Extracting Resilience Statistics from Utility Data in Distribution Grids

Nichelle'Le K. Carrington, Shanshan Ma, Ian Dobson, Zhaoyu Wang
Department of Electrical and Computer Engineering
Iowa State University, Ames, Iowa, USA
nkcarrin, sma, dobson, wzy @iastate.edu

***Abstract*—It is useful to quantify electrical distribution system resilience based on historical performance. This paper systematically extracts resilience curves from historical utility outage data, extracts resilience metrics such as duration, average recovery rates, and maximum number of simultaneously outaged components, and examines the statistics of these resilience metrics for small, medium, and large events. The resilience metrics and their typical variabilities are expected to be helpful in predicting and bounding the likely outcomes of future resilience events. For example, we can calculate the restoration time that will be achieved with 95% confidence.**

***Index Terms*—power distribution, restoration, resilience, risk**

I. Introduction

Maintaining a continuous energy supply to customers is the goal of utilities, but threats such as aging infrastructure, bad weather, and indigenous wildlife are the most common culprits for unplanned disruption of electrical services in electrical distribution systems [1]. When the outage happens, customers invariably want to know when the power supply will come back. The current method of estimating restoration time is based on the field reports of various crews after an event has occurred, and the utility makes a "best guess" estimate of the restoration time [1], [2]. But this method is approximate and delayed as the utility has to wait for the damaged areas to be safe for crews entering and reports from the inspection crews. We think that the determination of the restoration time can be usefully augmented by suitably processing historical data from previous events. Our approach uses detailed outage data that is routinely collected by many utilities.

To aid utilities with providing quicker estimation of the outage duration and restoration time, many researchers have used resilience curves as a credible method to model and evaluate system vulnerability and the ability to recover from hazards or adverse events. In [3], the entire life cycle of failure and recovery of large scale power failures is considered, but lesser events are ignored, which tends to exaggerate the typical impacts of events. The work in [4] statistically analyzed factors that affect outage duration but did not predict the duration or its variability. Authors in [1] used the text from inspection reports (without considering the number of outages) to predict outage duration to facilitate customer preparation. In these studies, extreme weather events are considered as isolated events and are used to observe the impacts of extreme conditions on the system's performance. But they omit the less extreme and more typical events that also contribute to the system's overall resilience.

In this paper, we systematically detect and extract the resilience curves from historical outage data of one United States utility. These resilience curves represent all the events that have disrupted the normal condition of the distribution system. The events are detected by processing the cumulative number of outages as a function of time. The cumulative number of outages is the number of outages present at a given time that have not yet been restored. Threshold values on the cumulative number of outages are used to define the beginning and end of the resilience events and classify the events into small, medium and large. The classification into small, medium and large events allows the resilience metrics of each size of event to be calculated, as well as the variability of the metric. Resilience curves of the same magnitude are grouped together into a large set of curves for analysis instead of re-sampling from one isolated event repetitively. The outage propagation and restoration stages are evaluated using resilience triangles on a large data set to assess: (a) variability of duration for outage and recovery processes to help with predictions; (b) average outage and recovery rates during events to help with assessment and predictions. By systematically detecting resilience curves, we are able to gain better insight on the overall performance of the system and generate statistics of the duration of outage and recovery processes from multiple events instead of focusing on one single event.

The rest of the paper is organized as follows: section II describes the resilience analysis framework and event extraction procedure and the metrics calculated. The utility data set and its metrics and their variabilities are presented in section III, and section IV concludes the paper.

