Topology and Voltage Aware Fragility and Data-Driven Risk Modeling of Distribution Feeders

Apsara Adhikari, Charlotte Wertz, Anamika Dubey
Department of Electrical Engineering and Computer Science
Washington State University

Arslan Ahmad, Ian Dobson
Department of Electrical and Computer Engineering
Iowa State University

***Abstract*—The increasing frequency and intensity of wind-related outages underscore the need for data-driven resilience planning in power distribution systems. Traditional resiliency analyses often rely on single system-level average fragility information, which overlooks variations in feeder topology and voltage class, leading to less effective hardening strategies. This study presents a topology and voltage-aware framework for modeling fragility and risk in distribution networks under wind stress events. Outage and weather data for two different areas in Massachusetts—urban Worcester and rural northwestern Massachusetts are considered, and separate fragility curves are developed by feeder type (main vs. lateral) and voltage class (5 kV vs. 15 kV). Results show that fragility varies significantly across different regions for both main and lateral events, as well as for different voltage levels, with distinct onset and rate of growth. Lateral feeders exhibit up to approximately six times higher outage rate than mains across the study regions. The fragility results are further translated into a customer-focused risk metric, the Expected Customers Not Served Rate (ECNSR), linking technical vulnerability to customer impact for targeted, risk-informed resilience investment.**

***Index Terms*—Power grid resilience, fragility curves, outage modeling, data-driven models, power distribution system.**

I. Introduction

The frequency and intensity of weather-related outages have increased in recent years, highlighting the need for utilities to enhance their systems through resilience planning. Effective resiliency planning requires a quantitative understanding of system fragility and risk, as well as how it varies across different network topologies and voltage classes. However, most existing resilience planning studies apply system-level single fragility information that masks structural heterogeneity, leading to suboptimal hardening strategies and inefficient resource allocation [1]. Stratifying the distribution network by feeder type such as main and lateral and voltage class enables utilities to accurately quantify differentiated risk exposure and implement risk-proportionate resilience strategies that optimize investment. This approach also provides insight into how physical construction and network configuration influence wind-induced failures, offering a deeper understanding of system-specific vulnerabilities.

Fragility curves are widely employed in power systems to evaluate the vulnerability of assets to natural hazards and to guide resilience planning. They estimate potential outage risk under forecasted weather conditions by translating hazard intensity into probabilistic failure likelihood. In [2], [3], failure rates were found to have an exponential relationship with wind speed, establishing a quantitative foundation for weather-dependent modeling. In [4], empirical fragility curves for overhead electrical lines were developed using an extensive fault database for vulnerability assessment. Similarly, [5] presented regionally normalized fragility curves to reflect differences in regional weather exposure and asset condition. The authors in [6] created fragility curves for high-voltage transmission lines (220 kV and 132 kV) to support network-level risk estimation. These studies demonstrate that fragility modeling provides a powerful mechanism for translating natural hazards into resilience risk.

While fragility curves quantify the probability of failure of power system components at given hazard intensities, risk assessment extends this analysis by incorporating both the likelihood of failure and the associated impact. This transforms fragility analysis into a decision-oriented framework for risk-informed resilience planning and investment prioritization, allowing utilities to allocate their limited resources where they yield the most significant improvement in overall service continuity. However, most existing research still applies a single, system-level fragility curve representing the aggregated vulnerability of the entire network, without accounting for variations in topology or voltage class. In reality, distribution systems are highly heterogeneous: main feeders that form the high-capacity backbone differ in construction, exposure, and protection from lateral branches that deliver power to end customers, while voltage levels follow distinct design standards and mechanical robustness. As a result, the probability of failure under identical wind conditions can vary significantly between feeder types and voltage classes. Despite this, prior studies integrating fragility modeling with risk-based planning such as [7], [8] have not incorporated stratification by topology or voltage class, potentially obscuring localized vulnerabilities and limiting the effectiveness of resilience planning.

To fill this gap, this study proposes a framework for modeling fragility and risk in distribution systems affected by wind events. We create separate fragility curves for two geographic regions, based on feeder type (main versus lateral) and voltage class, using outage and weather data. These curves show how the physical layout and design features influence the network's vulnerability to increasing wind intensity. We

Funding from the Power Systems Engineering Research Center (PSERC) Project S110 is gratefully acknowledged.

