Capturing Infrastructure Interdependencies for Power Outages Prediction During Extreme Events

Sangkeun Matthew Lee, Avishek Bose, Narayan Bhusal, and Supriya Chinthavali
Oak Ridge National Laboratory, 1 Bethel Valley Rd, Oak Ridge, TN 37830, United States
(emails: {lees4, bosea, bhusaln, chinthavalis@ornl.gov}

***Abstract*— As extreme weather events such as hurricanes, severe thunderstorms, and floods grow in frequency and intensity, the disruption of power grid systems poses significant challenges, including widespread electrical outages, economic losses, and threats to public safety. This paper presents a forward-looking approach that leverages geographical graph-based machine learning models to predict county-level maximum power outages during such events. By capturing the intricate interdependencies within power system networks, our approach aims to provide precise and actionable predictions that can optimize emergency response efforts and enhance grid resilience. Through the integration of real-world data, including hurricane advisories and power outage records, we have trained and benchmarked multiple machine learning models, demonstrating the feasibility and potential of this method. While our initial results are promising, this paper also charts a course for advancing these models, addressing the remaining challenges, and ultimately transforming how we anticipate and respond to the impacts of extreme weather on power systems.**

Keywords—power system, power outage, forecast, prediction, machine learning

I. Introduction

In recent years, the frequency and severity of extreme weather events—including hurricanes, severe thunderstorms, and floods—have been on the rise [1]. These events can severely impact power systems, often leading to widespread electrical outages and extensive damage to critical infrastructure, such as substations, transmission and distribution lines, and power generation plants. These disruptions not only hinder essential services like healthcare, transportation, and national security but also cause substantial economic losses and negatively affect community well-being [2][3][4].

The ability to accurately predict power outages resulting from these extreme weather events is crucial. Effective outage prediction can optimize emergency response efforts, enhance grid resilience, and mitigate the adverse effects of these disruptions.

However, predicting power outages in this context presents significant challenges. One of the primary challenges is that there are numerous factors that may influence power outages, and our understanding of which features are strongly correlated with outages remains incomplete. One of very crucial features is the interdependency across power system infrastructure and its related components—where the failure of one component can trigger failures in others—further complicate the prediction process. Accurately modeling these interdependencies requires detailed network data, which is often difficult to obtain and integrate.

Existing approaches have evolved from traditional statistical models like generalized linear models (GLMs) [15][16] and random forests [17][18] to more advanced techniques such as graph neural networks [19] to address the challenges of predicting power outages during extreme weather events. These models have progressively improved in their ability to capture the non-linear relationships between various factors, including weather conditions and power system characteristics. However, two key limitations persist: (1) they have largely overlooked the critical infrastructure interdependency of the power system, particularly how outages can propagate or cascade through interconnected infrastructure. Accurately modeling these cascading failures is crucial for predicting widespread power failures, but modeling interdependency is challenging due to complexity, lack of data, proprietary restrictions and security concerns; and (2) most models primarily rely on data from small geographic areas and are often dependent on proprietary data that are required to be provided by utilities, limiting their generalizability and applicability to other regions

In this paper, we envision that leveraging and combining high-quality, diverse real-world data—including historical outages (e.g., EAGLE-I historic power outage data [5]), detailed weather records (e.g., Hurricane Mapping [13] or National Weather Service archives [14]) , and power grid geographical information (e.g., Homeland Infrastructure Foundation-Level Data (HIFLD) [12])—presents an unprecedented opportunity to develop accurate predictive models, capturing critical infrastructure interdependencies.

More specifically, we heuristically infer interdependencies across power systems and other components based on their geographic proximity and metadata, then model these relationships as a large-scale geographical graph, where vertices represent infrastructure components and edges represent their interdependence. A graph operation, k-hop neighborhood searches from vertices in disrupted areas, are then used to generate data that augments the training data composed of power outage and weather-related information. This augmented data is subsequently used to train the model. Similar graph operations are performed to augment the input data and used for prediction.

To demonstrate the feasibility of our approach, we trained models using hurricane advisory data obtained from Hurricane Mapping [13], covering 11 hurricanes from 2016 to 2023. We evaluated the performance of various machine learning models, including Random Forest [6], SVR [7], XGBoost [8], Gradient Boosting [9], k-Nearest Neighbor [28] and neural networks [29], using historical power outage data [5] and advisory data [13] associated with these hurricanes. We applied a leave-one-out cross-validation approach, where 10 hurricanes were used for training and 1 for testing, and observed promising accuracy in

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

predicting county impact level(i.e., maximum power outage) for the next 72 hours in Florida from the advisory timestamp.

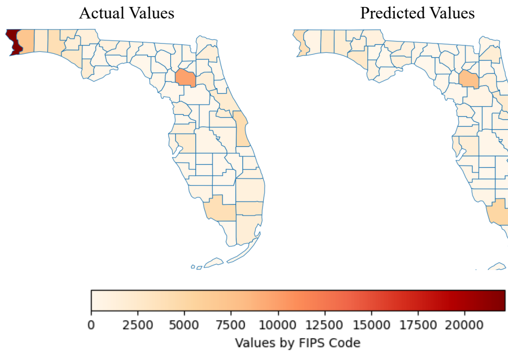


Fig. 1. Actual vs. Predicted County Impact Level within 72 Hours from the Hurricane Advisory Timestamp Using XGBoost. Hurricane Zeta (2020), Advisory #10.

Fig. 1 illustrates an example output generated by the XGBoost model, comparing actual and predicted maximum power outages (number of customers impacted) within 72 hours of Hurricane Zeta's (2020) advisory #10. While our current model is relatively simple and intended as a foundational step rather than a complete solution—given that we did not incorporate various meteorological variables and focus mostly on interdependency—the results are promising. They demonstrate the potential of this approach to significantly enhance existing power outage prediction models. The accuracy of the tested models varied depending on the hurricane and the model used; however, the XGBoost and Neural Network models exhibited better accuracy than the other techniques, with Mean Absolute Errors (MAE) of 531.316 and 665.143, respectively, for Hurricane Zeta (Category 3 hurricane). A detailed evaluation of the results will be discussed in Section IV.

Building on these initial successes, we also explore the remaining challenges and outline future research directions, emphasizing the importance of further advancing these models to enhance emergency preparedness and improve power system resilience in response to increasingly severe weather events.