II. Resilience and Systematic Detection

A. The Three Stages of Resilience

Resilience is the ability to prepare for, absorb, adapt to, and/or rapidly recover from adverse events [2], [5], [6], [7]. Resilience curves are used to model a system's response before, during, and after a hazard [2], [8], [9], [10]. The three-stage framework depicted on the resilience curve in Fig. 1 is a generalized model to capture a system's recovery process as it progresses in each stage [2], [11], [12]. The first stage focuses on hazard prevention ($0 \leq t \leq t_S$). This is when the

978-1-7281-5508-1/20/$31.00

system maintains normality before a hazardous event starts. The second stage is the outage propagation ($t_S \leq t \leq t_N$), where the successive outages occur at a faster rate while the hazard is absorbed by the system. The third stage of restoration ($t_N \leq t \leq t_E$) is the recovery of the system back to normal operation. We regard the second and third stage as a resilience event. We now explain how the resilience event is extracted from real data.

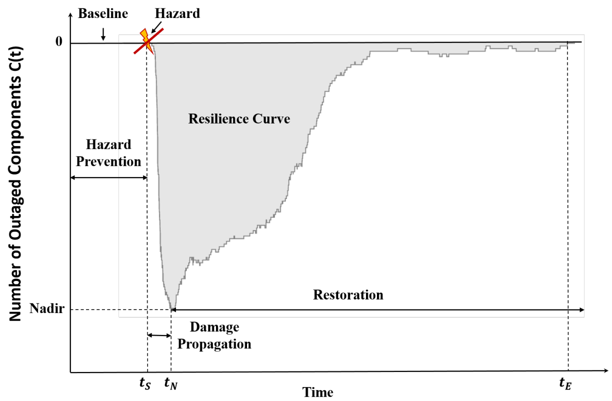

Fig. 1. The three stages of resilience shown on an extracted resilience curve.

B. Extracting Events

The cumulative number of outages varies with time as outages occur and are restored. Under normal conditions the cumulative number of outages stays near zero because outages are generally infrequent and are restored quickly. But under stressed conditions, outages are more frequent and accumulate before they can be restored, and the cumulative number of outages has excursions away from zero. These accumulations of outages are the resilience events. The resilience events are extracted from the data by detecting when the cumulative number of outages passes and returns to a threshold number of outages. We now give more details of this extraction.

Since the utility data includes the outage and restore time for each component, it is straightforward to sort these times by their order of occurrence and then calculate

$$\begin{aligned} C(t) &= -(\text{cumulative total outages at time } t \textit{ minus} \\ &\qquad \text{cumulative total restores at time } t) \\ &= -(\text{number of simultaneous outages at time } t) \end{aligned} \quad (1)$$

The threshold number of outages is zero or a small negative number of outages C_{base}; we use $C_{\text{base}} = 0$ as a simple case. Under normal conditions $C(t)$ is at or above C_{base}. The start time t_S of an event is defined by $C(t)$ decreasing below C_{base} and the end time t_E of an event is defined by $C(t)$ increasing to C_{base}. Then $C(t)$ for $t_S \leq t \leq t_E$ is a resilience curve for the event as shown in Fig. 1. The minus sign in (1) ensures compatibility with standard resilience curves.

C. Nadir, Resilience Triangles, and Metrics

In Fig. 1, the lowest point of the resilience event curve $C(t)$, called the nadir, occurs at t_N. The nadir $C(t_N)$ corresponds to the maximum number of simultaneously occurring outages

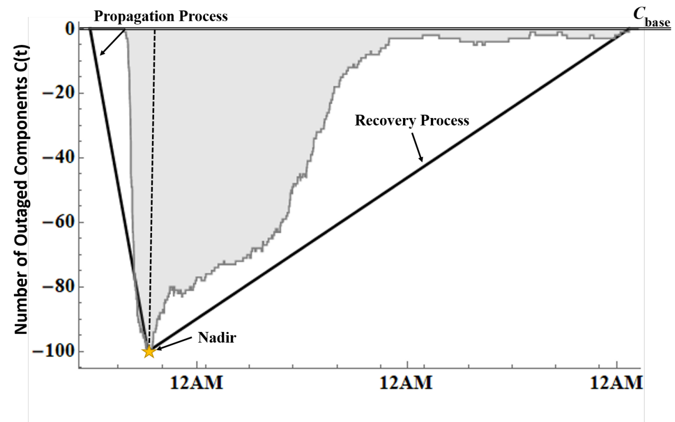

Fig. 2. The outage propagation process begins when the curve decreases from the baseline at zero and ends at the nadir. The recovery process starts at the nadir and ends when the curve increases to the baseline.

in an event. (In the exceptional case of several lowest points occurring at exactly the same level in the same event, we choose the last one to be the nadir.) We use the nadir to locate the end of the propagation process and the beginning of the restoration process. Note that dividing the event time into separate propagation and restoration processes in this way is idealized, since in real data these processes overlap somewhat as shown in Fig. 3. We will address this in future work.