979-8-3315-8138-1

then convert the fragility information into a risk calculation focused on customers through the Expected Customers Not Served Rate (ECNSR). This combines the likelihood of failure due to weather with the impact on customers. This connection between technical reliability and customer impact gives a precise risk measure to inform decisions on resilience investment.The empirical analysis is used to illustrate the methodology under contrasting but meteorologically similar conditions, so the numerical fragility and risk estimates are system-specific, whereas the modeling framework and topology/voltage stratification are intended to be applicable to other distribution networks.

The rest of this paper is organized as follows. Section II describes the datasets used for analysis, including outage and weather data. Section III outlines the method for creating wind-driven fragility curves and calculating the related risk metric. Section IV presents and discusses the main results. Finally, Section V concludes the paper with final remarks, limitations, and directions for future research.

II. Data Description

A. Outage Datasets

The outage data used in this analysis were obtained from the Massachusetts government website [9]. The dataset includes detailed outage records that specify the city or town of occurrence, the number of customers affected, outage duration, time of occurrence, cause of the outage, and other related attributes. For the analysis, we used outage data from year 2016 to 2024 for two study areas; Area 1 is the urban city of Worcester, and Area 2 is rural northwestern Massachusetts. These areas were selected to capture the contrast between urban and rural distribution networks while maintaining comparable weather characteristics. Although both regions experience similar overall weather patterns, Area 1 has higher wind gust intensities than Area 2, which is valuable for analyzing how network topology and voltage class influence fragility and risk under different exposure conditions. To calculate the stratified fragility curves, the neighboring substations were merged to ensure sufficient outage data- eight in Area 1 and five in Area 2.

B. Weather Datasets

For weather data, this study used observations from the Iowa Environmental Mesonet (IEM) dataset [10], which compiles records from ASOS and AWOS weather stations in hourly resolution. Only wind speed and wind gust variables were utilized in this analysis.

III. Methodology

A. Data Processing and Cleaning

The outage data processing began with geocoding raw outage records containing the street addresses of protection devices that experienced outages. For records with missing or ambiguous location data, additional information from the outage comments was used to infer and fill in the missing locations. Records for which location data could not be determined were excluded from the dataset—approximately 1.5% for Area 1 and 1.2% for Area 2.

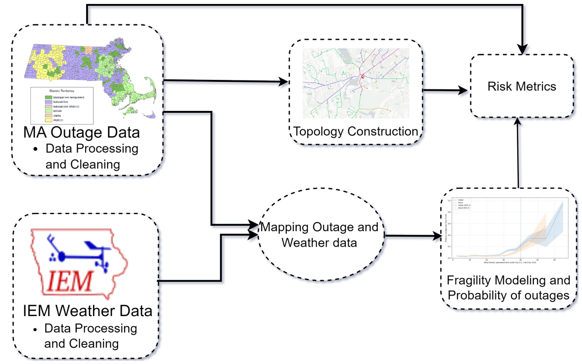

Fig. 1. Framework for topology and voltage class aware fragility and risk modeling

For the weather data, nearby weather stations were identified based on their proximity to electrical substations within the study area. A new variable combining wind speed and gust, denoted as $\tilde{w}(t)$, was calculated as:

$$\tilde{w}(t) = \begin{cases} \text{wind gust speed}, & \text{if wind gust observed}, \\ \text{sustained wind speed}, & \text{if wind gust not observed}. \end{cases}$$

To relate weather and outage observations, we define a combined wind exposure variable $\tilde{w}(t)$ evaluated at the start of the hour in which the outage is recorded. When gust data are available, the gust value timestamped at the beginning of that hour corresponds to the maximum gust reported over the hour.

B. Topology Construction

The distribution system topology was approximated from the sparse outage data using a set of practical assumptions. It was assumed that all protection devices operated at least once during the fault and that distribution circuits follow existing street layouts—an approach commonly adopted in synthetic feeder construction [11].

A substation (S) and its related points (P) were selected. An OpenStreetMap (OSM) graph was created over the bounding box that contained P. The nodes S and P were connected to the nearest OSM nodes. Dijkstra's algorithm was used to find the shortest network distances from S to all nodes. The points were then labeled as MAIN or LATERAL. They were grouped by circuit number and branch label. Corresponding S→MAIN and MAIN→LATERAL paths were traced, or S→LATERAL if no main exists, as shown in Fig. 2. To assess sensitivity to geocoding uncertainty, we perturbed 10% of outage locations with Gaussian noise and repeated the outage-to-network mapping, main/lateral classification, and fragility estimation. The main/lateral labels remained stable and the fragility curves showed no significant change, indicating that, although the reconstructed topology does not reproduce exact feeder routing, it preserves the key structural features needed for fragility analysis, such as main–lateral branching patterns and approximate line lengths.

