Understanding Bulk Power Reliability:
The Importance of Good Data and A Critical Review of Existing Sources

Emily Fisher
Lawrence Berkeley National
Laboratory
esfisher@lbl.gov

Joseph H. Eto
Lawrence Berkeley National
Laboratory
jheto@lbl.gov

Kristina Hamachi LaCommare
Lawrence Berkeley National
Laboratory
kshamachi@lbl.gov

Abstract

Bulk power system reliability is of critical importance to the electricity sector. Complete and accurate information on events affecting the bulk power system is essential for assessing trends and efforts to maintain or improve reliability. Yet, current sources of this information were not designed with these uses in mind. They were designed, instead, to support real-time emergency notification to industry and government first-responders. This paper reviews information currently collected by both industry and government sources for this purpose and assesses factors that might affect their usefulness in supporting the academic literature that has relied upon them to draw conclusions about the reliability of the US electric power system.

1. Introduction

The reliability of the bulk power system is of critical importance to the electricity system in the U.S. because it relies on high-voltage power transportation. Understanding the reliability of the bulk power system, how to measure it, and how that reliability is changing over time, is important for developing and assessing technologies and policies that seek to affect reliability.

Meaningful assessment of reliability depends on comprehensive and accurate data. Currently, there are two major data collection efforts that focus on the reliability of the bulk power system in the U.S. The Department of Energy (DOE) collects data through the Electric Emergency Incident and Disturbance Report (via Form OE-417). The North American Electric Reliability Corporation (NERC) collects data through the reporting requirements contained in the Disturbance Reporting reliability standard, EOP-004, which is known as the Interconnection Reliability Operating Limit and Preliminary Disturbance Report.

Many authors have used these data to assess characteristics of and longitudinal trends in bulk power system reliability. By and large, these authors take the completeness and accuracy of these data as given. Yet both sources of data were developed originally to support emergency notification. It is not evident from their design and implementation that they were intended to serve as comprehensive repositories of definitive information on bulk power system reliability. Closer examination of these data for completeness and accuracy, therefore, is warranted to help us understand what kinds of analysis can be performed using these data, and evaluate the claims made by authors who use the data to support their analysis. Examining current reliability datasets can also indicate potential changes in future data collections that may better support reliability analysis.

The objective of this paper is to begin exploring what can and cannot be learned from analyzing the two current sources of data on the reliability of the bulk power system. To do this, we first review the history, purpose, and current reporting requirements for the DOE and NERC reliability data. In the third section, we examine the completeness and accuracy of DOE and NERC data, first by comparing them to one other, and second by comparing the NERC data to more detailed investigations that were subsequently conducted by NERC on a handful of events. We then, in the fourth section, review the literature that has relied on these data and use our findings to draw preliminary conclusions regarding the claims that have been made in the literature. In the fifth section, we outline possible directions for improving the comprehensiveness and accuracy of data on the reliability of the bulk power system.

2. Current DOE and NERC bulk power system reliability emergency reporting

The DOE began collecting data on the reliability of the bulk power system through the EIA-417 form in the late 1970's. At that time the Energy Information Administration (EIA) administered the form. When the DOE's Office of Electricity Delivery and Energy Reliability (OE) was established in 2003, the office, among other things, took over collection and validation of this bulk power system reliability data and renamed the form OE-417.

The purpose of the form is "to meet [DOE's] overall national security and Federal Emergency Management Agency's National Response Framework responsibilities," by "obtain[ing] current information regarding emergency situations on U.S. electric energy supply systems." [1] Another way to define the form's purpose is to alert the Federal government when the bulk power system has transitioned from a normal state of operation to an emergency state of operation. The information submitted on the form is circulated to various government agencies and to the public. For example, the data are used in the OE Spot Report, which is circulated only to other government agencies, and the OE Situational Report, which is publically available. Annual data is also available publically on the OE website.[1]

Entities that are required to submit OE-417 include reliability coordinators, balancing authorities, utilities and some generating entities. An initial report is required either one or six hours following a triggering event (described below as reporting criteria) and a final report is due 48 hours following the triggering event.[2] OE staff take a variety of measures to improve the quality and completeness of the reporting, including contacting respondents to quality-check data and utilities that staff learn have experienced an event but have not filed an OE-417 form. According to OE staff, data posted on the OE website are the best available, developed from the initial and final reports as well as the follow-up work by OE staff. However, staff also note that there are limited resources available for enforcement of the reporting requirements and do not know how many events go unreported.

