Attachment B

LBNL-5268E

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

An Examination of Temporal Trends in Electricity Reliability Based on Reports from U.S. Electric Utilities

Joseph H. Eto, Kristina Hamachi LaCommare, Peter Larsen,
Annika Todd, and Emily Fisher

January 2012

The work described in this report was funded by the Office of Electricity Delivery and Energy Reliability of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Disclaimer

An Examination of Temporal Trends in Electricity Reliability Based on Reports from U.S. Electric Utilities

Joseph H. Eto, Kristina Hamachi LaCommare,
Peter Larsen, Annika Todd, and Emily Fisher

Ernest Orlando Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90-4000
Berkeley CA 94720-8136

The work described in this report was funded by the Office of Electricity Delivery and Energy Reliability of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Abstract

Since the 1960s, the U.S. electric power system has experienced a major blackout about once every 10 years. Each has been a vivid reminder of the importance society places on the continuous availability of electricity and has led to calls for changes to enhance reliability. At the root of these calls are judgments about what reliability is worth and how much should be paid to ensure it.

In principle, comprehensive information on the actual reliability of the electric power system and on how proposed changes would affect reliability ought to help inform these judgments. Yet, comprehensive, national-scale information on the reliability of the U.S. electric power system is lacking.

This report helps to address this information gap by assessing trends in U.S. electricity reliability based on information reported by electric utilities on power interruptions experienced by their customers. Our research augments prior investigations, which focused only on power interruptions originating in the bulk power system, by considering interruptions originating both from the bulk power system and from within local distribution systems. Our research also accounts for differences among utility reliability reporting practices by employing statistical techniques that remove the influence of these differences on the trends that we identify.

The research analyzes up to 10 years of electricity reliability information collected from 155 U.S. electric utilities, which together account for roughly 50% of total U.S. electricity sales. The questions analyzed include:

1. Are there trends in reported electricity reliability over time?
2. How are trends in reported electricity reliability affected by the installation or upgrade of an automated outage management system?
3. How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003?

Acknowledgments

The work described in this report was funded by the Office of Electricity Delivery and Energy Reliability (OE) of the U.S. Department of Energy (DOE) under Contract No. DE-AC02-05CH11231. The authors are grateful to Joe Paladino for his support of this research.

We thank the staff at the state regulatory agencies and individual utilities who provided the reliability metric information analyzed in this report: Don Lamontagne and Jeffrey Smith (Arizona Public Service), Salt River Project (Arizona), Clark Cotton and Lynn Morgan (Arkansas Public Service Commission), Stephen Brown (Colorado Public Utilities Commission), Beverly Barker (Idaho Public Utilities Commission), Harry Stoller and Roy Buxton (Illinois Commerce Commission), Jim Sundermeyer (Iowa Utilities Board), Brian McManus (Louisiana Public Service Commission), Derek Davidson (Maine Public Utilities Commission), Richard Miller (Maryland Public Service Commission), Donald Nelson (Massachusetts Department of Public Utilities), Eric Dahlgren (Montana Public Service Commission), NV Energy (Nevada), Steve Mullen (New Hampshire Public Utilities Commission), Nanik Aswani (New Jersey Board of Public Utilities), Jack Sidler (New Mexico Public Regulatory Commission), Howard Lowdermilk (North Carolina Utilities Commission), Jerry Lein and Cara DeSaye (North Dakota Public Service Commission), Jason Cross, John Williams and Arla Cahill (Ohio Public Utilities Commission), Darren Gill (Pennsylvania Public Utilities Commission), Al Contente (Rhode Island Public Utilities Commission), Donald Neumeyer (Wisconsin Public Service Commission), and Joshua Jones (Wyoming Public Service Commission).

We also thank members of the IEEE Distribution Reliability Working Group for their thoughtful comments on various drafts of this report, including their generous sharing of experiences in analyzing similar data: James Bouford, Heide Caswell, Jim Cole, Jane Hammes, Robert Jones, Mark Konya, Don Lamontagne, David Lankutis, Thomas Menten, William Ranken, Rodney Robinson, and Val Werner.

And finally we thank the following reviewers for their helpful suggestions to improve the presentation of our findings and the discussion of their significance: Seth Blumsack, Peter Cappers, Cha Chi Fan, Meredith Fowlie, Alex Hoffman, Nathan Mitchell, Dave Mohre, Michael Perry, Marie Rinkoski-Spangler, Josh Schellenberg, Richard Schmalensee, and Kim Wissman.

All errors and omissions are the sole responsibility of the authors.

Table of Contents

Abstract i

Acknowledgments iii

Table of Contents v

List of Figures and Tables vii

Acronyms and Abbreviations ix

Executive Summary xi

1. Introduction 1

2. Data Collection and Review 5
   - 2.1 Sources of Data 5
     - 2.1.1 Utility-reported reliability data 5
     - 2.1.2 Temperature-related weather data 6
     - 2.1.3 Retail electricity sales 6
   - 2.2 Review of Utility-Reported Reliability Data 6
     - 2.2.1 Geographic representation and coverage 6
     - 2.2.2 Completeness of reported reliability data by utility and over time 8
     - 2.2.3 Distributions of reported SAIDI and SAIFI over time 9
     - 2.2.4 Automated outage management systems (OMS) 11
     - 2.2.5 Major events 12

3. An Initial Review of Time Trends in Reported Electricity Reliability 17
   - 3.1 Time Trends in Reported Electricity Reliability Based on Descriptive Statistics 17
   - 3.2 Discussion of Time Trends Based on Descriptive Statistics 23

4. Findings from the Statistical Analysis of Reliability Data Reported by Electric Utilities 25
   - 4.1 Introduction to the Statistical Methods Used in the Analysis 25
   - 4.2 Application of the Statistical Models 26
   - 4.3 Findings 28
     - 4.3.1 Are there utility-specific differences in reported electricity reliability? 28
     - 4.3.2 Are there trends in reported electricity reliability over time? 29
     - 4.3.3 How are trends in reported electricity reliability affected by the installation or upgrade of the automated OMS? 31
     - 4.3.4 How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003? 33

5. Summary and Interpretation of Findings, and Next Steps 35

References 39

Appendix A. Customer Weighted Average Comparison to IEEE DRWG Benchmarking Analysis41

Appendix B. Why a Log-Normal Distribution? 43

Appendix C. Examination of Outliers 47

Appendix D. Detailed Results from Regression Analysis 49

List of Figures and Tables

Figure 1. Map of U.S. by NERC Region 7
Figure 2. Number of Utilities with SAIDI and SAIFI Data 8
Figure 3. Completeness of Time Series 9
Figure 4. Box-Plot of SAIDI by Year with and without Major Events 10
Figure 5. Box-Plot of SAIFI by Year with and without Major Events 10
Figure 6. Number of Utilities by Year and NERC Region that Installed or Upgraded their OMS 12
Figure 7. Percentage Difference in SAIDI Without Major Events Included between Using and Not Using IEEE Standard 1366-2003 14
Figure 8. Percentage Difference in SAIFI Without Major Events Included between Using and Not Using IEEE Standard 1366-2003 15
Figure 9. SAIDI with Major Events – All Reported Reliability Data 18
Figure 10. SAIDI without Major Events – All Reported Reliability Data 18
Figure 11. SAIFI with Major Events – All Reported Reliability Data 19
Figure 12. SAIFI without Major Events – All Reported Reliability Data 19
Figure 13. SAIDI with Major Events – Reported Reliability Data from Same Utilities for Every Year – N=28 20
Figure 14. SAIDI without Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67 20
Figure 15. SAIFI with Major Events – Reported Reliability Data from Same Utilities for Every Year – N=28 21
Figure 16. SAIFI without Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67 21

Figure A- 1. Customer-weighted SAIDI w/o Major Events – Comparison of LBNL Findings to those of the IEEE DRWG 2011 Benchmarking Survey 42
Figure A- 2. Customer-weighted SAIFI w/o Major Events – Comparison of LBNL Findings to those of the IEEE DRWG 2011 Benchmarking Survey 42

Figure B- 1. Graphical analysis of SAIDI with and without major events included. 45
Figure B- 2. Graphical analysis of SAIFI with and without major events included 46

Table 1. 2009 Sales of Utilities for which Data were Collected, by NERC Region 7
Table 2. Summary of Utilities with an OMS 11
Table 3. Summary of Utilities Using IEEE 1366-2003 13
Table 4. Summary of Numerical Best Fit of Trends in SAIDI 22
Table 5. Summary of Numerical Best Fit of Trends in SAIFI 22
Table 6. F-test of the Hypothesis that there are No Utility-Specific Effects 28
Table 7. One-way Random Effects Regression (Model 1): The Effect of Sales, HDD, CDD, Time, and OMS on Frequency and Duration of Interruptions (with Major Events Included) 30
Table 8. One-way Random Effects Regression (Model 1): The Effect of Sales, HDD, CDD, Time, OMS, and OMS “Learning” on Frequency and Duration of Interruptions (Without Major Events Included) 32

Table 9. No Utility Fixed Effects Regression (Model 2): Effect of IEEE, sales, HDD, CDD, Time, and OMS on Frequency and Duration of Interruptions (*without major events*)... 33

Table B- 1. Statistical Tests for Normality ... 44
Table B- 2. Statistical Tests for Log-normality ... 44

Table C- 1. Summary Explanation of Identified Outliers ... 47
Table C- 2. Excluding Outliers and their Effect on the Pooled Regression Results ... 48

Table D- 1. One-way Fixed Effects Regression for the Effect of Sales, HDD, CDD, time, and OMS on Frequency and Duration of Grid Disruptions ... 49
Table D- 2. Two-way Fixed Effects Regression for the Effect of Sales, HDD, CDD, and OMS on Frequency and Duration of Grid Disruptions ... 50

Acronyms and Abbreviations

ASCC	Alaska Systems Coordinating Council
CAIDI	Customer Average Interruption Duration Index
CDD	cooling degree-day
EIA	Energy Information Administration
FRCC	Florida Reliability Coordinating Council
HDD	heating degree-day
HICC	Hawaiian Islands Coordinating Council
IEEE	Institute of Electrical and Electronics Engineers
MED	major event day
MRO	Midwest Reliability Organization
MWh	megawatt (10^6 watts) hour
NCDC	National Climatic Data Center
NERC	North American Electric Reliability Corporation
NOAA	National Oceanic Atmospheric Administration
NPCC	Northeast Power Coordinating Council
OMS	outage management system
RFC	Reliability First Corporation
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SERC	Southeast Electric Reliability Council
SPP	Southwest Power Pool
TRE	Texas Regional Entity
WECC	Western Electricity Coordinating Council

Executive Summary

Since the 1960s, the U.S. electric power system has experienced a major blackout about once every 10 years. Each has been a vivid reminder of the importance society places on the continuous availability of electricity and has led to calls for changes to enhance reliability. At the root of these calls are judgments about what reliability is worth and how much should be paid to ensure it.

The goal of this study is to inform discussions of the reliability of the U.S. electric power system by assessing trends in power interruptions experienced by U.S. electricity consumers. Our analysis is based on up to 10 years of electricity reported reliability information collected from a convenience sample of 155 U.S. electric utilities, which together account for roughly 50% of total U.S. electricity sales.

We built on prior investigations, which focused only on power interruptions originating in the bulk power system, by considering interruptions originating both from the bulk power system and from within local distribution systems. We also accounted for differences among utility practices for collecting information on and reporting power interruptions by employing statistical techniques that remove the influence of these differences on the trends we identify.

We sought to answer three questions:

1. Are there trends in reported electricity reliability over time?

We first conducted an examination relying on descriptive statistics (mean, median, customer-weighted mean) and find that reported reliability has been decreasing over time. With minor exceptions, we observed this trend for all three descriptive statistics when considering all utility reports taken together, as well as only those utility reports for which we had a complete record of 10 years of data. We point out that descriptive statistics alone mask the effects of utility-specific effects that may introduce bias into our findings.

Next, we applied rigorous statistical methods both to confirm that there were utility-specific differences among electricity reliability reports and to take explicit account of these differences in exploring correlations between reported reliability metrics and other factors. Applying these methods, we find that there are statistically significant temporal trends. We find that reported average duration and average frequency of power interruptions has been increasing over time at a rate of approximately 2% annually. In other words, reported reliability is getting worse.

While our findings are highly statistically significant, it is important to place them in appropriate context. The average annual trends we find are modest in comparison to the routinely larger year to year variations in the average duration and frequency of power interruptions experienced by utility customers. For example, in Appendix A, we present a simple analysis of trends over the most recent four years and find reported reliability has been improving over this period.

In addition, we make no claims regarding the applicability of our findings to the reliability of the U.S. electric power system as a whole. Strictly speaking, our findings apply only to the convenience sample of primarily investor-owned utilities for which we were able to collect reported reliability information. In any given year, these utilities represented roughly 50% of total U.S. electricity sales.

2. How are trends in reported electricity reliability affected by the installation or upgrade of an automated outage management system (OMS)?

A principal contribution of our work has been to examine potential sources of measurement error that could influence apparent trends in reported reliability. We find statistically significant evidence that installation or upgrade of an OMS is correlated an increase in the reported duration of power interruptions. This finding confirms anecdotal evidence long been known within the industry that reliance on prior (manual) measurement methods under-reports reliability. We also found preliminary but not statistically significant evidence for a so-called "learning effect" by which reported reliability gradually improves in years subsequent to the initial decrease in reported reliability.

Our findings might suggest that it is simply more accurate measurement of reliability, rather than lower actual reliability, which "explains" the statistically significant trend of decreasing reported reliability over time. However, our analysis takes this factor into account explicitly and still finds statistically significant secular trends toward lower reported reliability over time. Our findings, therefore, highlight the importance of taking into account the means by which reliability information is collected when examining trends in reported reliability.

3. How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003?

We also examined a potential source of measurement bias in the form of utility reporting practices. We find that reliance on IEEE Standard 1366-2003 is correlated with higher reported reliability on average compared to reported reliability not using the IEEE standard and that this correlation is statistically significant. Nevertheless, taking this correlation into account, the secular trend of decreasing reported reliability over time remains statistically significant and at approximately the same magnitude as was found earlier (i.e., decreasing at roughly 2% annually). We caution that it is premature to attribute reliance on the IEEE standard as "causing" higher reported reliability because we could not separate the effect of reliance on the IEEE standard from other utility-specific factors (which we did not account for separately) that may also be correlated with reliance on the IEEE standard.

Next Steps

This study finds that there has been a modest, yet statistically significant secular trend of decreasing or declining reported reliability over the past 10 years. In making this finding, we summarize what our analysis to date has and has not accomplished, and outline the directions for next steps in this line of inquiry.

We wish to state clearly that, at this point, we cannot say what has caused the observed decreasing trends in reported reliability or why it is taking place. Our work has considered and characterized the influence of potential sources of measurement error or bias and found that taking these considerations into account changes neither the direction of these trends nor their statistical significance. These findings are important because they allow us to focus on potential causal factors that would help us explain the trends we observe.

To begin this process, we considered potential correlations with highly aggregated measures of weather variability and a simple measure of utility size but found neither to be statistically significant. We believe it is extremely appropriate to continue exploring differences among utilities to better understand the sources or causes of the secular trends in reliability that we observe. Some of the factors we believe should be considered include more disaggregate measures of weather variability (e.g., lightning strikes and severe storms), utility characteristics (e.g., the number of rural versus urban customers, and the extent to which transmission and distribution lines are overhead versus underground), and utility spending on transmission and distribution maintenance and upgrades, including advanced ("smart grid") technologies. It is our hope that the analysis we have conducted to date will help pave the way for these investigations and that they will be used to help ground future decisions about U.S. reliability policy, practices, and technology on a more solid factual base.

1. Introduction

Since the 1960s, the U.S. electric power system has experienced a major blackout about once every 10 years. Each has been a vivid reminder of the importance society places on the continuous availability of electricity and has led to calls for changes to enhance reliability. At the root of these calls are judgments about what reliability is worth and how much should be paid to ensure it.

In principle, information on the actual reliability of the electric power system and how proposed changes would affect reliability ought to help inform these judgments. Use of this type of information in local decision making, for example between an investor-owned utility and its state public utilities commission, is common. Yet, comprehensive, national-scale information on the reliability of the U.S. electric power system is lacking.

This report helps to address this information gap by assessing trends in U.S. electricity reliability based on information reported by electric utilities on power interruptions experienced by their customers. Our research augments prior investigations, which focused only on power interruptions originating in the bulk power system, by considering interruptions originating both from the bulk power and from within local distribution systems. Our research also accounts for differences among utility reliability reporting practices by employing statistical techniques that remove the influence of these differences on the trends that we identify.

The focus of prior published investigations of U.S. electric power system reliability has been primarily on the reliability of the bulk power system. For example, Amin (2008) suggests that the reliability of the bulk power system has been declining over time based on a review of the frequency and size of reported events. The response by Hines et al. (2009) rejects that hypothesis based on a rigorous statistical examination of the same data.[1]

At the same time, interruptions originating on the bulk power system represent only a small fraction of the power interruptions experienced by electricity consumers, as indicated in Hines et al. (2009) and Eto and LaCommare (2008). The vast majority of interruptions experienced by electricity consumers are caused by events affecting primarily the electric distribution system. Thus, analyses of power interruptions originating in the bulk power system alone address only a small portion of electricity consumers' total reliability experience.

Utilities routinely collect information on reliability of electric service provided to their customers. This information almost always includes all power interruptions experienced by their customers, both those originating in the bulk power system and those originating from within the electricity distribution system. The main metrics that utilities use to report this information focus separately on the frequency and the duration of power interruptions. (See text box.)

[1] Others observe that the data on bulk power system reliability relied on by studies such as these are sometimes inconsistent, incomplete and inaccurate (Fisher, et al. 2012).

Defining SAIDI and SAIFI

The System Average Interruption Duration Index (SAIDI) and the System Average Interruption Frequency Index (SAIFI) are metrics for the average duration and average number, respectively, of sustained power interruptions experienced by the population of customers served by a utility over the course of a year. SAIDI and SAIFI are two of the most commonly used metrics by utilities and industry experts when reporting on the continuity of electricity service to customers.

