Correlating Power Quality Indices with System Reliability Indices

Technical Report

Correlating Power Quality Indices with System Reliability Indices

1008508

Final Report, March 2005

EPRI Project Manager
R. Schainker

Electric Power Research Institute • 3412 Hillview Avenue, Palo Alto, California 94304 • PO Box 10412, Palo Alto, California 94303 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THIS DOCUMENT WAS PREPARED BY THE ORGANIZATION(S) NAMED BELOW AS AN ACCOUNT OF WORK SPONSORED OR COSPONSORED BY THE ELECTRIC POWER RESEARCH INSTITUTE, INC. (EPRI). NEITHER EPRI, ANY MEMBER OF EPRI, ANY COSPONSOR, THE ORGANIZATION(S) BELOW, NOR ANY PERSON ACTING ON BEHALF OF ANY OF THEM:

(A) MAKES ANY WARRANTY OR REPRESENTATION WHATSOEVER, EXPRESS OR IMPLIED, (I) WITH RESPECT TO THE USE OF ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT, INCLUDING MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, OR (II) THAT SUCH USE DOES NOT INFRINGE ON OR INTERFERE WITH PRIVATELY OWNED RIGHTS, INCLUDING ANY PARTY'S INTELLECTUAL PROPERTY, OR (III) THAT THIS DOCUMENT IS SUITABLE TO ANY PARTICULAR USER'S CIRCUMSTANCE; OR

(B) ASSUMES RESPONSIBILITY FOR ANY DAMAGES OR OTHER LIABILITY WHATSOEVER (INCLUDING ANY CONSEQUENTIAL DAMAGES, EVEN IF EPRI OR ANY EPRI REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES) RESULTING FROM YOUR SELECTION OR USE OF THIS DOCUMENT OR ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT.

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

EPRI Solutions, Inc.

ORDERING INFORMATION

Requests for copies of this report should be directed to EPRI Orders and Conferences, 1355 Willow Way, Suite 278, Concord, CA 94520, (800) 313-3774, press 2 or internally x5379, (925) 609-9169, (925) 609-1310 (fax).

Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc. EPRI. ELECTRIFY THE WORLD is a service mark of the Electric Power Research Institute, Inc.

Copyright © 2005 Electric Power Research Institute, Inc. All rights reserved.

CITATIONS

This report was prepared by

EPRI Solutions, Inc.
942 Corridor Park Blvd.
Knoxville, TN 37932

Principal Investigators
A. Maitra
A. Gaikwad
P. Sutherland
T. Short
W. Sunderman

This report describes research sponsored by EPRI.

The report is a corporate document that should be cited in the literature in the following manner:

Correlating Power Quality Indices with System Reliability Indices, EPRI, Palo Alto, CA: 2005. 1008508.

PRODUCT DESCRIPTION

Energy companies can use better knowledge about existing power quality to target maintenance efforts, to establish a baseline for offering premium power services, and to use as a selling tool to sell sites with high power quality. Unfortunately, most energy companies do not have widespread, long-term records of power quality needed to provide this information. But energy companies often have good historical records of reliability indices for all circuits on their system (System Average Interruption Frequency Index, or SAIFI, and System Average Interruption Duration Index, or SAIDI, for long-duration interruptions). Areas with more voltage sags normally have more momentary interruptions; also, areas with more voltage sags normally have more long-duration interruptions. This project quantifies this effect and determines how effectively SAIFI and SAIDI predict power quality indices. The starting point for the analysis was data from two utilities ("Utility A" and "Utility B" in the report) in the Distribution Power Quality (DPQ) II data set expanded to include reliability data.

Results & Findings

This project leverages analysis work conducted in 2003 and extends this to determine how effectively reliability measures (for example, SAIFI, SAIDI, and number of outages) predict power quality indices such as momentary interruptions and sags. Characteristics of the site (either at the feeder level or at the substation level) help identify major influences on the System Average RMS Variation Frequency Index ($SARFI_x$) and the Momentary Average Interruption Frequency Index (MAIFI). This allows energy companies to predict the level of quality at a location. Among other uses, utilities can predict the quality at a given site to find good sites to locate sensitive customers, to determine if a site is under-performing based on the given infrastructure, or to quantify the risk of offering premium power services.

The starting point for the 2004 R&D activity was to expand the DPQ II dataset and include reliability data as well. As a part of the 2004 R&D activity, the project team obtained additional new data from two utilities that participated in DPQ Phase II. The historical power quality data ($SARFI_x$ and MAIFI from DPQ II) were correlated with feeder SAIFI, SAIDI, and number of outages. The effect of other variables—including overall feeder length, feeder backbone length, number of feeders per bus, substation transformer MVA, and load density on the correlation—were examined. Regression models from the Voltage Sag Prediction Model were refined.

Challenges & Objective(s)

While utilities widely record reliability indices such as SAIFI, SAIDI, and number of outages, most have few power quality monitors. The 2004 R&D activity focused on investigating the possibility of refining voltage-sag predictions using reliability indices. In a statistical prediction model, reliability indices would become another site characteristic. This could improve power quality predictions and also may reveal interesting practical ties between power quality and reliability. These are statistical regression models built from optimizations of models containing several of the site characteristics available for the utility datasets used. The main challenges were finding good models, dealing with sites that had unknown site characteristics, and analyzing variability in the models. Data from one utility (Utility A) was for the period 1998-2003, while data from the second utility (Utility B) was from 2001-2003.

Applications, Values & Use

This report provides utility engineers with a set of straightforward equations for estimating the rate of voltage sags at a location. Only a few site parameters are needed: substation transformer size, number of feeders on the bus, distribution voltage level, SAIFI, and total feeder length. The prediction model also includes an estimate of variability, which can be portrayed as prediction limits.

One universal prediction model for all utilities is not possible. The best prediction model varies between utilities based on the utility's substation design and service territory. Substation design and geographical service territory and other aspects specific to a distribution system design will determine the best model. Utility A had a more uniform substation design across their region. The geographic conditions for Utility A also were uniform. For these reasons, multi-linear-regression-based prediction models worked better for Utility A. The error between the observed and predicted values for SARFI appears to follow a normal Gaussian distribution. The best prediction model for Utility B was linear with a gamma error distribution. However, the coefficient derived using data from Utility A is quite similar to that derived using Utility B.

EPRI Perspective

Estimating the frequency and characteristics of power quality events, which are part of the normal electrical environment, is critical for designing immunity into end-use equipment. This information also is important to define baseline levels of power quality that can be expected by a customer and sets the stage for providing premium power services. The prediction models will help utilities enhance the compatibility of sensitive equipment with the electrical environment.

Approach

The main components of the prediction modeling are model fitting, variable selection, and handling missing data and quantifying variability. The project team analyzed the variability and uncertainty in various models. A relatively straightforward set of equations for the voltage sag as well as momentary interruption based indices was found. From this analysis, if appropriate sets of site descriptors and reliability indices for predictions of MAIFI are chosen, all major SARFI indices can be obtained using the developed predictor model.

Keywords

Power quality
Distribution
Voltage sag
Power quality monitoring

ACKNOWLEDGMENTS

This report contains material reprinted with permission from:

T. A. Short, *Electric Power Distribution Handbook*, CRC Press, Boca Raton, FL, 2004.

CONTENTS

---

1 INTRODUCTION ... 1-1

How Concerned Are We With the Quality of Our Supply ... 1-1

Previous Efforts ... 1-2

Using Monitoring To Gauge System Performance ... 1-2

A Need for Improved Data Analysis ... 1-5

Project Objectives and Scope ... 1-8

Data Collected from Two Candidate Utilities ... 1-10

Monitor Types ... 1-10

Monitor Settings ... 1-10

System Locations/Levels/Data Type ... 1-10

2 EVALUATING TIME AND SPATIAL VARIABILITY IN POWER QUALITY AND RELIABILITY INDICES ... 2-1

Measures for Distribution Reliability and Power Quality ... 2-2

Power Quality Prediction ... 2-5

SARFI Statistics ... 2-5

Probability Distributions of Power Quality and Reliability Variables ... 2-9

Goodness-of-Fit Tests ... 2-10

3 VARIABLES AFFECTING DISTRIBUTION RELIABILITY INDICES ... 3-1

Outage Management Systems ... 3-2

Storm Normalization and Defining Major Events - Impact on Indices ... 3-2

Distribution Automation ... 3-4

Geography, Circuit Exposure, and Supply Configuration ... 3-5

Customer Perception of Service Quality ... 3-7

4 VARIABLES AFFECTING POWER QUALITY AND MOMENTARY INTERRUPTIONS ... 4-1

Parameters Available in the Datasets ... 4-1

Electrical Parameters That are Important for Sags ... 4-2

Correlation Of SARFI$_x$ and Site Parameters 4-8
Feeder Correlation 4-8
Substation/Bus Level Correlation 4-10
Correlations Between Power Quality and Reliability Variables 4-19
Variations with Parameters 4-29
Bin Level Correlations with Site Parameters. 4-32

5 **PREDICTION APPROACH** **5-1**
Model Fitting 5-2
Predicting Variability 5-4

6 **PREDICTIONS FOR SARFI** **6-1**
Predicting Sags and Momentary Interruptions Based on Predictions Derived From Utility Data 6-1
Need to Decide Whether to do MLR or GLM 6-2
Histogram for Residuals 6-3
Prediction Model Development Using Data from Utility A and Utility B 6-3
Other Ways to Estimate Voltage Sags 6-12

7 **SUMMARY** **7-1**
Improving Predictions from the Correlation Analysis 7-2
Future Work 7-3

A **DISTRIBUTION RESULTS FOR UTILITY A** **A-1**

B **DISTRIBUTION RESULTS FOR UTILITY B** **B-1**

C **PROPERTIES OF BUS DATA FOR UTILITY A** **C-1**

D **PROPERTIES OF BUS DATA FOR UTILITY B** **D-1**

LIST OF FIGURES

Figure 1-1 Flow Chart for Selecting Appropriate Index 1-6

Figure 1-2 Location Categories for the Sample Sites 1-12

Figure 1-3 Monitor Connection for Location B 1-13

Figure 2-1 Cumulative Distribution of $\text{SARFI}_{\text{ITIC}}$ by Monitoring Location 2-9

Figure 2-2 A Lognormal Probability Plot 2-11

Figure 2-3 Graphical Comparisons of CDFs Distribution of $\text{SARFI}_{\text{ITIC}}$ and MAIFI for Utility B 2-15

Figure 2-4 Graphical Comparisons of CDFs Distribution of $\text{SARFI}_{\text{ITIC}}$ and SAIFI for Utility B 2-15

Figure 2-5 Graphical Comparisons of CDFs Distribution of $\text{SARFI}_{\text{ITIC}}$ and SAIDI for Utility B 2-16

Figure 2-6 SAIFI Gamma Distribution, Bus Level, Including Storms for Utility A 2-17

Figure 2-7 SAIDI Gamma Distribution, Bus Level, Including Storms for Utility A 2-17

Figure 2-8 Outages Gamma Distribution, Bus Level, Including Storms for Utility A 2-18

Figure 2-9 SARFI_ITIC Gamma Distribution, Bus Level, Including Storms for Utility A 2-18

Figure 2-10 SARFI_50 Weibull Distribution, Bus Level, Including Storms for Utility A 2-19

Figure 2-11 SAIFI Gamma Distribution, Bus Level, Including Storms for Utility B 2-19

Figure 2-12 SAIDI Gamma Distribution, Bus Level, Including Storms for Utility B 2-20

Figure 2-13 MAIFI Gamma Distribution, Bus Level, Including Storms for Utility B 2-20

Figure 2-14 SARFI_ITIC Gamma Distribution, Bus Level, Including Storms for Utility B 2-21

Figure 2-15 SARFI_70 Gamma Distribution, Bus Level, Including Storms for Utility B 2-21

Figure 2-16 SARFI_50 Weibull Distribution, Bus Level, Including Storms for Utility B 2-22

Figure 3-1 Variables Impacting Distribution Reliability Indices 3-1

Figure 3-2 Impact of Major Events on Indices [EEI Survey 1999] 3-4

Figure 3-3 Effect of Circuit Length on Indices 3-5

Figure 4-1 Voltage Divider Equation Giving the Voltage at the Bus for a Fault Downstream 4-3

Figure 4-2 Substation Voltage Profile for Faults at the Given Distance (Single-Phase and Three-Phase Faults Are Shown for Each Voltage) 4-4

Figure 4-3 Comparison of Substation Voltage for Faults on Overhead Circuits and Cable Circuits at the Given Distance (Single-Phase and Three-Phase Faults Are Shown) 4-5

Figure 4-4 Cumulative Distribution of Substation Bus Voltage Sags per Year for the Given (25-MVA, 10% Transformer, 500-kcmil Feeder, n=2 or 4 Feeders off the Bus, λ = 1 Faults/Phase/Mile of Mains/Year, Assumes Line-to-Ground Faults Only) 4-7

Figure 4-5 Feeder Level Correlations for SARFI_ITIC at Utility A – (Lack of Correlation Between SARFI_ITIC Measured by Bus Monitor and Feeder Characteristics).................4-8

Figure 4-6 Feeder Level Correlations for SARFI_70 at Utility A – (Lack of Correlation Between SARFI_70 Measured by Bus Monitor and Feeder Characteristics) 4-9

Figure 4-7 Feeder Level Correlations for SARFI_80 at Utility A – (Lack of Correlation Between SARFI_80 Measured by Bus Monitor and Feeder Characteristics) 4-9

Figure 4-8 Feeder Level Correlations for SARFI_50 at Utility A – (Lack of Correlation Between SARFI_50 Measured by Bus Monitor and Feeder Characteristics) 4-10

Figure 4-9 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Various Parameters. Bus Level – Utility A.............4-11

Figure 4-10 SAIRFI_{70} Site Variations With Various Parameters. Bus Level – Utility A............4-12

Figure 4-11 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Various Parameters. Bus Level (15KV) – Utility B...4-13

Figure 4-12 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Various Parameters. Bus Level (25KV) – Utility B...4-14

Figure 4-13 SARFI_{50} Site Variations With Various Parameters. Bus Level (15KV) – Utility B...4-15

Figure 4-14 SARFI_{50} Site Variations With Various Parameters. Bus Level (25KV) – Utility B...4-16

Figure 4-15 MAIFI Site Variations With Various Parameters. Bus Level (15KV) – Utility B...4-17

Figure 4-16 MAIFI Site Variations With Various Parameters. Bus Level (25KV) – Utility B...4-18

Figure 4-17 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Reliability Variables. Bus Level – Utility A............4-22

Figure 4-18 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Reliability Variables. Bus Level – Utility B............4-23

Figure 4-19 $\text{SARFI}_{\text{ITIC}}$ Site Variations With Reliability Variables. Bus Level – Utility B............4-24

Figure 4-20 Correlations Between Variations of SARFI. Bus Level – Utility A........................4-25

Figure 4-21 Correlations Between Variations of SARFI. Bus Level – Utility B........................4-26

Figure 4-22 Correlations Between Variations of SARFI. Bus Level – Utility B........................4-27

Figure 4-23 Correlations Between Variations of SARFI. Bus Level – Utility B........................4-28

Figure 4-24 Correlations Between Variations of SARFI. Bus Level – Utility B........................4-29

Figure 4-25 $\text{SARFI}_{\text{ITIC}}$ Bin Variations With SAIFI and Feeder Length. Bus Level – Utility A....4-32

Figure 4-26 $\text{SARFI}_{\text{ITIC}}$ Bin Variations With SAIDI and Feeder Length. Bus Level – Utility A..4-33

Figure 4-27 $\text{SARFI}_{\text{ITIC}}$ Bin Variations With Outages and Feeder Length. Bus Level – Utility A...4-34

Figure 4-28 $\text{SARFI}_{\text{ITIC}}$ Bin Variations With SAIFI and Feeder Length. 15 kV Bus Level – Utility B...4-35

Figure 4-29 $\text{SARFI}_{\text{ITIC}}$ Bin Variations With SAIFI and Feeder Length. 25 kV Bus Level – Utility B...4-36

Figure 4-30 MAIFI Bin Variations With SAIFI and Feeder Length. 15 kV Bus Level – Utility B 4-37

Figure 4-31 MAIFI Bin Variations With SAIFI and Feeder Length. 25 kV Bus Level – Utility B 4-38

Figure 5-1 Comparison of the SARFI$_{ITIC}$ Data With Normal and Gamma Distribution Fits (Cumulative Distributions and Probability Densities)....5-3

Figure 6-1 Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using MLR Approach (Utility A)....6-4

Figure 6-2 Residual Distribution for SARFI$_{ITIC}$ Derived Using MLR Approach (Utility A) 6-5

Figure 6-3 Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using GLM Approach (Utility A)....6-6

Figure 6-4 Observed SARFI$_{ITIC}$ Versus Predicted SARFI$_{ITIC}$ Derived Using GLM Approach (Utility A)....6-7

Figure 6-5 Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using GLM Approach (Utility B)....6-8

Figure 6-6 Observed SARFI$_{ITIC}$ Versus Predicted SARFI$_{ITIC}$ Derived Using GLM Approach (Utility B)....6-9

Figure 6-7 Model Coefficients and Results for MAIFI Derived Using GLM Approach (Utility B)....6-10

Figure 6-8 Observed MAIFI Versus Predicted MAIFI Derived Using GLM Approach (Utility B)....6-11

Figure A-1 SAIFI Gamma Distribution, Bus Level, Excluding Storms.... A-2

Figure A-2 SAIDI Gamma Distribution, Bus Level, Excluding Storms A-2

Figure A-3 Outages Gamma Distribution, Bus Level, Excluding Storms A-3

Figure A-4 SARFI_90 Gamma Distribution, Bus Level, Including Storms A-3

Figure A-5 SARFI_80 Gamma Distribution, Bus Level, Including Storms A-4

Figure A-6 SARFI_10 Gamma Distribution, Bus Level, Including Storm A-4

Figure B-1 SAIFI Gamma Distribution, Bus Level, Including Storm.... B-2

Figure B-2 SAIDI Gamma Distribution, Bus Level, Including Storm B-2

Figure B-3 MAIFI Gamma Distribution, Bus Level, Including Storm B-3

Figure B-4 SAIRFI_ITIC Gamma Distribution, Bus Level, Including Storm B-3

Figure B-5 SAIRFI_70 Gamma Distribution, Bus Level, Including Storm B-4

Figure B-6 SAIRFI_50Gamma Distribution, Bus Level, Including Storm B-4

LIST OF TABLES

Table 2-1 SARFI Site Medians for Utility A and Utility B (15KV and 25KV Class) 2-6

Table 2-2 $\text{SARFI}_{\text{ITIC}}$ Site Statistics for Utility A and Utility B by Survey and Location (15KV and 25KV Class) 2-7

Table 2-3 SARFI Site Medians for Utility B (15KV and 25KV Class) 2-7

Table 2-4 MAIFI (Including and Excluding Storms) Site Statistics for Utility B by Survey and Location (15KV and 25KV Class) 2-7

Table 2-5 $\text{SARFI}_{\text{ITIC}}$ Site Statistics by Survey and Location for DPQ Phase I and II 2-8

Table 2-6 Bus Level Distributions of Power Quality and Reliability Variables for Utility A 2-12

Table 2-7 Bus Level Distributions of Power Quality and Reliability Variables for Utility B, 15 kV Class. 2-14

Table 3-1 Comparison of Reliability of Different Distribution Configurations [Settembrini et. al., 1991] 3-6

Table 3-2 Historical Trend of System Average Interruption Duration Index (SAIDI) for the Four Different Categories of Networks Maintained by Australian Utilities 3-7

Table 4-1 Coefficient of Correlation (Expressed as a Percentage) Between Power Quality and Reliability Variables– Utility A. 4-19

Table 4-2 Coefficient of Correlation (Expressed as a Percentage) Between Power Quality and Reliability Variables at Bus Level– Utility B 4-20

Table 4-3 Coefficient of Correlation (Expressed as a Percentage) Between Power Quality Variables – Utility A 4-20

Table 4-4 Coefficient of Correlation (Expressed as a Percentage) Between Bus Level Power Quality Variables – Utility B 4-21

Table 4-5 $\text{SARFI}_{\text{ITIC}}$ Site Statistics by Various Site Characteristics Utility A 4-30

Table 4-6 $\text{SARFI}_{\text{ITIC}}$ Site Statistics by Various Site Characteristics Utility B (15KV Class) 4-31

Table 6-1 Model Coefficients and Results for $\text{SARFI}_{\text{ITIC}}$ Derived Using MLR Approach (Utility A) 6-4

Table 6-2 Model Coefficients and Results for $\text{SARFI}_{\text{ITIC}}$ Derived Using GLM Approach (Utility A) 6-8

Table 6-3 Model Coefficients and Results for $\text{SARFI}_{\text{ITIC}}$ Derived Using GLM Approach (Utility B) 6-10

Table 6-4 Model Coefficients and Results for MAIFI Derived Using GLM Approach (Utility B) 6-12

Table 7-1 Coefficients for Prediction Equations for Various SARFI and MAIFI Indices Using Data from Utility A and Utility B 7-3

1

INTRODUCTION

How Concerned Are We With the Quality of Our Supply

The quality of electric power delivery (supply quality) consists of reliability of the supply (continuity of power delivery) and quality of the electric power. As electric utilities around the world are attempting to meet the increasing customer needs for supply quality, the concept of characterization and benchmarking is being embraced. In recent years, as many electric utilities wrestle with the idea of deregulation, their customers are becoming even more concerned about the possible decline in power quality and reliability due to cost-cutting measures. For one thing, what is “best” is likely to be different for each group. Customer satisfaction, therefore, is becoming very important to the electricity suppliers for maintaining a customer base.

However, customers, and especially other suppliers of competitive services, are not necessarily going to simply trust the distribution company to do what is best. Much of the responsibility for allowing the monopoly provider (distribution company) of one service (power delivery) to provide a related competitive service rests with the distribution company itself. With increased sensitivity and the industry’s dependency on sophisticated process control equipment in manufacturing, the first task involves an ability to access and quantify voltage sag and reliability levels throughout electrical systems. Utilities realize that they must understand the levels of supply quality provided throughout their power delivery system. Additionally, utilities must be able to determine if the provided levels are appropriate. This is certainly becoming more prevalent as customers become more knowledgeable and regulators in different parts of the world require threshold levels for a minimum quality of supply.

A recent regulatory trend is moving to performance-based rates, where performance is penalized or rewarded based on sustained/temporary interruptions and short-term voltage disturbances as quantified by power quality and reliability (PQR) indices. Some utilities also pay bonuses to managers or others based in part on indices. Some commercial and industrial customers ask utilities for their power quality indices before locating a facility within the utility’s service territory. Characterization of service quality also helps in managing customer expectation and opens the door for utilities to contract with specific customers to provide service of specified quality level over some period of time at a premium price.

Fortunately, there are ways to help address the concerns:

- Identify metrics to measure power quality (PQ) and reliability.
- Collect yearly data and apply advanced data analysis techniques to evaluate system performance

- Establish a baseline for normal service. The baseline will likely be different for various utilities and consumers.
- Evaluate the inherent variability in power system reliability and power quality
- Establish agreements and contracts (rate structure with incentives) for the electric utility companies to meet and exceed the power quality standards. Contracts will specify the quality of the electricity to be delivered and will have provisions for penalizing the utilities if they do not meet the specifications. One type of agreement and contract is the premium power-based contract for electric service providers and consumers. This type of contract is guaranteed performance-based, defining premium power as a service beyond the normal expectation.

Previous Efforts

An important difference in analyzing power quality and reliability benchmarks is their “spread” of influence. Premium power based options are highly localized to a specific customer or at the most to a group of feeders in the system. On the other hand, reliability benchmarks are measured at the system level and/or at sub-levels within the system. Sub-levels are identified based on the inherent differences in the network such as geographical topology, voltage service levels, population density, and load density. However, any efforts to establish benchmarks require quantification of the existing service.

Using Monitoring To Gauge System Performance

Direct monitoring of transmission as well as distribution systems seems to be the preferred approach for benchmarking power quality levels that many utilities are performing for their industrial and commercial customers. However, what should be measured in order to characterize power quality of electric service as satisfactory or otherwise?

Traditionally, reliability indices such as System Average Interruption Frequency Index (SAIFI) and System Average Interruption Duration Index (SAIDI) were assumed to be adequate to define the quality of power. With dramatic increases in computer-controlled loads, momentary interruptions and voltage sags have also become critical performance indices. Momentary Average Interruption Frequency Index (MAIFI) and System Average RMS variation Frequency Index ($SARFI_X$) are the indices used for quantifying momentary interruptions and sags, respectively. These two power quality events have also been recognized as the biggest concern for utilities and their customers because of their tremendous economic impacts on end users.

To measure the quality of the system, power quality monitors are often placed at different feeders across a power system to record RMS voltage variations. The recorded data can be used to calculate and characterize power quality and reliability performance in the system.

Technically, the most accurate and definitive assessment of quality power provided to the customer can be achieved through long-term monitoring at the customer’s point of service delivery. The customer that is requesting such an assessment should pay for monitoring and data analysis. However, it is impractical to wait several years before providing customers information regarding the quality level of the power that they are receiving. In such cases, if the utility has power quality monitoring data from a *statistically* valid system-wide benchmarking project, data from these selected monitoring locations can then be used to estimate the system-wide baseline (“average”) power quality for that particular utility.

In the early 1990s, EPRI initiated a project called Distribution Power Quality (DPQ) that resulted in power quality monitoring at 277 distribution sites, which were statistically selected throughout the United States, to gain valuable knowledge regarding the frequency and severity of power quality events. The data collected for a period of more than two years were analyzed to understand the frequency and severity of different types of power quality events. The data from DPQ—called DPQ I in this report—were used to compose the first and only comprehensive database that characterizes the power quality level in distribution systems. In the absence of customer-specific data, the results from a national survey such as the Distribution Power Quality (DPQ[1]) Project can be used to provide the baseline of power quality to the customer.

As part of the Premium Power Grades Project[2] in 2000 —called DPQ II in this report—the performance range of grid power was characterized using data from the DPQ Project to define base power quality levels in distribution systems. Since the completion of DPQ I in 1995, several utilities have implemented system-wide power quality monitoring programs both in distribution and transmission systems. The wealth of data collected by these utilities provides a unique opportunity to synthesize meaningful information regarding the variability of grid power based on system characteristics, especially information regarding the rate of voltage sags. It also presents a unique opportunity to compare the results of DPQ I with the data from these monitoring programs.

In 2001, there were two primary focuses for the Premium Power Grades Project, as outlined in the EPRI report *Premium Power Grades: Developing Site Selection Criteria for DPQ Phase II*, EPRI, Palo Alto, CA: 2002. EPRI 1005921. The first was to identify the technical characteristics of distribution and transmission sites with permanently installed power quality monitors. This was done through the use of an initial survey. From that survey, two characteristics were selected to stratify the monitor sites: voltage class and lightning flash density. The second focus of the Premium Power Grades Project was to develop site-selection criteria for sampling the available data. This sample would eventually be used to represent the population of monitored sites.

During 2002, we applied the site-selection criteria to the entire population of 1876 monitored sites in the service territories of 13 participating utilities. The utilities then randomly selected sites in each voltage class and lightning-flash-density strata. From this group of sites, we requested more detailed site characteristics and the actual power quality data. The power quality data were delivered in several different formats, so a strategy was devised and executed to integrate that data into a common database. The data were then analyzed in a manner similar to DPQ I. The data were also analyzed according to some of the more recent international standards. Finally, we performed some preliminary interpretation of the data with regard to system compatibility and energy-storage requirements.

---

[1] An Assessment of Distribution System Power Quality: Volumes 1-3, EPRI, Palo Alto, CA: 1996. TR-106294-V1, TR-106294-V2, TR106294-V3.

[2] Understanding Premium Power Grades, EPRI, Palo Alto, CA: 2000. TR-1000406

EPRI's Distribution Power Quality (DPQ) Phase II effort[3] in 2002 resulted in a database of voltage-sag[4] information at numerous sites, including transmission, distribution, and secondary networks. The DPQ II report provides the results of a comprehensive analysis of over 541,339 monitor-days of data that characterize the severity and frequency of short-duration variations in terms of voltage sags, swells, and interruptions. This project resulted in characterizing power quality in terms of short-duration variations such as voltage sags, voltage swells, and voltage interruptions. The characterization was based on analysis of data from 480 power quality monitors at different locations in a power system spanning a date range from August 30, 1993, through December 12, 2002 and included the following:

- Determine which utilities have implemented system-wide power quality monitoring programs.
- Determine what types of monitoring systems have been implemented and how much data are available.
- Determine what types of systems are being monitored, such as distribution and transmission.
- Identify the characteristics of the different data sources that reside within utility service areas and end-use customer sites.
- Determine a statistical approach for selecting monitoring locations from existing sites on the member utility systems.
- Determine what methods of data collection and aggregation are feasible for collecting monitoring data from large, independent monitoring systems.
- Develop a data-collection procedure that will enable data from these sources to be sanitized in order to develop a generic voltage-variation database without any particular reference to proprietary data sources of the utility or end-use customer.
- Categorize the data that are received from different sources into separate power quality parameters[5], as defined in IEEE Std. 1159-1995, *Recommended Practice for Monitoring Electric Power Quality.*
- Analyze the data in a manner similar to DPQ I.
- Analyze the data according to some of the more recent international standards.

The results of DPQ II confirm that events such as voltage sags, swells, and momentary interruptions are far more common than outages. If customer equipment is impacted by these short-duration events, then customer perceptions of quality of service will be different from classical electric utility reliability indices such as SAIFI and SAIDI, which describe the frequency and duration of long-term outages. Understanding the severity and frequency of these short-duration events will provide manufacturers of equipment a target for designing immunity

---

[3] Distribution System Power Quality Assessment: Phase II: Voltage Sag and Interruption Analysis, EPRI, Palo Alto, CA: 2003. 1001678

[4] Reliability data were not collected as a part of this effort

[5] Primary power quality variables of interest for this study were voltage sags and momentary interruptions.

to these events and also provide customers information for specifying equipment sensitivity. The results of DPQ II validate the earlier DPQ I study in terms of frequency of events for different voltage magnitudes and durations.

Besides the DPQ I report published in 1995, this is the only known source of such data in the entire world. The results are not only useful for electric service providers to understand their electrical environment in terms of these events but also very important for equipment manufacturers to understand the electrical environment in which their equipment has to perform.

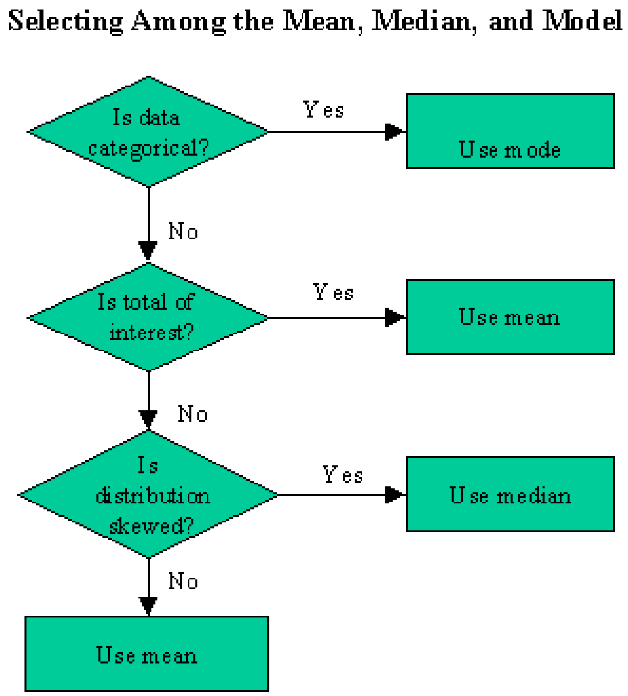
However, this is only the first step in the entire process. The next step is to analyze index data to benchmark the service quality before any contracts can be generated. Without getting into details about statistical distribution of power quality and reliability indices, it is safe to say that power quality and reliability indices in general do not follow a normal distribution, and therefore “average” is not a good representation of the population. In most cases, the distribution is skewed by the performance of a number of feeders. Overall, the median, or “what 50% of the customers will see as minutes-off-supply,” is not equal to the average minutes-off-supply performance of the network.

A Need for Improved Data Analysis

Because data is not information, it is important to ascertain how information and knowledge can be obtained from a set of data. Statistical analysis is often used to make sense of data sets. However, to define statistical data analysis, one must first define statistics. Statistics is a set of methods[6] that is used to collect, analyze, present, and interpret data. In general, statistics can be broadly defined as the study of numerical data to better understand the characteristics of a population or process. In the deterministic framework, inputs are single values. Utilities currently use deterministic methods to assess the risks associated with disturbances to electric power systems. However, these methods tend to produce overly conservative solutions. Traditionally, the average and/or median of an index are used to reflect performance. Data analysis using a single estimate such as average is called *deterministic approach*. To add to this misconception, it is a common mistake to specify the wrong index for central tendency. Figure 1-1 is a simple flow chart used in selecting the appropriate index.

---

[6] Tim Bedford and Roger Cooke, Probabilistic Risk Analysis: Foundations and Methods, Cambridge University Press, 2001



Selecting Among the Mean, Median, and Model

Figure 1-1
Flow Chart for Selecting Appropriate Index

There is an interesting fable about a statistician who drowned in a river with an average depth of three feet! In the deterministic approach, all inputs are assumed to have "most likely" or average values. However, the deterministic approach does not consider *variability* and *uncertainty* in the data. Once this misconception is cleared, the question then is, "How well does a single data parameter such as average assess the entire dataset?"

Variability represents heterogeneity or diversity, which is not reducible through further measurement or study. Fundamentally a *property of nature*, variability arises due to the differences in the value of a quantity among different members of a population. For example, variability might refer to different feeders in a distribution system having different performance in terms of voltage interruptions. In essence, some feeders will perform better than the others due to differences in topology, weather, and existing system conditions, and so on.

Additionally, some years will be more severe in terms of storms, lightning flashes, tornados, and so on than other years that can greatly impact the quality of supply. Regardless of the level of measurement accuracy, this variability cannot be reduced with further measurement. Variability is an intrinsic property of distribution systems that makes each feeder unique in its performance. Therefore variability is present across the system data (spatial variation) as well as over a period of time (temporal variation). Statistical indices such as CP05 (5th percentile), CP50 (50th percentile), CP95 (95th percentile) represents variability in a dataset.

Uncertainty represents lack of knowledge about a poorly characterized phenomenon that is sometimes reducible through further measurement or study. Fundamentally a *property of the risk* *assessment*, uncertainty might be reduced through further measurement. For example, one-year voltage sag measurements at a substation provide an indication of the expected voltage sag performance at the substation for that given year. However, it does not accurately represent the "expected" voltage sag rate in the future. Further measurement over a longer period of time will reduce, but not necessarily eliminate the uncertainty in quantifying the expected voltage sag rate at that substation.

As a result, uncertainty is unavoidable in any service quality dataset. In general, utilities do not have the resources to monitor all the feeders in its service territory. System performance therefore can and should be predicted from the limited information available. Also, monitored data at a feeder will be available only for a limited period of time. Uncertainty is quantified using a desired confidence level (C.I.) or probability band, which also is representative of risk (e.g., 95% C.I. level and 50% C.I. level).

While improvements in modeling the system behavior could lead to more accurate estimates of the *average* values of power quality indices, the deterministic approaches (based on power quality data) do not estimate the variability of these indices. Advanced data analysis techniques helps to see the full range of variability and uncertainty instead of being misled into thinking that service quality indices are point values. In this approach, inputs are treated as random variables coming from a probability distribution. The power quality indices found across the system or over a period of time can be fit into statistical distributions like Poisson, lognormal, or Weibull. The outcome is no longer a single answer but a distribution itself. Compare this with the deterministic approach, where the output is also a single value. The inability of deterministic approach to account for variability and uncertainty could give misleading results. The probabilistic risk assessment, therefore, enables risk managers to see the full range of variability and uncertainty instead of being mislead into thinking that exposure, effects, and eventually risk are point values. Knowledge of the range over which reliability and power quality indices (system as well as feeder level) are expected to vary would be helpful to distribution engineers in making appropriate allocation of the available resources towards the maintenance of distribution system. Such an assessment can be made using the probability distributions of power quality and reliability indices obtained from different statistical techniques.

The large amount of data gathered together as part of EPRI's Distribution Power Quality (DPQ) phase II study represents an opportunity to explore ways to predict voltage sags at locations without monitoring. Each of the DPQ phase II sites has a number of site characteristics that can be used in a statistical model to predict voltage sags. Such predictions can help utilities in many ways. Predictions and prediction limits can define risks associated with performance-based contract. Utilities can use predictions to decide if a given circuit area is under-performing relative to other sites with similar characteristics. Utilities can pass on expected performance numbers to their customers to help them fortify their facilities.

The work performed in 2003 and described in the EPRI report, *Voltage Sag Prediction Model using Distribution Power Quality Phase II Data* (EPRI, Palo Alto, CA: 2003. 1002199) leverages previous year's research and data analysis and provides approaches that can be used to develop and test prediction models for voltage sags based on the DPQ phase II dataset. The DPQ phase II database was also combined with the database from the original DPQ study to form a combined database from which prediction models for voltage sags were developed. The main challenges were finding good models, dealing with sites that had unknown site characteristics, and analyzing the variability in the models. These are statistical regression models built from optimizations of models containing several of the site characteristics available for the DPQ phase II and DPQ phase I sites

The variables most impacting SARFI_x were tried in various model formulations. Only a few site parameters are needed: substation transformer size, number of feeders on the bus, distribution voltage level, total feeder length, lightning ground flash density, and a crude estimate of tree coverage. The prediction model also includes an estimate of variability, which can be portrayed as prediction limits. The DPQ I/DPQ II dataset has a number of sites where some of the site characteristics are unknown. A multiple imputation-based technique was finally adopted to fill in missing data in a manner that allows us to make predictions without the missing data distorting the predictions. A relatively straightforward set of equations for the major SARFI_x voltage sag indices was found.

Project Objectives and Scope

Utilities widely record reliability indices such as SAIFI, SAIDI, number of outages, but most have few power quality monitors. The major focus of this year's activity is focused towards investigating if it may be possible to refine voltage-sag predictions using reliability indices. In a statistical prediction model, the reliability indices would become another site characteristic. This could improve power quality predictions and may also reveal interesting practical ties between power quality and reliability.

This project leverages the analysis work conducted in 2003 and extends this to determine how effectively reliability measures such as SAIFI, SAIDI, and number of outages, predict power quality indices such as momentary interruptions and sags. Our goal is to provide insight on how power quality varies by site, to find ways to predict quality at a site and what circuit parameters cause the variations. Also, can reliability measures along with other circuit parameters be adequate to predict the variations in power quality. Characteristics of the site (either at the feeder level or at the substation level) help identify the major influences on SARFI_x and MAIFI and allow us to predict the level of quality at a location. Among other uses, utilities can use a prediction of the quality at a given site to find good sites to locate sensitive customers, to determine if a site is under-performing based on the given infrastructure, or to quantify the risk of offering premium power services.