The remainder of this paper is organized as follows. Section II reviews existing studies on power outage prediction models and available datasets. Section III presents our initial models, explaining how we infer critical infrastructure interdependencies from available critical infrastructure data and incorporate them into power outage prediction model. In Section IV, we provide an evaluation of their performance. Finally, Section V offers concluding remarks and discusses the remaining challenges in developing accurate prediction models for power outages and outlines our vision for addressing these challenges.

II. Related Work

A. Existing Models

Numerous research studies have been conducted to develop models that predict power outages caused by extreme weather events, such as hurricanes, by leveraging various features related to weather conditions, environmental factors, and power system characteristics. Early efforts primarily employed generalized linear models (GLMs) [15][16], which utilized hurricane parameters like maximum wind speeds, environmental indicators such as tree types and soil moisture, and system-specific data, including the number of transformers and customers.

More recent studies have explored the use of ensemble learning methods, such as random forests, to enhance the predictive accuracy of outage models [17][18][22]. Guikema et al. [17] acknowledged the importance of using publicly available data set to make these models widely applicable, but authors admit that the models only using public dataset lead to lacking important features such as soil moisture data. Nateghi et. [18] el. focused on accurately power outage duration rather than occurrences. Lee et al. [22] benchmarked various machine-learning techniques for a power outage model during hurricanes using publicly available data. However, the features used in the model were limited, and the predictions were focused solely on the state level.

As research has progressed, there has been a gradual shift towards more advanced techniques, such as neural networks. Additionally, more recent approaches have started to incorporate graph-based methods, such as graph neural networks (GNNs), which are better suited to capturing the spatial and topological relationships within the power grid. For example, Owerko et al. [19] applied GNNs to predict power outages in New York City based on weather measurements, achieving higher accuracy by modeling the connections between weather stations as a graph.

While the evolution of power outage prediction models has seen significant advancements through the incorporation of more sophisticated techniques, existing models have not yet accounted for critical infrastructure interdependencies—an essential factor in estimating damage caused by hurricanes. For instance, damage or failure can cascade across transmission substations and transmission lines, leading to power outages at the transmission level during extreme events [20]. Incorporating this information into predictive models could be highly beneficial. However, no existing approaches have done so due to the complexity of modeling critical infrastructure interdependencies and the lack of publicly available datasets that accurately represent these relationships.

B. Available Data sets

Historical Power Outage Data: The historical power outage data used in most existing work is typically sourced from regional utility companies [15][16][18][19] or manually extracted from the DOE Situation Reports [17]. In contrast, we leverage the U.S. historical power outage data collected by Oak Ridge National Laboratory (ORNL) through the EAGLE-I™ (Environment for Analysis of Geo-Located Energy Information) system [5]. EAGLE-I, maintained by ORNL for the Department of Energy (DOE), is an operational and scalable data and information platform that provides real-time wide-area situational awareness of the energy sector. It serves as a centralized platform for monitoring power distribution outages, covering over 146 million customers, which accounts for over 92% of the U.S. and its territories. This platform has been collecting electric outage data since 2014, with records including the total number of customers without power at the county level, alongside details about the relevant utility company. The outage data represents the number of affected customers, with snapshots recorded every 15 minutes.

The EAGLE-I power outage data, spanning eight years from 2014 to 2022, is now publicly available for data discovery purposes [21]. This availability is particularly significant as it enables the development of hurricane-related power outage prediction models on a national scale. Researchers can now train models using this extensive dataset without relying on proprietary data provided by utility companies or the manual extraction of information from past DOE situation reports. This opens up new opportunities for more accurate and scalable predictions of power outages during extreme weather events. In our previous work [22], we demonstrated how EAGLE-I dataset could be used to predict power outage during extreme weather events. In this paper, we further expand our approach to incorporate interdependency information into the model.

Extreme weather event data: To develop effective power outage prediction models, incorporating accurate and comprehensive extreme weather event data is crucial. This data should include detailed historical information on the type, geographic location, and timing of each extreme weather event. In this paper, we used the dataset from Hurricane Mapping [13], which offers an organized archive of hurricane advisory data. Although Hurricane Mapping data requires a subscription fee and is not freely available, it is derived from publicly accessible sources like those provided by the National Hurricane Center (NHC) [11]. Since our primary focus in this paper is on demonstrating how critical infrastructure interdependencies can be incorporated into the model. As a result, we concentrated on hurricanes and a limited set of weather variables (e.g., Wind Probabilities). However, our approach can be naturally extended to include other types of extreme weather events. For weather events beyond hurricanes, the National Weather Service (NWS) [14] provides publicly available weather data that should be considered. This dataset contains information about the geography, type (e.g., hurricane, flood, thunderstorm), and advisory level (Watch, Warning, Advisory, etc.) of weather events that have occurred in the U.S. The dataset and its metadata are accessible in various formats.

Critical infrastructure data: As previously discussed, existing models often fail to account for interdependencies across heterogeneous layers of critical infrastructure, such as transmission lines, power plants, and substations. Incorporating these interdependencies into predictive models is a complex challenge, further complicated by the lack of publicly available datasets that accurately represent these relationships. Although straightforward datasets representing critical infrastructure interdependencies are not directly available, geographic data for various U.S. critical infrastructure is accessible through the Homeland Infrastructure Foundation-Level Data (HIFLD) program, managed by the U.S. Department of Homeland Security (DHS) [23].

HIFLD offers a comprehensive array of geospatial data on critical infrastructure sectors, including energy, transportation, communications, healthcare, and emergency services. The HIFLD program is divided into two categories: HIFLD Open, which is publicly accessible, and HIFLD Secure, which includes licensed and non-licensed datasets intended for official use only.

HIFLD Open data, which provides over 400 datasets containing detailed geographical information, is already rich in content. However, it is also highly heterogeneous—lacking standardization and functioning independently of one another. Crucially, these datasets do not inherently include interdependency information. As a result, although these datasets have significant potential, effectively utilizing them requires innovative approaches to integrate them into predictive models, particularly to address the critical interdependencies necessary for accurate damage estimation.

In Section III, we will explain how the HIFLD dataset can be used to infer critical infrastructure interdependencies and can be utilized for machine learning-based power outage prediction models in combination with EAGLE-I and hurricane advisory archives.

III. Data Processing and Model Training

A. *Inference of Interdependency from Geographic Datasets*

Critical infrastructure interdependencies refer to the relationships between various systems, where the functioning of one system depends on others. This interconnection means that the failure of one component can cascade through the critical infrastructure network, potentially leading to widespread disruptions. Capturing these interdependencies is crucial for creating accurate models. Real-world interdependencies can be highly complex, involving backup power systems, rerouting mechanisms, and specific combinations of component statuses within the network.

