IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Quantifying the Power System Resilience of the US Power Grid Through Weather and Power Outage Data Mapping

Sangkeun Matthew Lee[1], Supriya Chinthavali[2], Narayan Bhusal[3], Nils Stenvig[3], Anika Tabassum[1], and Teja Kuruganti[4]

[1]Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Rd, Oak Ridge, TN 37830, United States (e-mail: lees4@ornl.gov, tabassuma@ornl.gov)
[2]Geospatial Science and Human Security Division, Oak Ridge National Laboratory, 1 Bethel Valley Rd, Oak Ridge, TN 37830, United States (e-mail: chinthavalis@ornl.gov)
[3]Electrification and Energy Infrastructures Division, Oak Ridge National Laboratory, 1 Bethel Valley Rd, Oak Ridge, TN 37830, United States (e-mail: bhusaln@ornl.gov, stenvignm@ornl.gov)
[4]Computational Sciences and Engineering Division, 1 Bethel Valley Rd, Oak Ridge, TN 37830, United States (e-mail: kurugantipv@ornl.gov)

Corresponding author: Sangkeun Matthew Lee (e-mail: lees4@ornl.gov).

ABSTRACT Recent increases of extreme weather events such as severe thunderstorms, floods, and hurricanes are leading to destruction in power system equipment (transmission and distribution poles and lines, substations, power plants, etc.) and are causing widespread prolonged power outages. These outages often cause inconveniences in critical services (health care, transportation, national security, etc.) and significant losses in economy, leading to human suffering. Therefore, understanding the spatiotemporal correlation of these events with power systems is crucial to planning and for maintaining reliable operation and control under such events. However, developing such correlation requires several datasets, including weather event and power outage datasets, along with coordination from multiple entities (e.g., electric utilities, government agencies, and research organizations). Also, high-resolution data collection is a time-consuming and tedious task because different interest groups are involved in the process. To this end, we propose an automated data framework that maps severe weather events with power outages to quantify power system resilience. This framework uses the publicly available National Weather Service dataset and Oak Ridge National Laboratory's Environment for Analysis of Geo-Located Energy Information (EAGLE-I) power outage dataset to quantify the power system resilience. The proposed work can quantify power system resilience against extreme weather events at the county/state level for different weather event types (e.g., hurricanes, sever thunderstorms, and floods). The outcome of the proposed work will be useful to identify vulnerability hot spots, develop weather event–based planning strategies (planning strategies might change with events types), develop asset management strategies, and develop predictive analysis tools.

INDEX TERMS Data analytics, EAGLE-I, extreme weather events, power outages, power system resilience, resilience quantification

I. INTRODUCTION

Extreme weather events such as hurricanes, severe thunderstorms, floods, heat waves, and earthquakes have become increasingly frequent and severe in recent years [1]. These events are posing significant operational, control, and planning challenges for the power system worldwide. The power system—as the most complicated and interconnected machine with aging infrastructure in the United States [2]—is highly vulnerable to these events. Extreme weather events have been causing significant disruptions in the power grid system, resulting in widespread power outages and severe infrastructure (e.g., substations, transmission and distribution lines, and power generation plants) damage, leading to inconveniences in critical services (e.g., health care, transportation, and national security), severe economic losses, and adverse effects on the well being of the community [3, 4, 5]. Therefore, understanding the effects of different weather events on power outages and resilience in various regions is crucial, as it enables utility companies and emergency responders to plan for and respond to these events more effectively [6, 7].

However, creating a national-scale report on various kinds of weather events causing power outages and restoration time is a challenging and costly task. This process requires a significant amount of time and effort due to the labor-intensive data recording and collection process. Utility companies need to monitor and gather weather and power outage information, and this information from multiple utility companies needs to be gathered, standardized, and processed by a single entity. According to the 2022 annual electric power industry report by the US Energy Information Administration (EIA), 1,700 electric utilities were operating in the United States as of 2020 [8]. To accurately understand power outages and restoration times caused by weather events on a national scale with high-resolution geographic data, it would be ideal for utility companies to actively monitor individual outages and their corresponding restoration times for each weather event. Moreover, the data should be consolidated into a central repository for reporting purposes. However, gathering weather events and their effect on power outages and restoration times nationally is challenging. This undertaking necessitates the involvement of utility companies, which might be required to allocate substantial resources to monitor various weather types and collect detailed information. Implementing such data collection initiatives can be an extra burden for utility companies.

The US Department of Energy (DOE) and the EIA mandate US utility companies to submit their data related to power system resilience [9]. However, the utilities' emphasis is on extreme events only using a generic nationwide threshold (e.g., 50,000 customers affected, 300 MW power loss) or on aggregated data at the utility or state level rather than high-resolution data, such as at the county level, which can omit a considerable number of smaller scale but important events. The effect of the same extreme events can vary significantly in different counties, so the aggregation of data at the utility or state level might obscure important information (e.g., variation of power system resilience in the same state, vulnerability of a county to specific weather event types). Therefore, there is an urgent need for (1) good quality data on weather and power outages for the United States and (2) a reusable automated data framework that combines the best available datasets to enable high-resolution spatiotemporal analysis, allowing for a configurable threshold to detect significant weather events of various scales and types associated with power outages.

To address these needs, this paper leverages the publicly available National Weather Service (NWS) dataset in combination with Oak Ridge National Laboratory's (ORNL's) Environment for Analysis of Geo-Located Energy Information (EAGLE-I) power outage dataset. This work proposes an automated data framework that preprocesses and combines these datasets, detects extreme weather events that cause power outages over a configurable threshold, and quantifies strength and vulnerability (resilience) of the US power grids. A high-level conceptual framework of the proposed work is shown in Fig. 1.

Contributions: The proposed data framework can offer significant insights into improving power grid resilience and emergency response planning, saving time and effort, particularly in regions prone to extreme weather events. The major contributions of this paper can be summarized as follows.

- This work explains publicly available datasets, the NWS dataset, and the EAGLE-I dataset and presents how these datasets can be repeatedly and automatically preprocessed and combined together to quantify power system resilience to weather events in the United States. Mapping of the NWS and EAGLE-I power outage datasets is useful to identify vulnerability hot spots, develop weather event–based planning strategies (planning strategies might change with event types), develop asset management strategies, and develop predictive analysis tools.
- This work demonstrates how the proposed method can identify critical weather events and and their details, which cannot be captured in DOE and EIA reports at higher geographical resolution using a configurable threshold.
- This work also presents a spatiotemporal analysis of resilience to various types of weather events (e.g., severe thunderstorms, floods, and hurricanes) through NWS and EAGLE-I power outage dataset mapping. We explored several research topics including (1) the geographical pattern of US power system resilience, (2) comparison of resilience related to different weather event types, (3) changes in resilience over time, and (4) causes of continued power outages after the end of extreme events.

The remainder of this paper is organized as follows. Section II reviews related work, encompassing an overview of the US power outage and reliability reports, challenges, and other relevant studies. Section III introduces the datasets used in this study and explains the data processing framework accompanied by a discussion of the underlying assumptions and potential limitations. Section IV provides a national-scale analysis of US power system resilience associated with weather events. Finally, Section V provides concluding remarks with future directions.

II. RELATED WORK

Although power system reliability is well defined and widely accepted across the industry [10], power system resilience definitions have not yet been standardized [4, 7, 11, 12, 13, 14, 15, 16]. In this paper, we follow the power system

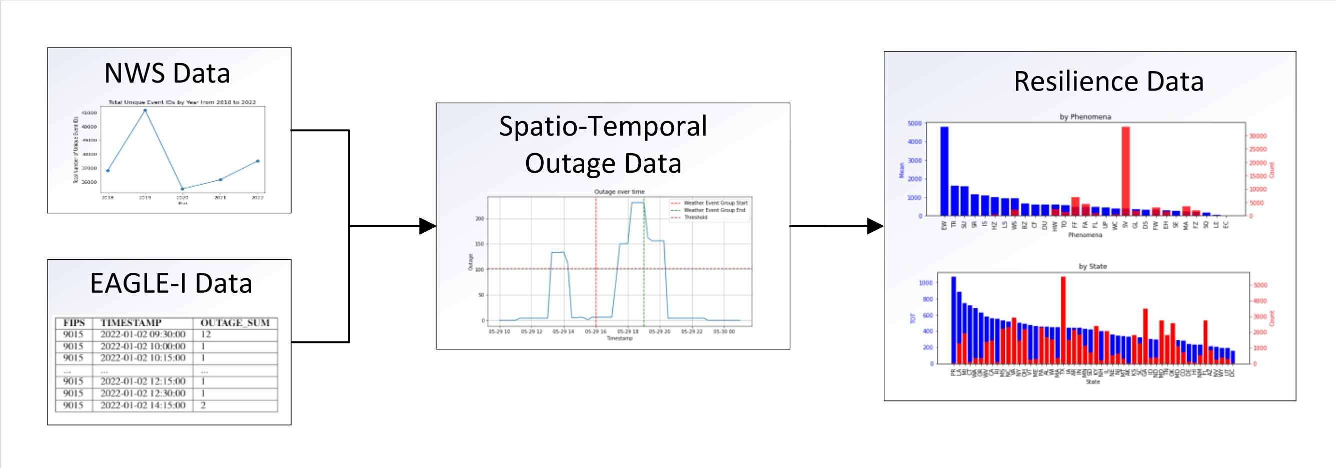


FIGURE 1: High-level conceptual framework of the proposed work.

resilience defined and adopted by the US National Infrastructure Advisory Council (NIAC) and the North America Electric Reliability Corporation (NERC). They define *resilience* of a power system as a power system's ability to withstand, adapt to, and recover from disruptions or unexpected events [17], such as severe weather events or cyberattacks. A resilient power grid system can absorb disturbances and damages so that effects can be limited and recovered from quickly.

In this paper, we particularly focus on the effect of severe weather events on power grid operation. To understand electric outages and power system resilience associated with extreme weather events, having reliable and accurate data is the first step [5]. The DOE collects information on major electric incidents and emergencies for national security, emergency management, and analytical purposes. Balancing authorities, reliability coordinators, and electric utilities are mandated to submit Form DOE-417 to the DOE [9]. Because mandating utility companies to report all power outages and restoration time is impractical, DOE has established various criteria for filing this form. For instance, an uncontrolled loss of 300 MW or more for more than 15 minutes or loss of electric service to more than 50,000 customers must be reported [18].

DOE publishes an annual summary report based on the collected data, including the events' dates and times, restoration dates and times, affected areas, event types, demand losses, and the number of affected customers. Because the report relies on manual submissions, the details of each event can vary because there are no strict standardizations or requirements on some of the data fields. For example, some reported events might have restoration times, while others might not. Some reports might include only state-level information and lack country-specific details. Some events lack the number of customers affected. The "event type" column in the published data provides information on whether the reported incident is related to severe weather. However, it lacks further details, such as specifying whether the event was a flood or a thunderstorm. In the 2022 report, 95 severe weather-related events were reported, but this number could have omitted many significant weather-related outage events that were not qualified as "major" events but that created a significant impact, as generalized criteria (e.g., 50,000 customers or 300 MW) might not capture significant impacts on smaller counties [5].

