IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 9 October 2024, accepted 30 October 2024, date of publication 29 November 2024, date of current version 16 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3509263

RESEARCH ARTICLE

Probabilistic Restoration Modeling of Wide-Area Power Outage

NICHOLAS WILLEMS[1], BANDANA KAR[2], SAMUEL LEVINSON[3], BENJAMIN TURNER[1], JOHN BREWER[4], AND MARIJA PRICA[1]

[1]National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, USA
[2]National Renewable Energy Laboratory, Golden, CO 80401, USA
[3]Duquesne Light Company, Pittsburgh, PA 15219, USA
[4]NETL, Morgantown, WV 26507, USA

Corresponding author: Bandana Kar (bandana.kar@nrel.gov)

This work was supported by United States (U.S.) Department of Energy's (DOE) Office of Cybersecurity, Energy, Security, and Emergency Response (CESER) and carried out at the National Energy Technology Laboratory and Oak Ridge National Laboratory, agencies of U.S. Government, through a support contract.

ABSTRACT The timely restoration of electricity services following extreme weather events is crucial to meet customer energy resilience as well as for the economic and national security of the United States. Electricity restoration plans are needed to monitor multi-state power restoration operations, undertake resource planning, and analyze system vulnerabilities. However, these plans are proprietary to utility companies and not readily available to first responders and decision-makers. The purpose of the Restoration of Power Outage from Wide-area Severe Weather Disruptions (RePOWERD) project was to (i) *determine which type of model* – empirical, statistical, or probabilistic-most accurately predicts restoration times for distribution-level power outages caused by Category 2 or higher hurricanes, and (ii) identify the impact on restoration times of *various predictor variables*, such as power outage impact (i.e., customers impacted), storm characteristics, land-use patterns, and baseline customer density at county-service-area resolution. Seven models were developed for hurricanes that made landfalls from 2017 - 2022 along the Southeast region of the United States (Irma, Michael, Harvey, Laura, and Zeta). Comparing methods for predicting the time to restore power to 95 % of impacted customers for these hurricanes revealed that: 1) outage magnitude (i.e., initial number of customers experiencing outages and their spatial distributions) is the strongest predictor of recovery time; 2) the performance of the log-linear regression model was similar to more complex, less interpretable models (e.g., accelerated failure time); and 3) the final log-linear regression model achieved strong overall performance, but it struggled with certain hurricanes (overall adjusted R^2 of 0.6730, with a minimum of 0.4006 for Harvey and maximum of 0.8636 for Zeta). Using the log-linear regression model to forecast restoration time is *viable*, as all input data are publicly available prior to or at storm onset; however, the model *reliability* would benefit from expanding the scope of predictors and training data.

INDEX TERMS Electricity restoration, energy resilience, probabilistic modeling, tropical storms, wide-area power outage.

I. INTRODUCTION

Access to reliable electricity is crucial for the operation of basic energy services needed for survival (e.g., heating, cooling, ventilation, critical plug loads) as well as energy resilience (i.e., the functioning of diverse and interdependent systems, such as fuel and energy systems, telecommunications networks, emergency response systems, and other critical assets). The energy infrastructure of the mainland United States are aging and struggling to handle changing demands due to electrification needs [1], [2]. The infrastructure are highly vulnerable to extreme weather events which could lead to reduced energy efficiency and more power outages [3]. Although physical improvements

The associate editor coordinating the review of this manuscript and approving it for publication was Mouquan Shen.

to the energy infrastructure and the expansion of renewables may reduce power outage duration and size, the push to electrify the building sector and the disproportionate impacts of power outages on communities [22], [23], [24] make rapid restoration of electric services following extreme events (e.g., hurricanes) vital.

Large-scale and long-term power outages are becoming more common in the U. An analysis of national power outage data since 2000 revealed that the average annual number of weather-related power outages increased by roughly 78% during 2011 through 2021, compared to 2000 through 2010 [4]. During 2000 to 2021, 83% of major U.S. power outages (impacting at least 50,000 customers) were caused by tropical storms (which include hurricanes), winter storms, wildfires, etc. [4], [5], [6], [7], [8]. Tropical storms and hurricanes are among the most common system-disrupting events the bulk electric system (BES) faces [19]. According to the Energy Information Administration (EIA), in 2020, the average U.S. customer experienced over eight hours of power outage due to extreme weather events [11].

Considering that extreme weather events are forecasted to be more frequent and intense [9], [10], major outages may occur both due to increased electricity demand and damaged infrastructure compromising electricity supply. During a tropical storm event, power outages occur due to damages to overhead power lines from high wind and flying debris as well as flooding, lightning, and heavy precipitation that cause secondary damages. Although transmission outages can have a significant impact on large numbers of customers, most of the power outages during storm events occur on the distribution systems rather than on the transmission system [20]. Frequent and long-duration power outages cost the U.S. economy billions of dollars and, in some cases, loss of life [12], [13], [14], [15], [16]. To reduce the impact and likelihood of power outages and improve restoration time, it is imperative to analyze the vulnerabilities of the bulk electric system (BES).

Assessing system vulnerabilities and potential risks to the BES from future extreme events is pivotal to supporting critical infrastructure protection programs and ensuring adequate resource allocation prior to extreme events. Such information is essential to expedite restoration, thus reducing adverse societal and financial impacts of power outages. Considering the multi-billion-dollar annual cost to the U.S. economy caused by extreme weather event-induced outages and the dependency of critical assets (e.g., hospitals) and human survival (as illustrated by the increasing heat stress mortality risk and other events) on electric service, the U.S. Department of Energy (DOE) launched the Building a Better Grid Initiative in early 2022 to upgrade and modernize the transmission systems to achieve grid resilience [21]. With ∼80% of the U.S. energy infrastructure being privately owned [17], utility companies tend to have proprietary models of their service areas to determine system damage pertaining to extreme events as well as estimate and expedite outage restoration. However, these disparate models and their outcomes are unavailable to policymakers (e.g., DOE), first responders, or emergency management agencies to support resource allocation and large-scale response efforts following outages. The Restoration of Power Outage from Wide-area Severe Weather Disruptions (RePOWERD) project fills this gap [18] by developing models to forecast restoration times for distribution-level power outages caused by Category 2 or higher hurricanes to assist policymakers and first responders with system vulnerability assessment and resource planning.

