Rapid Evaluation and Response to Impacts on Critical End-Use Loads Following Natural Hazard-Driven Power Outages: A Modular and Responsive Geospatial Technology

Patrick D. Royer1 · Wei Du1 · Kevin Schneider1

Accepted: 12 April 2022 / Published online: 25 May 2022 © The Author(s) 2022

Abstract The disparate nature of data for electric power utilities complicates the emergency recovery and response process. The reduced efficiency of response to natural hazards and disasters can extend the time that electrical service is not available for critical end-use loads, and in extreme events, leave the public without power for extended periods. This article presents a methodology for the development of a semantic data model for power systems and the integration of electrical grid topology, population, and electric distribution line reliability indices into a unified, cloud-based, serverless framework that supports power system operations in response to extreme events. An iterative and pragmatic approach to working with large and disparate datasets of different formats and types resulted in improved application runtime and efficiency, which is important to consider in real time decision-making processes during hurricanes and similar catastrophic events. This technology was developed initially for Puerto Rico, following extreme hurricane and earthquake events in 2017 and 2020, but is applicable to utilities around the world. Given the highly abstract and modular design approach, this technology is equally applicable to any geographic region and similar natural hazard events. In addition to a review of the requirements, development, and deployment of this framework, technical aspects related to application performance and response time are highlighted.

Keywords Electrical grid · Hurricane response · Puerto Rico · Resilience and geospatial technology · Software applications

1 Introduction

Across the world, the magnitude of blackouts is increasing concurrently with the frequency and intensity of natural hazard-induced disasters (Rudin et al. 2014; US Environmental Protection Agency 2016). Tropical cyclones and hurricanes are some of the most destructive natural hazard events, resulting in loss of power that affects millions of people each year. Extreme weather events cause extensive damage to dependent infrastructure systems that can cost billions of dollars to repair and lead to cascading effects on the workforce, healthcare, and economic health in general (Weems et al. 2007; Zimmerman and Restrepo 2009; Mukherjee et al. 2018; NOAA 2020). In 2017, Hurricanes Irma and Maria are notable extreme weather events that occurred within several weeks of each other that impacted Puerto Rico, with the latter causing more extensive damage to critical infrastructure. In January 2020, an earthquake of 6.4 magnitude impacted the area. In addition to the immediate damage to buildings, the loss of life, and the economic impact, much of the critical infrastructure that is necessary for the normal functioning of society has been damaged during these events. This includes the electric power systems operated by the local power authority. Disaster response management was significantly hindered by the lack of an integrated technical platform, drawing from a multitude of relevant data sources, to rapidly assist in the identification and designation of backup power sources.
Outage-forecasting models have become increasingly accurate in predicting hurricane-induced outages prior to landfall, and strong developments have been made recently in the application of machine learning to quantify uncertainty in outage duration (Han et al. 2009; Nateghi et al. 2014; Yang et al. 2020; Zhai et al. 2021). While outage prediction is a useful tool in natural hazard risk management and preparation, there is a need for emergency and response mangers to plan for response measures related to power backup systems and to manage risk on longer time scales. Preparation and rapid response, through the study of resiliency, may be used to mitigate adverse consequences and improve response time to power restoration (Tokgoz et al. 2017). Staid et al. (2014) developed a literature-based sensitivity analysis to simulate the impacts on a power system under 12 plausible climate-hurricane scenarios, providing information at the census tract scale for planning and mitigation strategies. Research by Staid et al. (2014) highlights important links between climate uncertainty and planning and mitigation strategies at local scales for future events. Jeffers et al. (2018) developed an application, the Resilient Node Cluster Analysis Tool (ReNCAT), that performs an analysis of microgrid locations for improving community response to major disruptions across the island and is used for hurricane response planning. ReNCAT is a native (desktop) application that integrates optimization algorithms into its core codebase (Jeffers et al. 2018). ReNCAT has probability functions for critical end-use load outages based on historic flood events, landside probability, and earthquake history. It also calculates a resilience factor, societal burden, which is the metric used for members of the community to meet basic needs.

In addition to outage prediction models and research related to system hardening and hazard risk management and assessment, ongoing and near real-time monitoring and damage assessment that uses remote sensing is playing an increasingly important role in disaster response. Remote sensing technology is rapidly evolving in terms of sophistication and resolution. The way remotely sensed data is accessed, transferred, and processed is improving and becoming exponentially more efficient in terms of applying outcome to critical decision-making tasks. Data obtained from unmanned aerial vehicles (UAV) in concert with both passive and active remote sensing from airborne and spaceborne satellite-based platforms are becoming more prevalent and targeted with respect to ongoing monitoring (Voigt et al. 2007; Ofli et al. 2016; Alam et al. 2020; Vyron and Potsiou 2020). Such approaches enable damage and/or vulnerability detection in a more synoptic manner, and potentially alert authorities and the public to damages before these can be detected in situ. Alam et al. (2020) used UAVs in concert with machine learning to estimate and evaluate utility pole inclination angles. This work focuses on automation of an existing process of monitoring and inspecting the health conditions of utility poles and provides ongoing evaluation of a critical metric within the overall assessment of system resilience. Vyron and Potsiou (2020) used deep learning and remote sensing imagery to identify optimal landing spots for aerial support following extreme events. This work highlights the value of machine learning with respect to the magnitude and scale of satellite data available, and its ability to provide effective and usable results in time-restricted and life-critical applications.