Similarly, the EIA gathers power system reliability data from distribution utilities and electricity power marketers through the EIA-861 form and publishes the annual electric power industry report [8]. This report encompasses statistics such as utility information, operational data, reliability indices both with and without major event days, energy efficiency, and more. The EIA report houses a wealth of invaluable information on reliability, featuring metrics such as the customer average interruption duration index (CAIDI), system average interruption duration index (SAIDI), and system average interruption frequency index (SAIFI). It is essential to note, however, that the reliability metrics included in the EIA report do not focus on power system resilience in the face of extreme weather events. Reliability and resilience are two distinct concepts. To assess resilience clearly, we need to analyze power outages dataset, weather data, and their spatiotemporal relation during extreme weather events. The aggregated data in the Annual Electric Power Industry Report poses limitations in understanding power system resilience. Furthermore, both the DOE and EIA reports rely exclusively on manual efforts by mandating the participation of utilities, a process that can be time consuming and costly.

The work presented in [5] demonstrates the usefulness of the EAGLE-I power outage data and shows how resilience metrics such as event duration, impact duration, recovery duration, and impact level can be calculated. The authors of [5] used a population-based threshold to identify extreme power outages and conducted a comprehensive analysis of power outages from 2014 to 2021. This analysis illustrated the usefulness of EAGLE-I data in detecting events that could not be found using DOE criteria. The study primarily focused on state-level outage behavior. However, it did not examine different types of weather phenomena (e.g., floods and severe thunderstorms) and their specific effects on power system resilience. Contrarily, we focus on analyzing power system resilience to extreme weather events at both the county and state levels. We also examine the impact of different types of extreme weather events (through weather and power outage data mapping) on the power system resilience, which helps to identify vulnerability hot spots. Information of weather event types is important as power system resilience planning strategies change with the type of weather events. The weather data mapped in the outage data also gives information about whether a power outage continued after the end of an extreme event. Continued power outage after an extreme event could be caused by delays in weather (NWS starts to collect data after weather alerts for extreme events) and power outage data collection (EAGLE-I platform collects power outage data through the utility company's data API or their website). Another possible reason for continued outages after an extreme event signifies weaker and aging power grid infrastructure (important information for asset management). We have quantified the continued power outage duration after the end of an extreme event as one of the resilience metrics for further analysis. Furthermore, information about a weather event will also help in developing the framework to predict power outages in the future against similar events (important for predictive analysis). Moreover, we have also calculated a more realistic threshold value—average power outages from regular causes—to distinguish power outages caused by extreme events from other regular causes (e.g., vegetation and system faults). The extreme event–related power outages determined using this threshold are more realistic than that of the arbitrary 25% value used by [5]. This calculation of threshold value is possible because of the mapping of NWS data on the EAGLE-I dataset, which is lacking in [5].

III. DATA PROCESSING

A. NATIONAL WEATHER SERVICE DATA

For weather information, we use weather event data collected by the NWS, a US government agency that provides weather, water, and climate data, forecasts, warnings, and impact-based decision support services. Specifically, we used the NWS Valid Time Extent Code (VTEC) archives data processed by the Iowa State University Iowa Environmental Mesonet (IEM). This dataset contains information about the geography and life cycle of weather events that occur in the United States, which include watches, warnings, advisories, and others. VTEC archives the dataset, and its metadata are publicly available in the Shapefile and Keyhole Markup Language (KML) formats. Weather events are classified using a two-character phenomena code (such as FL for flood and SV for severe thunderstorms). IEM daily updates the data at 2:00 a.m. Central time, and the dataset is publicly available. The start and end times of weather events can be modified over time and upgraded to higher-risk events (such as a watch being upgraded to a warning), and the most recent snapshot of the weather events is available in the dataset.

Table 1 displays the selective columns of the NWS VTEC archive dataset. Note that we excluded some of the columns available in the VTEC archive dataset because they are used only for certain types of weather events and are not necessary for our analysis (e.g., HAILTAG column for hail size tag in inches and HV_CAUSE for the cause of flood events). The complete metadata for the dataset is available on the VTEC archive website[1].

To preprocess the dataset, we initially filtered the data using the SIG and STATUS columns. Since our focus is on extreme weather events, we used the SIG column to filter out everything except either warnings or advisories (with values of "W" and "Y" for the SIG column, respectively). As per the NWS's definition, a warning is issued when a hazardous weather or hydrologic event is happening, imminent, or expected to occur. Similarly, an advisory is issued but with less severity than warnings. Both warnings and advisories can cause significant inconvenience and, if caution is not exercised, can lead to situations that can endanger life or property. Note that we excluded less serious events such as watches and statements but included the ones that were upgraded during the life cycle of those events. We also filtered out data entries whose STATUS column values were "CAN," indicating that the event was canceled and did not actually occur.

As the second step, we converted polygons to counties. The entries in the original dataset have a column GTYPE, which indicates the data entry is either for a county or a polygon. Since EAGLE-I power outage data is county-level data, to more easily join the datasets we converted all polygon-based data into county-based data. Since polygons can cover multiple counties, we identified all the counties that intersect with the polygons and replaced the original data entries with the county geometry.

Following that, we aimed to obtain a unique identifier for each weather event that we could use to identify a weather event with a specific phenomenon type occurring at a particular time frame for its affected regions. However, the original dataset did not have a clear identifier that we could use universally in this manner. Although there was a column named ETN for the event ID, the value was not unique across the dataset. In reality, a weather event is not a discrete object, so there needs to be an assumption to assign an ID to weather an event. Therefore, we assumed that a weather alert can be identified by when, where, and what. We identify weather events with their duration (when), significance, phenomenon type (what), and state (where). Hence, we combined the values of the ISSUED, EXPIRED, PHENOM, SIG, and STATE columns and used them as the ID value for each event. Note that the same event simultaneously happening in multiple states will have multiple identifiers and be considered to be

[1]https://mesonet.agron.iastate.edu/info/datasets

TABLE 1: Part of the list of columns of VTEC archive dataset

Column	Description
WFO	A three-character identifier for NWS Offices/Centers.
ISSUED	The timestamp represents the event's start time. During an event's lifecycle, the NWS can update this issued value. This value represents the last known state of the event start time.
EXPIRED	Like the ISSUED column above, this represents the event's end time and is the last updated value.
PHENOM	This is the two-character NWS code for VTEC events. (e.g., TO for Tornadoes, SV for Severe Thunderstorms.)
SIG	This is the one character NWS code for VTEC significance. (e.g., W for warnings, Y for advisory)
ETN	This is the VTEC event identifier, a unique value for the combination of an issuance center and a continuous spatial region for the event.
GTYPE	Either P for polygon or C for county/zone/parish.
STATUS	The VTEC status code indicates the stage of an event's life cycle. (e.g., NEW for new events, EXP for expired events)
Geometry	Area impacted by the event.

different events. Finally, using the geometry of the counties, we added a FIPS code and STATE column, computed the duration for each event, and added the YEAR column using the value of the ISSUED column.

Table 2 displays samples of the preprocessed dataset. The first two rows represent a gale event (with a PHENOM value of "GL"; see the full list in Table 3) that occurred in January 2018 and affected Delaware and New Jersey. Note that there can be multiple rows with the same event ID if a single event simultaneously affects multiple counties in the same state.

There is always the possibility of errors and missing data when it comes to large datasets, such as the NWS VTEC dataset. The calculated durations of events were unreasonable in some cases. In one instance, the EXPIRED time stamp was earlier than the ISSUED date, resulting in a negative event duration. In another instance, the duration was extremely long—over seven days. Out of a total of 447,266 data rows from 2018 to 2022, 992 rows (0.221%) had negative durations, and 1,503 rows (0.336%) had durations of more than seven days. We filtered out these outlier cases and retained the remaining 444,771 rows (99.442%). The number of unique EVENT_IDs was 187,073 on the retained data. Figure 2 displays the distribution of data rows with different event durations. Of the data, 419,122 rows (94.2%) had durations of less than 24 hours, and 17,943 rows (4.0%) lasted between 24 and 48 hours. Events lasting longer than 48 hours were rare.

Figure 3 presents the total number of unique weather events in the preprocessed dataset per year from 2018 to 2022. The highest number of events was observed in 2019, followed by a decline in 2020. The number of unique weather events experienced a steady increase from 2020 to 2022.

Figure 4 presents the average number of unique event IDs per year for each state from 2018 to 2022, sorted from largest to smallest. Although the total number of events per state varied annually, Texas and Florida had the highest average number of events, followed by Kansas. Figure 5 showcases the top 50 counties with the highest average yearly total number of events from 2018 to 2022. Notably, Florida had the most number of counties (18) in this top 50 ranking list, followed by Arizona (7). In contrast, Texas had only one county on the list. This observation highlights the importance of county-level analysis, as significant findings can be overlooked when data is aggregated at the state level.

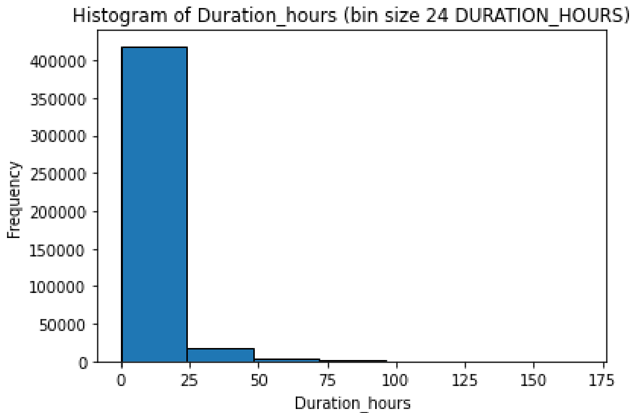


FIGURE 2: Distribution of data rows with different event duration (in hours). There can be multiple rows for the same event, so the numbers indicate the count of rows, not the count of distinct events.

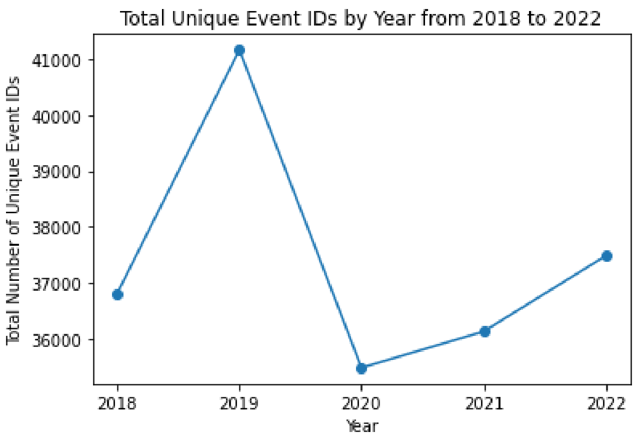


FIGURE 3: Total unique event IDs by year from 2018 to 2022.

For instance, Texas was ranked 1st and Arizona 12th when examining the number of unique event IDs by state.

