliers.

Outlier – An outlier is defined in context of the terms above as a value that is smaller than the lower fence, or larger than the upper fence; however, for this analysis, focus is only given to values larger than the upper fence. The data analyzed in this document is sparse and results in many lower quartile values being equal to zero, which will result in a lower fence < 0 which, of course, cannot occur.

Plots for each component group and failure mode (at the site level) were created to aid in identifying potential data outliers. The plots created contain the data, UF line, and MLE line. The slope of a line from any point to the origin is the occurrence rate or probability corresponding to that point. Points to the left of the UF or MLE line exceed the UF or MLE value. Points to the left of the UF lines are classified as potential data outliers based on the methodology outline above. Note in some plots, the UF line is below the MLE line. This is due to a large amount of entries having zero failures. In these scenarios, points to the left of the MLE line are classified as potential data outliers.

Although the procedure identifies some potential outliers, due to the sparse data available in this analysis, it is determined (generally) that there is insufficient evidence to exclude any data points. Leaving these points in the data set is useful as it is a conservative choice.

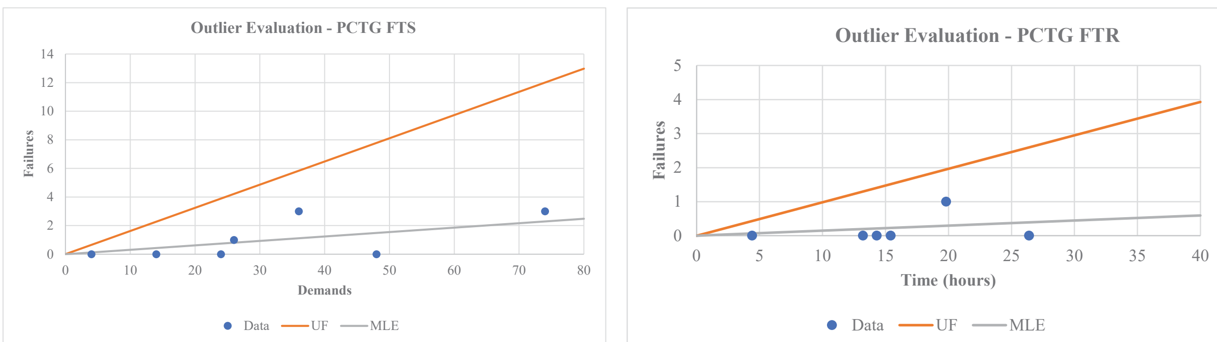


Figure B-1: PCTG Outlier Evaluation

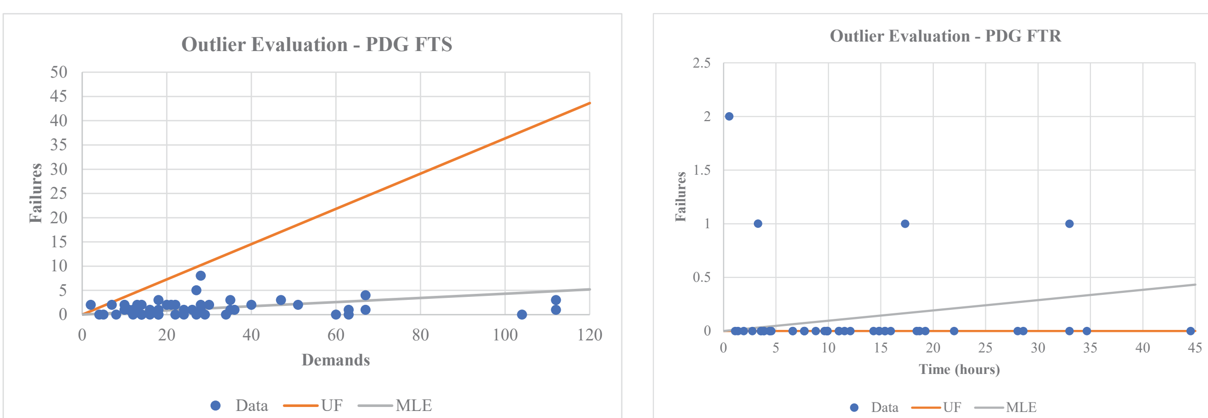


Figure B-2: PDG Outlier Evaluation

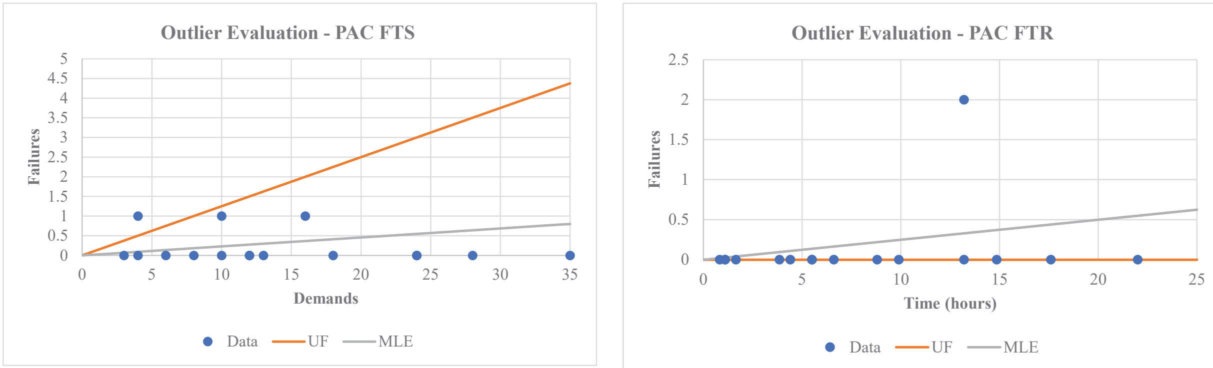


Figure B-3: PAC Outlier Evaluation

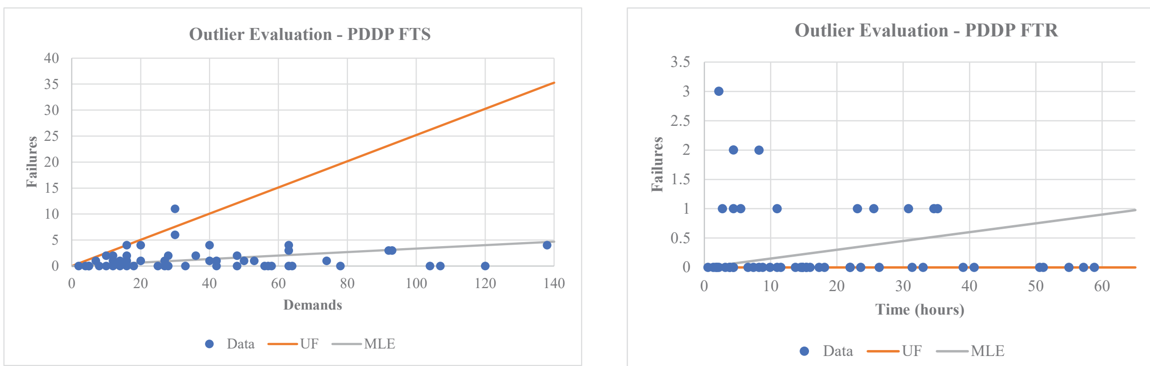


Figure B-4: PDDP Outlier Evaluation

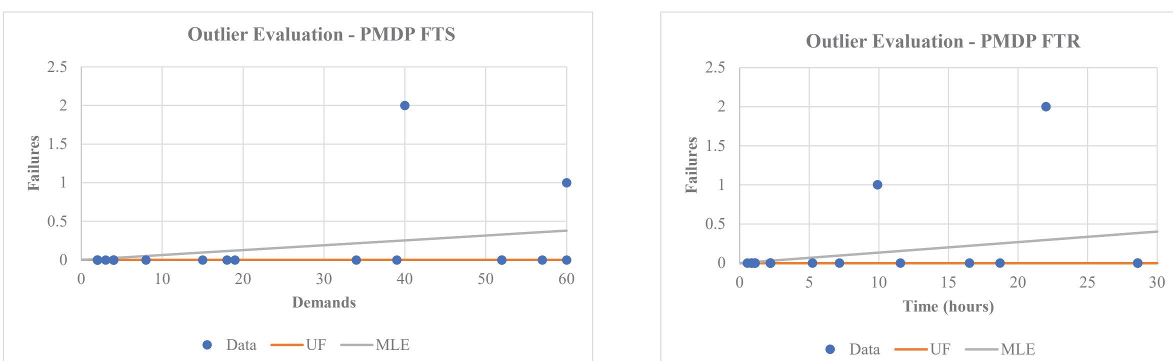


Figure B-5: PMDPDP Outlier Evaluation

APPENDIX C - Empirical Bayes Analysis

C.1 PDG-FTR EMPIRICAL BAYES ANALYSIS

The data for the PDG-FTR component/failure mode was processed through the EB analysis.