Once a power system is disrupted, timely restoration of electrical power for critical end-use loads can be achieved by using backup systems while the primary system is in the process of restoration. Backup systems encompass the use of one or more technologies, including microgrids, solar, and remote and controlled switches, which can be deployed at the distribution level to increase the resiliency of critical end-use loads (Table 1). These loads include seaports, airports, hospitals, police stations, and drinking water treatment plants, all of which have very distinctive and unique sets of requirements for identifying the most appropriate backup system. Identification and selection of an optimal backup source requires systematic and thorough review of each asset on a case-by-case basis and is tightly coupled with geospatial analytics and location intelligence sciences.

Although modern data media formats, such as geosocial media, can be used as a powerful and effective early alert mechanism to point authorities to the time and location of system failure (Granell and Ostermann 2016), expert judgment is required in order to define the best path to recovery from disruptions. Such an assessment for the sum of assets affected by a catastrophic type event with a large geographic footprint is time consuming and time prohibitive under emergency response scenarios. Automated end-to-end analysis and stepwise logical survey is necessary to inform and guide the recovery process and execute plans in a timely manner.

This article describes a methodology that was developed to identify a range of different backup systems and map them to critical end-use loads by applying broader categorical knowledge for each resource type and utilizing higher resolution spatial data and situational knowledge for each individual asset independently. This work builds on existing research and planning tools and offers a more comprehensive approach to backup systems of different types (Table 1). The capability described here can be accessed through a web-based application with an intuitive user interface for disaster managers and planners that is location aware and mobile friendly. A semantic data model was developed in concert with a data management and data mining solution, which includes three spatially explicit
Microgrids can support one or more critical loads locally. The difference between microgrids and stand-alone backup generation is that microgrids integrate the operation of multiple generation resources, which provides greater reliability, redundancy if a unit should fail, and greater operational flexibility. Microgrids are well-suited for large campus or facility-level deployments. Deploying a microgrid, or network of microgrids, takes extensive engineering analysis, has a high cost, and historically takes 12 or more months. This task will not determine the design of the microgrids, only identify where their deployment should be considered (Van et al. 2012).

### Table 1 Operational options to provide backup power sources

| Type                        | Description                                                                 |
|-----------------------------|-----------------------------------------------------------------------------|
| Remote-controlled switches  | Remote-controlled switches are individual switches at the primary distribution level (for example, 13.2 kV) that can be operated remotely from the distribution control center, or potentially other locations. Switches are designed to isolate equipment and/or to change the topology of the system; they do not have the ability to interrupt fault current. As a result, these devices provide a level of operational flexibility, but they must be coupled with protection devices to effectively reconfigure a system. These are commodity devices that can be installed in a short period of time. |
| Recloser Switches           | Reclosers have the same remote operating capability as an automated switch, but they have the ability to interrupt fault current. Due to the fault interrupting capability, they are able to be an active participant when reconfiguring a distribution circuit. Additionally, reclosers can also operate to protect against temporary and permanent faults. Reclosers typically have the ability to automatically open and close several times to try to clear momentary faults, such as vegetation-related faults. If the fault fails to clear, the recloser remains open to disconnect the faulted portion of the system. This momentary fault mitigation capability is especially useful in rural distribution systems, due to the long distance of distribution lines and their exposure to vegetation. Reclosers often do not provide as significant a mitigation for momentary faults in urban distribution systems, unless there is a corresponding high level of vegetation. These are commodity devices that can be installed in a short period of time. |
| Fault Location, Isolation, and Service Restoration (FLISR) | FLISR refers to a broad range of technologies. The basic FLISR scheme operates to automatically isolate a fault when it occurs, determine where it occurred, and reconfigure the system to restore the greatest number of customers back to service. A FLISR scheme can include the operation of two or more distribution circuits in a predefined teaming arrangement; these can be a stand-alone system or integrated into a distribution management system (DMS). Because of the need for at least one adjacent distribution circuit, these systems are typically not installed in remote rural areas. A two-circuit FLISR, also referred to as an auto-loop, is a two-circuit arrangement with approximately five reclosers that provides automated FLISR capabilities on a set of two circuits. This technology requires advanced engineering analysis and would take several months to a year to install (Akker 2014). |
| Self-Healing System         | A full self-healing system can be viewed as an extension of a simple FLISR scheme, which operates with predefined teaming arrangements. Numerous distribution circuits can be flexibly operated to achieve high-level goals. This capability typically includes the installation of a centralized DMS and the deployment of advanced applications. The deployment of a DMS, or an advanced DMS (ADMS), is a major undertaking that integrates multiple operational systems. Historically, it takes multiple years for a utility to deploy this technology, due to the need for a DMS/ADMS, and as such should be considered in future road maps. |
| Stand-alone Backup Generation | One of the most common methods to increase the resiliency of critical end-use loads is the installation of a stand-alone backup generation source. In these deployments, one or more fossil-fuel generators are connected using a manual or automated bus transfer (MBT/ABT) switch. When there is a loss of utility service, the MBT/ABT can operate to restore power. Traditionally, backup generation is never interconnected with the utility distribution system. Although backup generation is a technically feasible option, the cost of maintenance, operation, and fuel handling make this much more expensive than utility-delivered power. However, for critical end-use loads or remote locations, stand-alone backup generation can be effective. Backup generation is a commodity technology and its installations can be accomplished within 12 months, assuming there are no regulatory or permitting issues. |
| Microgrids                  | Microgrids can support one or more critical loads locally. The difference between microgrids and stand-alone backup generation is that microgrids integrate the operation of multiple generation resources, which provides greater reliability, redundancy if a unit should fail, and greater operational flexibility. Microgrids are well-suited for large campus or facility-level deployments. Deploying a microgrid, or network of microgrids, takes extensive engineering analysis, has a high cost, and historically takes 12 or more months. This task will not determine the design of the microgrids, only identify where their deployment should be considered (Van et al. 2012). |