Different weather event types exhibit varying levels of prevalence. Figure 6 displays the average number of unique event IDs per year for each weather event type. Severe thunderstorms were the most prevalent type of event, followed by

TABLE 2: Samples of the preprocessed NWS VTEC dataset

EVENT_ID	ISSUED	EXPIRED	FIPS	SIG	PHENOM	COUNTY	STATE	YEAR	DURATION
201801010227-201801011100-GL-W-DE	1/1/18 2:27	1/1/18 11:00	10005	W	GL	Sussex	DE	2018	0 days 08:33:00
201801010227-201801011100-GL-W-NJ	1/1/18 2:27	1/1/18 11:00	34009	W	GL	Cape May	NJ	2018	0 days 08:33:00
201801011327-201801021900-UP-W-MA	1/1/18 13:27	1/2/18 19:00	25007	W	UP	Dukes,MA	MA	2018	1 days 05:33:00
201801011405-201801030000-GL-W-FL	1/1/18 14:05	1/3/18 0:00	12031	W	GL	Duval,FL	FL	2018	1 days 09:55:00

TABLE 3: Phenomena codes and descriptions (some of them seem to be missing)

Code	Description	Code	Description	Code	Description
BZ	Blizzard	WI	Wind	AV	Avalanche
WS	Winter Storm	HW	High Wind	TS	Tsunami
WW	Winter Weather	LW	Lake Wind	MA	Marine
SN	Snow	FG	Dense Fog	SC	Small Craft
HS	Heavy Snow	SM	Dense Smoke	GL	Gale
LE	Lake Effect Snow	HT	Heat	SR	Storm
LB	Lake Effect Snow & Blowing Snow	EH	Excessive Heat	HF	Hurricane Force Wind
BS	Blowing/Drifting Snow	DU	Blowing Dust	TR	Tropical Storm
SB	Snow & Blowing Snow	DS	Dust Storm	HU	Hurricane
IP	Sleet	FL	Flood	TY	Typhoon
HP	Heavy Sleet	FF	Flash Flood	TI	Inland Tropical Storm Wind
HI	Inland Hurricane Wind	LS	Lakeshore Flood	CF	Coastal Flood
ZF	Freezing Fog	FZ	Freeze	FR	Frost
WC	Wind Chill	EC	Extreme Cold	AS	Air Stagnation
RH	Radiological Hazard	VO	Volcano	UP	Ice Accretion
LO	Low Water	SU	High Surf	FA	Flood
EW	Extreme Wind	EH	Extreme Heat	IS	Ice Storm

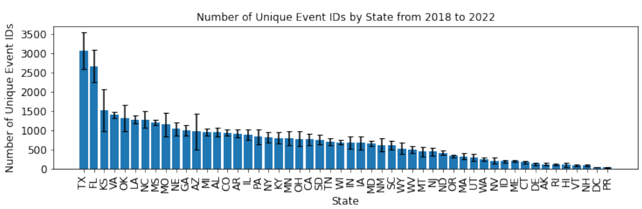


FIGURE 4: Average annual number of unique event IDs by state with error bars showing standard deviation (from 2018 to 2022).

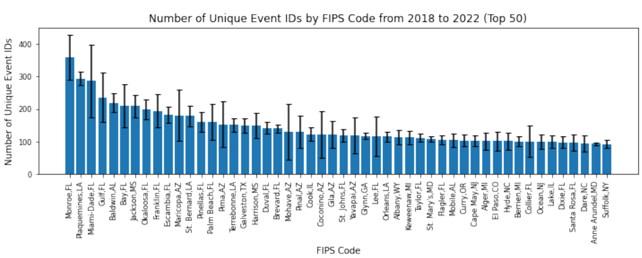


FIGURE 5: Average annual number of unique event IDs by county with error bars showing standard deviation (top 50, from 2018 to 2022).

marine (MA), flash flood (FF), tornado (TO), and flood (FA).

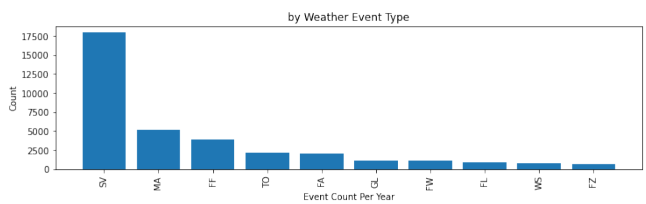


FIGURE 6: Average annual number of unique event IDs by weather event type (top 10).

B. EAGLE-I DATA

For US power outage information, we leverage the data collected by the ORNL's EAGLE-I system[2]. EAGLE-I is an interactive geographic information system (GIS) operated by ORNL and sponsored by DOE that enables users to visualize and map the nation's energy infrastructure, including assets at risk such as the electric grid, petroleum, and natural gas infrastructures. The EAGLE-I platform has been collecting US power outage information data since 2014. The historical dataset is available for academic use.

EAGLE-I's electricity outage data comprises records of the total number of customers without power in a geographical district at the county level, along with information about the associated utility company. Note that more than one utility company could report outages for a single county if multiple utility companies serve the county. Thus, we

[2]https://eagle-i.doe.gov/

aggregated the data at the county level and calculated the total outage count for each time stamp. Table 5 displays samples of the EAGLE-I power outage data. Each row contains the total number of outages for the county with the FIPS code at the time stamp. More specifically, the number of outages represents the number of customers affected. This outage information is recorded every 15 minutes; however, no data record is available if the outage number is zero. In this example data, the number of customers affected in the county with the FIPS code 9015 was 12 at 2022-01-02 09:30. Although the EAGLE-I power outage data does not track individual power outages and contains only the total number of customers affected in a county at a given time, we can infer restoration times from the records based on certain assumptions by tracking the number of power outages. For example, no data was recorded at 2022-01-02 09:45, indicating that the outages affecting 12 customers were fully resolved within 15 minutes.

The coverage ratio represents the proportion of Eagle-I coverage compared with the total number of customers in the respective state. Over time, the customer coverage provided by the EAGLE-I platform has been expanding, driven by the acquisition of data from more utility companies that serve a wider range of regions. The coverage ratio provides important information about the public reporting of the outage dataset. Figure 7 displays the maximum national coverage ratio annually from 2018 to 2022, considering all 50 states and Washington, DC. This figure shows the increasing trend of outage data coverage. On average, the coverage ratio was a commendable 0.871. Table 4 shows that EAGLE-I data reporting for all states except Colorado and Kansas from 2018 to 2022 has improved in terms of number of customers. This increased trend of the coverage ratio shows that more power outage data are publicly available every year for academic use. Although the "Total Customers" column shows the total number of customers for 2022, the coverage ratios are calculated based on the number of customers of the specific year.

Because the coverage ratio for EAGLE-I data is not perfectly 1.0, there is a possibility that some customers might experience power outages within regions that fall outside the EAGLE-I scope. This could lead to an underestimation of power outage numbers derived from the data. To address this, we correct the power outage figure by dividing it by the coverage ratio. This adjustment uses the 5-year national average coverage ratio of 0.871. Note that all power outage figures presented in subsequent sections are these adjusted to 0.871 coverage ratio numbers.

C. MAPPING EAGLE-I DATASETS AND THE NWS DATASETS

Our objective is to map the NWS data and EAGLE-I datasets, using the geographical and temporal information from both sources to study power outage data associated with severe weather events. Specifically, we aim to understand how power outages occur and are restored before, during, and after weather events in a given region.

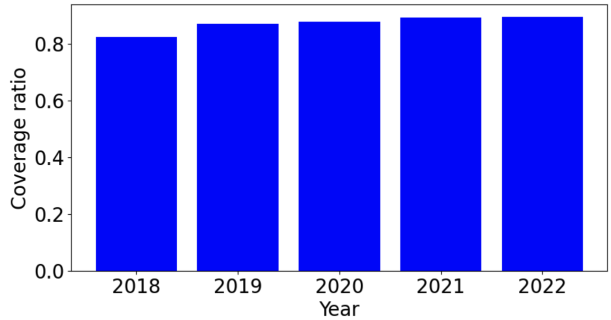


FIGURE 7: Eagle-I coverage ratio for 2018-2022.

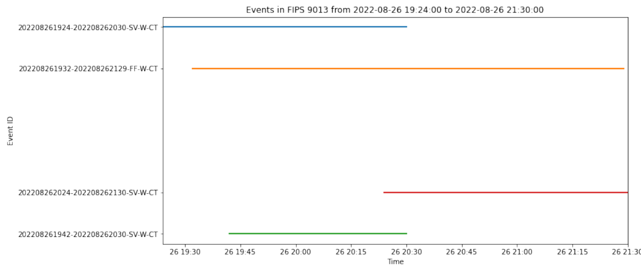


FIGURE 8: Visualization of a weather event group composed of five different weather events.

It is common for multiple related weather events to occur simultaneously or in close succession within a region, so instead of analyzing each weather event's impact to power systems, we need to identify groups of weather events that are highly relevant. Figure 8 illustrates four weather events from the NWS dataset spanning the timeline in Tolland County, Connecticut (FIP 9013), from 2022-08-26 19:24:00 to 2022-08-26 21:30:00. FF and SV represent flash floods and severe thunderstorms, respectively, and it is evident that correlated events can occur or co-occur in proximity to one another. We define a *weather event group* as a collection of events occurring in a specific county in which the group encompasses all weather events that overlap in duration with any other events within the group. For example, all events shown in Fig. 8 were grouped together and assigned the same group ID. Additionally, for each weather event group, we calculated the group duration, which is the duration from the start of the first event to the expiration of the last event in the group. We then assigned the type of the longest weather event in the group as the type for the entire weather event group.

Figure 9 displays a power outage pattern corresponding to a weather event group consisting of events depicted in Fig. 8. The red vertical line signifies the beginning of the weather event group, while the green vertical line indicates its end.

Although we cannot definitively state that the weather events directly caused the power outages, there seems to be a strong correlation between the weather event group and the increase in power outage numbers in this instance, as we

TABLE 4: EAGLE-I coverage over time (from 2018 to 2022)

