Reliability Data – A Review of Importance, Use, and Availability

Niklas Ekstedt
Carl Johan Wallnerström
Sajeesh Babu
KTH

Patrik Hilber
Per Westerlund
Jan Henning Jürgensen
KTH

Tommie Lindquist
Svenska kraftnät

SUMMARY

For reliability studies of power distribution systems availability and collection of data on reliability is a key aspect. The acquirement of data can be challenging, because it endures effort and experience to know where to obtain accessible types of data. This paper gives the reader a guide to why input data to reliability analyses and asset management are useful, which data that can be obtained, and how to access the different types of data. Also, how to measure data accurately and the quality needed are discussed in the paper.

After a general discussion on the benefits of data, we discuss the importance of knowing exactly what the data are measurements of. Furthermore, we argue that data from different contexts, even if seemingly similar, should be used with care. We also state and explain that the amount data restrict the type of analysis that can be conducted. The paper continues with a description of some examples of (to different degrees) open accessible data. Nationally collected reliability data from Swedish utilities, reported to authorities and interest organizations, are described and discussed. We discuss how Swedish weather data, which recently have become free and open, enable more studies on the weather related reliability effects, and some existing test systems are mentioned. A section follows that describes how failure and condition data are typically stored and utilized internally in organizations. Finally, we conclude that the paper is a potential guide and inspiration for anyone planning to conduct a reliability study in the future.

KEYWORDS

Asset management, DARWin, Data management, Input data, Reliability analysis, Weather Data.

1. INTRODUCTION

Any form of analysis leads to the fundamental need for input data. Producing accurate and useful data is often a long-term work and associated with costs and thus the benefits have to be well-motivated. Input data must be sufficient to satisfy selected methods requirements, but also limited to not hold unnecessary or irrelevant information. Input data sets the highest level of the analysis results regardless how accurate the method is, i.e. an analysis method is never better than its input data. Therefore, when planning a study, it is useful to have knowledge about possible data sources and how to access them. This knowledge can preferably be gathered in a reference guide, for widespread use.

All serious studies should explain their used data and sources of data and many reliability studies have been conducted during the years. Thus, many data sources have been used and are described in numerous different reports of different kinds, some of which are mentioned in later sections of this paper. Despite this, few attempts have been made to gather and discuss different data types and sources that can be used for reliability studies, in one single place. This paper aims to become a possible starting point for anyone requiring data, and knowledge about data, for reliability studies.

In this paper we have gathered and structured the authors' collective knowledge about reliability data and presented it in a structured way. Workshops and numerous discussions have been conducted, for the purpose of this paper, besides using the authors' experience from previous reliability studies. The rest of the paper is structured as follows: In section 2, the benefits and art of choosing the right detail level of data are discussed. In section 3, some sources of data, which are open (or at least possibly accessible for interested parties), are discussed and presented. In section 4, some types of data you may find within a network utility, are discussed and presented. Finally, section 5 concludes the work.

2. BENEFITS AND THE ART OF CHOOSING RIGHT DETAIL LEVEL

2.1 On the benefits of data

H. James Harrington once said "Measurement is the first step that leads to control and eventually to improvement". Transferring this quote to the reliability analysis domain of power systems, reliability data is necessary to maintain and improve reliability standards, and effectively measure an organization's effort of accomplishing it.

Blackouts as in the United States and Canada in 2003 or the impact of the storms Gudrun and Per in Sweden depict the vulnerability of power systems, which resulted in political attention and responsibility to report reliability data such as frequency and duration of customer interruptions [1]. The public and political pressure towards a more reliable grid could to some extent be viewed as a pedagogical problem that could be reduced with more data that depicts a progress of a continuous improvement of the power system. Furthermore, reliability data enable the comparison of utilities, regulators effort and performance tracking and trends can become visible [1]. Reference [2] suggests for the American energy sector additional mandatory reports to further collect information about reliability events. This could have positive influences on policy-making and improved analysis. Especially, policy-making depends on accurate and complete data, so that policy makers have a good understanding of what actions are realistic and what has to be revised.