In our previous work [24][25], we demonstrated how a critical infrastructure interdependency network can be created from multiple geographical datasets. This approach models the infrastructure as a graph network, where nodes represent components like substations, transmission lines, power plants, or gas pipelines, and directed edges from node A to node B indicate that the failure of A can negatively impact or disrupt B. While this approach may not perfectly capture all real-world interdependency relationships, it could provide a reasonable estimate using the available datasets. In our previous work, we did not explore how this information could be utilized in power outage prediction models—a gap that this paper aims to address. In this study specifically, we used 31 HIFLD critical infrastructure datasets. These include 8 energy layers (Substations, Power Plants, Transmission Lines, Gas Pipelines, Delivery Points, Compressor Stations, Ethanol Plants, Dam Lines) and 23 non-energy layers.

The non-energy layers include:

- **Safety and Security:** Red Cross Chapter Facilities, Courthouses, EPA ER RMP Facilities, Fire Stations, Local Emergency Operations Centers, Local Law Enforcement
- **Transportation:** Aviation Facilities, Hurricane Evacuation Routes, Primary Roads, Railroads
- **Food-Water-Shelter:** National Shelter System Facilities, Public Refrigerated Warehouses

- **Health and Medical:** Hospitals, Nursing Homes, Urgent Care Facilities
- **Hazardous Material:** Wastewater Treatment Plants
- **Communication:** Cell Towers, FM Transmission Towers
- **Other:** Child Care Centers, Colleges and University Campuses, Banks, Credit Unions, Public Schools

The 31 critical infrastructure datasets, originally in Shapefile format, were downloaded from the HIFLD data portal and processed to create a large-scale geographical graph. While the original process is introduced in our previous work [24][25], we made significant updates on the workflow and the data we used. Here is a summary of the updated workflow.

First, each critical infrastructure component in the dataset is converted into a node (i.e., vertex). Each node represents a physical critical infrastructure entity, such as a substation, power plant, transmission line, public school, bank, and others. It's important to note that even transmission lines, which are typically represented as polylines (series of points), are converted into vertices rather than edges. As a result, we have 31 types of nodes, where each node type corresponds to a specific category of critical infrastructure component, as defined by the original dataset. Each node is assigned its own geographical information (e.g., coordinates in the case of a point, or a series of coordinates in the case of a line or polygon) along with property-value pairs that describe the node (e.g., a substation might have a name, while a transmission line might have a maximum voltage). These property values are inherited from the original dataset. Each layer retains its own identifier, which uniquely identifies a component within that layer, also inherited from the original data. Additionally, we create a global node ID that is unique across the entire network.

Second, we infer and create interdependency edges between different node types by applying rules defined using a predefined set of operations. Below are examples of these operations and how they were used to infer critical infrastructure interdependencies across components in different layers:

- **Within-Distance**: This operation creates interdependency edges between nodes of type A and nodes of type B if B is located within a specified distance (e.g., 10 km) from A. If multiple B-type nodes are within this range, 1:N edges are created. For example, we used this operation to create edges from Cell Towers to Hospitals, as the disruption of mobile communication could potentially disrupt emergency services, coordination with first responders, and critical communications with staff.
- **Within-Area**: This operation creates edges from A to B if B is located within specific boundaries corresponding to A. This is applicable when additional boundary information is available. If multiple B-type nodes are within this boundary associated with A, 1:N edges are created. For instance, let us say that service area data is available for distribution substations, all components within a substation's service area can be connected to the substation. In our case, we used this operation to create edges from distribution substations to non-energy components being served by them. To this end, for each distribution substation, we estimated substation service area using a cost-distance algorithm [26], and used the areas to create between substation to other non-energy components.
- **Nearest**: This operation creates edges from A-type nodes to B-type nodes if B is the nearest neighbor to A. For example, edges can be created from Substation A to Transmission Line B if B is the nearest transmission line to A, implying that a disruption of the substation could negatively impact the transmission line. However, in some cases, both directions or the reverse direction of the edges may make more sense, so the rules need to be applied carefully. For instance, while connecting a substation to a transmission line, it is also important to consider that a transmission line failure may impact a distribution substation, necessitating bidirectional edges under the same conditions.
- **Intersection**: This operation creates edges from A nodes to B nodes if A and B intersect. For example, transmission lines that are connected to each other may intersect, and if so, we create edges to reflect that a disruption in one transmission line can cascade to others.

As a result, we created a total of 146 different dependencies across 31 node types. The constructed Critical Infrastructure Interdependency (CII) graph comprises 1,068,727 nodes and 4,846,678 edges. It is important to note that our goal is not to perfectly capture all existing critical infrastructure interdependencies; rather, we focus on representing the heterogeneous dependencies using simple but reasonable assumptions and heuristics. Next, we will introduce how this graph can be used to estimate cascading impacts and how we leverage it for power outage prediction during extreme events.

B. *Estimating Cascading Impact Using k-hop Neighborhood Search on the CII Graph*

Each node in the CII graph is assigned geographical coordinates, allowing us to locate and identify critical infrastructure nodes within the impact zone of an extreme weather event. If the impact area is known, we can estimate the cascading effects by performing a k-hop neighborhood search (*k-hop search* for short) on the graph. A *k-hop search* with k=1 identifies direct impacts (i.e., direct A⟶B interactions), while a search with k>1 includes nodes that are indirectly affected by the initial disruption. These identified nodes are considered to be at risk due to the disruption of the original nodes within the impact zone.

For example, when a hurricane advisory is issued, from NHC or Hurricane Mapping, we can obtain data that specifies areas where wind speeds exceeding 50 mph are expected, along with associated probabilities over the next 72 hours. Both NHC and Hurricane Mapping offer data with various time windows such as 24, 72, 120 hours, but in this paper, we focus on the 72 hour time frame. Figure 2 illustrates the transmission lines within the area predicted to experience 50 mph winds with a 70% probability over 72 hours during Hurricane Laura (Advisory #22). We then perform a 3-hop neighborhood search on the graph, identifying all nodes reachable within three hops from these transmission lines. This allows us to determine the extent of cascading impacts and identify a range of infrastructure components that may be affected by the disruption.

TABLE I. PEARSON'S CORRELATION COEFFICIENT BETWEEN THE NUMBER OF COMPONENTS IDENTIFIED BY K-HOP SEARCH AND MAXIMUM POWER OUTAGES IN EACH COUNTY BASED ON HURRICANE ADVISORY INFORMATION.