Equations (Eqs. 1, 2, and 3) and binary decision tree approach that is applied to each critical end-use load to determine optimal backup source(s). The semantic model represents the data layer at the core of this framework and is novel in the way that it integrates information from physical data stores from a variety of disparate and previously unrelated native formats. Data sources are related and integrated into a larger data model to mirror the real-world onset and response to extreme weather events. The data model is then used as the core for a broader software...
development framework that was developed using modern cloud-based, serverless design principles for robustness and increased throughput.

The proposed methodology for identifying optimal backup sources described in this article is applicable to other types of extreme weather threats but is not necessarily applicable to every type of threat (for example, physical cyber-attack threats). As hurricanes involve damage to infrastructure from high wind speeds directly, and falling debris from high winds, flooding, and landslides indirectly, this capability is relevant to other types of natural hazards. The proposed methodology also assumes total failure of critical infrastructure resulting in loss of power to end-use load. It may be useful in situations where damage is incurred but does not result in total failure and/or partial loss of power, yet the method is not explicitly designed to account for such events. The underlying data model can be abstracted to include other data types and is novel in that sense, but the current model assumes these data are available, accessible, and migratable in perpetuity as updates to the system are made, which may not be a realistic expectation.

1.1 Development of a Comprehensive Data Model for Recovery from Complete or Partial Failure of Distribution System

The Geospatial Science Program Management Office maintains that the compilation of critical infrastructure geospatial data is core to mission objectives. To maximize utility for the broader research community, the data should be interrelated in a manner that is inherently semantic and configured in a manner that is agile and fully extensible (Hammer and McLeod 1978; Codd 1979). The approach to data management and stewardship in modern day data science should recognize the need to be able to synergize a multitude of different data types and data standards, which typically include conventional horizontally scalable schema-bound Relational Database Management Systems (RDBMS), schemaless NoSQL (for example Mongo and AWS DynamoDB), and unstructured data types. Utility systems are largely managed as separate entities, spanning both private and public sectors, ranging from city to metropolitan area in scale, and other levels of aggregation, adding additional challenges with data integration. As utility systems are gaining access to more data than ever before, the challenge of digital transformation from legacy systems into a more modern framework, and integration of this framework into instantaneous data-driven solutions, has become more notable (Wong et al. 2009).

Historically, electrical power systems have been operated with a sparse level of data. This is true for both the physical planning data and the near real-time data available to operators (Wood and Wollenberg 1996). When a power system is built, the physical characteristics of the system are recorded. These include, but are not limited to, conductor and cable type, physical arrangements of conductors on poles, and the conductors to which the end-use loads are connected. Prior to the 1980s, this information was traditionally recorded in a paper format, and digitization was rare. Additionally, over time the accuracy of the information would degrade as the physical characteristics of the conductors varied with age, and end-use load connections changed due to activities such as storm restoration. Today most utilities digitize their data in a Geographic Information System (GIS) or a similar system, but the industry continues to grapple with errors in those data that are introduced over time.