Operating a power grid requires complex electric equipment that has to be controlled, maintained and operated efficiently. Energy companies also need to focus on reducing costs, improving customer service and strengthen the reliable operation. One method that could achieve all the aforementioned points is condition monitoring (CM) of power system components. CM in the reliability domain is the monitoring of component characteristics over a determined time interval to gain information about the actual condition of the observed component to predict the need for maintenance or replacement [3]. Reference [3] underlines the importance of CM techniques for electrical equipment and descripts the potential benefits for utilities. A challenge of CM is the initial investment cost in new measurement methods and online-monitoring or alternative labor cost of gathering the data off-line. CM depicts the demand for a high quantity of accurate data and emphasizes the potential of saving costs in the long run. Reference [4] compares preventive maintenance with reliability centered maintenance and condition-based maintenance and discusses advantages and disadvantages.

Responsible research is based on reliable data to evaluated ideas, methods and models of the real world. Especially, utilities could benefit directly from providing data to researchers. The support of research with relevant data enables often directly access to the results before it is been published. The researcher could also been seen as a consultant who uses the data to execute an objective analysis that could be utilized to compare it with internal results. This supports important decision making and contribute to research as well. In sum, the benefits of qualitative data are widespread and as essential to an organization or researcher as the researcher itself. Data provide trends, depicts the result of actions and helps to understand the impact of certain events.

2.2 Importance of data sources and their processes

It is important to know what the data are based on and which assumptions that has been made and which definitions used. Even if two sources seem to have the same data or that e.g. the word "failure" seems to be easily defined, this is not necessary true. Are the data for example based on annual statistics over a timeframe of several years, industry averages or own standard values etc.? The definition of failures is often not standardised and which faults in statistics are included is unclear. For example, in some circumstances only failures that disconnect customers more than 3 minutes are included. What does fault length refer to: repair time, average outage time (mean value between fault when switching is possible and when it is not), a template value or something else? Which components are assumed to be perfect and which kinds of events are neglected? If input data are uncertain, performing sensitivity analyses adds value.

2.3 Own resources versus external resources

Reliability data differ significantly between different power distribution systems. This is due to different maintenance strategies, populations, environments, and as mentioned above different definitions and recordings of the failures. Hence, if possible, it is often recommended to primarily use data from the own system and secondly data from similar networks. A combination is often a good options since the own data often are based on fewer data points and perhaps not cover all kind of components that should be included in the analysis.

2.4 Choosing risk analysis method based on existing data

Figure 1 – Illustrative example of methods used in risk assessment [5]

Robust reliability analysis requires many historical fault events and is therefore not applicable in many risk analysis contexts with low probabilities and major consequences, such as airplane and nuclear [5]. Reliability analysis can be described as an extreme risk analysis method when relatively large amounts of historical data are available and hence, in other contexts than power distribution, other risk analysis methods are better to use. A commonly used solution when no or only a few historical "sharp/real" events exist is to study incidents and assume the dependency between incidents and "real" events. Customer outages in power supply are less severe but more likely to occur than in many other fields; hence, reliability analysis is a possible approach.

2.5 An approach of how to validate analysis results as a function of the amount of statistical data

Using reliability indices in different contexts applied to power distribution is well-established. However, these indices could be based on few measurements followed by a high statistical uncertainty. Therefore, a statistical model was proposed in [6] on evaluating reliability indices as a function of amount of input data. If m is the number of reported outages of a certain category that a reliability index is based on, then:

$$c_V = 100\frac{\frac{\mu_Z}{\sqrt{m}}}{\mu_Z} = \frac{100}{\sqrt{m}}[\%] \quad (1)$$

c_V is the coefficient of variation, which is defined as the ratio of standard deviation and expected value. For example, if the result is 10 [unit] and c_V is equal to 20 %, then the standard deviation is 2. Statistical analysis can be performed, such as, for instance, a 95 % confidence interval can be calculated as:

$$[calculated\ reliability\ index] \pm 1.96c_V[calculated\ reliability\ index] \quad (2)$$