Name	**Year**	**Advisory**	**Wind Speed (Knot)**	**Probability (%)**	**Pearson's Correlation Coefficient at k**				
					k=1	**k=3**	**k=5**	**k=7**	**k=9**
Irma	2017	40	50	90	0.094	0.345	0.826	0.820	0.877
Irma	2017	40	50	70	0.647	0.799	0.847	0.855	0.858
Irma	2017	40	34	90	0.698	0.859	0.855	0.860	0.857
Irma	2017	40	34	70	0.631	0.864	0.860	0.858	0.857
Laura	2020	22A	50	70	0.676	0.674	0.748	0.840	0.816
Laura	2020	22A	34	90	0.683	0.650	0.737	0.839	0.817
Laura	2020	22A	34	70	0.839	0.793	0.841	0.789	0.718
Ian	2022	21	50	70	0.612	0.707	0.605	0.473	0.450
Ian	2022	21	34	90	0.652	0.724	0.624	0.465	0.435
Ian	2022	21	34	70	0.561	0.455	0.449	0.436	0.427
Ida	2021	07	50	70	0.490	0.508	0.744	0.717	0.650
Ida	2021	07	34	90	0.726	0.760	0.750	0.698	0.619
Ida	2021	07	34	70	0.683	0.652	0.649	0.609	0.584
Idalia	2023	09	34	90	0.106	0.119	0.176	0.311	0.277
Idalia	2023	09	34	70	0.329	0.349	0.302	0.258	0.197

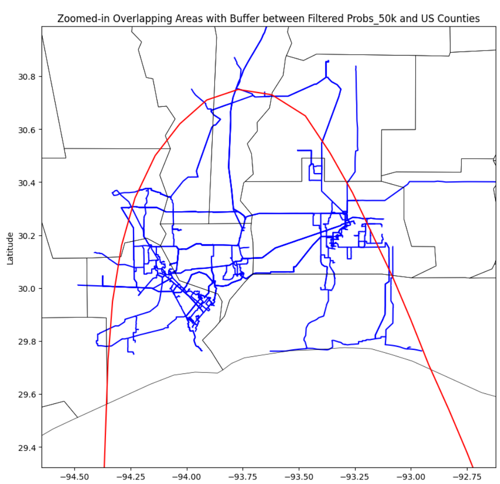


Fig. 2. Transmission lines identified within the area expected to experience wind speeds of 50 mph with a 70% probability over the next 72 hours, based on Hurricane Laura (Advisory #22).

Fig.3. shows the result where each blue dot represents the vertices that are identified by the 3-hop search. The blue dots in Fig. 3 represent components in the Critical Infrastructure Interdependency (CII) graph that could potentially be impacted by the hurricane within the next 72 hours, either directly or indirectly. We observe that these components, at risk of disruption, are located both within and outside the defined impact zone. This result is valuable for identifying vulnerable components, even those located outside the impact zone; however, it must be interpreted cautiously. First, the CII graph is constructed based on simplified assumptions and heuristic rules, which may not fully account for the complexity of interdependencies across different critical infrastructure layers. Second, not all components within the impact zone are equally affected, and various types of infrastructure (e.g., transmission lines, substations, power plants) may experience simultaneous disruptions. Finally, we used an arbitrary k-value of 3 in this example, but the optimal k-value for accurately modeling hurricane impacts remains uncertain.

To better understand and utilize the *k-hop search* results on the Critical Infrastructure Interdependency (CII) graph, we calculate the Pearson's correlation coefficient [27] between these search results and power outages during the same period (within 72 hours of the advisory being issued). Specifically, we count the number of impacted components (blue dots in Fig.3) within each county and correlate these figures with the county's maximum number of customer outages recorded within 72 hours during the hurricane's impact, using data from EAGLE-I. A higher correlation value indicates a stronger relationship between the k-hop search results and the maximum observed outage levels during the 72-hour period.

Fig. 4 resents an example pair of choropleth maps representing the data used to compute the correlation. It is important to note that the ranges of the values are different for each dataset. As, (a) shows the map based on the number of customers, but (b) shows the map based on the number of critical infrastructure components. Data in Fig.4 (b) cannot be directly used as a prediction of data shown Fig.4 (a).

Table I presents the correlations calculated for advisories from four hurricanes: Irma, Laura, Ian, and Idalia. It is important to note that there are multiple options for defining the initial impact zone, and data availability varies for each case. For example, during Hurricane Irma, wind speed data was available for 50-knot and 34-knot winds with 90% and 70% probability, respectively, and these zones overlapped with transmission lines, allowing us to calculate the correlation. However, for other cases, such as Hurricane Idalia, wind speed data for 50 knots was not available. One pattern is that correlations are generally higher when the k is greater than 1. This indicates that the k-hop search is effective in capturing not only the immediate impact within the storm's impact zone but also the indirect effects on nodes located beyond it. The cascading nature of infrastructure disruptions, particularly during hurricanes, becomes evident as correlations peak at larger k values, suggesting that indirect disruptions are significant contributors to overall outages.

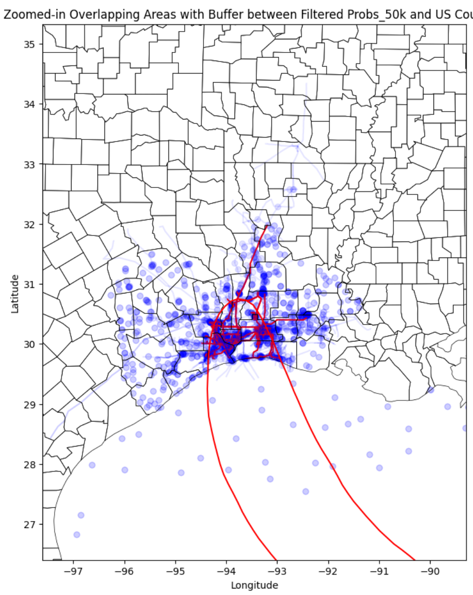
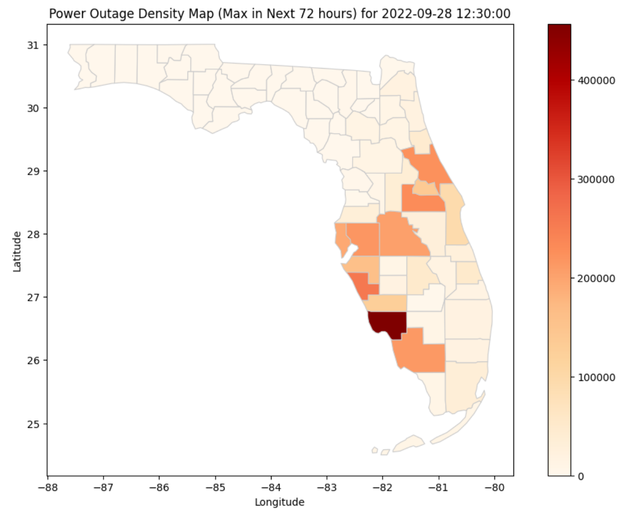