Historical changes in the requirements and procedures for filing have produced discontinuous jumps in the number of events reported. For instance, in the 1990's municipal utilities and cooperatives were added to the required respondents, resulting in an increase in respondents from around 50 to around 200.[3] According to OE staff, when NERC instituted mandatory reliability reporting in 2007, there was a large increase in reported events to DOE.[4]

Reporting criteria (i.e., triggering events) for filing form OE-417 fall into four main areas: intentional attacks on the system, system failures (e.g., blackout) or voltage events, load loss and/or emergency reduction, and fuel supply emergencies. For the more severe or serious events—actual physical and cyber attacks resulting in interruptions, impacts to critical infrastructure or operations; complete operational failure; separation of the system (i.e., islanding); uncontrolled firm load interruption of 300 MW or more; emergency load shedding of 100 MW or more; system-wide voltage reductions; and public appeals to reduce electricity use—initial reporting is due within one hour of the start of the event. For the less severe events, including suspected attacks, loss of electric service to 50,000 customers for one hour or more, and fuel supply emergencies, initial reporting is due within six hours of the start of the event. Any event reported on form OE-417 to DOE must also be submitted to NERC, under the NERC Disturbance Reporting reliability requirement, EOP-004.

NERC began collecting reliability data in 1984 through System Disturbance Reports. In 2007, submission of bulk power events became mandatory through the Disturbance Reporting requirement, EOP-004. [2] The NERC EOP-004 is one of many reporting and reliability requirements that were developed by NERC in their capacity as the Electric Reliability Organization.

The purpose of NERC EOP-004 is to make reliability event data available for study and analysis in order to avoid similar problems in the future. More types of entities are subject to the reporting requirements of EOP-004 than are required to submit form OE-417. In addition to the reliability coordinators, balancing authorities, utilities (load serving entities), and some generators that are required to submit form OE-417, NERC's reporting requirements also apply to transmission operators, all generator owners, and regional reliability organizations.

EOP-004 is mandatory and requires that all events reported to DOE under OE-417 must also be reported to NERC. In addition to the DOE criteria for reporting the following types of events also trigger mandatory EOP-004 reporting: loss of generation (2,000 MW or more in the Eastern or Western

Interconnection, 1,000 MW or more in ERCOT); loss of other components that result in loss of 300 MW of firm load (or a lesser amount for small entities), modification of operating procedure, under-frequency or under-voltage situations, or other serious outcomes; firm load shedding of 100 MW or more to maintain continuity of the bulk power system; system islanding; and any action resulting in voltage excursions, component damage, or interference with automated system protection. For most of these events a preliminary report must be submitted to NERC within 24 hours. A final report is due 60 days after the event. Event data reported under EOP-004 are available in annual files on the NERC website, beginning with 1992 data.[5] Year 2008 and 2010 data are being reviewed and processed by NERC, and, as of August 2011, have not been released publically.

OE-417 and the NERC Disturbance Report Form request nearly identical information, including a description of the triggering event, date and time of event and of restoration, whether the event originated in the reporting entity's system, and amount of load and number of customers interrupted. In fact, since 2007, DOE has allowed submission of the NERC Disturbance Report Form in lieu of OE-417. NERC requires submission of form OE-417 as part of the reporting requirements of standard EOP-004.[6]

There are, however, some slight differences in the reports. For instance the NERC Disturbance Report Form asks for system frequency just prior and following the event, and OE-417 does not. OE-417 asks for high voltage substations lost during the event, and the NERC Disturbance Report Form does not.

By design, the intent of these forms is to collect information on reliability events, but not necessarily the exact same events. At the same time, these datasets may not be complete or accurate, in the sense that events that should have been reported to DOE or NERC may not have been, or events that were reported may contain information that does not exactly reflect what occurred during the event. Researchers need to understand what these datasets contain and how complete and accurate they are before using them in reliability analysis. In the following section, we begin this examination.

3. Analysis of NERC and DOE reliability data

On the face of it, various factors suggest neither the DOE nor the NERC dataset presents a complete or fully accurate picture of bulk power system reliability. From the standpoint of completeness, data collection by NERC has not always been mandatory. Requirements for who must file the OE-417 have changed over time. While OE-417 reporting is mandatory, and DOE staff takes measures to contact utilities that may not have filed a report when required, DOE does not have the resources to enforce compliance strictly. NERC Disturbance Report data from 2008 have not been released publically, and (at the time this paper was written in the Summer of 2011) no anticipated release date is set.

From the standpoint of accuracy, the principal factor is the short reporting time frame (although efforts are made to validate information and release publically only what is considered to be the most accurate). Considering, for instance, the number of customers that experience an interruption, in the absence of automated outage management systems (which may or may not be the sources of information reported on the forms), or without an in-depth post-event investigation by the respondent, inaccuracy is possible.