The starting point for this year's activity, therefore, was to expand the Distribution Power Quality (DPQ) II dataset and include reliability data as well. As a part of this year's activity, the project team obtained additional new data from two utilities that participated in the DPQ Phase II. The overall objectives for this year's activity can be summarized as follows:

- Compare the shapes of the distributions of power quality and reliability measures to obtain cursory information about correlations between different power quality and reliability measures (for example SARFI_x versus SAIFI or SAIDI versus MAIFI or SARFI_x versus MAIFI, etc.)
  - Observe the goodness-of-fit for reliability and power quality measures and examine how the distributions of SARFI_x compare to those of SAIFI and SAIDI and MAIFI

  - Investigate whether the power quality and reliability indices in general follow a Normal distribution or are the distributions skewed depending on the yearly as well site variations[7] in system performance (type of system, voltage level, exposure, etc.).
  - Examine the variation in power quality and reliability at different locations. Identify if distribution of reliability and power quality changes at feeder level and bus level.
- Apply data from the two utilities that contributed to DPQ phase II to examine the variability and uncertainty of power quality and reliability measures and apply advanced statistical techniques that can capture the site variations as well as time variations that are inherent in power systems
- Evaluate relationships between site characteristics on voltage sags and reliability (including and excluding major storms). How does the correlations change with respect to different feeder characteristics namely total exposure (three-phase as well as single-phase), backbone length, number of feeders, transformer size, load density, percentage underground versus overhead, etc.
- Ascertain if correlation exists between PQ indices and reliability indices. Correlations between $SARFI_X$ (ITIC, 90, 80, 70, 50, 10) with respect to reliability measures such as SAIFI, SAIDI, and outages counts were investigated. Correlation between MAIFI with respect to reliability measures such as SAIFI, SAIDI, and outage counts were also investigated.
- Evaluate if correlations exist between PQ definition of momentary interruptions ($SARFI_{10}$) and reliability definition (MAIFI and $MAIFI_E$).
- Evaluate the effect of including and excluding major storms from the reliability data and how these effect the overall prediction.
- Develop and test prediction models[8] for voltage sags and momentary interruptions based on the two datasets. The variables most impacting $SARFI_X$ and MAIFI were tried in various model formulations. The prediction model also includes an estimate of variability, which can be portrayed as prediction limits
- Ascertain how strong are the correlations between SARFI (or MAIFI) and SAIFI (or SAIDI or Outages) relative to correlations between SARFI (or MAIFI) and one or more site characteristics namely total feeder length or backbone length or number of feeders or transformer size or load density or percentage underground versus overhead.
- Investigate the correlations of PQ indices and the reliability indices as well as the prediction models for voltage sags and momentary interruptions at different levels in the power system (feeder level versus the bus level).

---

[7] For example, variability might refer to different feeders in a distribution system having different performance in terms of voltage interruptions. In essence, some feeders will perform better than the others due to differences in topology, weather, and existing system conditions, and so on.

[8] These are statistical regression models built from optimizations of models containing several of the site characteristics available for the DPQ sites. The main challenges were finding good models, dealing with sites that had unknown site characteristics, and analyzing the variability in the models.

- Compare the linear models with more generalized models that predict the voltage-sag rate as well as momentary interruptions at a particular site, given the feeder reliability indices and site characteristics.

Data from one utility was for the period 1998-2003, while the second utility was from 2001-2003. The main components of the prediction modeling are model fitting, variable selection, and handling missing data and quantifying variability. Various models were analyzed, and the variability and uncertainty in the models was analyzed. A relatively straightforward set of equations for the voltage sag as well as momentary interruption based indices was found. From this analysis, if we chose appropriate sets of site descriptors and reliability indices predictions of Momentary Average Interruption Frequency Index (MAIFI) and all of the major SARFI (System Average RMS Variation Frequency Index) indices can be obtained using the develop predictor model.

Since current measurements were not available from the two utilities, understanding if inherent differences exist in correlations (as well as the prediction models) between power quality and reliability measures for transmission-based faults and distribution-based faults, could not be investigated in this year's activity.

Data Collected from Two Candidate Utilities

As a part of this year's activity, the project team obtained additional new data from two utilities that participated in the DPQ Phase II. The following points should be taken into consideration in the interpretation of the results presented in this report:

- Monitor types
- Monitor settings
- System levels

Monitor Types

Different types of monitors record event data differently depending on sampling rates, triggering algorithms, and data storage. The data collected from the two utilities in essence used similar monitoring devices (Dranetz BMI 8010 and 7100). These power quality monitors sample a waveform at 128 samples per cycle.

Monitor Settings

The triggering thresholds set for the monitors deployed by the two utilities were similar.

System Locations/Levels/Data Type

Data collected from the two utilities fell into two location categories (shown in Figure 1-2 and Figure 1-3):

- **Locations B:** The typical monitoring configuration for this location is to have current transformers (CTs) on the upstream of individual feeder breakers and potential transformers (PTs) on the distribution bus. An example of this type of configuration is shown in Figure 1-3. This is certainly an adequate way to measure voltage sags accurately ($10\%<V<90\%$). However, there will be a difference in interruption counts if the desire is to track interruptions on a particular feeder by voltage alone. The reason is that for PTs located on the distribution bus, the voltage does not necessarily go to zero for faults out on the feeder being monitored, even if the distribution feeder breaker opens.
- **Location E:** The typical monitoring configuration for this location is to have current transformers (CTs) measuring the total currents of the feeders (also referred as the "totalizer") and potential transformers (PTs) on the distribution bus. An example of this type of configuration is shown in Figure 1-2.

Additional information about these two dataset include:

Utility A:

- 19 substations at 13.8 kV
- Reliability data[9] from feeder monitors were obtained
- Power Quality data from bus monitors were obtained
- Data Collected ranged from 1998 through 2003
- Site Characteristics for which data was available include:
  - Total Feeder Exposure (single-phase as well as three-phase)
  - Backbone Feeder Length
  - Percentage Underground feeder
  - Number of Feeders
  - Transformer Rating and Impedance
  - Load Density (number of customers per mile)

Utility B:

- 32 Substations at 12.47 and 13.8 kV (referred to as 15 kV)
- 350 Substations at 22.8 and 24 kV (referred to as 25 kV)
- Reliability data[10] from feeder monitors were obtained
- Power Quality data from bus as well as feeder monitors[11] were obtained.

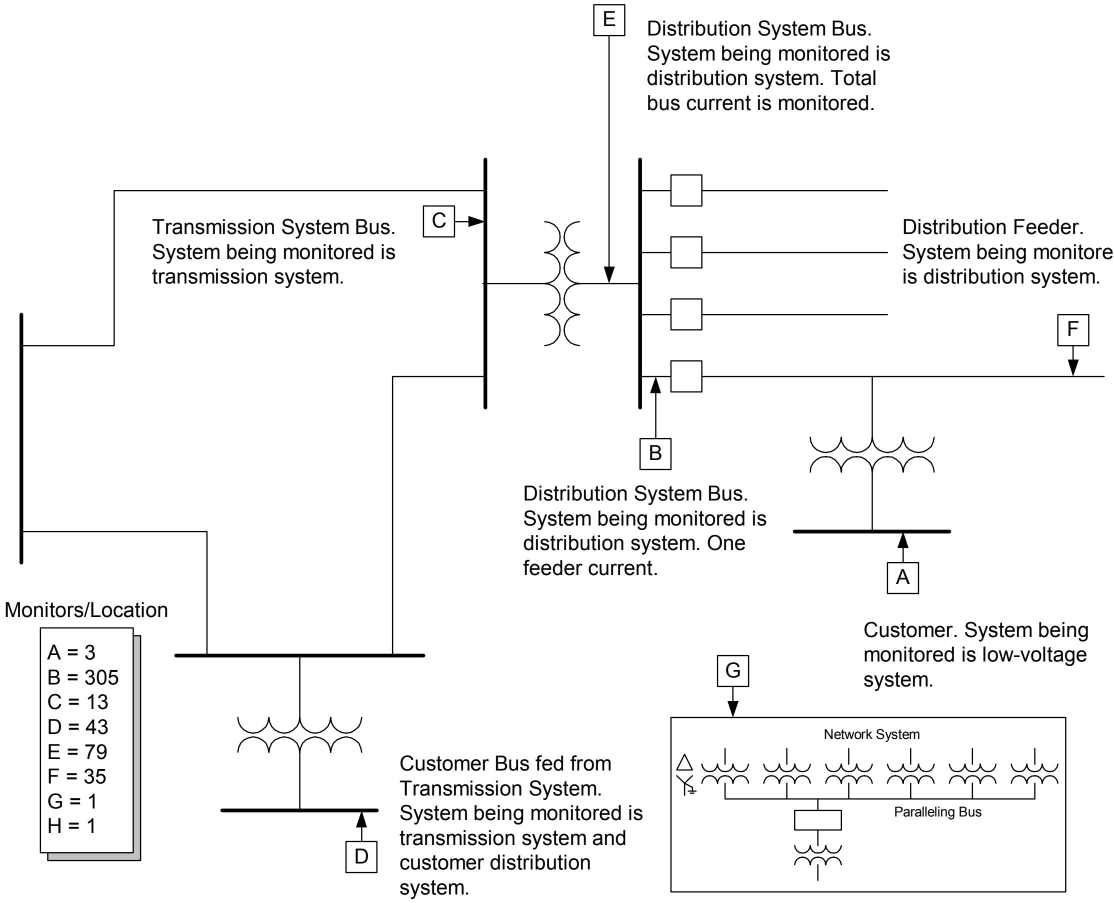
---

[9] MAIFI data was not obtained from Utility A. Only SAIFI, SAIDI, and number of outages were available

[10] SAIFI, SAIDI, and MAIFI were available

[11] Note, feeder monitors were placed upstream to the breaker

- Data Collected ranged from 2001 through 2003
  - Total Feeder Exposure (single-phase as well as three-phase)
  - Backbone Feeder Length
  - Percentage Underground feeder
  - Number of Feeders
  - Transformer Rating and Impedance
  - Load Density (number of customers per mile)



Distribution System Bus. System being monitored is distribution system. Total bus current is monitored.

Transmission System Bus. System being monitored is transmission system.

Distribution Feeder. System being monitored is distribution system.

Distribution System Bus. System being monitored is distribution system. One feeder current.

Customer. System being monitored is low-voltage system.

Customer Bus fed from Transmission System. System being monitored is transmission system and customer distribution system.

Monitors/Location

A = 3
B = 305
C = 13
D = 43
E = 79
F = 35
G = 1
H = 1

Figure 1-2
Location Categories for the Sample Sites

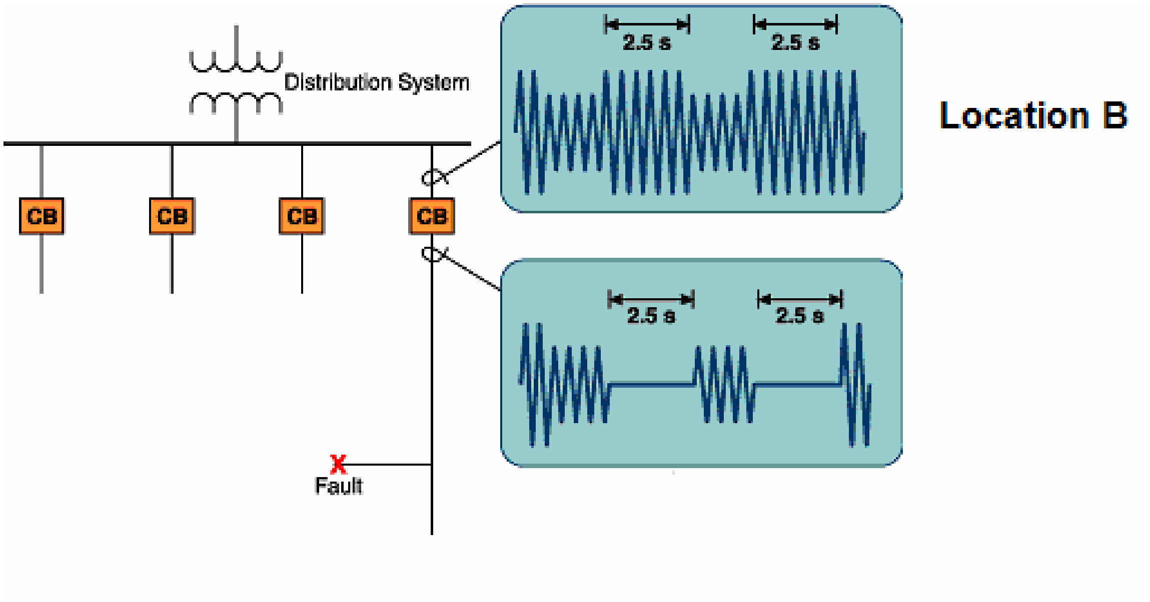


Figure 1-3
Monitor Connection for Location B

2

EVALUATING TIME AND SPATIAL VARIABILITY IN POWER QUALITY AND RELIABILITY INDICES

The first step in finding a model to predict voltage sags and momentary interruptions is to understand the spatial and temporal variations that exist in the dataset. It provides insight into the concepts of variability and uncertainty in data. These provide the basics of more in-depth analysis that can be used to assess power quality service from monitored data.

Topics covered in this chapter include the following:

- Compare the shapes of the distributions of power quality and reliability measures to obtain cursory information about correlations between different power quality and reliability measures (for example SARFI_X versus SAIFI or SAIDI versus MAIFI or SARFI_X versus MAIFI, etc.)
  - Observe the goodness-of-fit for reliability and power quality measures and examine how the distributions of SARFI_X compare to those of SAIFI and SAIDI and MAIFI
  - Investigate whether the power quality and reliability indices in general follow a Normal distribution or are the distributions skewed depending on the yearly as well site variations[12] in system performance (type of system, voltage level, exposure, etc.).
  - Examine the variation in power quality and reliability at different locations. Identify if distribution of reliability and power quality changes at feeder level and bus level.

  Apply data from the two utilities that contributed to DPQ phase II to examine the variability and uncertainty of power quality and reliability measures and apply advanced statistical techniques that can capture the site variations as well as time variations that are inherent in power systems

Step-by-step procedures that were followed to address these objectives are provided. Comprehensive evaluation of these procedures was carried out using two separate datasets. Data from one utility was for the period 1998-2003, while the second utility was from 2001-2003.

- ***Performance Indicators of Interest:*** Identify the metrics (SAIFI, SAIDI, Outages, MAIFI, SARFI_X) to measure supply quality.

---

[12] For example, variability might refer to different feeders in a distribution system having different performance in terms of voltage interruptions. In essence, some feeders will perform better than the others due to differences in topology, weather, and existing system conditions, and so on.

- ***Compare Shape and Distributions of Power Quality and Reliability measures using Advanced Analysis:*** If advanced statistical software is available, more advanced techniques can be used to account for spatial and temporal variability and uncertainty. Two types of approaches, namely parametric and nonparametric, can be used to account for data uncertainty and variability. One may ask, "Which statistical approach does one need to use?" The answer to this lies in assessing the distributional adequacy of a given dataset through the use of probability plots and "goodness-of-fit" tests. Simple steps that may be performed could include:
  - ***Applying Parametric-Based Techniques:*** If a given dataset fits a standard statistical distribution (Lognormal, Weibull, Gamma, etc.) with reasonable accuracy, use parametric methods to estimate variability and uncertainty by constructing the cumulative distribution function (CDF) and probability density function (PDF).
  - ***Applying Non Parametric-Based Techniques:*** If a given dataset does not fit a standard statistical distribution, apply non-parametric methods (Bootstrap, Monte-Carlo, etc.) to estimate variability and uncertainty by constructing the cumulative distribution function (CDF) and probability density function (PDF).

Measures for Distribution Reliability and Power Quality

Today, there are in excess of 40 distribution reliability indices, many of which are defined in various Institute of Electrical and Electronic Engineers, Inc. (IEEE), EPRI, and other industry-related documents. Reliability indices for distribution systems were defined as early as the 1970s when The Edison Electric Institute Transmission and Distribution Committee developed a *Guide for Reliability Measurement and Data Collection.* In general terms the measurement indexes that have proven to be most useful and meaningful in power distribution system design are: [13, 14]

- Load interruption frequency (number/unit time)
- Expected duration of load interruption events (time)

These indexes can be computed and then used to compute other indexes that are useful:

- Total expected (average) interruption time per year (or other time period)
- System availability or unavailability as measured at the load supply point in question
- Expected demanded, but unsupplied, energy per year

The two most commonly used indices based on sustained interruptions are the Service Average Interruption Frequency Index (SAIFI) and Service Average Interruption Duration Index (SAIDI). SAIFI is a measure of the number of sustained interruptions an average customer experiences each year and SAIDI is a measure of the total duration of sustained interruptions that a customer experiences every year. Also popular for sustained interruptions are indices on Customer

---

[13] Ayoub, A. K., and Patton, A. D., "A frequency and duration method for generating system reliability evaluation," IEEE Transactions on Power Apparatus and Systems, Nov. Dec. 1976, pp. 1929–1933.

[14] Billinton, R., and Allan, R. N., "Reliability Evaluation of Power Systems," Plenum Publishing Corp., 1983.

Average Interruption Duration (CAIDI) and Average Service Availability Index (ASAI). CAIDI and ASAI are indices that can be derived from SAIFI and SAIDI.

Based on the IEEE Standard 1366, “Trial-Use Guide for Electric Power Distribution Reliability Indices”, the following provides definitions for some of the indices. This document was originally published in 1998 and since revised and published in 2001 entitled, IEEE “Guide for Electric Power Distribution Reliability Indices”.

The following basic factors specify the data needed to calculate the indices:

i	An interruption event;
r_i	Restoration time for each interruption event;
E	Event;
T	Total;
ID_i	Number of interrupting device operations;.
ID_E	Interrupting device events during reporting period;
N_i	Number of interrupted customers for each interruption event during reporting period;
N_T	Total number of customers served for the area being indexed;
$_i$	Connected kVA load interrupted for each interruption event;
$_T$	Total connected kVA load served;
$CN_{(k>n)}$	Total number of customers who have experienced more than n sustained interruptions during the reporting period;
CN	Total number of customers who have experienced a sustained interruption during the reporting period;
$CNT_{(k>n)}$	Total number of customers who have experienced more than n sustained interruptions and momentary interruption events during the reporting period;
k	Number of interruptions experienced by an individual customer in the reporting period.

<u>SAIFI</u> System average interruption frequency index (sustained interruptions). This index is designed to give information about the average frequency of sustained interruptions per customer over a predefined area. In words the definition is:

$$SAIFI = \frac{Total\ number\ of\ customer\ interruptions}{Total\ number\ of\ customers\ served} \qquad \textbf{Eq. 2-1}$$

To calculate the index use the following equation:

$$SAIFI = \frac{\sum N_i}{N_T} \qquad \textbf{Eq. 2-2}$$

SAIDI System average interruption duration index. This index is commonly referred to as customer minutes of interruption or customer hours, and is designed to provide information about the average time the customers are interrupted. In words, the definition is:

$$SAIDI = \frac{\sum Customer\ interruption\ durations}{Total\ number\ of\ customers\ served} \qquad \textbf{Eq. 2-3}$$

To calculate the index, use the following equation:

$$SAIDI = \frac{\sum r_i N_i}{N_T} \qquad \textbf{Eq. 2-4}$$

CAIDI Customer average interruption duration index. CAIDI represents the average time required to restore service to the average customer per sustained interruption. In words, the definition is:

$$CAIDI = \frac{\sum Customer\ interruption\ durations}{Total\ number\ of\ customer\ interruptions} \qquad \textbf{Eq. 2-5}$$

To calculate the index use the following equation:

$$CAIDI = \frac{\sum r_i N_i}{N_T} = \frac{SAIDI}{SAIFI} \qquad \textbf{Eq. 2-6}$$

ASAI Average service availability index. This index represents the fraction of time (often in percentage) that a customer has power provided during one year or the defined reporting period. In words, the definition is:

$$ASAI = \frac{Customer\ hours\ service\ availability}{Customer\ hours\ service\ demand} \qquad \textbf{Eq. 2-7}$$

To calculate the index use the following equation:

$$ASAI = \frac{N_T \text{ x (No. of hours/year) - } \sum r_i N_i}{N_T \text{ x (No. of hours/year)}} \qquad \textbf{Eq. 2-8}$$

System indices for reporting momentary outages were also introduced in the IEEE Standard 1366. In the past many short-term outages were not reported. These are of particular consequence when determining service reliability for digital or highly automated process industries, where a very short interruption has about the same impact on the process as a sustained outage. The most used of the momentary outage indices is MAIFI.

MAIFI Momentary average interruption frequency index. This index is very similar to SAIFI, but it tracks the average frequency of momentary interruptions. In words, the definition is:

$$MAIFI = \frac{Total\ number\ of\ customer\ momentary\ interruptions}{Total\ number\ of\ customers\ served} \qquad \textbf{Eq. 2-9}$$

To calculate the index, use the following equation:

$$MAIFI = \frac{\sum ID_i N_i}{N_T} \qquad \textbf{Eq. 2-10}$$

Power Quality Prediction

SARFI Statistics

Several power quality indices have been introduced that are similar to the reliability index SAIFI (System Average Interruption Frequency Index). Utilities can use these for some of the same purposes as reliability indices: targeting areas for maintenance and circuit upgrades, tracking the performance of regions, and documenting performance to regulators. The most widely used power quality index is SARFI, defined as:[15,16]

> SARFI$_X$, System Average RMS (Variation) Frequency Index: SARFI$_X$ represents the average number of specified rms variation measurement events that occurred over the assessment period per customer served, where the specified disturbances are those with a magnitude less than *X* for sags or a magnitude greater than *X* for swells.

$$SARFI_x = \frac{\sum N_i}{N_T} \qquad \textbf{Eq. 2-11}$$

where,

X = rms voltage threshold; possible values - 140, 120, 110, 90, 80, 70, 50, and 10

It should be noted that system indices are calculated by weighting disturbance characteristics according to the number of customers or connected kVA experiencing the disturbance. Detailed data on the number of customers and connected kVA were not gathered as part of this project. Consequently, this information is not available for calculating SARFI$_x$. Also, the SARFI index has evolved over the years from a system-wide index to a more site-specific index where the

---

[15] Reliability Benchmarking Application Guide for Utility/Customer PQ Indices, EPRI, Palo Alto, CA: 1999. 113781.

[16] Sabin, D. D., Grebe, T. E., and Sundaram, A., "RMS Voltage Variation Statistical Analysis for a Survey of Distribution System Power Quality Performance," IEEE/PES Winter Meeting Power, February 1999.

$SARFI_X$ equation reduces to a count of the number of sags that have a magnitude below the specified RMS voltage threshold (X). Therefore, for this analysis, $SARFI_x$ is taken as raw count of the number of sags less than magnitude X.

The breakpoints were not chosen arbitrarily. The 90, 80, and 70% thresholds are boundaries of the ITI curve, the 50% threshold is a typical breakpoint for motor contactors, and 10% is the dividing line between a sag and an interruption. Two special variations of SARFI have also been defined. $SARFI_{ITIC}$ is the number of events below the lower ITI (Information Technology Industry Council) curve. In similar fashion, $SARFI_{SEMI}$ is the number of events below the SEMI (Semiconductor Equipment and Materials International) curve.

For this analysis, we used a one-minute aggregation for events, and we excluded events with duration longer than one minute.

The analysis in this report focuses mainly on $SARFI_{ITIC}$, but several other SARFI values are available in the combined dataset, including $SARFI_{90}$, $SARFI_{80}$, $SARFI_{70}$, and $SARFI_{50}$. Table 2-1 and Table 2-4 provide the statistics of the two datasets that were used in this year's analysis. The dataset for Utility B were divided based on the voltage levels. For Utility B, Feeders at 12.47 and 13.8Kv were combined and this dataset was referred to as "15KV Class". For Utility B, Feeders at 22.8 and 24 kV were combined and this dataset was referred to as "25KV Class".

Table 2-2 and Table 2-4 shows how the site median and upper and lower quartiles changed with location and with survey. Note the wide variation measured at different sites. For the purpose of comparison, Table 2-5 provides similar results that were obtained using DPQ phase I as well DPQ phase II dataset.

Table 2-1
SARFI Site Medians for Utility A and Utility B (15KV and 25KV Class)

	Utility A		Utility B	
	Feeder (15KV Class)	Bus (15KV Class)	Bus (15KV Class)	Bus (25KV Class)
$SARFI_{ITIC}$	9.00	8.00	21.0	45.0
$SARFI_{90}$	31.61	30.16	45.0	96.0
$SARFI_{80}$	13.00	12.00	33.0	74.0
$SARFI_{70}$	7.00	6.02	20.0	46.0
$SARFI_{50}$	3.00	2.23	6.0	17.0

Table 2-2
$SARFI_{ITIC}$ Site Statistics for Utility A and Utility B by Survey and Location (15KV and 25KV Class)

		1ST Quartile	Median	3rd Quartile
Survey	**Site Location**	**P(<25%)**	**P(<50%)**	**P(<75%)**
Utility A	Feeder (15KV Class)	6.0	9.0	12.0
	Bus (15KV Class)	6.0	8.0	11.8
Utility B	Bus (15KV Class)	11.0	21.0	37.0
	Bus (25KV Class)	31.0	45.0	69.0

Table 2-3
SARFI Site Medians for Utility B (15KV and 25KV Class)

	Bus (15KV Class)	Bus (25KV Class)
MAIFI (including storms)	9.8	10.3
MAIFI (excluding storms)	7.6	8.2

Table 2-4
MAIFI (Including and Excluding Storms) Site Statistics for Utility B by Survey and Location (15KV and 25KV Class)

		1ST Quartile	Median	3rd Quartile
Survey	**Site Location**	**P(<25%)**	**P(<50%)**	**P(<75%)**
Utility B	MAIFI (including storms) Bus (15KV Class)	5.7	9.8	15.4
	MAIFI (excluding storms) Bus (15KV Class)	7.4	10.3	14.1
	MAIFI (including storms) Bus (25KV Class)	4.7	7.6	12.3
	MAIFI (excluding storms) Bus (25KV Class)	5.99	8.2	11

Table 2-5
$\text{SARFI}_{\text{ITIC}}$ Site Statistics by Survey and Location for DPQ Phase I and II

			Median	
Survey	Site Location	P(<25%)	P(<50%)	P(<75%)
DPQ II	Transmission (N=53)	3.5	6.3	10.7
DPQ II	Substation (N=183)	5.7	9.4	16.1
DPQ II	Feeder (N=217)	14.4	22.1	35.9
DPQ I	Substation	9.8	17.9	32.6
DPQ I	Feeder	11.5	21.3	36.7

In general, we use the site median or actual distributions to characterize or compare the effects of site parameters. As an indicator, the average or mean misrepresents the typical site power quality. The median represents site data better, where by definition, 50% of sites have values higher than the median, and 50% have values lower. With balanced distributions such as the normal distribution, the average equals the median. In a skewed distribution like we see with the power quality data, the average is higher than the median. Additionally, poor sites and anomalies such as a severe storm skew the average upward.

In DPQ I, the feeder sites had a median that was 20% higher than substation sites (21.3 vs. 17.9). In DPQ II, the feeder site median was 2.4 times the substation median (22.1 vs. 9.4). Much of this difference is due to the difference in where the monitors were located in the two studies. In the DPQ I study, the substation monitors were all downstream of the substation breaker. Whether monitoring the substation bus or the substation feeder, the monitors in the DPQ II were located upstream of the substation breaker. An example of this type of configuration is shown in Figure 1-3. What is seen as an interruption (zero voltage) on the distribution feeder and by monitors in the DPQ I study is seen as just a voltage sag in the DPQ II study. The DPQ II study more accurately shows what happens on the substation bus; the DPQ I study more accurately shows what happens to voltages just outside of the substation.

Data collected from the two utilities (Utility A and Utility B) fell into two location categories (shown in Figure 1-2 and Figure 1-3):

- **Locations B:** The typical monitoring configuration for this location is to have current transformers (CTs) on the upstream of individual feeder breakers and potential transformers (PTs) on the distribution bus. An example of this type of configuration is shown in Figure 1-3. This is certainly an adequate way to measure voltage sags accurately (10%<V<90%). However, there will be a difference in interruption counts if the desire is to track interruptions on a particular feeder by voltage alone. The reason is that for PTs located on the distribution bus, the voltage does not necessarily go to zero for faults out on the feeder being monitored, even if the distribution feeder breaker opens.

- **Location E:** The typical monitoring configuration for this location is to have current transformers (CTs) measuring the total currents of the feeders (also referred as the "totalizer") and potential transformers (PTs) on the distribution bus. An example of this type of configuration is shown in Figure 1-2. The locations for the two datasets (Utility A and Utility B)

Figure 2-1 shows this variability with cumulative distributions by monitoring location. With this graph, we see that 80% of distribution feeder sites have a $\text{SARFI}_{\text{ITIC}}$ of at least 5 events per year.

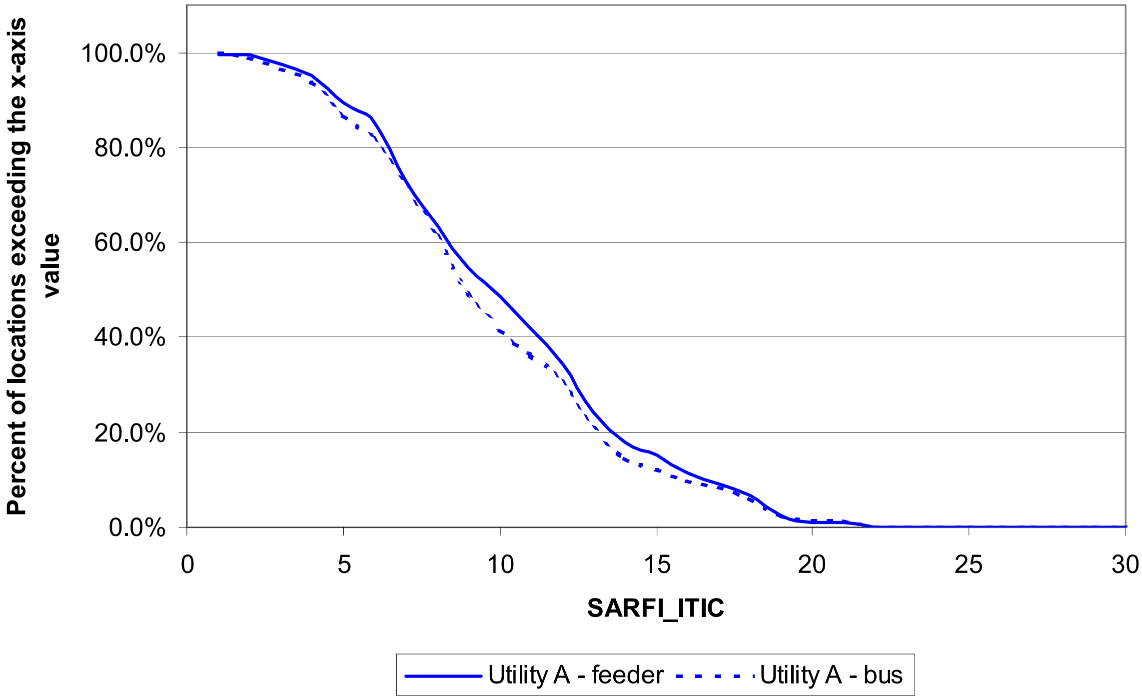


Figure 2-1
Cumulative Distribution of $\text{SARFI}_{\text{ITIC}}$ by Monitoring Location

Probability Distributions of Power Quality and Reliability Variables

In general, supply quality data have two types of variability and uncertainty: spatial and temporal. *Spatial* variability and uncertainty occur due to the fact that there will be differences even within a distribution system due to load profile, bus design, geographical conditions, and many other factors. Specifically, some feeders will perform better than the others. *Temporal* variability and uncertainty account for the fact that the performance of a distribution system will vary over a period of time. Temporal variation and uncertainty occur due to yearly variations in weather patterns, aging of equipment, and design changes, among other causes. Therefore, it is necessary to consider both spatial and temporal variability in order to accurately calculate the chosen quality of supply indices.

Also, based on the ongoing research activities in this area, it has been proven in the past that the distribution of most reliability and power quality indices do not appropriately fit into a normal distribution. Instead they have a large peak on the left and a long tail on the right, and exhibit strong kurtosis, or skew (see Figure 2-6). In a skewed distribution, the average is higher than the median. Additionally, poor performing sites and anomalies such as severe storms skew the average upward. Since the distribution is not normal, all the characteristics derived based on the assumption that the distribution is normal are invalid.

Subsequent sections in this chapter will focus on illustrating the use of some of these techniques to compare the shapes of the distributions of power quality and reliability measures for the two datasets. Aim is to obtain cursory information about correlations between different power quality and reliability measures (for example SARFI_x versus SAIFI or SAIDI versus MAIFI or SARFI_x versus MAIFI, etc.)

Goodness-of-Fit Tests

Advanced probability techniques such as parametric and nonparametric can be used to account for the inherent skewness that may exist for a dataset and thereby can provide a more accurate estimate of spatial and temporal variability and uncertainty. Advanced probability methods enable the reporting of indices based on probability/percentile representation that is more accurate than a simple deterministic use of numbers such as an *average* or *median* of the given index. Based on the ongoing research [17] [18] activities in this area, it is known that most of the reliability and power quality data can be fitted into a standard statistical distribution with reasonable accuracy. Examples of different techniques that can be utilized to evaluate uncertainties and variability include parametric bootstrapping, the Bayesian approach, and maximum likelihood estimation (MLE). A comprehensive explanation of all these methods can be found in most statistical reference books [19] [20] [21] [22] [23] [24] and advanced software manuals [25].

---

[17] An Assessment of Distribution System Power Quality: Volumes 1–3, EPRI, Palo Alto, CA: 1996. TR-106294-V1, TR-106294-V2, TR106294-V3.

[18] Nagaraj Balijepalli, “Advances in Distribution System Reliability Assessment,” Ph.D. Dissertation, Iowa State University, 2002.

[19] Tim Bedford and Roger Cooke, Probabilistic Risk Analysis: Foundations and Methods, Cambridge University Press, 2001

[20] Gerald J. Hahn and William Q. Meeker, Statistical Intervals: A Guide for Practitioners, John Wiley & Sons, Inc., 1998.

[21] William Q. Meeker and Luis A. Escobar, Statistical Methods for Reliability Data, John Wiley & Sons, Inc., 1998.

[22] Elsayed A. Elsayed, Reliability Engineering, Addison Wesley Longman Inc., 1996.

[23] Wallace R. Blischke and D.N. Prabhakar Murthy, Reliability: Modeling, Prediction, and Optimization, John Wiley & Sons Inc., 2000.

[24] A. C. Davison and D. V. Hinkley, Bootstrap Methods and Their Application, Cambridge University Press, 1997.

In parametric approaches, a statistical distribution is first assumed for evaluating variability and uncertainty bands. The data is first fitted to a statistical distribution (Lognormal, Weibull, and so on). Probability plots are utilized to test the goodness-of-fit for a given dataset. A probability plot (see Figure 2-2) can be used to obtain graphical estimates of parametric model parameters by fitting a straight line through the points on a probability plot. If the dataset deviates considerably from this straight line, a nonparametric framework may be more desirable.

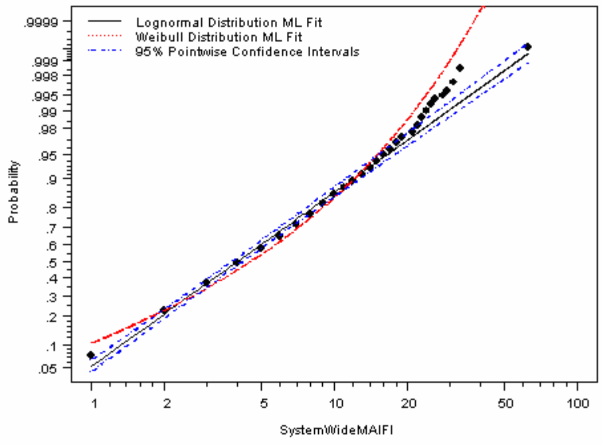


Figure 2-2
A Lognormal Probability Plot

Exploratory data analysis techniques such as goodness of fit tests, density plots, probability density curves, and cumulative distribution plots provide a better understanding of whether or not assumptions of statistical correlations and statistical differences in data sets are valid. Goodness of fit (GOF) tests were performed to formally test the specific distribution (Lognormal, Normal, Weibull, and so on) of the two utility's datasets. When goodness of fit correlations are performed, three methods are typically used:

- Chi-squared
- Kolmogorov-Smirnov
- Anderson-Darling

The latter two are used in this analysis. Detailed explanations of these methods may be found in any advanced statistics book. The results are shown in Tables 2-6 and 2-7.

---

[25] S-PLUS Resample Library User’s Manual, Beta Release 1, Insightful Corporation, 2002.

Table 2-6
Bus Level Distributions of Power Quality and Reliability Variables for Utility A

Variable	Storms	Method[26]	Co-efficient	Best Fit	Normal	Tri-angular	Log-Normal	Uniform	Expo-nential	Weibull	Gamma
SAIFI	Inc.	K-S	0.0485	Gamma	0.0984	0.1161	0.0884	0.2735	0.2055	0.0672	0.0485
		A-D	0.2462	Gamma	1.4039	1.1315	0.9594	8.5167	6.6814	0.4628	0.2462
SAIFI	Ex.	K-S	0.0772	Gamma	0.1390	0.2072	0.1671	0.3557	0.1975	0.0928	0.0772
		A-D	0.4685	Gamma	1.9626	3.8520	3.5747	14.8731	5.3647	0.7090	0.4685
SAIDI	Inc.	K-S	0.0652	Gamma	0.1074	0.3534	0.1033	0.5246	0.1552	0.0803	0.0652
		A-D	0.3452	Gamma	2.3259	15.0524	0.9879	37.8020	3.8305	0.6402	0.3452
SAIDI	Ex.	K-S	0.0518	Gamma	0.1062	0.1112	0.1661	0.2669	0.2180	0.0534	0.0518
		A-D	0.2976	Weibull[27]	0.9036	1.2310	3.5526	8.4896	4.7246	0.2976	0.3154
Outage	Inc.	K-S	0.0800	Gamma	0.1212	0.2723	0.1182	0.4584	0.1814	0.0850	0.0800
		A-D	0.3102	Gamma	2.2512	10.1483	0.7928	30.9443	4.3032	0.4278	0.3102
Outage	Ex.	K-S	0.0693	Weibull[2]	0.1106	0.2625	0.1205	0.4289	0.1758	0.0693	0.0736
		A-D	0.3370	Gamma	1.7083	8.1207	1.0788	26.3821	4.5973	0.3975	0.3370
$\text{SARFI}_{\text{ITIC}}$	Inc.	K-S	0.0792	Gamma	0.1099	0.1287	0.2882	0.2518	0.2931	0.0913	0.0792
		A-D	0.3535	Gamma	0.7522	1.5545	12.6062	7.2239	9.3706	0.5235	0.3535
SARFI_{90}	Inc.	K-S	0.0531	Gamma	0.0695	0.1840	0.0896	0.3429	0.2982	0.0645	0.0531
		A-D	0.2551	Gamma	0.4277	4.6450	1.0028	14.8407	12.3989	0.3112	0.2551

---

[26] K-S =Kolmogorov-Smirnov, A-D=Anderson-Darling.

[27] Next best fit is Gamma distribution.

Table 2-6 (cont.)
Bus Level Distributions of Power Quality and Reliability Variables for Utility A

Variable	Storms	Method[28]	Co-efficient	Best Fit	Normal	Tri-angular	Log-Normal	Uniform	Expo-nential	Weibull	Gamma
$SARFI_{80}$	Inc.	K-S	0.1004	Gamma	0.1401	0.2185	0.1640	0.2999	0.2920	0.1064	0.1004
		A-D	0.4951	Gamma	0.9721	4.6829	1.5702	12.6641	10.5878	0.5969	0.4951
$SARFI_{70}$	Inc.	K-S	0.0929	Triangular	0.1279	0.0929	0.2893	0.2155	0.2850	0.1013	0.1003
		A-D	0.6981	Gamma	1.0688	0.8272	11.5783	5.2940	8.2873	0.7285	0.6981
$SARFI_{50}$	Inc.	K-S	0.1586	Weibull[2]	0.1675	0.2102	0.4215	0.3261	0.1707	0.1586	0.2615
		A-D	2.3468	Normal	2.3468	8.4434	15.9031	22.4728	23.4769	4.2223	6.1325
$SARFI_{10}$	Inc.	K-S	0.4091	Normal	0.4091	0.7385	0.4637	0.7319	0.7439	0.7439	0.4760
		A-D	15.6145	Normal	15.6145	191.2005	18.6844	169.5863	423.1359	614.9885	16.9207

---

[28] K-S =Kolmogorov-Smirnov, A-D=Anderson-Darling.