Component/Failure Mode	5thº%	Mean	95thº%	α	β	Method
PDG-FTR	3.47E-04	1.03E-02	3.27E-02	0.856	82.9	EB

C.2 PDDP-FTR EMPIRICAL BAYES ANALYSIS

The data for the PDDP-FTR component/failure mode was processed through the EB analysis.

Component/Failure Mode	5thº%	Mean	95thº%	α	β	Method
PDDP-FTR	3.50E-10	3.95E-02	2.18E-01	0.15	3.79	EB

The resultant estimate for α is less than 0.3. Both the bet and gamma distribution can result in unrealistically low estimates for the 5th percentiles of the distributions as α decreases. NUREG/CR-6928 [4] has taken two approaches to handle low values of α:

1. Use a lower bound value of 0.3 for α and recalculate β and the mean for the distribution. This is based on observations that the 5th percentile drops dramatically as α is reduced from 0.3 to 0.2 and 0.1 for beta and gamma distributions. The 5th percentile values for these low values of α are considered unrealistic in terms of representing lower bounds on component unreliability.
2. When the difference between the 5th percentile and the mean is greater than 4 orders of magnitude, which happens to approximate the decision point of α < 0.3, instead of creating an arbitrary distribution, a JNI is used which is the same decision that is made when the EB does not return a result.

Approach one has been determined to be obsolete in newer revisions of NUREG/CR-6928 [4], and thus approach two is used. Consistent with this approach, a JNI is used to develop unreliability estimates for the PDDP-FTR failure mode.

Component/Failure Mode	5thº%	Mean	95thº%	α	β	Method
PDDP-FTR	9.79E-03	1.55E-02	2.22E-02	16.5	1065.2	JNI

APPENDIX D – NRC Audit