According to IEEE Standard 1366-2003, the metrics are defined as follows:

$$\text{SAIDI} = \frac{\sum \text{Customer Interruption Durations}}{\text{Total Number of Customers Served}} \quad \text{minutes of interruption per year}$$

$$\text{SAIFI} = \frac{\sum \text{Total Number of Customers Interrupted}}{\text{Total Number of Customers Served}} \quad \text{interruptions per year}$$

Larger values of SAIDI and SAIFI indicate less reliable electricity service meaning that customers, on average, experience longer or more frequent interruptions. In this report, we express this relationship by describing higher or increasing reported values of SAIDI or SAIFI as an indicator of lower or declining reported reliability.

Previous work examining electric utility practices for reporting reliability information revealed significant variation (Eto and LaCommare 2008). Despite the existence of standards - albeit voluntary ones - promulgated by the industry's professional society, the Institute for Electrical and Electronics Engineers (IEEE), differences in utilities' definition and classification of power interruption events make direct comparisons among data from different utilities problematic and potentially misleading.

In this paper, we analyze up to 10 years of reported electricity reliability information collected from a convenience sample of 155 U.S. electric utilities, which together account for roughly 50% of total U.S. electricity sales. Using these data sources, we quantify trends in electricity reliability and examine the relationship between these trends to the characteristics of the utilities, the climates in which their customers reside, utility reporting practices, and the adoption of advanced technologies for recording power interruptions. Our analysis uses statistical techniques that take into account differences in reliability reporting practices and other factors among electricity utilities, so that we can explore the effect of these differences.

The questions we examined and the motivations for examining them are as follows:

1. Are there trends in reported electricity reliability over time?

As noted above, Hines et al. (2009) concluded that there are no statistically significant trends over time based on a rigorous statistical examination of data on the reliability of the bulk power system. Taking explicit account of specific differences in utility reporting practices (and other factors) and using comprehensive information all power interruptions experienced by consumers, our analysis seeks to determine whether statistically significant temporal trends can be identified.

2. How are trends in reported electricity reliability affected by the installation or upgrade of an automated outage management system (OMS)?

McGrananghan (2006) speculated that adoption of OMS led one utility to report lower reliability because of under-reporting of customer power interruptions prior to adoption of the OMS. Our analysis explores the effect of installing or upgrading an OMS and how any such advanced reporting system is correlated with changes reported reliability over time.

3. How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003 (IEEE 2004)?

Eto and LaCommare (2008) compared reliability metrics reported by a convenience sample of 11 electric utilities using both historic company practices and IEEE Standard 1366-2003. Based on this small sample, those authors find no evidence of systematic bias resulting from use of the IEEE standard. The current analysis seeks to update the 2008 findings based on a larger sample of older and more recent data.

The remainder of this paper is organized as follows:

In Section 2, we describe the information we collected to conduct the analysis, including the electricity reliability metrics, the size of the utilities reporting the metrics, the weather experienced their customers, the adoption of automated technologies for recording power interruptions, and the practices for reporting power interruptions.

In Section 3, we present findings from our preliminary investigation of time trends in reported reliability based on means, medians, and customer-weighted means.

In Section 4, we describe and present findings from application of more advanced statistical methods to the reported reliability metrics, which take into account utility-specific differences that might influence time trends in reported reliability. The utility-specific differences include the size of utility, the weather their customers experienced, installation or adoption of an automated outage management system, and utility reporting practices vis-à-vis IEEE Standard 1366-2003.

In Section 5, we summarize our main findings and discuss next steps.

Four technical appendices follow. Appendix A compares a variant of the analysis of customer-weighted means presented in Section 3, which enables a direct comparison to a closely related analysis conducted by the IEEE Distribution Reliability Working Group. Appendix B presents the results from analyses we conducted to better understand the effect of a mathematical transformation of the dependent variables examined in Section 4 prior to conducting the regression analysis. Appendix C examines the statistical outliers identified in our statistical analysis and their impact on our findings. Appendix D provides information on the results from an alternative specification of the statistical model that is the basis for findings presented in Section 4.

2. Data Collection and Review

The data we collected for this study include:

- Utility-reported reliability metrics, focusing on the System Average Interruption Frequency Index (SAIFI) and the System Average Interruption Duration Index (SAIDI),
- Installation or upgrades of automated outage management systems (OMS),
- Adoption of IEEE Standard 1366-2003 for reporting reliability metrics,
- Temperature-related weather, and
- Retail electricity sales.

This section describes the sources for these data and reviews selected aspects of the data we collected on reliability metrics.

2.1 Sources of Data

2.1.1 Utility-reported reliability data

Our primary source for utility-reported reliability data was state utility regulatory commissions because investor-owned utilities routinely file these data with their commissions and these data are often made publicly available (Eto and LaCommare 2008). We contacted all the commissions that made these data publicly available. As a result of this approach, the sample of utilities for which we obtained reported reliability data are largely investor-owned utilities.

In addition, we also collected some data directly from individual utilities that we had identified through previous research. No formal statistical sampling procedures were employed in determining which utilities were contacted.

Two reliability metrics, SAIDI and SAIFI, were collected for the years 2000 to 2009. We requested SAIDI and SAIFI both with and without the inclusion of major events.[2] See Section 2.2.5 for a discussion of major events and the reason why utilities sometimes report reliability metrics both including and not including these events.

We also collected information on whether and in what year a utility installed or upgraded an automated OMS. An OMS provides an automatic and consistent means for collecting information on the frequency, extent, and duration of electric service interruptions. This automation technology generally replaces manual record keeping, which is widely recognized as a less reliable means of collecting service interruption information (LET Systems 2006).

Finally, we also collected information on whether the utility relied on IEEE Standard 1366-2003 in reporting its reliability metrics. Among other things, the IEEE standard features a heuristically derived, yet systematic and statistically based method for reporting SAIDI and SAIFI without major events.

---

[2] In some instances when SAIDI was not reported, the Customer Average Interruption Duration Index (CAIDI) was collected to derive SAIDI using the simple mathematical expression CAIDI = SAIDI/SAIFI.

2.1.2 Temperature-related weather data

We collected information on weather in the form of annual heating and cooling degree-days (HDD and CDD) for 2000 to 2009 from the National Oceanic & Atmospheric Administration (NOAA) National Climatic Data Center (NCDC) (NCDC 2011).[3] HDD and CDD are measures of the need for heating or cooling in a building. Thus, HDD is positive if ambient air temperature is cool and a building needs to be heated; CDD is positive if ambient air temperature is warm and a building needs to be cooled. HDD is defined as 65 minus the average of the daily high and low temperature where HDD is set to 0 if the average daily temperature is more than 65^{o} F. CDD is defined as the average of the daily high and low temperature minus 65 where the CDD is set to 0 if the average daily temperature observed is less than 65^{o} F. We assigned state-level HDD's and CDD's to each utility based on its location.

2.1.3 Retail electricity sales

We collected retail electricity sales data for each utility for the years 2000 to 2009 from information that is published annually by the U.S. Energy Information Administration (EIA 2010)[4].

2.2 Review of Utility-Reported Reliability Data

2.2.1 Geographic representation and coverage

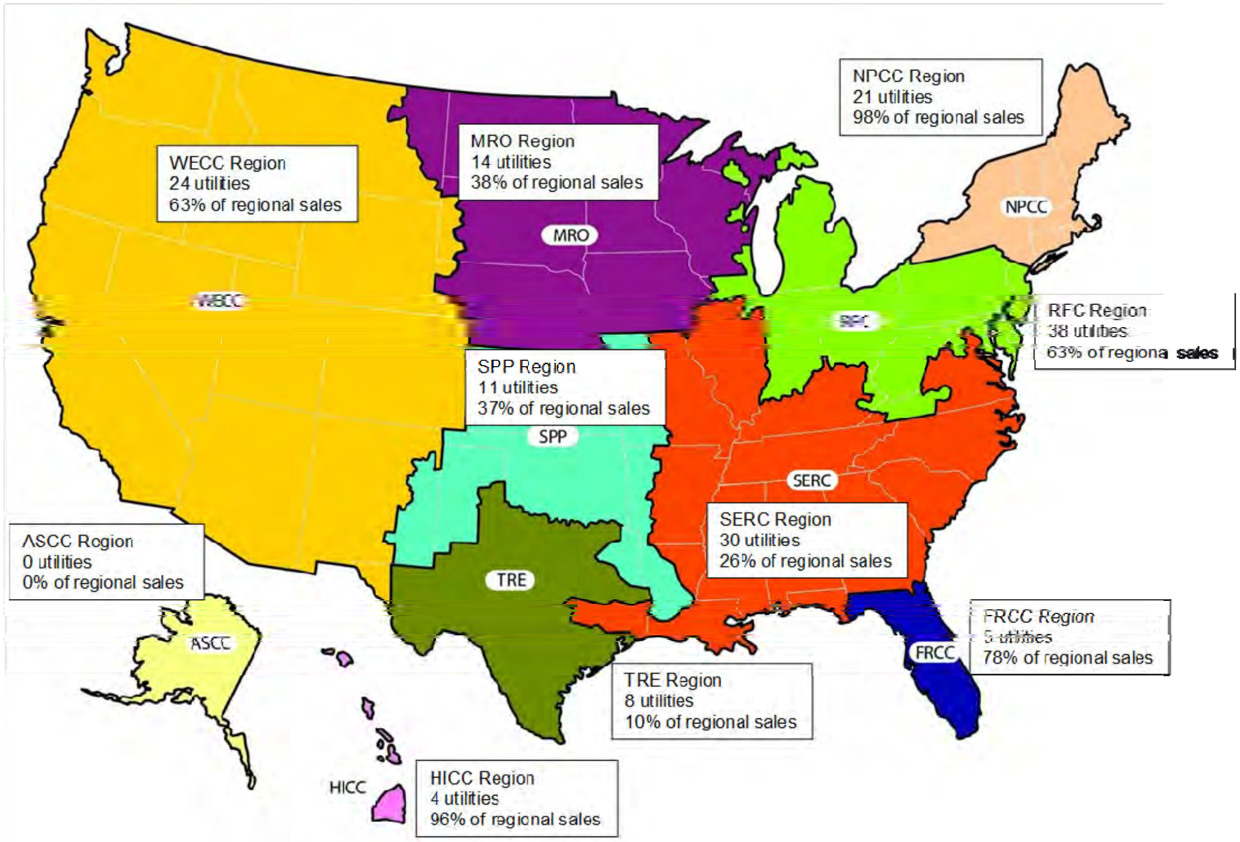
We collected reliability data reported by 155 different U.S. utilities. Of these, 139 are investor-owned utilities and 16 are either municipal utilities or electricity cooperatives. As noted earlier, the large number of investor-owned utilities included in our sample stems from our decision to collect their data through state public utility commissions, which routinely make these data publicly available.

Figure 1 shows the geographic distribution of these utilities by the North American Electric Reliability Corporation (NERC) region. The figure indicates, by NERC region, the number of utilities for which we collected reported reliability data and the percentage of total 2009 retail electricity sales within the region that were accounted for by these utilities.

Table 1 shows, by NERC region, the same information presented in Figure 1 as well as the percentages of total 2009 U.S. sales represented by the utilities for which we collected reported reliability data.

[3] Temperature records came from observation stations located in climatologically homogenous regions within a state. The station's observations are weighted by the area of its climate region as a proportion of the state's area. This produces a weighted average for temperature in the state. For further details on the weighting procedures, see NOAA National Climatic Data Center (2011).

[4] The electricity sales information from the EIA 861 form is also housed in a large database supported by Ventyx (Ventyx 2011).



Source: EPA EGrid 2010 Map

Figure 1. Map of U.S. by NERC Region

Table 1. 2009 Sales of Utilities for which Data were Collected, by NERC Region

NERC Region	Total Electricity Sold in 2009 (TWh)	Total Electricity Sold by Utilities for which Data were Collected (TWh)	Percentage of Region	Percentage of U.S. Total
Western Electricity Coordinating Council (WECC)	658.7	416.4	63%	12%
Midwest Reliability Organization (MRO)	205.5	77.6	38%	2%
Southwest Power Pool (SPP)	186.1	67.9	37%	2%
Northeast Power Coordinating Council (NPCC)	222.7	218.2	98%	6%
Reliability First Corporation (RFC)	919.7	579.2	63%	16%
Southeast Electricity Reliability Council (SERC)	876.3	231.2	26%	6%
Florida Reliability Coordinating Council (FRCC)	217.9	171.0	78%	5%
Texas Regional Entity (TRE)	271.4	28.4	10%	1%
Hawaiian Islands Coordinating Council (HICC)	10.1	9.7	96%	0%
Alaska Systems Coordinating Council (ASCC)	6.3	-	0%	0%
TOTAL	**3,574.7**	**1,799.6**	**50%**	**50%**

The reliability data we collected was reported by electric utilities that together represent half of total U.S. electricity sales in 2009. The percentages of sales represented vary by region, from a low of 0% (Alaska Systems Coordinating Council [ASCC]) to a high of 98% (Northeast Power Coordinating Council [NPCC]). Reliability data from utilities representing more than 50% of total regional sales were collected for the Hawaiian Islands Coordinating Council (HICC), Florida Reliability Coordinating Council (FRCC), NPCC, Reliability First Corporation (RFC), and the Western Electricity Coordinating Council (WECC).

2.2.2 Completeness of reported reliability data by utility and over time

Figure 2 shows, annually from 2000 to 2009, the number of utilities whose reported reliability data we collected. The figure shows a general increase from 2000 to 2006 in the number of utilities reporting SAIFI and SAIDI both with and without major events included. The trend declines after 2006 for SAIFI and SAIDI without major events and by 2009 for SAIFI and SAIDI with major events. This is likely because the most recent data were still being processed by the utilities or their regulators and were not available at the time this report was prepared.

Figure 3 shows the number of years of reported reliability data we collected from each of the 155 utilities. We were able to obtain a complete time series of 10 years of SAIFI and SAIDI without inclusion of major events for close to half of the utilities (70 utilities). We collected six or more years of reported reliability data for over 80% of the utilities (127 utilities).

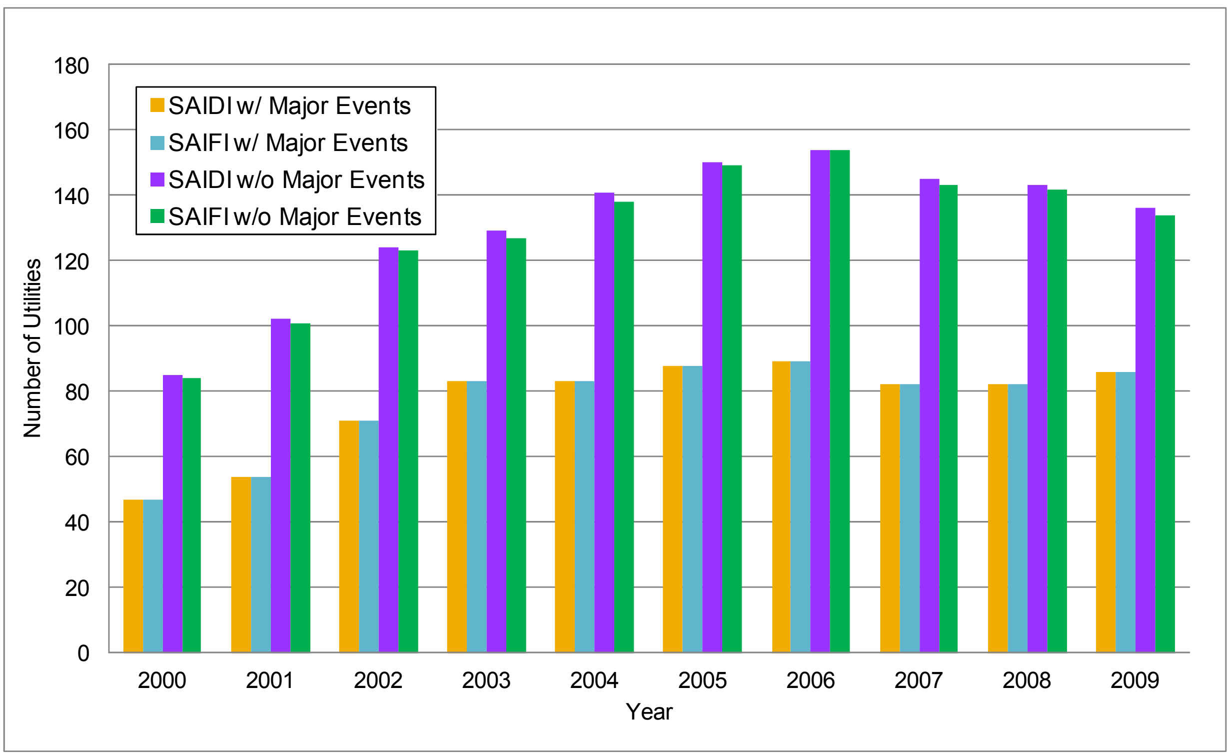


Figure 2. Number of Utilities with SAIDI and SAIFI Data

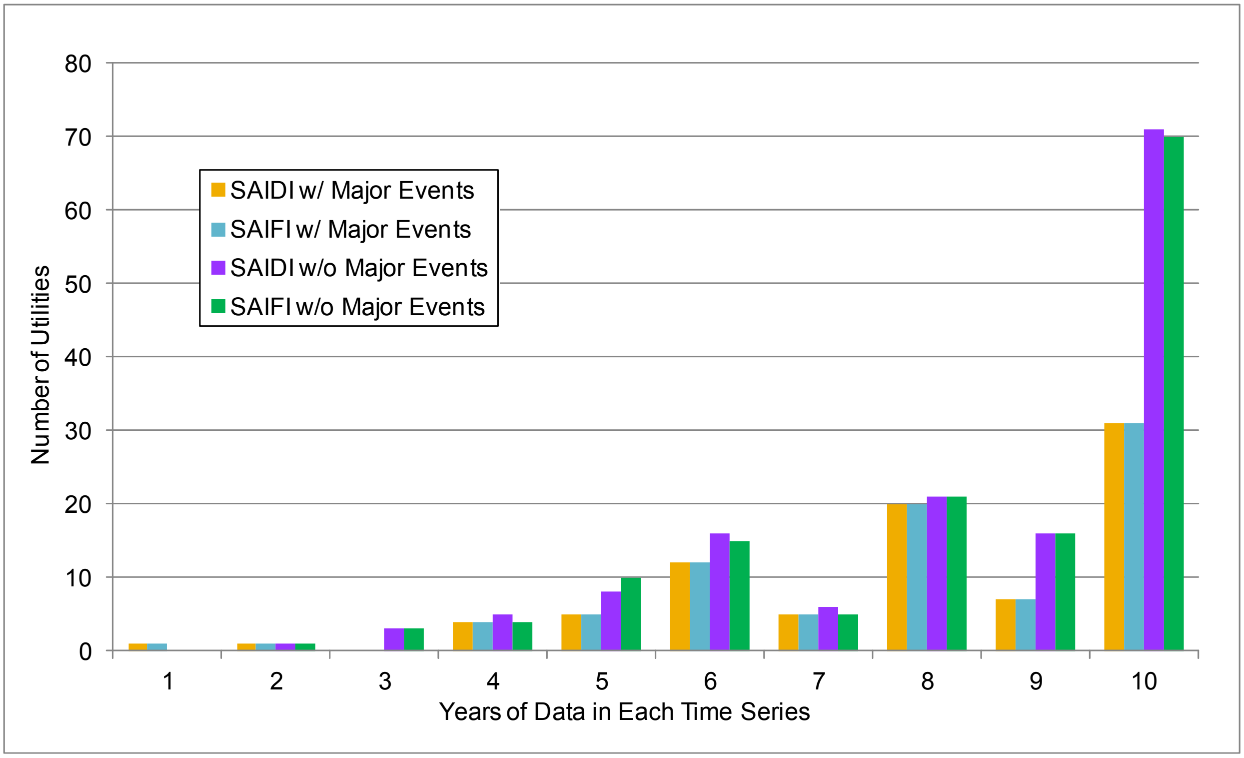


Figure 3. Completeness of Time Series

2.2.3 Distributions of reported SAIDI and SAIFI over time

Figures 4 and 5 summarize, by year in box-plot form, the reported SAIDI and SAIFI values. For each year, the box-plot shows the distribution of values, both with major events (left) and without major events (right). The top and bottom of each box represent the 75th and 25th percentiles, respectively and the line through the box is the median. The mean is indicated with a blue diamond. The end points of each vertical line are the minimum and maximum values in each data set.