Table 2-7
Bus Level Distributions of Power Quality and Reliability Variables for Utility B, 15 kV Class.

Variable	Storms	Method[29]	Co-efficient	Best Fit	Normal	Tri-angular	Log-Normal	Uniform	Expo-nential	Weibull	Gamma
SAIFI	Inc.	K-S	0.0504	Gamma	0.1175	0.3362	0.0806	0.5185	0.1610	0.0516	0.0504
		A-D	0.1725	Gamma	2.2412	17.5764	0.5368	45.2140	3.7943	0.2619	0.1725
SAIDI	Inc.	K-S	0.0469	Log-normal[30]	0.2951	0.6401	0.0469	0.7205	0.2326	0.1034	0.1238
		A-D	0.1994	Log-normal[4]	13.3411	90.4101	0.1994	120.0218	8.2317	1.4100	1.7369
MAIFI	Inc.	K-S	0.0722	Gamma	0.1836	0.3815	0.3493	0.5389	0.1261	0.1107	0.0722
		A-D	0.7088	Gamma	5.8170	25.8092	19.9585	59.1409	2.6839	1.3866	0.7088
$SARFI_{ITIC}$	Inc.	K-S	0.0663	Weibull[4]	0.1776	0.3078	0.2913	0.4559	0.0716	0.0663	0.1175
		A-D	0.9775	Weibull[4]	5.7299	21.5109	17.3104	53.7674	5.4300	0.9775	1.2814
$SARFI_{70}$	Inc.	K-S	0.0697	Weibull[4]	0.1748	0.3044	0.3060	0.4530	0.0788	0.0697	0.1148
		A-D	1.0581	Weibull[4]	5.6567	21.9590	18.0154	54.1384	6.9449	1.0581	1.4549
$SARFI_{50}$	Inc.	K-S	0.0960	Weibull[4]	0.1967	0.3647	0.3737	0.5133	0.1209	0.0960	0.1404
		A-D	1.8827	Weibull[4]	7.5688	33.9210	17.9540	68.4315	12.0623	2.0049	1.8827

---

[29] K-S =Kolmogorov-Smirnov, A-D=Anderson-Darling.

[30] Next best fit is Gamma distribution.

Statistical inference (q-q plots) based tests were performed to evaluate how the shape of the distribution amongst reliability and power quality indices compare. By evaluating the cumulative distribution functions (CDFs) for the 25th, 50th, and 75th percentiles, we compared variability amongst reliability and power quality indices. These are shown in Figure 2-3 through Figure 2-7.

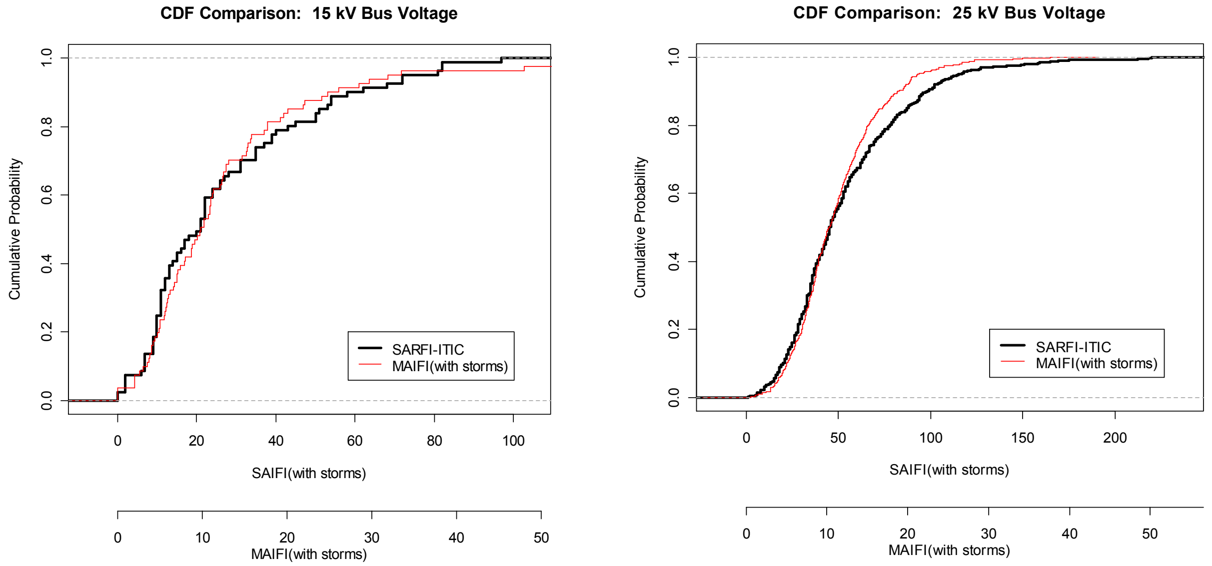


Figure 2-3
Graphical Comparisons of CDFs Distribution of SARFI$_{ITIC}$ and MAIFI for Utility B

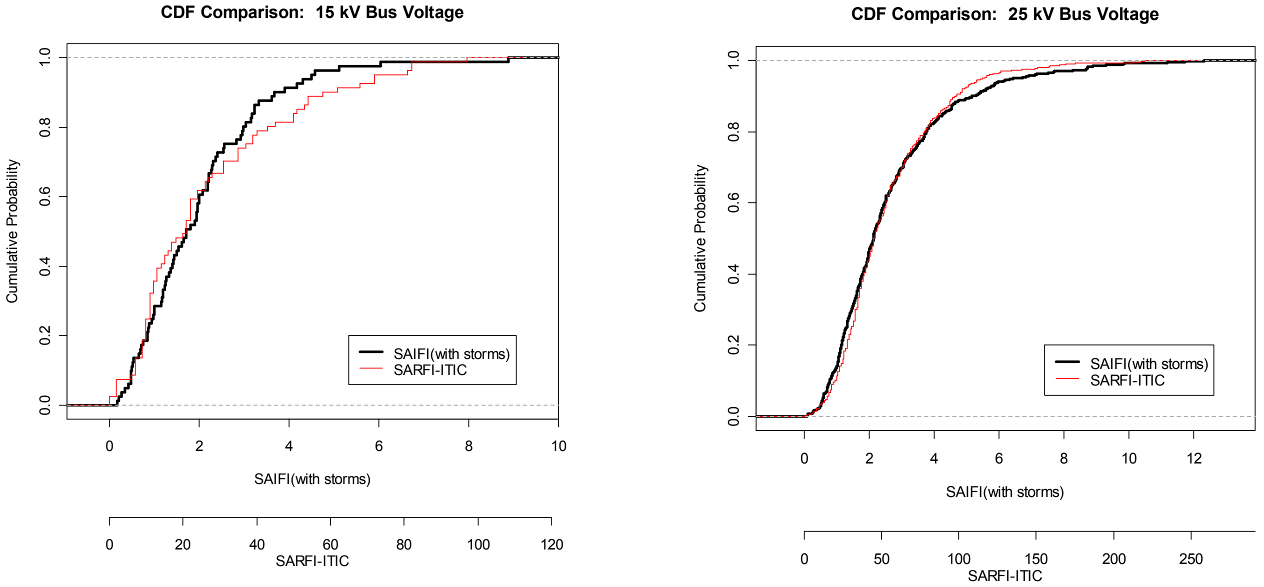


Figure 2-4
Graphical Comparisons of CDFs Distribution of SARFI$_{ITIC}$ and SAIFI for Utility B

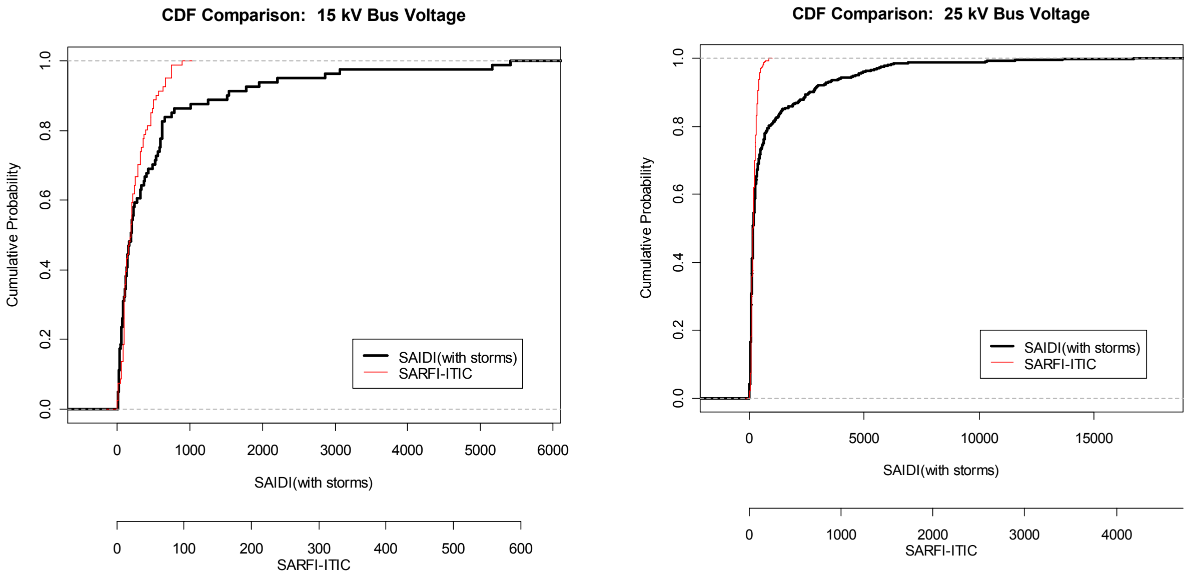


Figure 2-5
Graphical Comparisons of CDFs Distribution of SARFI$_{ITIC}$ and SAIDI for Utility B

Individual plots of the PDF and CDF distributions of SAIFI, SAIDI, outages, as well as SARFI were then calculated based on the fitted distribution. Figure 2-6 through Figure 2-16 show a right skew, which is best matched by a generalized exponential distribution, such as Gamma, Weibull and Lognormal. The figures shown are including storms, but similar results are seen with storms excluded (see Appendix). Similarly, with SARFI, the distributions for SARFI$_{ITIC}$, SARFI$_{70}$ and SARFI$_{50}$ are shown here, with other SARFI indices in the Appendix.

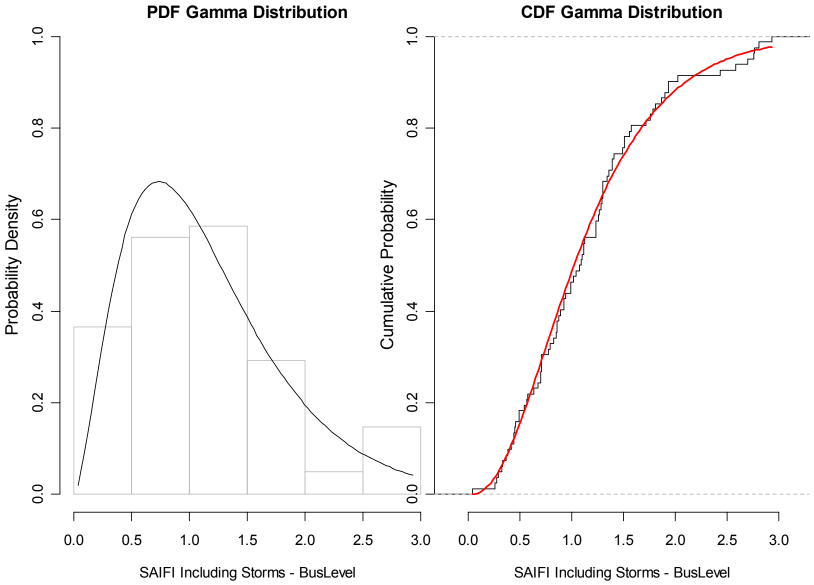


Figure 2-6
SAIFI Gamma Distribution, Bus Level, Including Storms for Utility A

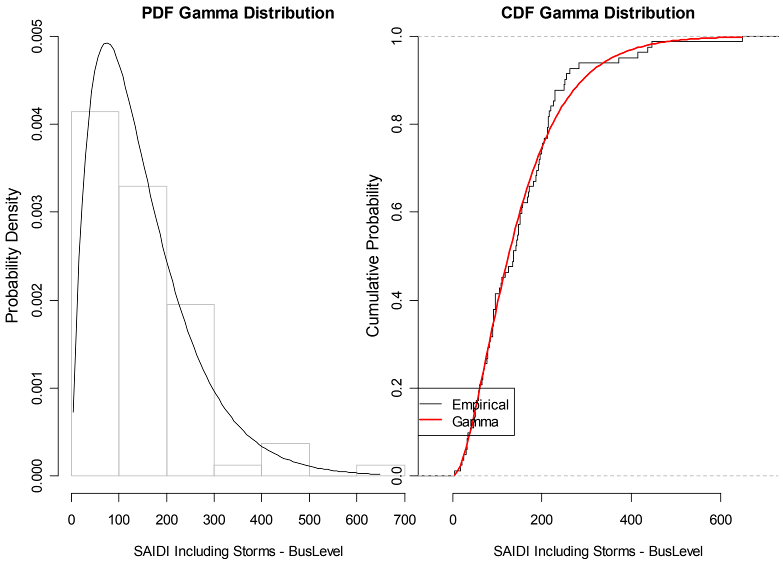


Figure 2-7
SAIDI Gamma Distribution, Bus Level, Including Storms for Utility A

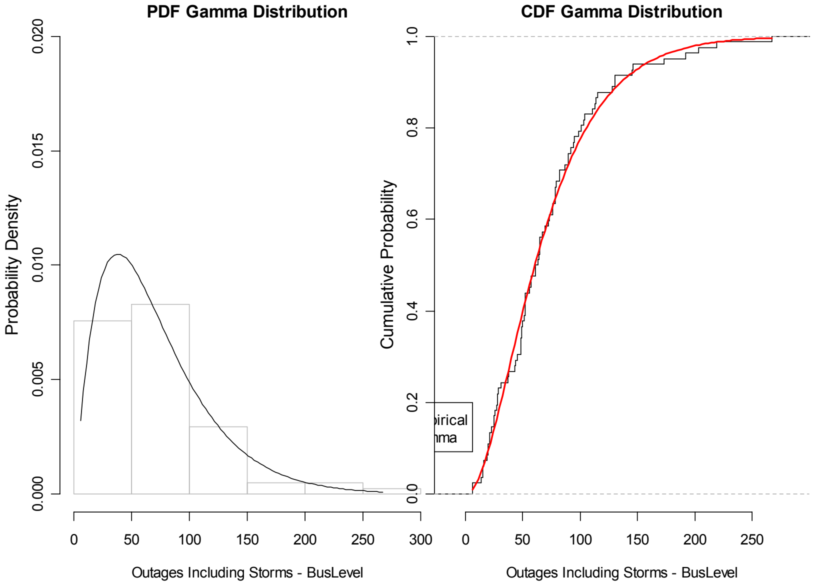


Figure 2-8
Outages Gamma Distribution, Bus Level, Including Storms for Utility A

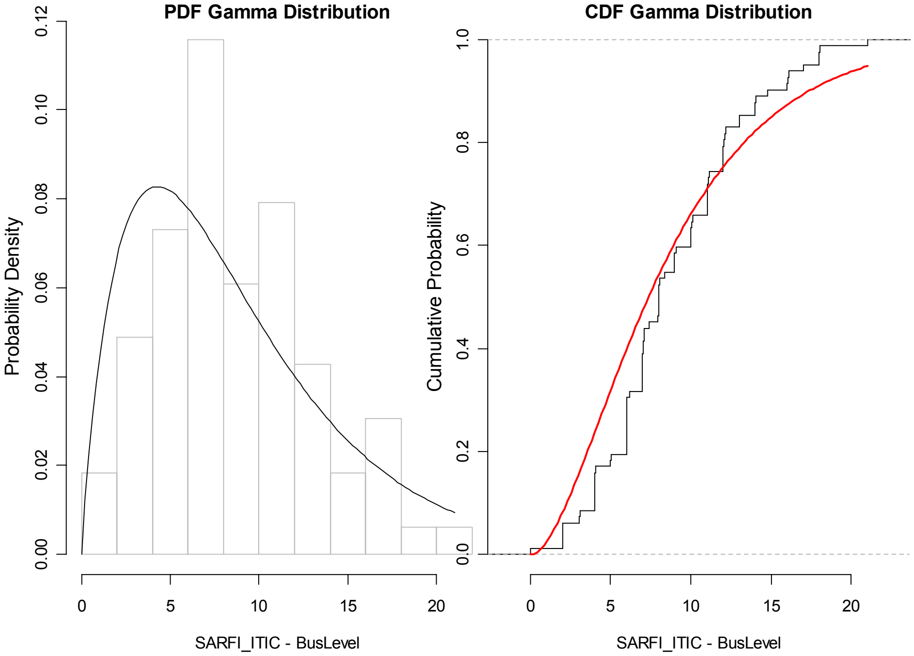


Figure 2-9
SARFI_ITIC Gamma Distribution, Bus Level, Including Storms for Utility A

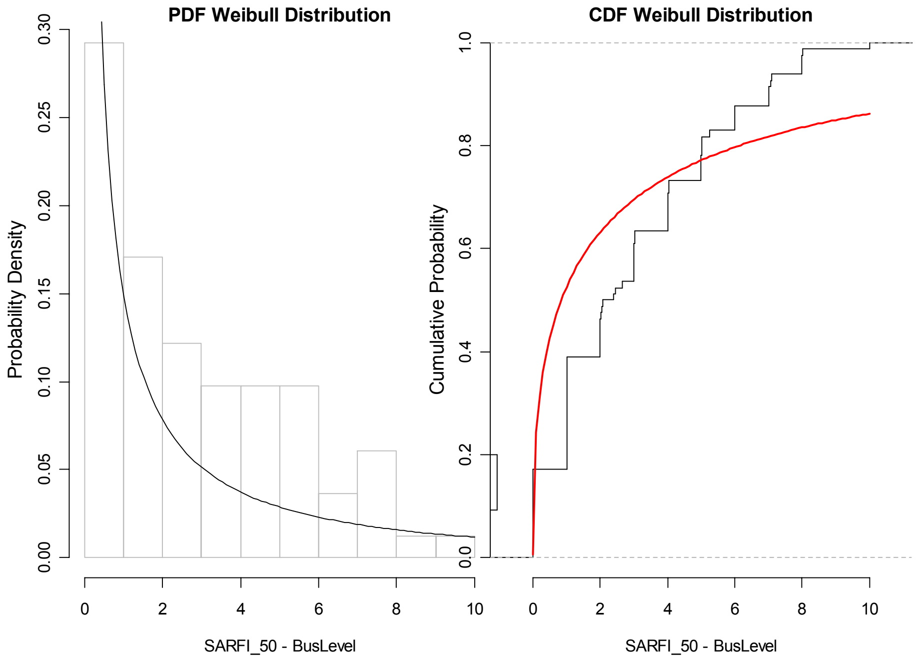


Figure 2-10
SARFI_50 Weibull Distribution, Bus Level, Including Storms for Utility A

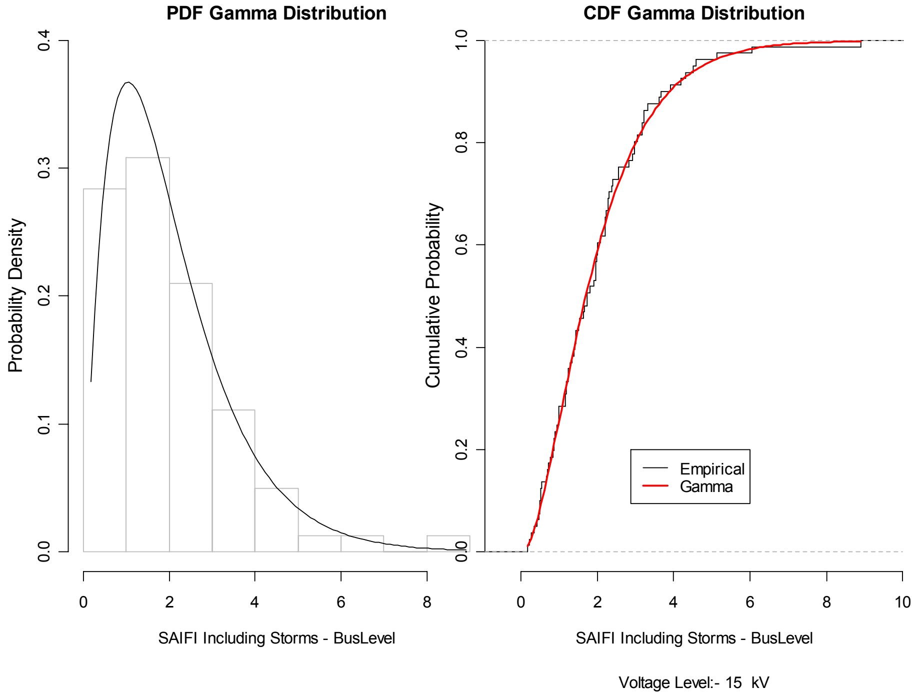


Figure 2-11
SAIFI Gamma Distribution, Bus Level, Including Storms for Utility B

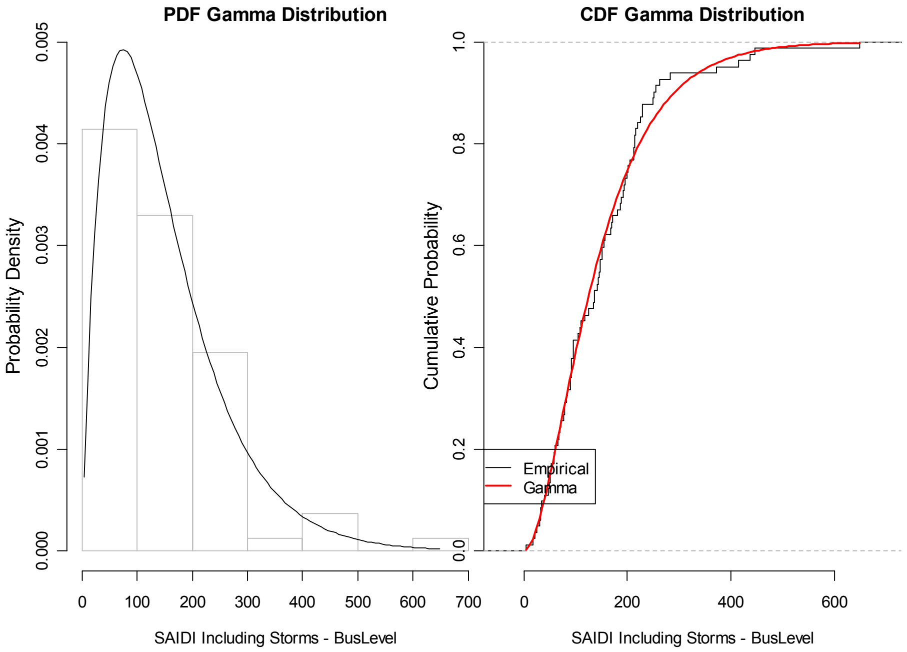


Figure 2-12
SAIDI Gamma Distribution, Bus Level, Including Storms for Utility B

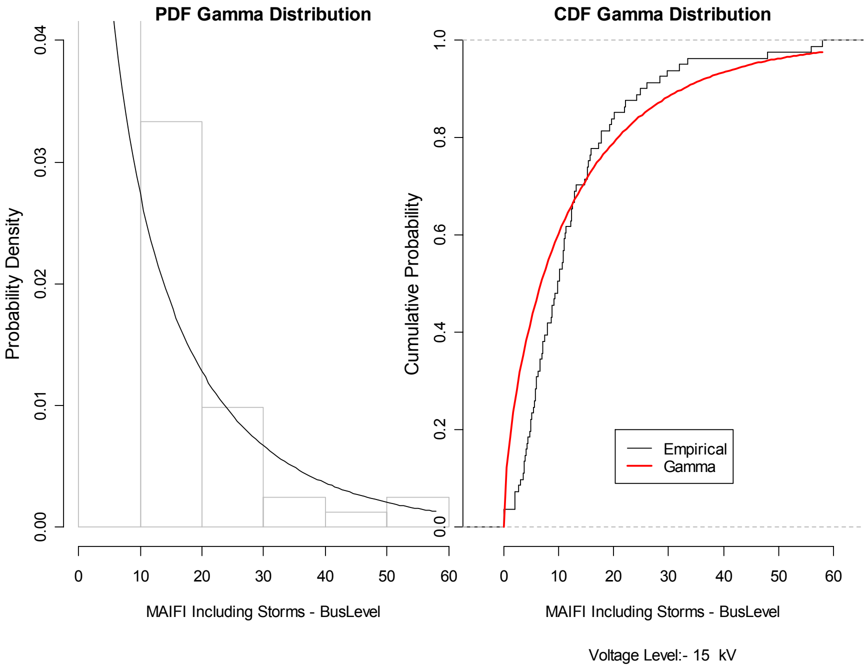


Figure 2-13
MAIFI Gamma Distribution, Bus Level, Including Storms for Utility B

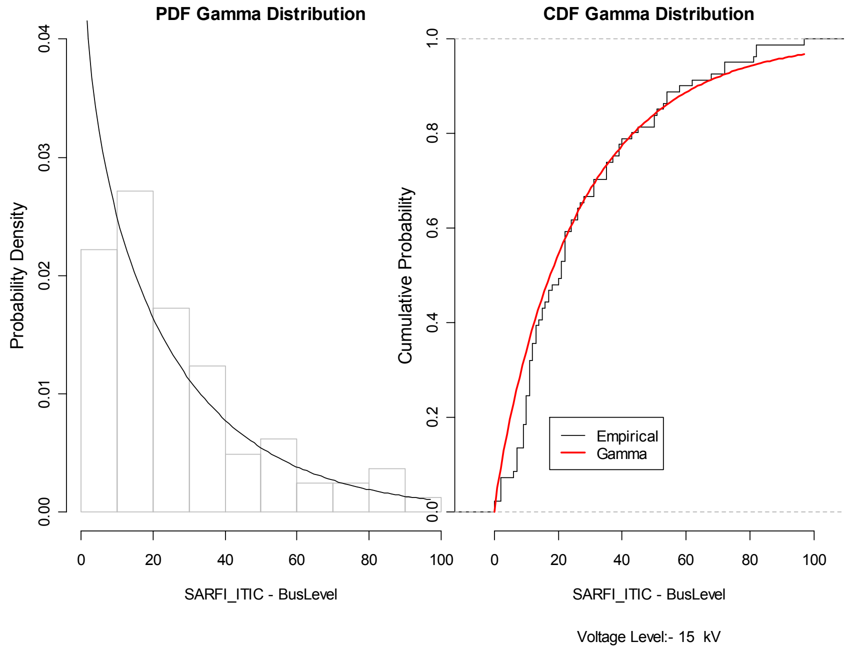


Figure 2-14
SARFI_ITIC Gamma Distribution, Bus Level, Including Storms for Utility B

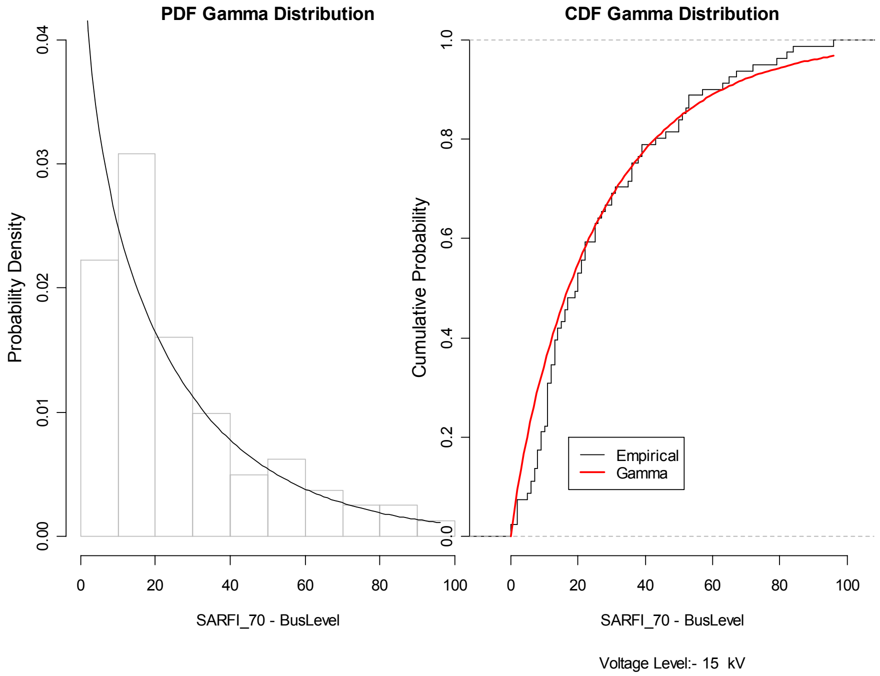


Figure 2-15
SARFI_70 Gamma Distribution, Bus Level, Including Storms for Utility B

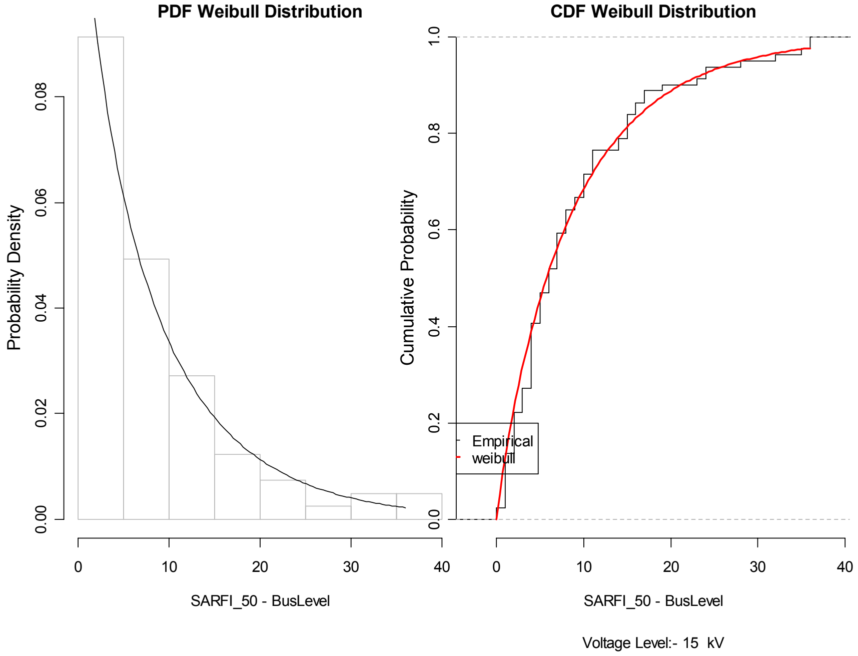


Figure 2-16
SARFI_50 Weibull Distribution, Bus Level, Including Storms for Utility B

3

VARIABLES AFFECTING DISTRIBUTION RELIABILITY INDICES

Even with the significant work and research that has been conducted to date regarding reliability indices, there are still a lot of unknowns and unanswered questions. One thing for certain is that there are numerous variables that can greatly impact the accuracy, uniformity, and consistency of reliability indices (refer to Figure 3-1).

Definition & Data Classification

- Major Events
- Interruption
- Planned/Unplanned
- Distribution/Transmission

Service Territory

- Geography
- Weather Pattern
- Vegetation Pattern
- Vehicle Access Pattern
- Animal Activity

Data Collection Process

- Outage Notification
- Outage Reporting
- Step Restoration Process
- Customer to Network Connectivity

System Design

- Urban/Rural/Downtown
- Load Characteristics
- Underground/Overhead
- Voltage Level
- Protection Scheme

Figure 3-1
Variables Impacting Distribution Reliability Indices

This chapter provides a qualitative overview of some of the important variables that can impact the accuracy, uniformity, and consistency of reliability indices.

Outage Management Systems

One question being raised is "How accurate is the data?" With any type of information and data collection, the analysis is only as good as the accuracy of the original data. With regards to reliability information, much of the previous data and information were based on estimates of customers impacted and estimated restoration times. This restoration time estimated the time from when a customer originally called in to notify the utility of an outage, to when a service truck was dispatched, to the final restoration of service.[31]

However, today, many utilities are deploying sophisticated systems that more accurately track various indicators such that the resulting reliability numbers are based on more accurate information and data. Utilities that have very sophisticated management information systems know exactly how many customers are connected to the power circuit, exactly where they are connected, and how they relate to one another. Progressive utilities that have automated and integrated systems often report the most accurate reliability indices. Such systems include outage management system (OMS), geographical information system (GIS), and automated mapping and facility management systems (AM/FM).[32]

Automated systems inherently provide for more accurate counting of customers served and customers interrupted. In the early years of automation, this improved accuracy often results in increased reliability indices that indicate an apparent degradation in service that is really only a data collection anomaly. The net effect is that the utilities that have invested in improved and automated systems are often penalized because of this perceived deterioration in service based on the increased indice values. Surveys conducted indicate that utilities may experience as much as a 75% increase in their reliability indices following automation.

Storm Normalization and Defining Major Events - Impact on Indices

Another question posed by policy makers and utilities is "Should we exclude storm events from the overall reliability indices?" In the U.S., most state public utilities commissions allow exclusion of major events. It is extremely difficult to use the indices as a valid utility reliability benchmark and track performance improvement without excluding events that are clearly out of control of the utilities responsibility. However, there are two primary rationales to be considered in not allowing exclusion of major events: 1) there is no uniform or consistent set of definitions for what a major event constitutes, therefore, all data should be maintained and kept "pure," and 2) not allowing exclusion of major events makes it simpler for the reliability-reporting scheme and removes any subjective bias, which may be inherent in defining major events.[33]

---

[31] David Schepers, "Reliability Surveying – Segmenting for Comparability," 2002

[32] Cheryl A. Warren Senior Member, IEEE, "Overview of 1366-2001 the Full Use Guide on Electric Power Distribution Reliability Indices," 2001

[33] David Schepers, "Reliability Surveying – Segmenting for Comparability," 2002

Some classify storm or major event based on a certain percentage of customers interrupted, while others use some form of weather-based classification. If storms are not excluded, reliability numbers increase as shown in Figure 3-2. The repair time (CAIDI) and the average total interruption time (SAIDI) increase the most if storm data is included. SAIFI is only moderately impacted.[34]

The major drawback in reporting reliability numbers without excluding major events is the difficulty in comparing results from year to year and utilizing the reliability indices as a benchmark for the utilities performance or lack thereof. The practice of exempting major events will result in data that indicates how the electrical system is maintained and managed, rather than one that shows how bad the winter weather was. Traditional definition of "Major Events" were based on a subjective criteria of adverse weather events such as:

> A major storm is a period of adverse weather during which service interruptions affect either 10 percent of customers in an operating area out of service for at least 2 hours, or more than 2 percent of customers out of service for 24 hours.

Various methods are used to classify storms. The two common categories are based on:

1. *Statistic*s—A common definition is 10% of customers affected within an operating area.
2. *Weather*—Common definitions are "interruptions caused by storms named by the national weather service" and "interruptions caused during storms that lead to a declaration of a state of emergency."

Some utilities exclude other non-weather related interruptions including those scheduled or those occurring from other parts of the utility system (normally substation or transmission caused interruptions). Both are done for the same reasons as storm exclusions: neither scheduled interruptions nor transmission-caused interruptions reflect the normal operating performance of the distribution system.[35]

Because of the inherent subjectivity in defining "Major Events," particularly as they relate to weather-related conditions, the IEEE has established a taskforce to further address this issue. The IEEE Working Group on System Design is engaged in proposing a statistically based definition for "Major Event" classification as a revision to the existing IEEE Standard 1366 "Guide for Electric Power Distribution Reliability Indices" published in 2001. Previous work conducted within the IEEE began laying a solid foundation for consistency in definitions for distribution reliability indices and factors that effect their calculation and published the IEEE Standard 1366, "Trial-Use Guide for Electric Power Distribution Reliability Indices" in 1998.

Presently under review and balloting for adoption, IEEE P1366 "DRAFT Full-Use Guide for Electric Power Distribution Reliability Indices" incorporates a methodology, 2.5 Beta Method,

---

[34] Tom Short, EPRI PEAC, "Reliability Indices," *White Paper for T&D World*, 2002

[35] Tom Short, EPRI PEAC, "Reliability Indices," *White Paper for T&D World*, 2002

for determination of major event days.[36] It is anticipated that once the document is balloted, the working group will continue to investigate the major event definition by reviewing catastrophic events and days with zero events to determine if a modification to the definition is warranted.

One argument against the statistical approach is that major substation or transmission outages should be considered as "major events" and be excluded from indices. From the customer's point of view, major event or no major event, an interruption is still a power outage impacting customers and manifested in loss of production, scrapped material, lost revenues, or a missed sporting event on Pay-Per-View television. And it is for this reason that some regulators hesitate to allow exclusions at all.[37]

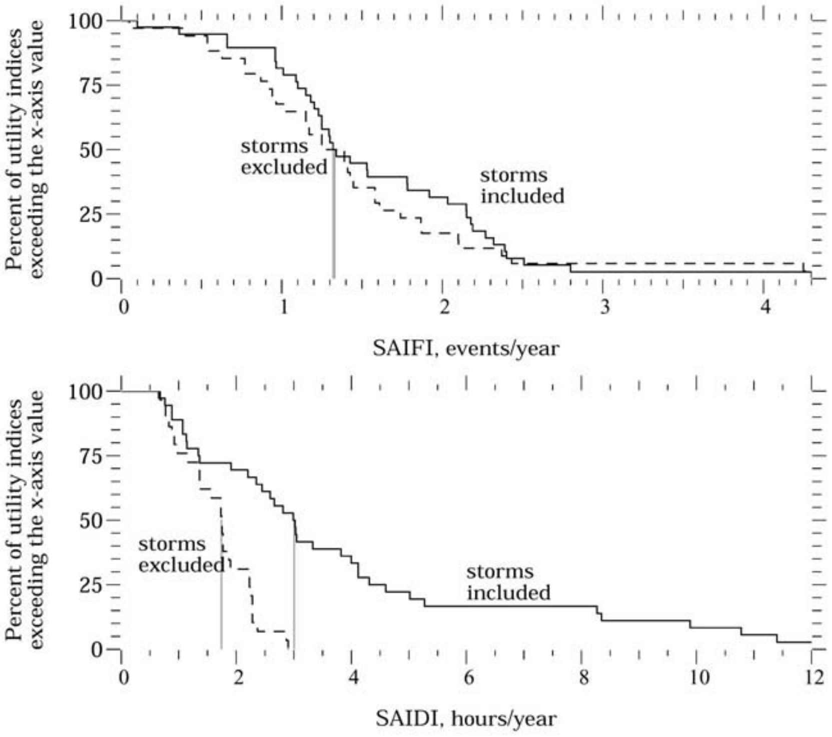


Figure 3-2
Impact of Major Events on Indices [EEI Survey 1999]

Distribution Automation

Another question that can significantly impact reliability indices is "To what level or sophistication has the utility implemented distribution automation and other supervisory control and data acquisition (SCADA)?" Larger utilities that have enhanced data and, therefore, higher numbers might be penalized. Conversely, smaller utilities that may not be able to financially

---

[36] IEEE P1366 – Draft 12, "Draft Full Use Guide for Electric Power Distribution Reliability Indices," February 2003

[37] Tom Short, EPRI PEAC, "Reliability Indices," *White Paper for T&D World*, 2002

justify such elaborate systems may object to being compared to others who have implemented automated distribution systems. In either case, to be able to appropriately compare indices, it is important to note which utilities have implemented such systems and the extent of implementation.[38]

Geography, Circuit Exposure, and Supply Configuration

"Physically, what type of region does a utility serve and with what type of supply configuration?" Answers to both of these questions can dramatically affect the reliability numbers. Service territories generally include urban, suburban, or rural, or more likely, some combination of all of these. Distribution systems designed for rural areas are generally made up of smaller substations with long radial circuits extending for many miles with little redundancy and few circuit ties. In contrast, systems in dense urban areas often deploy grid like systems and incorporate larger substations often with redundant facilities, shorter circuit lengths, and multiple feeds.