In this section we present our analysis of bulk power system reliability data collected by NERC and DOE from 2000 to 2009 (excluding 2008 because NERC has not released data from that year). We focus only on events that meet DOE's criteria of 50,000 or more customers interrupted and/or 300 MW of firm load shedding. Since 2007, all events reported to DOE under OE-417 criteria were required to be submitted to NERC as well; before 2007, while loss of 50,000 customers was not a specific criterion for NERC reporting, an event of that magnitude would likely meet other NERC criteria.

First we compare the NERC and DOE data to one another and show evidence that the datasets are incomplete and potentially inaccurate. We do this in two ways: first by identifying how often events are only reported to one institution or when the same event has different information reported describing the event; and second, by using each dataset to calculate common reliability metrics. Then we compare NERC emergency disturbance reporting with more detailed analyses for several events and discuss the types of corrections that were made following these more detailed post-event investigations.

The goal of our analysis is to determine whether the datasets contain accurate and complete

[5] http://www.nerc.com/page.php?cid=5|66 (Accessed August 2011)
[6] For events that fall under the NERC EOP-004 standard but do not trigger OE-417 reporting, the NERC Disturbance Report Form must be submitted to NERC

information. If the same event is reported with different load loss information, this is an indication that at least one of the datasets is not accurate. If an event is reported in one form but not another when it meets the criteria of both, it is an indication that the datasets are not complete. If the datasets are not complete, some reliability events are missing from the record (and from any subsequent analysis).

3.1. Comparison of events reported to NERC and DOE

To get a preliminary sense for the completeness and accuracy of the two datasets, we compare the reported events to determine if they were reported to both NERC and DOE, and whether the events were reported in the same way to both. First, we identify where NERC and DOE records referred to the same event, based on comparing event start date, and utility name or location. When there seems to be a match, we determine whether the event parameters (start date and time, restoration date and time, utility, number of customers interrupted and customer load lost) are consistent between the two reports. If all seven parameters match, the two event records are considered a "Match." Where at least one of these parameters is not the same, they are considered a "Match with differences." This includes differences that may have been a result of input error (either by the respondent or by DOE/NERC staff that entered the data into their respective databases). For instance, for one event the number of customers was reported as 159,019 to NERC and 157,019 to DOE. In a similar example, the number of customers interrupted was reported as 8,110 to NERC and 88,110 to DOE. In other cases, the event start times differ by 12-hours, which may be a result of incorrect interpretation of a 24-hour clock. Because it is difficult to determine whether there was an error in reporting or an intentional difference, these reports were considered "Match with differences."

For this analysis we exclude events that took place outside the continental US, because reporting requirements are not consistent.[7] In addition, we consider only events with 50,000 or more customers interrupted and/or loss of 300 MW or more of firm load. We apply this screening after we matched events, so in some cases an event reported to one institution does not meet the screening criteria but is a match to an event reported to the other institution that did. Our goal is to include as many reported events we could reasonably justify should be reported

to both DOE and NERC. If an event meets the criteria as reported to one of the institutions it is reasonable to assume the event should be reported to other institution.

This variation in reporting of the same event to NERC and DOE suggests that in some cases where an event is reported to one institution and does not meet our screening criteria, it may not meet the screen simply because the event parameters were entered incorrectly. As a result our screening strategy may remove significant events that should have been reported to both NERC and DOE because the number of customers interrupted or loss of load was not reported. We took this conservative approach in order to not over-state the mismatch between the datasets; in light of this, the "No Match" values below may be higher than we present, suggesting the datasets may be less complete than appears in this analysis.

Figure 1 summarizes the matching of data between NERC and DOE reports. The number of loss of load events reported to both institutions grows over time, from one or two events in the early 2000's, to 22 in 2009. However, in each year there are some events reported to NERC that are not reported to DOE, even though they meet the DOE reporting criteria. Also, in 2007 and 2009 when all OE-417 reports are required to be submitted to NERC, some events are not. This suggests each dataset is missing some events, and therefore both are incomplete. A careful look at 2009 data reveals three events in the NERC dataset with more than 50,000 customers interrupted not found in the DOE dataset. Similarly, nine events reported to DOE were not found in the EOP-004 dataset; there should have been none, because as of 2007 all OE-417 events are required to be reported to NERC.
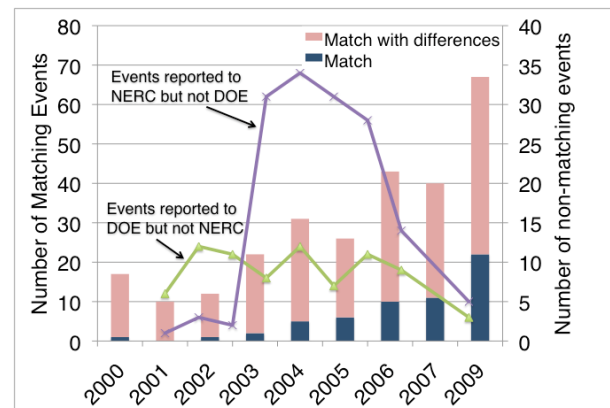