With the slight exception of SAIFI in year 2000, the mean values of SAIDI and SAIFI are greater when major events are included. This is to be expected. Removal of major events, which by definition are large, lowers the resulting SAIDI and SAIFI. The anomaly in the year 2000 SAIFI is due to the different mix of utilities for which we obtained SAIFI with versus without inclusion of major events. The large increase in variability in year 2008 was the result of a major hurricane.

We also examined the year to year variability in SAIDI and SAIFI for each utility. Considering the utilities for which we had a complete record of reliability metrics, we found that the mean of coefficient of variations (the ratio of the standard deviation to the mean) was more than 20% for both SAIDI and SAIFI (without major events), indicating considerable variability in the annual values of these metrics. (The means of coefficients of variation for SAIDI and SAIFI with major events were even larger.)

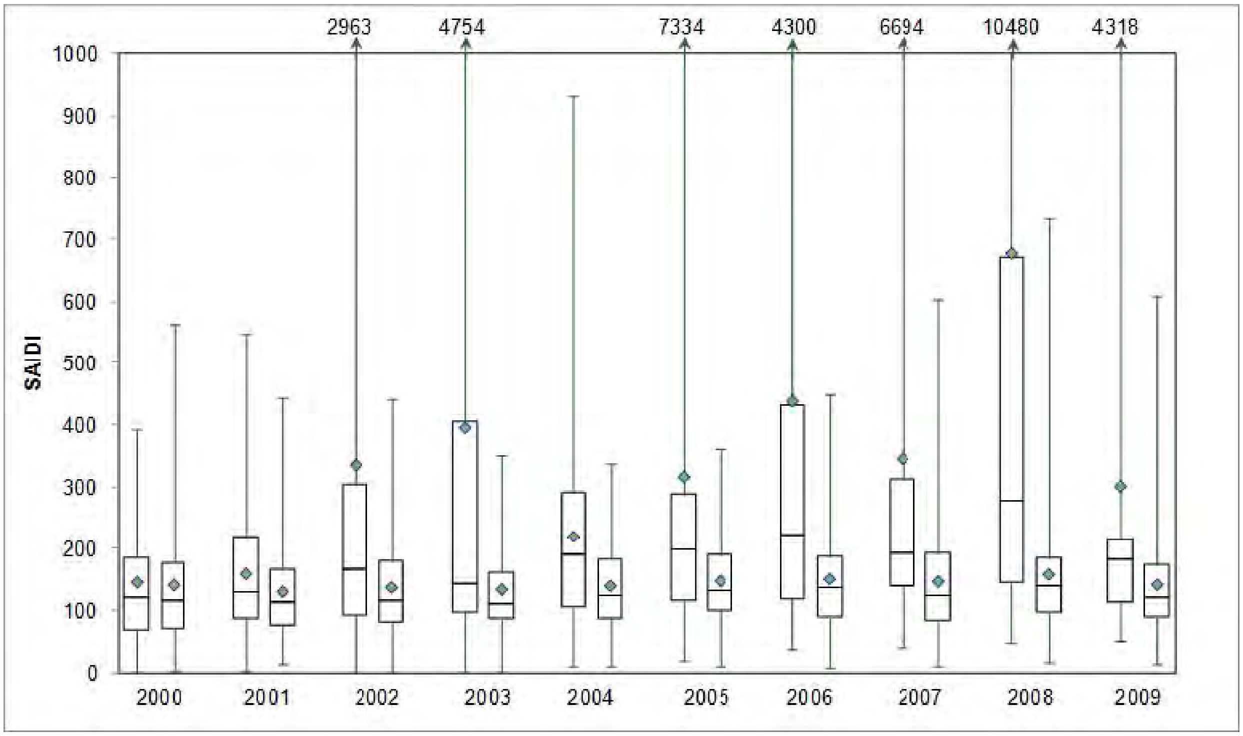


Figure 4. Box-Plot of SAIDI by Year with and without Major Events

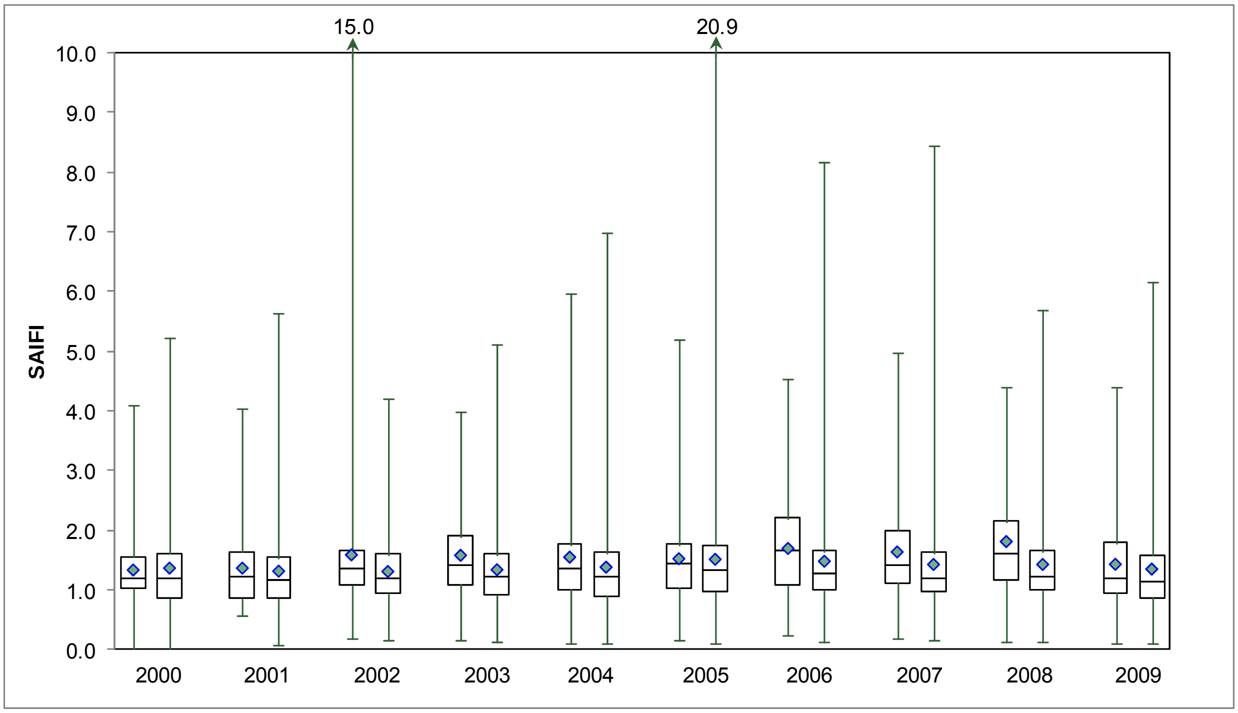


Figure 5. Box-Plot of SAIFI by Year with and without Major Events

2.2.4 Automated outage management systems (OMS)

Table 2 summarizes, by NERC region, the number of utilities that had installed or upgraded their OMS by 2010. We found that 110 utilities or 65% of the 155 utilities for which we collected reported reliability data, had installed or upgraded their OMS by 2010.

Table 2. Summary of Utilities with an OMS

NERC Region	# Utilities We Obtained From	# Utilities that Reported They Had Installed or Upgraded their OMS by 2009
WECC	24	21
MRO	14	9
SPP	11	5
NPCC	21	16
RFC	38	24
SERC	30	16
FRCC	5	5
TRE	8	3
HICC	4	1
ASCC	0	0
TOTAL	**155**	**100**

Figure 6 presents, by NERC region, the number of utilities that installed or upgraded their OMS in each year. The line spanning years represents the cumulative number of utilities that installed or upgraded their OMS up to and including each year. The figure shows that a significant number of utilities had installed or upgraded their OMS prior to first year of our analysis (i.e., prior to 2000). The "year unknown" column in the figure represents the number of utilities that reported they had installed or upgraded their OMS, but for which we could not determine the year of installation or upgrade. In reviewing the information we collected on OMS installation or upgrade, we found that none of the utilities installed or upgraded their OMS system more than once during the 10-year time period for which we collected reported reliability data.

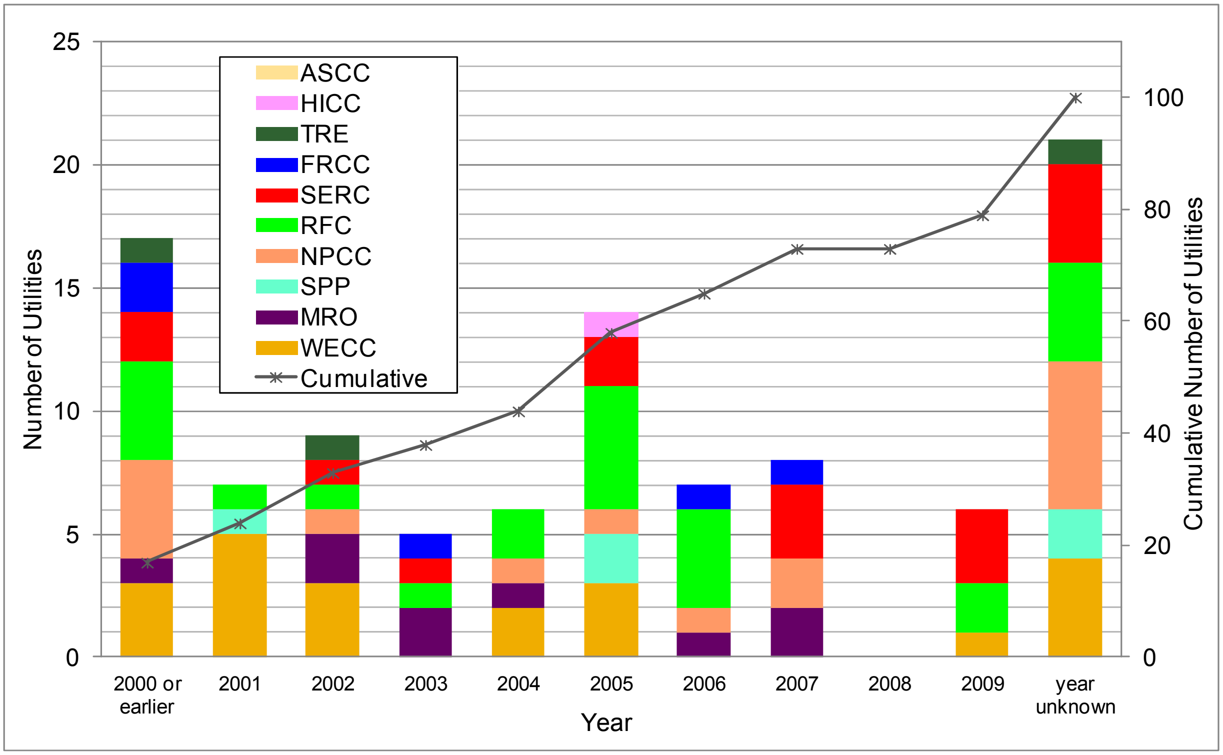


Figure 6. Number of Utilities by Year and NERC Region that Installed or Upgraded their OMS

2.2.5 Major events

Information on reliability is sometimes segmented using the concept of major events. Major events are extraordinary power interruptions and are defined by a variety of criteria to differentiate them from routine power interruptions. There are a number of different definitions for major events. (See Eto and LaCommare (2008)). IEEE Standard 1366-2003 is a voluntary industry standard that articulates a consistent set of definitions and procedures for measuring and reporting distribution reliability information, including a heuristically-derived and statistically-based definition of major events.

Adoption of IEEE 1366-2003 is in its early stages. In 2006, 14 utilities (of the 120 utilities whose data we obtained) reported reliability information to their state regulatory utility commission using this standard (Eto and LaCommare 2008). For the current study, we collected reliability data for 38 utilities (of the 155 utilities whose data we collected) that reported these data using the IEEE standard.[5] (See Table 3).

[5] Our sample is influenced by the decision to collect reliability information reported to state regulatory utility commissions because commission rules usually specify how data are to be reported and, in particular, whether the IEEE Standard 1366-2003 or another set of reliability data definitions will be used. Many utilities rely on IEEE Standard 1366-2003 for internal uses of reliability metrics and at the same time report reliability data to their state regulatory utility commissions using different definitions, as required by commission rules.

Table 3. Summary of Utilities Using IEEE 1366-2003

NERC Region	# Utilities For Which We Obtained Reported Reliability Data	# Utilities That Reported Reliability Using IEEE Std. 1366-2003
WECC	24	14
MRO	14	1
SPP	11	1
NPCC	21	4
RFC	38	12
SERC	30	6
FRCC	5	0
TRE	8	0
HICC	4	0
ASCC	0	0
TOTAL	**155**	**38**

The reported reliability data were prepared for all the years of data we collected either using IEEE Standard 1366-2003 or some other definition for the SAIDI and SAIFI reliability metrics.[6] In total, we collected data from 38 utilities that used the IEEE standard to report their reliability.

Of these 38 utilities, eight utilities also reported their reliability for some portion of the ten years using another set of definitions for the SAIDI and SAIFI reliability metric without major events. In preparation for the rigorous analysis of the relationship between reported reliability and reporting practices presented in section 4, we look specifically at the differences in reported SAIDI and SAIFI without inclusion of major events, as reported by these eight utilities.

Figures 7 presents the percentage differences between SAIDI (not including major events) reported using the IEEE standard and SAIDI (not including major events) reported using another set of definitions. Figure 8 presents the same comparison for SAIFI (not including major events). Each color represents the percentage differences for each year of data from a single utility. (Note that the utilities are not identified by name.)

Visual inspection of Figure 7 shows that SAIDI when reported using the IEEE standard is generally lower, on average, than SAIDI when reported using a method other than the IEEE standard; that is, the percentage differences are generally negative values. However, for a given utility, there is also significant variability in these values from year to year and these variations appear to be as large as, or even larger than, the average of the percentage differences over the years. Figure 8 indicates that the percentage differences for reported SAIFI using the two methods are less discernable (i.e., close to zero). In addition, the percentage variation from year to year for a given utility is also smaller, with a few notable exceptions.

[6] In every instance in which a utility relied on IEEE Standard 1366-2003 to report its reliability metrics, the standard was used to prepare the metrics for each year for which data were obtained. In many instances, this meant that the utility had recalculated its reliability metrics using the standards for the years prior to the utility's decision to use the standard.

It is difficult to draw definitive conclusions from such a small sample of utilities. In Section 4, we apply more sophisticated statistical methods to re-examine this topic using a much larger sample. Application of these methods will demonstrate the additional value they provide when compared to the simple comparisons presented in this section and in Section 3, which have formed the primary basis for prior analyses of these data.