State/Territory	Total Customers (2022)	Coverage Ratio					
		2018	2019	2020	2021	2022	Overall Change (2022-2018)
AK	364614	0.67	0.67	0.68	0.68	0.71	0.04
AL	2615212	0.79	0.81	0.81	0.81	0.81	0.02
AR	1630606	0.78	0.83	0.86	0.86	0.86	0.08
AZ	3188212	0.87	0.86	0.9	0.91	0.91	0.04
CA	15666677	0.91	0.93	0.94	0.95	0.95	0.04
CO	2768968	0.88	0.83	0.84	0.83	0.84	-0.04
CT	1680077	0.95	0.95	0.95	0.95	0.96	0.01
DC	317140	1.0	1.0	1.0	1.0	1.0	0.0
DE	499849	0.86	0.86	0.86	0.86	0.86	0.0
FL	10989517	0.97	0.97	0.97	0.98	0.98	0.01
GA	5059196	0.9	0.86	0.88	0.9	0.91	0.01
HI	502534	0.61	0.66	0.67	1.0	1.0	0.39
IA	1601073	0.89	0.89	0.91	0.91	0.91	0.02
ID	904872	0.91	0.92	0.95	0.9	0.91	0.0
IL	5906057	0.92	0.99	0.94	0.98	0.98	0.06
IN	3228957	0.79	0.82	0.86	0.86	0.86	0.07
KS	1980177	0.71	0.74	0.79	0.8	0.59	-0.12
KY	2316157	0.78	0.92	0.92	0.92	0.94	0.16
LA	2400685	0.89	0.89	0.9	0.9	0.91	0.02
MA	3243537	0.89	0.89	0.94	0.99	0.99	0.1
MD	2635074	0.98	0.97	0.98	0.99	0.99	0.01
ME	816054	0.96	0.96	0.96	0.96	0.96	0.0
MI	4933959	0.91	0.96	0.95	0.95	0.96	0.05
MN	2724798	0.71	0.71	0.8	0.8	0.8	0.09
MO	3160891	0.8	0.95	0.95	0.95	0.95	0.15
MS	1546679	0.66	0.74	0.77	0.77	0.77	0.11
MT	634592	0.16	0.75	0.76	0.76	0.76	0.6
NC	5297545	0.82	0.9	0.91	0.92	0.92	0.1
ND	463692	0.51	0.53	0.66	0.66	0.66	0.15
NE	1034499	0.45	0.46	0.48	0.48	0.48	0.03
NH	738754	0.88	0.99	0.99	0.99	0.99	0.11
NJ	4136585	0.99	0.98	0.99	0.99	0.99	0.0
NM	1045100	0.74	0.81	0.82	0.82	0.82	0.08
NV	1390235	0.95	0.95	0.96	0.97	0.97	0.02
NY	8309187	0.98	0.98	0.99	0.99	0.99	0.01
OH	5603381	0.93	0.93	0.93	0.93	0.93	0.0
OK	2047920	0.89	0.9	0.93	0.93	0.94	0.05
OR	2045351	0.75	0.86	0.9	0.9	0.9	0.15
PA	6111114	0.95	0.96	0.98	0.98	0.98	0.03
RI	505947	0.99	0.99	0.99	0.99	0.99	0.0
SC	2800722	0.99	0.99	0.99	0.99	0.99	0.0
SD	475661	0.46	0.87	0.79	0.79	0.89	0.43
TN	3433064	0.55	0.73	0.73	0.73	0.75	0.2
TX	14691490	0.9	0.93	0.63	0.94	0.93	0.03
UT	1252077	0.8	0.8	0.79	0.79	0.8	0.0
VA	3941323	0.96	0.97	0.97	0.97	0.97	0.01
VI	56133	1.0	1.0	1.0	1.0	1.0	0.0
VT	376122	1.0	1.0	1.0	1.0	1.0	0.0
WA	3532113	0.71	0.86	0.87	0.87	0.87	0.16
WI	3068156	0.83	0.84	0.84	0.85	0.85	0.02
WV	1010110	1.0	1.0	1.0	0.97	1.0	0.0
WY	329909	0.62	0.68	0.71	0.71	0.71	0.09

TABLE 5: Samples of EAGLE-I power outage data

FIPS	TIMESTAMP	OUTAGE_SUM
9015	2022-01-02 09:30:00	12
9015	2022-01-02 10:00:00	1
9015	2022-01-02 10:15:00	1
...	...	...
9015	2022-01-02 12:15:00	1
9015	2022-01-02 12:30:00	1
9015	2022-01-02 14:15:00	2

observe the increase in power outages immediately after the beginning of the weather event group.

In Fig. 9, the peak number power of outages reached is 2657; nevertheless, note that the peak number of outages was observed after the weather event group had expired and that the outages persisted many hours. So we need to consider not only power outages during weather events but also power outages after the events. The EAGLE-I data is updated at 15-minute intervals, but there is no guaranteed maximum delay of 15 minutes in the process. The EAGLE-I platform collects

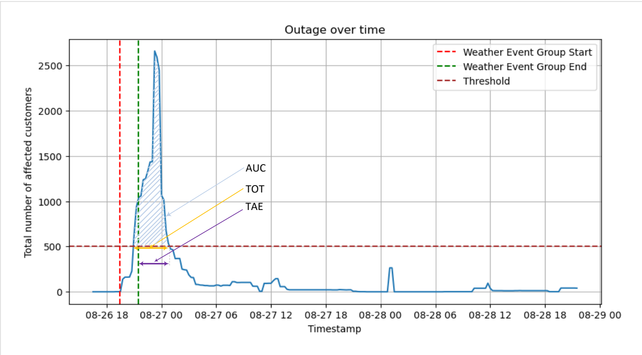


FIGURE 9: Power outage pattern from 3 hours before to 48 hours after a weather event group (FIPS=9013 from 2022-08-26 19:24:00 to 2022-08-26 21:30:00).

power outage data through the utility company's data API or their website, which could result in additional delays on the utility company's end. Another potential reason for observing an increase in power outages after the end of an extreme weather event is that the actual damage might occur after the weather event because of delayed infrastructure damage (probably because of weak and aging infrastructure), flooding aftermath, ice accumulation, and other related factors. We have quantified the continued power outage duration after the end of an extreme event as one of the resilience metrics for further analysis.

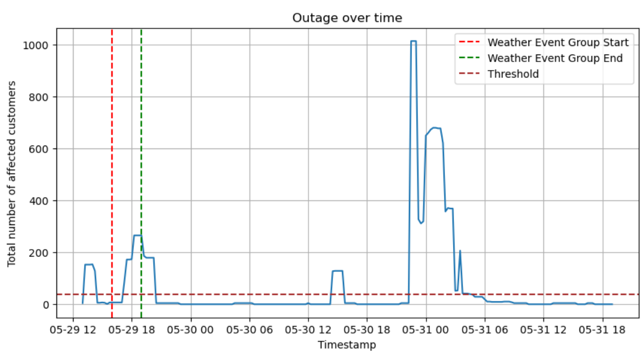


FIGURE 10: Power outage pattern from 3 hours before to 48 hours after a weather event group (FIPS=1057 from 2018-05-29 16:00:00 to 2018-05-29 19:00:00).

Figure 10 illustrates another power outage pattern for a different weather event group (comprising a single flood event) in Fayette County, Alabama, on May 29, 2018. We observe multiple peak outages before, during, and after the active weather event group. Power outages observed before the active weather alert can still be correlated to the weather alert, as weather events can be unpredictable, leading to delays in issuing weather alerts. Another possible explanation of the outage some time after (curve at around 05-31 00, as shown in Fig. 10) the end of the event is that the delayed infrastructure damage led to further power outages. We cannot be absolutely certain that power outages before, during, and after weather events are caused by these events; however, it is reasonable to assume a connection, especially if we observe unusually high power outage numbers.

Power is restored to the majority of customers within 48 hours after an extreme weather event [19]. Thus, we considered any outages that occurred within the time frame of 3 hours before (to accommodate the uncertainty with NWS data recording) to 48 hours after the expiration time of the weather event group to be correlated with the weather event group. In addition, we used this information to calculate the threshold value to distinguish the power outages caused by extreme events vs. power outages caused by other regular events (e.g., load change and system faults). This threshold helps us to determine the accurate power outages caused by extreme events, and it is more realistic than the arbitrary 25% value used by [5]. This calculation of threshold value is possible because of the mapping of NWS data on the EAGLE-I dataset. In this process, a regular outage dataset is created by filtering out all the outage data at the time of extreme weather events and up to 48 hours after the end of the events. The average value of the regular outage dataset is taken as the threshold to distinguish power outages caused by extreme events from outages due to other causes. For example, the horizontal dashed line in Fig. 9 is the threshold line. During extreme events, all the outages above this threshold are considered outages due to extreme events. Outages below the threshold are due to other regular causes and are therefore excluded from the resilience analysis.

Nevertheless, the threshold value is configurable, and, depending on the analytic purposes, the same threshold can be used for all counties. If we use a lower threshold, more power outages will be related to extreme weather.

The proposed approach can be summarized through a flow chart, as shown in Fig. 11.

IV. RESILIENCE ANALYSIS

This section provides assessment of the resilience of the US power systems to extreme weather events using the proposed automated data framework. To provide numerical assessment of the US power system, quantification metrics are necessary. Although there are several power system resilience quantification metrics in the literature, they have not yet been standardized or universally accepted [4, 13]. Several efforts have been made, for example [4, 11, 14, 15, 16, 20], to capture the resilience features (e.g., withstand, adapt, and restore or recover) for the assessment of power system resilience. Motivated by these conventions, we propose the following resilience metrics to study and quantify the power system resilience against extreme weather events.

- Power outage time over threshold (TOT): TOT represents the time between an outage curve crossing the threshold line, as shown in Fig. 9. TOT provides the information about the duration of power outages experienced by customers due to extreme events.
- Power outage area under curve (AUC): TOT gives information about the duration of outages; however,

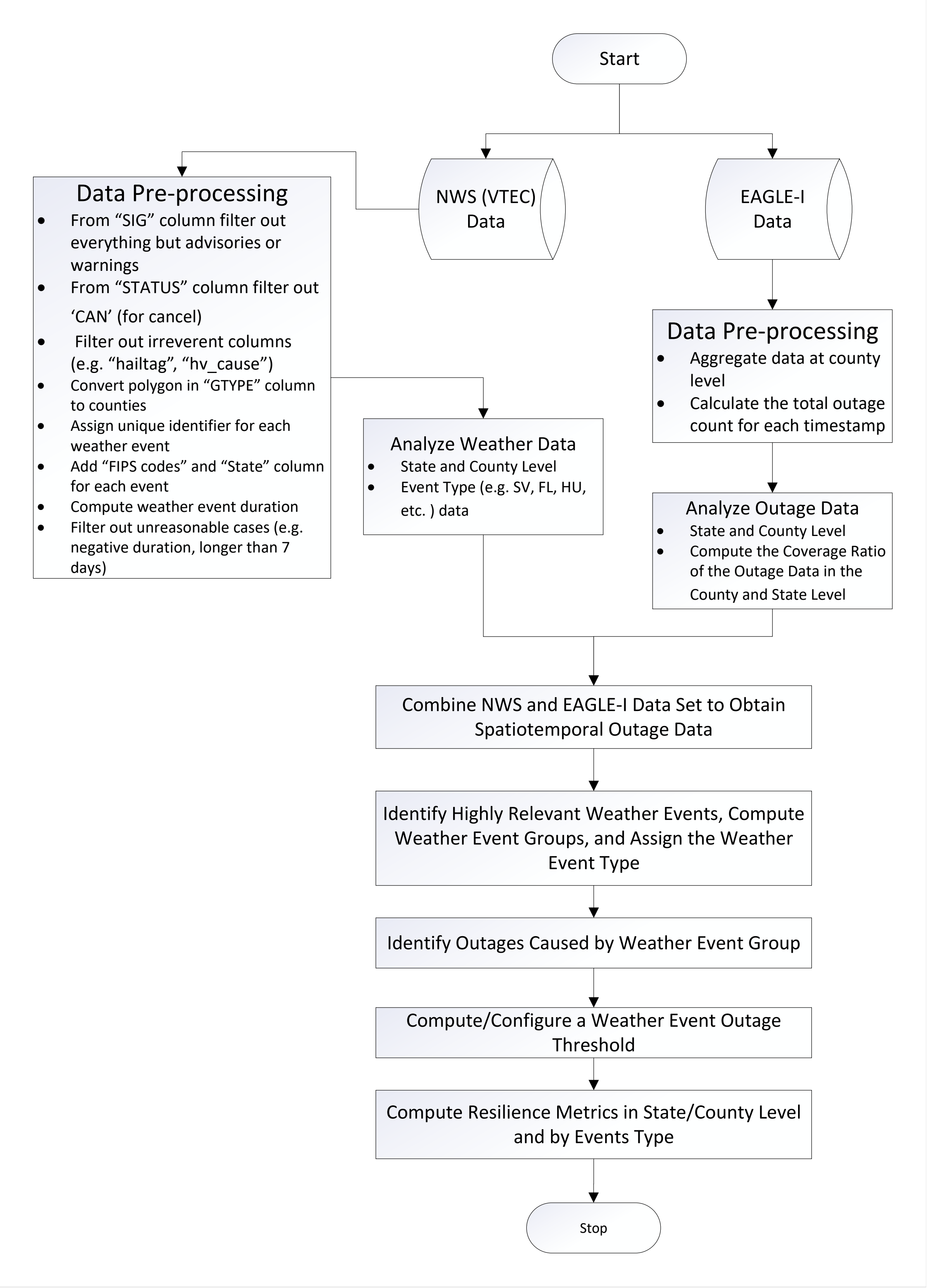


FIGURE 11: Flowchart of the Proposed Approach.

it does not convey information about the number of customers affected. To account for the number of customers affected, along with the total outage time, we calculated the area under the curve as another quantification metric. AUC is the area under the curve above the dotted threshold outage line, as shown in Fig. 9. AUC is normalized—outage impact per customer in minutes—(sum of the area under the curve for each event divided by the total number of customers) in this work to compare the AUC at the state and county levels and by event type.