After these definitions, metrics of duration follow easily from the widths of the resilience triangles. The duration of propagation is $t_N - t_S$, and the duration of restoration is $t_E - t_N$. The event duration is $t_E - t_S$. These duration metrics can aid in explaining the impact of a disruptive event and predict the impacts of future events [8].

Average rate metrics follow easily from the slopes of the resilience triangles. The average outage process rate is $-C(t_N)/(t_N - t_S)$ and the average recovery process rate is $-C(t_N)/(t_E - t_N)$.

The maximum number of simultaneously outaged components, the duration of an event, the restoration time, and the average outage and recovery rates are important metrics that we extract from from previous events to help a utility to estimate these metrics for an anticipated or ongoing event. In particular we calculate the statistics of these metrics from previous events to be able to estimate a 95% upper bound confidence interval for the restoration time. This can help provide the customers of the utility with an upper bound estimate of the restoration time with a reasonable certainty.

Events are grouped into small, medium, or large depending on the nadir $C(t_N)$ of their resilience curve:

Small events have $-3 \geq C(t_N) \geq -9$.
Medium events have $-10 \geq C(t_N) \geq -19$.
Large events have $-20 \geq C(t_N)$.

D. Customer resilience curves

Since the outage data also includes the number of customers outaged and restored, we can also form the cumulative number of customers out $C^{\text{cust}}(t)$ as a function of time, similarly to the definition of $C(t)$ in (1) except that "outages" are replaced by