Circuit distance alone is a substantial reliability issue; longer circuits have inherently more interruptions. This is due to the fact that more circuit length equates to more exposure and more points of potential failure. Most of the change is in SAIFI; the repair time (CAIDI) is less dependent on circuit lengths. Figure 3-3 shows the effect on SAIFI at one utility in the southwest.[39]

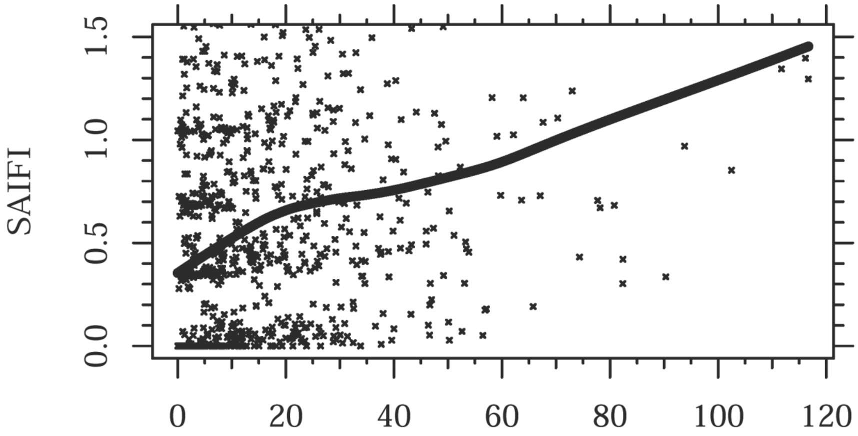


Figure 3-3
Effect of Circuit Length on Indices

---

[38] David Schepers, "Reliability Surveying – Segmenting for Comparability," 2002

[39] Tom Short, EPRI PEAC, "Reliability Indices," White Paper for T&D World, 2002

Another validation that urban areas have inherently better indices is the 2000 Indianapolis Power and Light survey that only included performance of utilities in large cities. As expected, the SAID and SAIFI results for urban areas resulted in a 28% improvement than utility surveys that included both urban and rural service territories.

The distribution supply and configuration also greatly impacts reliability, which is also related to geography. Radial circuits provide the poorest service; grid networks provide exceptionally reliable service. Table 3-1 provides estimated reliability for several common distribution supply configurations developed by New York City's Consolidated Edison. Note that the repair time (CAIDI) increases for the more urban configurations. Being underground and dealing with traffic increases the time for repairs.

Table 3-1
Comparison of Reliability of Different Distribution Configurations [Settembrini et. al., 1991]

	SAIFI Interruptions/Year	CAIDI Minutes/Interruption	MAIFI Momentary Interruptions/Year
Simple radial	0.3 to 1.3	90	5 to 10
Primary auto-loop	0.4 to 0.7	65	10 to 15
Underground residential	0.4 to 0.7	60	4 to 8
Primary selective	0.1 to 0.5	180	4 to 8
Secondary selective	0.1 to 0.5	180	2. to 4
Spot network	0.02 to 0.1	180	0 to 1
Grid network	0.005 to 0.02	135	0

Based on Table 3-1, simple radial systems have the highest SAIFI while grid network have the smallest SAIFI values.

Some countries, such as Australia, have undertaken further refinement of the averages by segmenting the network into different categories such as Commercial Business District (CBD), Urban, Rural, and Remote. This categorization is primarily an effort to distinguish the inherent differences in network topologies and an attempt to compare "apples to apples" when evaluating the performance of different distribution businesses (refer to Table 3-2).[40]

---

[40] EPRI, "Future Power Quality and Reliability Requirement for South Australia," 2001

Table 3-2
Historical Trend of System Average Interruption Duration Index (SAIDI) for the Four Different Categories of Networks Maintained by Australian Utilities

Network Category	Year				
	1995/96	1996/97	1997/98	1998/99	1999/00
CBD	-	-	23.5	5.1	12.4
Urban	-	83.0	85.8	93.1	99.5
Rural	-	164.0	143.4	130.6	141.9
Remote	-	300.0	244.2	230.6	237.0
Total SAIDI Average	116.0	118.4	112.6	110.4	118.6

Customer Perception of Service Quality

A primary question still remains – "How accurately do these system-wide average indices reflect the ***customer*** perception of reliability of electric power?" The answer to this question lies in the statistical distribution of the indices. At this point without getting into the details regarding statistical analysis, distribution of reliability indices, in general do not follow a normal distribution; and therefore, "average" is not a good representation of the population. In most cases, the distribution is skewed by the performance of a number of feeders and, overall, the median (what 50% of the customers will see as interruptions) is not equal to nor a close approximation of the average performance of the entire network.

4

VARIABLES AFFECTING POWER QUALITY AND MOMENTARY INTERRUPTIONS

The next step is to find out what site parameters and distribution reliability indices affect sags the most. This is an issue that's worth investigating on its own right and adds insight on ways to improve power quality to end-use customers.

Topics covered in this chapter include the following:

- Evaluate relationships between site characteristics on voltage sags and reliability (including and excluding major storms). How does the correlations change with respect to different feeder characteristics namely total exposure (three-phase as well as single-phase), backbone length, number of feeders, transformer size, load density, percentage underground versus overhead, etc.
- Ascertain if correlation exists between PQ indices and reliability indices. Correlations between SARFI_X (ITIC, 90, 80, 70, 50, 10) with respect to reliability measures such as SAIFI, SAIDI, and outages counts were investigated. Correlation between MAIFI with respect to reliability measures such as SAIFI, SAIDI, and outage counts were also investigated.
- Evaluate if correlations exist between PQ definition of momentary interruptions (SARFI_{10}) and reliability definition (MAIFI and MAIFI_E).
- Evaluate the effect of including and excluding major storms from the reliability data and how these effect the overall prediction.

Parameters Available in the Datasets

A number of site characteristics are available for the monitoring locations. Descriptions of these follow:

Monitoring location – It describes the location of the monitor as one of the following:

- Bus
- Feeder

Both the "Substation" and "Feeder" types are on the distribution system for the two utilities

***Load density* –** The load density is described as one of the following:

- Urban
- Suburban
- Rural

The load density is classified according to a 0-30, 31-70 and 71-100 percentile distribution.

***Feeder construction* –** Utilities were asked to describe the feeder construction as one of the following:

- Overhead
- Underground
- Mixed

Where utilities provided circuit lengths, a circuit was classified as overhead if it was more than 2/3 overhead, underground if it was more than 2/3 underground, or mixed.

***Nominal distribution base voltage* –** This is the nominal line-to-line voltage in kilovolts.

***Feeder length* –** The length is the total circuit miles on just the feeder monitored.

***Substation transformer size* –** This is the base MVA rating (the open-air or OA rating) of the transformer.

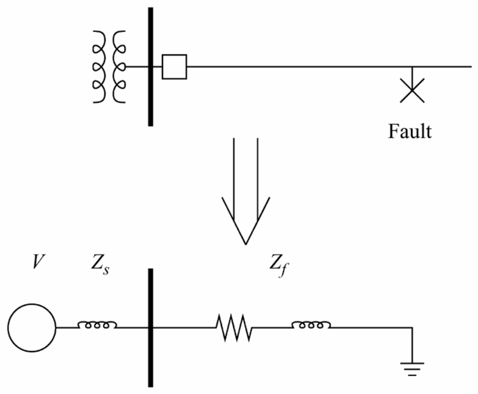
***Number of feeders* –** This is the number of feeders off the substation bus. If a utility uses parallel busses (closed bus ties), we divided the number of feeders by the number of busses to obtain a “feeders per bus” number. Therefore, you may notice a non-integer number of feeders in some of the graphs.

***Monitor days* –** This is the number of days that the monitor recorded. We sometimes used this to weight distributions or regression models.

The SARFI statistics at each site were analyzed separately. The site characteristics data did have many missing values, which we had to contend with.

Electrical Parameters That are Important for Sags

The voltage during the fault at the substation bus is given by the voltage-divider expression in Figure 4-1 based on the source impedance (Z_s), the feeder line impedance (Z_f), and the prefault voltage (V). This voltage can be at the substation bus or another location on the power system.

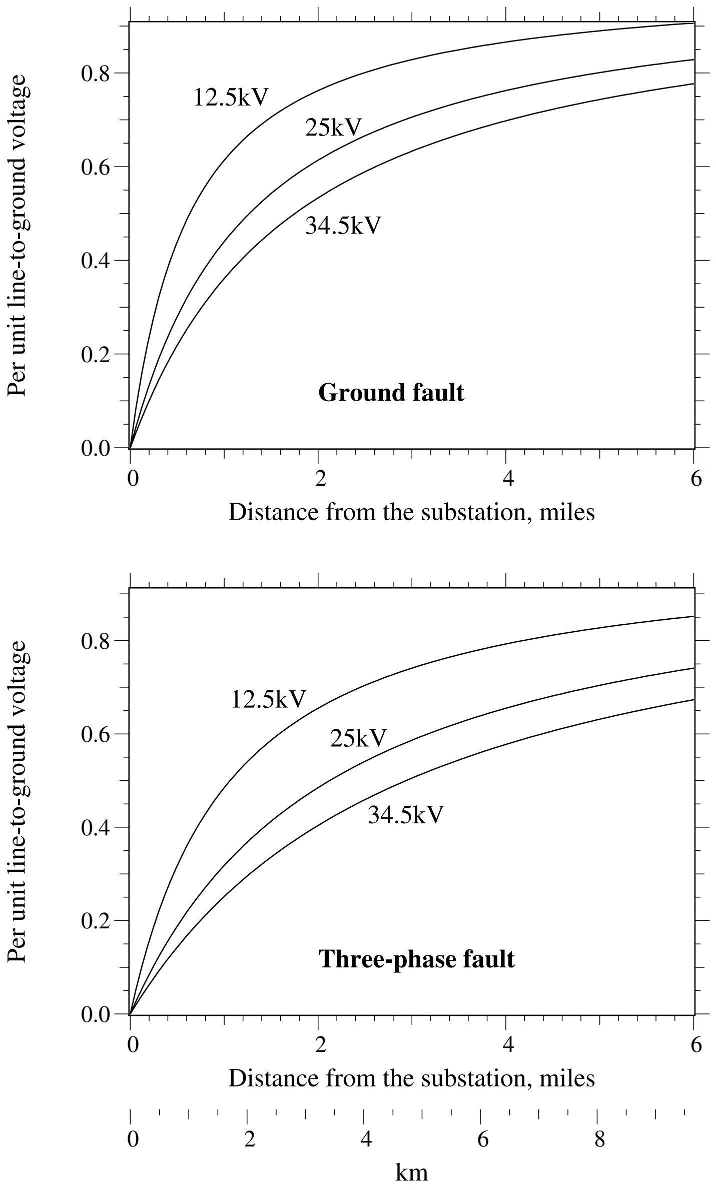


$$V_{bus} = \frac{Z_f}{Z_f + Z_s}$$

Figure 4-1
Voltage Divider Equation Giving the Voltage at the Bus for a Fault Downstream

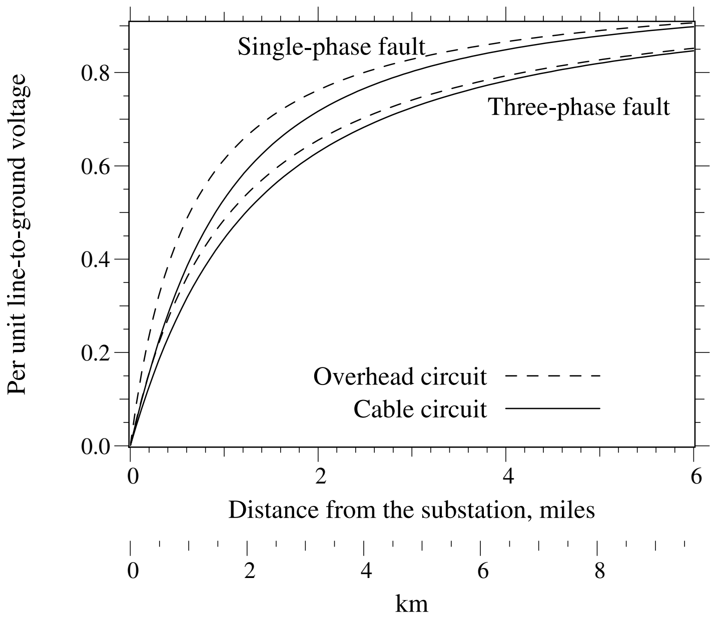
The voltage will sag deeper for faults electrically closer to the bus (smaller Z_f). Also, as the available fault current decreases (larger Z_s), the sag becomes deeper. The source impedance includes the transformer impedance plus the subtransmission source impedance (often, subtransmission impedance is small enough to be ignored). The impedances used in the equation depend on the type of fault it is. For a three-phase fault (giving the most severe voltage sag), use the positive-sequence impedance ($Z_f = Z_{f1}$). For a line-to-ground fault (the least severe voltage sag), use the loop impedance, which is $Z_f = (2Z_{f1} + Z_{f0})/3$. A good approximation is 1 ohm for the substation transformer (which represents a 7- to 8-kA bus fault current) and 1 ohm per mile (0.6 Ω/km) of overhead line for ground faults. For accuracy, use complex division because the impedances are complex, but for back-of-the-envelope, first-approximation calculations, use the impedance magnitude.

Note that this can be used for any type of fault as long as the appropriate fault values are used in the equation. If the angles are ignored, the equation is an approximation (which is usually acceptable). Figure 4-2 shows a profile of the substation bus voltage for faults at the given distance along the line for 12.47, 24.94, and 34.5 kV. The higher-voltage systems have more severe voltage sags for faults at a given distance. The graph also shows that three-phase faults cause more severe sags. Figure 4-3 compares sags on underground and overhead systems.



Phase characteristics: 500 kcmil, all-aluminum, GMD = 4.69 ft (1.43 m)
Neutral characteristics: 3/0 all-aluminum, 4-ft (1.22-m) line-to-neutral spacing
Z_1=0.207 + j0.628 Ω/mile (0.1286 + j0.3901 Ω/km)
Z_0=0.720 + j1.849 Ω/mile (0.4475 + j1.1489 Ω/km)
Z_s=0.378 + j1.035 Ω/mile (0.2350 + j0.6430 Ω/km)

Figure 4-2
Substation Voltage Profile for Faults at the Given Distance (Single-Phase and Three-Phase Faults Are Shown for Each Voltage)



500-kcmil aluminum conductor, 220-mil XLPE insulation,
1/3 neutrals, flat spacing, 7.5 in between cables
Z_1=0.3543 + j0.3596 Ω/mile (0.2201 + j0.2234 Ω/km)
Z_0=0.8728 + j0.2344 Ω/mile (0.5423 + j0.1456 Ω/km)
Z_s=0.5271 + j0.3178 Ω/mile (0.3275 + j0.1975 Ω/km)

Figure 4-3
Comparison of Substation Voltage for Faults on Overhead Circuits and Cable Circuits at the Given Distance (Single-Phase and Three-Phase Faults Are Shown)

The effect of feeder faults on voltage sags at the substation bus can be estimated with the following equation:

$$S(V_{sag}) = n_f \lambda \frac{V_{sag}}{1 - V_{sag}} \left(\frac{Z_s}{Z_f} \right)$$ **Eq. 4-1**

where,

S = annual number of sags per year where the voltage sags below V_{sag}

V_{sag} = per-unit voltage sag level of interest (in the range of 0 to 1, such as 0.7)

n_f = number of feeders off the bus

λ = feeder mains fault rate per mile (or other unit of distance) per phase, including faults on laterals and including both temporary and permanent faults

Z_f = feeder impedance in ohms per mile (or other unit of distance); usually use $Z_f = (2Z_1+Z_0)/3$ for ground faults

Z_s = source impedance in ohms

The distribution of voltage sags based on this equation is shown in Figure 4-4 for some common parameters. Several points are noted from this analysis on voltage sags:

- *Exposure* – For 15-kV circuits, we can ignore exposure beyond the first 2 or 3 miles (4 or 5 kilometers) for sags to the bus voltage. The first mile or two is most important as far as circuit improvement, maintenance, or application of current-limiting fuses.
- *System voltage* – Sags are more severe on higher voltage distribution systems (especially at 34.5 kV). A fault 4 miles from the substation sags the voltage much more on a 25-kV system than on a 12-kV system because the substation transformer is a higher impedance relative to the line impedance at higher system voltages. For 24.94 kV, exposure as far as five miles from the station is significant.
- *Single versus three-phase faults* – Three-phase faults cause more severe sags than single-line-to-ground faults. Three-phase faults farther away can pull the voltage down.
- *Underground versus overhead* – All-underground circuits have more exposure to sags because cables have lower impedance than overhead lines.
- *Number of feeders* – The number of sags on the station bus is directly proportional to the number of feeders off the bus.
- *Transformer impedance* – A lower station transformer impedance (a bigger transformer or lower percent impedance) improves voltage sags.
- *Bus tie* – It does not matter whether a substation bus tie is open or closed. If it is open, a fault affects only half of the feeders. A fault that does occur forces a deeper sag because of a higher effective source impedance. These two effects tend to cancel each other.
- *Voltage regulation* – Raising the nominal voltage improves the voltage seen by customers during a fault. Say that a fault drops the voltage to 0.8 per unit, and the prefault voltage was 1.0 per unit. If the prefault voltage were 1.1 per unit, the voltage during the sag is 0.88 per unit. This is not a big difference, but for equipment sensitive to sags to 0.7 to 0.85 per unit, higher voltages appreciably reduce the number of tripouts.

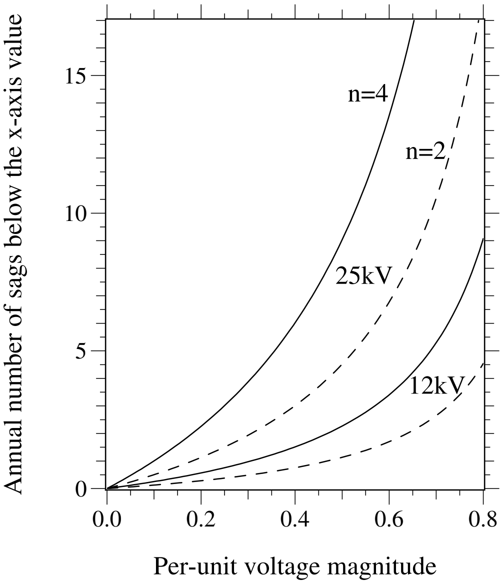


Figure 4-4
Cumulative Distribution of Substation Bus Voltage Sags per Year for the Given (25-MVA, 10% Transformer, 500-kcmil Feeder, n=2 or 4 Feeders off the Bus, λ = 1 Faults/Phase/Mile of Mains/Year, Assumes Line-to-Ground Faults Only)

Customers at the end of a circuit have more severe voltage sags because almost all faults upstream appear as little or no voltage (most actually fit the power quality definition of an interruption, a voltage to below 10%).

This section explains the theory behind some of the parameters. In the next section, we will investigate the impact of some of these variables on actual data. One of the parameters that we will investigate in more detail is:

$$\frac{n_f \cdot kV^2}{MVA_{xfmr}}$$ **Eq. 4-2**

The number of bus sags is directly proportional to n_f, the number of feeders off the bus, and to Z_S, the source impedance (a lower station transformer impedance, a bigger transformer or lower percent impedance, reduces the severity of voltage sags at the station bus). The transformer impedance is $Z_{\%}kV^2/MVA_{xfrmr}$; but because the per-unit impedance of station transformers is roughly constant (7 to 10%), we use kV^2/MVA_{xfrmr}.

Correlation Of SARFI$_x$ and Site Parameters

Feeder Correlation

Feeder monitors are typically triggered on current only and depends which side of the breaker it is. Therefore in order to account for what will be the true sag count that a customer on a feeder will see we have to use bus monitor's sag count especially if we are trying to make correlation of voltage sags with respect to site characteristics and reliability.

Note for Utility A we did not have SARFI$_x$ data from feeder monitors. The SARFI$_x$ were obtained from bus monitors. We took this number and assigned this to individual feeders.

For Utility B even though we had feeder monitors, these were also triggered on currents and does not account for what the true sag count will be (faults on the other feeders will also cause a sag on feeder 1. The customer on feeder 1 will also see the sag but the feeder monitor will not count that sag)

The plots in Figure 4-5 through 4-8 are evidence to support that we will not get any correlation if we take bus monitor sag count and correlate these with individual feeders. Note these plots are evidence to show that no correlation exists when we try to correlate SARFI$_x$ using data recorded by bus monitors with feeder characteristics

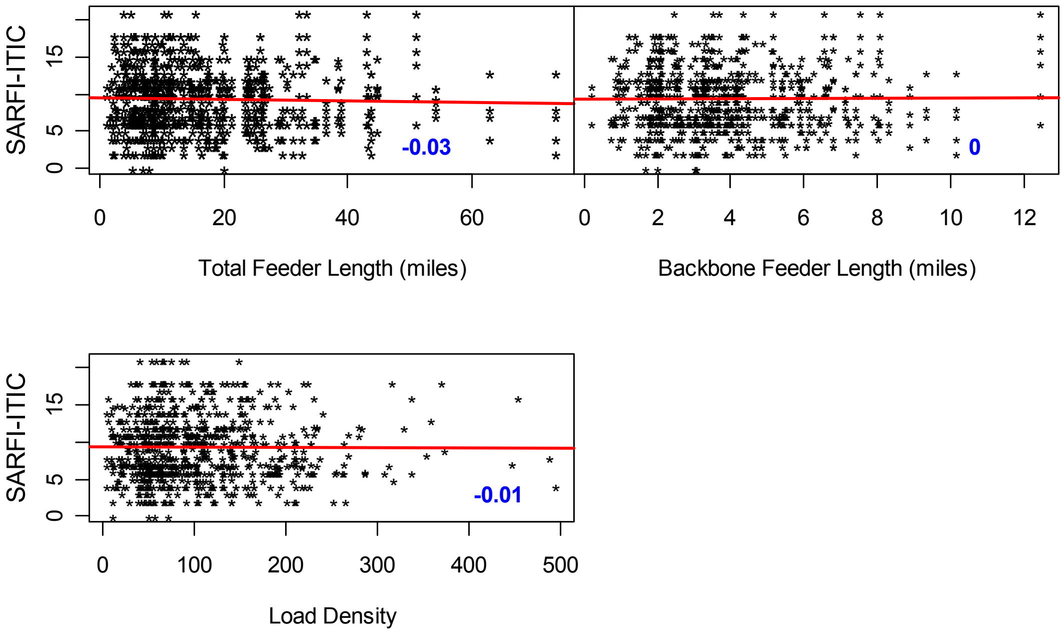


Figure 4-5
Feeder Level Correlations for SARFI_ITIC at Utility A – (Lack of Correlation Between SARFI_ITIC Measured by Bus Monitor and Feeder Characteristics)

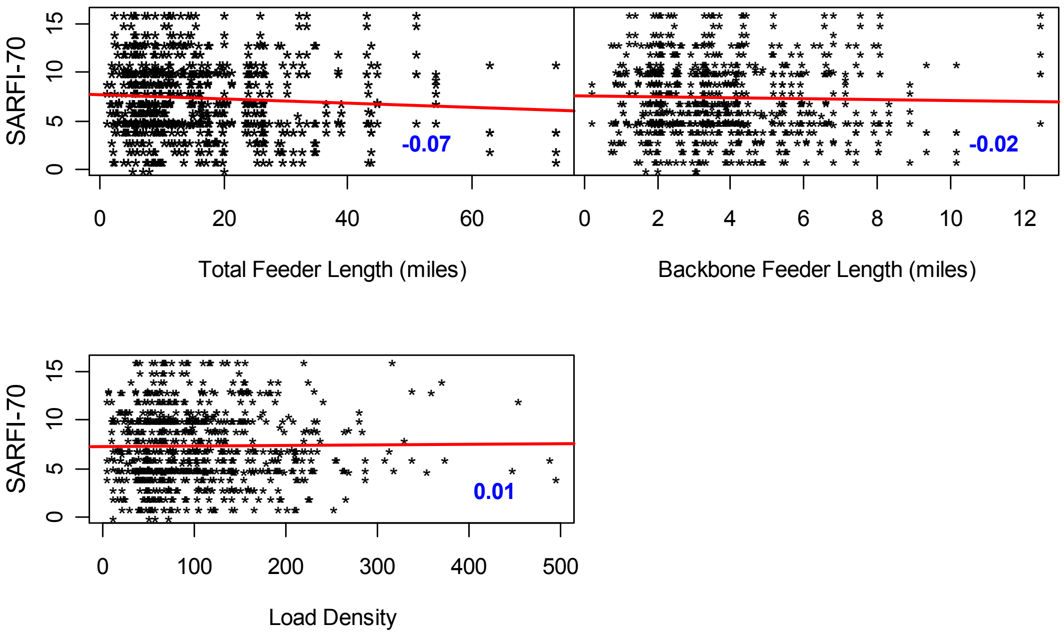


Figure 4-6
Feeder Level Correlations for SARFI_70 at Utility A – (Lack of Correlation Between SARFI_70 Measured by Bus Monitor and Feeder Characteristics)

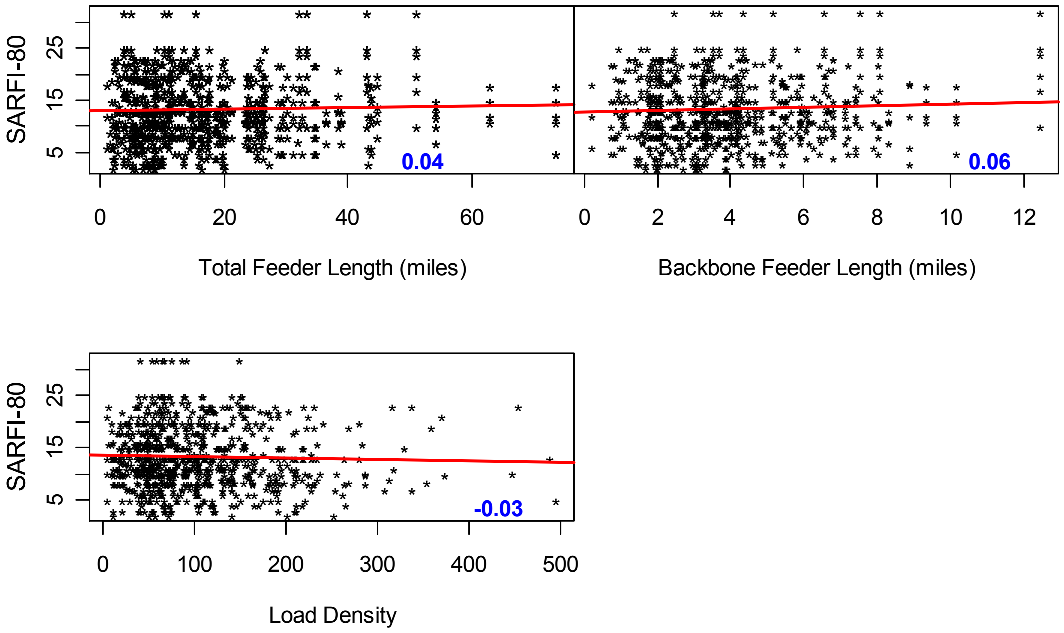


Figure 4-7
Feeder Level Correlations for SARFI_80 at Utility A – (Lack of Correlation Between SARFI_80 Measured by Bus Monitor and Feeder Characteristics)

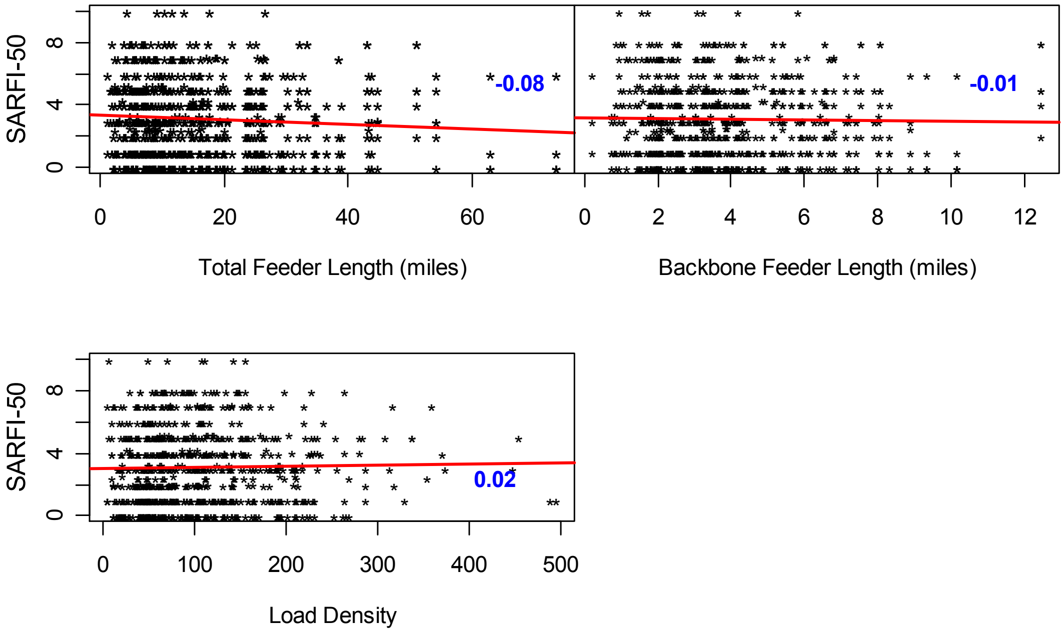


Figure 4-8
Feeder Level Correlations for SARFI_50 at Utility A – (Lack of Correlation Between SARFI_50 Measured by Bus Monitor and Feeder Characteristics)

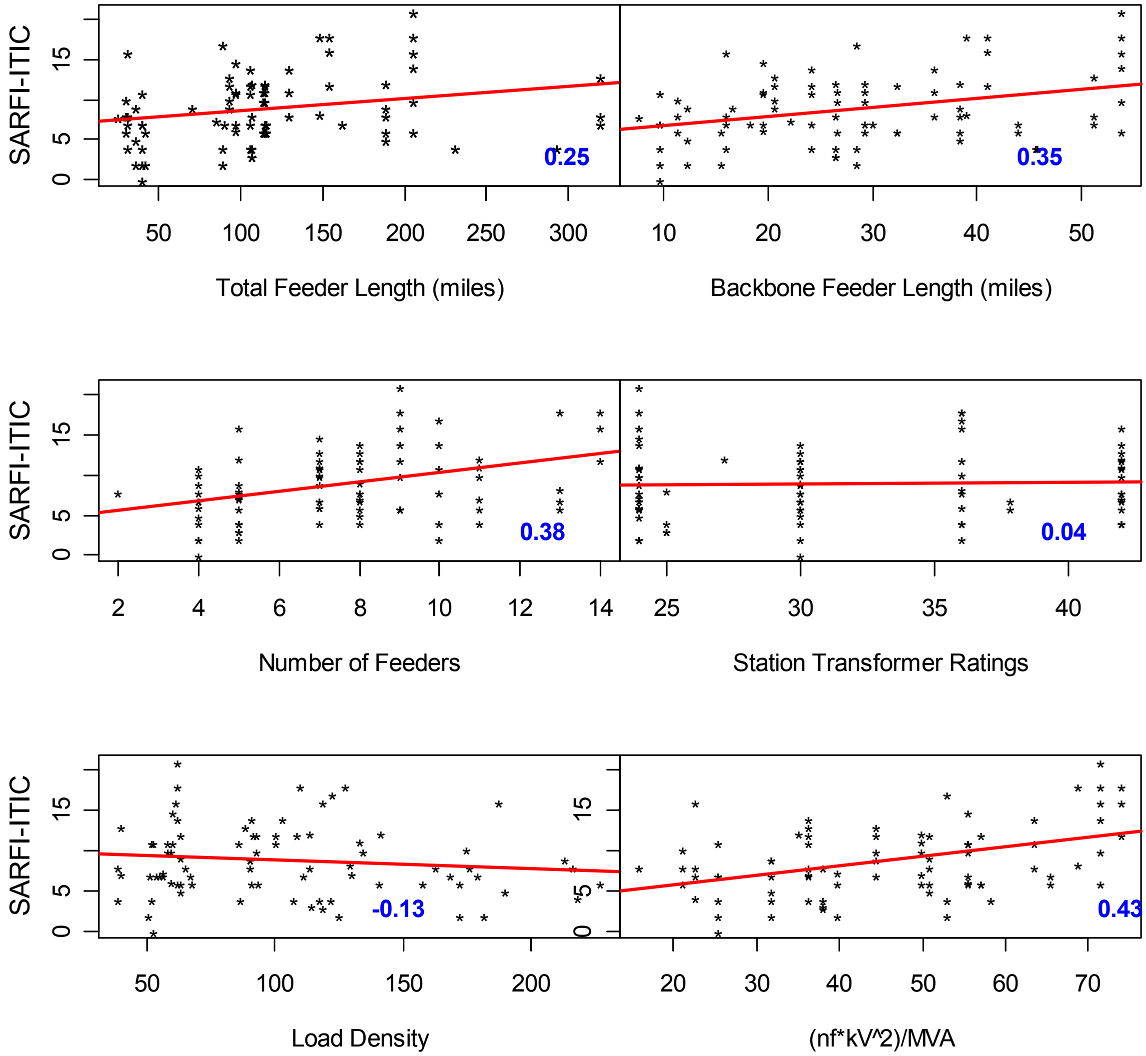
Substation/Bus Level Correlation

Figure 4-9 through Figure 4-16 show the effect of site parameters on $\text{SARFI}_{\text{ITIC}}$, SARFI_{70}, SARFI_{50}, and MAIFI using substation data. . The graph also shows a linear curvefit and a correlation coefficient. Note the high variability of data. Some of the site parameters have some correlation, but again, the data does not show any of the parameters dominating.

The numbers in blue are the coefficient of correlation (r). This shows that exposure and number of feeders have the highest influence on reliability and power quality indices. Correlation coefficients are higher for feeder length, backbone length and number of feeders and the product term $\frac{n_f kV^2}{MVA}$. Correlations for the product term were better for Utility A than for Utility B. For $\text{SARFI}_{\text{ITIC}}$, transformer ratings and load density do not have significant correlation. The same trend exists for SARFI 70, 80 and 90. In general, as the SARFI number decreases, the correlation worsens as well.Also, the circuit lengths in these tables and graphs are circuit miles of all feeders fed from the substation bus, whereas the circuit lengths for the feeder sites were just for the feeder where the monitor was.

Note however that the correlation cannot prove that the change in the value of one variable caused the change in the other variable. A strong correlation can be produced simply by chance, by the effect of other external variables not considered in the calculations, or by the cause-and-effect relationship. Additional analysis in the form of linear and generalized regressions is required.

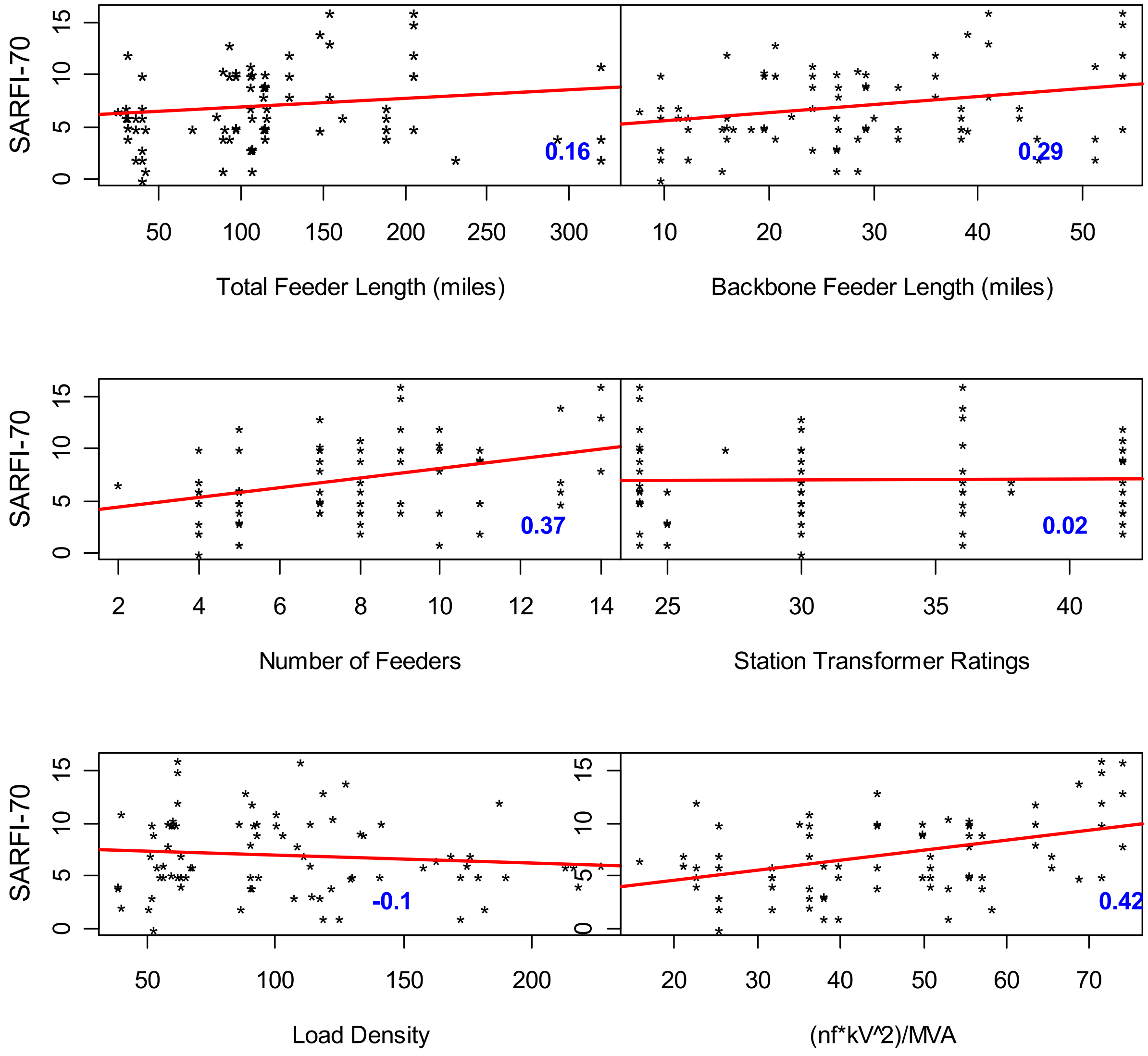
Substation Sites



The correlation coefficients (*r*) are given in the lower right corner of each plot.

Figure 4-9
SARFI$_{\text{ITIC}}$ Site Variations With Various Parameters. Bus Level – Utility A

Substation Sites



The correlation coefficients (*r*) are given in the lower right corner of each plot.