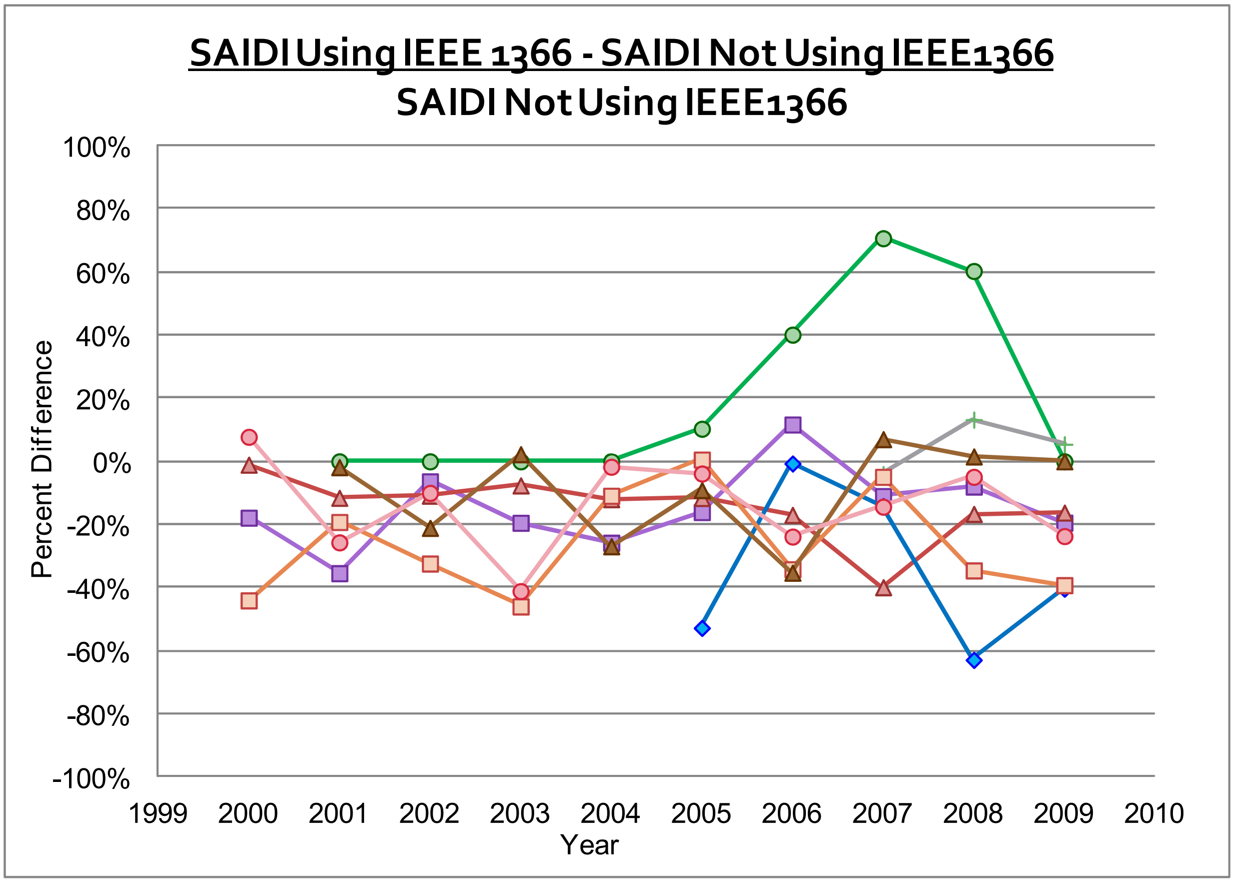


Figure 7. Percentage Difference in SAIDI Without Major Events Included between Using and Not Using IEEE Standard 1366-2003

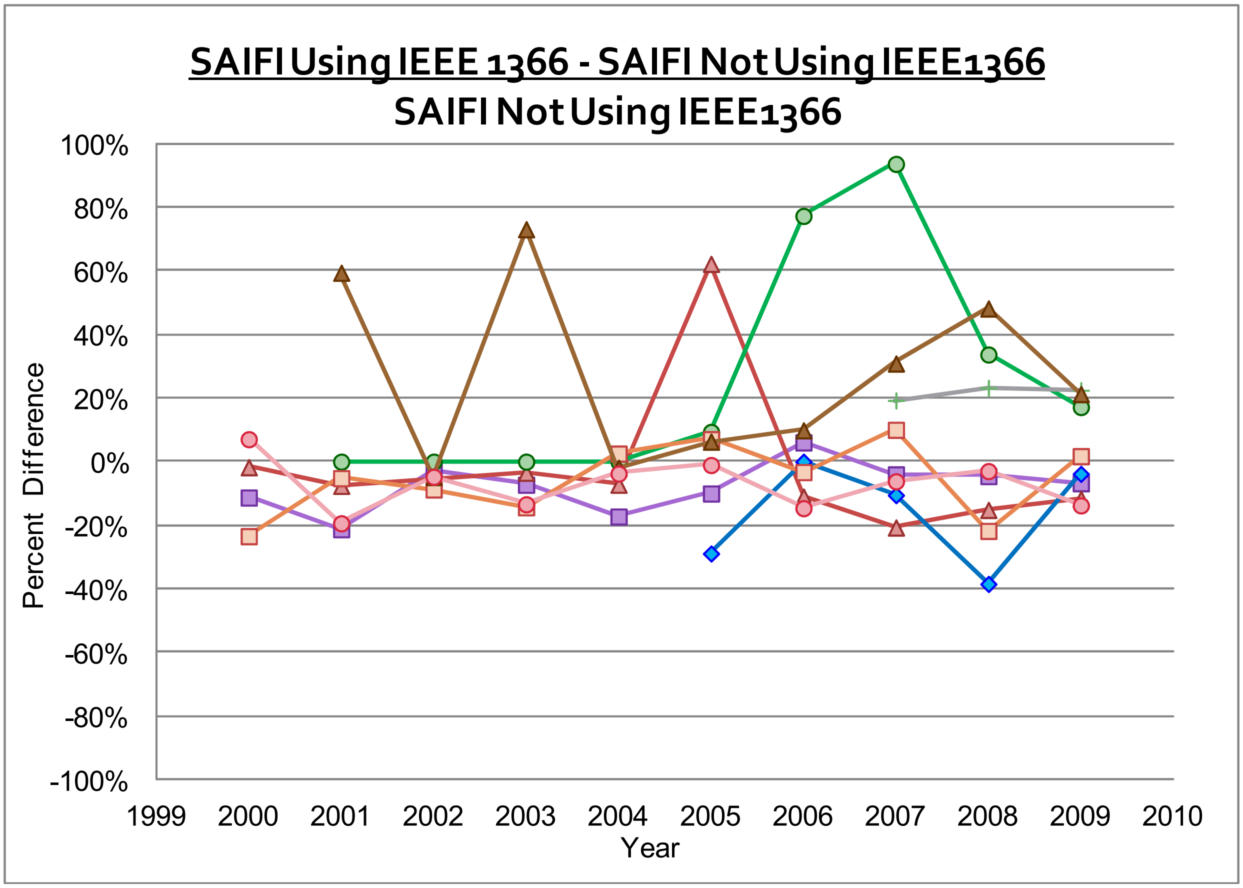


Figure 8. Percentage Difference in SAIFI Without Major Events Included between Using and Not Using IEEE Standard 1366-2003

3. An Initial Review of Time Trends in Reported Electricity Reliability

This section presents findings from several approaches to describing time trends in reported electricity reliability. The approaches are all based on descriptive statistics, including means, medians, and customer-weighted means. The presentation compares the trends to one another and discusses considerations that affect their interpretation. Appendix A compares a variant of the analysis of customer-weighted means, which enables a direct comparison to a closely related analysis conducted by the IEEE Distribution Reliability Working Group.

Descriptive statistics cannot take into account the influences of utility-specific factors, such as the location or size of a utility, or utility-specific sources of measurement bias, such as reliance on automated outage management systems to collect reliability data or use of IEEE Standard 1366-2003 to report reliability metrics. We present a multivariate statistical analysis, which seeks to take these factors into account in Section 4.

3.1 Time Trends in Reported Electricity Reliability Based on Descriptive Statistics

We developed time trends for each of the four reported reliability metrics (SAIDI and SAIFI both with and without inclusion of major events) using three descriptive statistics and two sets of the data we collected. The three descriptive statistics are the mean, the median, and the customer-weighted mean. The customer-weighted mean takes into account differences in utility size and can be thought of an aggregate SAIDI and SAIFI for the entire population of included utilities, taken as a whole.

The two sets of data on reported electricity reliability are: 1) all utilities for which we had reported reliability, which we label "All;" and 2) a subset of the full set, which includes only those utilities for which we had reported reliability data for every year in the time series (years 2000-2009), which we label "Same Utilities." [7]

Figures 9 through 12 plot the three descriptive statistics for SAIDI (both with and without inclusion of major events) and SAIFI (both with and without inclusion of major events), respectively, for all the utilities for which we had reported reliability data.

Figures 13 through 16 plot the three descriptive statistics for SAIDI and SAIFI (both with and without inclusion of major events), respectively, for only those utilities for which we had reported reliability data for every year in the time series (2000-2009). We had 10 years of data on SAIDI and SAIFI with major events for 28 utilities. We had 10 years of data on SAIDI and SAIFI without major events for 67 utilities.

Tables 4 and 5 present the numerical results from a best-fit linear regression for each of the trends plotted in Figures 9 through 16 for SAIDI and SAIFI, respectively.

[7] Figure 3 in Section 2 shows the number of utilities for which we had 10 years of reported reliability data for each of the four reliability metrics.

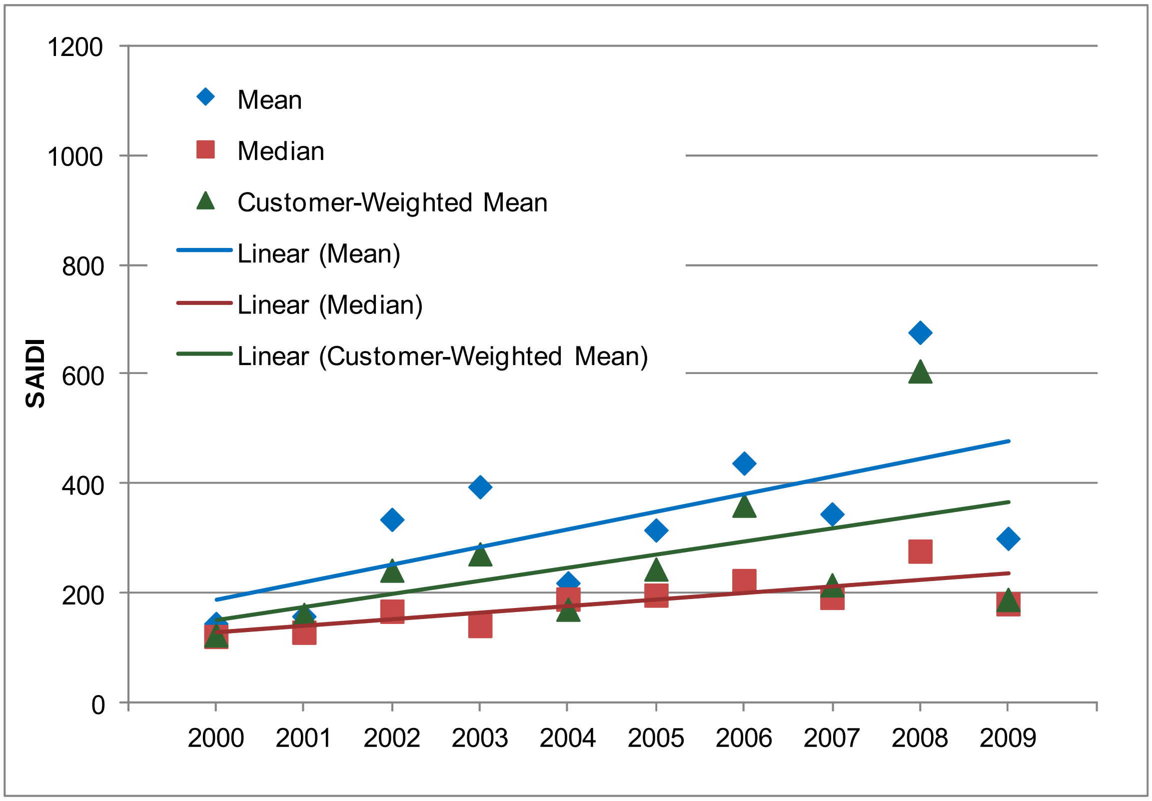


Figure 9. SAIDI with Major Events – All Reported Reliability Data

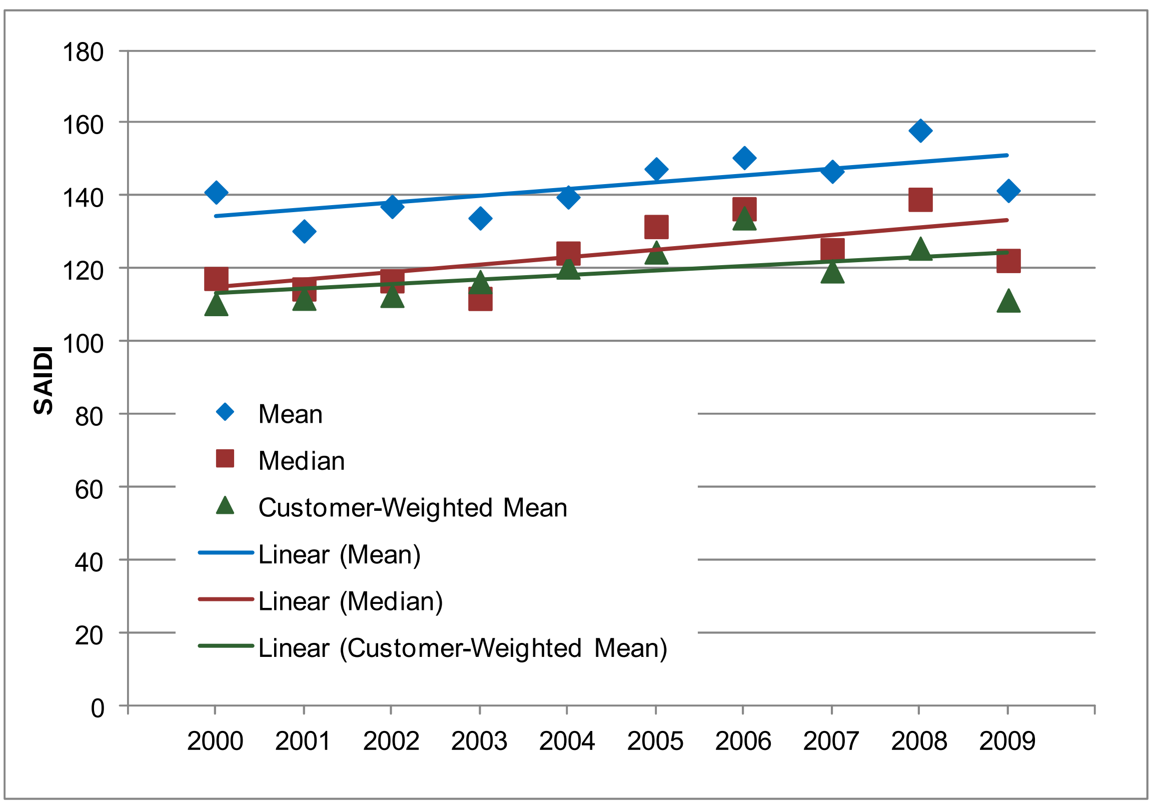


Figure 10. SAIDI without Major Events – All Reported Reliability Data

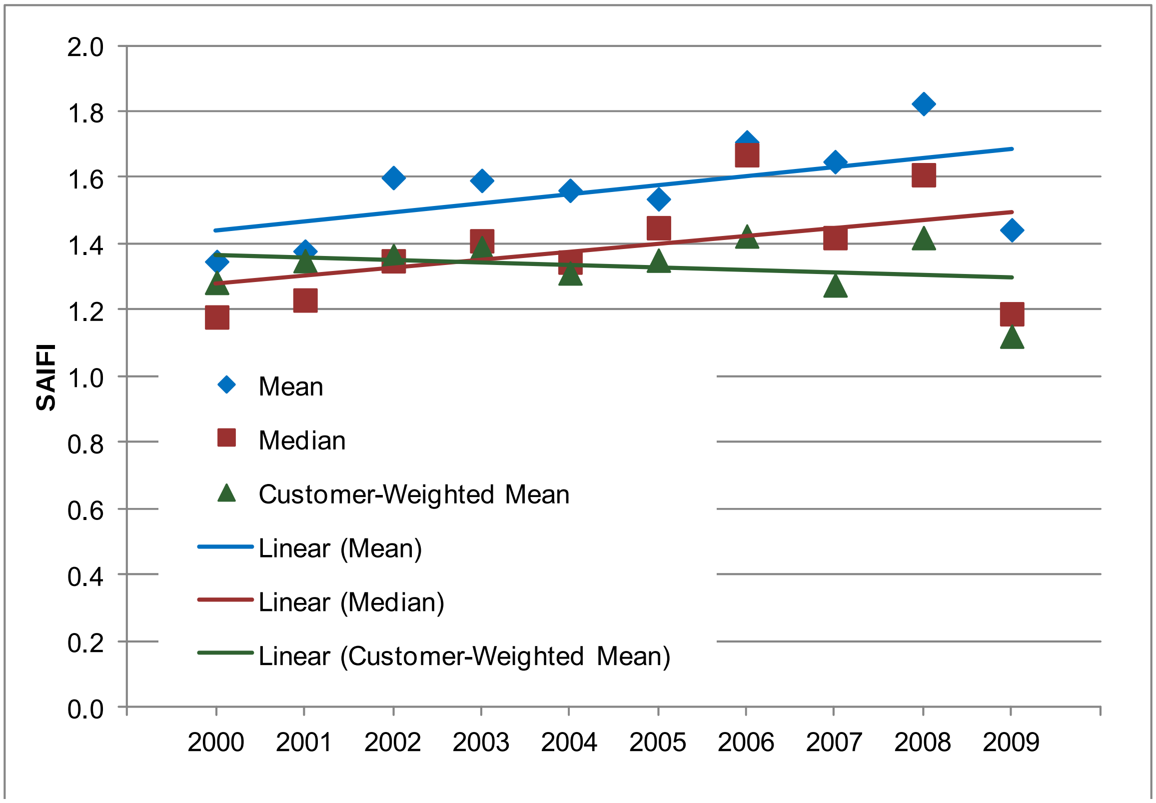


Figure 11. SAIFI with Major Events – All Reported Reliability Data

Figure 12. SAIFI without Major Events – All Reported Reliability Data

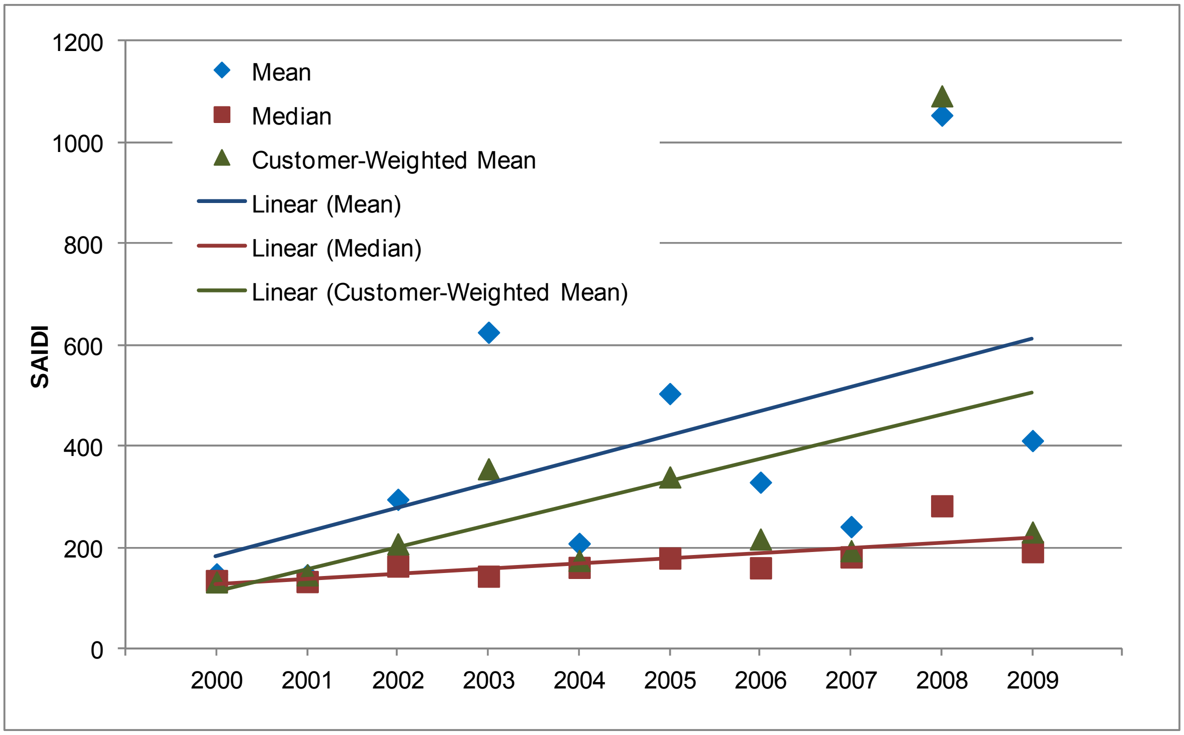


Figure 13. SAIDI with Major Events – Reported Reliability Data from Same Utilities for Every Year – N=28