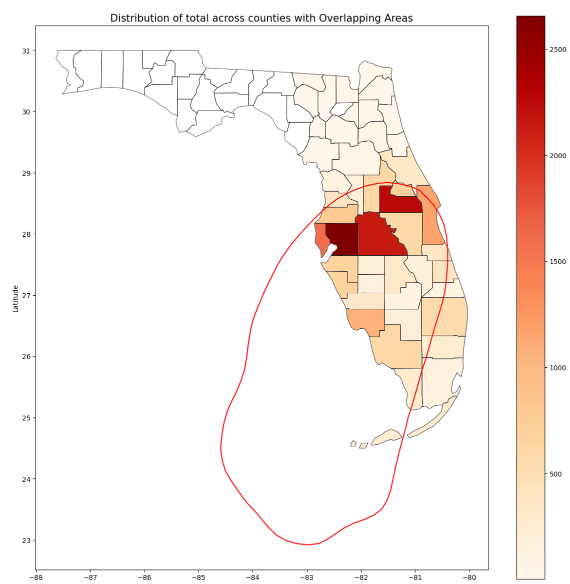
Fig. 3. Result of a 3-hop neighborhood search identifying vertices (represented by blue dots) that could be impacted by the disruption of transmission lines under the area (in Fig.2)

However, it's important to note that there is no one-size-fits-all approach when it comes to determining the optimal k value. For example, when using 50-knot area with 70% probability as impact area, Hurricane Irma's correlation increases steadily with k, peaking at k=9. In contrast, during Hurricane Laura, the correlation reaches its highest point at k=5, after which it begins to decline. This variability underscores that the cascading impact on infrastructure depends on the characteristics of each storm, such as its wind speed and the probability of high winds affecting transmission lines and other components. Different storms require different k values to best capture the relationship between the k-hop search and the maximum outages recorded, meaning that the optimal k must be tuned on a case-by-case basis.

Another interesting observation is that advisories with higher wind speeds or higher probabilities of impact tend to result in stronger correlations, as seen with Hurricane Irma's advisories at 90% probability. For instance, when considering the 50-knot wind speed advisory at 90% probability, the correlation values steadily rise and remain high, peaking at k=9. This trend suggests that the certainty of high-impact winds with less uncertainty (i.e., high wind probability) leads to a more accurate emulation of the cascading impact. Conversely, advisories with lower wind speeds or lower probabilities, such as in the case of Hurricane Idalia, show weaker correlations. This is reasonable, as low wind speeds combined with high uncertainty are more likely to result in lower disruptions within the impact area.



(a) Maximum number of customers without power per county within 72 hours of advisory issuance (Advisory #21)



(b) Total number of components per county from the k-hop search (k=3, 34-knot, 90% probability)

Fig. 4. Comparison of choropleth maps generated using k-hop search result and power outage data for Hurricane Ian Advisory #21

In summary, while k-search using the CII graph is not a direct method for predicting power outages, we observe a strong Pearson correlation coefficient (>0.8~0.9) with the maximum number of customers experiencing outages in affected areas when both the impact zone (including factors like wind speed and probability) and k are optimized. However, k-search result are not directly comparable to power outage values, thus, machine learning models capable of learning from k-search results and predicting power outages is essential. Next, we explain how k-hop search results are used as features for training machine learning models that predicts power outage numbers.

C. Machine Learning Models using k-hop search based features

Data Preparation: Our main objective is to train machine learning models to predict the maximum number of customers without power per county within 72 hours of an advisory issuance, as illustrated in Fig. 4(a). Accurately predicting these values enables county-level risk assessment both before and during hurricane landfall, benefiting various stakeholders, including utility companies, emergency management agencies, homeowners, and residents. We reorganize available data, such as hurricane advisory data, k-hop search results, and power outage data, to create the training dataset. In this paper, we specifically focus on Florida due to the frequency of historical hurricane events in this region; however, the same approach can be applied and generalized to other areas. The hurricanes used in our analysis include Matthew (2016), Irma (2017), Harvey (2017), Dorian (2019), Sally (2020), Eta (2020), Zeta (2020), Elsa (2021), Nicole (2022), Ian (2022), and Idalia (2023). We used hurricane advisories at intervals of 5, specifically advisories 5, 10, 15, 20, and so on, to capture key stages of the hurricane's development and impact.

The process for constructing our training dataset is as follows: For each hurricane advisory, we build the dataset with k values ranging from 1 to 10:

- **K-hop search configuration features**: These input features include wind speed, wind probability from the advisory, and the k value used for the k-hop search (3 features in total).
- **Direct impact count features:** These features represent the number of infrastructure components per county located within the direct impact zone (67 features, one for each county). These features remain constant across all k values.
- **Indirect impact count features**: These features represent the number of infrastructure components per county located within the k-hop search result (67 features, one for each county).
- **Target variables**: The maximum number of customers without power per county within 72 hours of the advisory issuance. (67 variables, one for each county)

With this configuration, we prepared a training dataset. The total number of input features and target variables are 137 and 67 respectively. We did not incorporate time-series power outage data into our training dataset for simplicity, rather, we solely focus on learning the relationship between k-hop search result and actual power outage values to evaluate if employing k-hop search, in other words, estimated critical infrastructure interdependency information, can play significant role in producing accurate result.

Training Machine Learning Models: With the prepared training dataset, we trained five machine learning models capable of predicting multiple target variables: Random Forest (RF) [6], Support Vector Regressor (SVR) [7], XGBoost [8], Gradient Boosting [9], k-Nearest Neighbor (kNN) [28], and Neural Networks [29]. For implementation, we used Python's scikit-learn library [30], applying the library's default hyperparameters.

Making Predictions with Trained Models: Before landfall or during a hurricane's impact, hurricane advisories are issued. Hurricane advisory data provide information such as the shape of the impact zone, wind speed, and wind probability, which were used for training. Using this information, we prepare k-hop search configuration features. Additionally, a value for k (ranging from 1 to 10) must be selected. We then perform a k-hop search on the CII graph to generate direct and indirect impact count features. These features are fed into a trained model (e.g., an XGBoost model) to predict the maximum number of customers without power per county within 72 hours of the advisory issuance (67 target variables, one for each county).