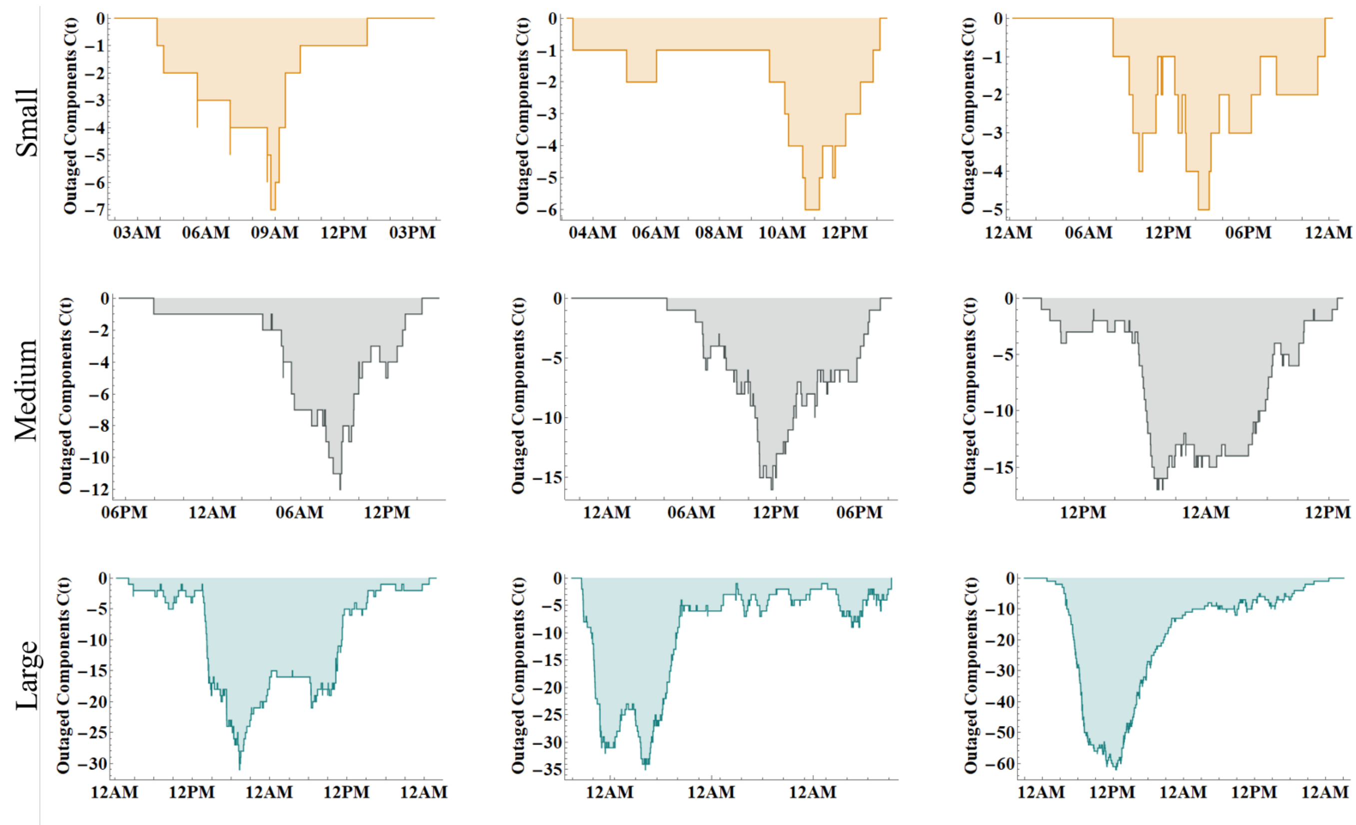

Fig. 3. Component resilience curve examples. Small event (Orange) causes are scheduled maintenance or minor physical damage. Medium event (Gray) causes are moderate weather/storm or moderate physical damage. Large event (Teal) causes are extreme or severe weather/storm or severe physical damage.

"customers". Then the customer resilience curves for an event occurring for $t_S \leq t \leq t_E$ is the portion of the cumulative customer curve $C^{\text{cust}}(t)$ for $t_S \leq t \leq t_E$ (see Fig. 4). The

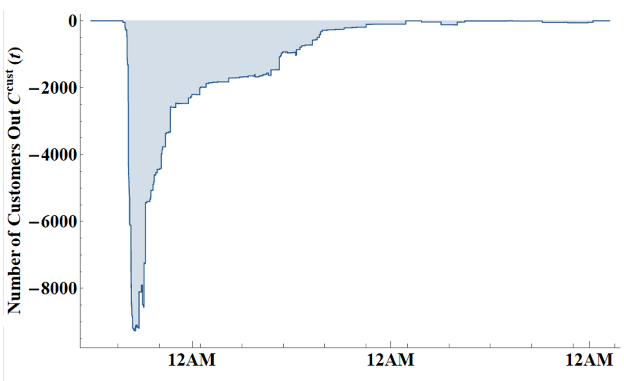

Fig. 4. The customer resilience curve shows the cumulative number of customers outaged during the event corresponding to Figure 2.

area above the customer resilience curve is the total customer hours outaged in the event. If the event were to be included in the SAIFI calculation, this customer area would directly add to the SAIFI numerator. The average customer recovery rate is $-C^{\text{cust}}(t_N)/(t_E - t_N)$.

III. Results

The historical outage data was gathered by one distribution utility from 2011 to 2016. 32,291 outages were reported during this time and 15,648 pole locations were identified. The start and end time of the outages were recorded by fuse cards equipped to record the time at which an outage begins and ends based on the loss of power. The raw data is private.

The probability density functions for propagation, restoration and total event durations for all events are shown in Fig. 5.