- Power outage time after the end of the event (TAE): TAE is calculated to determine how long a power outage event continues after the end of an extreme weather event. A power outage after the end of a event could be caused by a delay in weather and power outage data collection. Another important reason for continued power outage after an event signifies how quickly the system is bouncing back. A system's response depends on several factors (e.g., system condition and available resources), an important one being the physical condition of the grid. Physically weak and aging infrastructures have more probability of failure resulting longer response time due to repair and installation requirement. Therefore, TAE gives some prospective on the physical condition of the power grid and hardening requirement. TAE does not include the immediate (within the event duration) recovery response.

Both TOT and AUC information are important to determine the extent of the affects on customers. There are four conditions for customer impacts: (1) low TOT and low AUC signify that very few customers were impacted and that they experienced shorter power outage durations (this is characteristic of a more resilient power system); (2) low TOT and high AUC signify that many customers were impacted for a shorter duration of time; (3) high TOT and low AUC signify that a few customers experienced prolonged power outages; and (4) high TOT and high AUC signify that a large number (wide spread) of customers experienced prolonged outages (this indicates a less/poor resilient power system).

Note that the time difference between TOT and TAE gives information about the capability of a utility company to immediately restore power to customers (by repairing the damaged infrastructure, etc.) during the extreme event and to what extent other interrelated networks (communication and road network) are damaged. High TOT-TAE indicates, to some extent, damage to the electrical infrastructure (e.g., distribution and transmission poles and lines), the severity of the event, damage to the interrelated infrastructures (e.g., road network and communication network), and and unprepared or resource-scarce utility company.

All of the resilience metrics are quantified for each weather event group occurring within a county. For our analysis, we measured the duration of power outages over the threshold within a time window spanning from 3 hours before to 48 hours after weather event groups. This is because power outages might start before weather events and can persist afterward, as depicted in Fig. 9 and Fig. 10. (Note that an actual power outage due to an event does not occur before an event; nevertheless, because of the uncertainty in the NWS data collection, outages are seen before the start of the events in the outage datasets.)

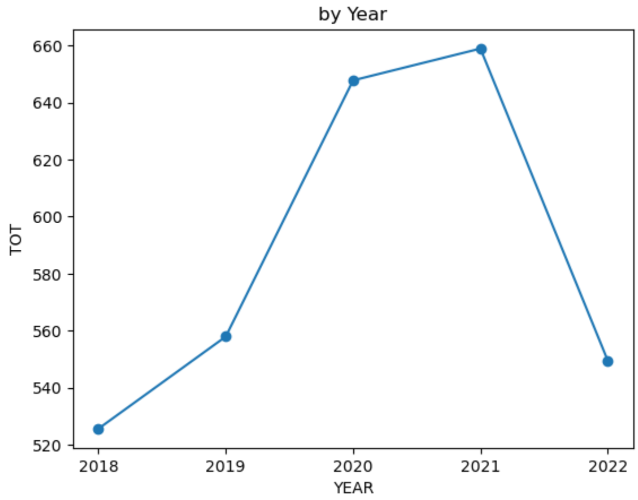


FIGURE 12: Annual average TOT from 2018 to 2022 (in minutes).

Figure 12 illustrates the changes in the yearly average TOT (in minutes) associated with weather events. We observe a continual increase from 2018 (525.591) to 2021 (658.932–peak value in the given range) and a decrease in 2022 (549.504). The 5-year average was 588.489.

Figure 13 illustrates the changes in the yearly average AUC—outage impact per customer in minutes—(the sum of the area under the curve for each event divided by the total number of customers) associated with weather events. We observe the average AUC (normalized) as 4.8 in 2018, 0.82 in 2019, 1.5 in 2020, 1.3 in 2021, and 0.92 in 2022. The 5-year average was 1.87 outage impact per customer in the United States.

Figure 14 illustrates the changes in the yearly average TAE (in minutes) associated with weather events. We observe a continual change from 2018 to 2022. The 5-year average TAE is 23.86 (minutes after the end of the events).

These figures show that there are no specific trends in therms of all the compared metrics.

As seen in Fig. 3, the number of weather events occurring in 2020 was the lowest, suggesting that there is no direct correlation between the number of weather events and the measured resilience metrics of the US power systems. Many other factors could have influenced this result, such as the size of weather events, the age of infrastructure, and maintenance practices.

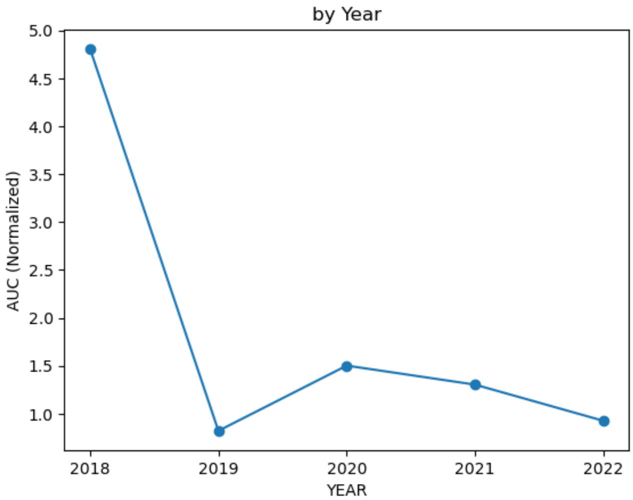


FIGURE 13: Annual average AUC (normalized) from 2018 to 2022 (cumulative customers impacted).

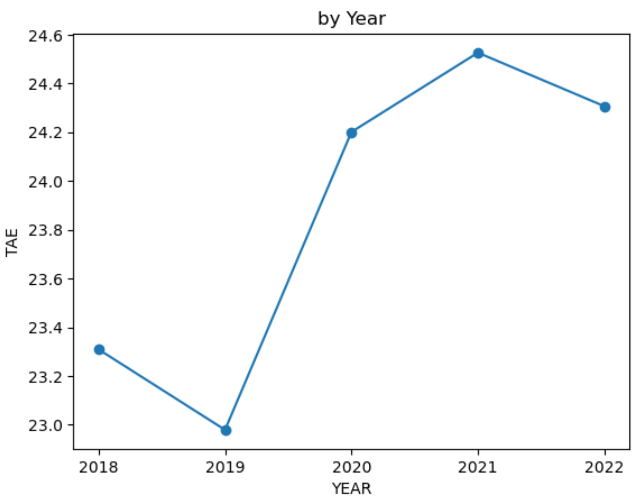


FIGURE 14: Annual average TAE from 2018 to 2022 (in minutes).

A. STATE AND COUNTY LEVEL ANALYSIS

The impact of weather events at the state and county level in terms of the proposed metrics can be presented as follows.

1) Power outage time over threshold

The blue bar in Fig. 15 depicts a comparison of the average TOT (sum of TOT of each event divided by the total number of events) across various US states and territories. California (CA) demonstrates the highest value of 1111.967, followed by Michigan (MI) of 979.428, and Louisiana (LA) of 835.603. These results show that events in these states are causing longer outages. Conversely, the states with the lowest average TOT are Wyoming of 271.242, Utah (UT) of 327.527, and Idaho (ID) of 357.438, indicating shorter outage duration per event in these states. Although TOT provides information about how long it takes for a power system to return to pre-event status after an extreme weather event, it alone cannot provide information about power system resilience level because it does not properly incorporate the number of customers. California has the most customers, which could be the reason for longer outage duration (outage impact per customer is analyzed with the AUC in the following section). The red bar chart in Fig. 15 represents the number of events exceeding the dotted outage threshold line in each state. A higher number indicates more events exceeding the outage threshold to the extreme weather events.

We did not observe a strong correlation between the number of reported cases exceeding the threshold line and the average TOT values. For instance, Oregon (OR) had less threshold exceeding cases (871), yet it had a high average TOT of 784.632, ranked at 5 out of 51 districts. States with high average TOT but low case numbers exceeding the threshold line—such as OR, WA, and HI—indicates that the events in these states are more impactful (i.e., have longer outages). Conversely, we noticed the high number of cases exceeding the threshold line in Texas (TX) with 15116 cases, Florida (FL) with 9540 cases, and Georgia (GA) with 6878 cases. However, their respective average TOT values: 639.221, 527.362, and 428.770, respectively, are either comparable with or lower than the overall average TOT value of 588.489. This means the the events in these states are less impactful (i.e., have shorter outages).

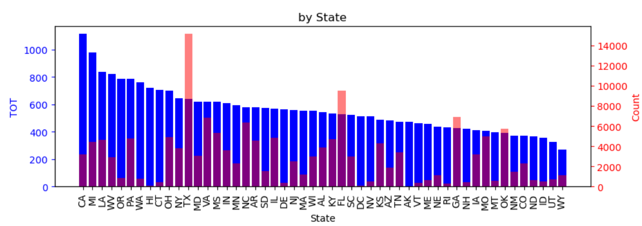


FIGURE 15: Average TOT by state.