Figure 4-10
SAIRFI$_{70}$ Site Variations With Various Parameters. Bus Level – Utility A

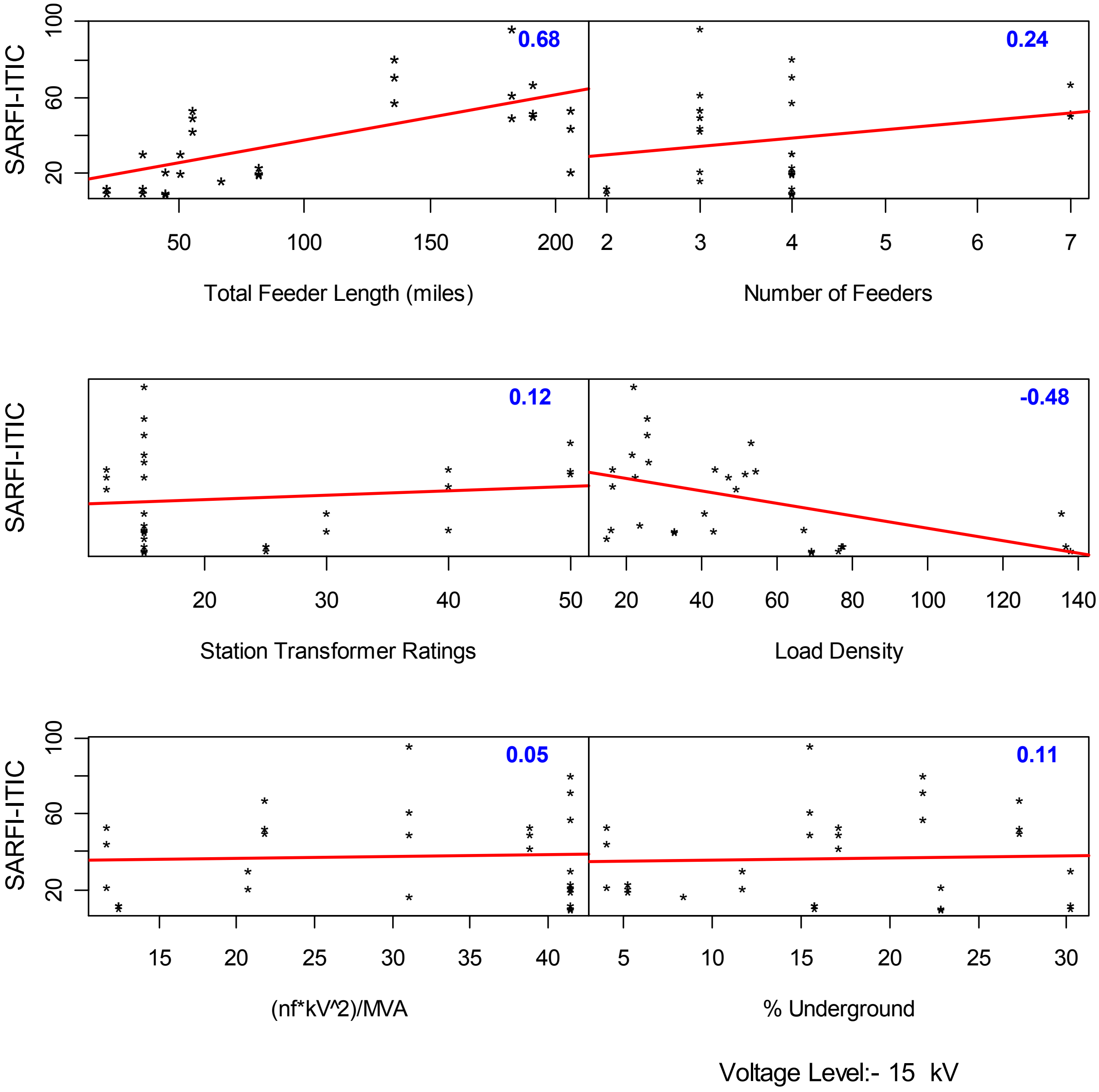


Figure 4-11
SARFI$_{ITIC}$ Site Variations With Various Parameters. Bus Level (15KV) – Utility B

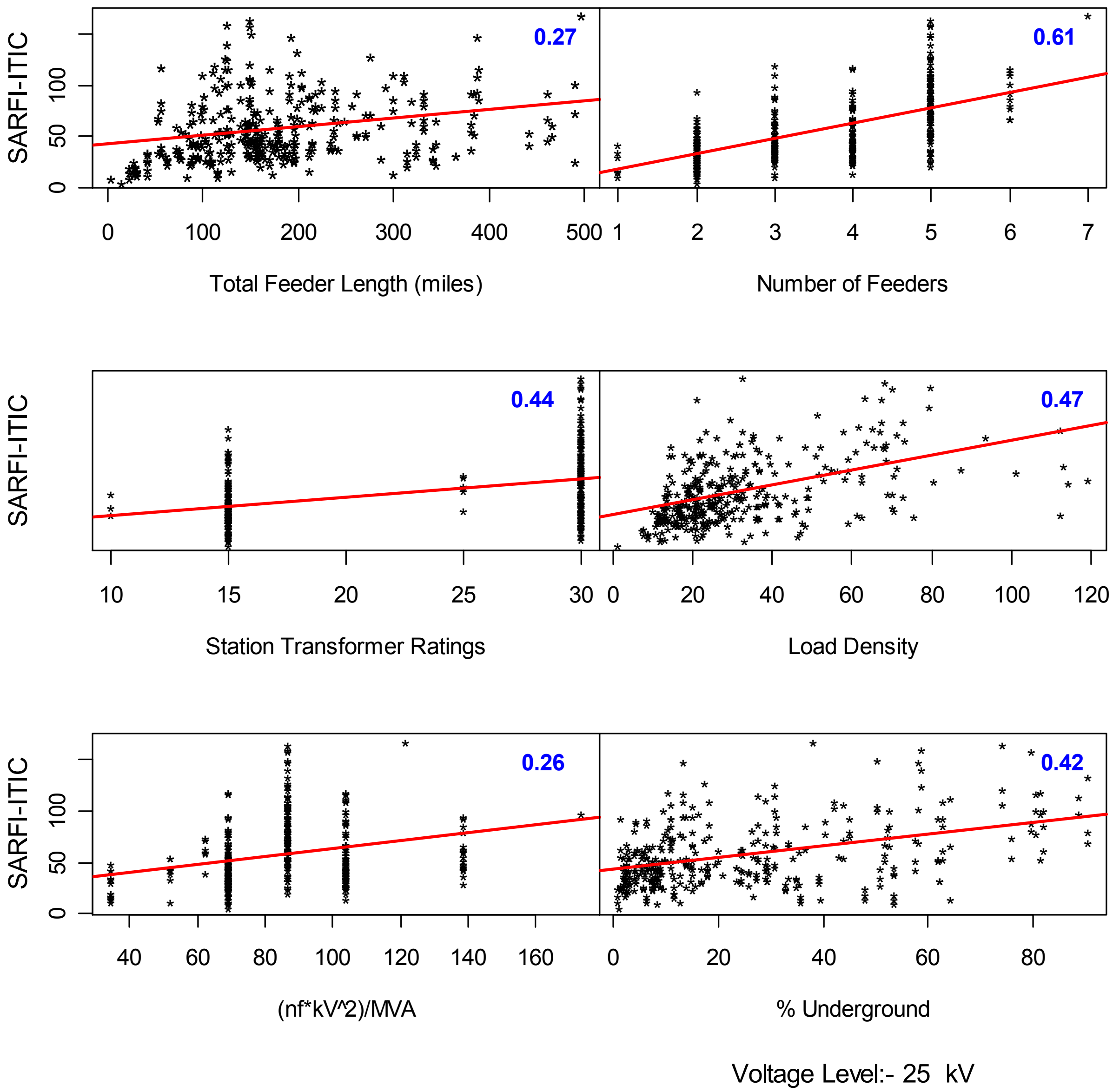


Figure 4-12
SARFI$_{ITIC}$ Site Variations With Various Parameters. Bus Level (25KV) – Utility B

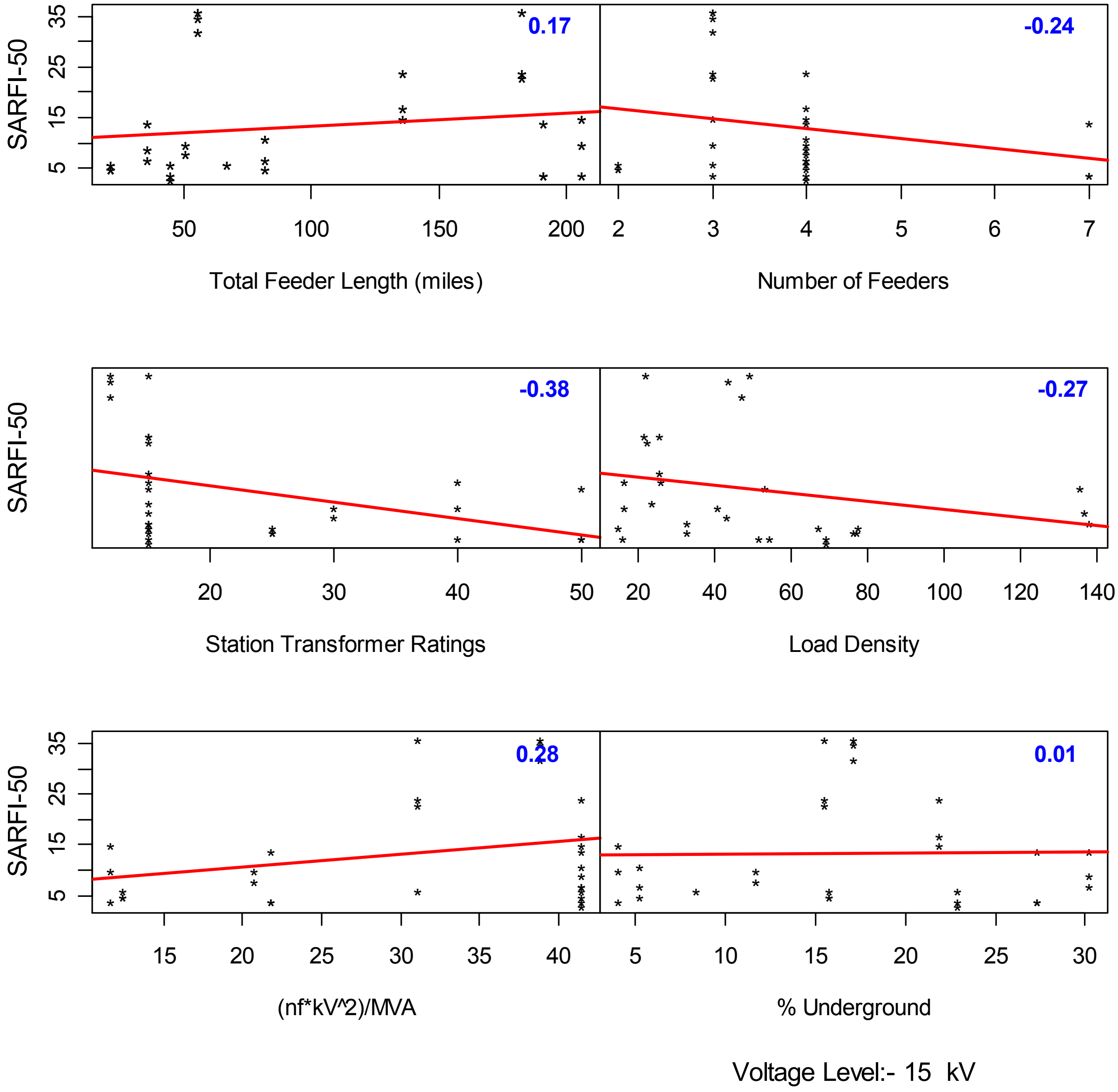


Figure 4-13
SARFI$_{50}$ Site Variations With Various Parameters. Bus Level (15KV) – Utility B

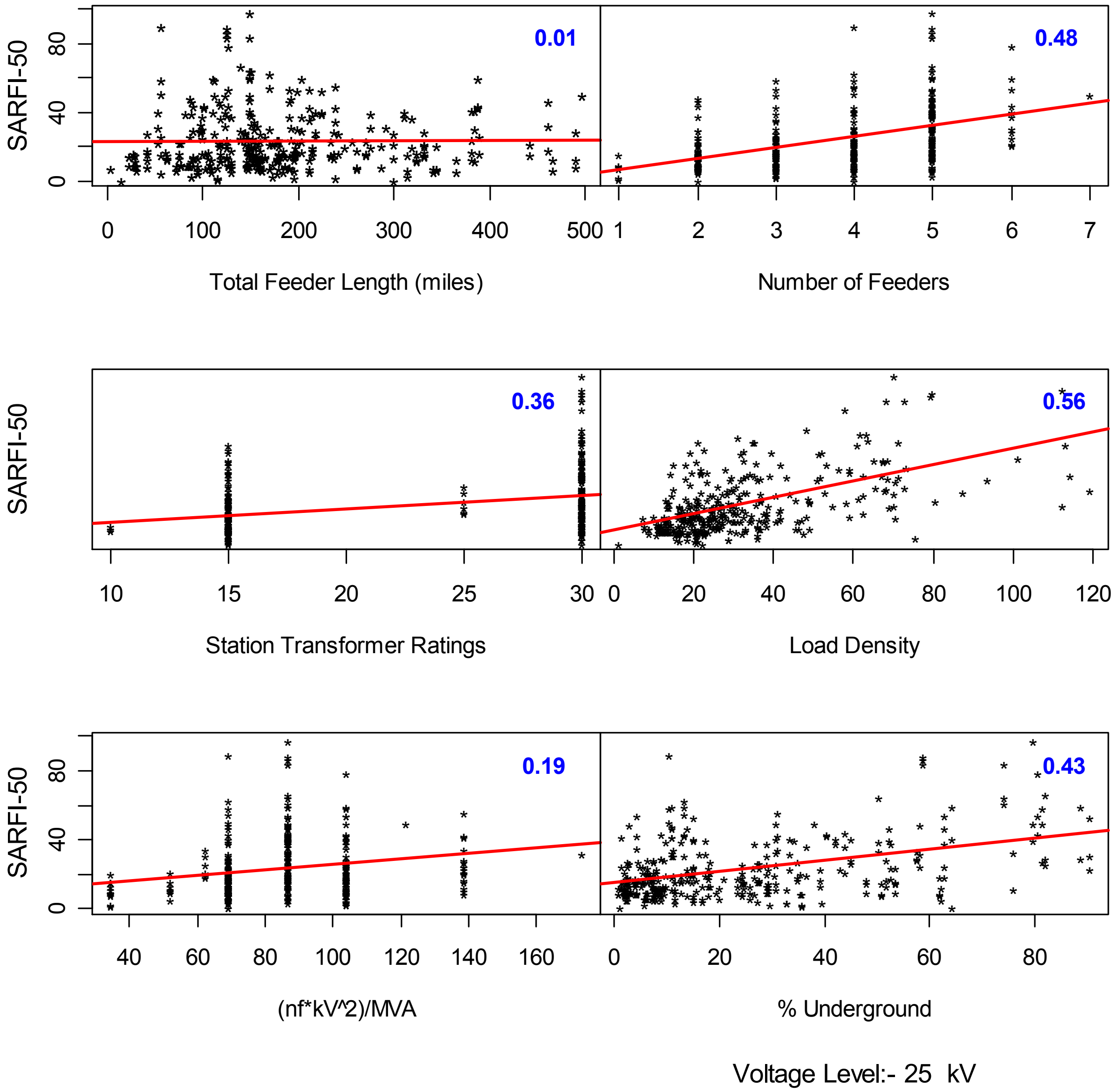


Figure 4-14
SARFI$_{50}$ Site Variations With Various Parameters. Bus Level (25KV) – Utility B

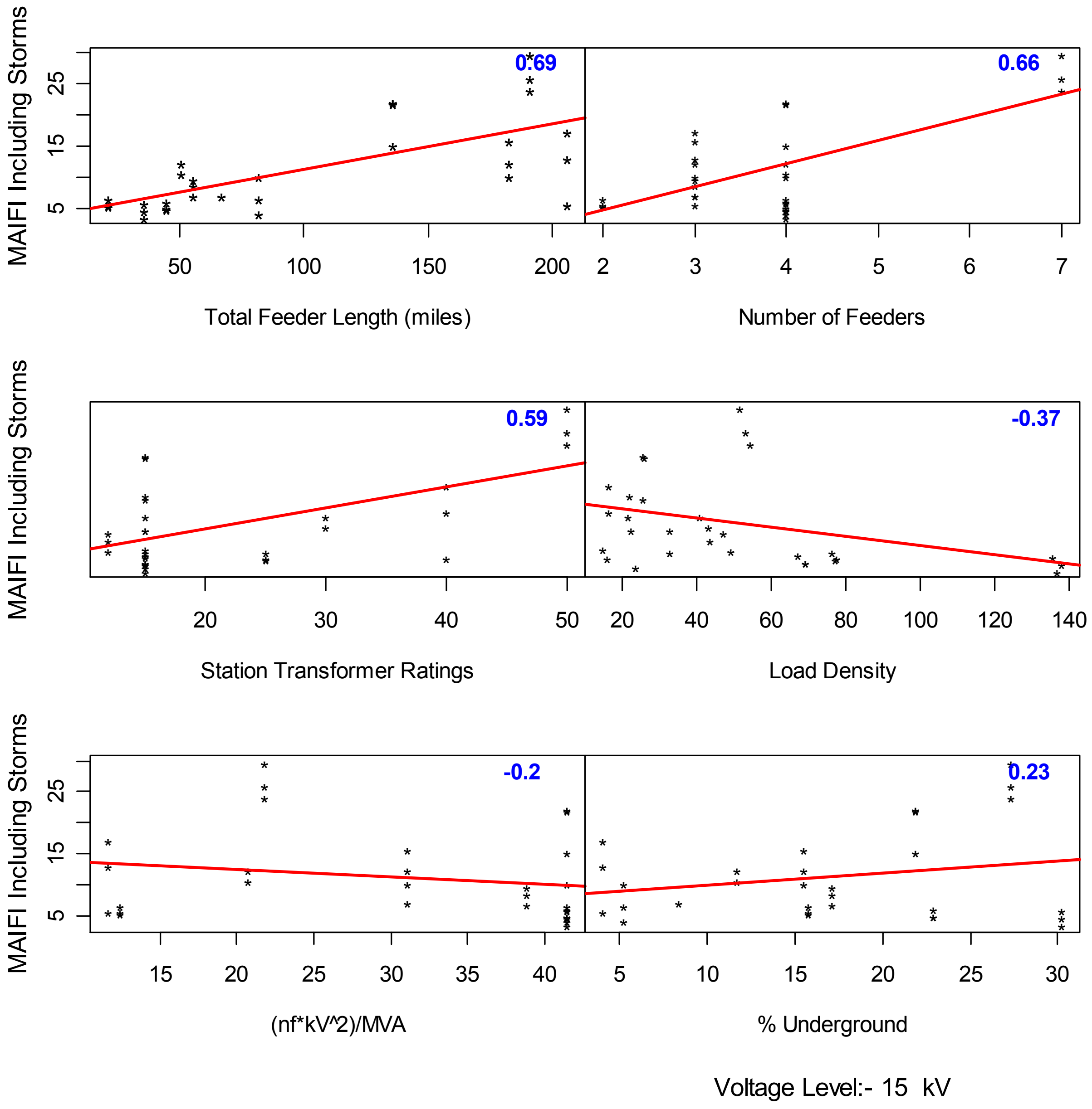


Figure 4-15
MAIFI Site Variations With Various Parameters. Bus Level (15KV) – Utility B

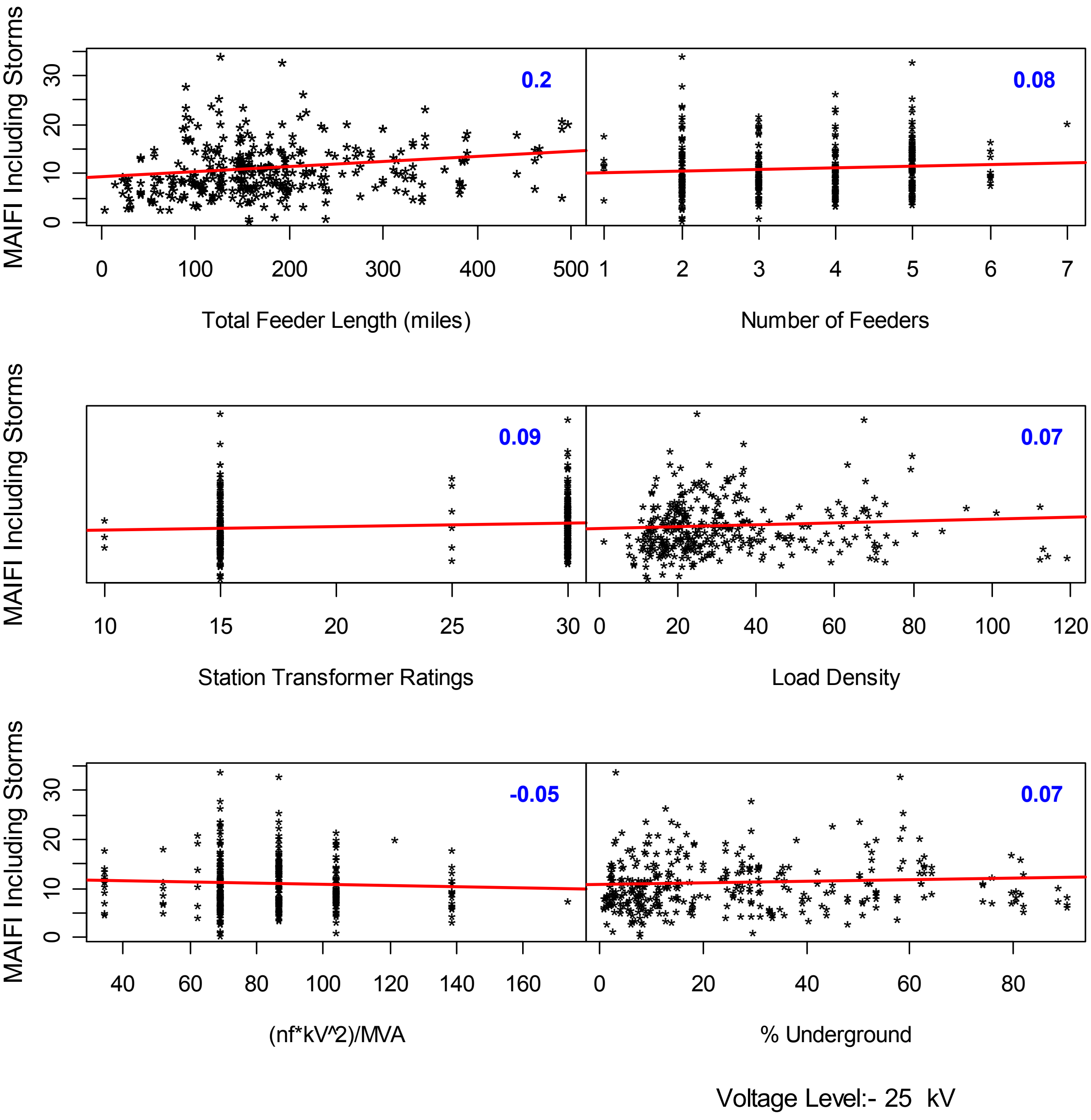


Figure 4-16
MAIFI Site Variations With Various Parameters. Bus Level (25KV) – Utility B

Correlations Between Power Quality and Reliability Variables

The next step is to ascertain if correlation exists between PQ indices and reliability indices. Correlations between $SARFI_X$ (ITIC, 90, 80, 70, 50, 10) with respect to reliability measures such as SAIFI, SAIDI, and outages counts were analyzed for the two utilities. Correlation between MAIFI with respect to reliability measures such as SAIFI, SAIDI, and outage counts were also investigated. The correlations were performed both with and without storms. The following figures and tables show the correlations at the feeder level and the bus level. The strength of relationship between the two variables of interest is quantified by **coefficient of correlation** (r). The number may be expressed either in per unit or as a percentage. Values of r can vary from –1 (-100%) for a perfect negative correlation up to +1 (+100%) for a perfect positive correlation. For example in Table 4-1, the coefficient of correlation between $SARFI_{70}$ and SAIFI with storms is 42% or 0.42 at bus level. In Tables 4-1 through 4-4, the highest SARFI correlations are with the Bus variables, rather than the Feeder variables. This is because outages near the end of a long feeder, while picked up as feeder reliability data would be registered as power quality data on the bus.

Table 4-1
Coefficient of Correlation (Expressed as a Percentage) Between Power Quality and Reliability Variables– Utility A

		With Storms			Without Storms		
Level	SARFI Variable	SAIFI	Number of Outages	SAIDI	SAIFI	Number of Outages	SAIDI
Feeder	$SARFI_{70}$	15	3	14	17	3	15
	$SARFI_{50}$	11	1	9	15	1	14
	$SARFI_{ITIC}$	23	9	20	22	9	19
Bus	$SARFI_{70}$	42	20	31	41	20	41
	$SARFI_{50}$	36	20	29	43	20	48
	$SARFI_{ITIC}$	54	40	39	48	30	45

Table 4-2
Coefficient of Correlation (Expressed as a Percentage) Between Power Quality and Reliability Variables at Bus Level– Utility B

Bus Level		With Storms		Without Storms	
	SARFI Variable	SAIFI	SAIDI	SAIFI	SAIDI
15 kV	SARFI$_{ITIC}$	38	5	40	-3
	SARFI$_{50}$	17	-1	18	-6
	MAIFI (with storms)	38	23	55	60
	MAIFI (w/o storms)	38	12	55	56
25 kV	SARFI$_{ITIC}$	40	28	26	11
	SARFI$_{50}$	32	25	20	7
	MAIFI (with storms)	54	37	35	27
	MAIFI (w/o storms)	20	7	32	24

Table 4-3
Coefficient of Correlation (Expressed as a Percentage) Between Power Quality Variables – Utility A

SARFI$_{70}$	SARFI$_{10}$	SARFI$_{50}$	SARFI$_{90}$	SARFI$_{ITIC}$
Feeder	26	73	60	92
Bus	31	77	60	93

Table 4-4
Coefficient of Correlation (Expressed as a Percentage) Between Bus Level Power Quality Variables – Utility B

Variable	Voltage Class	Storms	$SARFI_{10}$	$SARFI_{50}$	$SARFI_{70}$	$SARFI_{90}$	$SARFI_{ITIC}$
$SARFI_{ITIC}$	15 kV	Included	28	79	100	83	N/A
	25 kV	Included	36	86	100	86	N/A
MAIFI	15 kV	Included	-3	3	--	37	26
		Excluded	-2	3	--	36	25
	25 kV	Included	11	34	--	44	42
		Excluded	10	27	--	27	30

Figure 4-17 and Figure 4-19 shows correlations between SARFI_ITIC and reliability variables, while Figure 4-20 and Figure 4-24 shows the correlation with other forms of SARFI. SARFI_ITIC and SARFI_70 go hand in hand for both utilities.

The numbers in blue are the coefficient of correlation. This shows that the correlation of SARFI_ITIC is better with SAIFI and SAIDI as compared to outages. Note however that the correlation cannot prove that the change in the value of one variable caused the change in the other variable. A strong correlation can be produced simply by chance, by the effect of other external variables not considered in the calculations, or by the cause-and-effect relationship. Additional analysis in the form of linear and generalized regressions is required to verify presence of a cause-and-effect relationship.

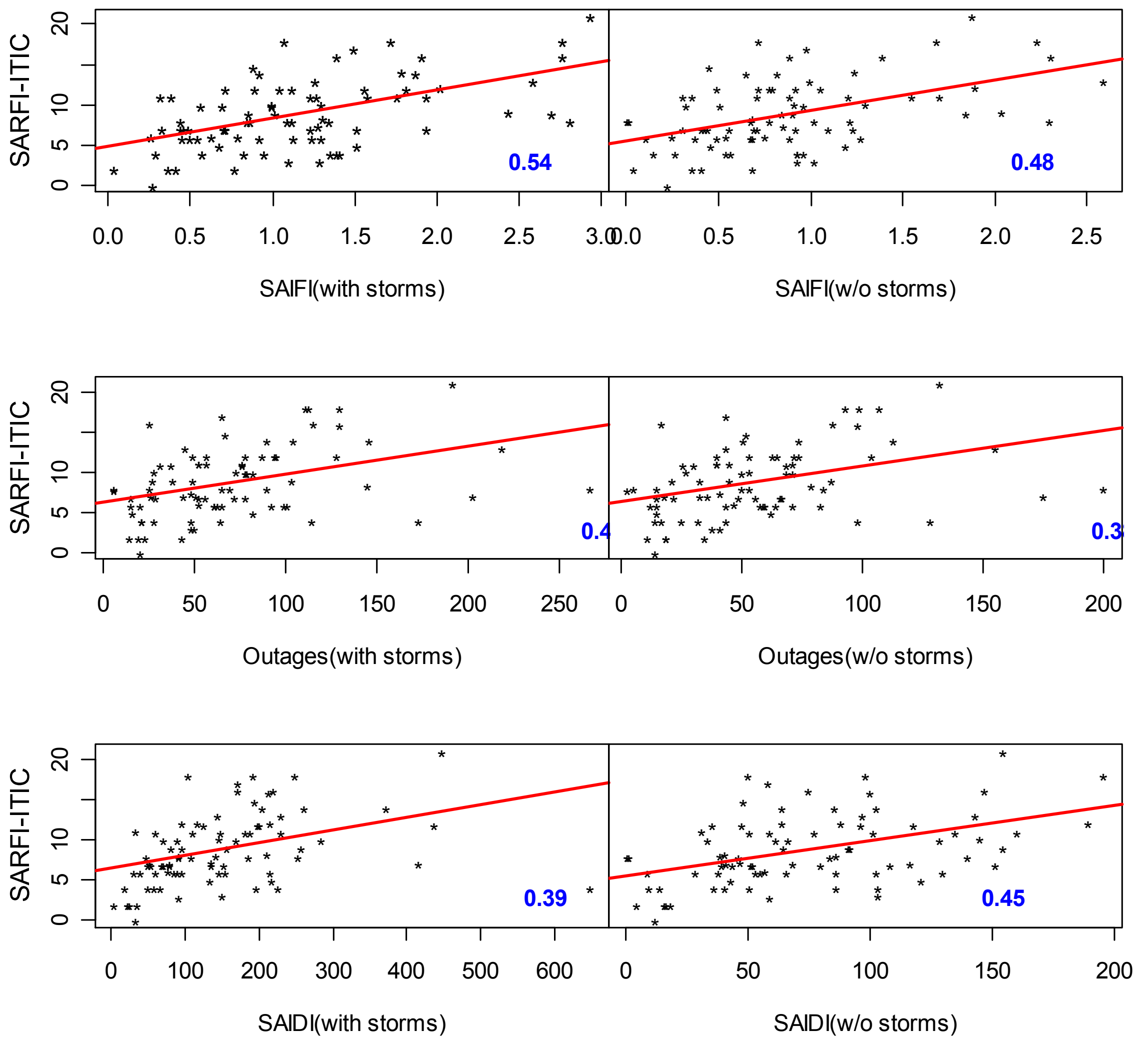


Figure 4-17
SARFI$_{ITIC}$ Site Variations With Reliability Variables. Bus Level – Utility A

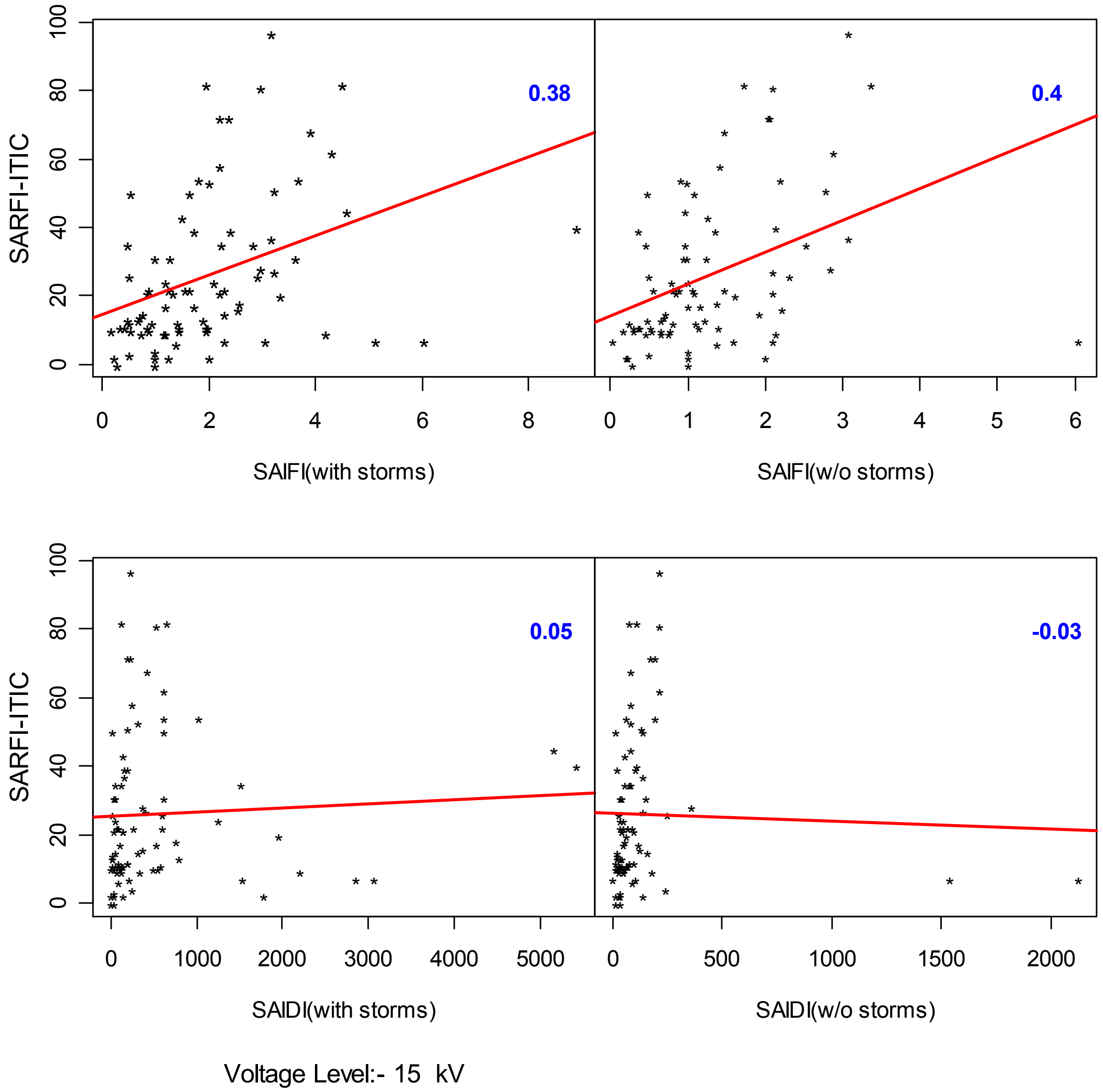


Figure 4-18
SARFI$_{ITIC}$ Site Variations With Reliability Variables. Bus Level – Utility B

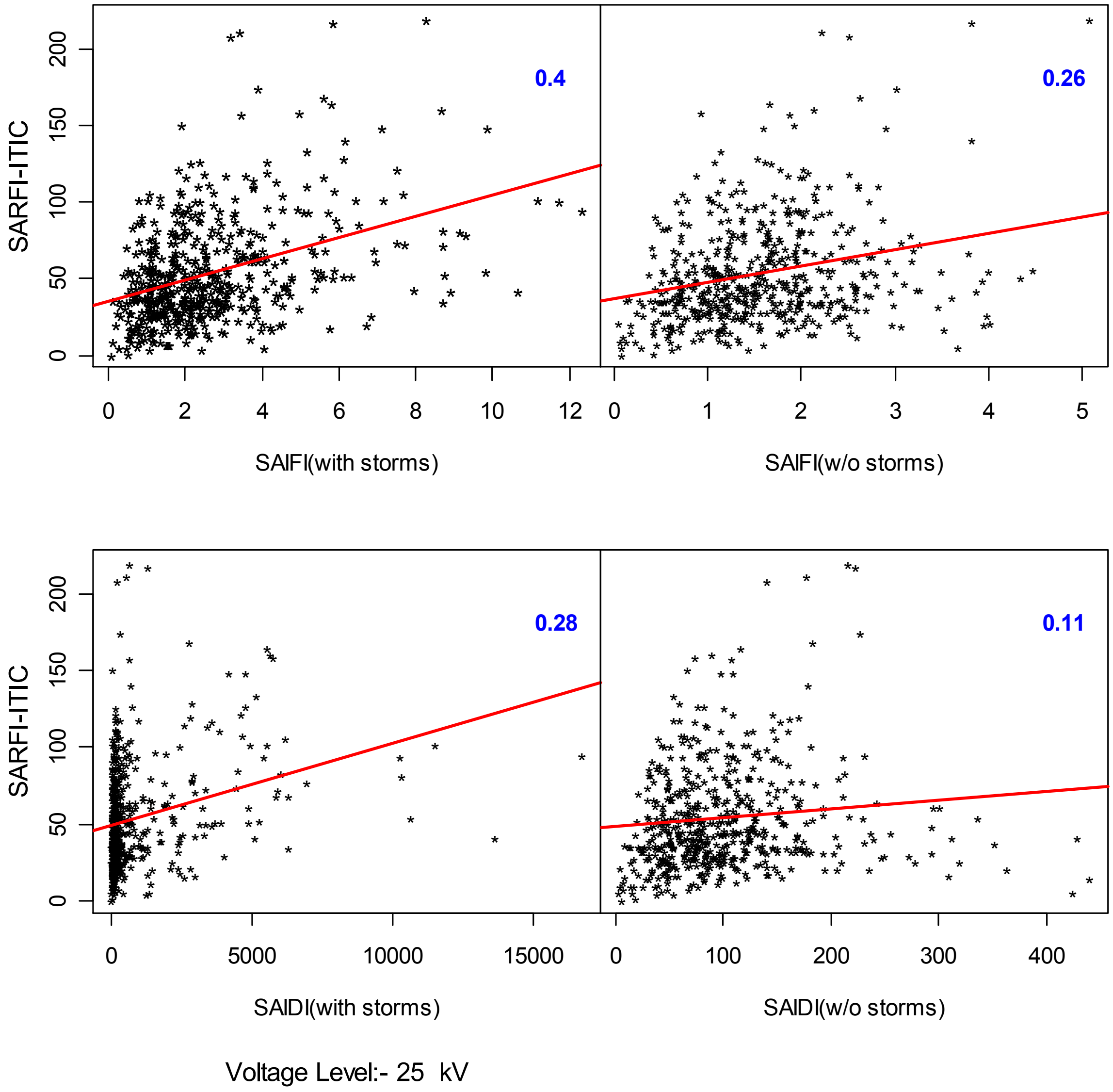


Figure 4-19
SARFI$_{ITIC}$ Site Variations With Reliability Variables. Bus Level – Utility B