Data access for system operators is still relatively sparse and routinely inaccessible. At the bulk transmission system level, it is common for high voltage (115-kV+) transmission lines to have monitoring at both ends; this includes breaker status, current, voltage, and power flow (Wood and Wollenberg 1996). For distribution systems, the data are much more limited. Typically, there is monitoring at the substation and a few points on the circuit where equipment, such as voltage regulators or shunt capacitors, is located; but overall, the system is not considered fully observable (Wood and Wollenberg 1996). While there are additional data sources, such as smart meters, these are typically not connected to operational systems. Connections between customer information systems and distribution management systems (DMS) are an ongoing area of research in the industry. In addition to the data sources that utilities oversee, there are other sources of data that can be beneficial within the spectrum of power reliability and redundancy. These can include weather data, population data, and information on the function of critical loads such as hospitals and police stations. As a result, the operations of electric power systems occur with a limited subset of available data, which often leads to suboptimal decision making. Electrical system management and resiliency plans would benefit immensely from the development of a broader semantic data model between various components of sources of structured and unstructured data. Such a model should surpass logical relationships between independent grid topology features (for example, substations, transmission lines, and towers) and extend to clients, power consumption sources, reliable backup scenarios, climate patterns, cell phone coverage, and other information (Geospatial Science—Program Management Office 2018). Although still in the process of developing a comprehensive global semantic model that includes all these

1 https://www.energy.gov/ig/downloads/special-report-doe-oig-20-58
components, we have endeavored, with the development of a web-based serverless framework, hereafter referred to as the Electrical Grid Resilience and Assessment System (EGRASS), to make a significant step start towards this end. This includes the development of semantic relationships between electrical infrastructure geospatial information, electrical infrastructure reliability, critical end-use loads, and population dynamics.

In Puerto Rico, as well as in other regions of the United States, much of the local and institutional knowledge gathered from observed site conditions and anecdotal information related to system health and system-customer identification is not captured systematically and/or is not easily accessible for analysis. This is the case for the relationship between clients and distribution feeders that is usually captured in a billing mechanism from the utility company. This mechanism may have service level distribution data per customer, but such information is not publicly available and/or is not available in a GIS database. Extensive data modeling was necessary to infill knowledge gaps from conflicting entity relationships within data and between data sets of different types. Figure 1 represents the modeled spatial and referential relationship between critical end-use loads and distribution circuits. Modeled relationships provide input for optimization of backup selection. This includes an estimate of population affected by a nearby critical resource, the number of substations and related distribution feeders linked to critical resources, and the number of nearby resources in the same category.

In a real-world scenario, a natural hazard event triggers rapid response that includes prioritization of critical end-use loads for which power should be restored first (Fig. 1). The sequence in which reestablishment of power critical to end-use loads occurs depends on ranking these resources by population affected, the number of nearby substations and similar resources, reliability indices, and an estimate of level of effort required to restore critical infrastructure. Based on this information, judgement rules can be processed to isolate optimal backup power sources for each unique end-use load scenario. Then coordinated rapid response can be executed. Integrated data results from physical data sources are highlighted by an example hospital site in Puerto Rico shown in Fig. 1 that is associated with a distribution feeder based on spatial proximity, a population estimate for each household in the surrounding area, and substation(s) that are associated with distribution feeders using Relational Database Management System (RDBMS) referential integrity to maintain logical relationships between tables. The hospital is shown in the center of the map related to feeder 1421-03 with symbols for substations and location of all other nearby end-use loads (Fig. 1). The inset map shows labeled distribution feeders, population estimate centroids for each building footprint, and other nearby critical end-use loads.

A data model was designed in consideration of the structural interrelationship between the “real world” and physical data stores to capture the meaning of this information in the application environment. In this work, physical data stores were available natively in varying format, some of which were transformed to a target format based on the ongoing nature of data ingestion. Physical data stores included structured, semi-structured, and unstructured data, captured and managed in the Structured Query Language (SQL) tabular and geospatial database (DB), noSQL DB, and unstructured DB storage respectively (Fig. 2). Format and management of these data types was entirely based on data availability and the foreseeable ongoing transfer of these data. Structured tabular data included grid topology network—the relationship between substations, distribution lines, circuits, and busses.Structured spatial relationships are the relationship between end-use loads and distribution circuits (Figs. 1 and 2). Semi-structured data include population, reliability indices, and user profiles. Unstructured data include chosen backup power sources, anecdotal information and site data that are not captured in any systematic format, and structure material and post hurricane restoration work. In Puerto Rico this information is known in a general sense by region but is not documented in a reproducible manner that can be reconciled with grid topology.

The data sources displayed in Fig. 2 are used collectively to inform judgement rules that are used in the identification of backup sources for each critical end-use load. Grid topology is managed in a conventional structured database with primary keys, foreign keys, and a schema that is commonly found in an entity relationship (ER) model. Semi-structured data, such as population data, reliability indices, and damage reports, are still managed in a database, but the database does not enforce a schema, which allows for flexibility in data types. Most semi-structured databases, such as MongoDB and DynamoDB are managed in keypair format. Unstructured data require no format and can come in the form of reports, images, spreadsheets, and other formats (Fig. 2b).