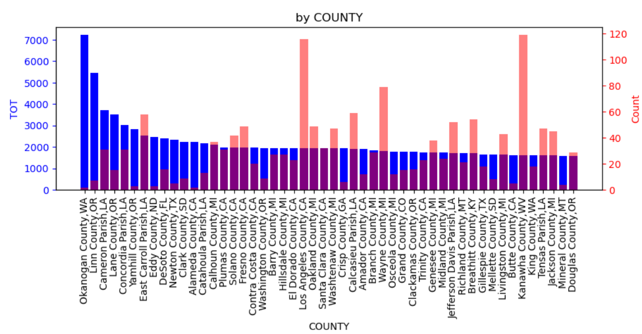
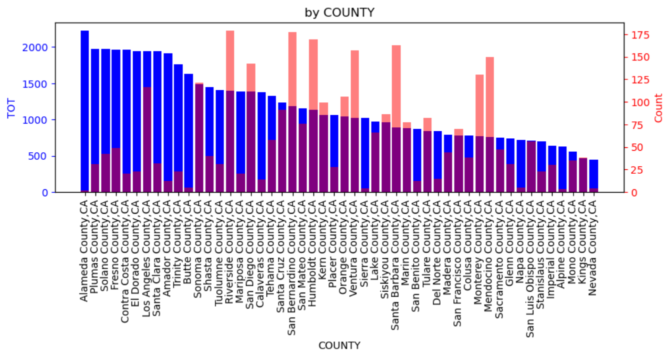
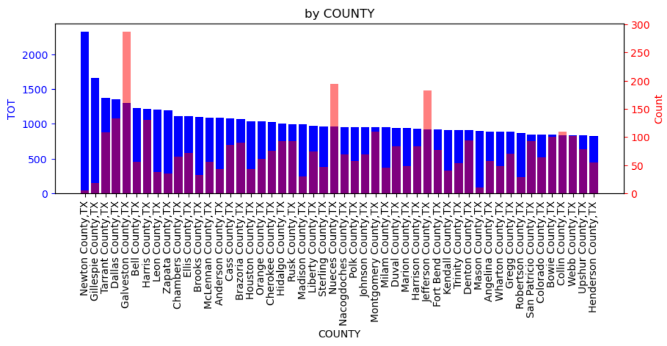


FIGURE 16: Average TOT by county.

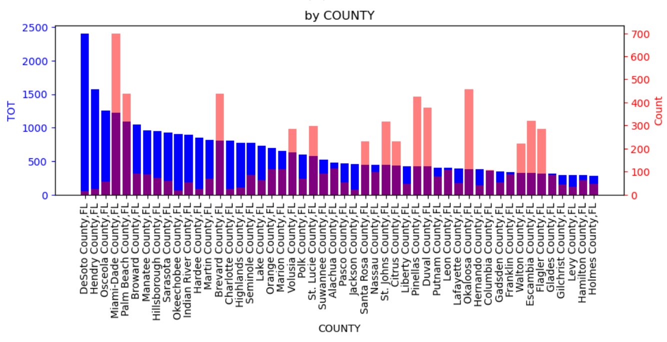
Figure 16 presents the top 50 counties with the highest average TOT. Okanogan County (WA) topped the list with an average TOT value of 7222.500, followed by Linn County (OR), with 5455.714, and Cameron Parish (LA), with 3720.483. The number of cases exceeding the threshold line in these counties was relatively low, with 2, 7, and 31 cases, respectively, indicating that power outages exceeding the threshold due to weather events were infrequent in these areas. However, when such events occur, they tend to result in prolonged outages.



(a) CA



(b) TX



(c) FL

FIGURE 17: Average TOT by county.

Figure 17 shows the average TOT of counties in CA, TX, and FL. As we can see, even within the same state, the variation of average TOT values is high. On average at the state level, CA had a higher TOT value compared with both TX and FL; however, there are many counties in TX and FL, which have higher average TOTs than those of counties in CA. This confirms that not only state-level resolution but also higher resolution analyses are needed.

In this work, we have compared the outage duration without incorporating the magnitude of the events (e.g., wind speed and flooding magnitude for flood); however, to compare the events more precisely, it is important to incorporate the the magnitude of the event in the resilience metric. This will be our future research work.

The state- and county-level results are presented in the US map in Fig. 18 and Fig. 19, respectively.

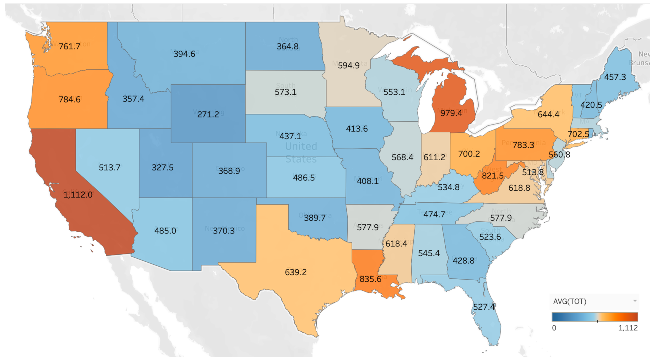


FIGURE 18: Average TOT by state.

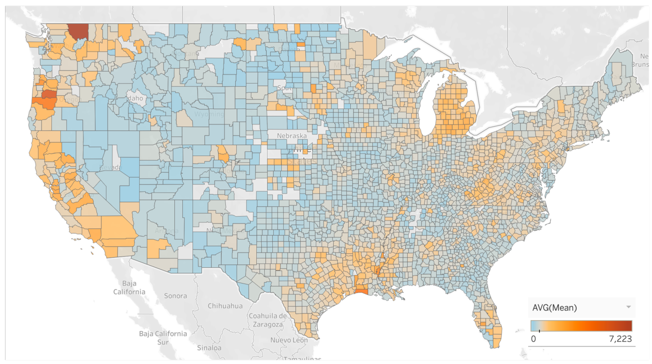


FIGURE 19: Average TOT by county.

2) Power outage area under curve

Table 6 provides the average AUC—outage impact per customer in minutes—(sum of AUC for each event divided by the total number of customers) by state associated with extreme weather events. This table indicates that the outage impact per customer in minutes (normalized average AUC) is highest in Virginia, followed by North Carolina, Louisiana, Texas, and so on. This shows that the individual customers in Virginia are the most heavily affected, indicating poor power grid resilience in terms of AUC. On the other hand, customers in Washington, DC, Delaware, and Arizona are the least affected, indicating relatively better power system resilience in terms of the AUC metric.

Table 7 provides the average AUC—outage impact per customer in minutes—(sum of the AUC for each event divided by total number of customers) by county (50 counties with the highest AUC) associated with extreme weather events. This table indicates that Waynesboro (VA) has seen the most per-customer impact, followed by Buena Vista (VA), Henderson (NC), Lexington (VA), and so on.

Although at the state level Virginia seems to be the worst performing state, this is mostly driven by three counties/cities (Waynesboro city, Buena Vista city, and Lexington city) because the average outage impact per customer of the rest of the counties is only 0.94 minutes. This indicates that Waynesboro city, Buena Vista city, and Lexington city need special attention to improve the power system resilience of the state of Virginia. Similarly, if we look at one of the least impacted states, for example Florida, DeSoto County is 46th among all the counties in the United states, demanding special attention in terms of outage impact per customer in minutes. Therefore, it is important to perform granular analysis of the outages to better prepare the most impacted areas.

3) Power outage time after the end of the event

The blue bar in Fig. 20 depicts a comparison of the average TAE (sum of TAE of each event divided by the total number of events in minutes) across various US states and territories. The bar chart shows that Hawaii has the highest TAE, followed by Delaware, California, Washington, DC, Florida, and so on. North Dakota, Montana, Idaho, Iowa, and Rhode Island are among the lowest TAE states. These results indicate that it takes the longest amount of time (despite of having very few events exceeding the threshold line) to repair the infrastructure in Hawaii and the shortest amount of time in North Dakota.

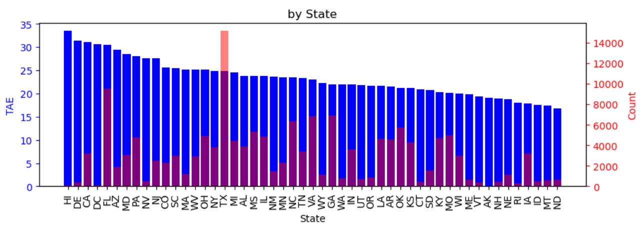


FIGURE 20: Average TAE by state.

The blue bar in Fig. 21 depicts a comparison of the average TAE (sum of TAE of each event divided by the total number of events in minutes) across the top 50 counties in the United States. The bar chart shows that Towner County (ND), Broward County (FL), Miami-Dade County (FL), Alameda County (CA), and Palm Beach County (FL) have the highest TAE. These results indicate that it takes longer (despite of having very few events exceeding the threshold line) to repair the infrastructure in these counties.

Longer TAE means more damaged electrical infrastructure and possibly other correlated infrastructure (e.g., communication and road network); however, we cannot ignore the magnitude of the events in each case. Therefore, our future work is to incorporate the magnitude of the event in the resilience metrics calculations.

TABLE 6: Normalized AUC (outage impact per customer in minutes) and number of events exceeding threshold by state

STATE/TERRITORY	AUC (Normalized) Average	Event Counts
VA	16.59527437	6856
NC	14.35562963	6113
LA	1.486094215	4603
TX	1.464276922	15116
MI	0.982493214	4415
MS	0.82509542	5344
CA	0.755563057	3168
WV	0.68075874	2935
SD	0.650614191	1520
MT	0.620723784	560
RI	0.565342403	285
AR	0.526769223	4552
ID	0.510142759	527
ME	0.503731345	628
KY	0.500728472	4705
NY	0.471289467	3837
VT	0.451072603	410
CT	0.45011935	468
AL	0.44600515	3886
IA	0.431426086	3196
ND	0.424015878	643
WY	0.422947932	1090
MA	0.415828686	1180
GA	0.40435656	6878
WA	0.396937028	781
WI	0.39030809	3018
OR	0.338603471	871
MN	0.330480555	2321
KS	0.311283907	4324
IN	0.303854302	3586
OH	0.303216771	4884
NE	0.293938124	1134
PA	0.259884987	4747
OK	0.24978558	5727
MO	0.245036721	4962
IL	0.24251287	4823
SC	0.223674508	2994
NJ	0.222577386	2518
NM	0.217991348	1470
FL	0.213642574	9540
NH	0.20174064	424
TN	0.185489274	3372
HI	0.177588582	115
AK	0.171831129	63
CO	0.140916608	2296
NV	0.138750568	541
UT	0.131288415	740
MD	0.107413659	3070
AZ	0.106808477	1907
DE	0.070587611	382
DC	0.021646113	130

B. WEATHER EVENT TYPES

This section provides the analysis of how different weather event types impact power systems in terms of the proposed metrics.