979-8-3315-8138-1

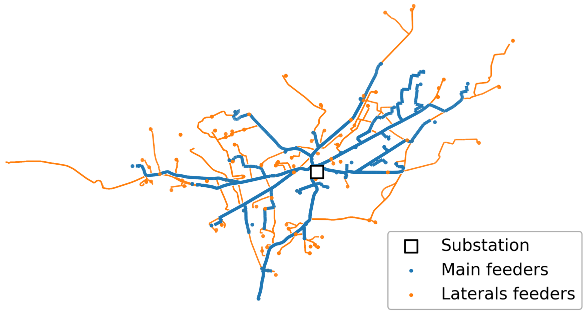

Fig. 2. Topology for the Webster St. substation in Worcester city constructed from the outage dataset, labeled by main and laterals feeder.

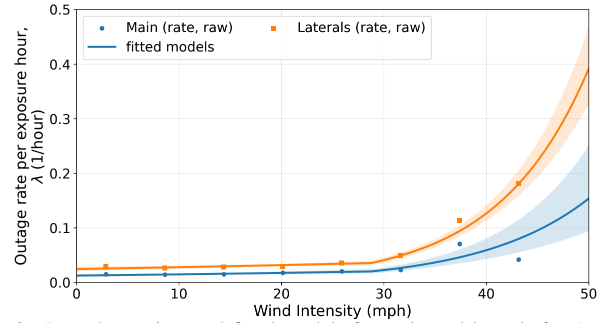

Fig. 3. Raw data points and fitted models for main and laterals for Area 1.

C. Fragility Modeling and Probability of outage

For calculating the probability of failure, a secondary dataset was constructed in which each observation was labeled as 1 if an outage occurred and 0 otherwise. For each one-hour interval, the corresponding wind exposure at the beginning of the hour was recorded to capture the relationship between weather conditions and outage occurrence.The wind data for each hour was binned with approximately 5 mph bin widths up to 40 mph, followed by 10 mph intervals for the stronger winds. These bins ensure sufficient resolution in wind at lower wind intensities and sufficient outage data in each bin at higher wind speeds. The hourly rate of failure in each bin is estimated as the number of outages in that bin divided by the total number of hours of wind speeds observed in that bin. This yields the average outage rate as a function of wind intensity, as indicated by the dots in figures 3-6.

After computing the empirical outage rates for each wind-speed bin, the relationship between wind intensity and failure rate was modeled as

$$\lambda(x_b) = \exp\{\beta_0 + \beta_1 x_b + \beta_2 (x_b - k)\}, \qquad (1)$$

where x_b is the midpoint of the wind-speed bin and $\lambda(x_b)$ is the corresponding average outage rate. On the log scale, the model is continuous and piecewise linear, with

$$\frac{d}{dx} \log \lambda(x) = \begin{cases} \beta_1, & x < k, \\ \beta_1 + \beta_2, & x \geq k, \end{cases}$$

so that the hinge at $x = k$ captures the change in fragility behavior at higher wind speeds.

In our cases, k represents the critical wind speed, which is approximately 28 mph for both study areas. This value is determined in a data-driven manner by treating k as a hyperparameter optimized to minimize the Akaike Information Criterion (AIC). The optimal value (k = 28 mph) aligns with the observed breakpoint in the empirical outage-rate curves, where failure rates begin to increase more sharply with wind intensity. The model parameters $(\beta_0, \beta_1, \beta_2)$ are then estimated using maximum likelihood under the Poisson assumption.

D. Expected Customers Not Served Rate (ECNSR)

To translate the fragility curves into a customer-oriented measure of risk, we calculated the *Expected Customers Not Served Rate (ECNSR)*. This metric quantifies the expected number of customer outages per hour at a given design wind intensity w, combining both outage rate and customer impact.