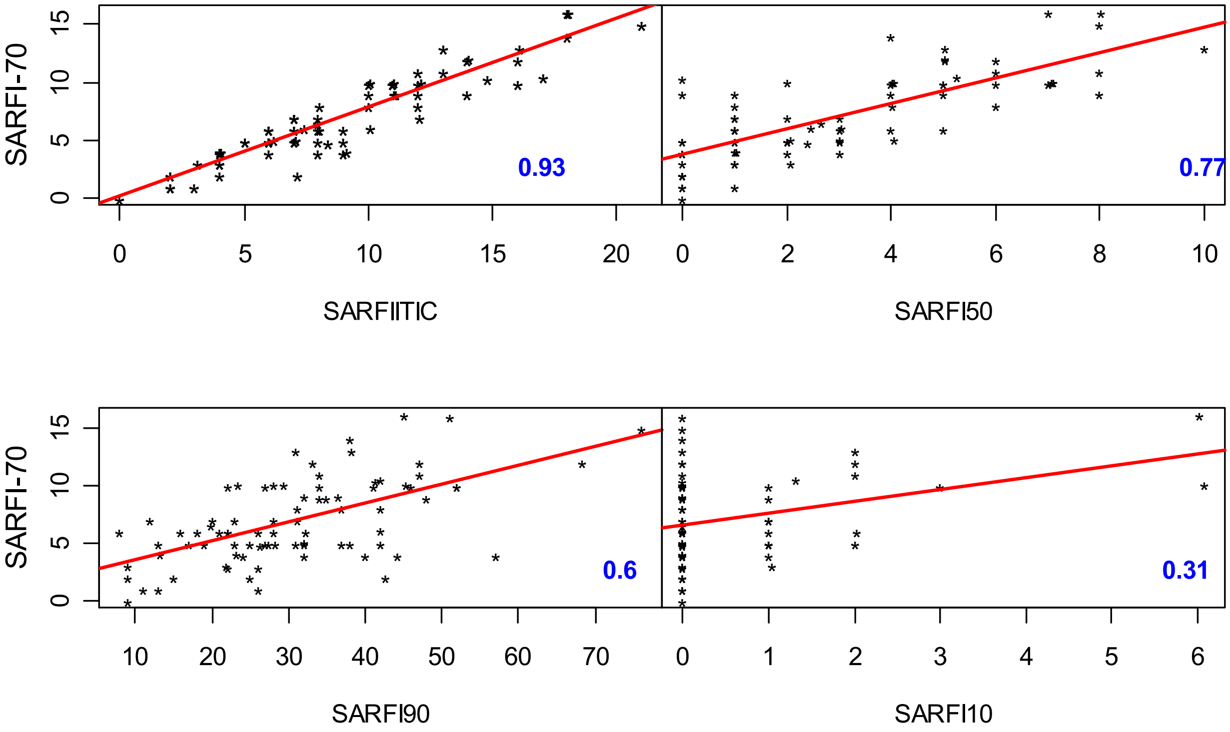


Figure 4-20
Correlations Between Variations of SARFI. Bus Level – Utility A

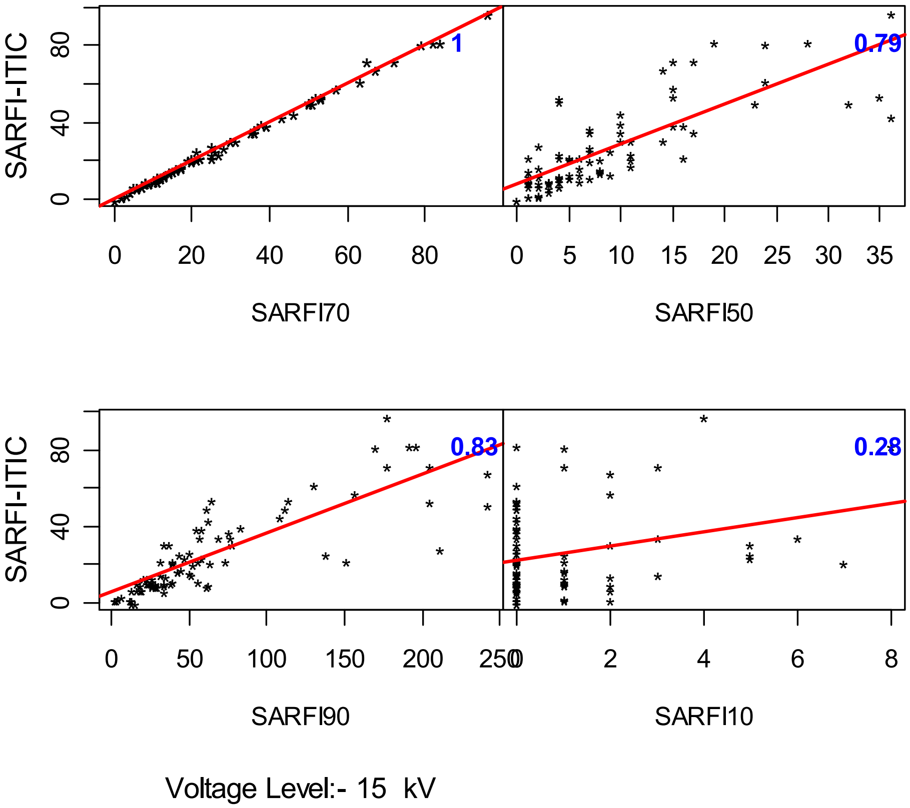


Figure 4-21
Correlations Between Variations of SARFI. Bus Level – Utility B

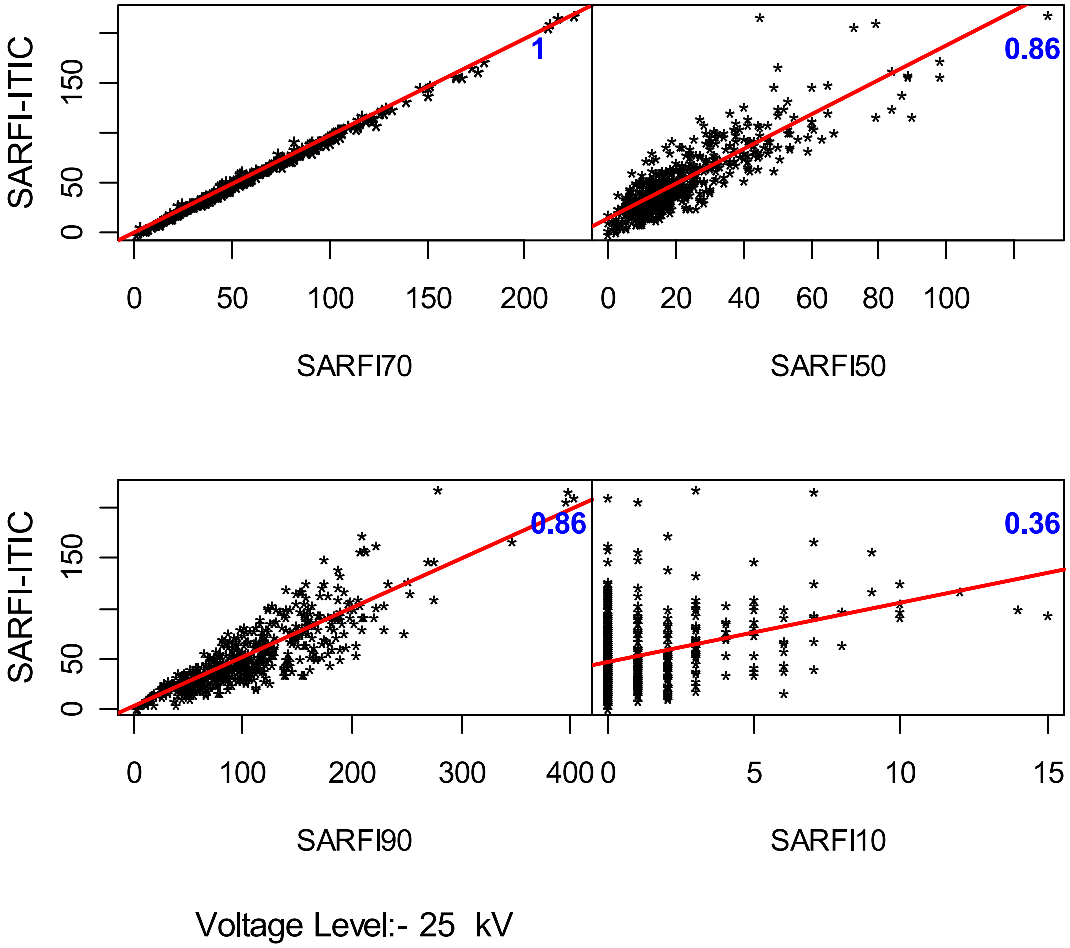


Figure 4-22
Correlations Between Variations of SARFI. Bus Level – Utility B

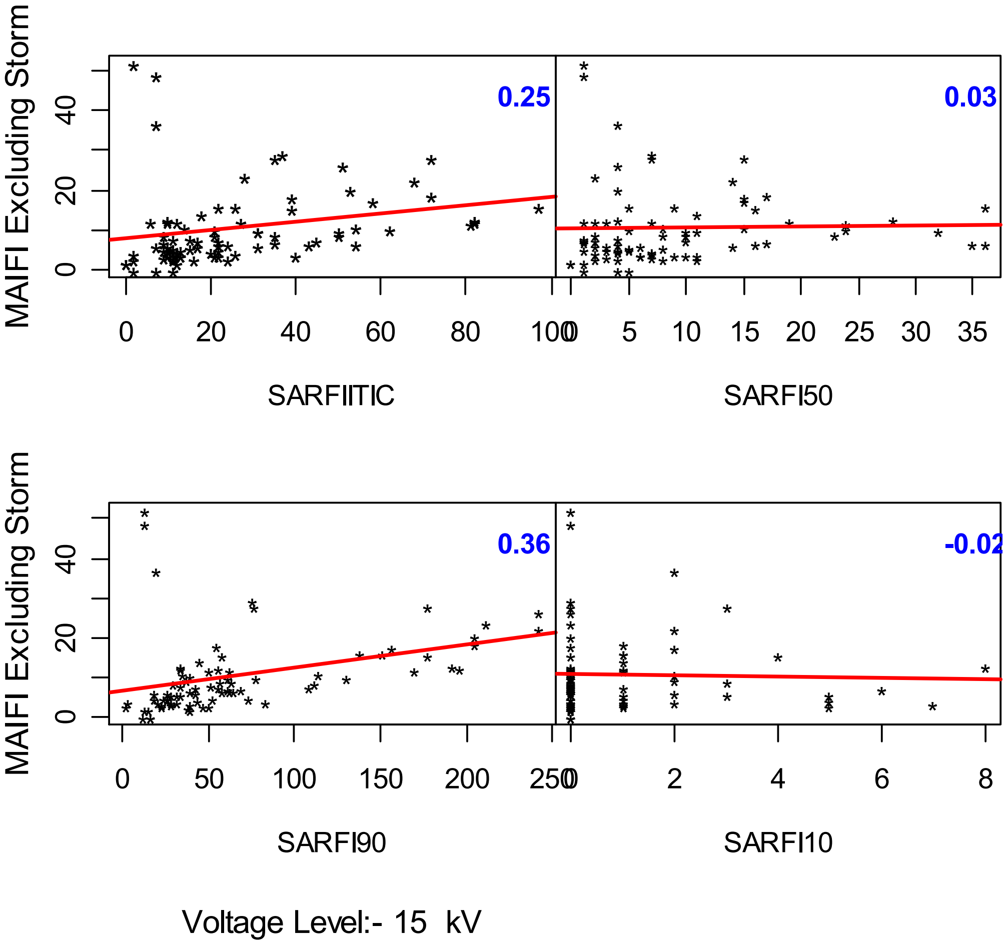


Figure 4-23
Correlations Between Variations of SARFI. Bus Level – Utility B

Note the lack of correlation between MAIFI and $SARFI_{10}$ for Utility B. $SARFI_{10}$ count for Utility B did not include sags due momentary interruptions caused by breaker operation because the monitors were placed upstream of feeder breakers (refer Figure 1-3). A momentary interruption is unlikely to cause deep sag at this monitor location.

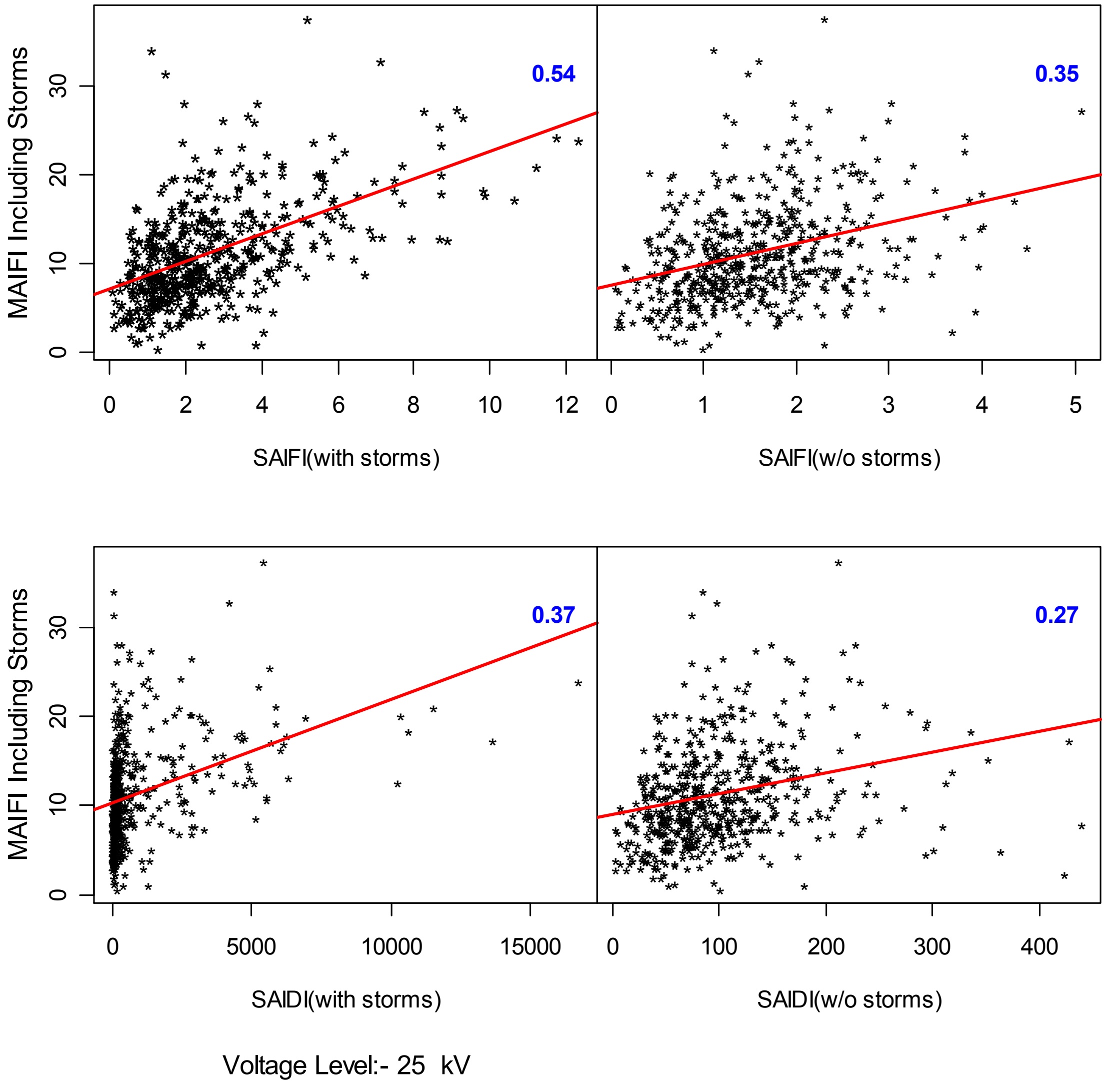


Figure 4-24
Correlations Between Variations of SARFI. Bus Level – Utility B

Variations with Parameters

Table 4-5 and Table 4-6 shows the impact of various site characteristics on $SARFI_{ITIC}$ using substation data. Many trends are expected. Urban sites tend to have fewer sags than suburban or rural sites. Higher voltage systems have more sags than lower voltage systems. No single site characteristic dominates.

Table 4-5
$\text{SARFI}_{\text{ITIC}}$ Site Statistics by Various Site Characteristics Utility A

$\text{SARFI}_{\text{ITIC}}$ Bus Level – Utility A					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	6.000	8.000	8.847	11.780
Total Feeder Length					
	0-30%	4.000	7.000	6.748	8.000
	30%-70%	6.012	9.500	8.840	11.340
	70-100%	7.117	10.000	10.960	14.080
Backbone Length					
	0-30%	5.250	7.010	7.242	9.000
	30%-70%	6.000	9.000	8.798	12.000
	70-100%	6.500	9.100	10.730	15.040
Number of Feeders					
	$n_f < 5$	4.011	7.000	6.519	8.022
	$>5\ n_f < 9$	6.797	10.000	9.933	12.010
	$n_f > 9$	6.500	10.000	10.100	13.070
Station Transformer (MVA)					
	$\text{Xfmr}_{\text{size}} < 25$	6.000	7.667	8.721	11.020
	$>25\ \text{Xfmr}_{\text{size}} < 36$	6.000	8.689	9.003	12.000
	$\text{Xfmr}_{\text{size}} > 36$	6.750	7.558	8.732	11.340
Load Density					
	0-30%	7.000	8.000	9.349	11.060
	30%-70%	6.000	9.000	9.117	12.000
	70-100%	6.000	7.078	8.000	10.000

Table 4-6
$SARFI_{ITIC}$ Site Statistics by Various Site Characteristics Utility B (15KV Class)

$SARFI_{ITIC}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	11.0	21.0	26.6	37.0
Total Feeder Length					
	0-30%	6.3	11.0	12.1	13.0
	30%-70%	12.0	22.0	27.4	35.0
	70-100%	21.3	42.5	42.6	60.0
Percent Underground					
	0-30%	9.3	14.0	18.9	23.5
	30%-70%	10.0	17.0	28.0	43.0
	70-100%	15.0	26.0	33.7	47.0
Number of Feeders					
	$n_f < 5$	6.0	9.0	8.0	11.0
	$>5\ n_f <9$	18.0	26.0	33.7	43.0
	$n_f > 9$	52.0	53.0	57.3	60.5
Load Density					
	0-30%	7.0	10.0	15.0	22.0
	30%-70%	12.0	20.5	32.0	52.0
	70-100%	13.0	33.0	31.6	40.0

Bin Level Correlations with Site Parameters.

More detailed examination of the correlations with site parameters was performed by breaking the parameters into bins of 0-30%, 31-70% and 71-100%. Using Utility A data (Figures 4-25 through Figure 4-27) we observe that better correlation exists between SARFI_ITIC and SAIFI, than correlations between SARFI_ITIC and SAIDI or outages. Utility B data (Figure 4-28 through Figure 4-31) extends this to MAIFI, showing good correlations with SAIFI and poorer correlations with SAIDI (not shown).

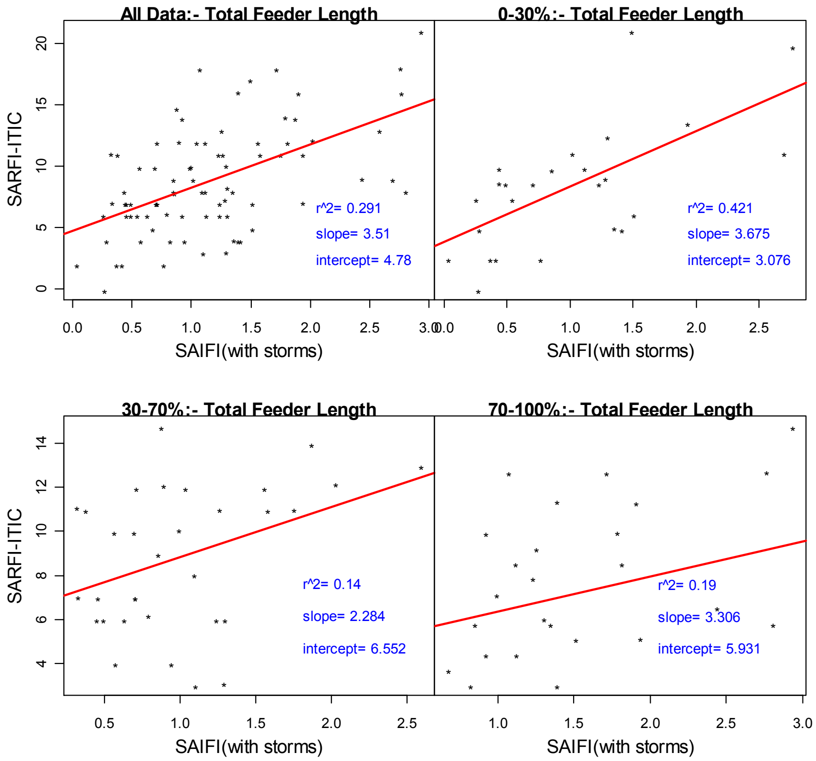


Figure 4-25
SARFI$_{ITIC}$ Bin Variations With SAIFI and Feeder Length. Bus Level – Utility A

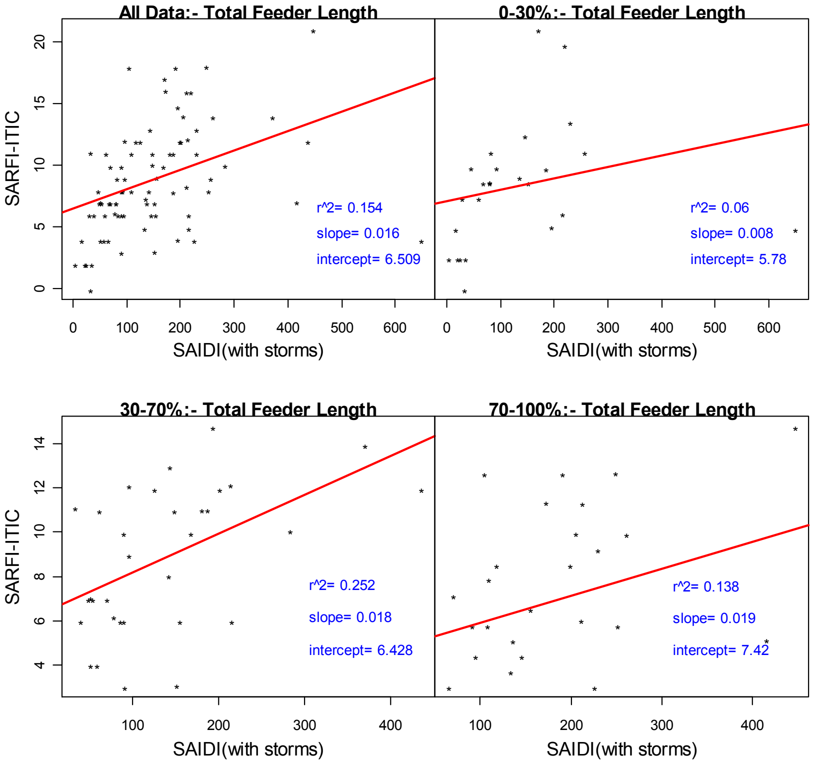


Figure 4-26
SARFI$_{ITIC}$ Bin Variations With SAIDI and Feeder Length. Bus Level – Utility A

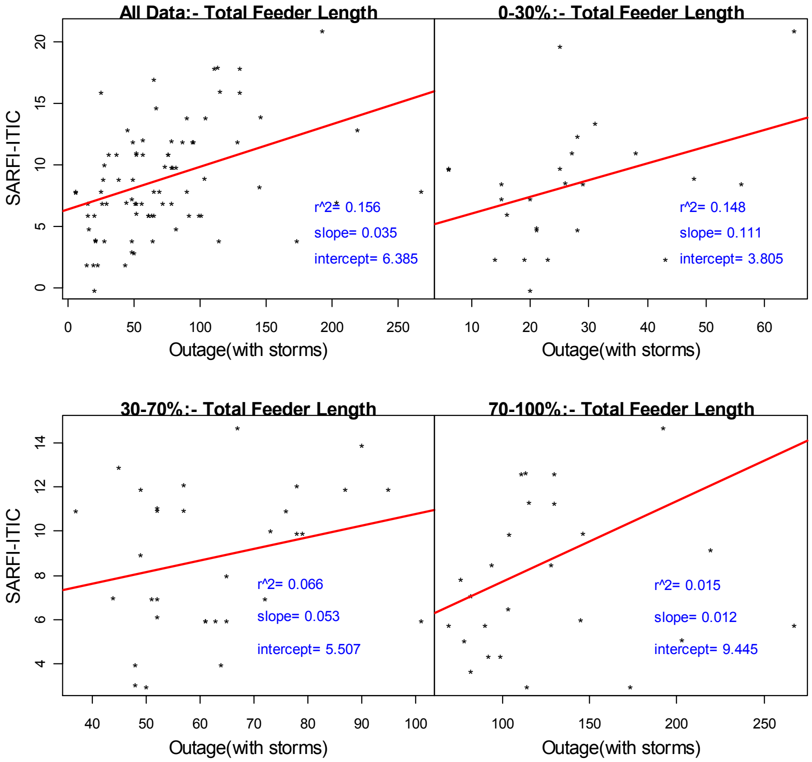


Figure 4-27
SARFI$_{ITIC}$ Bin Variations With Outages and Feeder Length. Bus Level – Utility A

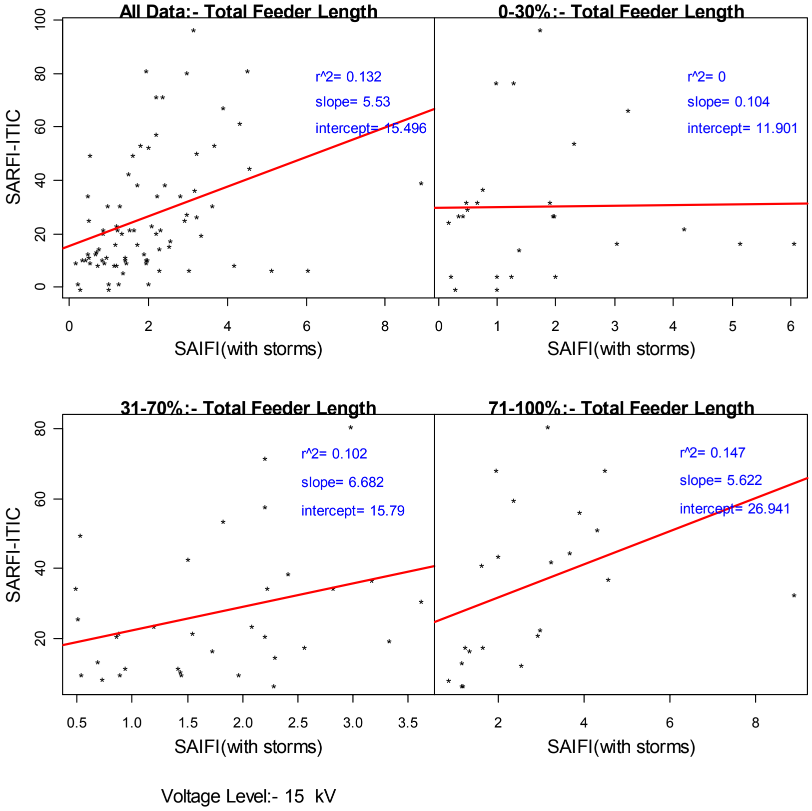


Figure 4-28
SARFI$_{ITIC}$ Bin Variations With SAIFI and Feeder Length. 15 kV Bus Level – Utility B

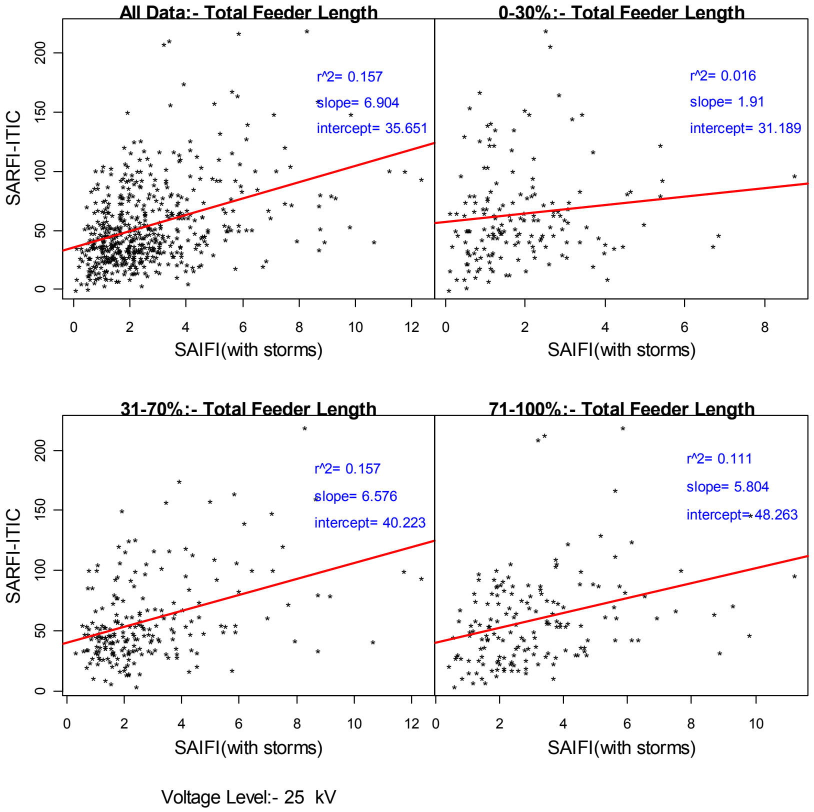


Figure 4-29
$\text{SARFI}_{\text{ITIC}}$ Bin Variations With SAIFI and Feeder Length. 25 kV Bus Level – Utility B

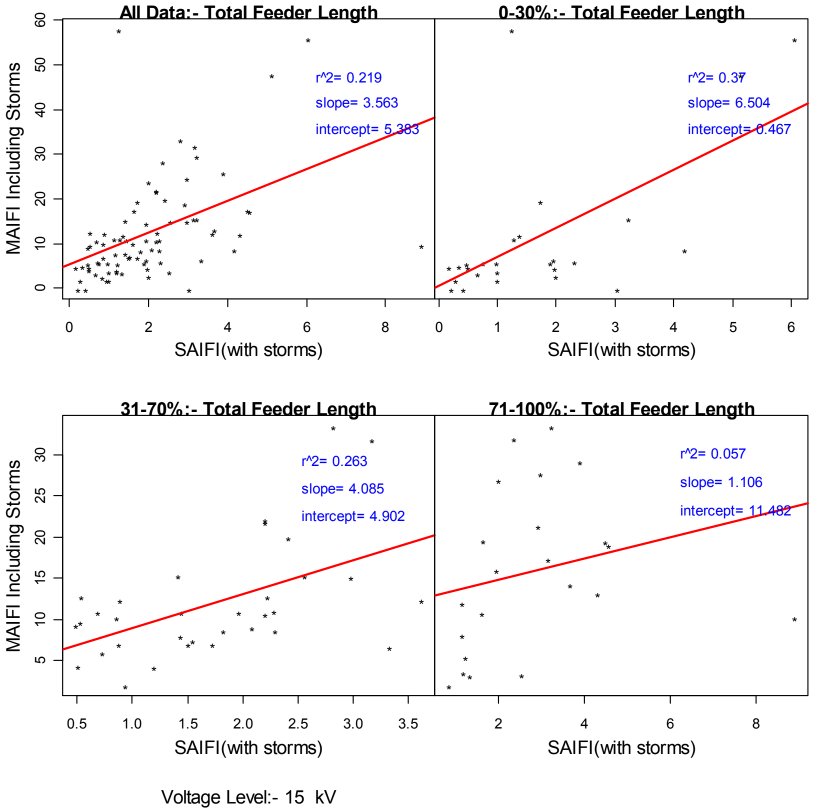


Figure 4-30
MAIFI Bin Variations With SAIFI and Feeder Length. 15 kV Bus Level – Utility B

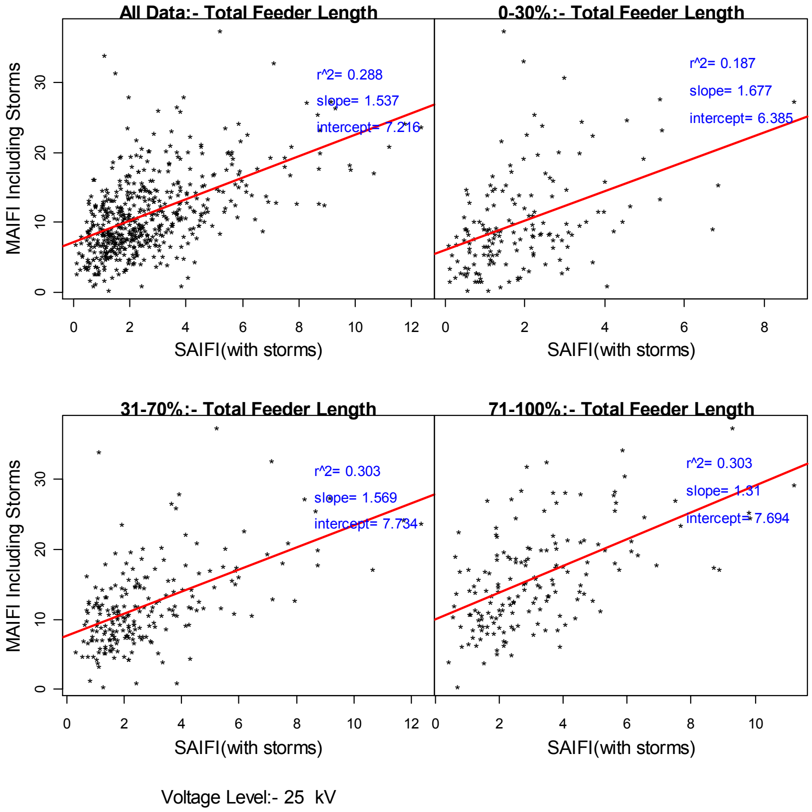


Figure 4-31
MAIFI Bin Variations With SAIFI and Feeder Length. 25 kV Bus Level – Utility B

5
PREDICTION APPROACH

Utilities widely record reliability indices such as SAIFI, SAIDI, number of outages, but most have few power quality monitors. Another major focus of this year's activity is geared towards investigating if it may be possible to refine voltage-sag predictions using reliability indices. In a statistical prediction model, the reliability indices would become another site characteristic. This could improve power quality predictions and may also reveal interesting practical ties between power quality and reliability.

There are no magic procedures to get you the "best model". In some sense model selection is like data mining. The R-square statistic and the multiple correlation coefficients, discussed in the earlier chapters, are descriptive measures of how strong the linear association is between the observed and fitted Y values, but they are not tests of goodness of fit per se. Other measures of fit such as the adjusted R-square and Akaike information criterion (AIC) are designed to take into account the number of X variables in the model. Because R-square can never decrease as new X variables are added, the adjusted R-square or AIC may give a better idea of how the strength of the association between the observed and fitted Y values has changed as X variables are added to or deleted from the model. One method of deciding which model is more appropriate than the other is therefore based on Akaike™s information criterion (AIC). The model with the lowest AIC value gives the best model. The adjusted R-square may in fact decrease if a new X variable does not substantially increase the amount of variation in Y explained by the X variables.

The next step therefore includes:

- Develop and test prediction models[41] for voltage sags and momentary interruptions based on the two datasets. The variables most impacting SARFI_X and MAIFI were tried in various model formulations. The prediction model also includes an estimate of variability, which can be portrayed as prediction limits
- Ascertain how strong are the correlations between SARFI (or MAIFI) and SAIFI (or SAIDI or Outages) relative to correlations between SARFI (or MAIFI) and one or more site characteristics namely total feeder length or backbone length or number of feeders or transformer size or load density or percentage underground versus overhead.
- Investigate the correlations of PQ indices and the reliability indices as well as the prediction models for voltage sags and momentary interruptions at different levels in the power system (feeder level versus the bus level).

---

[41] These are statistical regression models built from optimizations of models containing several of the site characteristics available for the DPQ sites. The main challenges were finding good models, dealing with sites that had unknown site characteristics, and analyzing the variability in the models.

- Compare the linear models with more generalized models that predict the voltage-sag rate as well as momentary interruptions at a particular site, given the feeder reliability indices and site characteristics.

This chapter briefly provides theoretical discussions on the prediction approach used in this study.

Model Fitting

Regression techniques are commonly used to find a model prediction formula. Several model types are available. We explored various linear models. For more information on the modeling strategy used here, see Venables and Ripley[42] and McCullagh and Nelder[43].

A linear model is a fit to an equation of the following form:

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \varepsilon \qquad \textbf{Eq. 5-1}$$

The *x*'s are site characteristics (such as base voltage or lightning flash density), and the *a*'s are coefficients fitted to the model. The ε is the error distribution, which is Gaussian in the traditional linear model.

A generalized linear model is somewhat different from a standard linear model; we used a generalization where the distribution of the error ε is assumed to be a Gamma distribution rather than a Normal distribution in a strictly linear model. A Gamma distribution skews to the right, like a lognormal distribution. The asymmetric distribution improved the predictions over strictly linear models with a Normal (Gaussian) distribution.

The Gamma error distribution is an important feature of the prediction model. It allows more accurate predictions, and it provides a better estimate of the variability and uncertainty in the answer. A comparison of the distributions of the DPQ I/DPQ II feeder data for $\text{SARFI}_{\text{ITIC}}$ along with fits to Normal and Gamma distributions are shown in Figure 5-1.

---

[42] Venables, W. N. and Ripley, B. D., *Modern Applied Statistics with S-PLUS*. Third Edition, Springer, 1999.

[43] McCullagh, P. and Nelder, J., *Generalized Linear Models*, Chapman & Hall, London, U.K., 1989.

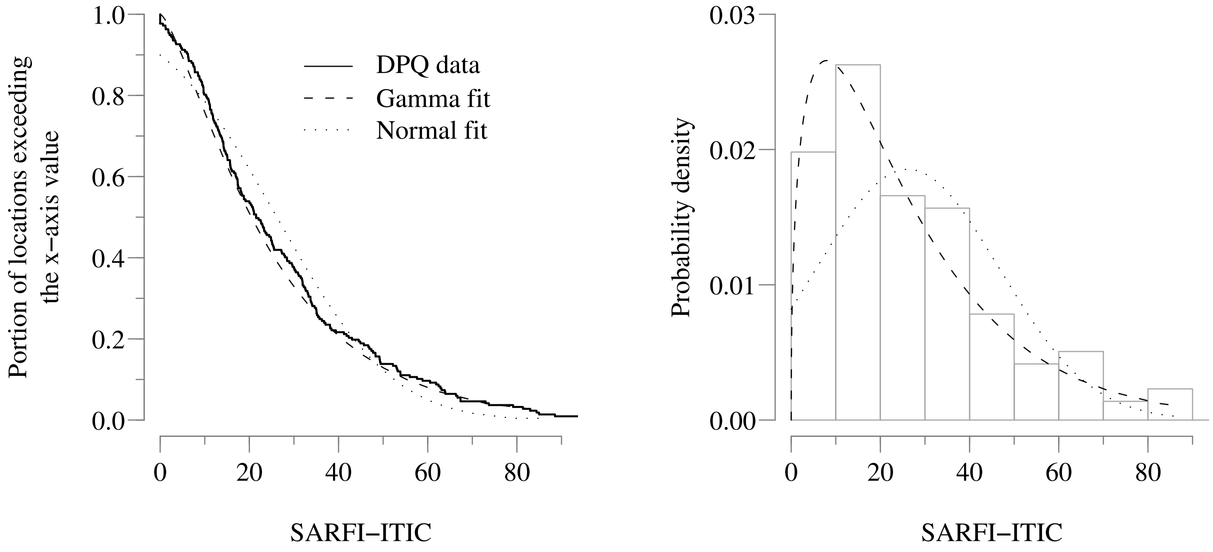


Figure 5-1
Comparison of the $\text{SARFI}_{\text{ITIC}}$ Data With Normal and Gamma Distribution Fits (Cumulative Distributions and Probability Densities)

Another common generalized linear model that is often used in reliability modeling is the Gamma distribution with a logarithmic link. This model is of the following form:

$$y = \exp(a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n) = k_0 \exp(a_1 x_2) \exp(a_1 x_2) \cdots \exp(a_n x_n) \quad \textbf{Eq. 5-2}$$

In this case, the explanatory variables contribute to the prediction multiplicatively rather than additively. We found some logarithmic-link models that were almost as good as the additive linear model with the Gamma distribution. But because they were no better than strictly linear models, the linear models were used.

Several different methods are available for choosing models. The most straightforward is the deviance, the measure that is being optimized. For a standard linear model, the deviance is the residual sum of squares. For a generalized linear model, the deviance is the log-likelihood. Another more general way to grade models is with the AIC (Akaike Information Criterion). The AIC includes a penalty on the number of terms as well as the log likelihood (adding more terms to a model reduces the degrees of freedom). We chose models with the lowest AIC score.

Most of the modeling on this project was developed using the *S* language,[44] either the open-source version known as *R* (www.r-project.org) and/or the commercial S-PLUS software (www.insightful.com).

---

[44] Becker, R. A., Chambers, J. M., and Wilks, A. R., *The NEW S Language*, Chapman & Hall, 1988.

Predicting Variability

One of the most important results of a prediction is an estimate of the variability and accuracy of the model. Two measures are often used to quantify the precision:

- *Confidence interval* — Of all sites with the given site characteristics, the average SARFI of all sites will fill within this interval at a given percentage of confidence (say 95%).
- *Prediction interval* — Of all sites with the given site characteristics, a certain percentage (again say 95%) of them will fall within the interval. Or said another way, there is a given probability that a specific site will fall within the prediction interval.