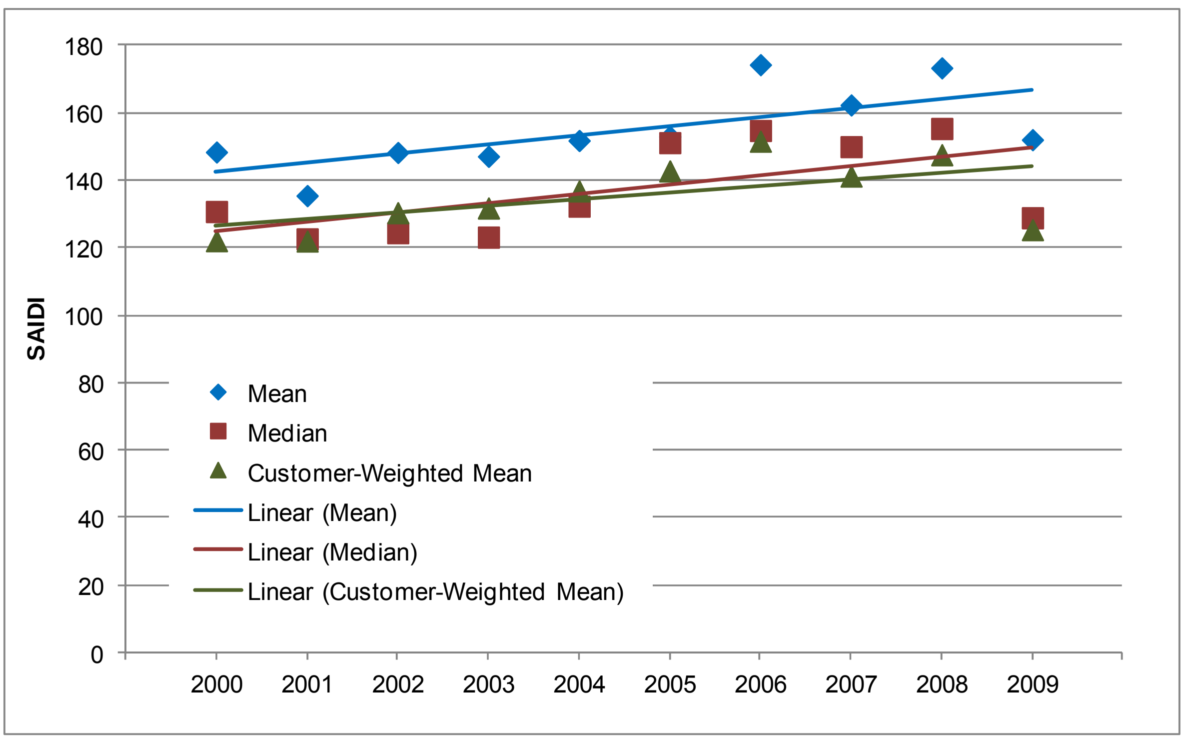


Figure 14. SAIDI without Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67

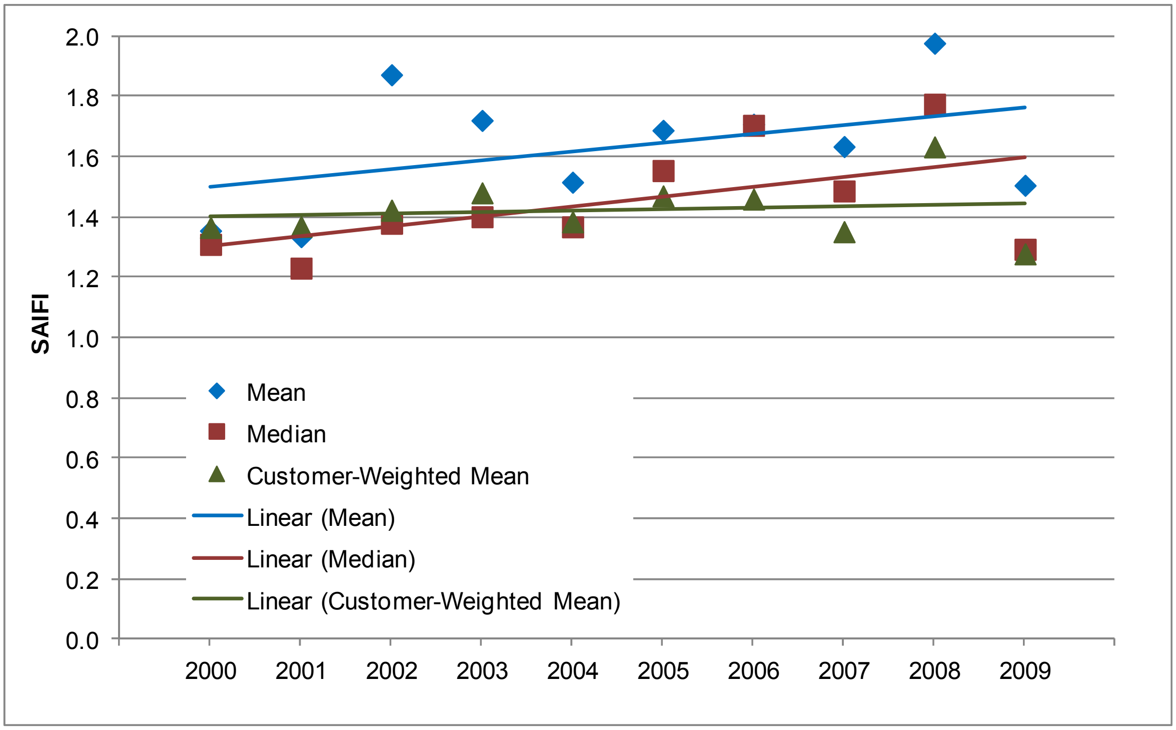


Figure 15. SAIFI with Major Events – Reported Reliability Data from Same Utilities for Every Year – N=28

Figure 16. SAIFI without Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67

Table 4. Summary of Numerical Best Fit of Trends in SAIDI

	intercept	slope	R squared
SAIDI w/ Major Events - All Reported Reliability Data			
Average	157	32.1	0.40
Median	118	11.8	0.59
Customer Weighted Average	129	23.6	0.27
SAIDI w/o Major Events - All Reported Reliability Data			
Average	133	1.84	0.46
Median	113	2.06	0.45
Customer Weighted Average	112	1.18	0.22
SAIDI w/ Major Events - Reported Reliability Data from Same Utilities for Every Year – N=28			
Average	134	47.7	0.27
Median	116	10.4	0.53
Customer Weighted Average	69	43.6	0.21
SAIDI w/o Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67			
Average	140	2.73	0.47
Median	122	2.75	0.37
Customer Weighted Average	124	1.99	0.33

* Note intercept, slope and R-squared assume x=1 (2000) to x=10 (2009)

Table 5. Summary of Numerical Best Fit of Trends in SAIFI

	intercept	slope	R squared
SAIFI w/ Major Events - All Reported Reliability Data			
Average	1.41	0.0275	0.32
Median	1.25	0.0241	0.20
Customer Weighted Average	1.37	-0.0077	0.07
SAIFI w/o Major Events - All Reported Reliability Data			
Average	1.35	0.0109	0.22
Median	1.21	0.0015	0.01
Customer Weighted Average	1.24	-0.0146	0.59
SAIFI w/ Major Events - Reported Reliability Data from Same Utilities for Every Year – N=28			
Average	1.47	0.0290	0.18
Median	1.27	0.0320	0.29
Customer Weighted Average	1.40	0.0044	0.02
SAIFI w/o Major Events – Reported Reliability Data from Same Utilities for Every Year – N=67			
Average	1.37	0.0148	0.24
Median	1.27	0.0059	0.04
Customer Weighted Average	1.28	-0.0056	0.13

* Note intercept, slope and R-squared assume x=1 (2000) to x=10 (2009)

3.2 Discussion of Time Trends Based on Descriptive Statistics

One dominant theme emerges when considering these time trends taken as a whole: the best-fit linear slopes associated with the time trends are generally positive. That is, trends in reported reliability metrics, whether assessed by considering means, medians, or customer-weighted means, indicate that the value of the metrics is increasing over time. Increasing values of SAIDI and SAIFI suggest that reported reliability is getting worse, on average, over the 10 years.[8]

This finding holds regardless of whether major events are included in the calculation of SAIDI and SAIFI. The finding also holds considering both sets of data: all reported reliability data and reported reliability from the same utilities for all 10 years.

The limited exceptions to this are the trend for customer-weighted SAIFI calculated using all reported reliability data both with and without inclusion of major event days, as well as the trend from the group of utilities for which we had a complete set of values for all 10 years (the "same" utilities) without inclusion of major event days.

Additional themes emerge from sub-groupings of these time trends.

The impacts of major events on reliability appear to have increased over time. Generally speaking, the slopes are larger for the time trends based on SAIDI and SAIFI with major events than they are for the time trends based on SAIDI and SAIFI without inclusion of major events (i.e., steeper slopes mean that reliability is getting worse faster).

On average, larger utilities, as measured by numbers of customers, would appear to be more reliable than smaller utilities. Both the intercept and slope terms are higher for the time trends based on means than they are for the time trends based on customer-weighted means.

The time trends discussed in this section are all subject to important caveats that temper the significance of these themes.

First and foremost, the statistical representativeness of the data we have collected with respect to the reliability experience of the U.S. as a whole has not been established. The findings presented in this section, while reflective of a significant portion of total U.S. electricity sales, can only be said to capture the collective reliability experience of these utilities alone and not the entire U.S. In section 2, we noted that our data are composed primarily of data from investor-owned utilities and are not drawn evenly from all regions of the U.S.

Second, the trends we examine focus on averages estimated over a period of 10 years. In Appendix A, we consider only the most recent four years of this period and find generally that, on average, reliability has improved continuously. While this reversal is not large enough to off-set the overall trend for the entire 10 years, it is notable and should be acknowledged when seeking to draw conclusions regarding the significance of the overall 10-year trend.

---

[8] Appendix A compares a variant of this analysis of customer-weighted means, which enables a direct comparison to a closely related analysis conducted by the IEEE Distribution Reliability Working Group.

Third, the slopes are modest in size compared the year-to-year variability that exists in the reliability metrics reported by individual utilities. In Section 2, we found that the means of the coefficients of variation (i.e., the standard deviation divided by mean) for SAIDI and SAIFI without major events from utilities for which we had all 10 years of data (67 utilities) was 20% or more. The 10-year change in values for these same reliability metrics is generally 20% or less.

Fourth, as will be examined directly in Section 4, trends based solely on descriptive statistics cannot take into account utility-specific influences that may introduce bias. Potential, yet unaccounted for, sources of bias include the means by which reliability data were collected (e.g., using an OMS versus using more manual forms of recording the frequency, extent, and duration of power interruptions), and the means by which they were reported (e.g., using IEEE Standard 1366-2003 versus individualized, state PUC-mandated reporting conventions), which are just two of several that we examine in Section 4.

4. Findings from the Statistical Analysis of Reliability Data Reported by Electric Utilities

This section describes the statistical methods we used to analyze reported electric reliability. The purpose of these methods is to take explicit account of utility-specific effects that might otherwise introduce bias into our findings. The trends presented in Section 3 were all based on descriptive statistics that cannot take these factors into account. After introducing the statistical methods, we present our findings from application of them to identify reliability trends and to correlate these trends with the factors we considered. The questions we explore in this analysis include:

1. Are there trends in reported electricity reliability over time?
2. How are trends in reported electricity reliability affected by the installation or upgrade of an OMS?
3. How are trends in reported electricity reliability affected by use of IEEE Standard 1366-2003?

4.1 Introduction to the Statistical Methods Used in the Analysis

As described in Section 2, the reliability data we analyzed consists of up to 10 years of two reliability metrics, SAIDI and SAIFI both with and without major events, collected from up to 155 electricity distribution utilities. The data have both a cross-sectional (i.e., multiple utilities) and time-series (i.e., multiple years) element. This type of data is commonly referred to as an analysis of panel data because the methods and results involve data that have these features.

The structure and completeness of the panel data influenced our choice of tests for specifying the statistical models and the methods for estimating the model parameters and standard errors. Cameron and Trivendi (2009) refer to the specific type of panel data we analyzed as "short" because the data structure has many entities (i.e., utilities), but only a few time periods (compared to the number of entities). In addition, our panel data are unbalanced because they do not contain reliability metrics for every year from all utilities (Wooldridge 2002). In sum, our analysis is of a short, unbalanced panel data set.

The conventional statistical method used to analyze short, unbalanced panel data is multivariate regression. Multivariate regression models provide quantitative estimates of the strength of the correlation between an outcome variable (i.e., the reliability metric SAIDI or SAIFI) and a set of explanatory variables.

The specific forms of the multivariate regression models we estimated are called either "fixed effects" or "random effects" models. Fixed and random effects models are particularly useful for this type of analysis because they enable the regressions to explicitly account for differences in the outcomes (i.e., SAIDI and SAIFI) that are correlated with differences in the sources of the data for these outcomes (i.e., the utilities). As noted earlier, utilities follow different practices in reporting reliability (e.g., whether or not they use IEEE Standard 1366-2003). Fixed and random effects models can explicitly account for these correlations and thereby remove the influence of these differences from the other correlations under consideration (i.e., correlations with the other explanatory variables).

4.2 Application of the Statistical Models

Application of the statistical methods involved four sequential steps. First, we transformed the reliability metrics by expressing them as natural logarithms. Second, we conducted F-tests on the transformed reliability metrics to confirm the appropriateness of using statistical models that consider utility-specific effects. Third, we used Hausman's tests to determine whether it was more appropriate to estimate a fixed effects model versus a random effects model to capture these utility-specific effects. Fourth, we estimated two sets of models; the first consisting of a set of random effects models, and a second consisting of models that do not consider utility-specific effects. We briefly describe below each of these steps.

We decided to transform the reliability metrics by expressing them as natural logarithms for two reasons. First, it is well known that the metrics themselves tend to follow a log-normal distribution; transforming them results in a normal distribution. Second and perhaps more importantly, expression as a natural logarithm allows for a natural interpretation of the estimated coefficients from the regression equations as percentages. For example, if an estimated coefficient for an explanatory variable has a value of 0.02, the natural interpretation is that a step change in that variable correlates to a 2% increase in the reliability metric. For more information on the transformation of the data, please see Appendix B.

The F-test is a standard statistical test to determine the appropriateness of estimating fixed and random effects models. The F-test is a test of the null hypothesis that there are no fixed or random effects. If these null hypotheses can be rejected with some degree of statistical confidence, it means there may be fixed or random effects, which means the use of fixed and random effects models to estimate these effects is warranted.

We estimated two sets of log-linear models. First, we estimated a set of models that include utility-specific effects; we called this set "Model 1." The models in this set control for systematic differences across utilities, such as time, region (climate), system size, and installation or upgrade of an OMS. The set consists of separate models for SAIFI and SAIDI, both with and without inclusion of major events. We estimated both fixed and random effects versions of Model 1.

The specification of Model 1 is as follows:

$$y_{it} = \alpha + \beta_1 Sales_{it} + \beta_2 HDD_{it} + \beta_3 CDD_{it} + \beta_4 YR_t + \beta_5 OMS_{it} + \beta_6 POST\ OMS_{it} + \mu_i + \varepsilon_{it} \quad (1)$$

where:

y_{it} is the natural log of the reliability metric (SAIDI or SAIFI) for utility i=1,2,…,N in year t=1,2,…,T;
$Sales_{it}$ is annual electricity sales in Millions of MWh;
HDD_{it} is heating degree-days;
CDD_{it} is cooling degree-days;
YR_t is a time trend in years;

OMS_{it} is an indicator variable that takes the value 1 if utility i has an OMS installed in year t, and 0 otherwise
$POST\ OMS_{it}$ takes the value 1 for the first year after utility i installs an OMS, 2 for the second year after an OMS is installed, etc, and 0 for earlier years prior to the installation of an OMS; and
μ_i is the utility-specific error.

We applied the Hausman (1978) specification test to Model 1 to determine whether the fixed or random effects version of this model was more appropriate. The Hausman test examines whether, under the null hypothesis, the individual utility effects are uncorrelated with the other regressors in the model. If the null hypothesis is not rejected, both the random effects and the fixed effects models are consistent, but only the random effects model is efficient. This means that fixed and random effects models will have the same expected values, but the random effects model will have much smaller standard errors. Using a fixed effects model when the random effects model is consistent may lead to an erroneous interpretation of the statistical significance of coefficients. See Greene (2000) for a more detailed discussion of the difference between fixed and random effects. The Hausman test did not reject the null hypothesis of random effects in six of the eight regressions we ran (see Tables 7 and 8). We therefore concluded−in general−that the random effects model was consistent and more efficient than the fixed effects version.