1) Power outage time over threshold

Figure 22 depicts the average TOT for major weather event group types in the United States. Extreme Wind (EW) ranked first with a value of 10786.666, followed by Tropical Storm (TR) at 5145.000, High Surf (SU) at 2109.545, Storm (SR) at 1421.982, and Ice Storm (IS) at 1330.421. If we ignore

TABLE 7: Normalized AUC (outage impact per customer in minutes) and number of events exceeding threshold by county/city (top 50 counties/cities in terms of AUC)

County/City	AUC (Normalized) Average	Event Counts
Waynesboro city,VA	3137.351364	28
Buena Vista city,VA	1107.848555	17
Henderson,NC	782.0821355	85
Lexington city,VA	160.3547439	13
Jeff Davis,TX	106.1521333	26
New Hanover,NC	43.94577692	186
Eddy County,ND	42.45527929	3
Brewster,TX	33.18640674	43
Alamance,NC	30.1593694	83
Edwards,TX	21.0306568	50
Niobrara,WY	20.3530807	17
Hamilton,NY	18.45565422	28
Pamlico,NC	18.31150337	54
Calaveras,CA	15.0333035	14
Bandera,TX	14.83818733	30
Clay,GA	13.59388089	22
Vance,NC	13.18136239	41
Jones,NC	12.77636036	30
Hyde,NC	12.40177629	95
Robertson,TX	12.03211373	29
Butte,ID	11.85758594	6
Caswell,NC	10.78120772	36
Austin,TX	10.06645712	56
Terrell,TX	9.807070503	29
Madison,TX	9.475823648	30
Gillespie,TX	9.424912583	18
Lac qui Parle,MN	8.950037584	8
Cameron Parish,LA	8.738896875	31
Mineral,MT	8.456057107	4
Wharton,TX	7.969166093	48
Tuolumne,CA	7.919039827	31
Real,TX	7.810912253	32
Lunenburg,VA	7.445799494	20
Amador,CA	7.351623007	12
Kinney,TX	7.242181213	25
Alpine,CA	7.185482113	3
Webster,GA	7.076485858	13
Plumas,CA	6.93486849	31
Waller,TX	6.852374185	55
Alcona,MI	6.62276117	16
Nantucket,MA	6.615854555	20
Mason,TX	6.610827325	10
Sutton,TX	6.529470975	32
Lafourche Parish,LA	6.443906556	36
DeSoto,FL	5.988746355	16
Orange,TX	5.931742994	61
Pender,NC	5.907983287	48
St. Helena Parish,LA	5.861437846	75

the magnitude of the weather event type, these weather event types are causing longer outages.

We observe that weather event types with high average TOT values (more impactful) are less frequent in terms of the number of events exceeding the threshold line (red bar shows the number of events). On the other hand, severe thunderstorms (SV) was the most frequent cause of TOT (in terms of the number of events exceeding the threshold line); however, the average TOT for SV is 595.171, which is lower than many other less frequent weather event types.

Figure 23 and Fig. 24 provide the average TOT by weather event type for California and Florida. These figures show

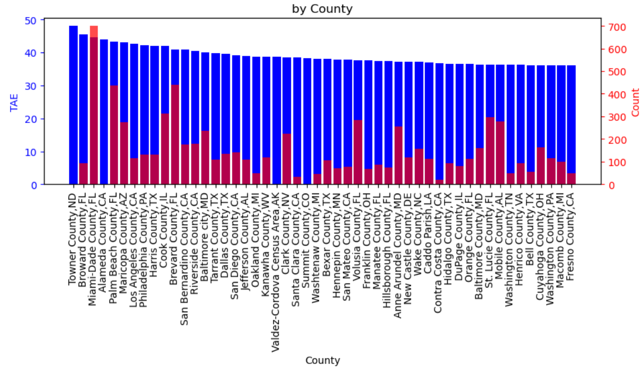


FIGURE 21: Average TAE by county.

that the most impactful in terms of TOT weather event is storm (SR) in the state of California and tropical storm (TR) in Florida. This suggests that California needs to put more emphasis on developing strategies against SR while Florida needs to develop strategies against TR to improve TOT.

Note that although there is some observation that less frequent events are beginning to have more impact in terms of the duration of power outages, less frequent events are not always the more impactful cases. Since the impact of weather event type heavily depends on the magnitude of the event types, the magnitude of the event types needs to be incorporated in the power outages to properly compare the impact of weather events on power systems. This will be our future research work.

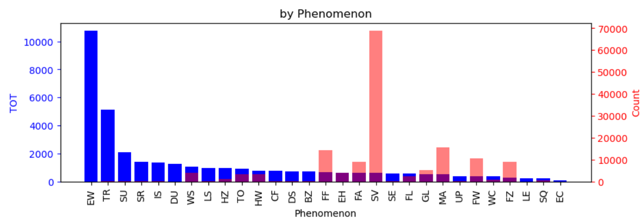


FIGURE 22: Average TOT by weather event type. Refer to Table 3 for descriptions of weather event types.

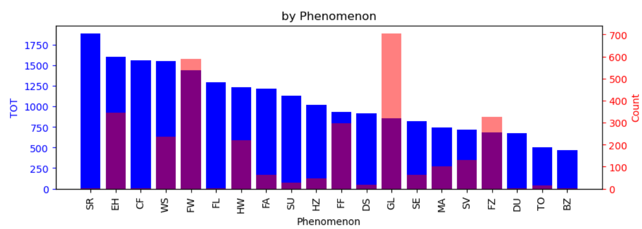


FIGURE 23: Average TOT by weather event type for California. Refer to Table 3 for descriptions of weather event types.

2) Power outage area under curve

Table 8 provides the average AUC—outage impact per customer in minutes—(sum of AUC for each event divided

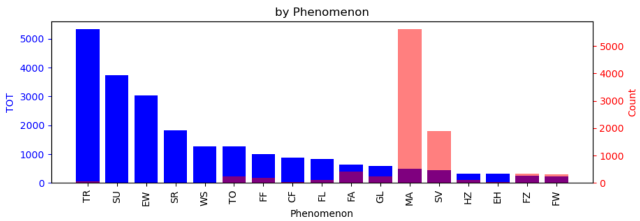


FIGURE 24: Average TOT by weather event type for Florida. Refer to Table 3 for descriptions of weather event types.

by the total number of customers) by weather event types. This table shows that extreme wind (EW) seems to be most impactful in terms of normalized AUC, followed by wind storm (WS), tropical storm (TR), high surf (SU), and so on.

Figure 25 and Fig. 26 provide the normalized AUC by weather event type for California and Florida. These figures show that the most impactful weather event in terms of normalized AUC is storm (SR) in the state of California and extreme wind (EW) in Florida. This suggests that California needs to put more emphasis on developing strategies against SR while Florida needs to develop strategies against EW to improve AUC.

As mentioned previously, the magnitude of events needs to be incorporated into these assessments, and this will be our future research work.

TABLE 8: Normalized AUC (outage impact per customer in minutes) and number of events exceeding threshold by weather event type. Tefer to Table 3 for descriptions of weather event types.

Weather Event Type	AUC (Normalized) Average	Event Counts
EW	39.15921714	9
WS	27.67849325	3882
TR	17.08418539	162
SU	6.399028663	55
WC	4.78087939	714
HZ	4.367593496	1205
FL	4.013990338	2402
FF	2.443149496	14218
DU	2.137192453	22
FA	2.097149557	8875
SR	1.889431622	115
IS	1.744756457	190
TO	1.163922854	3463
CF	1.112326013	324
HW	1.025952819	3450
SV	1.002797422	68847
BZ	0.607286806	253
FZ	0.603463098	8938
DS	0.513772292	146
LS	0.496723425	146
MA	0.357315886	15607
FW	0.290463951	10396
GL	0.272037721	5295
EH	0.264263687	3942
UP	0.195987502	140
SE	0.138693018	217
SQ	0.095396101	627
LE	0.01406229	14
EC	0.000663974	1

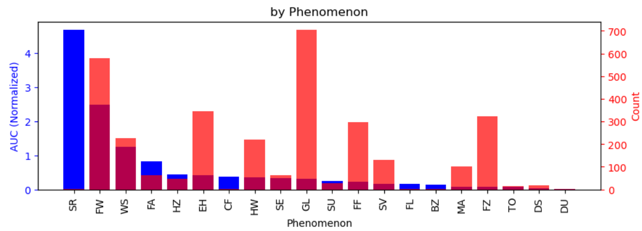


FIGURE 25: Normalized AUC by weather event type for California. Refer to Table 3 for descriptions of weather event types.

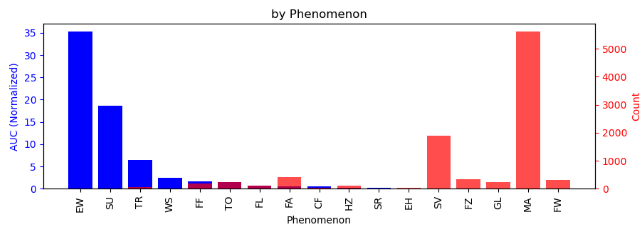


FIGURE 26: Normalized AUC by weather event type for Florida. Refer to Table 3 for descriptions of weather event types.

3) Power outage time after the end of the event

The blue bar in Fig. 27 depicts a comparison of the average TAE (sum of the TAE of each event divided by the total number of events in minutes) by weather event types. The blue bar chart shows that extreme wind is most impactful in terms of average TAE followed by tropical storm (TR), marine (MA), high surf (SU), and so on. This shows that the power outage continues for longer duration after an extreme wind. Extreme wind causes several types of damage to power system infrastructure, leading to longer power outages. The presence of vegetation exacerbates the situation as extreme wind causes vegetation damage, which could damage the power system infrastructure.

Figure 28 and Fig. 29 provide the average TAE by weather event type for California and Florida. These figures show that the most impactful weather event in terms of TAE is blowing dust (DU) in the state of California and extreme wind (EW) in Florida. This suggests that California needs to put more emphasis on developing strategies against DU while Florida needs to develop strategies against EW to improve TAE.

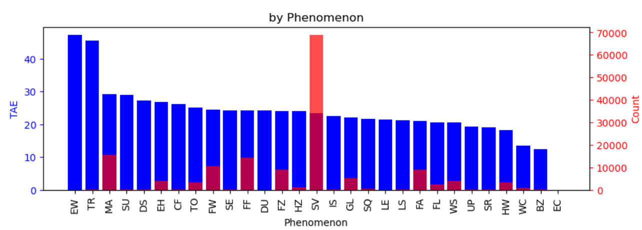


FIGURE 27: Average TAE by weather event type. Refer to Table 3 for descriptions of weather event types.

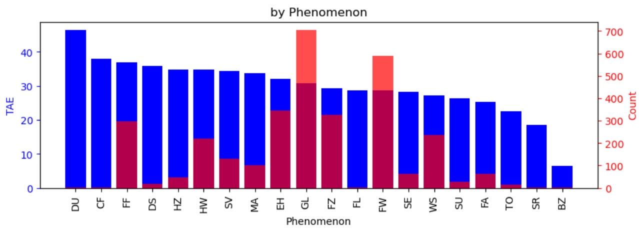


FIGURE 28: Average TAE by weather event type for California. Refer to Table 3 for descriptions of weather event types.

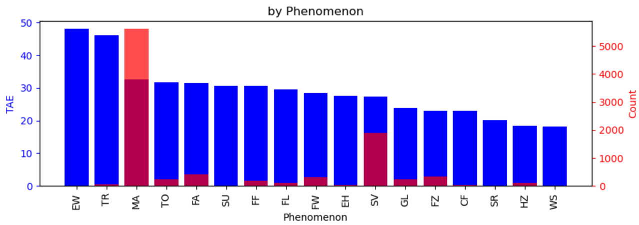


FIGURE 29: Average TAE by weather event type for Florida. Refer to Table 3 for descriptions of weather event types.

V. CONCLUSIONS

This paper presents an automated data framework to analyze the resilience of power systems against extreme weather events. For this work, 2018 to 2022 data was collected from the NWS dataset and the ORNL EAGLE-I power outage dataset to conduct a thorough power resilience analysis from various perspectives. Our methodology involved preprocessing, error handling, grouping near-time weather events, mapping the EAGLE-I power outage dataset on the NWS dataset, configuring the outage event threshold, and quantifying the resilience of the power system against extreme weather events. We defined TOT, AUC, and TAE as the metrics to uncover significant disparities in resilience associated with various types of extreme weather events and across diverse geographic locations.