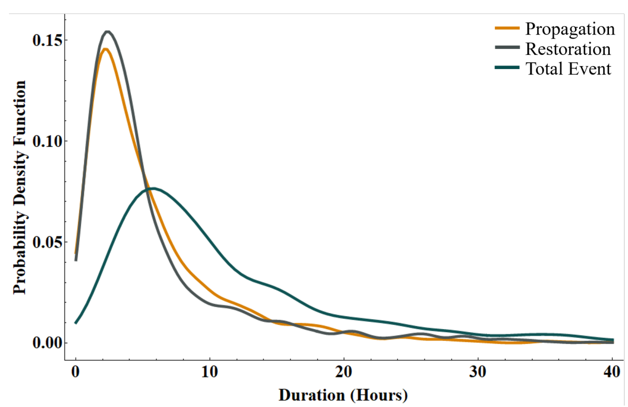

Fig. 5. Distributions of propagation, restoration and total event durations.

The cause code of each outage and restore within an event was tracked to determine the most common cause of outages for the event. There were 34,945 outages with causes reported and 63 types of cause codes. About 5% of those causes were weather-related, and 14% were animal-related. The remaining causes were mostly component malfunctions, tree limbs, and debris. The top causes for all events were tree limbs near the clearance zone of lines and squirrels. The top three weather-related causes were wind, rain, and lightning.

A. Small, Medium and Large Events

1486 events were extracted and grouped by size using the methods of Section II. There were 910 small events, 75 medium events, and 50 large events found in the data. The events in Fig. 3 are samples of these events. The large events had more outages caused by weather than the small and medium events. A common pattern of many large events is that outages caused by weather are followed by outages caused by tree limbs and other debris.

B. Nadir

The statistics for the resilience curve nadirs are summarized in Table I. In Table I, it can be seen that the utility can expect at least 5 simultaneously outaged components for any event. The average number of simultaneously outaged components expected in large events is over twice the average for medium events and over 8 times the average for small events.

TABLE I
NADIR $C(t_N)$

Events	Mean	Median	Std.Dev.
Small	–5.25	–5	1.36
Medium	–13.81	–13	2.75
Large	–42.52	–34	22.13
All	–7.67	–5	9.59

C. Event Duration

The average event duration and their variabilities are summarized in Table II, and the survival function of the event duration for each event size is shown in Fig. 6. The average event duration is 13 hours, but the variability is high. Small events are over within 24 hours with 95% confidence, but the corresponding upper bounds for the medium events are two times longer and the large events are four times longer. The utility will be able to make an assessment of the event duration for the event and also scale the estimate of the duration up or down if necessary for sudden changes to conditions.

TABLE II
EVENT DURATION (HOURS)

Events	Mean	Median	Std.Dev.	95%CI
Small	9.50	7.65	6.74	24
Medium	32.11	21.63	51.99	55
Large	49.48	41.42	23.31	99
All	13.07	8.5	19.07	36

D. Outage Propagation Process

The estimation and variability of the duration for the outage propagation process are shown in Table III, and the survival function of the distribution of the propagation process for each group is in Fig. 7. The duration of propagation during any event is less than 18 hours with 95% confidence.

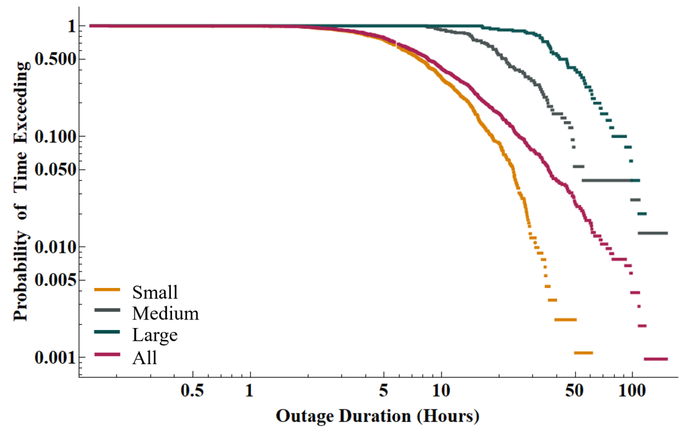

Fig. 6. Survival functions of event duration.