The real-world response to the impact of extreme weather events, and the catastrophes those extreme events cause, is based on the principle that the most important use for recovery resources should be to reestablish quickly those institutions and infrastructure elements that provide essential services to the largest number of people. The semantic data model involves identification of the most populated regions and collocated critical end-use loads (structured spatial data) coupled with grid topology to identify nearby circuits and referential substations (structured data) and understanding of site conditions to identify
Fig. 1 Schematic depicting application workflow and semantic model logic for extreme weather event and response (upper). Map depicting spatial modeling relating feeders, substations, and total population estimate to critical end-use loads (lower).
optimal backup power sources (semi-structured and unstructured data). The sequence in which data form stores are used is mapped to the flow of the real world in a semantic model (Figs. 1 and 2).

1.2 Decision Tree Logic and Relevant Equations for Identifying Optimal Backup Sources for Critical End-Use Loads

The underlying logic for identifying optimal backup resources for critical end-use loads categorically and individually relies on data access and accessibility patterns driven by the development of a data model described in the previous section. The semantic data model, and the encapsulation of this model in a flexible model framework, was necessary in order to execute logic programatically against multitiered, disparate data sources with varying storage and access mechanisms.

Provided here is a list of variables, spatially explicit equations, and a sample of underlying logic used in algorithms for identifying optimal backup source(s) for end-use load category and individual assets within each category described in Table 1. Binary tree logic involves nested calculations that account for the overall reliability and integrity of a system, demand, accessibility to site, existing redundancy, and current orientation and location of existing grid infrastructure. Assessed in EGRASS are six different critical end-use loads with 1–6 possible different backup sources for each asset, varying across categories and individual assets within each category. A sample of this logic (Fig. 3) and the variables used (Table 2) are provided in the section. The logic in its entirety is encompassed in the code base for the application and database layers of the application. The code repository is not publicly hosted, but access to the code for research purposes is granted; this includes the entire body of nested query statements highlighted in Fig. 3.
Population density, accessibility by vehicle, and substation count, were calculated as follows:

\[
\text{POP} = \sum_{i=1}^{n} P(d)_i = P(d)_1 + P(d)_2 + \ldots + P(d)_n  \quad (1)
\]

\[
\text{ACCESS} = \sum_{i=1}^{n} R(d)_i = R(d)_1 + R(d)_2 + \ldots + R(d)_n  \quad (2)
\]

\[
\text{SBCOUNT} = \sum_{i=1}^{n} S(d)_i = S(d)_1 + S(d)_2 + \ldots + S(d)_n  \quad (3)
\]

where \(d\) is the radius (Euclidean search distance) for the circular area around each critical end-use load type (shown in Fig. 1 for an example hospital), \(POP\) is the total population in the vicinity and includes household count \(P\) for every residence within the specified search distance (Fig. 1). The substation count \(SBCOUNT\) is calculated as the total number of substations \(S\) within search distance (example shown in Fig. 1). \(ACCESS\) is a measure of accessibility calculated by the total linear length of roads \(R\) intersecting

Fig. 3 Decision tree logic used for backup type Reclosure Switches, applies to all critical end-use load types. ACCESS Accessibility by vehicle; SAIDI System average interruption duration index; SAIFI System Average interruption frequency index
the search area. In addition to accessibility measured by linear roadway count, terrain ruggedness was calculated from a 30 m digital elevation model (DEM) using equations developed by Riley et al. (1999). In this calculation, a terrain roughness index (TRI) is assigned by comparing variation in adjacent elevation values assigned to pixels surrounding each asset. This provides a relative level of effort and potential difficulty in terms of accessing the site following an event, which is important to consider in Puerto Rico given the intermountain area in the center of the island and the propensity for mudslide in this area. Population density was normalized and categorically binned as densely, moderately, or sparsely populated for each end-use load based on distribution of population counts for all sites. Accessibility by road and terrain roughness is normalized using linear scale normalization and used collectively to determine ease of access following a potential event.

Standardized reliability indices were used to evaluate the integrity of the electrical distribution system collocated with critical end-use loads. The System Average Interruption Duration Index (SAIDI) is commonly used as a reliability indicator by electric power utilities (Heydt and Graf 2010). The SAIDI is the average outage duration for each customer served. The System Average Interruption Frequency Index (SAIFI) is commonly used as a reliability