For any feeder group G (e.g., mains, laterals, or voltage classes), let $\lambda_G(w)$ be the hourly outage rate from the fitted fragility curve at wind intensity w. If the group contains N_G feeders and feeder i serves C_i customers, the average number of customers per feeder and *ECNSR* are ;

$$\bar{C}_G = \frac{1}{N_G} \sum_{i \in G} C_i, \qquad \mathrm{ECNSR}_G(w) = \lambda_G(w)\, \bar{C}_G, \qquad (2)$$

which represents the expected number of customers lost per hour at wind intensity w. This formulation allows ECNSR to be evaluated for different categories of feeders. For instance, main and lateral feeders can be compared by substituting their respective fragility-derived outage rates $\lambda_{\mathrm{main}}(w)$ and $\lambda_{\mathrm{lat}}(w)$, together with their average customer counts. The same approach applies when comparing voltage levels, such as 5 kV and 15 kV feeders, enabling a direct assessment of risk exposure across different segments of the distribution network. By linking the probabilistic fragility model to customer impact, ECNSR captures both technical vulnerability and the social consequences of weather-driven outages. This offers a practical basis for identifying which feeder categories contribute most to system-wide risk and for guiding risk-proportional resilience planning.

IV. Results

This section presents the estimated outage rates and risk results and compares wind-induced outage rates across voltage levels and feeder types, highlighting differences in onset thresholds, growth rates, and risk.

A. Fragility Modeling

1) Main vs. Lateral Fragility: Figures 3 and 4 show the fitted fragility curves, representing mean outage rates per hour with 95% confidence intervals, for main and lateral feeders in both study areas. In both areas, the fitted Poisson hinge GLM captures a clear nonlinear transition in outage rates around the critical wind threshold of approximately 28 mph. Across all wind speeds, laterals exhibit higher outage rates than mains, with the gap widening at higher winds, indicating greater fragility and faster deterioration of lateral performance as wind intensity increases.This behavior is consistent with laterals being more sensitive to wind exposure due to smaller

979-8-3315-8138-1

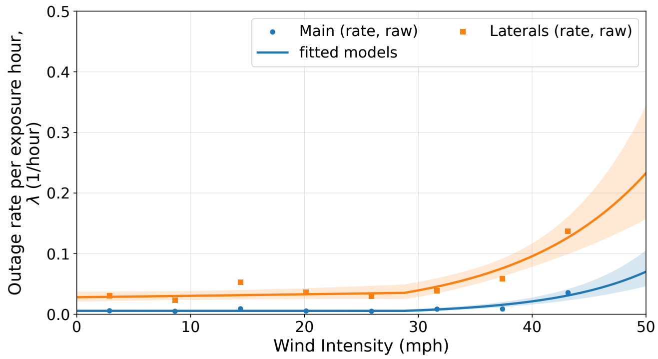

Fig. 4. Raw data points and fitted models for main and laterals for Area 2.

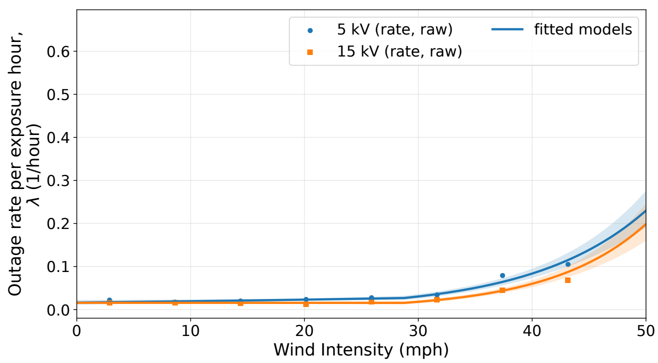

Fig. 5. Raw data points and fitted models for 15KV and 5KV voltage class feeders for Area 1

conductor sizes, closer proximity to vegetation, and lighter construction standards.

From the outage rate curves in the figures, we observe that in Area 1, laterals fail about 1.76 times more often than mains at lower wind intensities (25 mph), and at higher wind intensities (45 mph) the outage rate increases to about 2.63 times that of mains. In Area 2, the effects are even more pronounced, with laterals exhibiting approximately 6.4 times higher fragility than mains at 25 mph and about 4.5 times higher fragility at higher wind intensities. These differences highlight how network configuration and surrounding environment modulate the relative vulnerability of mains and laterals. Overall outage rates per hour are higher in Area 1 compared to Area 2, consistent with the urban denser built infrastructure, more obstructed corridors, and stronger gust exposure. Area 1 experiences around 12% higher gust speeds than Area 2. By contrast, Area 2 exhibits a slightly higher onset threshold and lower overall outage rates, reflecting its more open terrain, sparser feeder layout, and reduced gust exposure, which is consistent with rural operating environment.To assess robustness to incomplete outage reporting, we randomly removed 10% of outage events in both study areas and re-estimated the fragility curves. The maximum deviation from the baseline was 8.95% for laterals and 12.99% for mains; however, the overall curve shape and trends were preserved, and all deviations remained within the corresponding confidence intervals.