3. OPEN SOURCES OF DATA

3.1 Nationally collected reliability data from Swedish utilities

DARWin is a project, within the Swedish industry organization Swedish Energy (Svensk Energi), which collects and annually presents outage data from most of the Swedish electricity system operators. The project has its roots in the 80's and the project got its current form in the late 90's [7]. According to the annual DARWin report [8] the reporting utilities represent more than 90% of Sweden's 5.2 million electricity customers.

The voluntarily participating utilities report all their events, planned and unplanned, which have caused an outage more than 3 minutes for any customer. All those events generate a single report each, and according to [8] the number of reported events to DARWin in year 2012 was approximately 80,000. 20,000 of them were planned events. Some of the information included in each event report are: interruption time, number of affected customers, failing component and its voltage level, and cause of failure.

The data in the DARWin project have been used in many reliability studies. Swedish Energy annually publish an open accessible report [8], based on the DARWin data, which presents numerous reliability results. For example, all unplanned events for a single year are divided in voltage level and failure cause, as shown in Figure 2. Another study that has used the DARWin data in its analysis is [9], where an RCM analysis is performed for the Stockholm city sub-transmission and distribution system. Another example of a study using DARWin data is the study in [10], where a weather dependent reliability model is proposed. The annual reports [8] from Swedish Energy is open accessible and can be downloaded from the Swedish Energy website. However, the actual underlying data in the DARWin database are not directly accessible, but according to the authors' experience, the anonymized data can be retrieved upon request.

Felorsak	**24 kV**	**12 kV**	**<10 kV**	**0,4 kV**	**Totalt**
Åska	882	1901		764	3547
Övrigt väder	2344	4886	11	2110	9351
Åverkan	365	1132	17	4241	5755
Material/metod	1155	3160	13	8075	12403
Personal	97	273		465	835
Överlast	24	105	2	678	809
Återvändande last	18	13		9	40
Säkringsbrott	255	828	6	4031	5120
Okänd	2741	8926	26	9142	20835
Summa	7881	21224	75	29515	58695

Figure 2 - A table in [8], presenting the unplanned events from 2012 divided in voltage level and failure cause. In Swedish.

The Swedish Energy Market Inspectorate (EI) is collecting interruption reports from Swedish network utilities on customer level every year since 2011 [11]. The annual reports and some reported underlying data are openly published on EI's website. For example, a datasheet with each named network utility's individual key figures is open accessible [12]. Approximately 20 key figures are presented and some of them are SAIDI, SAIFI, and percentage of customers with more than 4 outages. Furthermore, a datasheet with each named network utility's component information is also open accessible [13]. In this datasheet, some of the information are the total length of uninsulated overhead lines, number substations, and totally installed transformer effect, among other.

In a comparison between the DARWin data and the EI data, it can be noted that some of the information overlaps, like the number of failures and reliability indices. Nevertheless, the two sources of data are complementing each other, since the DARWin data are more focused on failure causes and failed components, whereas the EI data are more focused on the individual utilities and customers.

3.2 Swedish standard cost data regarding investments, re-investments, maintanence etc.

Since the 1960's, Swedish Energy has annually published a standard cost publication entitled EBR (elbyggnadsrationalisering, ~electrical construction rationalization), aimed to rationalize planning, investments and maintenance of power systems, i.e., to provide rational "standards" when constructing power systems. EBR consists of six aggregation levels (P1-P6), affecting each other from bottom to top, see Figure 3. P1-P3 and parts of P4 are reported in the annual EBR publications (nowadays both physical and electronic versions) [5].