Figure 1. Matching of event reported to NERC and DOE

This matching exercise also provides insight on whether the datasets are accurate. As Figure 1

---

[7] Some Canadian and Mexican utilities report to NERC but not DOE; Puerto Rico utilities report to DOE but not NERC.

illustrates, the number of events that fall under the "Match with differences" is not decreasing over time. This means there are still many reports of events in the DOE and NERC datasets that are inconsistent in one of the important parameters: number of customers interrupted, amount of load lost, or length of interruption. These are critical parameters for analyzing reliability, and if they are inconsistently reported, it means at least one of the datasets contains inaccuracies.

3.2. Comparison of bulk power system SAIFI and SAIDI using NERC and DOE reliability data

One important measure of reliability is how often an end-use customer experiences power interruptions, and what the duration of those interruptions are. To determine whether the DOE and NERC reliability datasets are telling a consistent story about customer interruptions, we calculate standard metrics used to report distribution system reliability for each dataset and compare them. We do not seek to draw conclusions about the reliability of the bulk power system or how it is changing over time, but rather to examine whether it is appropriate to use the existing NERC or DOE data to examine customer interruptions.

System Average Interruption Frequency Index (SAIFI) and System Average Interruption Duration Index (SAIDI) are two common metrics of reliability of power systems. [3] SAIFI measures how many times - on average - an end-use customer experiences a power interruption over a given period of time. SAIDI measures how long the average customer is without power over a given period of time. The equations to calculate SAIFI and SAIDI are as follows:

$$\text{SAIFI} = \frac{\sum \text{Total Number of Customers Interrupted}}{\text{Total Number of Customers Served}}$$

$$\text{SAIDI} = \frac{\sum \text{Customer Interruption Durations}}{\text{Total Number of Customers Served}}$$

We calculate annual SAIFI by summing all events in the NERC and DOE datasets that have a value for number of customers interrupted. We obtain the total number of customers in the country for that year (taken from EIA form 861).

If NERC and DOE reliability data are consistent and accurate with respect to customer interruptions originating on the bulk power system, the annual SAIFI would be the same. However, as can be seen in Figure 2, they are different. There is as much as a

54 percent difference between the two SAIFI values in a given year. This difference is due to two factors: the two datasets contain at least some different events, and a different number of interrupted customers reported for matched events. The difference in SAIFI cannot be explained by a consistent bias in the datasets, e.g., one dataset being more "comprehensive" than the other, as suggested in the literature [4], because in some years SAIFI is higher for one dataset and in other years it is lower.
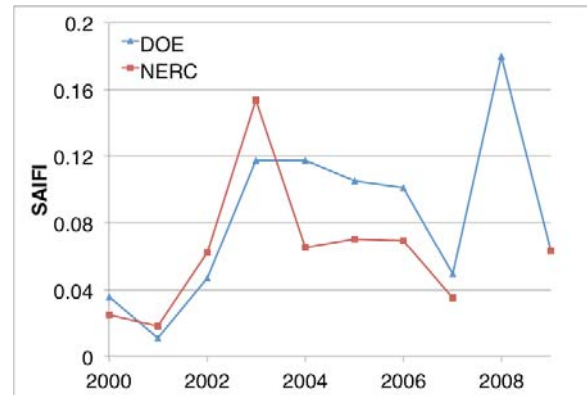

Figure 2. Annual SAIFI calculated from NERC and DOE bulk power system reliability data

We calculate SAIDI with all events in the NERC and DOE datasets for which we can determine the event duration; in other words, all events with interruption start and end date and time. However, not all customers are without power for the entire time between the start and end of an event. As the utility responds to the event, power is restored sequentially to affected customers, until all customers are completely restored and the event is considered over. We estimate this function, customers restored as a function of event duration, in two ways: linear restoration and exponential restoration. Linear restoration assumes there is a constant rate at which service is restored to customers.

$$\text{Customer Mins Interrupted (linear)} = \frac{[\text{\# Customers Affected}] * [\text{Mins of Interruption}]}{2}$$

Exponential restoration assumes customers are restored slowly at first, but at an increasing rate.

$$\text{Customer Mins Interrupted (exp)} = \frac{[\text{\#Customers Affected}]}{\{[-\ln(0.01)]/[\text{Mins of Interruption}]\}}$$

We calculated SAIDI using both methods, and the results are shown in Figure 3.