This paper discusses and compares seven empirical, statistical, and probabilistic models. These models were developed for Hurricanes Irma, Michael, Harvey, Laura, and Zeta, which made landfall from 2017 - 2022 along the Southeast region of the United States. Power outage data (outage magnitude), storm characteristics, land-use distribution, and customer density were all included. The paper also identifies (i) *which type of model* - empirical, statistical, or probabilistic-most accurately predicts restoration times for distribution-level power outages caused by Category 2 or higher hurricanes and (ii) the impact on restoration times of *various predictor variables*, such as wind speed and land use at a county-service-area resolution.

The remainder of the manuscript is organized into five sections. Section II provides an overview of the restoration models that are either discussed in the literature or used in practice. Section III introduces the case study tropical storm events and associated data sets that were used to develop the models. A discussion of the seven models - an empirical model (exponential decay function), five statistical models (regression tree, random forest, log-linear regression, and two accelerated failure time models), and a spatial regression model - is presented in Section III. The model results and a comparison of their performance based on accuracy in estimating restoration time are presented in Section IV. The significance of the results from a practitioner perspective, concluding remarks, and future directions for expanding this process-oriented model are presented in Section V.

II. BRIEF OVERVIEW OF RESTORATION MODELS

A multitude of studies have modeled restoration times for critical infrastructure systems – including electric power grids and water distribution systems – following extreme events such as hurricanes and ice storms. These models can be grouped into the following categories:

Markov Models use discrete-state and discrete-transition Markov processes to model the restoration process. Zhang modeled post-earthquake recovery of multiple lifelines as a discrete-state and discrete-transition Markov process [29]. The model assumed that the lifelines compete with each other for restoration resources, which is not typically true for real-world processes. Zhang illustrated the difficulty of using Markov modeling for restoration time estimation [29]. This method requires state vectors and transition probabilities, which are often difficult to obtain.

Deterministic Resource Constraint Models are used to model the actual restoration process based on assumptions regarding the number of repairs that can be made within an hour given a certain number of repair crews and/or materials. Compared to Markov modeling techniques, these models are based closely on real-world processes. Wanik et al. estimated power recovery times using information provided by a utility on the number of electrical faults repaired and number of repair crews working per day after three major outage events (Tropical Storm Irene, the 2011 nor'easter, and Hurricane Sandy) in Connecticut [28]. The restoration time was determined as a function of the total outages on the system, the average daily number of active repair crews, and the estimated number of daily restorations for each crew. These models require extensive proprietary data describing utility repair crew activities.

Restoration Simulation Models estimate power system equipment failure rates during storms and average repair times. Balijepalli et al. modeled reliability of power systems to lightning storm events by relating empirically defined rates of lightning strike failures on overhead distribution equipment to a probability density function for intensity and duration of lightning storm events obtained via the bootstrap method [30]. The authors quantified reliability of the system using average system interruption duration and frequency, momentary average interruption duration and frequency reliability indices. The models, however, required detailed information about the system topology and construction (i.e., power line height and shielding). Even relative to deterministic resource constraint models, restoration simulation models tend to be data-intensive, time consuming, and may only be applied to specific systems.

Empirical Curve Fitting fits an empirical model, such as an exponential decay function, to historic data to estimate the restoration time. There have been several variations of this model. Reed used a gamma distribution to relate the number of outages remaining with a duration less than the corresponding duration on the x-axis [25]. To estimate restoration time, Duffey and Ha [26], Duffey [27] implemented exponential regression to estimate a decay coefficient for power outage restorations following a variety of historic events including hurricanes, wildfires, and ice storms. This technique is generally effective at fitting historic power outage restoration data with covariates (e.g., wind speed) [26], [27]. While these models depend less heavily on proprietary data, the variables used in this technique are generally not available before or during power restoration operations. Being purely backward-looking limits these specific models' utility for pre-storm resource planning and situational awareness.

Statistical Models tend to describe the relationships between outage duration and covariates to predict restoration time across space [31]. They can be seen as a generalization of the model proposed by Duffey [26], [27], which expands the scope of predictive variables. These models are generally used for restoration planning in case of weather-induced disasters [26], [27], [33], [34]. Liu et al. [35] used customer outage data from three major East Coast power companies to predict the number of power outages likely to occur from hurricanes and ice storms in a 3-kilometer square grid cell using a spatial, negative-binomial generalized linear mixed model. The authors also used wind speed and duration, land cover data, and soil drainage levels as well as utility-specific information including equipment inventories, number of protective devices, and company indicator covariates to develop an outage prediction model for a specific utility service area. Guikema et al. implemented a spatially generalized hurricane outage prediction model to determine the cumulative proportion of customers without power due to a hurricane at the census tract level [36]. This model also used topography, land use, soil moisture levels, hurricane wind characteristics, customer density, and land use/cover type in a multivariate model to demonstrate the impact of several variables on widespread power outages due to hurricanes. Davidson et al. [31] implemented a failure time model that incorporated utility-specific data sets, a wind field simulation model, and rainfall data to estimate county-level restoration times for hurricane and ice storm events. This multivariate model effectively predicted the shape of the restoration curves with an average overestimation of 7.2 hours (35 percent) [31].

Several restoration models have been developed to estimate restoration time for different extreme weather events. Considering that energy restoration is a complicated process, each model has its advantages and limitations depending upon the assumptions and input variables. In this study, seven models were developed using data for five historic hurricanes. The objective was to compare and contrast the performance of the models, identify the significant variables influencing restoration time, and produce a reliable model with high accuracy that could be used by non-utility decision-makers.

III. METHODOLOGY

Because of the high vulnerability of energy infrastructure to tropical storm events (which include hurricanes), it is crucial to develop a process-oriented restoration model to provide situational awareness information for resource planning and recovery. The characteristics of hurricanes, including their eye size, wind speed, amount of rainfall, pressure, and the area affected, can differ greatly. As a result, some hurricanes can lead to widespread power outages, whereas others may remain stationary over a particular area, leading to more concentrated severe flooding and extended power outages. To account for this variability, the power outage data for the five hurricanes (Irma, Michael, Harvey, Laura, and Zeta) that significantly affected the Southeast region from 2017 through 2022 (Figure 1) were analyzed and used to build the seven models discussed below. These storms led to power outages affecting hundreds of thousands to millions of customers and required substantial mobilization of repair crews.