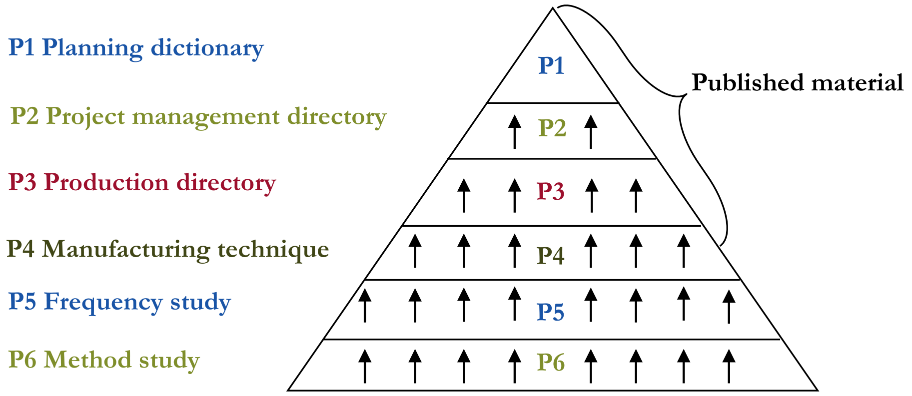

Figure 3 – The structure of EBR [5]

Underlying frequency and time studies (P5 and P6) are partly based on interviews with different DSOs, partly on methodological studies in which Swedish Energy conducted field studies. The number of underlying studies differs, but is often only a few. According to interviews with Swedish Energy [14], the reason is the difficulty finding voluntary DSOs because it is both time- and resource consuming. In EBR, the costs are divided into the following categories: (a) work, (b) material, (c) machine, (d) equipment and (e) other. "Other" includes building permit costs. EBR does not include common costs such as overall administration and research. The recommendation is to add 3-8%.

3.3 Tendency towards more free public data including an example of Swedish weather data

In many countries there is a tendency towards free open public data. For example, within the European union the PSI Directive (Directive 2003/98/EC on the re-use of public sector information) encourage that as much public information as possible should be free for re-use. An idea is to simulate innovation, new services and products.

Utilities and companies within the power and energy sector often have the opportunity to utilize the increasing amount of free data. One example is weather data. The benefits of good weather data have and will increase further due to more renewable energy (better prediction), smart grid concepts, and technologies such as dynamic rating (that require such data) etc. The Swedish Meteorological and Hydrological Institute (SMHI) is one example of a governmental body that very recently has made a huge amount of both historical and real-time data freely available for use. Even before that, such data has been used in research (discounted cost for use in non-commercial research purposes) and by companies. Examples of publication with results from research projects within the RCAM group, based on data from SMHI, are [14] and [6].

3.4 Test systems

As stated in section 2.3, there are large differences between different power systems and hence there is often a benefit to use own data or at least data from a similar system. However, there are several benefits to develop and use more general test systems in some cases, for example:

- It saves time when testing methods in development
- It facilitates the comparison of different methods
- It enables comparison of software, if the same input data are used
- Research results can be reproduced

Examples of test systems:

- RTS, a larger international (IEEE) test systems, primarily for the transmission [15]
- RBTS, smaller systems that contain more data even at the distribution level [16] and [17]
- Representative test systems for Swedish power distribution: a typical urban system and a typical rural system [18]

4. UTILITY INTERNAL SOURCES

This section concentrates on how data, which is needed for reliability evaluations, are collected and used by utilities. Providing the consumers with safe and reliable supply of electricity, power utilities have the responsibility of maintaining the parts of the grid that they are in charge of. These tasks should be regularly supervised and documented for different purposes [19]. They can be used to:

- Identify the congestions, hotspots, or week-links in the network
- Assess if restructuring or new extensions are required
- Evaluate if the performance is satisfying the required efficiency and quality
- Know if the operations abide the regulations from inspectoral agencies
- Identify the degree of performance, so as to participate in the electricity market

Utilities consolidate and maintain a “fault register” which is a listing of the relevant failures and related events in concerned network, with available details and information on each event. The basic information in a typical registry entry can be the time of the event, key component, local network specification, voltage, type and description of the event along with correspondence to the concerned technician [20]. The registers can circulate internally in the utility within the different departments concerned with the performance of the network. Departments of the utility, such as operations, control, power quality, maintenance etc. can read, analyze and make decisions based on the registered events.