IV. Evaluation

Hurricane advisories are time-dependent, as wind speed, probability, and impacted areas are continuously updated. Therefore, including future information from the same hurricane in both model training and testing must be avoided to prevent misleading performance. To address this, we assume that each hurricane is an independent event, though not within the same hurricane. We performed leave-one-out cross-validation, where we tested on one hurricane while training the model on all others, excluding the one being tested. This approach ensures that the model does not overfit to the specific characteristics of a single hurricane and generalizes well across independent events. By isolating each hurricane in this way, we avoid the risk of using future information from the same hurricane during training, creating a more robust model for predicting power outages during new, unseen hurricanes. At the time of prediction in practice, the available impact zones for the k-hop search to create input features may vary (e.g., some hurricanes may have a 5-knot 90% option available, while others may only have a 30-knot 50% option). For our evaluation, we used the measured average MAE, generating input features from the k-hop search results based on all available options for each hurricane, with k varying from 1 to 10.

Fig. 5 provides a detailed comparison of the Mean Absolute Error (MAE), broken down by both hurricane and machine learning model. This allows for a clear understanding of how each model performs for individual hurricanes. While Random Forest performs well across many hurricanes, its MAE spikes significantly for Hurricane Irma, suggesting that it may struggle with predicting outages for more severe storms. In contrast, XGBoost and Neural Networks show relatively lower MAEs across various hurricanes, with Neural Networks often delivering the lowest errors, as shown in Fig. 6 confirming its robustness in diverse storm conditions.

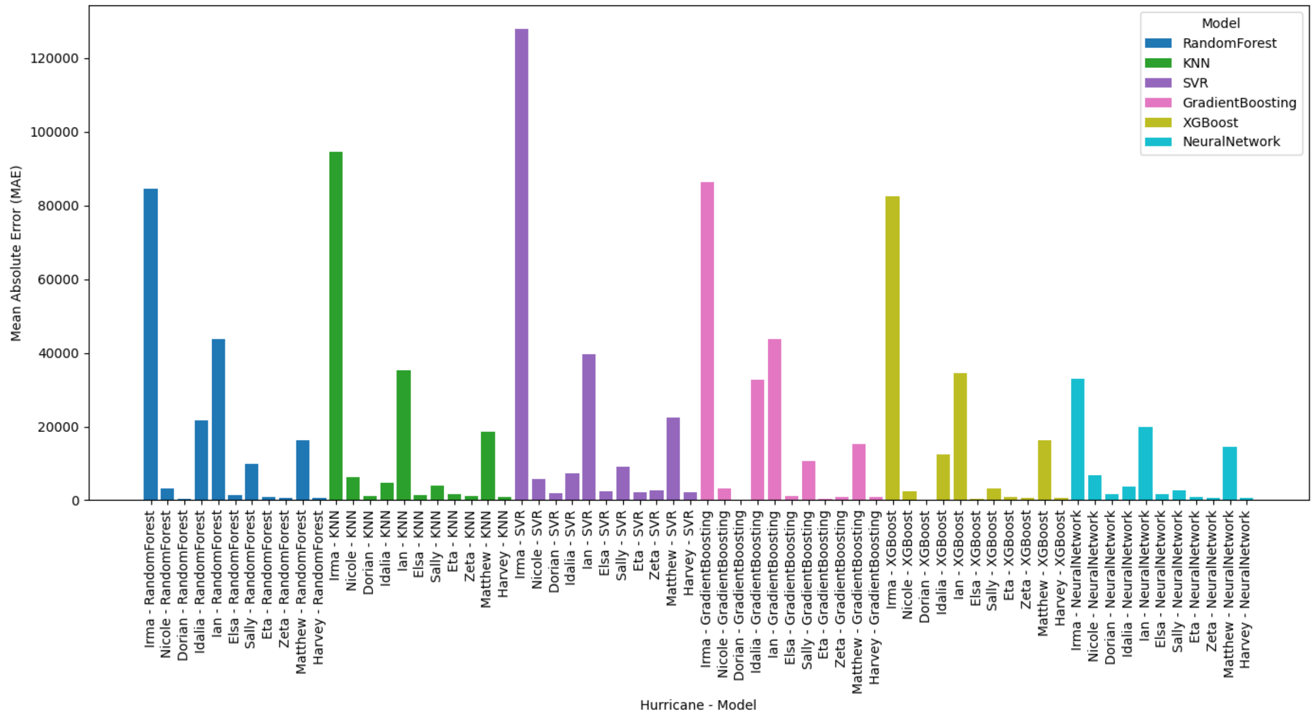


Fig. 5. MAE Comparison for Each Model and Hurricane

However, no single model consistently outperforms all others in every case, highlighting the complexity of these predictions. Since all models used scikit-learn's default hyperparameters, further tuning could enhance accuracy. Given Neural Networks' strong performance and potential for improvement with additional data and tuning, these results are promising for future development.

It is also noteworthy that for hurricanes like Harvey, Nicole, and Zeta, predictions show relatively low MAE, indicating higher accuracy. However, for hurricanes such as Irma and Ian, the MAE is significantly higher, especially for Irma. This suggests that predicting outages for these storms was more challenging. This could be due to the complexity of their impact patterns, wind speeds, or other factors not accounted for in this study.