Hurricane Irma, a Category 4 storm, was roughly the size of Texas and impacted southern Florida in August 2017. Its slow movement meant that some areas were affected for over 24 hours, leading to power outages for nearly 60 percent of Florida's approximately 10 million electric

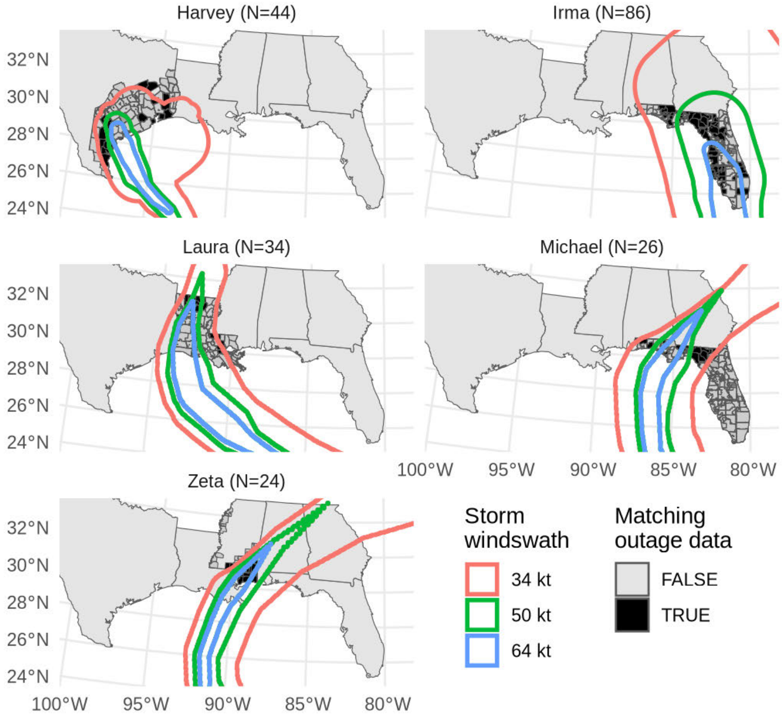


FIGURE 1. **Comparison of storm path and data coverage. N is the number of county-service-areas with matching outage data. Dark polygons are the intersection of a county and utility service area for which matching outage data are available.**

customers [37], [38]. Hurricane Michael, a Category 5 storm, struck the Florida Panhandle in October 2018 with record-setting sustained winds of 151 miles per hour. The hurricane caused devastating damage over a narrow 100-mile band, and left about 400,000 customers without power, mainly in the Florida Panhandle; power restoration took around 23 days [39], [40]. Hurricane Harvey hit a narrow band in Texas in August 2017 as a Category 4 storm, causing significant rainfall and flooding [41]. This led to over 300,000 customers without power due to damage to power plants and transmission and distribution infrastructure in Texas and Louisiana [42]. In August 2020, Hurricane Laura, one of the strongest hurricanes to make landfall as a Category 4 storm in Louisiana in 164 years, not only disrupted crude oil production along the Gulf Coast for 15 days but also caused significant damage to transmission and distribution systems [43], [44]. As a result, more than 900,000 customers in Louisiana and Texas were left without power [45]. Lastly, Hurricane Zeta, a Category 2 storm, made landfall in southern Louisiana and Mississippi in October 2020 [46]. Despite being a Category 2 hurricane, it caused significant power outage, leaving more than 2.5 million customers in the Southeast region without power [47].

A. DATA ACQUISITION AND PROCESSING

Several datasets capturing impacted area characteristics, hurricane characteristics, and outages were acquired to inform the models (Table 1). Two considerations drove the selection of data sets and their sources. First, the predictors should be generally available preceding or at the onset of a storm event. Although other models (discussed in Section II) have achieved strong predictive accuracy, they rely on ex-post analyses, which are not useful for pre-event resource planning and decision-making activities. Second, data should not be proprietary as it can be expensive to acquire and can lead to licensing restrictions. Though historical outage data for the five hurricanes were purchased from PowerOutage.us, real-time outage data can be accessed for free online from this source. Another alternative is to use outage data from Environment for Analysis of Geo-located Energy Information (EAGLE-I). The combined spatial and outage datasets provide restoration time (etr_{95}, time for the outage peak to reduce by 95 percent) alongside various predictors, resolved at a county-service-area resolution.

1) SPATIAL DATA

The utility service area boundary layer is publicly available from the Homeland Infrastructure Foundation-Level Data (HIFLD) dataset. Each boundary polygon represents electric power retail service territories for residential, industrial, and commercial customers. However, this data layer includes significant overlap of service areas between utilities. For instance, in Florida, Duke Energy (responsible for energy generation and transmission) covers more than half of the state, with multiple distribution utilities functioning within its service area boundary. Because hurricane outages

TABLE 1. Data sources and vintages. *No energy justice variables were identified as significant in the analysis that follows.

Data	Source	Vintage
County boundaries	U.S. Census Bureau	2020
Utility service areas	HIFLD	2022
Land use cover	Commission for Environmental Cooperation	2020
Land use change	Commission for Environmental Cooperation	2015-2020
Energy justice*	DOE J40	2022
Storm paths and wind speed	National Hurricane Center	Various, corresponding to storm dates
Outage number time-series	Power Outage US	Various, corresponding to storm dates

primarily result from distribution system failures, this dataset was processed in an attempt to better represent the most-local utility, presumed to be responsible for addressing the distribution system outage restoration. Specifically, spatial differences between service areas were computed, starting with the smallest service area and subtracting it from all larger areas; this process was repeated for each state across all service areas. While distribution-level outage restoration is within the purview of utilities, emergency response activities and resource planning efforts are generally undertaken by counties and cities. To minimize the issue stemming from redundant counting due to overlapped service boundary areas and assist first responders with resource planning, the 2020 county boundary layer from the U.S. Census Bureau was intersected with the processed service areas to generate the functional spatial unit for analysis, which is the county-service-area.

Storm data, specifically, the wind swath files, were acquired from the National Hurricane Center's (NHC) "best track" projections. The NHC also produces both forecast and retrospective hurricane tracks. For each storm, the wind swath data records the positions of a storm at 6-hour intervals and at three wind speed thresholds (greater than 34 knots, greater than 50 knots, and greater than 64 knots). For each wind speed threshold, the wind-swath was intersected with the county boundary and utility service area boundary to determine the fraction of a polygon (a county or utility service area) being subjected to each wind speed (Figure 1). By relating wind speed with outage numbers for each county over a duration, predictions can be generated for future storms using the NHC storm track predictions as inputs.

2) OUTAGE DATA

For each hurricane, the outage data were procured from PowerOutage.us for the state(s) that was directly in the storm's path (Figure 1). These data provide a baseline of total customers and number of customers experiencing an outage (i.e., outage counts from here onward) for cities, counties, and utility service areas at varying temporal resolutions, which settles around 20-minute intervals during high outage times. For each storm and county/utility service area, the maximum outage number magnitude and time were calculated. In many cases, the outage does not decay monotonically to zero, whether due to real phenomena or noise in the data. To account for this, the outage was considered to be resolved only when consecutive observations spanning at least 1 hour showed that the outage size had been reduced by 95 percent. This effectively filters out cases where outages are reported as resolved at one instant, only to rebound in subsequent observations.