For the same percent probability, the prediction interval is wider than the confidence interval. For predictions, the prediction interval is much more useful and is used in this project.

For a linear model with normal error distribution of the form:

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \varepsilon \qquad \textbf{Eq. 5-3}$$

the prediction interval calculation can be directly solved. For a generalized linear model, the prediction intervals are more difficult to find. As a first approximation, the prediction intervals may be found from the distribution of the error term, which we've assumed to have a Gamma distribution. We refer to these as "simple prediction limits" in this report. The problem with these simple prediction limits is that they do not include any variability in the prediction due to uncertainty in the model parameters (a_1, a_2, …).

One way to estimate prediction intervals is with a method called the *bootstrap*.[45] To bootstrap, take a random sampling of data with replacement. The resampling *with replacement* is the key to the bootstrap—by randomly taking a set of values from the original data, we get slightly different versions of the data; note that some of the data will be repeated because of sampling with replacement. If we take enough bootstrap copies of the data set (say 1000), perform some analysis of each of the 1000 data sets, then the variability of the answers (whether it's a prediction or some other analysis) gives us an estimate of the variability of the answer relative to the entire population. Bootstrapping assumes that the original data is a good representation of the overall distribution. Bootstrapping does not make any assumption about the underlying distribution. It is also very flexible and can be adapted to many types of problems. A major downside to bootstrapping is that it requires specialized software, and even with modern computers, the processing time for each run can be several minutes.

---

[45] Davison, A. C. and Hinkley, D. V., *Bootstrap Methods and their Applications*. Cambridge University Press, Cambridge, 1997.

As it turns out, the bootstrap approach to prediction levels is not vital for predicting voltage sags. The Gamma error distribution (ε) dominates, so it is possible to directly use the Gamma distribution to estimate prediction intervals of SARFI predictions. This makes it possible to produce useful estimates of variability in a spreadsheet-type model; advanced statistical software is not needed. If we have some power quality monitoring at a site, it would be nice to use that data to improve our prediction about voltage sags. There is a way to use bootstrapped prediction models along with measured data to estimate a new prediction. The approach relies on Bayesian statistics, where probability of a new event can be calculated based on an earlier probability estimate.

The basic idea is to use a bootstrap of model coefficients as a prior distribution, then to use the fundamental Bayesian formula to obtain a new distribution, which will be in the form of a weighted distribution.

6

PREDICTIONS FOR SARFI

Now, we get to the heart of the matter—predictions. In this section, prediction models are developed using data from two utilities. We also discuss extensions that can incorporate measurement data into a prediction. The variability in the data and the variability and accuracy of the predictions are also discussed.

For this work, we concentrate on predictions of voltage sags at *substation* sites. Since SARFI is a measure of voltage

Feeder monitors are typically triggered on current only and depends which side of the breaker it is. Therefore in order to account for what will be the true sag count that a customer on a feeder will see we have to use bus monitor's sag count especially if we are trying to make correlation of voltage sags with respect to site characteristics and reliability. Note for Utility A we did not have feeder monitors. The SARFI_X were obtained from bus monitors. We took this number and assigned this to individual feeders. For Utility B even though we had feeder monitors, these were also triggered on currents and goes not account for what the true sag count will be (faults on the other feeders will also cause a sag on feeder 1. The customer on feeder 1 will also see the sag but the feeder monitor will not count that sag). The plots provided in Chapter 4 are evidence to support that we will not get any correlations if we take bus monitor sag count and correlate these with individual feeders. Note those plots were evidence to suggest that no correlations exist when we try to correlate SARFI_X with feeder characteristics and reliability indices.

Predicting Sags and Momentary Interruptions Based on Predictions Derived From Utility Data

The starting point for predictions is a formula for $\text{SARFI}_{\text{ITIC}}$ or MAIFI based on utility data. It is a linear equation as follows:

$$N = k_0 + k_1 R + k_2 l + k_3 \frac{n_f \cdot kV^2}{MVA_{xfmr}} \qquad \textbf{Eq. 6-1}$$

where,

N = predicted variable: $\text{SARFI}_{\text{ITIC}}$ or MAIFI

l = total exposure (including three-phase and single-phase portions) on the circuit in miles (multiply kilometers by 1.609)

R = Reliability variable: SAIFI, SAIDI or Number of Outages

kV = base line-to-line voltage in kV

n_f = total number of feeders off the substation bus

MVA_{xfmr} = station transformer base rating (open-air rating) in MVA

There are no magic procedures to get you the "best model". In some sense model selection is like data mining. The R-square statistic and the multiple correlation coefficients are descriptive measures of how strong the linear association is between the observed and fitted Y values, but they are not tests of goodness of fit per se. Other measures of fit such as the adjusted R-square and Akaike information criterion (AIC) are designed to take into account the number of X variables in the model Because R-square can never decrease as new X variables are added, the adjusted R-square or AIC may give a better idea of how the strength of the association between the observed and fitted Y values has changed as X variables are added to or deleted from the model. One method of model selection is based on Akaike™s information criterion (AIC). The model with the lowest AIC value gives the best model. The adjusted R-square may in fact decrease if a new X variable does not substantially increase the amount of variation in Y explained by the X variables.

Other informal signs of multi co-linearity are

- Regression coefficients change drastically when adding or deleting an X variable.
- A regression coefficient is negative when theoretically Y should increase with increasing values of that X variable, or the regression coefficient is positive when theoretically Y should decrease with increasing values of that X variable.
- None of the individual coefficients has a significant t-statistic, but the overall F-test for fit is significant.

Need to Decide Whether to do MLR or GLM

If all the assumptions for the multiple linear regression hold, all the residuals should come from the same normal distribution with mean 0. Departures from normality can suggest the presence of outliers in the data, or of a non-normal distribution.

The residual distribution will give an indication of whether error in Y appears to be normally distributed, but will not indicate the cause(s) of the non-normality. A residual is the difference between the observed value of a response measurement and the value that is fitted under the hypothesized model. If the residuals do not appear to be close to following a normal distribution, then transforming the Y variable may be a reasonable alternative.

Histogram for Residuals

The histogram for residuals provides a reference for detecting gross non-normality.

Prediction Model Development Using Data from Utility A and Utility B

The main components of the prediction modeling are model fitting, variable selection, and handling missing data and quantifying variability. Two types of curve fitting approaches and the variability and uncertainty in the models were tried. These include:

Multi-linear Regression based Model

Generalized Linear Model

The model described in Equation (6-1) was used for both the Multiple Linear Regression (MLR) and Generalized Linear Model (GLM). Prediction models were derived using data from the two utilities.

Model coefficients and results obtained using the dataset of Utility A and Utility B are listed in Tables 6-1 through 6-4. It was determined that the best results were found with SAIFI. Prediction models for SAIDI and Outages were also made, but are not shown. Figure 6-1 through 6-8 shows predicted values versus site observations for the Utility A and B data.

It is clear from these results that the best prediction model varies between utility-to-utility depending on the utility’s sub station design and service territory. Substation design and geographical service territory and other aspects specific of a distribution system design will determine the best model. Utility A had more uniform substation design across their region. Also, the geographic conditions for Utility A were similar. Therefore it was deduced that the MLR based predicted models seemed to work better for them. The histogram for residuals, shown in Figure 6-2, suggests that the error between the observed and predicted values for SARFI follows a normal distribution. The best prediction model for Utility B was linear with a Gamma error distribution. However, the coefficient derived using data from Utility A are quite similar to that derived using Utility B.

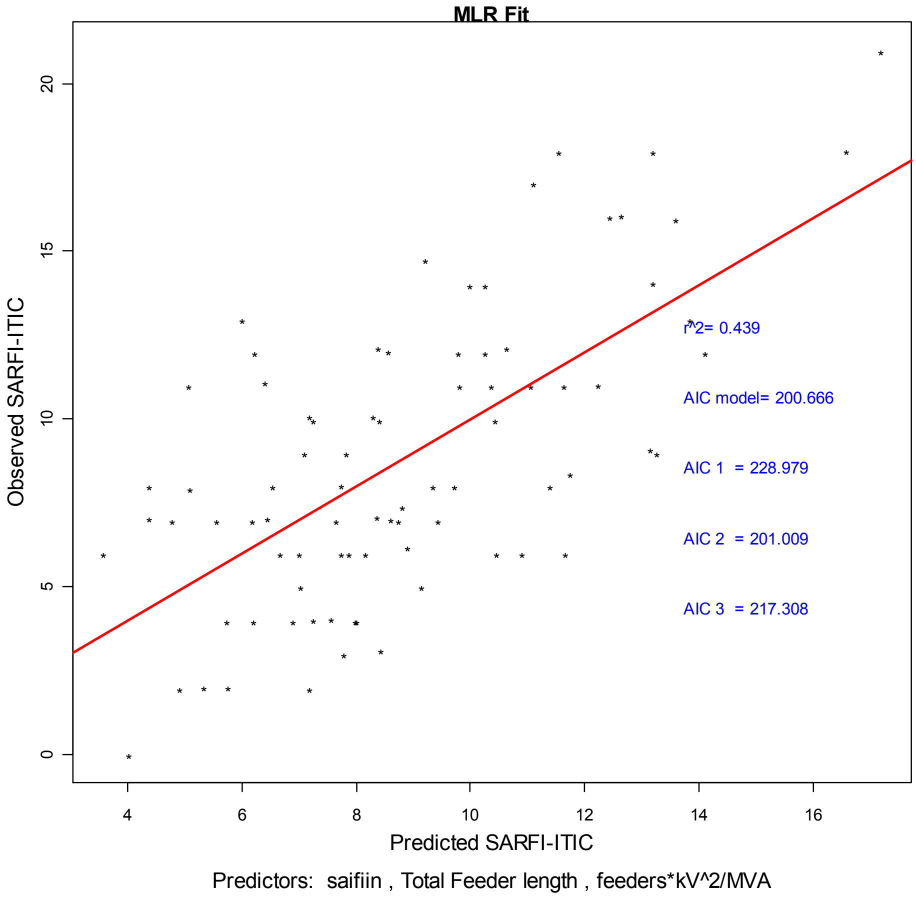


Figure 6-1
Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using MLR Approach (Utility A)

Table 6-1
Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using MLR Approach (Utility A)

Independent Parameters	Estimate	Std. Error	t value	Pr(>|t|)	Significance[46]
(Intercept)	4.38E-01	1.24	0.354	72.4%	
SAIFI (with storms)	3.47E+00	0.59	5.906	0.0%	***
Total Feeder Length (mile)	-1.86E-06	0.00	-1.504	13.7%	
nf*Kv^2/MVA	1.20E-01	0.03	4.462	0.0%	***

---

[46] Signification: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

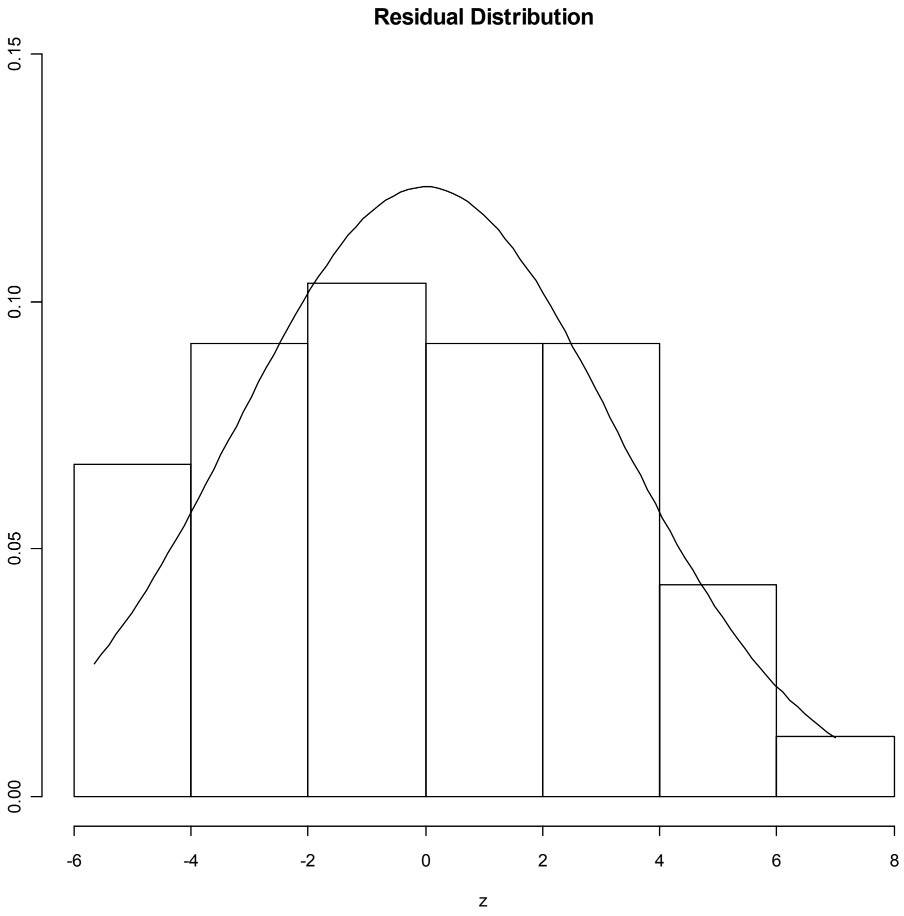


Figure 6-2
Residual Distribution for $\text{SARFI}_{\text{ITIC}}$ Derived Using MLR Approach (Utility A)

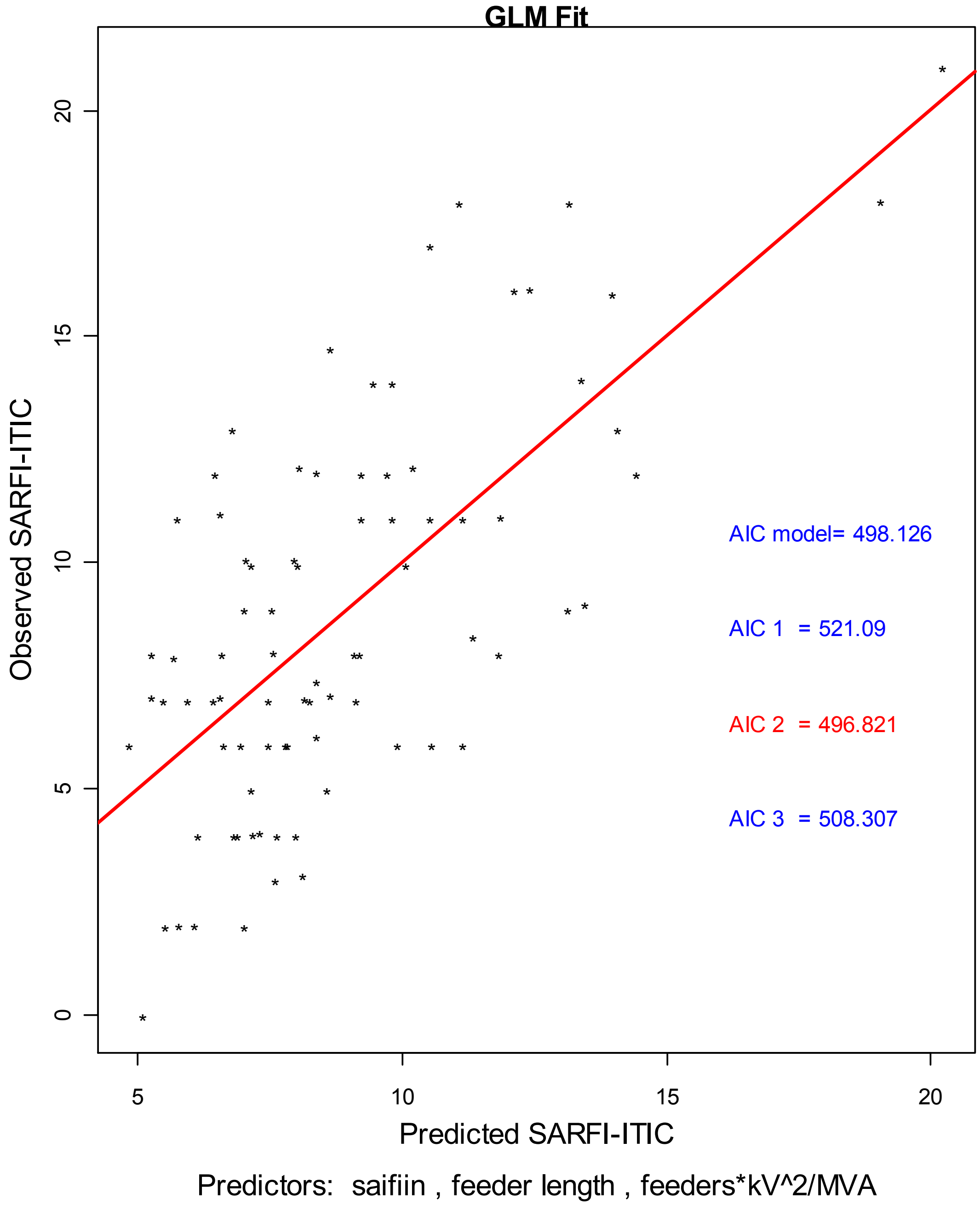


Figure 6-3
Model Coefficients and Results for $\text{SARFI}_{\text{ITIC}}$ Derived Using GLM Approach (Utility A)

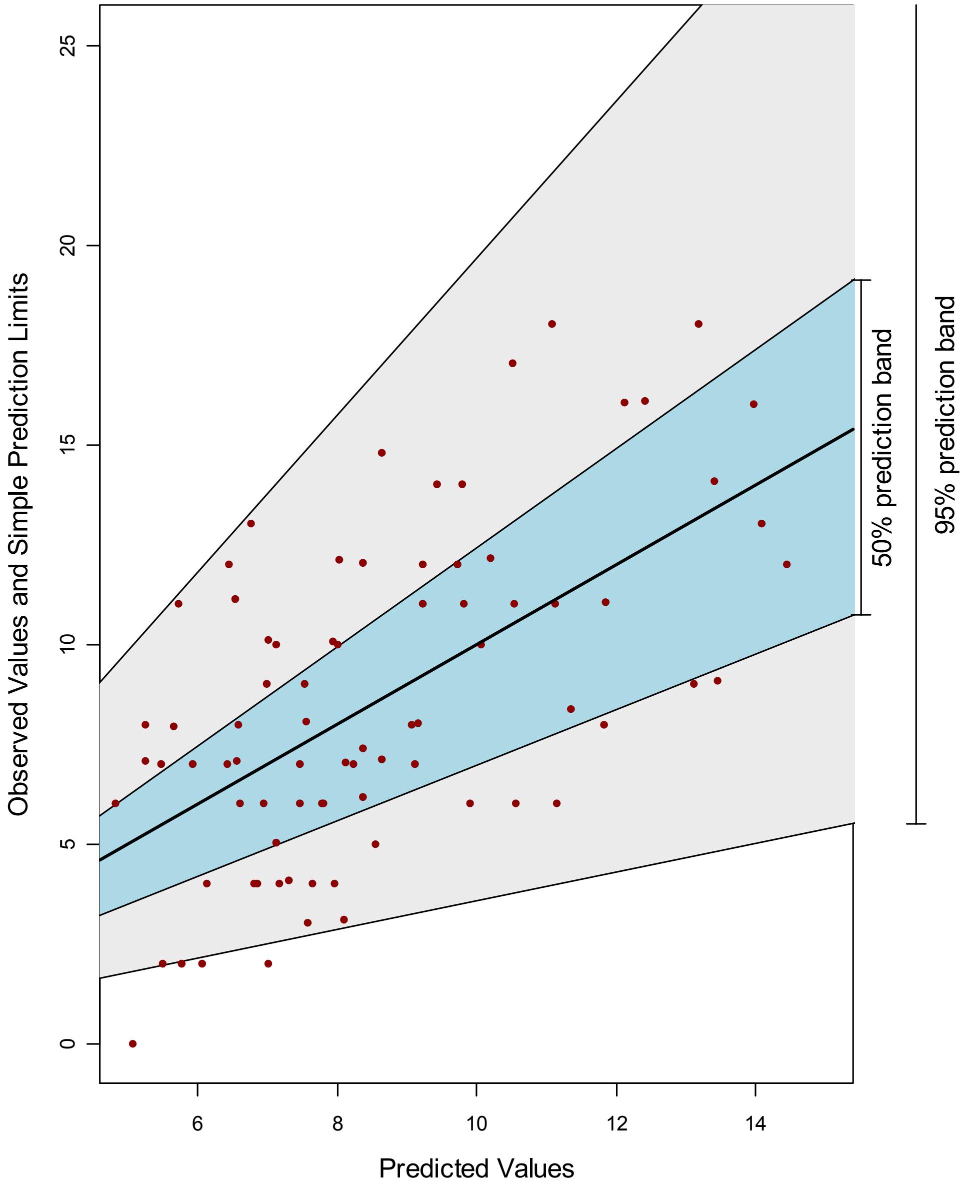


Figure 6-4
Observed $\text{SARFI}_{\text{ITIC}}$ Versus Predicted $\text{SARFI}_{\text{ITIC}}$ Derived Using GLM Approach (Utility A)

Table 6-2
Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using GLM Approach (Utility A)

Independent Parameters	Estimate	Std. Error	t value	Pr(>|t|)	Significance[47]
(Intercept)	1.26E+00	0.16	8.08	0.00%	***
SAIFI (with storms)	3.59E-01	0.07	4.87	0.00%	***
Total Feeder Length (mile)	-1.31E-07	0.00	-0.84	40.11%	
nf*Kv^2/MVA	1.18E-02	0.00	3.48	0.08%	***

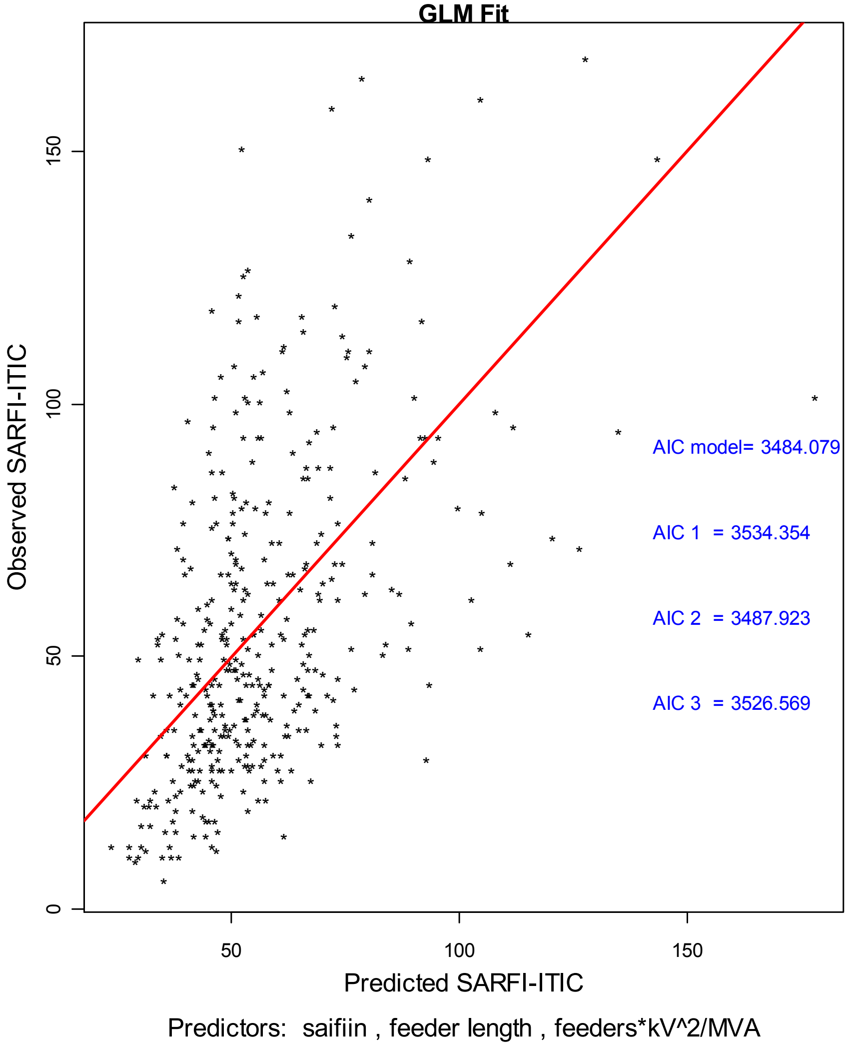


Figure 6-5
Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using GLM Approach (Utility B)

---

[47] Signification: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

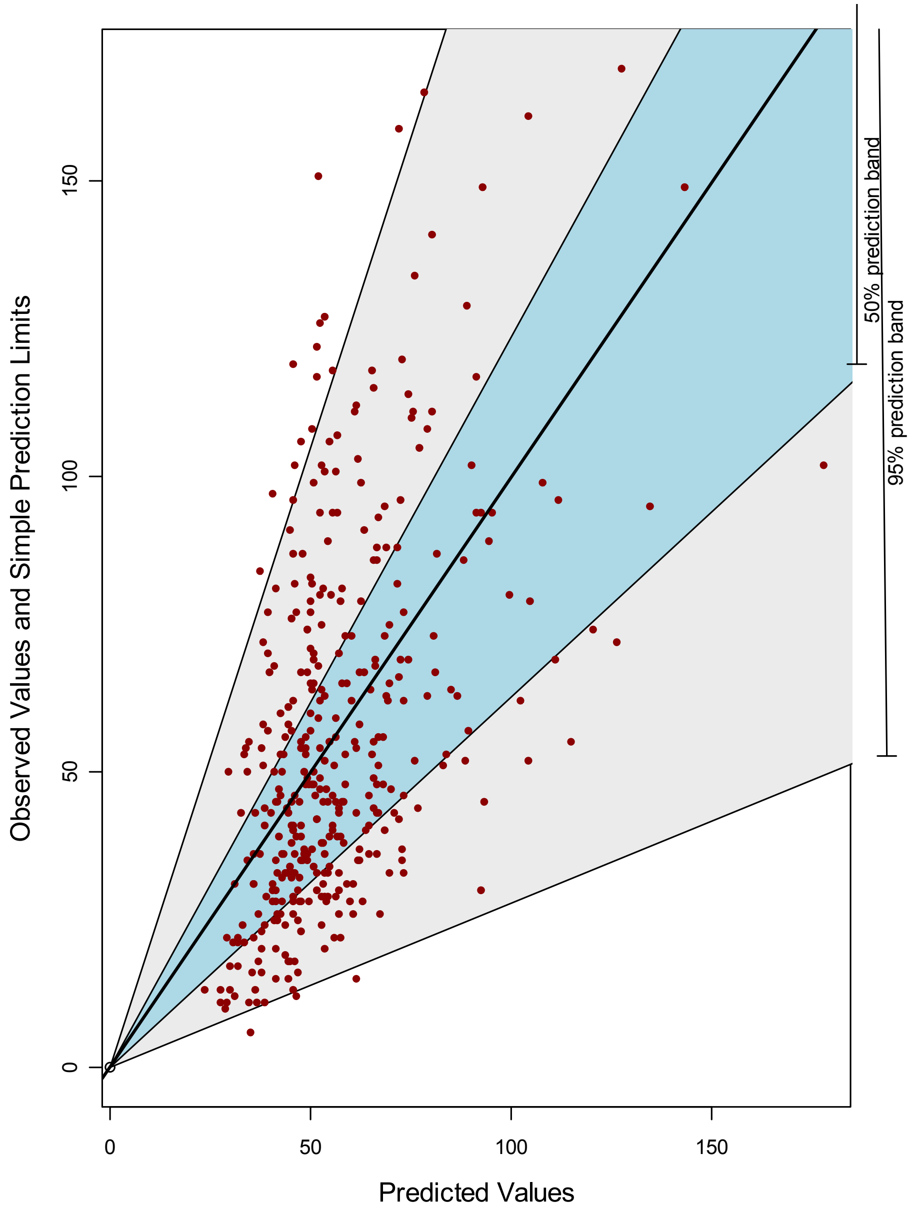
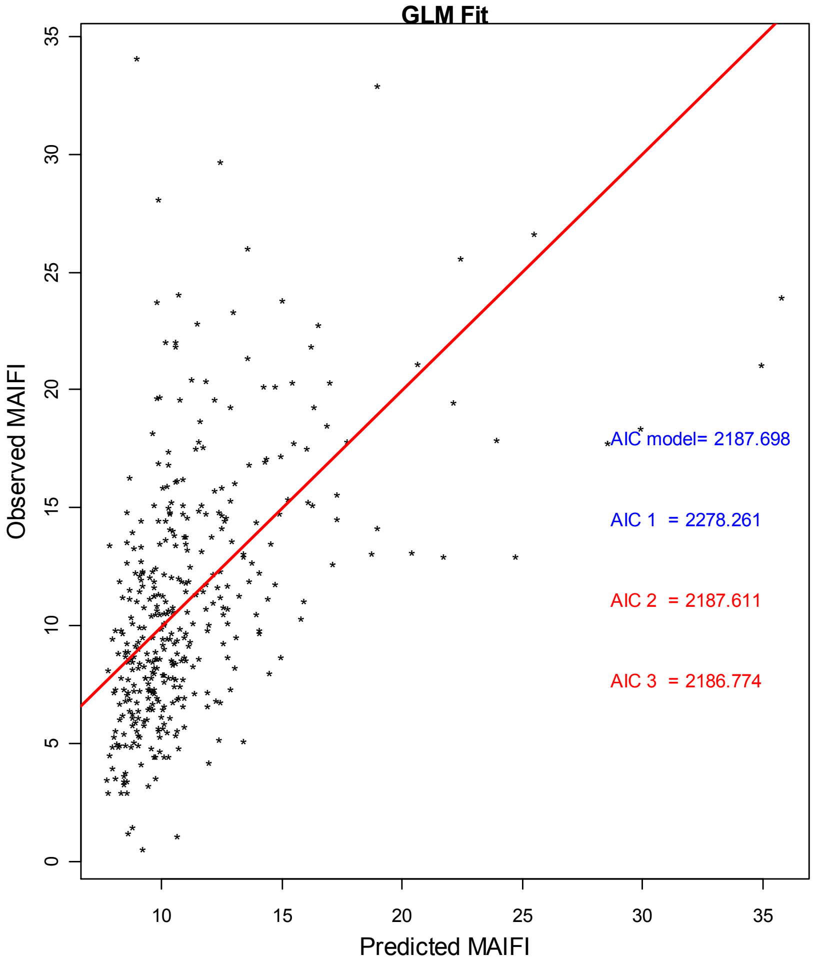


Figure 6-6
Observed $SARFI_{ITIC}$ Versus Predicted $SARFI_{ITIC}$ Derived Using GLM Approach (Utility B)

Table 6-3
Model Coefficients and Results for SARFI$_{ITIC}$ Derived Using GLM Approach (Utility B)

Independent Parameters	Estimate	Std. Error	t value	1-Pr(>|t|)	Significance[48]
(Intercept)	3.01E+00	9.14E-02	32.891	100.00%	***
SAIFI (with storms)	1.06E-01	1.49E-02	7.12	100.00%	***
Total Feeder Length (mile)	7.27E-04	2.82E-04	2.582	98.98%	*
nf*Kv^2/MVA	7.32E-03	1.04E-03	7.066	100.00%	***



Predictors: saifiin , feeder length , feeders*kV^2/MVA

Figure 6-7
Model Coefficients and Results for MAIFI Derived Using GLM Approach (Utility B)

---

[48] Signification: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

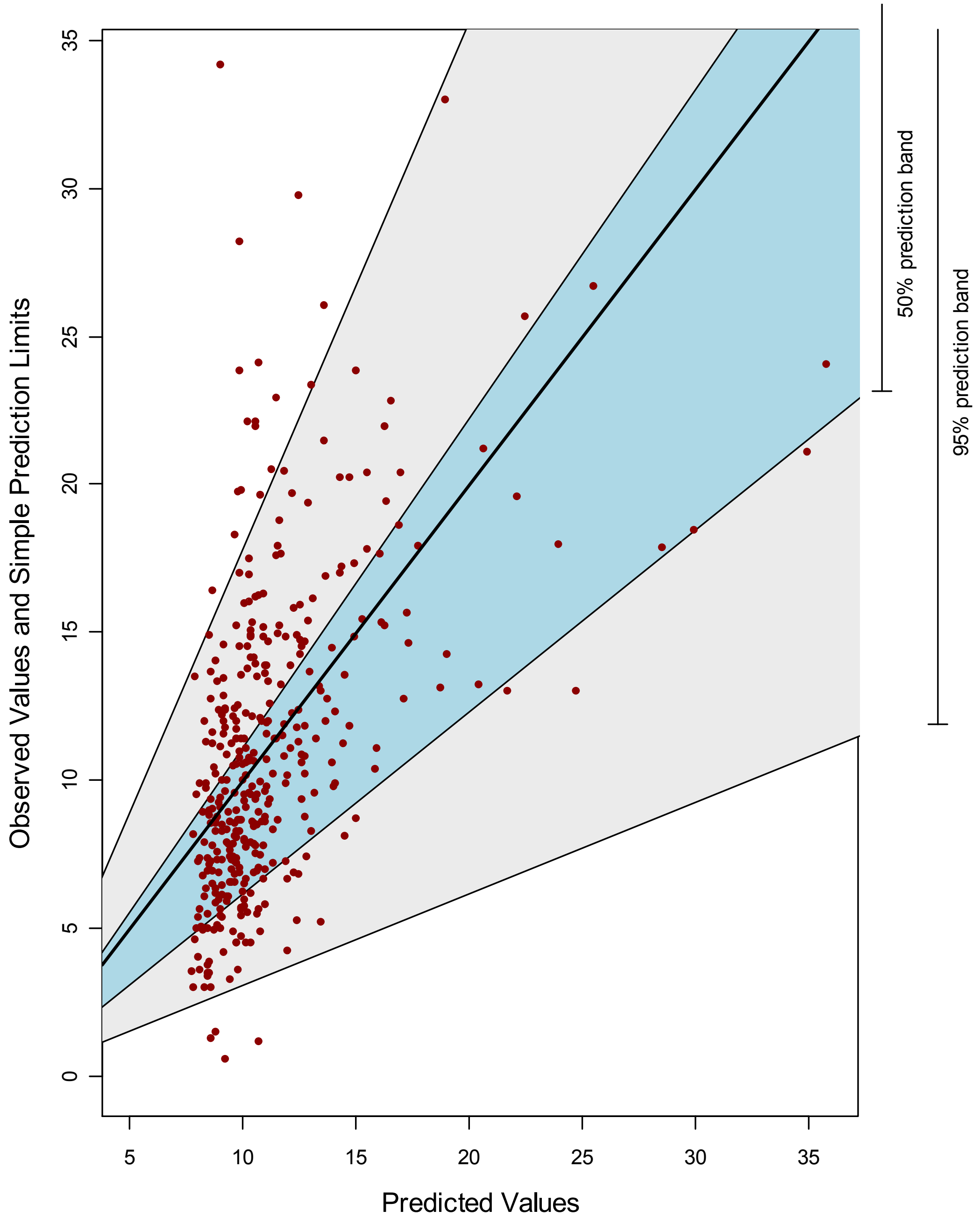


Figure 6-8
Observed MAIFI Versus Predicted MAIFI Derived Using GLM Approach (Utility B)

Table 6-4
Model Coefficients and Results for MAIFI Derived Using GLM Approach (Utility B)

Independent Parameters	Estimate	Std. Error	t value	1-Pr(>|t|)	Significance[49]
(Intercept)	1.08E+00	8.03E-02	25.86	100.00%	***
SAIFI (with storms)	1.24E-01	1.30E-02	9.486	100.00%	***
Total Feeder Length (mile)	3.54E-04	2.47E-04	1.43	84.60%	
nf*Kv^2/MVA	-9.42E-04	9.10E-04	-1.035	69.90%	

Note the high degree of variability in the predictions—these are not precise estimates, but they do well in predicting observations within a given prediction band.

Other Ways to Estimate Voltage Sags

The procedure documented in this report is not the only way to estimate the frequency of voltage sags at a site. The prediction models described here are useable with a few easily available inputs. Keep in mind that even though the site parameters are important electrical characteristics, the prediction model is not based on an electrical model—it is a derived regression model. As such, it is difficult to use for analyzing the effect of some changes on voltage sags such as faster relaying, more tree trimming, animal guards, and so on. Another common way to estimate voltage sags is using a model of the system and a short-circuit program. We briefly discuss two variations of this.

The calculation of the voltage magnitude at various points on a system during a fault at a given location is easily done with any short-circuit program. We make the fairly accurate assumption that the fault impedance is zero. The engineer or computer program finds the duration of the sag using the time-current characteristics of the protective device that should operate along with the fault current through it.

Based on a short-circuit program, the *fault positions method* repeatedly applies faults at various locations and tallies the voltages at specified locations during the faults. The runs, which may apply thousands of fault locations, result in predictions of the number of voltage sags below a given magnitude at the specified locations. This procedure is well documented in the IEEE Gold Book[50] (see also Conrad et. al.[51]).

---

[49] Signification: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

[50] IEEE Std. 493-1997, *IEEE Recommended Practice for the Design of Reliable Industrial and Commercial Power Systems (Gold Book).*

[51] Conrad, L., Kevin, L., and Cliff, G., "Predicting and Preventing Problems Associated with Remote Fault-Clearing Voltage Dips," *IEEE Transactions on Industry Applications*, vol. 27, pp. 167–72, 1991.

The faults are applied along each line in a system. The end results are scaled by the fault rate on the line, which can be based on historical results or typical values for the voltage and construction.

We need considerable detail for the fault-positions analysis, especially a complete system model including proper zero-sequence impedances and transformer connections (these are left out of many transmission system load-flow models).

Another simpler method for voltage sags is the *method of critical distances* (Bollen[52]). The approach is to find the farthest distance, the *critical* distance, to a fault that causes a sag of a given magnitude. Pick a sag voltage of interest, 0.7 per unit for example. Find the critical distance for the chosen voltage. Using a feeder map, add up the circuit lengths within the critical distance. Multiply the total exposed length by the fault rate—this is the number of events expected. This method is not as accurate as the fault positions method, but is much simpler: We can calculate the results by hand, and the process of doing the calculations provides insight on the portions of distribution and transmission system that can cause sags to the given customer. We can also target this *area of vulnerability* for inspection or additional maintenance or apply faster protection schemes covering those circuits (to clear faults and sags more quickly).

Whether using the critical distance or the fault positions method, the key input—and the most unknown parameter—is the fault rate. Rates of voltage sags are directly proportional to rates of faults. Good results from the circuit modeling approaches require good estimates of fault rates. Estimates of the uncertainty and variability associated with a prediction must come from knowledge of the variability of fault rates.

Both the circuit modeling approach and the statistical prediction approach described in this report have their uses. The statistical approach has the advantage that it is fast and easy (at least the plain prediction equation without the bootstraps). We can also crudely use it to evaluate some simple changes in circuits such as increasing or decreasing the number of feeders from the bus. The statistical approach is also based on real data and easily provides a good idea of the prediction band within a given confidence level, and with some extensions, the statistical approach can incorporate measurement data. The circuit modeling approach has the advantage that we can use it to estimate the effect of more precise circuit changes: moving circuit sections, adjusting voltage regulation, increasing source stiffness, and so on. The circuit modeling approach is more work and requires much more data input. For some applications, both methods may be appropriate—the statistical prediction gives a quick answer and a good idea of the variability; the circuit modeling approach can be used to evaluate a wider range of solution options.

---

[52] Bollen, M. H. J., Understanding Power Quality Problems: Voltage Sags and Interruptions, IEEE Press, New York, 2000.