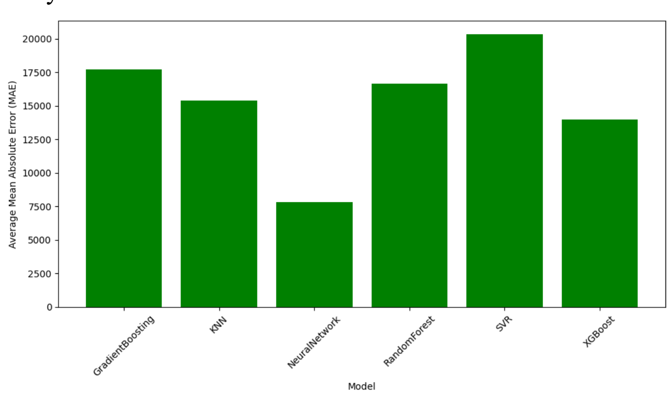
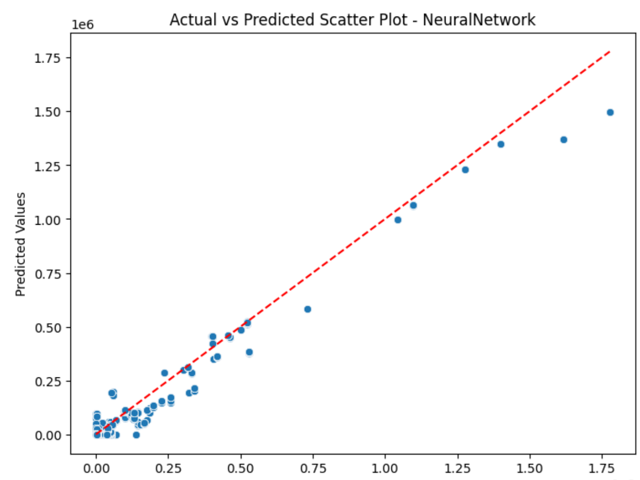
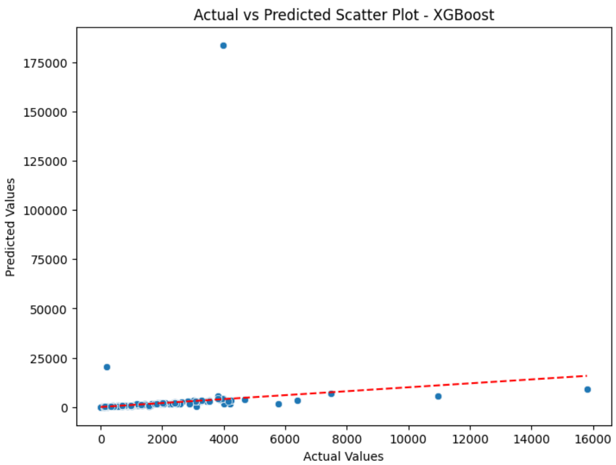


Fig. 6. Average MAE per Hurricane Across All Models

It's important to interpret the Mean Absolute Error (MAE) with caution. Fig. 7(a) displays the scatter plot for Hurricane Irma using the Neural Network model, where the MAE was 32,863.753. Despite the relatively high MAE, the model captured the overall trend quite well, demonstrating its ability to predict the general impact of the hurricane. Fig. 7(a) confirms that our approach was able to produce meaningful results, even when the prediction error was higher. Another example is shown in Fig. 7(b), which displays the XGBoost results for Hurricane Zeta, with a lower MAE of 611.011. While the model performed exceptionally well overall, a few outliers significantly inflated the MAE. These examples highlight the limitations of relying solely on MAE as a performance metric for power outage prediction, emphasizing the need for more specialized metrics that can better account for the nuances in these forecasts. Developing such metrics will be an important direction for future research.

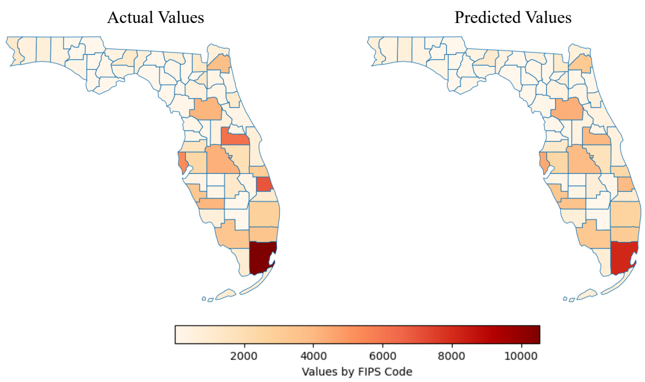


(a) Hurricane Irma – Neural Network (MAE: 32863. 753)

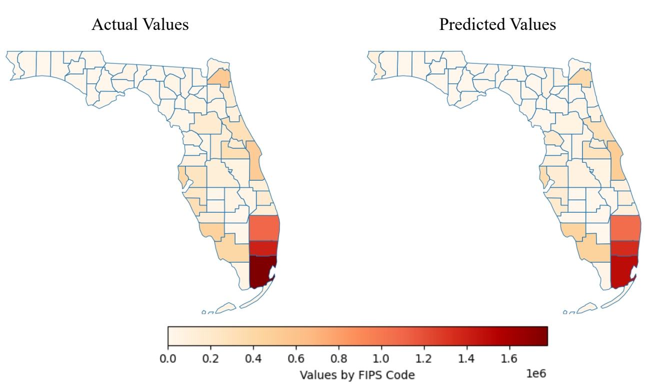


(b) Hurricane Zeta – XGBoost (MAE: 611.011)

Fig. 7. Actual vs. Predicted Scatter Plot



(a) Neural Network Prediction for hurricane Dorian (Advisory #9) (Using the result of k-hop search with k=3, wind speed 34-knot, wind probability 70%)



(b) Neural Network Prediction for hurricane Irma (Advisory #40) (Using the result of k-hop search with k=3, wind speed 50-knot, wind probability 90%)

Fig. 8. Example Predictions

Fig. 8 shows a choropleth map visualization of two predictions for Hurricanes Dorian and Irma. We observe that our predictions accurately reflect the geographical pattern of how power outages spread within the state of Florida. Unlike the results in Fig. 4, the values generated by the trained models are directly comparable to the actual values. As a result, these two maps share the same color scale at the bottom, which explains the ranges of outage values.

V. Conclusions

In this paper, we presented machine learning models designed to predict county-level maximum power outages during hurricane events. Specifically, we demonstrated how to construct a large-scale geographical Critical Infrastructure Interdependency (CII) graph that captures complex relationships across various critical infrastructure systems. We showed how performing a k-hop neighborhood search on this graph can be both insightful on its own and useful as features for machine learning models. Despite not incorporating a wide range of weather and geospatial features, our models were able to accurately predict county-level maximum power outages, effectively capturing how outages cascade across different counties. Given that advisories are issued in advance, this capability is particularly valuable for subject matter experts such as emergency responders, utility operators, and policymakers as they plan for potential impacts. This study points to several directions for future research and improvements:

Incorporating additional features: As mentioned, we did not include many known geospatial and meteorological variables that could affect power outages. In future work, we can expand our models to incorporate a wider range of features, such as soil moisture, temperature, and other weather conditions. These additional inputs could help further improve model accuracy. Other related output variables should also be considered, such as outage values after given time, restoration time, minimum outage values, and average outage values, rather than focusing solely on the maximum.

Time-series prediction: Power outages evolve over time. In this initial step, we focused on predicting maximum outages within a 72-hour window. However, future models can incorporate time-varying inputs and outputs to provide more granular, real-time predictions. This would allow us to track how outages develop over time, which would be incredibly useful for improving emergency response strategies.