B. RESTORATION MODELS

Two of the most prominent models of restoration time are empirical curve fitting and statistical models. While the former is based on observations of restoration patterns, statistical models tend to be regression models that include features beyond the outage profile to forecast restoration time. To inform decision-making, the first iteration of the restoration model was built using only outage data and profile (Figure 2, top). Progressively, additional features were added, increasing model complexity and utility in identifying underlying variables that influence restoration time (Figure 2). A comparison of the model performance across all storms is presented in Table 2); a discussion of the model performance and the best-performing model structure is presented in Section IV.

TABLE 2. Comparison of performance between model formulations including root mean squared error (RMSE) and adjusted R^2. *Estimated using the number of variables alone (not including spatial variation).

Model	RMSE	Adjusted R^2
Exponential Decay	57.43	0.3296
Log transform	62.01	0.5052
Regression Tree	53.91	0.6467
Random Forest	59.80	0.6621
Log linear Regression	57.60	0.6730
Log linear AFT	56.76	-
Exponential AFT	53.79	-
Spatial Regression	54.96	0.7052*

1) MODEL BUILDING BLOCKS

One of the simplest models of outage restoration time is based on empirical observation. Duffey et al. suggested that the number of customers experiencing an outage ($N(t)$) decays exponentially over time from its initial peak (N_0) toward a background or irreducible outage (N_i) [27].

$$N(t) = N_i + (N_0 + N_i)e^{-\alpha t} \tag{1}$$

Assuming the irreducible outage rate is zero, this implies that the 95 percent restoration time (etr_{95}) can be expressed as:

$$\mathrm{etr}_{95} = \frac{1}{\alpha} \ln \frac{N_0}{0.05 \cdot \#\,\mathrm{Customers}} \tag{2}$$

A model with this form (with only log-max-outages and log-number-of-customers as predictors) achieves an

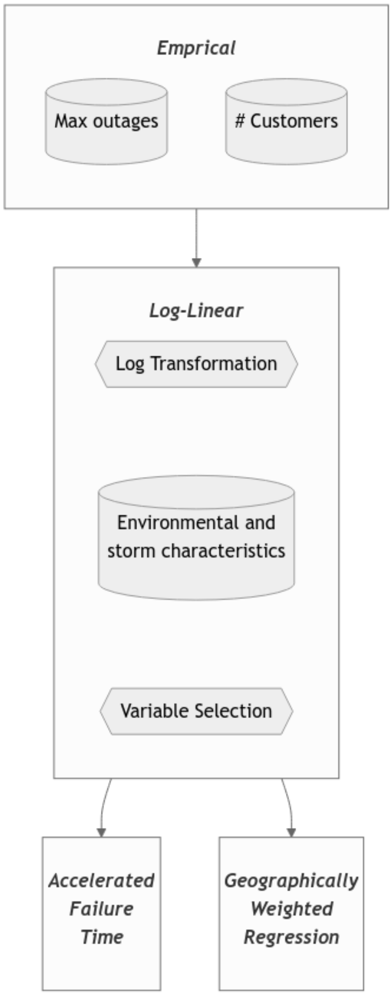


FIGURE 2. Flow diagram illustrating the addition of model building blocks. New model structure (hexagons) and data (cylinders) are added from top to bottom, increasing model complexity.

adjusted R^2 of 0.3296. The data quality and scope (different areas experiencing different storm conditions) are possible reasons why this model performed poorly with these data. Although slightly more difficult to interpret, it makes mathematical sense to log-transform the response variable (restoration time) to ensure that it is strictly greater than zero. Using $ln(\text{etr}_{95})$ as the response variable in equation 2 increases the adjusted R^2 to 0.5052.

Building on this foundation, a more complex model could include other predictors, such as geographic location or storm characteristics that influence restoration time. Given that storm outages occur mainly because of distribution-level failures [31], it was hypothesized that the outage severity and restoration time will be driven by utility characteristics (i.e., restoration crew size and availability), storm strength, and distribution system vulnerability (Table 3).[1]

[1] Unlike Davidson et al. [31], a single-stage model was formulated that predicts recovery time given the peak outage.

TABLE 3. Hypothesized drivers and proxies for outages stemming from the distribution system during tropical storms.

Driver	Proxy
Restoration crew availability	Storm size (larger storms limit resource availability in neighboring areas) Utility size (larger utilities have more resources) Utility and county recovery
Storm strength	Wind speed Flooding
Distribution system vulnerability	Network size Presence of trees Presence of storm buffers (e.g., watersheds)

TABLE 4. Model parameters and significant terms across data subsets. P-values are reported for the full dataset. Xs denote terms that are still significant when fitting the model to a single storm.

		Storm:	*Harvey*	*Irma*	*Michael*	*Laura*	*Zeta*
Parameter	**Coefficient**	**P-value**					
(Intercept)	-0.63897	5.70e-01					X
Crew availability							
log(max_outages)	1.334648	7.32e-22	X	X		X	X
log(customer_count)	-1.27529	1.05e-22	X	X	X	X	X
log(utility_outages)	0.16746	1.75e-03		X			
Storm strength							
windswath_50	1.502427	3.48e-10		X	X		
month(storm_start)	0.165474	1.24e-01					
System vulnerability							
lu_frac_6	2.279749	1.36e-01					X
lu_frac_18	1.033322	1.45e-02					
luc_frac_603	11.73669	6.00e-04		X			X
luc_frac_1403	-191.69	5.49e-10		X			

2) LOG-LINEAR

While forecasted restoration time is crucial for resource planning, it is also critical to determine the underlying factors that influence restoration time. For instance, if crew size and crew availability delay restoration, then both the city and the utility whose service area is experiencing outages would have to coordinate with other utilities for personnel and resource sharing. Hence, to identify which variables are most predictive of restoration time given a log-linear form, the following modeling techniques were deployed: correlation matrices, least absolute shrinkage and selection operator (LASSO) regression, regression trees, and random forest. From the union of variables identified across techniques, only the variables with significant p-values (<0.01) were selected. The variables selected could be grouped together based on the categories originally assumed (Tables 3 and 4).