Each division investigating the event of its choice can receive the related information and follow up the further actions. The required actions can be of different priorities. The major faults in the network, affecting a large number of consumers for a long period of time, demand attention from different departments and need maintenance to be dispatched at the earliest. In Sweden, the evaluation reports explaining the fault, affected customers and actions taken corresponding to major outage events are to be submitted to energy market inspectorate following the events [21]. The judgments from the skilled engineers of the utility decide the time and scale of the corrective actions. These corrective actions are also recorded along the registered events.

The categorization of registered faults as per the type of failure provides some indication to the causal observation of faults in a network. For instance, the average number of failures due to material failures, faulty maneuvers, unknown reasons etc. can be figured out. Reference [22] is a work by

SINTEF Energy research, Norway, based on the protection system failure statistics for 10 years, and it investigates the major reasons of failure. Another highlight of fault registers is that they can provide categorical input to reliability evaluations. For example, the statistics on the number of failures recorded over certain region or even per components can be isolated from the registers and used for calculation [23].

The use of reliability data at a utility can be divided into two cases: planning the network and planning the maintenance of the built network. During planning, the components are regarded as equal and their availability is estimated according to the tests of the manufacturer and the experience from their users. For planning of maintenance of a specific component, obviously its history should be recorded, manufacturer, year of production, type, year of installation, the failures with their cause and so on.

An important task for a utility is to monitor the condition of the equipment. This includes both direct indicators as indirect ones. A direct indicator of the condition of an electric contact is the resistance, which can be calculated when measuring the current and the voltage. An indirect indicator in the same case is the temperature rise, which has to be related to the current, in order to give an indication of the condition. Another division of the indicators is according to the purpose of collection. Both the resistance and the temperature rise are measured for maintenance purpose, where there are other measurements carried out for the operation of the electric network, such as currents and voltages as well as the number of operations (the openings and closings of a breaker) and recording during disturbances. Some of this indicators are collected continuously, whereas others more seldom, normally during preventive maintenance. In the last case are the indicators that require a disconnection of the equipment. The availability of this data is varying. There is normally a database about the apparatus, where the history with maintenance is recorded. The voltages and currents are recorded in another system.

5. CONCLUSION

In this paper we have discussed the importance, use, and availability of reliability data. Several examples of data sources, with different degree of availability, have been presented. Experiences from utility internal registers have been used to explain and discuss how this data are and can be used. The paper is suited for usage as a guide, for anyone planning a reliability analysis. It can also be used as an inspiration for future reliability studies. Hopefully, the gathered knowledge in this paper will help improve the quality of future reliability studies. Since we believe that reliability analyses are important for further improvements of the power grid, we hope that the construction of this paper is step toward a more reliable power system.

6. ACKNOWLEDGEMENT

This work was sponsored by SweGRIDS and the risk analysis program at Elforsk.

BIBLIOGRAPHY

[1] J. H. Eto, “Tracking the Reliability of the U.S. Electric Power System: An Assessment of Publicly Available Information Reported to State Public Utility Commissions,” Ernest Orlando Lawrence Berkeley National Laboratory, October 2008

[2] E. Fisher, J. H. Eto, and K. H. LaCommare, "Understanding Bulk Power Reliability: The Importance of Good Data and a Critical Review of Existing Sources," *System Science (HICSS), 2012 45th Hawaii International Conference on* , vol., no., pp.2159,2168, 4-7 Jan. 2012

[3] Y. Han and Y. H. Song, "Condition monitoring techniques for electrical equipment-a literature survey," *Power Delivery, IEEE Transactions on* , vol.18, no.1, pp.4,13, Jan 2003

[4] A. H. C. Tsang, (1995) "Condition-based maintenance: tools and decision making", Journal of Quality in Maintenance Engineering, Vol. 1 Iss: 3, pp.3 – 17