Table 2  A description of variables used in assessment

| Variable | Variable type | Description | Significance |
|----------|---------------|-------------|--------------|
| Population density (POP) | Categorical | Categorized as sparse, moderate, or dense | Quantifies demand |
| Accessibility by vehicle (ACCESS) | Distance in meters, normalized over all measurements | Quantification of total linear roadway | Accessibility to site for repairs following disruption, and for implementation of backup solutions for strengthening prior to disruption |
| Terrain roughness index (TRI) | Elevation in meters, normalized over range of data | Terrain Roughness Index | Used in combination with the ACCESS variable to quantify ease of access following a disruption |
| Substation count (SBCOUT) | Count (integer) | Total number of collocated substations | Characterizes resiliency in terms of redundancy |
| Similar asset count | Count (integer) | The total number of end-use loads of the same type (for example, the number of other hospitals for end-use load hospital type) | Characterizes the redundancy of the same service type |
| Distribution line placement and orientation | Categorical | Overhead or underground | Used in concert with reliability indices, and pertains directly to the logistics of implementing backup solutions |
| Distance to feeder | Distance in meters | Distance between end-use load and vertex of the nearest feeder | Characterizes resiliency and redundancy |
| Number of feeders | Count (integer) | Total number of collocated feeders | Applicable to the use of FLISR and Reclosure Switches, which require >= 2 feeders |
| System Average Interruption Duration Index (SAIDI) | Average time in milliseconds (decimal) | The System Average Interruption Duration Index during a year | System reliability index |
| System Average Interruption Frequency Index (SAIFI) | Average count (integer) | System Average Interruption Frequency Index during a year | System reliability index |
| Customer Average Interruption Duration Index (CAIDI) | Average time in milliseconds (integer) | The customer average interruption duration index | System reliability index |
| Hospital trauma level | Categorical (1–5) | Refers to the kinds of resources available in a trauma center and the number of patients admitted in a year. Applies only to hospitals | Quantifies importance of specific hospitals |
| Backup generator | Boolean (True/False) | Whether or not there is already a backup generator at the facility. This information is not available at every site | Characterize system redundancy and suitability for microgrids |
indicator by electric power utilities. The SAIFI is the average number of interruptions that a customer would experience. The Customer Average Interruption Duration Index (CAIDI) is a reliability index commonly used by electric power utilities. It is related to SAIDI and SAIFI. The CAIDI gives the average outage duration that any given customer would experience. It can also be viewed as the average restoration time.

Figure 3 conveys the use of variables (Table 2) in the decision tree for backup system type Reclosure Switches, defined in Table 1. The solution shown here is unique in that for Reclosure Switches the same logic applies to all critical end-use load types, which is not the case for other backup power types. The use of microgrids, as an example, applies differently to hospitals than it does for other types of end-use loads.

2 A Computational Framework for Emergency Response and Disaster Preparedness

Although data modeling was the initial focus of this article, the overarching objective was to develop a robust framework for rapid assessment, disaster response, and decision-making tools. The data model is an important first tier of this framework. The remainder of this article describes the development of the application and the techniques used to automate this logic and the development of a tool for intuitive execution of application functionality as well as the outcome on a select subset of end-use loads.

A serverless philosophy was adopted early on in moving this work to the cloud for a multitude of reasons, but several virtual machines were used due to serverless memory constraints, operating system level dependencies, and lack of a stable container (that is, Docker) solution. An approach to Application Programming Interface (API) execution was developed involving downsampling of larger data sets and invoking geospatial indices to optimize efficiency and throughput speed. Also discussed are implications for moving computing workloads from on premise to the cloud.

2.1 Hybrid Approach to Cloud Architecture Using Virtual Machines and Serverless Framework

The Electrical Grid Resilience and Assessment System (EGRASS) is a fully customized, single-page application (SPA) with progressive web application (PWA) functionality developed as part of a broader effort to develop a risk-based framework based on outage definitions with associated probabilities of occurrence from hurricane events. The client component of EGRASS was developed using React API, a lightweight, highly abstract, and modular framework.

The backend component was built primarily using a serverless centric framework with Amazon Web Services (AWS) technology layered on top of a multitude of data and data storage combinations (Fig. 2). Extensive effort was taken to streamline and simplify the user interface while also optimizing architecture and configuration for the backend for maximum efficiency and optimal response time. Data access and visualization available in EGRASS surpass the capacity of commercial off-the-shelf programs relevant to domain specific analytic objectives, which are useful when little is known about large data sets and exploration goals are vague (Keim 2002; Yi et al. 2007; Aigner et al. 2008).

The architecture is primarily serverless with multiple copies of a virtual machine (VM) running a docker container that hosts a PostgreSQL database with PostGIS spatial extension and Geoserver open-source geospatial data server. The VM is set up through a load balancer for distributing resources. The application was built within a virtual private cloud (VPC), with data archived in a PostgreSQL server, Dynamo databases, and S3 unstructured data. Authentication is handled using AWS Cognito, REpresentational State Transfer (REST) functionality is developed using AWS Lambda functions and exposed via API Gateway. API Gateway is the central system for creating RESTful API for interacting with the client. Lambda is the scalable, serverless computing service that runs code in response to events and automatically manages the underlying computing resources. Underlying data resources include Dynamo semi-structured/noSQL data, PostgreSQL relational database system, and S3 unstructured data sources. A spatially enabled RDBMS, in conjunction with NoSQL, was used to leverage aspects of the application that are storage driven, normalized, relational, and good for online analytical processing.