3) ACCELERATED FAILURE TIME

Davidson et al. [31] used an accelerated failure time (AFT) model with utility and storm characteristics as predictors. An AFT model is statistical model that uses a log-transformed response variable and can assume a non-normal distribution of residuals. The AFT structure makes physical sense because the restoration time is strictly greater than zero [31]. For

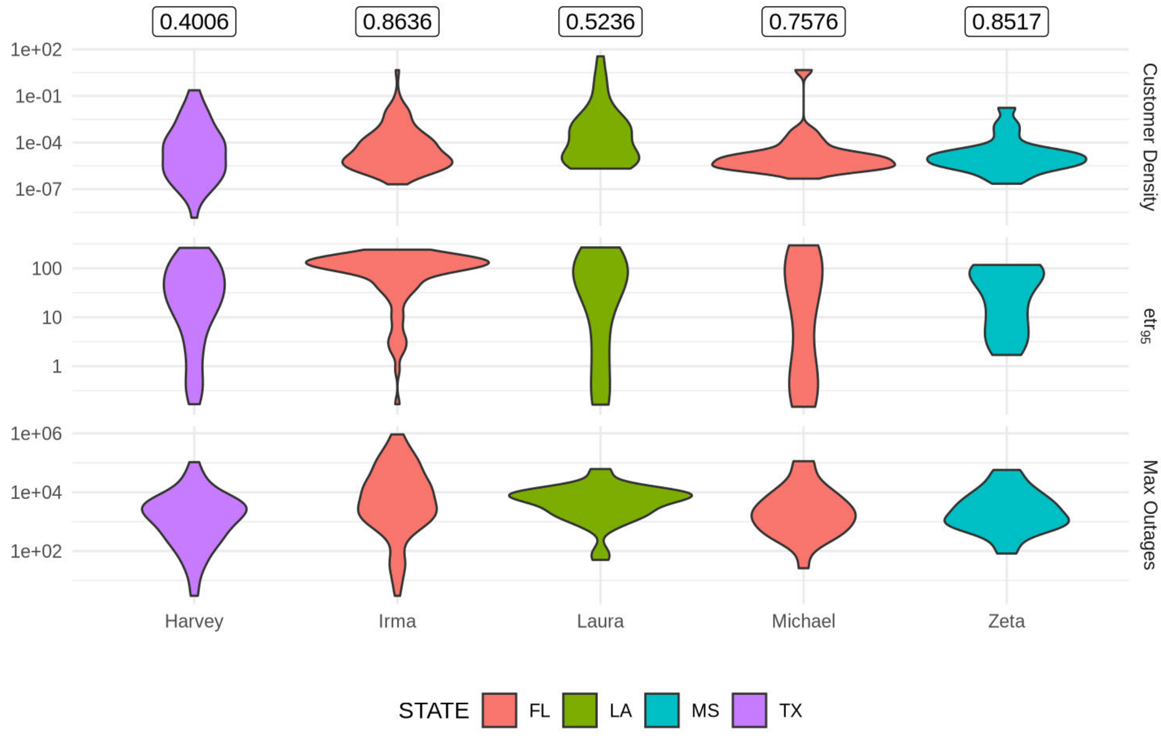


FIGURE 3. Comparison of storm and outage characteristics between events. The width of each element is proportional to the number of observations at the corresponding y-value.

comparison purposes, two AFT models with the same set of predictors as the log-linear model were used. The distribution of residuals that produced the lowest Akaike Information Criterion (a comparative measure of model performance given the same dataset) score were log-linear and exponential. The AFT performs slightly better but is less interpretable compared to the log-linear model (Table 2).

4) GEOGRAPHICALLY WEIGHTED REGRESSION

Geographically weighted regression (GWR) is an emergent technique that allows regression coefficients to vary in space. It has been used for exploratory purposes but is less accepted as a predictive tool [32]. For comparison, the R package 'spgwr' was used to produce a spatially-weighted regression model, which performs slightly better than the standard log-linear regression model. Like the AFT models, the geographically weighted model includes the same set of predictors as the log-linear model.

IV. RESULTS AND DISCUSSION

The results of the variable selection process point to system characteristics important to restoration time. Surprisingly, few variables associated with storm strength were selected. Storm strength is likely a strong predictor of outage size, but that contribution is erased by the inclusion of variables related directly to the maximum number of outages. Land use categories 1-6 (Table 4) each denote the coverage of different kinds of trees. Interestingly, whereas the coverage of mixed forest (lu_fra_6) was the most important land-use characteristic, differences in tropical or sub-tropical broadleaf evergreen forest coverage were identified as the most important land-use changes (luc_frac_603, luc_frac_1403). Different tree types may have opposite effects on restoration time, or this may be a regional phenomenon (i.e., the changes are only significant in areas where one type of tree cover exists). Perhaps counter intuitively, shrinking of wetland areas was negatively associated with recovery time (luc_frac_1403 is the amount of wetland (code 14) converted to broadleaf evergreen forest (code 3)), possibly because of improved access to infrastructure. Demographic and income variables did not increase predictive power. Filtering by outage size had little impact on accuracy. Adding a dummy variable representing the county, however, caused the largest improvement in predictive performance (adjusted R^2 increased to 0.83 for the log-linear regression model). This indicates that other county-specific features might be relevant to predicting restoration times.

Performance of the log-linear regression model was found to vary significantly across storms (Figure 3), with especially poor performance for Hurricane Harvey in Texas. The model struggles more with fast restoration time areas (Figure 4), which are less common in the data. Although the underlying cause for performance degradation is unclear, storm characteristics might be one explanation for model

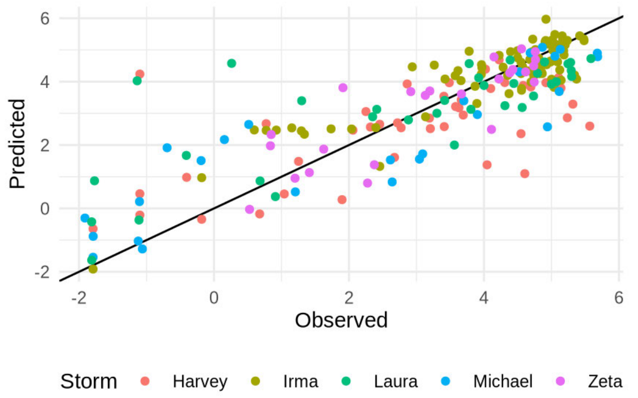


FIGURE 4. Comparison of predicted and observed values of $\ln etr_{95}$ using the log-linear regression model.

performance. For instance, Hurricane Harvey was notable for a significant amount of rainfall and flooding while its wind speed dropped from a Category 4 to a Category 1 hurricane. The outage data quality may also be a factor in model performance. Both Hurricanes Harvey and Laura have relatively sparse data coverage and a smaller proportion of regions within the high wind speed path of the storm (Figure 1). It is also notable that the two storms for which the log-linear regression model performed worst caused significant damage to transmission infrastructure rather than distribution infrastructure, which was the focus of this work.