The linear model of customer interruption minutes always results in a higher calculated SAIDI than that calculated with the exponential model; thus these two models can be seen as reasonable bookends. These are not the only two methods for estimating how long each customer was without power, but they are adequate for our purposes. For this analysis we are interested in the comparison of SAIDIs calculated with the NERC and DOE datasets, not the actual values of SAIDI resulting from the calculated restoration time.

As with the SAIFI comparison, if NERC and DOE reliability data are complete and accurate, the annual SAIDIs should be the same for a given year and customer restoration methodology. However, as can be seen in Figure 3, they are not. This disparity arises because there is a different number of events in each dataset as well as variation in the number of reported customers interrupted and duration for the matching events. As with the SAIFI analysis, we do not observe a consistent bias in the data. In some years DOE data results in a higher SAIDI, in other years NERC does. The percent difference between the two SAIDI values is as high as 192 percent.
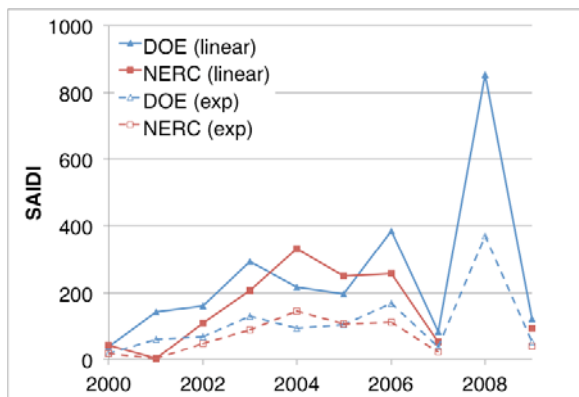

Figure 3. Annual SAIDI calculated from NERC and DOE bulk power system reliability data

3.3. Selected comparison of NERC emergency and power-event analysis event reports

From the standpoint of assessing trends in or impacts of technologies and policies on bulk power system reliability, one weakness of the reliability data collected by NERC and DOE is that they must be reported very quickly after an event begins. There is not enough time in the reporting schedule to allow for detailed analysis of the causes or extent of the event; as noted earlier, subsequent analysis of these topics is not the purpose these reports were intended to serve. Indeed, submitting a report containing hurriedly estimated values can be challenging enough when

utility personnel are primarily, and justifiably, concerned with restoring power to customers.

Conducting post-event analysis does not guarantee complete or more accurate values for event descriptors like number of customers interrupted or amount of load lost. However, if these values are revised as a result of post-event analysis, that must mean there was evidence that the original estimates could be improved. To explore the kinds of changes that can be made to an event record when there is time to do a detailed analysis, we compare emergency reports with detailed analysis performed by NERC's Disturbance Analysis Working Group for 13 events from 2000 to 2002 (the most recent year these detailed analyses are available).[8] We compare the following parameters from the two reports: number of customers, time event began, and time of restoration. Not every event has data in each category.

Table 1. NERC emergency and detailed disturbance reports: number of customers

Event	Date	# Customers	
		System Disturbance Report	DAWG Event Analysis
Public Service Co. NM Fire	3/18/00	600,000	660,000
TransEnergie short circuit	1/17/01	234,000	284,671
Saskatchewan Power Co. generation separation	3/10/01	246,000	248,588
ConEd NY - Terrorist Attack	9/11/01	12,000	13,000
WECC-AZNMSNV contractor accident	8/2/02	350,000	348,360
WECC-NWPP severe weather	12/26/02	Not reported	140,000

The number of customers interrupted is reported in five emergency reports and seven detailed reports. The difference in these values range from 1,640 (a

---

[8] Two events with both emergency and detailed NERC reports were not included here because the emergency reports had very little information (e.g., no amount of generation lost, no customers interrupted).

near-zero percent change) to 50,671 (a 22 percent change). Number of customers lost is an input to both the SAIDI and SAIFI calculations, so inaccuracies in this number can affect calculation of these metrics.

Table 2. NERC emergency and detailed disturbance reports: time start

Event	Date	Time-Start	
		System Disturbance Report	DAWG Event Analysis
Public Service Co. NM Fire	3/18/00	15:57	15:50
VA Power Co. Substation Fire	4/1/00	16:47	16:46
NE WI, Up MI Peninsula Islanding	11/14/01	6:35	6:36
Jacksonville Electric Authority equipment failure	4/29/02	16:21	15:50

Table 3. NERC emergency and detailed disturbance reports: time restored

Event	Date	Time-Restored	
		System Disturbance Report	DAWG Event Analysis
New Brunswick Power Salt and Rain	12/20/00	17:28	16:59
TransEnergie short circuit	1/17/01	20:13	19:44
WECC line fault	8/1/01		6:44
Jacksonville Electric Authority equipment failure	4/29/02	4/29/02 23:59	4/30/02 1:30
MAIN - gen loss	7/26/02	19:00	16:33
WECC-AZNMSNV contractor accident	8/2/02	Not reported	9:54
WECC-NWPP severe weather	12/26/02	Not reported	15:13