[5] C. J. Wallnerström and P. Hilber, ”Reliability Analysis and Asset Management Applied to Power Distribution”, 2014

[6] C. J. Wallnerström and P. Hilber, “Vulnerability Analysis of Power Distribution Systems for Cost-effective Resource Allocation”, IEEE Transaction on Power Systems, IEEE Trans. on power systems, vol. 27, pp. 224-232, 2012.
[7] M Tapper, private conversation, May 2014.
[8] M. Tapper, (2014, February 21). *DARWin Driftstörningsstatistik 2012* [Online]. Available: http://www.svenskenergi.se/Global/Statistik/DARWin/DARWin%202012.pdf
[9] L. Bertling, R. Eriksson, R. Allan, L.-Å. Gustafsson, and M. Åhlén, "Survey of causes of failures based on statistics and practice for improvements of preventive maintenance plans," in 14th Power Systems Computation Conf., 2002.
[10] K. Alvehag and L. Söder, "A stochastic weather dependent reliability model for distribution systems," in Proc. 10th Int. Conf. Probabilistic Methods Applied Power Systems (PMAPS), Rincón, 2008, pp. 1-8.
[11] K. Granath and B. Gustavsson, (2014). *Ei R2014:04 Leverenssäkerheten i elnäten 2012* [Online]. Available: http://ei.se/Documents/Publikationer/rapporter_och_pm/Rapporter%202014/Ei_R2014_04.pdf
[12] Energimarknadsinspektionen, (2014). *2012_ars_nyckeltal_over_elnatsforetagens_elavbrott* [Online]. Available: http://ei.se/Documents/Publikationer/statistik/2012_ars_nyckeltal_over_elnatsforetagens_elavbrott.xlsx
[13] Energimarknadsinspektionen, (2014). *Sarskild_rapport_teknisk_data* [Online]. Available: http://www.energimarknadsinspektionen.se/Documents/Publikationer/arsrapporter/el/sammanstallningar_over_elnatsforetagens_arsrapporter_lokalnat/Sarskild_rapport_teknisk_data.xlsx
[14] C. J. Wallnerström, A. Isenberg, J. Setréus and P. Hilber, “The Potential of Using Equivalent Comparison Standards to judge Effectible Costs in Electrical Distribution Tariff Regulation”, CIRED 2011, June 6-9 2011, Frankfurt, Germany
[15] IEEE Subcommittee report. “IEEE Reliability Test System”, IEEE Transaction PAS-98 pp2047-2054, 1979.
[16] R. N. Allan, R. Billinton, I. Sjarief, L. Goel, and K. S. So, “A Reliability Test System for Educational Purposes – Basic Distribution System Data and Results”, IEEE Transaction Power System, Vol. 6, No. 2, May 1991, pp. 813-820.
[17] R. Billinton and S. Jonnavithula, “A Test System Teaching Overall Power System Reliability Assessment”, IEEE Transaction Power System, Vol. 11, Issue 4, Nov. 1996 pp1670 – 1676.
[18] Elforsk report 8:42, free to download at www.elforsk.se
[19] J. E. Kwoka Jr, "Electric power distribution: economies of scale, mergers, and restructuring." Applied Economics 37.20 (2005): 2373-2386.
[20] J. Öckerman, Regionnät Skydd och Kontroll, Vattenfall Eldistribution AB, Stockholm, Sweden, email correspondence, May 2014.
[21] J. Olsson, "Fault report", message from Fortum to the Swedish Energy Markets Inspectorate 2012-12-19, in Swedish.
[22] G. H. Kjølle, Y. Aabø, and B. T. Hjartsjø. Fault Statistics as a Basis for Designing Cost-Effective Protection and Control Solutions [Online]. Available: http://www.sintef.no/project/Vern_kontroll%20og%20automatisering/cigre_2002.pdf
[23] Y. He, "Study and Analysis of Distribution Equipment Reliability Data." Elforsk Report (2010).