TABLE III
PROPAGATION PROCESS DURATION (HOURS)

Events	Mean	Median	Std.Dev.	95%CI
Small	5.01	3.5	4.66	15
Medium	17.79	7.82	50.77	43
Large	14.78	10.92	11.91	40
All	6.41	3.93	15.06	18

E. Recovery Process

The estimation and variability of the duration for the recovery process are shown in Table IV and the survival functions of the recovery process duration for each events size are shown in Fig. 8. The expected duration of recovery during any event is less than 22.6 hours with 95% confidence. This upper bound on duration for any event is about four times lower than the upper bound on duration for large events and twice that of small events.

TABLE IV
RESTORATION DURATION (HOURS)

Events	Mean	Median	Std.Dev.	95%CI
Small	4.5	3.35	3.97	13
Medium	14.31	11.9	10.89	30
Large	34.7	28.37	21.77	85
All	6.67	3.73	9.58	22.6

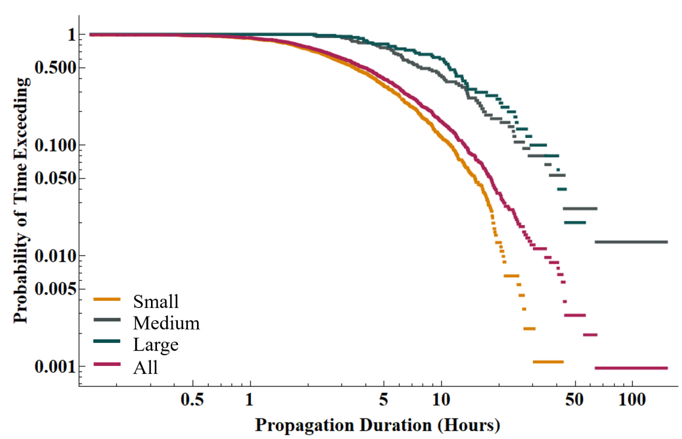

Fig. 7. Survival functions of the propagation process duration.

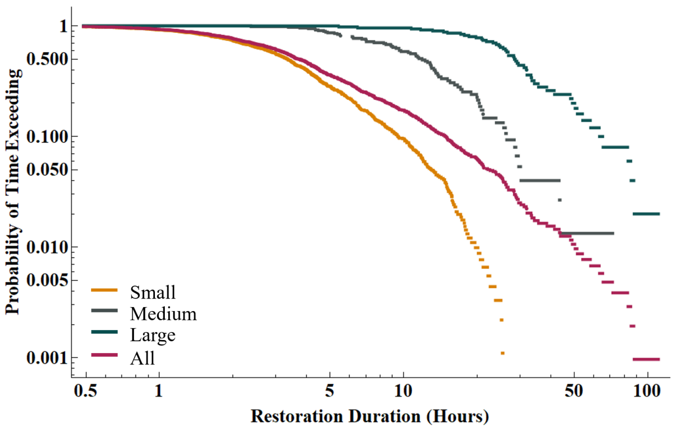

Fig. 8. Survival functions of recovery process duration.

F. Customer Impact

The customer resilience curves for each event and the customer area under each curve were computed, and the statistics are shown in Table V. Table V shows that for all events, the customer hours of outage are less than 2399 with 95% confidence.

TABLE V
CUSTOMER AREA (CUSTOMER HOURS)

Events	Mean	Median	Std.Dev.	95%CI
Small	405	168	812	1466
Medium	1442	982	1944	5297
Large	2501	1972	204	5534
All	581	209	1145	2399

G. Average Outage and Recovery Rates

The average recovery process rate and the average outage process rate are calculated from the resilience triangles by dividing the magnitude of the nadir by the duration of the process. Table VI shows that for medium and large events, the average recovery rate is slower than the average outage rate. The 95% confidence interval for the average recovery rate shown in Table VI is a one-sided lower confidence interval. That is, the probability that the average recovery rate is more than the given value is 0.95. After the nadir of a resilience event, when the damage has been inspected and the current number of outages is known, the average recovery rate and its 95% confident lower bound can be multiplied by the number of outages to estimate the expected recovery time and its 95% confident upper bound recovery time.