Through this proposed approach we analyzed the affect of extreme weather events on power systems at the state and county levels. The analysis shows that states are impacted differently in terms of the defined metrics. For example, California is the most impacted in terms of TOT followed by Michigan and Louisiana. On the other hand, in terms of normalized AUC (outage impact per customer in minutes), Virgina is the most impacted state followed by North Carolina, Louisiana, and Texas. However, when we look from the prospective of TAE, Hawaii, Delaware, California, Washington, DC, and Florida are the among the most impacted states. Therefore, there is no single trend that follows all the metrics together. County-level analysis shows that power system resilience varies significantly in terms of presented metrics. County- and city-level granularity is crucial to better understand the unique resilience challenges faced by individual communities.

This work also analyzed the impact of different weather event types on power systems. Weather events such as extreme wind, tropical storms, and high surf need to be concerned more in terms of TOT as their average TOT was very high. Similary, wind storm, tropical storm, and high surf were most impactful in terms of normalized AUC (outage impact per customer in minutes). In terms of TAE, extreme wind is most impactful, followed by tropical storm, marine, and high surf. This shows that extreme wind challenges power systems most in terms of all the quantified metrics. Note that the magnitude of the events are not incorporated properly in this study.

Our automated data workflow has the added advantage of not requiring active utility company participation, which can often be a significant barrier to data collection in this field. This approach could pave the way for more extensive and more frequent analyses of power outages caused by extreme weather events, ultimately aiding in improving power grid resilience and reducing the societal and economic impacts of power outages. Also, the proposed approach is useful to identify vulnerable hot spots (helpful for asset management) to know the impact of specific event types, which is important for developing planning strategies and the predictive analysis framework.

In the future we will develop the framework to incorporate the magnitude of the weather events to better analyze the specific impacts across different states and counties. The magnitude of the weather event types needs to be normalized and incorporated into the quantification metrics to further analyze their impact. Another important prospective of this research is analyzing the impact of climate change on power outages, as is studying the impact on maintenance schedule. These analyses will become more useful and trustworthy in the future as the coverage of EAGLE-I data increases.

ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude to our colleagues Sarah Tennille and Melissa Allen, for their support in identifying relevant datasets that includes weather data and modeled customer counts at county level for computing percentages of outages. Their expertise and insights about these datasets were invaluable in helping us to complete this work.

SANGKEUN MATTHEW LEE obtained his Ph.D. degree in computer science and engineering from Seoul National University, South Korea in 2012. He is currently serving as an R&D Associate in the Computer Science and Mathematics Division at Oak Ridge National Laboratory. His research focuses on big data, data science, and machine learning, and he has applied state-of-the-art data analysis technologies to various application domains. Over the past few years, Dr. Lee has collaborated with scientists from various disciplines, including power systems domain, material science, building science, and mechanical engineering.

SUPRIYA CHINTHAVALI is a group leader for the Built Environment Characterization group within ORNL's Geospatial Sciences and Human Security Division at Oak Ridge National Laboratory with over 16 years of experience in data analytics and visualization specifically for understanding and protecting critical infrastructures. She has been leading multiple DOE programs and projects related to the energy sector's situational awareness (EAGLE-I, NAERM and URBANNET), policy analysis, and smart grid related to the energy sector's situational awareness, policy analysis, and smart grid, and Outage Data Initiative (ODIN) as a co-lead and a principal investigator. She has a master's degree in Computer Science and Engineering and as well as Automotive Embedded Systems. Previously, she worked for Delphi Automotive systems as an advanced software engineer for four years.

NARAYAN BHUSAL received the B.E. degree in electrical engineering from Pulchowk Campus, Tribhuvan University, Nepal, in December 2015 and M.S. degree in electrical engineering from the University of Nevada, Las Vegas in May 2018. In 2021, he obtained a Ph.D. degree in electrical engineering from the University of Nevada, Reno. Currently, he is working at Oakridge National Laboratory (ORNL) as a Research Scientist. Prior to joining ORNL, Narayan worked as a Principal Engineer at Quanta Technology LLC.

His research interest includes resilience, reliability, and stability of cyber-physical energy system; application of machine/deep learning in power system; virtual inertia; and community-based energy storage.

NILS STENVIG (Senior Member, IEEE) received the B.S. degree in electrical engineering in 2008, the M.B.A. degree in 2010, and the M.S. degree in electrical engineering in 2011, each from from Michigan Technological University, Houghton, MI. He is currently a R&D Staff Member and the Group Leader for Power Systems Resilience with Oak Ridge National Laboratory, Oak Ridge, TN, USA.

His primary areas of research are advanced grid modeling, interdependent energy infrastructures, and power systems reliability and resilience.

ANIKA TABASSUM is a Postdoctoral research associate at Oak Ridge National Laboratory, where she is contributing towards developing deep Learning for multi-scale and multimodal data evolved from energy research. Her research interest broadly lies in robust and scalable domain-guided ML for scientific discovery. She has been selected as RISING STAR 2023 by UT Austin and an outstanding postdoctoral award from her division at ORNL in 2022. She received her Ph.D. from the Department of Computer Science at Virginia Tech where she worked on bringing domain-guided ML to address multiple challenges to prepare and mitigate power system failures and disaster vulnerabilities. Her Ph.D. research work was funded by NSF Urban Computing fellowship. She won 1st prize in designing the COVID-19 forecasting model for Facebook-CDC challenge. She has published in multiple venues as NeuRIPS, AAAI, ACM SigKDD, CIKM, IEEE BigData, IAAI, and journal like ACM TIST and Elsevier.

TEJA KURUGANTI (Senior Member, IEEE) received M.S. and Ph.D. degrees in electrical engineering from the University of Tennessee, Knoxville, TN, USA, 2003 and 2012, respectively. He is currently a Distinguished R&D Staff Member and the Section Head for the Advanced Computing Methods for Engineered Systems with Oak Ridge National Laboratory, Oak Ridge, TN, USA.

His research interests include wireless sensor networks, modeling and simulation of communication and control systems, electric grid modeling, and novel techniques for enabling grid-responsive building loads.

REFERENCES

[1] Hassan Haes Alhelou, Mohamad Esmail Hamedani-Golshan, Takawira Cuthbert Njenda, and Pierluigi Siano. A survey on power system blackout and cascading events: Research motivations and challenges. Energies, 12(4):682, 2019.

[2] Thomas Loveland, Rezaul Mahmood, Toral Patel-Weynand, Krista Karstensen, Kari Beckendorf, Norman Bliss, and Andrew Carlton. National climate assessment technical report on the impacts of climate and land use and land cover change. 2012.

[3] Alan H. Sanstad, Qianru Zhu, Benjamin Leibowicz, Peter H. Larsen, and Joseph H. Eto. Case studies of the economic impacts of power interruptions and damage to electricity system infrastructure from extreme events. DOI 10.2172/1725813.

[4] Narayan Bhusal, Michael Abdelmalak, Md Kamruzzaman, and Mohammed Benidris. Power system resilience: Current practices, challenges, and future directions. IEEE Access, 8:18064–18086, 2020. DOI 10.1109/ACCESS.2020.2968586.

[5] Michael Abdelmalak, Jordan Cox, Sean Ericson, Eliza Hotchkiss, and Mohammed Benidris. Quantitative resilience-based assessment framework using eagle-i power outage data. IEEE Access, 11:7682–7697, 2023. DOI 10.1109/ACCESS.2023.3235615.

[6] Elizabeth L Hotchkiss and Alex Dane. Resilience roadmap: a collaborative approach to multi-jurisdictional resilience planning. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2019.

[7] Fauzan Hanif Jufri, Victor Widiputra, and Jaesung Jung. State-of-the-art review on power grid resilience to extreme weather events: Definitions, frameworks, quantitative assessment methodologies, and enhancement strategies. Applied Energy, 239:1049–1065, 2019. DOI 10.1016/j.apenergy.2019.02.017.

[8] U.S. Energy Information Administration (EIA). Annual Electric Power Industry Report, Form EIA-861 detailed data files. [Online]. Release Date: October 6, 2022, 2020 Data Re-released: November 3, 2021, Link to Data.

[9] U.S. Department of Energy. The Electric Emergency Incident and Disturbance Report (Form DOE-417). [Online]. Electric Disturbance Events (DOE-417), Link to Data.

[10] Gregg Edeson and Aleka Stevens. Why resiliency in the electrical grid should be measured from the customer's perspective. Power Magazine, 2021.

[11] Amin Gholami, Tohid Shekari, Mohammad Hassan Amirioun, Farrokh Aminifar, M. Hadi Amini, and Arman Sargolzaei. Toward a consensus on the definition and taxonomy of power system resilience. IEEE Access, 6:32035–32053, 2018. DOI 10.1109/ACCESS.2018.2845378.

[12] Akhtar Hussain, Van-Hai Bui, and Hak-Man Kim. Microgrids as a resilience resource and strategies used by microgrids for enhancing resilience. Applied Energy, 240:56–72, 2019. DOI 10.1016/j.apenergy.2019.02.055.

[13] Eric D. Vugrin, Andrea R Castillo, and Cesar Augusto Silva-Monroy. Resilience metrics for the electric power system: A performance-based approach. 2 2017. DOI 10.2172/1367499.

[14] Saeedeh Abbasi, Masoud Barati, and Gino J. Lim. A parallel sectionalized restoration scheme for resilient smart grid systems. IEEE Transactions on Smart Grid, 10(2):1660–1670, 2019. DOI 10.1109/TSG.2017.2775523.

[15] Saeed Mousavizadeh, Mahmoud-Reza Haghifam, and Mohammad-Hossein Shariatkhah. A linear two-stage method for resiliency analysis in distribution systems considering renewable energy and demand response resources. Applied Energy, 211:443–460, 2018. DOI 10.1016/j.apenergy.2017.11.067.

[16] Han Zhang, Zhaohong Bie, Gengfeng Li, and Yanling Lin. Assessment method and metrics of power system resilience after disasters. The Journal of Engineering, 2019(16):880–883, 2019. DOI 10.1049/joe.2018.8661.

[17] National Infrastructure Advisory Council (US). Critical infrastructure resilience: Final report and recommendations. National Infrastructure Advisory Council, 2009.

[18] Sayanti Mukherjee, Roshanak Nateghi, and Makarand Hastak. Data on major power outage events in the continental u.s. Data in Brief, 19:2079–2083, 2018. DOI https://doi.org/10.1016/j.dib.2018.06.067.

[19] Naveed Taimoor, Ikramullah Khosa, Muhammad Jawad, Jahanzeb Akhtar, Imran Ghous, Muhammad Bilal Qureshi, Ali R. Ansari, and Raheel Nawaz. Power outage estimation: The study of revenue-led top affected states of u.s. IEEE Access, 8:223271–223286, 2020. DOI 10.1109/ACCESS.2020.3043630.

[20] Lynne M Stevens. Htgr resilient control system strategy. Technical report, Idaho National Lab.(INL), Idaho Falls, ID (United States), 2010.