Both R-tree spatial indexing and geohashing were used to return geospatial queries as quickly and efficiently as possible (Zhang and Yi 2010). More complex spatial functions in the application are routed through one of several different approaches, depending on both the inherent structure of the data, and computational performance of the calculations. These approaches include: (1) stored database functions written in PL/pgSQL (Procedural Language/PostgreSQL) for postGIS; (2) noSQL spatial functions and operands written for Dynamo DB that use geoJSON format; and (3) python spatial libraries, including Shapely, Fiona, and Rasterio, integrated directly into the AWS Lambda functions.

2.2 Software Interaction Pattern, Behavior, and User Journey

The EGRASS software was designed with two primary modules: (1) identification and isolation of candidate
technology deployments to increase the resiliency of end-use loads that are critical to the normal operation of society; and (2) risk-based dynamic contingency analysis based on hurricane simulation. Both were designed to inform investment decisions in Puerto Rico’s transmission grid with identification of judgement rules. Electrical engineering judgement rules are migrated to a decision tree matrix and codified in program logic. The risk-based framework is structured on outage definitions with associated probabilities of occurrence from hurricane events, in combination with impact assessment derived from detailed dynamic cascading analysis. The interaction pattern and workflow for both use modes are described in this section, and detailed output for the first use mode as well as identification of candidate technology are explained in subsequent sections.

The EGRASS application has a user pool consisting of both power users and nonpower users, such as site managers and regulatory program managers. To accommodate a multitiered audience, the API was directly exposed and available for those in the user pool with programmatic knowledge and underlying subject matter expertise. The EGRASS user interface for the nonpower users was designed and optimized for speed and ease of use, and required little to no training. For nonpower users accessing it via the interface, higher level information is readily accessible, and a much greater level of detail is available by scoping in further. In candidate technology mode, the user selects an asset category of interest, and by following the population of assets in that category, the user selects a specific asset—for example, a hospital, a police station, or a drinking water treatment plant. At this stage, the zoom and extent are established based on the asset selected and five pieces of relevant information are returned in user panels: (1) a list of candidate technologies for that specific asset; (2) the population at risk; (3) the number of adjacent assets in the same category; (4) the number of adjacent substations; and (5) the reliability index of the longest transmission line. Items 2 –5 are all spatially queried within a predefined Euclidean search distance from the asset selected (Figs. 1 and 5). At this point in the interaction pattern, the user can narrow the search distance to zoom in on the extent of interest. All geospatial and tabular responses from the backend are updated in perpetuity. Reliability indices for each distribution line related to collocated substations can be individually assessed by clicking on the distribution line of interest (Fig. 4).

In the risk-based dynamic contingency mode, the user begins by selecting a historical hurricane track, assets of interest, and (optionally) a time resample interval and windspeed threshold. The hurricane simulation is run, all assets within the projected hurricane swath footprint are depicted by windspeed at each timestep, and a tabular output is provided in the lower panel (Fig. 5). The user can then cycle through each time step and assets are highlighted.

2.3 A Response Time-Optimized Approach to Execution of Underlying Geospatial Queries—Estimating Population at Risk and Intersection of Transmission Lines and Spatially Collocated Infrastructure

Quantification of population at risk and characterization of population statistics has emerged as a major challenge in areas where census data are sparse, difficult to retrieve, and/or nonexistent, and there are limited sources of streaming geospatial intelligence (GEOINT) data. As it relates to electrical power supply in Puerto Rico, population serviced and population at risk of loss of service was perhaps the single-most important piece of information for the overall utility of the application. This is due in part to the time involved in recovering and restoring power to critical assets, and also to the extent of population affected by loss of power. Developing a geospatial population data set at appropriate spatial resolution and developing a mechanism to execute geospatial queries against these data in a responsive manner were key metrics in the usability and value of this work as applied research related to disaster response. Furthermore, development of efficient and sustainable approaches to analyzing large data sets, geospatial and otherwise, is critical to every modern-day approach to algorithm design and underlying architecture.

Lacking well vetted data at the census block level in the area, population counts by municipality district were used. Residential homes were identified by using Microsoft building footprints. Population values were distributed among residential footprints proportionally by area in each respective municipality (Fig. 6). The centroid from each housing unit was used in subsequent queries and in the API for improved response times and to develop efficient queries.