Event start time is reported in emergency and detailed reports for four of the events. Of these, the difference in start time ranged from 1 minute to 31 minutes. Restoration time is reported in four emergency reports and seven detailed reports. The differences in restoration times range from 29 minutes to 147 minutes. Revision of event duration between the emergency and detailed reports can be calculated for six events. These revisions range from a change of one minute (a 0.3 percent change in duration) to 147 minutes (67 percent difference in duration). Interruption duration is an input to SAIDI, so a revision in this value would change that metric.

The differences identified here between events reported on emergency reports versus a more detailed analysis are significant in some cases. For all events we looked at, the values for customers interrupted, start time and/or restoration time were revised, at least to some degree, upon more detailed investigation. This indicates that emergency reporting, while important for emergency response, may not be appropriate as an archival record.

3.4. Summary of analysis

This section presented our analysis of the NERC and DOE datasets and found they are not complete nor fully accurate from an archival perspective.

To some degree, the NERC and DOE reliability data are intentionally not complete: they include events that meet the applicable reporting criteria and exclude events that do not. However, as our comparison in section 3.1 indicates, events that should have been reported to both institutions in some cases were not. Over the time period we examine, the number of events missing from the DOE data are as high as 34 (in 2004); events missing from NERC data are as high as 12 (in 2001 and 2004). In section 3.2, we find that reliability metrics derived from the two datasets are inconsistent, in part because each contains at least some distinct reliability events not found in the other. In addition, policy shifts (e.g., expanding the pool of required respondents) and anecdotal evidence (e.g., noted increase in DOE submissions after NERC reporting was made mandatory) suggest the datasets do not contain a complete record of reliability events. Additionally, NERC has not released data from some periods of time, including 2008 and 2010.

Our analysis also finds the NERC and DOE reliability data are not fully accurate either. In many cases, reports of the same event submitted to NERC and DOE contain different number of customers interrupted, start and restoration times. We also find

events where these parameters filed on emergency reports are revised after more detailed analysis. Because these event parameters are critical for analyzing reliability (in particular for calculating customer interruption metrics) using data that are not accurate may produce unreliable or misleading results.

However, there are many examples in the literature of reliability analysis conducted with these data. In the following section we examine this literature and assess the possible impact of incomplete or inaccurate source data may have on the findings.

4. Implications of our findings for the literature

Many researchers use NERC reliability data, sometimes in concert with DOE data, to analyze the reliability of the bulk power system. We review their findings by organizing them under two categories: ones that conduct straightforward time-series or simple statistical analysis as evidence of bulk power system reliability, and ones that examine the nature of interruptions in the system in order to identify an explanatory or descriptive model.

In the first category, several publications have used NERC and DOE reliability data as evidence that the bulk power system is becoming less reliable. [4-6] Many point to the increase from 1991 to 2005 in the *reported* number of significant interruptions and interruptions affecting more than 50,000 customers as evidence. [7] claims that patterns in NERC reliability data is evidence NERC's new reliability requirements have not improved reliability "yet." [8] shows NERC data (1984-2000) in cumulative probability of transmission and generation-related failures as motivation for their paper, encouraging focus on continuation of services in a blackout rather than only aiming to reduce or eliminate blackouts. [9] constructs time series-based statistical analysis with NERC data from 1990-2004 as input to risk management policies. The authors regress against time and customers to identify what factors affect the length of an interruption and how many customers are affected.

[10] conducts a statistical analysis of NERC interruption data. The authors make several claims: the frequency of large blackouts has not decreased between 1984 and 2006, there may be correlation between interruptions and season and time of day, there has been an increase in peak-hour and peak-season interruptions, there is evidence of a power-law relationship between interruption size and frequency (in support of SOC), and there is no correlation between interruption size and duration. The authors explain how they "scrub" the data, by focusing on only large events; controlling for demand growth, supply shortages and extreme natural events; and extrapolating incomplete data entries. Even still they make the assumption that, "all of the significant cascading failures are included in the NERC records, as required by regulation." [10]

Our findings from section 3—that the source data are incomplete and inaccurate—have the most potential ramifications for the time-series analysis since the consistency of reporting over time is fundamental to that kind of effort. Any finding based on a time-series analysis of the NERC or DOE reliability data is likely based on incomplete data, which may embed a bias in the analysis. The frequency of large blackouts could be quite different than what is evident in the data, if events have been underreported. Also, and more importantly, if the rate of reporting has increased (because of policy or other changes), a seeming increase in interruptions over time may be the result of improved reporting, not degrading reliability. Several papers conclude that the frequency of interruptions has increased over time, between the mid 1980s and early 1990s to the 2000's [4-9]. In these papers, the authors point to the increase in blackout frequency in the NERC System Disturbances Reports, without questioning whether this dataset is complete.