On March 24th and 25th, 2020, staff from the Office of Nuclear Reactor Regulation, Region I, and the Office of Nuclear Regulatory Research conducted a remote audit of PWROG-18043-P Revision 0. The audit included participation by INL staff and contractors, who are responsible for collecting and analyzing the non-FLEX data for use in developing component reliability parameters used in the NRC Standardized Plant Analysis Risk models. The audit and the associated observations from the audit are documented in ML20155K827 and ML29155K835. A summary of those observations and the changes made to the report are described below.

Process

1. The NRC audit team had several questions regarding the overall process, for example, the need for, but lack of, a failure definition. An explicit definition helps ensure a consistent approach and would limit the subjectivity and personal biases of individual reviewers. For example, some diesel generator start tests considered successes based on the availability and use of battery charging cables. It is uncertain that the analyzed data had verified that this contingency had been proceduralized to require a jumper source to remain at the location of the FLEX equipment in order to recover from a failed start attempt which would allow crediting the test as a successful start.

Response:

Section 5.2 of the report documents the failure criteria definitions used in classifying events. Following the NRC Audit, Section 5.1 of the report was updated with additional guidance on classifying events. Recoverability of components was considered in the classification of failure events but was ultimately not incorporated into the analysis.

2. As part of the American Society of Mechanical Engineers/American Nuclear Society PRA standard requirement to use current, up-to-date data, a formal update process should be established and included in the overall data handling process. This becomes more important as licensees start relying more on their PRA models in risk-informed regulatory processes (i.e., 10 CFR 50.69, Technical Specification Task Force-505). Along with a formalized update process, a formal data collection process would ensure a consistent approach to gathering FLEX OpE data.