To assist with the selection of a forecasting model that could be used by first responders and emergency managers, a methodological survey of forecasting methods was conducted with the objective of motivating continued research. The log-linear regression model does well for several scenarios (adjusted R^2>0.75 for three of five storms: Irma, Michael, and Zeta), but performance is significantly lacking for others (Harvey and Laura). While including rainfall totals may improve model performance for Hurricane Harvey, this is unlikely to improve performance across the board as rainfall and precipitation were not significant for the second worst performing hurricane (Laura). Another approach would be to filter transmission outages from the outage data. The restoration characteristics for these types of outages are likely to be different from diffuse, distribution-level outages.

V. CONCLUSION AND FUTURE DIRECTIONS

Power restoration times were modeled for a range of geographies and storm events. Every storm event exhibited a wide range of observed restoration times: from less than 1 hour to over 100 hours (Figure 3). The outage size (i.e., spatial distribution of outages) was not clearly related to the distribution of restoration times. The number of customers experiencing outages and outage size were the most important predictors across all storm events (Table 4). A larger outage size was associated with an increase in restoration time, but a larger number of customers was associated with a decrease in restoration time. Other predictors were significant in the multi-storm model but not when modeling individual events. This may be a result of the data subset (e.g., land use is more similar within a state and therefore not a good predictor of restoration time) or dependent on the storm itself; although Irma and Michael both impacted Florida, land use change variables were only significant in Irma.

The decay function was able to estimate restoration time and rate of restoration, the two parameters essential for resource planning at both the county and utility service area levels. The log-linear regression model identified the variables contributing to the restoration time; however, the performance of both of these models is influenced by the event characteristics, the total number of customers experiencing outages, the impacted area, and the baseline number of customers. For instance, the decay function performed well in terms of computing the estimated time of restoration and rate of restoration (with high accuracy >90 percent based on R^2) in the case of wide-area large-scale outages, but the model performed poorly in the case of events with fluctuating outages due to continuous disruptions to the system over a period (e.g., the conditions during Hurricanes Harvey and Laura). A similar trend was observed in the log-linear regression model, and the model did not identify counties and utility service areas where restoration would take longer depending on the underlying impacts, base customer, county type, and extent of damage.

Overall, the models succeeded in forecasting restoration time based on the underlying variables and processes that are essential for restoration. Among all the models, GWR, log-linear, and random forest performed best. Because restoration time is generally influenced by county characteristics (urban versus rural, vegetative versus non-vegetative), from a planning perspective, the log-linear regression model outputs could be used to geotarget counties where utilities should undertake pruning and vegetation growth management efforts to minimize power line damages due to hurricane winds and expedite restoration. With increased focus on deploying energy storage systems (e.g., battery, thermal) and renewables (e.g., photovoltaics), both on the grid and behind-the-meter, the model outputs could be used by utilities and residents alike in identifying counties where energy storage and renewables could be used to meet demand flexibility, increase reliability, maintain energy resilience, and increase passive survivability. Moving forward, the following steps will be taken to expand and enhance the models.

1) Differentiate between transmission and distribution outages.
2) Include Category 1 and 2 storm events, as well as events with fluctuating outages.
3) Incorporate utility characteristics (i.e., crew size, crew members, and resource availability for staging).

4) Incorporate behind-the-meter (e.g., battery and thermal storage), grid-edge, and community-scale energy solutions.
5) Replicate and expand the models to forecast restoration time considering the increase in number of outages due to climate extremes such as extreme heat, cold snaps, and wildfires.
6) Design and deploy an Agent Based Model (ABM) built on the variables used in these models to simulate the processes/agents involved in the restoration process (e.g., changing crew size, staging, accessibility, outage size, impact areas, etc.). This model will be based on the first version of ABM that was developed by Chen et al. [48] to assist with resource planning and expedite restoration time.

ACKNOWLEDGMENT

The authors would like to thank Dr. Olufemi A. Omitaomu, Dr. Yang Chen, Dr. Arun Iyengar, and Nicholas Roberts for their contributions to the RePOWERD Project and the development of the first version of the restoration model using empirical curve fitting. They also like to thank Jim Glass (manager, smart grid development at EPB Chattanooga) and Knoxville Utilities Board for sharing their insight into restoration planning following extreme events. Finally, they would like to thank the Environment for the Analysis of Geo Located Energy Information (EAGLE-I) team for their assistance in identifying the situational awareness information needed for EAGLE-I to support DOE-CESER with emergency response planning.

Neither the U.S. Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof. The findings and opinions presented here are those of the authors and do not reflect the policy or position of any of the afore-mentioned institutions. The U.S. government and the publisher, by accepting the article for publication, acknowledges that the U.S. 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 U.S. government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (https://energy.gov/downloads/doe-public-access-plan). A part of this research presented in this manuscript was carried out at by Dr. Bandana Kar during her tenure at the Oak Ridge National Laboratory (ORNL).

REFERENCES

[1] U.S. Department of Energy (DOE). (2023). *What Does it Take to Modernize the U.S. Electric Grid*. Accessed: May 12, 2024. [Online]. Available: https://www.energy.gov/gdo/articles/what-does-it-take-modernize-us-electric-grid

[2] American Society of Civil Engineers. (2021). *Energy*. Accessed: May 12, 2024. [Online]. Available: https://infrastructurereportcard.org/wp-content/uploads/2020/12/Energy-2021.pdf

[3] McKinsey Global Institute. (2020). *Climate Risk and Response: Physical Hazard and Socioeconomic Impacts Will Infrastructure Bend or Break Under Climate Stress*. Accessed: May 12, 2024. [Online]. Available: https://www.mckinsey.com/capabilities/sustainability/our-insights/will-infrastructure-bend-or-break-under-climate-stress

[4] Climate Central. (2022). *Surging Weather-Related Power Outages*. Accessed: May 12, 2024. [Online]. Available: https://www.climatecentral.org/climate-matters/surging-weather-related-power-outages

[5] J. Schuppe. (2021). *Hurricane Ida Power Grid Failure Forces a Reckoning Over Entergy's Monopoly in the South*. NBCNews. Accessed: May 12, 2024. [Online]. Available at: https://www.nbcnews.com/news/us-news/hurricane-ida-power-grid-failure-forces-reckoning-over-entergy-s-n1279971

[6] Energy Information Administration (EIA). (2021). *Hurricane Ida Caused At Least 1.2 Million Electricity Customers to Lose Power*. Accessed: May 12, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=49556

[7] J. W. Busby, K. Baker, M. D. Bazilian, A. Q. Gilbert, E. Grubert, V. Rai, J. D. Rhodes, S. Shidore, C. A. Smith, and M. E. Webber, "Cascading risks: Understanding the 2021 winter blackout in Texas," *Energy Res. Social Sci.*, vol. 77, Jul. 2021, Art. no. 102106, doi: 10.1016/j.erss.2021.102106.