2) Voltage-Level Fragility: Figs. 5 and 6 present the fitted fragility curves for the 5 kV and 15 kV feeders for the study ar-

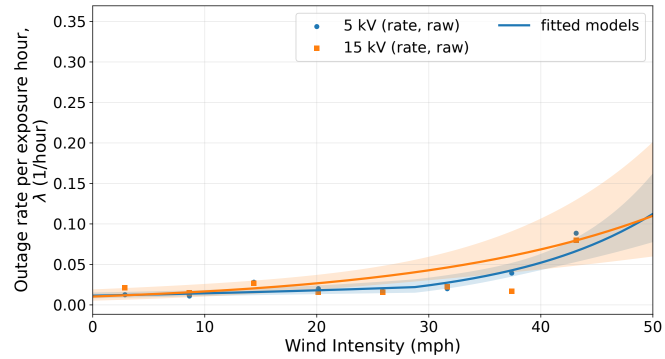

Fig. 6. Raw data points and fitted models for 15KV and 5KV voltage class feeders for Area 2

eas. In both cases, the curves exhibit the expected exponential behavior, with outage rates remaining low at moderate wind speeds and increasing sharply once wind intensity exceeds the critical threshold of approximately 28 mph. In Area 1, both voltage classes display similar onset patterns; however, the 5 kV feeders experience a slightly faster increase in failure rate at higher wind speeds. This is consistent with their smaller conductor sizes and greater exposure to roadside vegetation, which makes them more susceptible to wind-related in urban environment.

In contrast, in Area 2, the 15 kV feeders appear more fragile than the 5 kV feeders, exhibiting an earlier onset of failures and a steeper escalation beyond 28 mph. This behavior can be attributed to the longer average line lengths of the 15 kV circuits, which are commonly found in rural areas to supply sparse populations, resulting in greater exposure to wind loading and a higher probability of component failure. The Poisson hinge GLM models effectively capture these voltage-dependent differences, with wider confidence intervals at higher wind speeds reflecting the more limited data. These values should be viewed as system-specific empirical quantities rather than universal benchmarks, as they are shaped by the particular construction standards, vegetation conditions, and operating practices of the studied utility. Overall, these findings suggest that voltage-level fragility is influenced by both physical design characteristics and geographic context, underscoring the importance of incorporating voltage stratification into resilience and hardening strategies.

B. Risk Analysis

While the fragility curves provide valuable insight into the probability of wind-induced failures, they capture only the average rate of outages, not their overall impact on customers or system performance. For instance, although lateral feeders exhibit higher outage rates at a given wind speed compared to mains, the outage of a main feeder typically affects a much larger number of customers. Hence, fragility alone cannot fully represent the operational risk or the broader reliability implications of severe weather events. To quantify the combined effect of failure probability and outage impact, a risk-based metric, ECNSR, is computed.

TABLE I
ECNSR BY FEEDER TYPE AND BY VOLTAGE CLASS.

Region	Feeder	Wind (mph)	Rate (1/hr)	Risk (cust/hr)
Area 1	Main	25	0.0185	18.92
	Lateral	25	0.0326	4.93
	Main	35	0.0325	33.22
	Lateral	35	0.0701	10.58
	Main	45	0.0571	58.33
	Lateral	45	0.1504	22.70
Area 2	Main	25	0.0051	5.25
	Lateral	25	0.0326	4.92
	Main	35	0.0107	10.95
	Lateral	35	0.0470	7.10
	Main	45	0.0224	22.85
	Lateral	45	0.0978	10.24
Region	**Voltage**	**Wind (mph)**	**Rate (1/hr)**	**Risk (cust/hr)**
Area 1	5 kV	25	0.0309	8.81
	15 kV	25	0.0203	6.04
	5 kV	35	0.0551	15.71
	15 kV	35	0.0468	13.89
	5 kV	45	0.0983	28.02
	15 kV	45	0.1076	31.96
Area 2	5 kV	25	0.0180	5.31
	15 kV	25	0.0336	17.16
	5 kV	35	0.0186	5.48
	15 kV	35	0.0423	21.60
	5 kV	45	0.0192	5.67
	15 kV	45	0.0532	27.19

In both Area 1 and Area 2, laterals exhibit higher outage rates than mains, yet main outages produce substantially higher risk because each main serves many more customers (about 1022 vs. 151 in Area 1). At 25 mph in Area 1, laterals fail nearly twice as often as mains (0.0326 vs. 0.0185 outages/hour), but the main risk (18.92 customers/hour) is about four times the lateral risk (4.93 customers/hour), and the gap widens at higher winds (58.33 vs. 22.70 customers/hour at 45 mph). Similar patterns occur in Area 2 with generally lower absolute risk: at low wind speeds, main and lateral risks are comparable, whereas at higher wind speeds main feeder risk is nearly twice that of laterals.