Second, we estimated a set of models that did not include utility-specific effects, which we called "Model 2." The models sought to examine how reporting differences, specifically utilization of IEEE Standard 1366-2003, along with other unobserved correlates with utilization of the IEEE standard, correlate with reported reliability. The specification of Model 2, which does not include utility-specific effects, μ_i, is as follows:[9]

$$y_{it} = \alpha + \beta_1 IEEE_i + \beta_2 Sales_{it} + \beta_3 HDD_{it} + \beta_4 CDD_{it} + \beta_5 YR_t + \beta_6 OMS_{it} + \beta_7 POST\ OMS_{it} + \epsilon_{it} \quad (2)$$

where, in addition to the variables defined above:
$IEEE_i$ takes the value 1 if utility i reports interruptions using IEEE Standard 1366-2003, and 0 otherwise.

Note that for both Model 1 and Model 2, the time trend, YR_t, enters as a linear time variable rather than as a year-specific effect. This additional assumption was deemed reasonable because the cost of including it as a year-specific effect is high, in terms of degrees of freedom, for such a short dataset.[10] See Appendix D for a model that includes the time trend as a year-specific effect; the time trend is similar, but each individual year is not statistically significant.

We estimated standard errors for the one-way unbalanced data model using a specialization (Baltagi and Chang 1994) of the approach proposed by Wansbeek and Kapteyn (1989) for unbalanced two-way models. The Wansbeek and Kapteyn method for estimating variance

[9] In order to understand the effect on reported reliability using IEEE Standard 1366-2003, utility effects cannot be included in the model because utility effects take account for all systematic differences among utilities, including whether or not the utility used IEEE Standard 1366-2003.

[10] While reasonable for the purposes of this report, we plan to explore the assumption of linearity in future research.

components is the default approach used by SAS in the one-way random effects estimation of unbalanced panel data (SAS 2011b).

4.3 Findings

4.3.1 Are there utility-specific differences in reported electricity reliability?

Table 6 presents the results from the application of the F-test to the reliability metrics. The table indicates that both one-way (utility only) and two-way (utility and year) effects are statistically significant (at the 0.01% confidence level) for all four reliability metrics – SAIDI and SAIFI both with and without major events. That is, there are very strong correlations between the utility and the values of the reliability metrics as well as between the utility and the year when correlated to the values of the reliability metrics.

Table 6. F-test of the Hypothesis that there are No Utility-Specific Effects

	One-way Fixed Effect (Utility)			Two-way Fixed Effects (Utility and Year)		
Reliability Metric	F Value	Degrees of Freedom (among/within)	Prob. > F	F Value	Degrees of Freedom (among/within)	Prob. > F
ln SAIDI (w/o MEs)	15.29	143/1037	< 0.0001	14.80	152/1029	< 0.0001
ln SAIFI (w/o MEs)	15.67	143/1034	< 0.0001	15.08	152/1026	< 0.0001
ln SAIDI (w/MEs)	5.32	85/595	< 0.0001	5.74	94/587	< 0.0001
ln SAIFI (w/MEs)	9.61	85/595	< 0.0001	9.73	94/587	< 0.0001

Note: ME = major event, ln = natural logarithm,
Note: The SAS software test for effects of cross-level interactions (utility or utility and year) reports two types of degrees of freedom: 1) "among" and 2) "within". The "among" value is equal to k‐1 degrees of freedom where k is the number of cross-sections per effect. The "within" value is equal to N‐k degrees of freedom, where N is the total number of observations and k is the number of cross-sections per effect.

The strong correlation between the utility alone and the reliability metrics indicates that it is important to take this correlation into account when examining correlations between the reliability metrics and other correlated (or explanatory) variables. In other words, there are strong utility-specific effects that are systematically correlated with the reliability metrics.

At this point, we cannot determine the exact or complete set of sources or causes of these effects, but they are consistent with the existence of utility-specific differences in reporting practices. Hence, taking this correlation into account appropriately means that subsequent correlations with other variables will not be “contaminated” by these differences in reporting practices (by any other utility-specific effects).

The correlation between utility plus year to the reliability metrics means that year-to-year correlations are also important to take into account when examining correlations with other variables. This finding supports examining the reliability metrics, by utility, as a time series, rather than as a handful of observations randomly drawn from different years.

The bottom two rows of Table 7 include the results from applying the Hausman test to Model 1. The test does not reject the null hypothesis. We therefore conclude that a random effects model is consistent and more efficient than the fixed effects version in this case.[11] Accordingly, we present results for the random effects model only.[12]

4.3.2 Are there trends in reported electricity reliability over time?

We estimated each model separately for each of the four different reliability metrics: SAIFI and SAIDI, both with and without major events. We estimated Model 1 with two specifications for the treatment of installation or upgrade of an OMS. The first version considers only the differences in reported reliability before and after installation or upgrade. The second version considers a "learning" effect in which the model estimates the correlation with installation or upgrade in subsequent years.

Tables 7 and 8 present the results for the Model 1. Table 7 presents the results for SAIFI and SAIDI with and without major events for the initial version of Model 1 (i.e., without OMS learning). Table 8 presents the results for SAIFI and SAIDI with and without major events for the second versions of Model 1 (i.e., with OMS learning).

Both tables show evidence of a secular trend of increasing frequency and duration of interruptions on average over the years 2000-2009.[13] In Table 7, SAIFI and SAIDI without major events, columns III and IV, the coefficients for *YR* are 0.018 and 0.022 for SAIFI and SAIDI, respectively. Both of these estimated coefficients are statistically significant at the 1% confidence level. The natural interpretation of these coefficients is that SAIFI and SAIDI are increasing annually, by about 2% for both SAIFI and SAIDI.

It is useful to note that 2% annual decreases in reported reliability are roughly consistent with the simple linear trends presented in Tables 4 and 5 (and in Figures 9 through 16) in section 3. In this regard, the observation first made in section 3 – that these trends are modest in comparison to the year-to-year variability in these reported reliability metrics – also apply equally to these findings.

These trends are also confirmed when major events are not included in SAIFI and SAIDI. In Table 7, columns I and II, the coefficients for *YR* are 0.022 and 0.047 for SAIFI and SAIDI, respectively. Again both are statistically significant at the 1% level. The natural interpretation of these coefficients is that SAIFI is increasing annually at about 2% and SAIDI is increasing annually at about 5%.

---

[11] Note that because this is an unbalanced data set, the Breusch-Pagan test for random effects is not appropriate here (SAS 2011a).

[12] See Appendix D for fixed effect results; as expected, the coefficient estimates are similar but less efficient (that is, they are not statistically significant).

[13] Model 1 restricts the time trend to be linear. In the appendix, we present results from a model that includes year fixed effects rather than a linear time trend. The estimates of the time fixed effects increase in a relatively linear fashion, but the loss in degrees of freedom results in estimates that are not statistically significant. As noted in an earlier footnote, we plan to explore the assumption of a linear time trend in future research.

Table 7. One-way Random Effects Regression (Model 1): The Effect of Sales, HDD, CDD, Time, and OMS on Frequency and Duration of Interruptions (with Major Events Included)

	With Major Events Included				Without Major Events Included			
	I		II		III		IV	
	ln SAIFI		**ln SAIDI**		**ln SAIFI**		**ln SAIDI**	
Intercept	-42.6706	***	-89.3893	***	-35.9339	***	-39.1642	***
	(13.3111)		*(25.7755)*		*(8.4430)*		*(9.3735)*	
Sales	-0.00153		-0.00133		-0.00112		-0.00246	
	(0.0023)		*(0.0035)*		*(0.0020)*		*(0.0023)*	
HDD	-0.00002		6.65E-06		-0.00002		-0.00003	
	0.0000		*(0.0001)*		*0.0000*		*0.0000*	
CDD	-7.01E-06		-0.00006		0.000133	**	0.000072	
	(0.0001)		*(0.0001)*		*(0.0001)*		*(0.0001)*	
YR	0.021496	***	0.04716	***	0.01797	***	0.02194	***
	(0.0067)		*(0.0129)*		*(0.0042)*		*(0.0047)*	
OMS	0.004575		0.287343	***	-0.04346		0.136875	***
	(0.0561)		*(0.1020)*		*(0.0404)*		*(0.0452)*	
POST OMS								
Utility Effects	Yes		Yes		Yes		Yes	
R-square	0.0205		0.0581		0.0265		0.0632	
Hausman Test (m Value)	3.21		2.49		2.22		1.75	
Hausman Interpretation	Fail to Reject Null		Fail to Reject Null		Fail to Reject Null		Fail to Reject Null	

Note: A generalization of the R-square measure is reported and is based on Buse (1973). This generalized goodness-of-fit measure is the proportion of the transformed sum of squares of the dependent variable that is attributable to the influence of the independent variables exclusive of utility-specific random effects. Hausman (1978) tests the null hypothesis that random effects are preferred over fixed effects. Standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

It is important to observe that, in making these estimates, Model 1 also takes into account the potential for correlations with electricity sales and climate (as well as the effect of utility-specific differences, as discussed in Section 4.3.1). In fact, the model finds that these external factors are generally not at all correlated with the increasing secular trends observed for both SAIFI and SAIDI. One exception is that correlation between SAIFI without major events and CDDs is statistically significant at the 5% level. The natural interpretation of this correlation is that SAIFI is very slightly higher when there are more CDDs.

It is premature to speculate or draw conclusions about the causes underlying these trends without more explicit treatment of potential sources of bias in reported reliability.[14] In the next subsection, we focus on measurement error as one potential source of bias.

4.3.3 How are trends in reported electricity reliability affected by the installation or upgrade of the automated OMS?

The estimates for the *OMS* coefficients in Table 7, describe the strength of the correlation between installation or upgrade of an OMS and reported SAIFI and SAIDI. The results differ for SAIFI and SAIDI, both in terms of the direction and strength of the correlation.

The correlation of SAIFI (both with and without major events) to the installation or upgrade of an OMS is mixed and not statistically significant, even at a 10% confidence level. However, the correlation with SAIDI is always positive and is statistically significant at the 1% confidence level. The natural interpretation of this correlation is that utilities that install or upgrade their OMS report higher SAIDI by nearly 29% when major events are included and by nearly 14% when major events are not included compared to utilities that did not install or upgrade their OMS.

Table 8 further explores the relationship between installation or upgrade of an automated OMS and reported reliability by introducing a time-element, *POST OMS*, which tracks how the correlation with SAIFI and SAIDI changes in the years following installation or upgrade of the system.

The results are suggestive, but not conclusive. The results are suggestive because there is evidence that installation or upgrade of an OMS is correlated with an initial increase in SAIFI or SAIDI, but that SAIFI and SAIDI decrease in the years following installation or upgrade. The results are not conclusive because the estimated coefficients are not consistently statistically significant.

The only SAIFI coefficient that is statistically significant is *POST OMS* with major events included, at the 5% confidence level. The natural interpretation is that there is an annual reduction in SAIFI of 2.4% following installation or upgrade of an OMS compared to the SAIFI reported by utilities that did not install or upgrade their OMS.

The SAIDI coefficients that are statistically significant include *OMS* (1% level) and *POST OMS* (10% level) with major events included, and *OMS* (1% level) when major events are not included. The natural interpretations are that when major events are included, there is a one-time increase in SAIDI of 38% followed by an annual decrease of 4%, compared to SAIDI reported utilities that did not install or upgrade their OMS. When major events are not included, there is a one-time increase in SAIDI of 16%.

[14] A high-level analysis of identified outliers was also performed to assess the impact on the regression results and to understand the circumstances behind the outlier. Please see Appendix C.

Table 8. One-way Random Effects Regression (Model 1): The Effect of Sales, HDD, CDD, Time, OMS, and OMS "Learning" on Frequency and Duration of Interruptions (Without Major Events Included)

	With Major Events Included				Without Major Events Included			
	I		**II**		**III**		**IV**	
	ln SAIFI		**ln SAIDI**		**ln SAIFI**		**ln SAIDI**	
Intercept	-65.2081	***	-125.533	***	-42.9671	***	-46.763	***
	(17.1806)		*(31.8647)*		*(9.7397)*		*(10.8662)*	
Sales	-0.00114		-0.00081		-0.00087		-0.00221	
	(0.0023)		*(0.0035)*		*(0.0020)*		*(0.0023)*	
HDD	-0.00002		2.69E-07		-0.00002		-0.00003	
	0.0000		*(0.0001)*		*0.0000*		*0.0000*	
CDD	-0.00002		-0.00009		0.000131	**	0.00007	
	(0.0001)		*(0.0001)*		*(0.0001)*		*(0.0001)*	
YR	0.032765	***	0.065237	***	0.021482	***	0.025734	***
	(0.0086)		*(0.0159)*		*(0.0049)*		*(0.0054)*	
OMS	0.044586		0.380678	***	-0.01665		0.163991	***
	(0.0593)		*(0.1129)*		*(0.0444)*		*(0.0493)*	
POST OMS	-0.02411	**	-0.04141	*	-0.01257		-0.01337	
	(0.0117)		*(0.0216)*		*(0.0087)*		*(0.0097)*	
Utility Effects	Yes		Yes		Yes		Yes	
R-square	0.0292		0.0477		0.031		0.0491	
Hausman Test (m Value)	4.76		11.54	**	5.68		10.78	*
Hausman Interpretation	Fail to Reject Null		Reject Null		Fail to Reject Null		Reject Null	

Notes: A generalization of the R-square measure is reported and is based on Buse (1973). This generalized goodness-of-fit measure is the proportion of the transformed sum of squares of the dependent variable that is attributable to the influence of the independent variables exclusive of utility-specific random effects. Hausman (1978) tests the null hypothesis that random effects are preferred over fixed effects. Standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

The effects of installation or upgrade of an OMS in the years following installation or upgrade are not consistently statistically significant. However, the coefficients on the magnitude of the year-to-year changes in the reliability metrics remain highly statistically significant at the 1% level. In fact, SAIFI and SAIDI with major events increase annually at faster rates (of about 3% and 6.5%, respectively) than the estimates that do not consider this effect. SAIFI and SAIDI without major events increase annually at rates of about 2% and 2.5%, respectively, which is roughly consistent with the estimates that do not consider this effect.[15]

[15] We plan to explore other means for measuring a learning effect, such as consideration of additional years following installation or upgrade of an OMS.

4.3.4 How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003?

Table 9 presents the results for Model 2, which removes γ_i, the company-specific effect, and replaces it with $IEEE_i$, which indicates whether the company relied on IEEE Standard 1366-2003 in reporting its reliability metrics. Table 9 was developed for only SAIFI and SAIDI *without major events* because reliance on IEEE Standard 1366-2003 involves implementing a specific method for not including these events.

Table 9 reports that the coefficient for reliance on IEEE Standard 1366-2003 is not statistically significant for SAIFI and is statistically significant at the 5% level for SAIDI. The natural interpretation of the latter result is that reliance on IEEE Standard 1366-2003 is correlated with a lower reported SAIDI of about 11%. However, we caution the reader that, strictly speaking, this interpretation is premature. The most that can be said is that reliance on the IEEE standard, *along with all other utility-specific effects that are highly correlated with reliance on the IEEE standard*, is correlated with reported reliability in this manner. We leave it to future work to develop specifications that would separate the effect of reliance on the IEEE standards from these other correlates to isolate the impact of this effect uniquely.

The removal of utility-specific effects also affects the values and statistical significance of other correlates in the model, compared to the values estimated for them in Model 1. For SAIFI, the coefficients on *Sales* and *CDD* are also statistically significant in Model 2. For SAIDI, the coefficient on *CDD* is statistically significant in Model 2. The statistical significance of these correlates is likely reflective of a utility-specific effect because these correlations were not at all or less statistically significant when utility-specific effects were taken into account (in Model 1).

Table 9. No Utility Fixed Effects Regression (Model 2): Effect of IEEE, sales, HDD, CDD, Time, and OMS on Frequency and Duration of Interruptions (*without major events*).

	ln SAIFI		ln SAIDI	
Intercept	-34.5795	**	-43.9178	***
	(13.5039)		*(14.9314)*	
IEEE	0.023445		-0.10797	**
	(0.0425)		*(0.0470)*	
Sales	-0.00226	**	-0.00446	
	(0.0010)		*(0.0011)*	
HDD	-0.00002		-0.00008	
	(0.0000)		*(0.0000)*	
CDD	0.000109	***	-0.00011	***
	(0.0000)		*(0.0000)*	
YR	0.01734	**	0.024603	***
	(0.0068)		*(0.0075)*	
OMS	-0.00033		0.041122	
	(0.0602)		*(0.0663)*	
POST OMS	-0.01839		-0.01123	
	(0.0114)		*(0.0126)*	
Utility Fixed Effects	No		No	

Notes: Standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

In contrast, the time trend, *YR*, for both SAIFI and SAIDI is, like Model 1, statistically significant. The natural interpretation is that SAIFI and SAIDI are increasing at slightly less than 2% and nearly 2.5% annually, which is roughly consistent with the interpretation of the coefficients for Model 1 presented in Table 8.

5. Summary and Interpretation of Findings, and Next Steps

The goal of this study is to inform discussions of the reliability of the U.S. electric power system by assessing trends in power interruptions experienced by U.S. electricity consumers. Our analysis is based on up to 10 years of electricity reported reliability information collected from a convenience sample of 155 U.S. electric utilities, which together account for roughly 50% of total U.S. electricity sales.

We built on prior investigations, which focused only on power interruptions originating in the bulk power system, by considering interruptions originating both from the bulk power system and from within local distribution systems. We also accounted for differences among utility practices for collecting information on and reporting power interruptions by employing statistical techniques that remove the influence of these differences on the trends we identify.

We sought to answer three questions:

1. Are there trends in reported electricity reliability over time?

We first conducted an examination relying on descriptive statistics (mean, median, customer-weighted mean) and find that reported reliability has been decreasing over time. With minor exceptions, we observed this trend for all three descriptive statistics when considering all utility reports taken together, as well as only those utility reports for which we had a complete record of 10 years of data. We point out that descriptive statistics alone mask the effects of utility-specific effects that may introduce bias into our findings.