[8] University of Houston. (2021). *The Winter Storm of 2021*. Accessed: May 12, 2024. [Online]. Available: https://uh.edu/hobby/winter2021/storm.pdf

[9] Intergovernmental Panel on Climate Change, "Summary for policymakers," in *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Cambridge, U.K.: Cambridge Univ. Press, 2022, doi: 10.1017/9781009325844.001.

[10] NOAA National Centers for Environmental Information (NCEI). (2024). *U.S. Billion-Dollar Weather and Climate Disasters*. Accessed: May 12, 2024. [Online]. Available: https://www.ncei.noaa.gov/access/billions/

[11] Energy Information Administration (EIA). (2021). *U.S. Electricity Customers Experienced Eight Hours of Power Interruptions in 2020*. Accessed: May 12, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=50316

[12] U.S. Department of Energy (DOE). (2018). *What Does it Take to Modernize the U.S. Electric Grid*. Accessed: May 12, 2024. [Online]. Available: https://www.energy.gov/gdo/articles/what-does-it-take-modernize-us-electric-grid

[13] J. Skarha, L. Gordon, N. Sakib, J. June, D. J. Jester, L. J. Peterson, R. Andel, and D. M. Dosa, "Association of power outage with mortality and hospitalizations among Florida nursing home residents after hurricane irma," *JAMA Health Forum*, vol. 2, no. 11, Nov. 2021, Art. no. e213900, doi: 10.1001/jamahealthforum.2021.3900.

[14] The White House (WH). (2021). *FactSheet: The Bipartisan Infrastructure Deal*. Accessed: May 12, 2024. [Online]. Available: https://www.whitehouse.gov/briefing-room/statements-releases/2021/11/06/fact-sheet-the-bipartisan-infrastructure-deal/

[15] C. Hauser and E. Sandoval. (2021). *Death Toll From Texas Winter Storm Continues to Rise*. New York Times. Accessed: May 12, 2024. [Online]. Available: https://www.nytimes.com/2021/07/14/us/texas-winter-stormdeaths.html

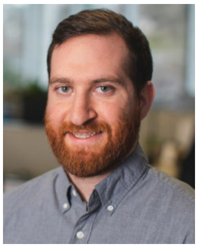
[16] A. Garcia. (2021). *Winter Storm Uri Cost Texas Between \$80 and \$130 Billion*. CHRON. Accessed: May 12, 2024. [Online]. Available: https://www.chron.com/politics/article/Texas-winter-storm-freeze-deaths-financial-cost-16585329.php

[17] U.S. Department of Homeland Security (DHS). (2021). *Energy Sector*. Accessed: May 12, 2024. [Online]. Available: https://www.cisa.gov/energy-sector