Hyperparameter tuning and model optimization: In this study, we did not extensively tune hyperparameters. For example, our neural network was implemented using the default settings in scikit-learn. There is significant potential to improve performance by optimizing the architecture for this specific problem. Given the nature of the data, exploring Graph Neural Networks (GNNs) could be especially beneficial, as they are well-suited to problems involving complex network relationships.

Expanding to other extreme weather events: While our study focused on hurricanes, the approach can be applied to other extreme weather events that cause cascading power outages, such as wildfires, floods, and ice storms. Expanding the scope of our research to these events would allow us to better understand the broader applicability of our model across different types of natural disasters.

More experiments and feature analysis: Further experiments are needed to explore how different features and configurations (such as wind probability, wind speed, k-hop, or weather event type) affect power outage predictions and model accuracy. This would deepen our understanding of the most critical factors influencing outages and how our models respond to different conditions.

Developing improved metrics: As discussed, relying solely on MAE may not provide the most comprehensive assessment of model performance. We need metrics that better capture the geographical patterns of power outages, account for outliers, and consider factors such as population density and the severity of the weather event. Developing these metrics will be crucial for accurately evaluating and improving the model's performance.

These future steps will help create more robust and reliable power outage prediction models that can assist decision-makers in mitigating the impacts of extreme weather on critical infrastructure.

VI. References

[1] H. H. Alhelou, M. Hamedani-Golshan, T. Njenda, and P. Siano, "A survey on power system blackout and cascading events: Research motivations and challenges," *Energies*, vol. 12, no. 4, p. 682, Feb. 2019.

[2] A. H. Sanstad, Q. Zhu, B. Leibowicz, P. H. Larsen, and J. H. Eto, "Case studies of the economic impacts of power interruptions and damage to electricity system infrastructure from extreme events," Lawrence Berkeley Nat. Lab. (LBNL), Berkeley, CA, USA, Tech. Rep. 3221384, 2020.

[3] N. Bhusal, M. Abdelmalak, M. Kamruzzaman, and M. Benidris, "Power system resilience: Current practices, challenges, and future directions," *IEEE Access*, vol. 8, pp. 18064–18086, 2020, doi: 10.1109/ACCESS.2020.2968586.

[4] M. Abdelmalak, J. Cox, S. Ericson, E. Hotchkiss, and M. Benidris, "Quantitative resilience-based assessment framework using EAGLE-I power outage data," *IEEE Access*, vol.

[5] Oak Ridge National Laboratory, "Environment for analysis of geo-located energy information (EAGLE-I) system," accessed May 2, 2023. [Online]. Available: https://eagle-i.doe.gov/.

[6] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.

[7] M. Awad and R. Khanna, "Support vector regression," in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, 1st ed. Berkeley, CA, USA: Apress, 2015, pp. 67–80.

[8] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785-794, doi: 10.1145/2939672.2939785.

[9] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, no. 5, pp. 1189-1232, 2001, doi: 10.1214/aos/1013203451.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.

[11] National Hurricane Center. [Online]. Available: https://www.nhc.noaa.gov/.

[12] Homeland Infrastructure Foundation-Level Data. [Online]. Available: https://hifld-geoplatform.hub.arcgis.com/.

[13] Hurricane Mapping. [Online]. Available: http://hurricanemapping.com.

[14] National Weather Service (NWS) Valid Time Extent Code (VTEC) Archives. Available: https://mesonet.agron.iastate.edu/info/datasets.

[15] H. Liu, R. A. Davidson, and T. V. Apanasovich, "Statistical forecasting of electric power restoration times in hurricanes and ice storms," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2270-2279, 2007.

[16] H. Liu, R. A. Davidson, and T. V. Apanasovich, "Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms," *Rel. Eng. Syst. Safety*, vol. 93, no. 6, pp. 897-912, 2008.

[17] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, "Predicting hurricane power outages to support storm response planning," *IEEE Access*, vol. 2, pp. 1364–1373, 2014.

[18] R. Nateghi, S. D. Guikema, and S. M. Quiring, "Forecasting hurricane-induced power outage durations," *Natural Hazards*, vol. 74, pp. 1795-1811, 2014.

[19] D. Owerko, F. Gama, and A. Ribeiro, "Predicting power outages using graph neural networks," in *Proc. IEEE GlobalSIP*, Nov. 2018, pp. 743-747.

[20] S. Lee, L. Chen, S. Duan, S. Chinthavali, M. Shankar, and B. A. Prakash, "URBAN-NET: A network-based infrastructure monitoring and analysis system for emergency management and public safety," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2016, pp. 2600-2609.

[21] C. Brelsford, S. Tennille, A. Myers, S. Chinthavali, V. Tansakul, M. Denman, M. Coletti, et al., "A dataset of recorded electricity outages by United States county 2014–2022," *Scientific Data*, vol. 11, no. 1, p. 271, 2024.

[22] S. Lee, J. Choi, G. Jung, A. Tabassum, N. Stenvig, and S. Chinthavali, "Predicting power outage during extreme weather events with EAGLE-I and NWS datasets," in *Proc. IEEE IRI*, 2023, pp. 211-212.

[23] Homeland Infrastructure Foundation-Level Data. [Online]. Available: https://hifld-geoplatform.hub.arcgis.com/.

[24] S. Lee, L. Chen, S. Duan, S. Chinthavali, M. Shankar, and B. A. Prakash, "URBAN-NET: A network-based infrastructure monitoring and analysis system for emergency management and public safety," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2016, pp. 2600-2609.

[25] S. Lee, S. Chinthavali, S. Tennille, J. Chae, A. Tabassum, V. Tansakul, D. Redmon, R. Moncrief, and A. Myers, "Graph-based cascading impact estimation for identifying crucial infrastructure components," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2022, pp. 6749-6751.

[26] M. R. Allen, S. J. Fernandez, J. S. Fu, and M. M. Olama, "Impacts of climate change on sub-regional electricity demand and distribution in the southern United States," *Nature Energy*, vol. 1, no. 8, pp. 1-9, 2016.

[27] P. Sedgwick, "Pearson's correlation coefficient," *BMJ*, vol. 345, p. e4483, 2012.

[28] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.

[29] A. D. Dongare, R. R. Kharde, and A. D. Kachare, "Introduction to artificial neural network," *Int. J. Eng. Innov. Technol. (IJEIT)*, vol. 2, no. 1, pp. 189–194, 2012

[30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

Acknowledgment

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (https://www.energy.gov/doe-public-access-plan).