Response:

The current scope of this project involves an initial data collection and development of failure rates for the equipment as specified in Section 4 of the report. The intention of the PWROG is to periodically update this analysis as more data becomes available. It is expected that an update will occur in the next 2-3 years, followed by periodic updates consistent with the update frequency of the permanently installed equipment data updates (approximately every 4 years). It is noted that, while this is the intention of the PWROG, no formal scope has been approved at this current time.

3. Another observation discussed with the PWROG was the choice of the component boundary definitions when compared to the NRC's current definitions. The PWROG stated that they intended

to use similar boundary definitions as the NRC's analyses and agreed to look at identified differences to see if errors were made. Because of inherent differences between installed and FLEX equipment boundary definitions, attention will be required for licensees and staff to appropriately model component and system failure probabilities in both the NRC and the licensees' PRA models.

Response:

Component boundary definitions were updated to be generally consistent with the NRC's current definitions as specified in the NRC 2015 parameter update documentation. Note that starter batteries are the exception since they are included in the component boundary for the generators and other diesel-driven equipment.

<u>**Data Collection**</u>

1. The NRC audit team had several questions and discussions regarding the overall data pedigree. There were discussions regarding the basis of the preventive maintenance (PM) frequencies and the possibility that these frequencies may not appropriately represent authentic equipment starts, which are necessary in order to generate reliability parameters. Specifically, the NRC audit team asked if the PMs provided by the licensees represented actual equipment starts and were not merely checklist-type rundowns of equipment availability which omitted running the equipment. The PWROG confirmed that, to the best of their understanding, the PMs did represent actual equipment starts but agreed that confirming this would be beneficial. There was also discussion about overlapping PMs listed for the same equipment and potential for double counting of equipment starts. For example, if one start and run fulfills the requirements for both a monthly and annual test, only one start and run should be recorded. These concerns are partially a reflection of the fact that the PWROG process is in its infancy compared to the established INPO process used by licensees since 1998, with the introduction of the Equipment Performance and Information Exchange database. As a result, individual licensee responses may not be consistent due to differing interpretations of the data request.

   Response:

   Following the NRC audit, the calculation of demands and run hours were updated to ensure double counting of PMs was not occurring. It is industry practice to credit tests as satisfying multiple PMs (e.g., a bi-annual and annual test). Additionally, the PMs were reviewed to ensure only credit for PMs that represent actual equipment starts are credited in the analysis. The Database used in the analysis was updated to list each PM with its associated run time, and a description of the PM.

2. Another data pedigree observation was that some of the data was collected before licensees' declarations or order compliance. There was some concern that failures occurring during these periods may not have been consistently captured in a licensees' corrective action program. This was an area of ongoing concern for the NRC audit team and was discussed with the PWROG. The audit team believes through discussions with PWROG that the intent was to utilize test data after a licensee had made their declaration of order compliance and submittal of the final integrated plan. The audit team believed that approach would ensure the integrity of the data used.

Response:

Following the NRC audit, the data were reviewed and a subset of equipment was identified that had PM start dates prior to the declaration of order compliance for FLEX. Discussions were held with the utilities to confirm the appropriate start dates consistent with the declaration of order compliance. It was noted that the majority of inconsistencies were related to B.5.b equipment which were typically on site prior to the declaration of order compliance for FLEX, and then subsequently pulled into the FLEX program. PM start dates were updated for this set of equipment to match the FLEX compliance date, and the failure rate calculations were updated.

3. Another topic that the NRC audit team focused on was whether the equipment "run" data represented loaded conditions. There were several conversations on the importance of limiting test results to fully loaded runs and questions regarding the likely inclusion of partial or non-loaded runs which would have the effect of artificially reducing failure rate in runtime calculations. Specifically, unloaded runs may not identify some failures which could be observed with the machines loaded (i.e., engine control modules, cooling systems, etc.)