The voltage-level risk results summarized in Table I reveal distinct differences between 5 kV and 15 kV feeders across both study areas. In Area 1, average customers per feeder are similar (285 at 5 kV, 297 at 15 kV), so overall risk is comparable: 5 kV feeders have slightly higher risk at low wind speeds, while 15 kV feeders dominate at higher winds. In Area 2, where 15 kV feeders serve many more customers on average (511 vs. 295), they consistently exhibit higher risk across all wind speeds, driving roughly 4 to 5 times more customer impact, especially at stronger winds. This disparity primarily reflects the longer line lengths and larger service territories of the 15 kV feeders, which increase their exposure to wind loading and fault potential.

These results demonstrate how outage rates and risk vary by voltage level across different geographical settings and infrastructure characteristics. They highlight that feeder topology and voltage class strongly influence outage impacts, with higher-voltage feeders contributing more to overall system risk, despite having comparable or lower fragility, which depends on geographical conditions and infrastructure. Integrating risk-based metrics, such as ECNSR, enables utilities to prioritize resilience investments where they deliver the most significant reliability benefits, supporting targeted, data-driven hardening in response to increasing weather-related outages.

V. CONCLUSION

This study presented a data-driven framework for modeling wind-induced fragility and translating it into customer-centric risk metrics, segregated by feeder topologies (mains vs. laterals) and voltage classes. The results highlight the importance of risk-proportional resilience planning to guide targeted investments in distribution systems. We introduced the Expected Customers Not Served Rate (ECNSR), a customer-focused risk metric that links the vulnerability computed from fragility curves to the risk of power interruptions to customers. Future work will extend this analysis to other regions, improve robustness under limited data, and incorporate an energy-based metric such as Expected Consumers Not Served Energy (ECNSE) by accounting for outage duration and customer/load-class energy. While the analysis is limited to two Massachusetts utility regions with merged substations and thus provides system-specific numerical estimates, the topology- and voltage-stratified framework and ECNSR metric are designed to be transferable to other distribution networks given suitable outage and weather data.

REFERENCES

[1] S. N. Jahromi, E. Hajipour, and M. Ehsan, "Optimal resilience-oriented microgrid formation considering failure probability of distribution feeders," *Electric Power Systems Research*, 2022.

[2] L. Xu and R. Brown, "Undergrounding assessment phase 3 report: Ex ante cost and benefit modeling," Quanta Technology, Raleigh, NC, Tech. Rep., 2008, Technical Report.

[3] A. Ahmad and I. Dobson, "Towards using utility data to quantify how investments would have increased the wind resilience of distribution systems," *IEEE Trans. Power Syst.*, 2024.

[4] S. Dunn et al., "Fragility curves for assessing the resilience of electricity networks constructed from an extensive fault database," *Natural Hazards Rev.*, 2018.

[5] D. L. Donaldson et al., "Enhancing power distribution network operational resilience to extreme wind events," *Meteorological Applications*, 2023.

[6] M. R. Jamieson, G. Strbac, and K. R. Bell, "Quantification and visualisation of extreme wind effects on transmission network outage probability and wind generation output," *IET Smart Grid*, 2020.

[7] A. Poudyal, S. Poudel, and A. Dubey, "Risk-based active distribution system planning for resilience against extreme weather events," *IEEE Trans. Sust. Energy*, 2023.

[8] E. L. Barrett et al., "A risk-based framework for power system modeling to improve resilience to extreme events," *IEEE Open Access J. Power & Energy*, 2023.

[9] Massachusetts Emergency Management Agency (MEMA), *Power outages*, https://www.mass.gov/info-details/power-outages.

[10] Iowa State University, *Iowa Environmental Mesonet: ASOS/AWOS/METAR Data Download*, https://mesonet.agron.iastate.edu/request/download.phtml.

[11] H. Li et al., "Building highly detailed synthetic electric grid data sets for combined transmission and distribution systems," *IEEE Open Access J. Power Energy*, 2020.

979-8-3315-8138-1