Findings or observations based on simple statistical analysis (e.g., mean, median) would potentially change if the dataset were more complete and accurate. For instance, [9] observes that the median customer loss for U.S. and Canadian interruptions between 1990 and 2004 were, respectively, 67,765 and one. This vast difference could in part be explained by the different populations in the countries, but most likely there is under-reporting of events as well. However, it is impossible to hypothesize whether the statistical measures would be higher or lower if a more accurate and complete dataset were available.

In the second category, there have been several papers examining whether the bulk power system displays characteristics of self-organized criticality (SOC). Systems that display SOC have non-linear dynamics and tend to be close to major disruption even when in an average system state. The SOC literature relies in part on time-series analysis (the implications of which are discussed above) but also on analysis of frequency and magnitude of interruptions.

[11] and [12] both analyze a time series of NERC reliability data for evidence of SOC in bulk power interruptions. The authors of these papers

examine long-term correlation of interruptions for evidence of SOC characteristics. These authors recognized that "[i]t is not clear how complete this [NERC] data is, but it is the best-documented source that we have found for blackouts in the North American power transmission system." [12]

[13] presents a similar analysis as [11], but with an extended time-frame of data. [14] also examines SOC of interruptions, but arrived at a different conclusion—that SOC is not an appropriate model for power systems—using the same NERC data.

Table 4. Literature that uses NERC or DOE emergency disturbance data

Paper	Type of Analysis	Dataset
Carreras 2000 [11]	SOC	NERC 1993-1998
Chen 2001 [13]	SOC	NERC 1984-1999
Talkudar 2003 [8]	Time-series	NERC 1984-2000
Carreras 2004 [12]	SOC	NERC 1984-1998
Weron 2005 [14]	SOC	NERC 1984-2000
Amin 2007 [5]	Time-series	NERC 1991-2005
Simonoff 2007 [9]	Time-series Simple statistical	NERC 1990-2004
Amin 2008 [4]	Time-series	NERC 1991-2006 DOE 2001-2005
Lyons 2008 [7]	Time-series	NERC 2002-2007
Hines 2009 [10]	Time-series SOC	NERC 1984-2006
Amin 2011 [6]	Time-series	NERC 1995-2005 DOE 2001-2005

The SOC studies include some time-series analysis, specifically investigation of time between interruptions [10-14]. For SOC systems, probability of an interruption increases exponentially with the waiting time since the last interruption. Some authors found exponential decay in the inverse probability of waiting time [11-13], and some did not [14]. This type of analysis depends on the completeness of the dataset, and incomplete interruption data could affect the accuracy of the findings. If the data are incomplete, time between interruptions could appear longer than in a complete dataset, which suggests a quicker drop-off in the probability may better reflect reality. This may strengthen the argument for SOC in bulk power systems.

The SOC literature includes other kinds of analyses, including examination of the probability of

interrupted customers, and frequency of interruptions based on number of customers affected [11-14]. Analysis of interruption frequency may be skewed by an incomplete dataset: a more complete dataset would contain more events, and therefore a higher frequency of events. However, it is not clear what the size of these missing (or incorrectly reported) events are, and, hence, how their inclusion would change the finding of SOC characteristics.

Our analysis of existing literature is a preliminary assessment. We cannot say with certainty whether and how the findings would change with a different dataset. This is merely a sketch of why having accurate and complete data is important for analysis of bulk power system reliability.

5. Suggestions for improving the completeness and accuracy of bulk power system reliability data

Because statistical analysis of bulk power system reliability based on complete and accurate data is important, changes in the current data collection could provide much value to the electricity sector. In addition to the existing emergency reporting, which is vital for immediate response to events, we recommend an additional mandatory reporting mechanism that would collect final information about reliability events. This would provide a solid basis for analysis, for improved policy-making, and, as a result, for a more reliable bulk power system. In addition, having multiple sources of reliability data containing potentially redundant information is useful because they would offer researchers a way to validate, or at least calibrate, the results of reliability analysis from one source against another.

One way to collect more complete and accurate data would be to expand the existing emergency reliability reporting structure, but require an additional "archival" report with a much longer reporting time (on the order of weeks not hours). As evident in the examination of NERC detailed reports in section 3.3, when more time is allowed and a careful investigation is undertaken, it is possible to correct and update the information on number of customers interrupted, start time and duration, and possibly add new information not available at the time of the initial report. This reporting could be similar to or perhaps a part of NERC new Event Analysis initiative.[9]

---

[9] In October 2010 NERC began a pilot of this initiative, which requires either a Brief or Event Analysis report on all events within 10 to 120 days. The template for the Phase 1 and Phase 2 field trials includes customer interruptions, as well as other information.