TABLE VI
AVERAGE OUTAGE AND RECOVERY RATES (PER HOUR)

	Average outage rate			Average recovery rate			
Events	Mean	Median	StdDev	Mean	Median	StdDev	95%CI
Small	0.48	0.68	0.42	0.45	0.66	0.35	0.37
Medium	0.50	0.55	0.70	0.68	0.97	0.95	0.22
Large	0.19	0.32	0.18	0.63	0.66	0.95	0.27
All	0.45	0.66	0.38	0.47	0.67	0.37	0.37

IV. CONCLUSIONS

This paper systematically detects and extracts resilience curves from 5 years of distribution utility outage data for each resilience event in which outages accumulate. For each event, the resilience curve is divided into an outage propagation period and a recovery period using the nadir of the resilience curve. The events are classified into small, medium, and large, and the statistics of resilience metrics such as the durations of the event, propagation and recovery and the customer hours lost are computed for each size of event. In contrast to previous work, we compute statistics of groups of typical resilience events rather than focusing on single resilience events. The statistics for the average durations and their variability should be helpful in estimating event and recovery times for future events before or while they are occurring. Further work studying the events and their associated cause codes in more detail is indicated.

ACKNOWLEDGEMENTS

We gratefully acknowledge support in part from NSF grants 1609080 and 1735354, and U.S. Department of Energy Office of Electricity under DE-OE0000875.

REFERENCES

[1] A. Jaech, B. Zhang, M. Ostendorf, and D. S. Kirschen, "Real-time prediction of the duration of distribution system outages," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 773–781, 2018.
[2] M. Ouyang, L. Dueñas-Osorio, and X. Min, "A three-stage resilience analysis framework for urban infrastructure systems," *Structural safety*, vol. 36, pp. 23–31, 2012.
[3] Y. Wei, C. Ji, F. Galvan, S. Couvillon, G. Orellana, and J. Momoh, "Non-stationary random process for large-scale failure and recovery of power distributions," *Applied Mathematics*, 2012.
[4] M.-Y. Chow, L. S. Taylor, and M.-S. Chow, "Time of outage restoration analysis in distribution systems," *IEEE Transactions on Power Delivery*, vol. 11, no. 3, pp. 1652–1658, 1996.
[5] S. Ma, N. Carrington, A. Arif, and Z. Wang, "Resilience assessment of a self-healing distribution systems under extreme weather events," in *IEEE PES General Meeting*. IEEE, 2019.
[6] N. Yodo and P. Wang, "Resilience modeling and quantification for engineered systems using bayesian networks," *Journal of Mechanical Design*, vol. 138, no. 3, p. 031404, 2016.
[7] B. M. Ayyub, "Practical resilience metrics for planning, design, and decision making," *ASCE-ASME Journal Risk & Uncertainty Engineering Systems, Part A: Civil Engineering*, vol. 1, no. 3, p. 04015008, 2015.
[8] B. M. Ayyub, "Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making," *Risk Analysis*, vol. 34, no. 2, pp. 340–355, 2014.
[9] R. Francis and B. Bekera, "A metric and frameworks for resilience analysis of engineered and infrastructure systems," *Reliability Engineering & System Safety*, vol. 121, pp. 90–103, 2014.
[10] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, "A review of definitions and measures of system resilience," *Reliability Engineering & System Safety*, vol. 145, pp. 47–61, 2016.
[11] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatziargyriou, "Power systems resilience assessment: hardening and smart operational enhancement strategies," *Proceedings IEEE*, vol. 105, no. 7, pp. 1202–1213, 2017.
[12] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017.