Response:

The majority of the data represents un-loaded equipment configurations for diesel generators. Typically, the more frequent tests are performed in un-loaded configurations, with less frequent tests being performed in loaded configurations. The current periodicity of un-loaded and loaded tests is developed through industry representatives with consultation from equipment experts. The basis for these frequencies has been vetted and a consensus approach is used throughout the industry. A review was performed on events classified as fail-to-run for the Portable Diesel Generator component group. At most, one event was identified as a potential fail-to-load event. This finding, as well as the current failure rate results, show that a fail-to-load failure mode would not be an important failure mode for this class of DGs given the current set of data.

<u>Data Analysis</u>

1. The NRC audit team reviewed the three data analysis methods that the PWROG used to develop the FLEX equipment failure probabilities. Out of the three methods, the audit team's review had concerns regarding the use of the WIP method. The WIP method was used for equipment categories that contain less than 50 demands or less than 100 hours of run time (i.e., in cases where the test results were sparse). The audit team inquired if there were other ways to better categorize the equipment where the challenges associated with limited data could be eliminated. The PWROG indicated that these categories were established early in their study before they knew that these categories would possess such limits. The PWROG displayed an openness to investigating a revision to this approach, should they decided to update their analysis.

Response:

The WIP methodology was removed from the report based on the comments from the NRC and INL. Then, the equipment categories were grouped together instead of having various operating ranges (e.g., PDG-HV, PDG-MV, and PDG-LV were combined into one group PDG). Following

this categorization, limited data was still an issue. Instead of the WIP method, the constrained noninformative prior distribution (CNID) was considered for situations in which there was limited data.

2. The NRC audit team received significant input from INL staff and contractors, specifically with respect to the technical report INL/EXT-20-58327, "Evaluation of Weakly Informed Priors for FLEX Data." (ADAMS Accession No. Ml20155K835). The concerns expressed by the INL statisticians centered around the WIP method's heavy reliance on engineering judgment. For example, there were discussion on: the proper "range factor" to use for new portable equipment, which attempts to measure the component's failure rate uncertainty; the proper additional scaling factors "Em" values to use; and what other scaling factor should be chosen? As such, the NRC audit team voiced their overall concern with the use of a method that relies heavily on engineering judgement.

Response:

The WIP methodology was removed from the report based on the comments from the NRC and INL. Instead, the Constrained noninformative prior distribution (CNID) was considered for situations in which there was limited data, but was not used.

3. The NRC audit team questioned the use of the permanently installed equipment as the foundation for the inference function used to create the failure rates and commented if other equipment (e.g., other FLEX equipment) would provide a more accurate foundation to reflect FLEX performance as standby equipment.

Response:

Parameters for the permanently installed equipment are still considered applicable when assessing potential prior distributions for the FLEX equipment. Specifically, the constrained noninformative prior distribution (CNID) represents a method in which the permanently installed parameters are "adjusted" by some validated factor. This method is discussed in the report, but ultimately not used in derivation of the failure rates.

4. Confirmatory calculations provided by INL revealed potential errors in the PWROG's WIP results. Specifically, it appears the PWROG used the beta equation when calculating alpha for the four fail-to-start WIP distributions.

Response:

The WIP methodology was removed from the report based on the comments from the NRC and INL. Instead, the Constrained noninformative prior distribution (CNID) was considered for situations in which there was limited data.

5. The NRC audit team also noted that the current PWROG analysis lacks the common-cause failure (CCF) analysis as well as the unavailability analysis, which would be important elements to be

included in PRA models reflecting FLEX equipment. The PWROG acknowledged the limitations and stated that such analyses might be added in the future when updating the FLEX data analysis.

Response:

The intent of the PWROG is to periodically update this analysis in the future (see response to Item #2 under Process related observations). It is the intent of the PWROG to update the independent component failure rate estimates, as well as to consider the development of common cause parameters for the FLEX equipment at a future date. It is noted that, while this is the intention of the PWROG, no formal scope has been approved at this current time.

Approval Information
Author Approval Degonish Matthew M Feb-16-2022 12:25:08
Reviewer Approval Kiper Kenneth L Feb-16-2022 14:31:15
Manager Approval Davis Stacy A Feb-16-2022 14:34:37
Manager Approval Mirizio Damian S Mar-03-2022 10:39:34

Files approved on Mar-03-2022