Another possible way to develop a statistically-valid reliability dataset would be to require SAIDI and SAIFI reporting from all utilities or other service providers that serve end-use customers. Reliability metrics that measure customer interruptions can be satisfactorily calculated with data from these entities alone, since the focus here is on reliability events that result ultimately in interruptions of power to end-use customers. Requiring a report from service providers of annual SAIDI and SAIFI, distinguishing between the origin of the events they report on (e.g., bulk power versus distribution system)**.** This would have the added benefit of avoiding the double-reporting that can arise when multiple layers of entities report, as happens often with existing NERC and DOE emergency reliability reporting. This would, however, not be adequate for the NERC and DOE emergency reporting that track many kinds of reliability events, not just ones that result in customer interruptions.

The frequency, detail and granularity of whatever new reliability reporting is developed needs to support the goal of the desired analysis. As noted above, one important measure of bulk power system reliability is number, frequency and duration of average customer interruptions. For that measure of reliability, an annual SAIDI and SAIFI value reported by utilities would be sufficient. However, for the SOC analysis reviewed above, researchers need to know the average length of time between outages—information not conveyed in annual SAIDI and SAIFI.

6. Conclusion

Reliability of the bulk power system is of vital importance to the economic and general welfare of the U.S. Knowing how reliable the system is, and how that reliability changes over time, is critical for developing and assessing technologies and policies that affect reliability. Currently, however, we lack a source of reliability data that is complete and accurate, and appropriate for use in analysis of trends or impacts of policies and technologies.

In this paper we begin to assess the bulk power system emergency reliability reporting to NERC and to DOE. We find evidence that these two datasets may not be accurate and complete. This places in question results of reliability analysis conducted with these data. We critically review the existing literature that analyzes these data, and hypothesize how the inaccuracy and incomplete nature of the data may affect the findings. Finally, we present some thoughts on what kind of data collection activity may improve the available reliability data.

8. References

[1] DOE, "OE-417 Electric Emergency Incident and Disturbance Report", revised January 2011.
[2] NERC, "Standard EOP-004-1 --- Disturbance Reporting," January 2007.
[3] 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 Electrical and Electronics Engineers, Inc. May 14. 35 pages.
[4] S.M. Amin, "For the Good of the Grid", IEEE Spectrum Magazine, 6(6), November/December 2008, pp 48-59.
[5] S.M. Amin and P. F. Schewe, "Preventing Blackouts", Scientific American, May 2007, pp. 60-67.
[6] S.M. Amin, "U.S. Electrical Grid Gets Less Reliable", IEEE Spectrum, January 2011.
[7] C. Lyons, J. Jacobi, and R. Starkweather. "NERC Standards and Standards Compliance: Still a Work in Progress?", The Electricity Journal, 21(3), April 2008, pp. 29-39.
[8] S.N. Talkudar, J. Apt, M. Ilic, L.B. Lave, M.G. Morgan, "Cascading Failures: Survival versus Prevention", The Electricity Journal, 16(9), November 2003, pp. 25-31.
[9] J.S. Simonoff, C.E. Restrepo, R. Zimmerman, "Risk-Management and Risk-Analysis-Based Decision Tools for Attacks on Electric Power", Risk Analysis, 27(3), June 2007, pp. 547-570.
[10] P. Hines, J. Apt, and S. Talkudar, "Large Blackouts in North America: Historic trends and Policy implications", Energy Policy, 37(12), December 2009, pp. 5249-5259.
[11] B.A. Carreras, D.E. Newman, I. Dobson, and A.B. Poole, "Evidence for Self-Organized Criticality in a Time Series of Electric Power System Blackouts", Proceedings of the 33th Hawaii International Conference on System Science, January 2000.
[12] B.A. Carreras, D.E. Newman, I. Dobson, and A.B. Poole, "Evidence for Self-Organized Criticality in a Time Series of Electric Power System Blackouts", IEEE transactions on Circuits and Systems I: Regular Papers, 51(9), September 2004, pp 1733-1740.
[13] J. Chen, J.S. Thorpe, and M. Parashar, "Analysis of Electric Power System Disturbance Data", Proceedings of the 34th Hawaii International Conference on System Science, January 2001.
[14] R. Weron and I. Simonsen, "Blackouts, risk, and fat-tailed distributions", Proceedings of the 3rd Nikkei Econophysics Symposium, Springer-Tokyo, 2005/2006.

Disclaimer