Next, we applied rigorous statistical methods both to confirm that there were utility-specific differences among electricity reliability reports and to take explicit account of these differences in exploring correlations between reported reliability metrics and other factors. Applying these methods, we find that there are statistically significant temporal trends. We find that reported average duration and average frequency of power interruptions has been increasing over time at a rate of approximately 2% annually. In other words, reported reliability is getting worse.

While our findings are highly statistically significant, it is important to place them in appropriate context. The average annual trends we find are modest in comparison to the routinely larger year to year variations in the average duration and frequency of power interruptions experienced by utility customers. For example, in Appendix A, we present a simple analysis of trends over the most recent four years and find reported reliability has been improving over this period.

In addition, we make no claims regarding the applicability of our findings to the reliability of the U.S. electric power system as a whole. Strictly speaking, our findings apply only to the convenience sample of primarily investor-owned utilities for which we were able to collect reported reliability information. In any given year, these utilities represented roughly 50% of total U.S. electricity sales.

2. How are trends in reported electricity reliability affected by the installation or upgrade of an automated outage management system (OMS)?

A principal contribution of our work has been to examine potential sources of measurement error that could influence apparent trends in reported reliability. We find statistically significant evidence that installation or upgrade of an OMS is correlated an increase in the reported duration of power interruptions. This finding confirms anecdotal evidence long been known within the industry that reliance on prior (manual) measurement methods under-reports reliability. We also found preliminary but not statistically significant evidence for a so-called "learning effect" by which reported reliability gradually improves in years subsequent to the initial decrease in reported reliability.

Our findings might suggest that it is simply more accurate measurement of reliability, rather than lower actual reliability, which "explains" the statistically significant trend of decreasing reported reliability over time. However, our analysis takes this factor into account explicitly and still finds statistically significant secular trends toward lower reported reliability over time. Our findings, therefore, highlight the importance of taking into account the means by which reliability information is collected when examining trends in reported reliability.

3. How are trends in reported electricity reliability affected by the use of IEEE Standard 1366-2003?

We also examined a potential source of measurement bias in the form of utility reporting practices. We find that reliance on IEEE Standard 1366-2003 is correlated with higher reported reliability on average compared to reported reliability not using the IEEE standard and that this correlation is statistically significant. Nevertheless, taking this correlation into account, the secular trend of decreasing reported reliability over time remains statistically significant and at approximately the same magnitude as was found earlier (i.e., decreasing at roughly 2% annually). We caution that it is premature to attribute reliance on the IEEE standard as "causing" higher reported reliability because we could not separate the effect of reliance on the IEEE standard from other utility-specific factors (which we did not account for separately) that may also be correlated with reliance on the IEEE standard.

Next Steps

This study finds that there has been a modest, yet statistically significant secular trend of decreasing or declining reported reliability over the past 10 years. In making this finding, we summarize what our analysis to date has and has not accomplished, and outline the directions for next steps in this line of inquiry.

We wish to state clearly that, at this point, we cannot say what has caused the observed decreasing trends in reported reliability or why it is taking place. Our work has considered and characterized the influence of potential sources of measurement error or bias and found that taking these considerations into account changes neither the direction of these trends nor their statistical significance. These findings are important because they allow us to focus on potential causal factors that would help us explain the trends we observe.

To begin this process, we considered potential correlations with highly aggregated measures of weather variability and a simple measure of utility size but found neither to be statistically significant. . However, these examinations are preliminary and far from complete. For example, with respect to the influence of weather variability, we can only conclude that annual HDDs and CDDs as a measure of yearly weather are not well-correlated with the reported reliability metrics. However, these are only two measures of yearly weather variability; there are others that could be studied. Similarly, utility size is only one measure of the many potential differences among utilities that might be correlated with reported reliability.

We believe it is extremely appropriate to continue exploring differences among utilities to better understand the sources or causes of the secular trends in reliability that we observe. Some of the factors we believe should be considered include more disaggregate measures of weather variability (e.g., lightning strikes and severe storms), utility characteristics (e.g., the number of rural versus urban customers, and the extent to which transmission and distribution lines are overhead versus underground), and utility spending on transmission and distribution maintenance and upgrades, including advanced ("smart grid") technologies.

It is our hope that the analysis we have conducted to date will help pave the way for these investigations and that they will be used to help ground future decisions about U.S. reliability policy, practices, and technology on a more solid factual base.

References

Amin, M. 2008. "Challenges in Reliability, Security, Efficiency, and Resilience of Energy Infrastructure: Toward Smart Self-Healing Electric Power Grid." *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*. pp.1-5, July 20-24.
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4596791&isnumber=4595968

Baltagi, B. H. and Y. Chang. 1994. "Incomplete Panels: A Comparative Study of Alternative Estimators for the Unbalanced One-Way Error Component Regression Model." *Journal of Econometrics,* 62(2): 67-89.

Buse, A. 1973. "Goodness of Fit in Generalized Least Squares Estimation," American Statistician, 27. pp. 106-108.

Cameron, A. C., and P. Trivedi. 2009. Microeconometrics Using Stata. TX: Stata Press.

Energy Information Administration (EIA). 2010. "Form EIA-861 Final Data File for 2009." DOE/EIA. http://ei-01.eia.doe.gov/cneaf/electricity/page/eia861.html

Environmental Protection Agency (EPA) 2011. "Emissions & Generation Resource Integrated Database (eGRID)". http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html

Eto, J. E. and K. H. LaCommare. 2008. *Tracking the Reliability of the U.S. Electric Power System: An Assessment of Publicly Available Information Reported to State Public Utility Commissions*. Berkeley CA: Lawrence Berkeley National Laboratory Report LBNL-1092E. October. http://certs.lbl.gov/certs-rtina-pubs.html

Fisher, E., J. Eto, K. Hamachi-LaCommare. 2012. "Understanding Bulk Power Reliability: The Importance of Good Data and A Critical Review of Existing Sources." *Proceedings of the 45th Annual Hawaii International Conference on System Sciences*. Wailea, HI. Jan. 1-4, 2012.

Greene, W. 2000. Econometric Analysis (Fourth Edition). Upper Saddle River NJ: Prentice-Hall.

Hausman, J. A. 1978. "Specification Tests in Econometrics," Econometrica, 46: 1251-1271.

Hines, P., J. Apt, and S. Talukdar. 2009. "Large Blackouts in North America: Historical Trends and Policy Implications." *Energy Policy*, v. 37, pp. 5,249-5,259.

IEEE Power Engineering Society. 2004. *IEEE Std 1366-2003 IEEE Guide for Electric Power Distribution Reliability Indices.* ISBN 0-7381-3890-8 SS95193. New York: Institute of Electric and Electronics Engineers, Inc. May 14. 35 pages.

LET Systems. 2006. "Requirements for the Implementation of an Outage Management System (OMS) Whitepaper." January.
http://www.letsys.com/img/oms_implementation_requirements_whitepaper.pdf

McGranaghan, M., A. Maitra, C. Perry, A. Gaikwad. 2006. "Effect of Outage Management System Implementation on Reliability Indices." *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*. pp.1,208-1,211, May 21-24.
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1668677&isnumber=34941

National Climatic Data Center (NCDC). 2011. U.S. Department of Commerce/NOAA: U.S. Climate Normals; Area-Weighted State, Regional, and National Temperature. Accessed May 2011 at http://cdo.ncdc.noaa.gov/cgi-bin/climatenormals/climatenormals.pl

SAS. 2011a. Specification Tests. Accessed June 27 at:
http://support.sas.com/documentation/cdl/en/etsug/63348/HTML/default/viewer.htm#etsug_panel_sect039.htm

SAS. 2011b. The One-way Random Effects Model. Accessed June 21 at:
http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/viewer.htm#etsug_panel_sect031.htm

UCLA. 2011. Introduction to SAS. UCLA: Academic Technology Services, Statistical Consulting Group. Accessed September 2011at: http://www.ats.ucla.edu/stat/sas/notes2/

Ventyx. 2011. Velocity Suite Database System, Boulder. Accessed May 23.

Wansbeek, T., and A. Kapteyn. 1989. "Estimation of the Error-Components Model with Incomplete Panels," Journal of Econometrics, 41, 341-361.

Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data, MIT Press.

Appendix A. Customer Weighted Average Comparison to IEEE DRWG Benchmarking Analysis

The IEEE Distribution Reliability Working Group (DRWG) conducts an annual benchmarking analysis based on reliability metrics that are submitted on a voluntary basis. At the IEEE Summer 2011 General Meeting, the DRWG presented a customer-weighted time trend that we can compare to a variant of the customer-weighted time trend presented in Section 3.

We made several adjustments to the customer-weighted time trend presented in Section 3 in order to facilitate a more direct comparison with the time trends developed by the DRWG. First, we compare only the years 2006 through 2009, which are the same years for which the DRWG developed its time trend. Second, we compare only those utilities that relied on IEEE Standard 1366-2003 to report their reliability metrics, which are the only utilities that DRWG considers in developing its time trend. Third, we compare only SAIDI and SAIFI without inclusion of major events, again, to be consistent with DRWG. Fourth, we develop our trends based only on those utilities for which we had all four years of reported reliability, again, to be consistent with DRWG. We label the adjusted customer-weighted time trends “Revised LBNL”

Figures A-1 and A-2 present both the original customer-weighted means and trend line best fit equations for all reported reliability data from Section 3 along with the DRWG’s and the revised LBNL customer-weighted means for SAIDI without inclusion of major events and SAIFI without inclusion of major events, respectively.

We find that, the DRWG and the revised LBNL trend lines for both SAIDI and SAIFI are consistent with one another. Both are downward sloping, indicating that over the period 2006-2009, reported reliability, on a customer-weighted basis, has been improving (i.e., reported reliability metrics indicate that reliability is getting better).

This finding contrasts with the time trends presented in Section 3, which found that over the ten-year period from 2000 to 2009, reported reliability metrics were generally increasing (i.e., reliability was getting worse) over time.

As a reminder, the same caveats applied to the time trends presented in Section 3 also apply to the findings presented in this Appendix. First, neither the statistical representativeness of the samples of reported reliability data examined by LBNL nor those included in the DRWG 2011 Benchmarking Survey compared to the U.S as a whole have been established. Second, the influence of utility-specific effects, such as biases that may have been introduced by the decision to report reliability data following IEEE Standard 1366-2003 and reliance on automated outage management system to collect reliability data, among other unexamined sources of potential bias, are not taken into account.

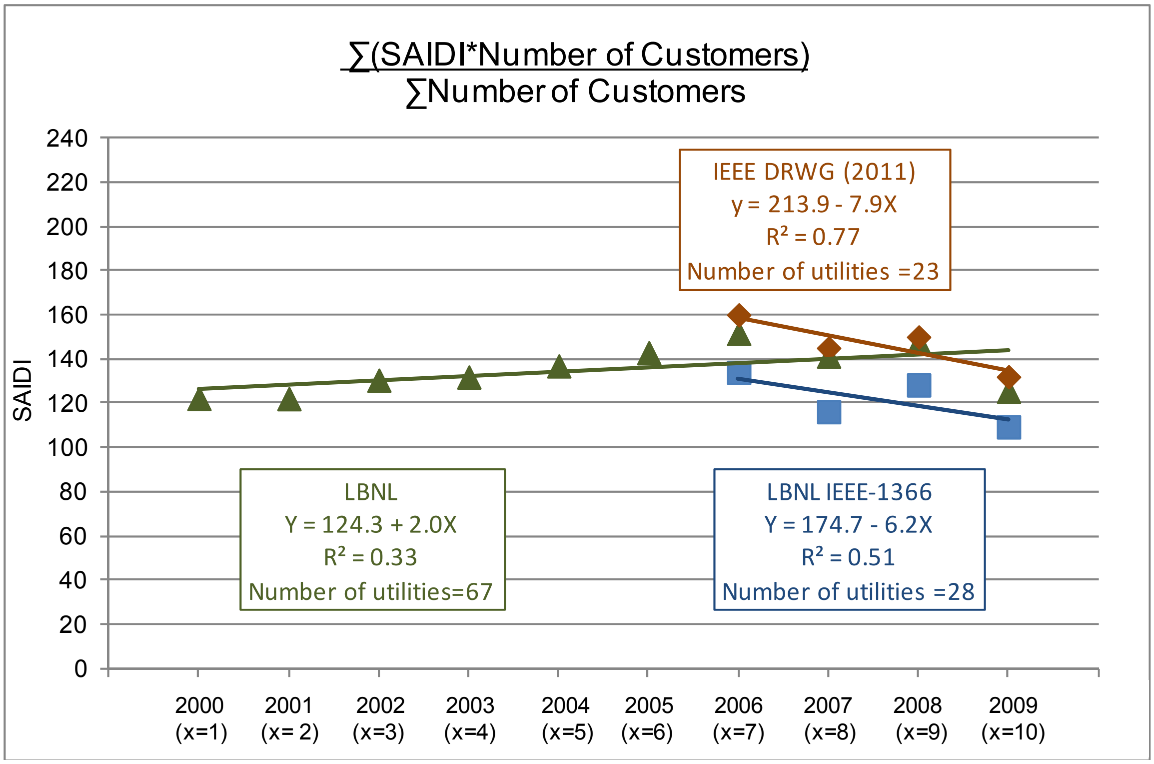


Figure A- 1. Customer-weighted SAIDI w/o Major Events – Comparison of LBNL Findings to those of the IEEE DRWG 2011 Benchmarking Survey

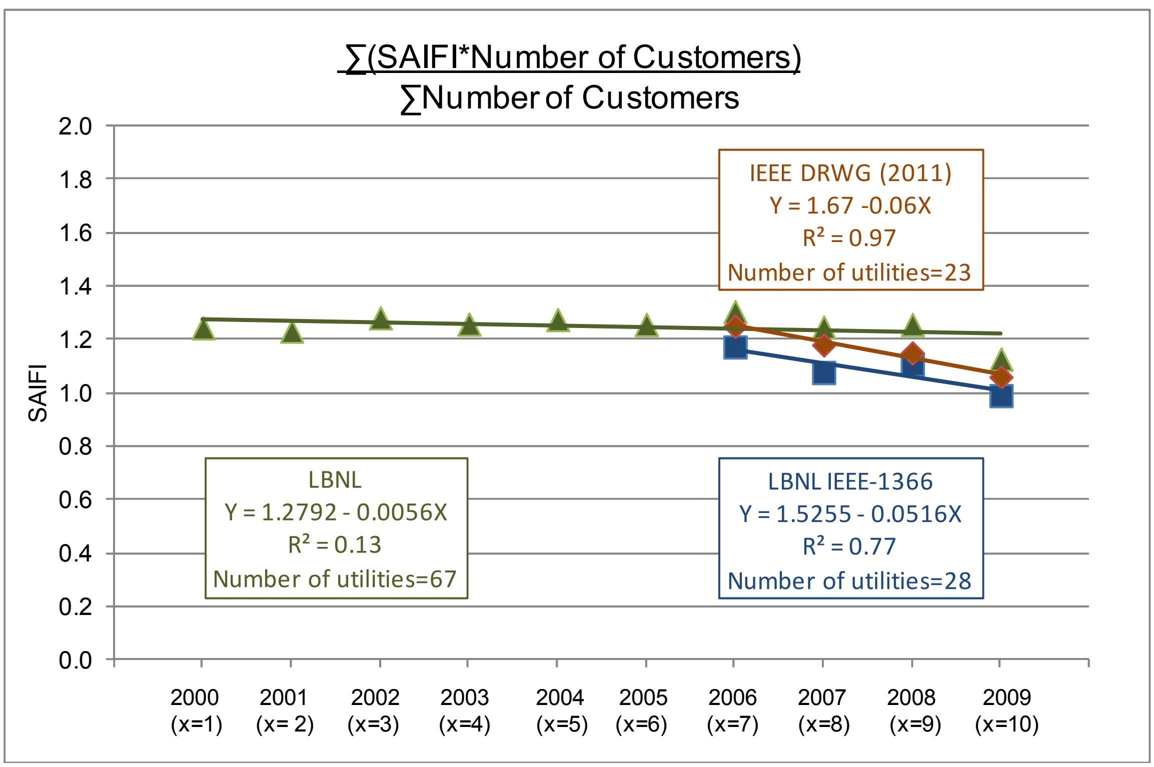


Figure A- 2. Customer-weighted SAIFI w/o Major Events – Comparison of LBNL Findings to those of the IEEE DRWG 2011 Benchmarking Survey

Appendix B. Why a Log-Normal Distribution?

Section 4.2 describes, among other things, our decision to transform the reliability metrics (annual SAIDI and SAIFI) examined in the regression analysis by expressing them as natural logarithms. This appendix describes and documents the rationale for this decision. The rationales all involve describing why we found it desirable to utilize the reliability metrics re-expressed as natural logarithms rather than utilize them un-transformed.

To summarize, we decided to transform the dependent variables for three reasons. First, we found encouraging visual evidence that, when expressed as natural logarithms, the distribution of annual values of SAIDI and SAIFI we collected appeared to follow a normal distribution better than the un-transformed annual values. Second, we performed statistical tests that gave a positive indication that expression of SAIDI without inclusion of major events as a natural logarithm followed a normal distribution far better than did the untransformed version of this variable. For the other three variables (SAIFI with major events, SAIFI without major events, and SAIDI with major events), the statistical tests indicated that neither the transformed nor the untransformed variables conclusively followed a normal distribution. Third, finding no evidence that using the variables in their un-transformed state was superior to using them in their transformed state, the ability to provide an easy-to-explain interpretation of the regression coefficients led to decide to use the variables in their transformed state.

Visual Evidence that Transformed Variables Follow a Normal Distribution Better than Untransformed Variables

From the standpoint of the regression analysis we sought to conduct, it is desirable that the dependent variables used in the analysis follow a normal distribution. Figures B-1 and B-2 show results from a graphical analysis that compares the observed data (i.e., the histogram bins) with theoretical normal (and log-normal) distributions (i.e., the curves shaded in blue). By visual inspection, we find that all four annual reliability metrics are more accurately represented by a log-normal distribution than by a normal distribution.