7
SUMMARY

Data from two utilities that participated in the DPQ II project was used for the correlation analysis between power quality and reliability indices. Inferences reached from this year's activity are:

- The PQ indices recorded by bus monitors do not correlate with feeder characteristics.
- Most of the PQ and reliability indices follow a right skewed distribution. Most of the indices for both the utilities followed Gamma distribution.
- MAIFI and $SARFI_{90}$ are highly correlated. Similar correlations exist between MAIFI and SAIFI. Most faults are likely to fall in the $SARFI_{90}$ category. Faults close to the substation will fall in $SARFI_{10}$ category.
- SAIFI gave the best correlation with PQ indices (SARFI as well as MAIFI). SAIDI and outage count did not have significant correlations with PQ indices.
- Adding SAIFI to the prediction model will improve the existing prediction models.
- Feeder length does not influence the prediction model at bus level although coefficient of correlation between feeder length and $SARFI_{70}$ and $SARFI_{ITIC}$ is significant (Refer figures 4-9 through 4-12). The strong correlation is merely because of the effect of other independent variables such as transformer size and voltage level that are not considered in the calculation of correlation. Regression analysis shows that there is no direct "cause-and-effect" relationship between feeder length and $SARFI_{70}$ and $SARFI_{ITIC}$.
- The product term (number of feeders*kV^2/MVA) has little effect on the prediction model. However, this product term is a better indicator than the individual characteristics (number of feeders, voltage level and transformer MVA) in the product.
- There is no universal prediction model for all the utilities. The best prediction model varies between the utilities based on the utility's sub station design and service territory. Substation design and geographical service territory and other aspects specific of a distribution system design will determine the best model. Utility A had more uniform geographic conditions and substation design across their region. The MLR based predicted models seemed to work better for Utility A. The histogram for residuals, shown in Figure 6-2, suggests that the error between the observed and predicted values for SARFI follows a normal distribution. The best prediction model for Utility B was linear with a Gamma error distribution. However, the coefficients derived using data from Utility A are quite similar to those from Utility B.

The correlations of power quality indices with site variables:

- Decrease with load density at 15 kV, but increase at 25 kV.
- Are highly variable with voltage levels.
- Increase with feeder length, the effect is greater at 15 kV than at 25 kV.
- Increase with percent underground unless storms are included.
- Usually increase with numbers of feeders on a bus, except no correlation with SAIFI or SAIDI.
- Are highly variable with transformer size.
- Feeder level correlations are less significant than bus level.

Improving Predictions from the Correlation Analysis

Inferences obtained from the correlation analysis are used to improve SARFI/MAIFI prediction fit. The SARFI/MAIFI predictions fit the following equation:

$$N = k_0 + k_1 R + k_2 l + k_3 \frac{n_f \cdot kV^2}{MVA_{xfmr}} \qquad \textbf{Eq. 7-1}$$

where,

N = predicted variable: $\mathrm{SARFI}_{\mathrm{ITIC}}$ or MAIFI

l = total exposure (including three-phase and single-phase portions) on the circuit in miles

R = Reliability variable: SAIFI, SAIDI or Number of Outages

kV = base line-to-line voltage in kV

n_f = total number of feeders off the substation bus

MVA_{xfmr} = station transformer base rating (open-air rating) in MVA

Data from two utilities that participated in the DPQ II project was analyzed. These are not precise estimates, but they help to more accurately define the probability range at a given site. The dispersion term of the model can be used to find prediction limits from the Gamma distribution.

Table 7-1
Coefficients for Prediction Equations for Various SARFI and MAIFI Indices Using Data from Utility A and Utility B

	k_0	k_1	k_2	k_3
	Intercept	R	I	$\frac{n_f \cdot kV^2}{MVA_{xfmr}}$
$\text{SARFI}_{\text{ITIC}}$ (Utility A)	1.26	0.359	.001	0.0118
$\text{SARFI}_{\text{ITIC}}$ (Utility B)	3.007	0.106	.001	0.007
MAIFI (Utility B)	2.08	0.124	.001	0.000

Future Work

Additional work in the area of prediction of voltage sags is possible. Some areas ripe for exploration are:

Variability – One of the big concerns when deriving a prediction at a location is estimating the variability. The prediction models derived in this report have a variability term. This term includes time variability, site-to-site variability, and uncertainty in the modeling. It would help if we could derive a better estimate of the time variability from the data and from the predictions. That would help when quantifying the ranges of possibilities when considering things like performance-based contracts. One way to advance this work is to use bootstrapping (resampling) of the combined DPQ dataset to quantify the time variability of voltage sags.

Combining circuit modeling approaches with statistical regression – Circuit modeling approaches such as the fault-positions method have the advantage that we can use them to evaluate many system reconfigurations. One way to improve the quality of predictions from circuit modeling methods is to use statistical predictions for fault rates—the primary input into the circuit modeling method. A portion of the DPQ Phase II data has some very good fault-rate data as well as voltage-sag data. If we can perform the same type of regression based on various site characteristics, we could predict fault rates in the area where we want to predict voltage sags. Then we could use that prediction as the input to the fault-positions method.

User-accessible implementation – The basic prediction models derived in this study can be easily implemented in a spreadsheet, so they are accessible to a wide range of users. But some of the more advanced prediction options are difficult or impossible to implement in a spreadsheet. If a user has unknown data for some of the site parameters or wants to incorporate measurements, these are beyond simple spreadsheet implementation. A more user-accessible implementation of the more advanced predictions is possible. The best option for this is probably a Web-based user interface to a calculation engine running R or S-PLUS, where the codes for the advanced modeling are already implemented.

Since current measurements were not available from the two utilities, understanding if inherent differences exist in correlations (as well as the prediction models) between power quality and reliability measures for transmission-based faults and distribution-based faults, could not be investigated in this year's activity.

A

DISTRIBUTION RESULTS FOR UTILITY A

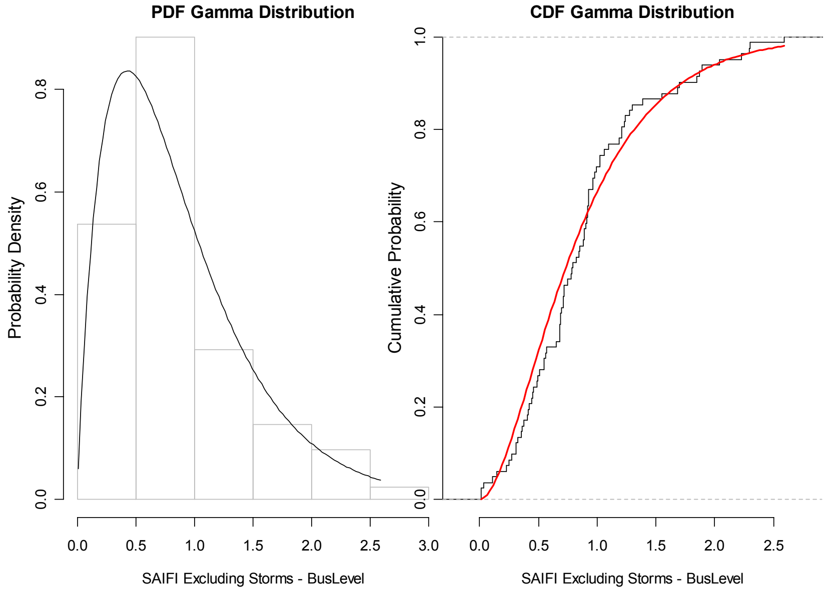


Figure A-1
SAIFI Gamma Distribution, Bus Level, Excluding Storms

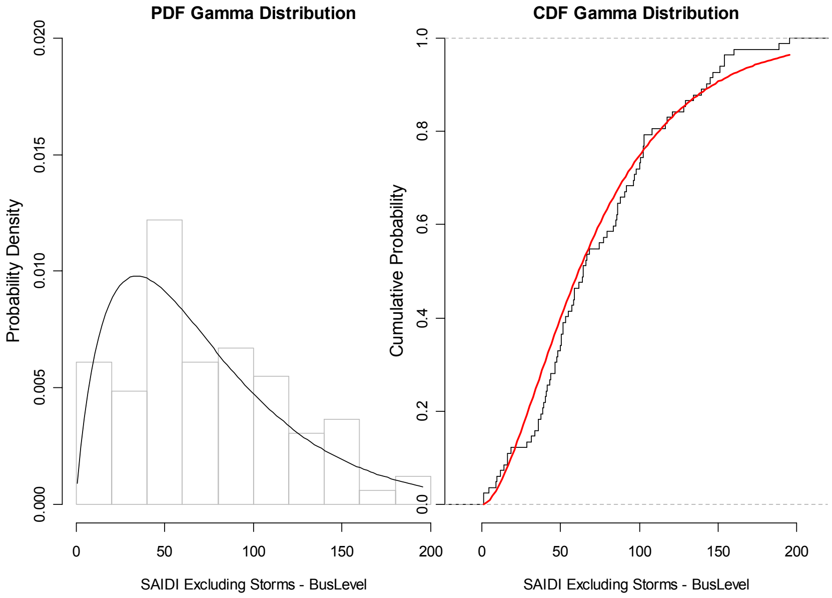


Figure A-2
SAIDI Gamma Distribution, Bus Level, Excluding Storms

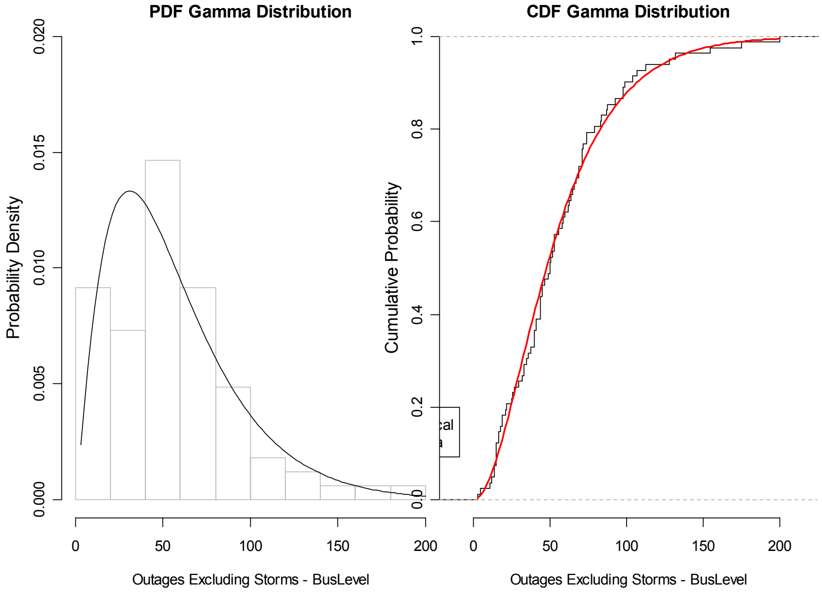


Figure A-3
Outages Gamma Distribution, Bus Level, Excluding Storms

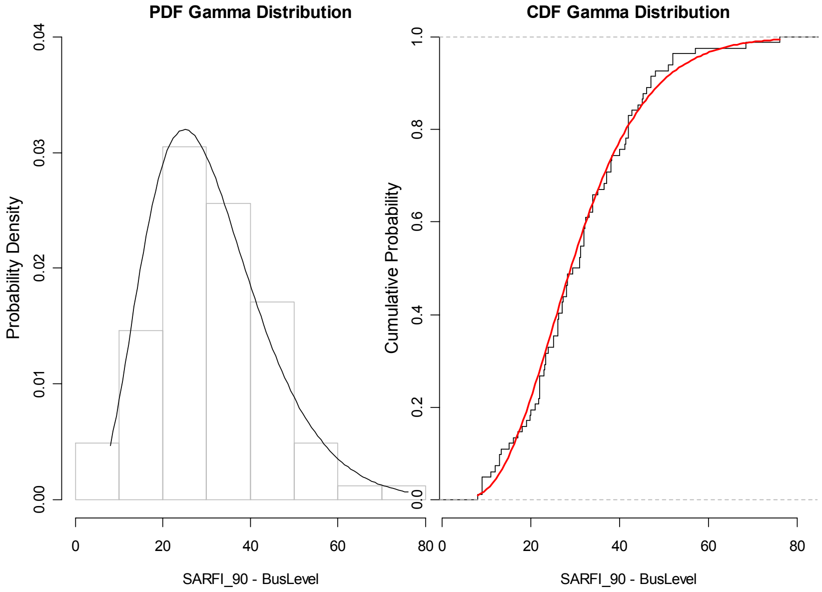


Figure A-4
SARFI_90 Gamma Distribution, Bus Level, Including Storms

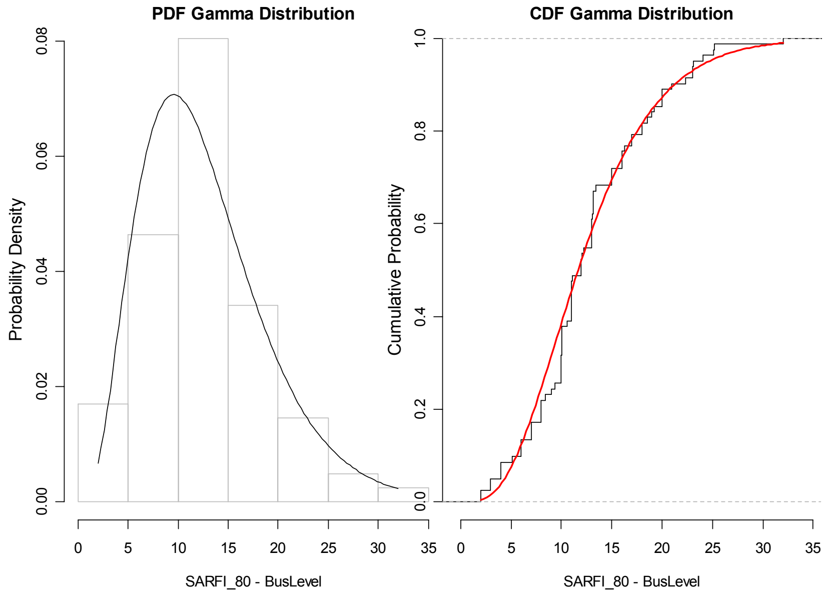


Figure A-5
SARFI_80 Gamma Distribution, Bus Level, Including Storms

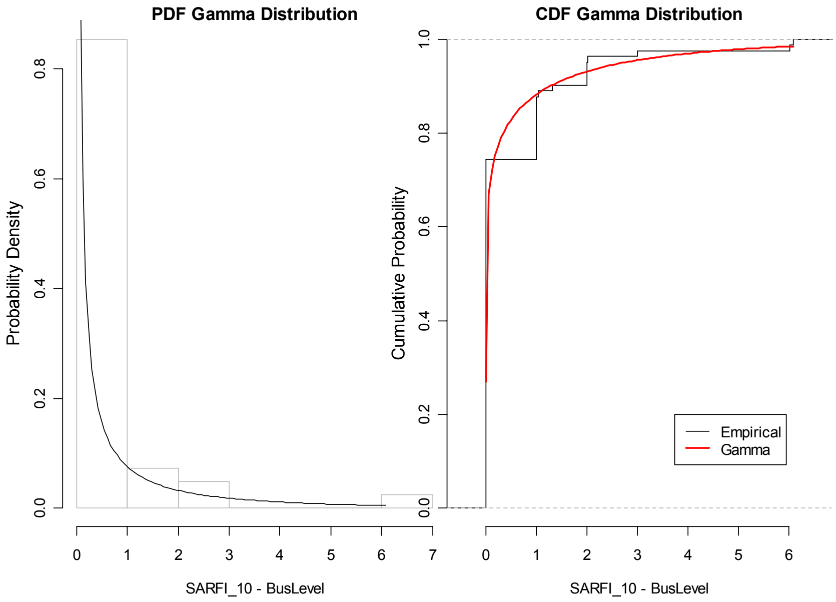


Figure A-6
SARFI_10 Gamma Distribution, Bus Level, Including Storm

B

DISTRIBUTION RESULTS FOR UTILITY B

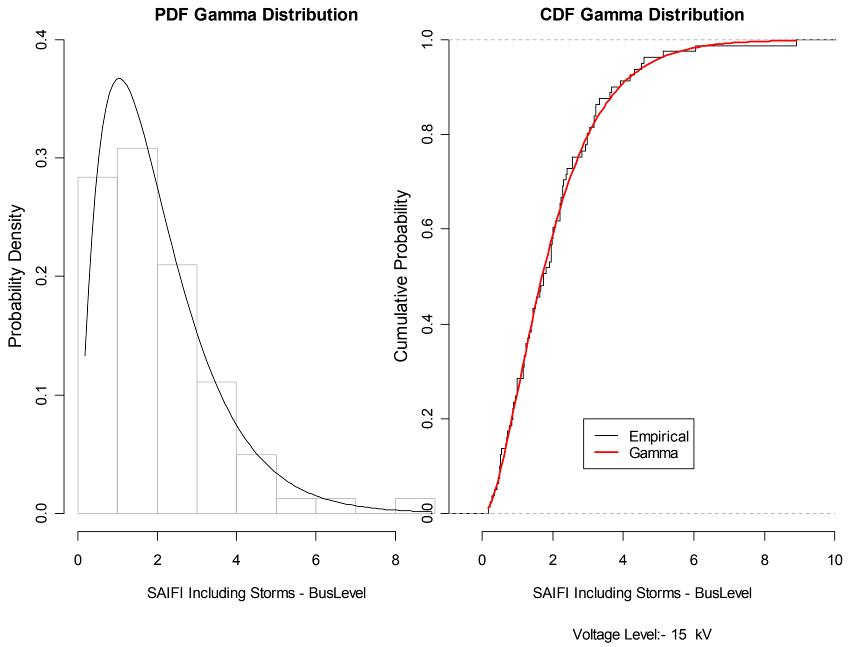


Figure B-1
SAIFI Gamma Distribution, Bus Level, Including Storm

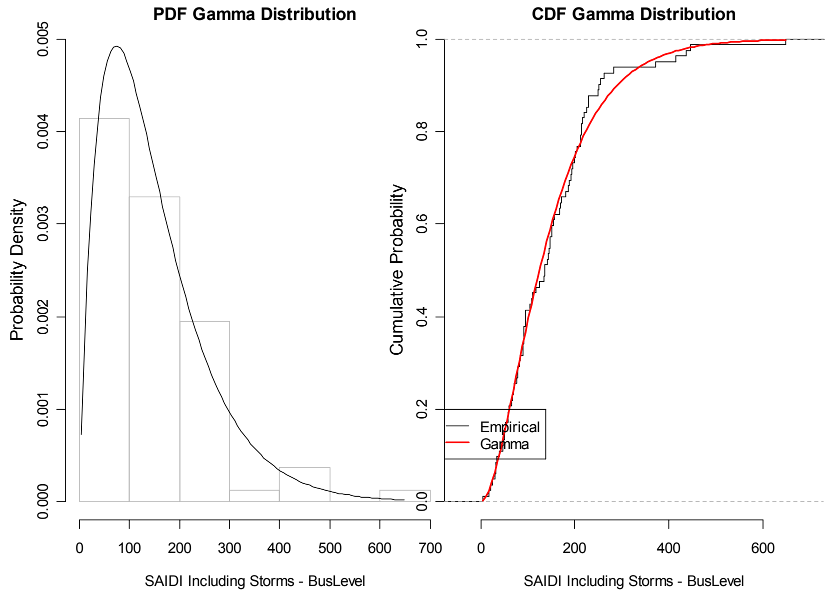


Figure B-2
SAIDI Gamma Distribution, Bus Level, Including Storm

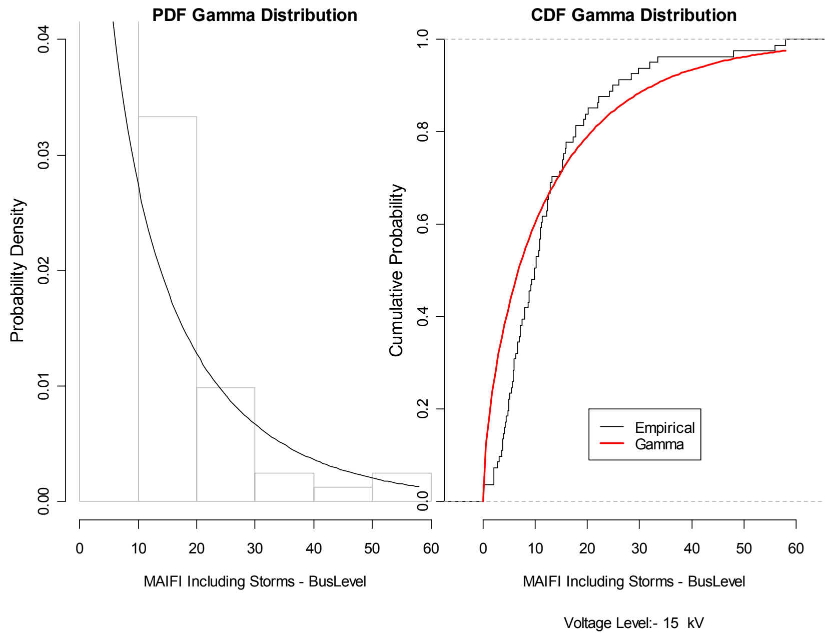


Figure B-3
MAIFI Gamma Distribution, Bus Level, Including Storm

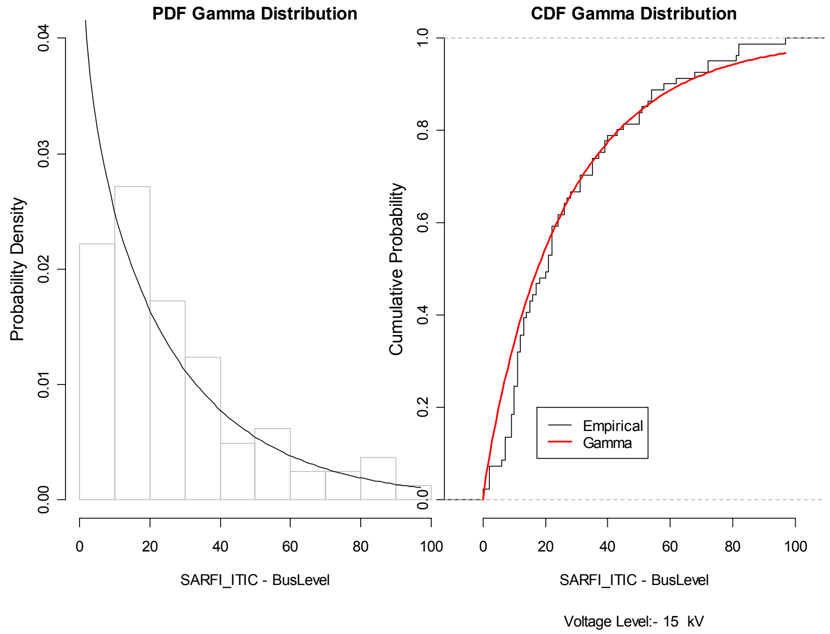


Figure B-4
SAIRFI_ITIC Gamma Distribution, Bus Level, Including Storm

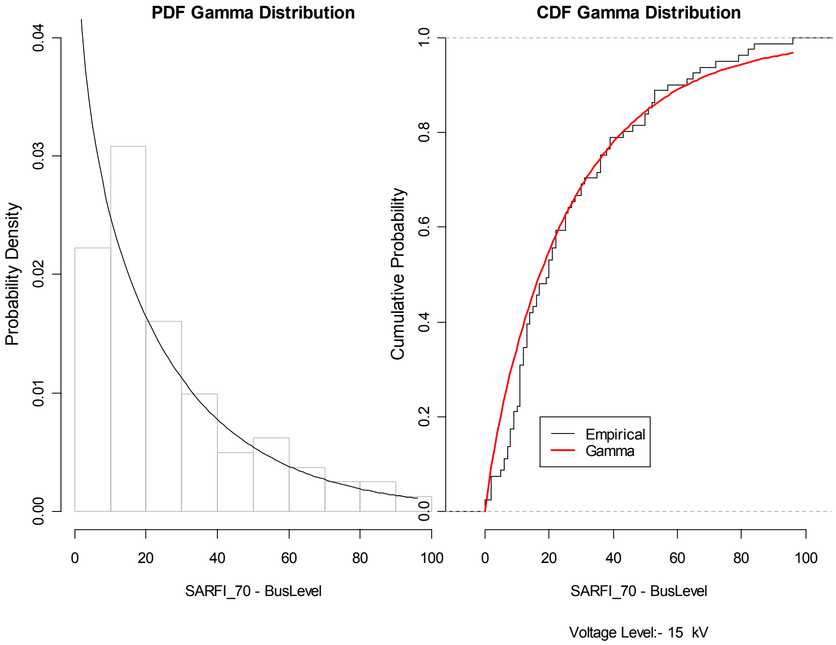


Figure B-5
SAIRFI_70 Gamma Distribution, Bus Level, Including Storm

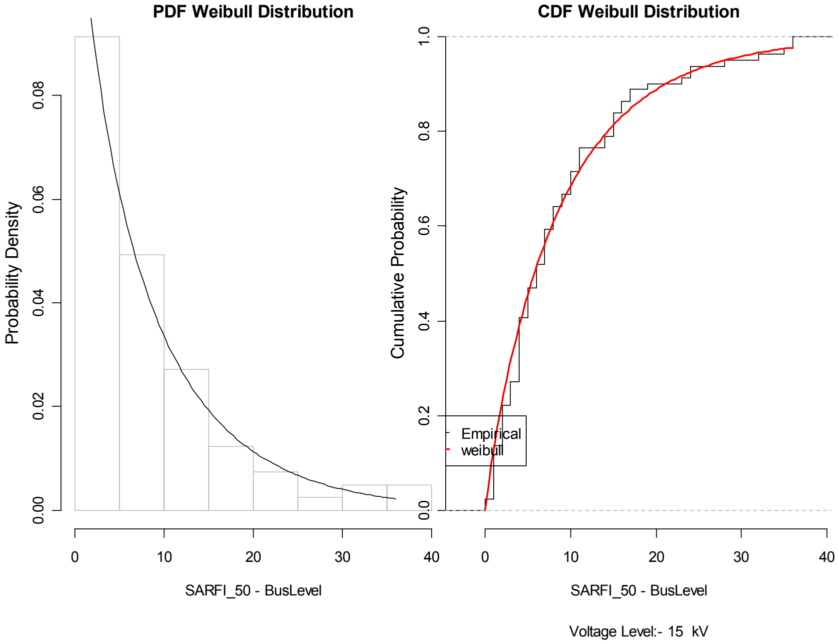


Figure B-6
SAIRFI_50Gamma Distribution, Bus Level, Including Storm

C

PROPERTIES OF BUS DATA FOR UTILITY A

SARFI$_{ITIC}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	6.000	8.000	8.847	11.780
Total Feeder Length					
	0-30%	4.000	7.000	6.748	8.000
	30%-70%	6.012	9.500	8.840	11.340
	70-100%	7.117	10.000	10.960	14.080
Backbone Length					
	0-30%	5.250	7.010	7.242	9.000
	30%-70%	6.000	9.000	8.798	12.000
	70-100%	6.500	9.100	10.730	15.040
Number of Feeders					
	$n_f < 5$	4.011	7.000	6.519	8.022
	$>5\ n_f <9$	6.797	10.000	9.933	12.010
	$n_f > 9$	6.500	10.000	10.100	13.070
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	6.000	7.667	8.721	11.020
	$>25\ Xfmr_{size} <36$	6.000	8.689	9.003	12.000
	$Xfmr_{size} > 36$	6.750	7.558	8.732	11.340
Load Density					
	0-30%	7.000	8.000	9.349	11.060
	30%-70%	6.000	9.000	9.117	12.000
	70-100%	6.000	7.078	8.000	10.000
SARFI$_{70}$ Bus Level					
	All Data	5.000	6.025	6.999	10.000
Total Feeder Length					
	0-30%	4.000	5.041	5.261	6.166
	30%-70%	5.010	7.500	7.258	10.000
	70-100%	5.000	8.000	8.406	12.000
Backbone Length					
	0-30%	5.000	5.578	5.786	6.903
	30%-70%	5.000	8.000	7.164	10.000
	70-100%	4.893	7.000	8.134	11.530
Number of Feeders					
	$n_f < 5$	3.064	5.041	5.160	6.116
	$>5\ n_f <9$	5.000	7.510	7.776	10.010
	$n_f > 9$	5.028	8.022	8.140	10.300
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	5.000	6.017	6.949	10.000
	$>25\ Xfmr_{size} <36$	4.601	6.041	7.050	10.010
	$Xfmr_{size} > 36$	4.770	6.500	6.973	9.325
Load Density					
	0-30%	5.000	7.000	7.437	10.000
	30%-70%	4.761	6.517	7.206	10.010
	70-100%	5.000	6.000	6.296	7.000

$SARFI_{50}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	1.000	2.234	2.944	5.000
Total Feeder Length					
	0-30%	0.000	2.000	1.938	3.000
	30%-70%	1.003	3.000	3.482	5.250
	70-100%	1.003	4.000	3.260	5.027
Backbone Length					
	0-30%	0.250	2.000	1.993	3.000
	30%-70%	1.000	3.008	3.487	5.247
	70-100%	1.003	3.000	3.239	5.014
Number of Feeders					
	$n_f < 5$	0.500	2.000	1.978	3.000
	$>5\ n_f < 9$	1.003	2.531	3.423	5.271
	$n_f > 9$	1.000	4.011	3.409	5.014
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	1.003	2.037	2.705	3.752
	$>25\ Xfmr_{size} < 36$	1.000	2.697	3.053	5.007
	$Xfmr_{size} > 36$	1.000	3.012	3.057	5.000
Load Density					
	0-30%	1.000	2.000	2.904	5.028
	30%-70%	1.002	3.000	3.322	5.007
	70-100%	1.000	2.645	2.499	4.000

SAIFI$_{including\_storms}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	0.697	1.082	1.159	1.472
Total Feeder Length					
	0-30%	0.437	0.860	0.999	1.355
	30%-70%	0.614	0.885	1.001	1.270
	70-100%	1.073	1.344	1.520	1.810
Backbone Length					
	0-30%	0.500	0.869	1.049	1.396
	30%-70%	0.573	0.923	0.986	1.263
	70-100%	1.095	1.344	1.533	1.858
Number of Feeders					
	n_f < 5	0.514	1.091	1.085	1.325
	>5 n_f <9	0.770	1.179	1.337	1.763
	n_f > 9	0.473	0.892	0.926	1.299
Station Transformer (MVA)					
	$Xfmr_{size}$ < 25	0.807	1.113	1.262	1.561
	>25 $Xfmr_{size}$ <36	0.809	1.093	1.146	1.431
	$Xfmr_{size}$ > 36	0.451	0.706	1.048	1.364
Load Density	0-30%	0.789	1.263	1.386	1.787
	30%-70%	0.839	1.056	1.161	1.355
	70-100%	0.415	0.562	0.929	1.302
SAIFI$_{excluding\_storms}$ Bus Level					
	All Data	0.4893	0.7865	0.8765	1.049
Total Feeder Length					
	0-30%	0.2526	0.6855	0.7232	0.9764
	30%-70%	0.4905	0.7174	0.8053	0.927
	70-100%	0.7174	0.9638	1.121	1.24
Backbone Length					
	0-30%	0.2908	0.6971	0.7627	1.123
	30%-70%	0.4911	0.7138	0.7867	0.9291
	70-100%	0.7547	0.9638	1.134	1.331
Number of Feeders					
	n_f < 5	0.329	0.8517	0.8249	1.107
	>5 n_f <9	0.6486	0.8869	1.001	1.125
	n_f > 9	0.4553	0.679	0.7138	0.8391
Station Transformer (MVA)					
	$Xfmr_{size}$ < 25	0.5479	0.9084	0.921	1.223
	>25 $Xfmr_{size}$ <36	0.6309	0.751	0.8916	1.067
	$Xfmr_{size}$ > 36	0.419	0.6223	0.7916	0.9347
Load Density	0-30%	0.6855	0.8906	0.9998	1.24
	30%-70%	0.671	0.8333	0.9139	1.022
	70-100%	0.3103	0.4887	0.7054	0.9764

SAIDI$_{including\_storms}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	72.31	135.8	149.2	200.3
Total Feeder Length					
	0-30%	34.93	81.58	129	185.9
	30%-70%	67.51	110.5	137.6	182.8
	70-100%	109.6	171.8	184.1	226
Backbone Length					
	0-30%	61.33	86.98	136.9	192.3
	30%-70%	60.78	109.6	135.3	170
	70-100%	112.5	171.8	183	219
Number of Feeders					
	$n_f < 5$	58.59	92.37	136	190.8
	>5 n_f <9	95.64	151.9	178.6	217.7
	$n_f > 9$	50.01	95.71	112	170.9
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	80.38	149.4	148.7	192.3
	>25 $Xfmr_{size}$ <36	84.14	129.4	151.4	193.1
	$Xfmr_{size} > 36$	51.06	92.58	145.6	221
Load Density	0-30%	79.49	181.2	192.3	229.4
	30%-70%	90.65	135.1	151	193.1
	70-100%	39.63	89.45	103.8	170
SAIDI$_{excluding\_storms}$ Bus Level					
	All Data	41.68	64.2	74.3	102.2
Total Feeder Length					
	0-30%	13.82	46.49	60.31	86.01
	30%-70%	46.44	58.76	73.85	94.5
	70-100%	61.5	91	88.86	102.7
Backbone Length					
	0-30%	15.97	54.1	65.85	112.1
	30%-70%	41.16	58.91	69.74	91.8
	70-100%	57.36	91	90.38	103
Number of Feeders					
	$n_f < 5$	16.09	58.91	70.55	111.9
	>5 n_f <9	52.82	85.48	85.82	104.5
	$n_f > 9$	37.98	49.87	57.78	70.45
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	46.83	60.21	76.41	102.9
	>25 $Xfmr_{size}$ <36	42.61	69.39	74.32	98.61
	$Xfmr_{size} > 36$	39.92	65.26	71.52	98.53
Load Density	0-30%	47.85	77.15	84.06	116.3
	30%-70%	50.51	79.02	78.93	97.59
	70-100%	18.38	41.16	58.6	86.01

Outage$_{including\_storms}$ **Bus Level**					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	37.25	62	71.2	91.5
Total Feeder Length					
	0-30%	19	25	26.6	29
	30%-70%	50.75	61	63.06	73.75
	70-100%	92	113	126.2	145
Backbone Length					
	0-30%	19.25	25	30.15	36.25
	30%-70%	49	61	62.21	76
	70-100%	96.5	114	130.5	145.5
Number of Feeders					
	$n_f < 5$	19.5	25	29.52	43
	$>5\ n_f < 9$	60	80	95.11	100
	$n_f > 9$	64	78	85.11	112.5
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	48	54	64.35	77.5
	$>25\ Xfmr_{size} < 36$	30.5	65	66.61	92.5
	$Xfmr_{size} > 36$	42.25	68	88.35	96.5
Load Density	0-30%	48	67	92	130
	30%-70%	51.5	71	75.28	92.5
	70-100%	20	28	45.16	65
Outage$_{excluding\_storms}$ **Bus Level**					
	All Data	30.5	50	56.18	71
Total Feeder Length					
	0-30%	15	18	21.56	26
	30%-70%	40.75	49.5	50.69	59.25
	70-100%	71	88	97.84	107
Backbone Length					
	0-30%	15	18.5	24.54	31.5
	30%-70%	41	49	49.61	59
	70-100%	76.5	93	101.4	110
Number of Feeders					
	$n_f < 5$	15	19	24.04	32.5
	$>5\ n_f < 9$	46.5	63	74.08	84.75
	$n_f > 9$	49	69	67.95	85.5
Station Transformer (MVA)					
	$Xfmr_{size} < 25$	33	44	50.73	66.5
	$>25\ Xfmr_{size} < 36$	26.75	53	53.17	74
	$Xfmr_{size} > 36$	32.25	55.5	68.7	72.5
Load Density	0-30%	36	52	71.32	99
	30%-70%	44	55.5	59.97	72.5
	70-100%	15	25	36.2	60

D

PROPERTIES OF BUS DATA FOR UTILITY B

$\text{SARFI}_{\text{ITIC}}$ Bus Level

Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	11.0	21.0	26.6	37.0
Total Feeder Length					
	0-30%	6.3	11.0	12.1	13.0
	30%-70%	12.0	22.0	27.4	35.0
	70-100%	21.3	42.5	42.6	60.0
Percent Underground					
	0-30%	9.3	14.0	18.9	23.5
	30%-70%	10.0	17.0	28.0	43.0
	70-100%	15.0	26.0	33.7	47.0
Number of Feeders					
	$n_f < 5$	6.0	9.0	8.0	11.0
	>5 n_f <9	18.0	26.0	33.7	43.0
	$n_f > 9$	52.0	53.0	57.3	60.5
Load Density					
	0-30%	7.0	10.0	15.0	22.0
	30%-70%	12.0	20.5	32.0	52.0
	70-100%	13.0	33.0	31.6	40.0

$\text{SAIFI}_{\text{including\_storms}}$ Bus Level

Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	0.986	1.724	2.011	2.563
Total Feeder Length					
	0-30%	0.534	1.265	1.700	1.992
	30%-70%	0.881	1.724	1.741	2.279
	70-100%	1.410	2.456	2.784	3.565
Percent Underground					
	0-30%	0.878	1.288	2.004	2.755
	30%-70%	1.282	1.820	2.123	2.290
	70-100%	0.736	1.957	1.852	2.515
Number of Feeders					
	$n_f < 5$	0.724	1.248	1.679	1.964
	>5 n_f <9	1.187	1.944	2.109	2.826
	$n_f > 9$	2.620	3.229	3.050	3.569
Load Density	0-30%	1.155	1.444	2.271	2.973
	30%-70%	0.925	2.013	2.015	2.654
	70-100%	0.626	1.861	1.735	2.331

$\text{SAIDI}_{\text{including storms}}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	80.4	194.3	546.6	587.0
Total Feeder Length					
	0-30%	36.8	96.3	576.2	598.3
	30%-70%	80.4	170.4	356.3	532.2
	70-100%	165.3	328.8	797.0	615.2
Percent Underground					
	0-30%	68.31	125.5	747.1	342.6
	30%-70%	104.5	229.1	524.7	620.9
	70-100%	44.33	239.6	342.3	579.9
Number of Feeders					
	$n_f < 5$	66.8	140.8	587.2	493.0
	$>5\ n_f < 9$	86.1	207.4	540.3	598.0
	$n_f > 9$	263.9	322.0	318.7	375.2
Load Density	0-30%	101.0	271.8	981.8	1020.0
	30%-70%	88.4	200.8	405.5	603.8
	70-100%	38.2	159.1	281.3	409.4

$\text{MAIFI}_{\text{including storms}}$ Bus Level					
Site Parameters	Category	P (25%)	Median	Mean	P (75%)
	All Data	5.65	9.81	12.55	15.39
Total Feeder Length					
	0-30%	3.64	5.63	11.52	10.64
	30%-70%	7.46	10.68	12.01	12.82
	70-100%	8.39	13.93	14.56	18.88
Percent Underground					
	0-30%	2.91	6.35	8.83	11.01
	30%-70%	7.07	10.68	14.63	15.39
	70-100%	5.96	11.86	13.82	18.88
Number of Feeders					
	$n_f < 5$	3.00	6.00	12.01	11.28
	$>5\ n_f < 9$	6.69	10.23	12.01	15.68
	$n_f > 9$	25.11	26.07	26.65	27.91
Load Density	0-30%	3.00	6.00	12.01	11.28
	30%-70%	6.69	10.23	12.01	15.68
	70-100%	25.11	26.07	26.65	27.91

Export Control Restrictions
Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is our obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case by case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI)

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, CA, and Charlotte, NC, was established in 1973 as an independent, non-profit center for public interest energy and environmental research. EPRI brings together member organizations, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of power generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents over 10% of EPRI's total R&D program.

Together…Shaping the Future of Electricity

Program: 1008508

Power Quality Solutions for Transmission and Distribution

© 2005 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc. TOGETHER...SHAPING THE FUTURE OF ELECTRICITY is a service mark of the Electric Power Research Institute, Inc.

Printed on recycled paper in the United States of America