[18] B. Kar, J. Brewer, O. A. Omitaomu, B. Turner, Y. Chen, N. Roberts, S. Levinson, M. Prica, and A. Iyengar. *Restoration of Power Outage From Wide-Area Severe Weather Disruptions (RePOWERD)*. U.S. Department of Energy Office of Scientific and Technical Information. Accessed: May 12, 2024, doi: 10.2172/1887668.
[19] North American Electric Reliability Corporation (NERC). (2022). *2022 State of Reliability*. Accessed: May 12, 2024. [Online]. Available: https://www.nerc.com/pa/RAPA/PA/Performance%20Analysis%20DL/NERC_SOR_2022.pdf
[20] R. J. Campbell. (2013). *Underground vs. Overhead Distribution Wires: Issues to Consider*. Edision Electric Institute. Accessed: May 12, 2024. [Online]. Available: https://sgp.fas.org/crs/misc/R42696.pdf
[21] U.S. Department of Energy (DOE). (2022). *Building a Better Grid Initiative*. Accessed: May 12, 2024. [Online]. Available: https://www.energy.gov/gdo/building-better-grid-initiative
[22] U.S. Department of Energy (DOE). (2024). *Decarbonizing the U.S. Economy by 2050—A National Blueprint for the Buildings Sector*. Accessed: May 12, 2024. [Online]. Available: https://www.energy.gov/eere/articles/decarbonizing-us-economy-2050
[23] J. Kim. (2023). *Increasing Power Outages Don't Hit Everyone Equally*. Scientific American. Accessed: May 12, 2024. [Online]. Available: https://www.scientificamerican.com/article/increasing-power-outages-dont-hit-everyone-equally1/
[24] J. Kim. (2024). *Low-Income Communities Most at Risk for Power Outages Linked to Extreme Weather*. Finance and Commerce. Accessed: May 12, 2024. [Online]. Available: https://finance-commerce.com/2024/05/low-income-communities-most-at-risk-for-power-outages-linked-to-extreme-weather
[25] D. A. Reed, "Electric utility distribution analysis for extreme winds," *J. Wind Eng. Ind. Aerodynamics*, vol. 96, no. 1, pp. 123–140, Jan. 2008.
[26] R. B. Duffey and T. Ha, "The probability and timing of power system restoration," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 3–9, Feb. 2013, doi: 10.1109/TPWRS.2012.2203832.
[27] R. B. Duffey, "Power restoration prediction following extreme events and disasters," *Int. J. Disaster Risk Sci.*, vol. 10, no. 1, pp. 134–148, Mar. 2019, doi: 10.1007/s13753-018-0189-2.
[28] D. Wanik, E. Anagnostou, B. Hartman, and T. Layton, "Estimated time of restoration (ETR) guidance for electric distribution networks," *J. Homeland Secur. Emergency Manage.*, vol. 15, no. 1, pp. 1–13, Apr. 2018.
[29] R. H. Zhang, "Lifeline interaaction and post-earthquake urban system reconstruction," in *Proc. 10th World Conf. Earthquake Eng.*, 1992, pp. 5475–5480.
[30] N. Balijepalli, S. S. Venkata, C. W. Richter, R. D. Christie, and V. J. Longo, "Distribution system reliability assessment due to lightning storms," *IEEE Trans. Power Del.*, vol. 20, no. 3, pp. 2153–2159, Jul. 2005, doi: 10.1109/TPWRD.2005.848724.
[31] R. A. Davidson, H. Liu, and T. V. Apanasovich, "Estimation of post-storm restoration times for electric power distribution systems," in *Advances in Electric Power and Energy Systems: Load and Price Forecasting*. Hoboken, NJ, USA: Wiley, 2017, pp. 251–284.
[32] R. Bivand and D. Yu. (2023). *Spgwr: Geographically Weighted Regression*. R Package Version 0.6-36. [Online]. Available: https://CRAN.R-project.org/package=spgwr
[33] M.-Y. Chow, L. S. Taylor, and M.-S. Chow, "Time of outage restoration analysis in distribution systems," *IEEE Trans. Power Del.*, vol. 11, no. 3, pp. 1652–1658, Jul. 1996.
[34] R. Nateghi, S. D. Guikema, and S. M. Quiring, "Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes," *Risk Anal.*, vol. 31, no. 12, pp. 1897–1906, Dec. 2011, doi: 10.1111/j.1539-6924.2011.01618.x.
[35] 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. Saf.*, vol. 93, no. 6, pp. 897–912, Jun. 2008.
[36] 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, doi: 10.1109/ACCESS.2014.2365716.
[37] National Weather Service. (2017). *Hurricane Irma*. Accessed: May 12, 2024. [Online]. Available: https://www.weather.gov/mfl/hurricaneirma
[38] Energy Information Administration (EIA). (2021). *Hurricane Irma Cut Power to Nearly Two-Thirds of Florida's Electricity Customers*. Accessed: May 12, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=32992
[39] National Hurricane Center (NHC). (2019). *Hurricane Michael*. Accessed: May 12, 2024. [Online]. Available: https://www.nhc.noaa.gov/data/tcr/AL142018_Michael.pdf
[40] Florida Public Service Commission (FPUC). (2019). *Hurricane Outage Reports*. Accessed: May 12, 2024. [Online]. Available: http://www.psc.state.fl.us/Home/HurricaneReport
[41] National Weather Service (NWS). (2017). *Hurricane Harvey and Its Impacts on Southeast Texas*. Accessed: May 12, 2024. [Online]. Available: https://www.weather.gov/hgx/hurricaneharvey
[42] Energy Information Administration (EIA). (2017). *Hurricane Harvey Caused Electric System Outages and Affected Wind Generation in Texas*. Accessed: May 12, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=32892
[43] National Weather Service (NWS). (2020). *Hurricane Laura*. Accessed: May 12, 2024. [Online]. Available: https://www.weather.gov/lch/2020Laura
[44] Energy Information Administration (EIA). (2020). *Hurricane Laura Shut in More Gulf of Mexico Crude Oil Production Than Any Storm Since 2008*. Accessed: May 12, 2024. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=45376
[45] Entergy. (2020). *Entergy Provides Update on Hurricane Laura*. Accessed: May 12, 2024. [Online]. Available: https://www.entergynewsroom.com/news/entergy-provides-update-on-hurricane-laura
[46] National Hurricane Center (NHC). (2021). *Hurricane Zeta*. Accessed: May 12, 2024. [Online]. Available: https://www.nhc.noaa.gov/data/tcr/AL282020_Zeta.pdf
[47] J. Masers. (2020). *Hurricane Zeta's High Winds Cause Second Largest U.S. Power Outage of 2020*. Yale Climate Connections. Accessed: May 12, 2024. [Online]. Available: https://yaleclimateconnections.org/2020/10/hurricane-zetas-high-winds-cause-largest-u-s-power-outage-of-2020/
[48] Y. Chen, O. A. Omitaomu, N. Roberts, and B. Kar, "Multi-agent simulation based framework for power restoration time estimation at distribution level," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2023, pp. 2849–2859, doi: 10.1109/WSC60868.2023.10407135.

NICHOLAS WILLEMS received the B.S. degree from the Mechanical Engineering Department, The University of Texas at Austin, USA, 2016, the M.S. degree in dynamic systems and controls from the University of California, San Diego, USA, in 2018, and the Ph.D. degree in thermal fluid systems from the Mechanical Engineering Department, The University of Texas at Austin, in 2022. In 2022, he joined KeyLogic, Pittsburgh, where he is currently a Site Support Contractor with the DOE's National Energy Technology Laboratory (NETL). He is involved at the intersection of data science, energy systems modeling, and the low-carbon transition.

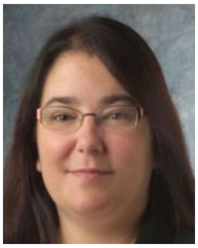
BANDANA KAR received the Ph.D. degree in geography (geoinformatics and remote sensing) from the University of South Carolina, Columbia, USA, in 2008. She is currently a Researcher V with the Energy Security and Resilience Center, National Renewable Energy Laboratory. She has been funded by DOE, DHS, NASA, and NSF to develop data driven solutions at the intersection of science, technology, and policy to improve energy resilience. With over 15 years of experience in academia, national laboratory, and the federal government, her expertise includes risk and resilience modeling, siting of onsite generation and storage, supply chain modeling of energy sources, geoinformatics, and remote sensing.



SAMUEL LEVINSON received a B.S. degree in materials science and engineering, the B.S. degree in engineering and public policy, and the M.S. degree in energy science, technology and policy from Carnegie Mellon University. He is currently with the Advanced Grid Solutions Team, Duquesne Light Company (DLC), Pittsburgh, where he leads the development of load forecasting models, the implementation of integrated distribution planning, and the development of data science methods. Prior to DLC, he was with KeyLogic, as a Site Support Contractor to NETL.

BENJAMIN TURNER received the B.S. degree in electrical engineering from the University of Pittsburgh, USA. He is currently a Senior Consultant with Deloitte Risk and Financial Advisory supporting NETL as a part of the Energy Markets Analysis Team. His areas of work include power market design and operation, regional power systems modeling, and carbon-free energy integration.

JOHN BREWER received the B.S. degree in mechanical engineering and the B.S. degree in aerospace engineering from West Virginia University. He has been with NETL, Morgantown, WV, USA, since 2015, where he is currently the Markets/Infrastructure Interdependence Lead.



MARIJA PRICA received the B.S. and M.S. degrees in electrical engineering from the University of Novi Sad, Serbia, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA. With over 20 years of experience in the power sector across industry, academia, and consulting, her expertise includes transmission and distribution planning, operation and protection, renewable energy and energy storage integration, and energy market operations. She recently joined NETL, Pittsburgh, as a General Engineer.