Statistical Tests for the Normality of the Distributions of Transformed and Untransformed Variables

Tables B-1 and B-2 report results from three statistical testing methods—(1) Kolmogorov-Smirnov, (2) Cramer-von Mises, and (3) Anderson-Darling—commonly used to evaluate the assumed shape of a distribution. Table 1 shows that all tests conducted for all of the reliability metrics rejected the null hypothesis of normality with a high degree of confidence. Table B-2 shows that the null hypothesis of log-normality was rejected for SAIFI (with and without major events included) and SAIDI (with major events), but we fail to reject the null hypothesis for two of the three tests of SAIDI (without major events).

To summarize, formal statistical testing indicated that SAIDI (without major events included) was best fit with a log-normal distribution. However—with the exception of SAIDI (without major events)—statistical testing rejected the null hypothesis of both normality and log-normality at a 99% or greater confidence level.

Table B- 1. Statistical Tests for Normality

Reliability Metric	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
	Reject Null Hypothesis of Normality?		
SAIDI (w/o major events)	Yes***	Yes***	Yes***
SAIDI (with major events)	Yes***	Yes***	Yes***
SAIFI (w/o major events)	Yes***	Yes***	Yes***
SAIFI (with major events)	Yes***	Yes***	Yes***

Note: *** Rejects the null hypothesis at the .01 significance level.

Table B- 2. Statistical Tests for Log-normality

Reliability Metric	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
	Reject Null Hypothesis of Log-normality?		
SAIDI (w/o major events)	No	No	Yes**
SAIDI (with major events)	Yes***	Yes***	Yes***
SAIFI (w/o major events)	Yes***	Yes***	Yes***
SAIFI (with major events)	Yes***	Yes***	Yes***

Note: *** Rejects the null hypothesis at the 0.01 significance level; ** Rejects the null hypothesis at the 0.05 significance level.

The Easy-to-Explain Interpretation of Regression Coefficients when Variables are Transformed

The visual and statistical tests reported above indicate limited support in favor of using transformed variables. Importantly, they offer no support for the superiority of using un-transformed variables.

Transformation of reliability metrics into logarithmic format allows for a natural interpretation of the estimated coefficients from the regression equations as percentages. For example, if an estimated coefficient for an explanatory variable has a value of 0.02, the interpretation is that a step change in that variable is correlated to a 2% increase in the reliability metric.

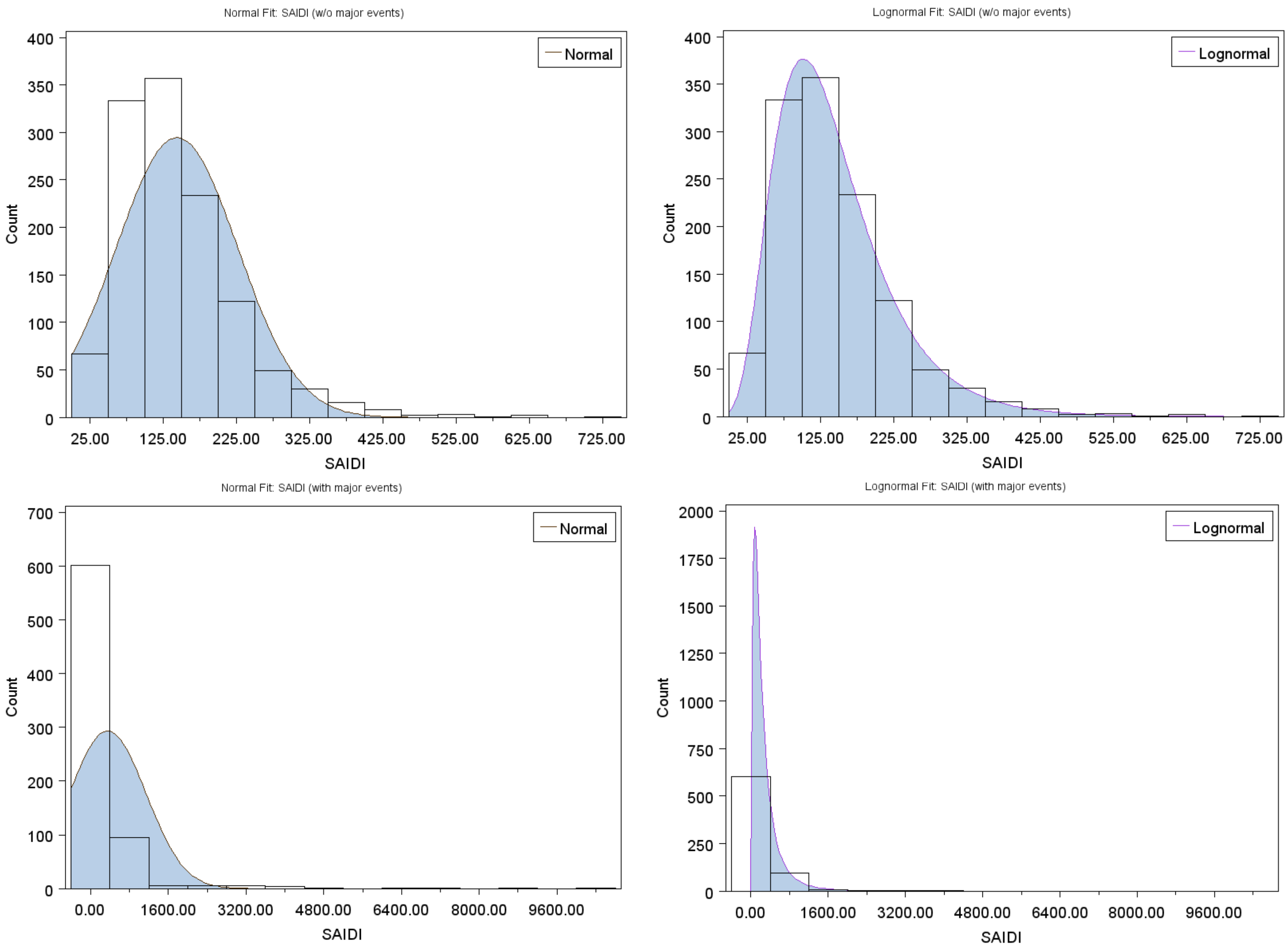


Figure B- 1. Graphical analysis of SAIDI with and without major events included.

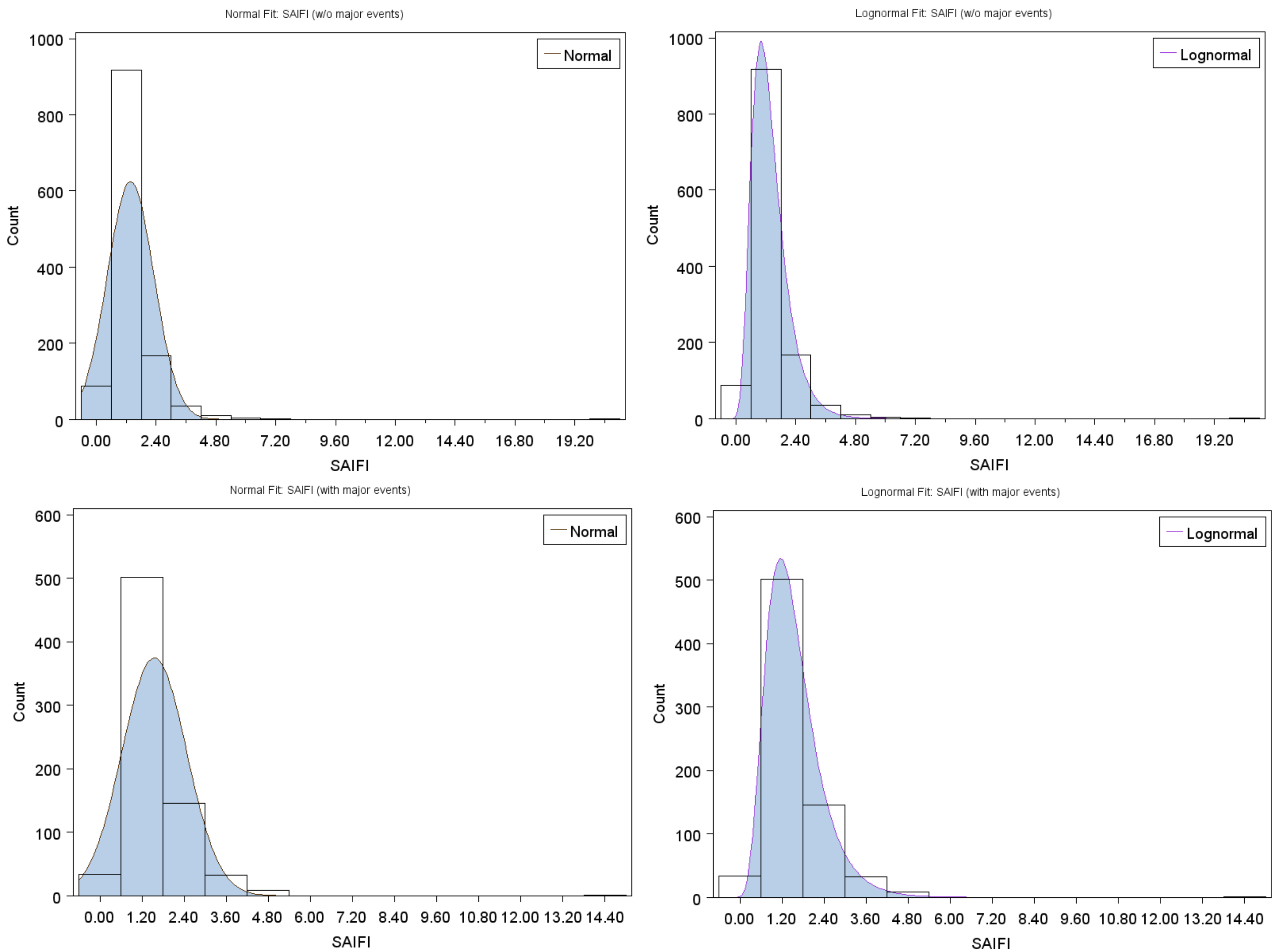


Figure B- 2. Graphical analysis of SAIFI with and without major events included

Appendix C. Examination of Outliers

We conducted a three-step analysis of the regression results to identify extreme and influential outliers, research circumstances that might have affected the reporting (values) of SAIFI and SAIDI, and tested the effect of removing outliers on the model results.

First, we flagged any observations that had a studentized residual that exceeded a pre-defined threshold of plus or minus three (UCLA 2011). In parallel, we carried out a Cook's statistical test to evaluate the size of the residual and influence (leverage) of the individual observations on the model results. Any observation with a Cook's D statistic greater than four divided by the sample size (n) was flagged as an influential and extreme outlier (UCLA 2011). We compiled a list of outliers that were flagged by *both* the studentized and Cook's statistical method simultaneously and carried out a deeper analysis on these observations.

Next, we reviewed the reliability event information to determine if any errors in information transcription occurred when we were collecting the data. No errors were found in the process of entering the data into our database from the source of the reliability event data.

We then looked into possible explanations for the extreme values to help us understand what was happening during these specific years and at these utilities. Table C-1 summarizes the number of extreme and influential outliers that were identified in the preceding steps. As shown in the table below, most of the outliers are explained by either a rare weather occurrence or by the characteristics of the utility service territory with these events leading to lower or higher values of the metrics when compared to other utilities.

Table C- 1. Summary Explanation of Identified Outliers

Dataset	Number of Extreme and Influential Outliers	Number of Utilities	Explanation
SAIDI with major events included	17	12	• 9 outliers due to severe storms, including two hurricanes • 1 outlier due to increased use of troubleshooting personnel that impacted the reliability metrics • No information on the remaining seven outliers
SAIDI without major events included	17	5	• 14 outliers attributed to characteristics of the service territory including a small territory size and increased use of troubleshooting personnel that impacted the reliability metrics • No information on the remaining three outliers
SAIFI with major events included	10	4	• Seven are from a single utility that attributes their anomalous metric values to the high concentration of distribution networks and large customer base • One was due to a large wind and snow storm • No information on the remaining two outliers
SAIFI without major events included	27	8	• 20 outliers attributed to characteristics of the service territory including things like a small territory size, high concentration of distribution networks, and representation of a large number of customers • No information on the remaining 7 outliers

Finally, we ran the regressions again using two methods to exclude outliers: 1) without the lowest and highest 1% of SAIDI (SAIDI) values (i.e., the 1%/99% exclusion method) and 2) without the extreme and influential outliers we identified in the Studentized and Cook's statistical analysis discussed. We found that the regression results did not significantly change when the outliers were removed according to these two methods. Table C-2 is a summary of the effects on the regression results after removing outliers using two different methods.

Table C- 2. Excluding Outliers and their Effect on the Pooled Regression Results

Pooled Regression (i.e., No Utility or Time Effects)	Outlier Exclusion Method	Effect of Excluding Outliers on Regression Results
SAIDI without major events included	1%/99% Method	R^2 slightly increased from 0.04 to 0.05; No sign changes on regressors; No regressors lost or gained significance at the 10% level.
SAIDI without major events included	Cook's and Studentized Residual Tests	R^2 slightly increased from 0.04 to 0.06; No sign changes on regressors; No regressors lost or gained significance at the 10% level.
SAIFI without major events included	1%/99% Method	R^2 slightly increased.; After outlier exclusion, heating degree-days (HDD) became significant at the 10% level and year became marginally insignificant at the 10% level (p=0.109); Post OMS regressor was marginally significant at 10% level before excluding the outliers (p=.106), but definitely not significant after the outliers were removed (p=0.43).
SAIFI without major events included	Cook's and Studentized Residual Tests	R^2 slightly increased from 0.07 to 0.09. No sign changes on regressors. After outliers were removed, heating degree-days (HDD) became significant at the 10% level. No other regressors lost or gained significance at the 10% level.

As a result of these three steps, we concluded that the outliers identified in our statistical analysis are valid observations. We also determined that their removal did not significantly affect the pooled regression results. For these reasons, we chose not to remove any of these outliers from the statistical regression analysis presented in the main body of the report.

Appendix D. Detailed Results from Regression Analysis

Table D- 1. One-way Fixed Effects Regression for the Effect of Sales, HDD, CDD, time, and OMS on Frequency and Duration of Grid Disruptions

	With Major Events				Without Major Events			
	I		**II**		**III**		**IV**	
	ln SAIFI		**ln SAIDI**		**ln SAIFI**		**ln SAIDI**	
Intercept	-61.9241	***	-123.951	***	-45.221	***	-46.4962	***
	(19.3904)		*(38.0343)*		*(10.2925)*		*(11.5011)*	
Sales	0.00295		0.006712		0.001054		0.002057	
	(0.0044)		*(0.0087)*		*(0.0036)*		*(0.0041)*	
HDD	2.109E-06		0.000028		-0.00002		-0.00004	
	(0.0001)		*(0.0001)*		*(0.0000)*		*(0.0000)*	
CDD	5.867E-06		-0.00005		0.000182	*	0.000236	**
	(0.0001)		*(0.0003)*		*(0.0001)*		*(0.0001)*	
YR	0.030683	***	0.063937	***	0.022749	***	0.025635	***
	(0.0097)		*(0.0190)*		*(0.0051)*		*(0.0057)*	
OMS	0.055153		0.381655	***	-0.02486		0.173814	***
	(0.0636)		*(0.1248)*		*(0.0464)*		*(0.0517)*	
POST OMS	-0.02157		-0.03947		-0.0128		-0.01033	
	(0.0133)		*(0.0262)*		*(0.0092)*		*(0.0103)*	
Utility Effects	Yes		Yes		Yes		Yes	
R^2	0.59		0.47		0.71		0.69	

Notes: Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table D- 2. Two-way Fixed Effects Regression for the Effect of Sales, HDD, CDD, and OMS on Frequency and Duration of Grid Disruptions

	With Major Events				Without Major Events			
	I		II		III		IV	
	ln SAIFI		ln SAIDI		ln SAIFI		ln SAIDI	
Intercept	-0.41598		4.50384	***	0.542495		4.874969	***
	(0.3332)		*(0.6499)*		*(0.3664)*		*(0.4091)*	
Sales	0.003169		0.007662		0.001042		0.001522	
	(0.0044)		*(0.0085)*		*(0.0037)*		*(0.0041)*	
HDD	0.000041		-5.51E-06		-0.00002		-0.00003	
	(0.0001)		*(0.0001)*		*(0.0001)*		*(0.0001)*	
CDD	-0.00016		-0.00035		0.000051		0.000215	*
	(0.0002)		*(0.0003)*		*(0.0001)*		*(0.0001)*	
POST OMS	-0.01241		-0.02826		-0.00842		-0.00889	
	(0.0133)		*(0.0260)*		*(0.0093)*		*(0.0104)*	
Year 1	-0.21793	**	-0.33234	*	-0.16384	***	-0.13058	**
	(0.1028)		*(0.2005)*		*(0.0585)*		*(0.0654)*	
Year 2	-0.08249		-0.29269		-0.13054	**	-0.15683	**
	(0.0980)		*(0.1912)*		*(0.0573)*		*(0.0640)*	
Year 3	0.050256		0.050184		-0.04065		-0.11338	*
	(0.0897)		*(0.1750)*		*(0.0556)*		*(0.0621)*	
Year 4	0.085979		0.130981		-0.01804		-0.08814	
	(0.0841)		*(0.1640)*		*(0.0514)*		*(0.0573)*	
Year 5	0.038003		-0.06511		-0.03443		-0.01921	
	(0.0804)		*(0.1569)*		*(0.0496)*		*(0.0552)*	
Year 6	0.093184		0.027743		0.045398		0.018774	
	(0.0801)		*(0.1562)*		*(0.0519)*		*(0.0580)*	
Year 7	0.193096	**	0.296795	*	0.062783		0.039727	
	(0.0835)		*(0.1629)*		*(0.0559)*		*(0.0625)*	
Year 8	0.133398	*	0.215998		0.025579		-0.01463	
	(0.0715)		*(0.1394)*		*(0.0476)*		*(0.0531)*	
Year 9	0.232577	***	0.572835	***	0.100408	**	0.138323	***
	(0.0648)		*(0.1264)*		*(0.0436)*		*(0.0488)*	
Utility Effects	Yes		Yes		Yes		Yes	
R^2	0.61		0.50		0.71		0.70	

Notes: Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01