A ceiling API response time of 1000-milliseconds over a range of network speeds was targeted and designed, and the API was structured based on this criterion, using dynamic population downsampling to query against data averaged over larger areas as search extent increased. There are over 1.5 million buildings in Puerto Rico, and obtaining target response required an iterative and comprehensive approach. Building footprint centroids (Fig. 7a) were dynamically resampled to coarser resolution as a function of search distance from point of interest: 500 m, 1000 m,
and ≥ 1800 m (Figs. 7b, c, and d respectively). Total population counts derived from downsampled population centroids were evaluated for accuracy by comparing these counts to raw counts from original building footprint estimates, which included all data. Distances further than 1800 m from centroid were not considered relevant in terms of response time as a search distance of approximately 1 km is most relevant with respect to critical end-use loads and electrical distribution circuits that can provide power to these end-use loads. The potential error was evaluated against the benefit of increased API response times at each 100 m interval. Mean raw population values from 100 locations were compared to the same locations with resampled data at each 100 m interval. Mean values from population estimates were compared with a paired t-test.

Similar geospatial queries included estimating the total number of collocated substations, similar asset types (for example, hospitals), and transmission lines within the search distance, as depicted by the intersecting search distance (Figs. 6c, 7). This information was also assessed geospatially in the context of the known attributes of the critical end-use loads and the historical transmission line reliability indices.

3 Optimal Back Source Selection by Resource Type by Planning Unit and for Selected Sites

The results are summarized for the methodology in three different formats: (1) a table representing the proportion of optimal backup systems for each end-use resource type for
a synoptic assessment of Puerto Rico backup power options (Table 2); (2) a geographic summary by planning unit for hospitals as an example of resource type by region (Fig. 8); and (3) four separate individual site selections to highlight the expected use of the interface (Fig. 9). Planning units are geographic entities established by the Puerto Rico Electric Power Authority (PREPA), which are used to associate critical infrastructure with specific regions of Puerto Rico. Importantly, PREPA was dissolved in 2021 and replaced by LUMA, which at the moment uses the previous entity’s administrative procedures without major change.

In Fig. 8 the pie charts are proportionally sized by the total number of hospitals in each planning unit and colors categorically reflect the proportion of each backup type for each planning unit respectively. San Juan planning unit, having the most hospitals, is shown in a separate map with street map as reference. Hospitals are used as an example resource type in this data summary, but similar results could be depicted for all other resource types.

3.1 Case Studies, Adaptation to Critical End-Use Loads in Puerto Rico

Several example scenarios are described here and shown in the context of spatial reference (Fig. 9). For simplification, and to protect sensitive geolocation data, site names are anonymized, and background map context is not shown. The text highlights the associated decision matrix for various backup sources; the logic to execute this selection is integrated into the API. The examples below represent a small subsample of end-use loads.

Scenario 1: Airport—The airport in Fig. 9a is served by three substations. Because this is a major airport that is already served by three substations, there are three...
Fig. 6 Geographic and census components involved in the development of population estimates. Island of Puerto Rico and municipal boundaries with population counts for each municipality (a). Building footprints for residential buildings only, eliminating nonresidential footprints (b). Illustration of population estimate function of search distance, as rendered in EGRASS application (c).

Fig. 7 Stepwise illustration of population downsampling based on search distance from end-use resource load of interest. Images in sequential order (a–d) depict decreasing sample resolution as a function of search distance. Subplot a depicts the centroids at native resolution for building footprints. Subplot b depicts coarser sampling of footprints within a radius of 500 m. Subplot c depicts coarser sampling of footprints within a radius of 1000 m. Subplot d depicts coarser sampling of footprints greater than, or within a radius of, 1600 m.
candidate technologies that could be considered. First, automated switches and/or reclosers could be used to provide redundant paths to the airport. Second, due to the large number of distribution circuits in a nearby urban area, a self-healing scheme that integrates the operation on multiple circuits could be considered. This could be an option unless the airport is supplied by dedicated distribution circuits. If it is supplied by dedicated circuits, a self-healing scheme would not be as useful; this information was not available at the time this article was prepared. The third option would be the installation of backup generation, if it does not already exist, and/or creation of a microgrid if there is sufficient on-site generation.

Scenario 2: Hospital—The hospital in Fig. 9b is the largest hospital in a densely populated area, with a number of collocated hospitals, and serves the highest number of inpatients and outpatients in the area and has
the highest bed count. Microgrids are recommended based on the importance of this hospital and the presence of backup generators in the general area. Hospitals are typically equipped with backup generation and fuel for 72 h, per National Fire Protection Association Standard 110.4 “Standard for Emergency and Standby Power Systems.” Because there are more than a half-dozen

---

**Table 3** Recommended power backup by critical end-use load category

| Load Category                  | Remote controlled switches (%) | Reclosures (%) | Fault location, isolation, and service restoration (FLISR) (%) | Self-healing systems (%) | Backup generation (%) | Microgrid (%) |
|-------------------------------|--------------------------------|----------------|---------------------------------------------------------------|--------------------------|-----------------------|---------------|
| Airports (10)                 | 53.6                           | 11.1           | 22.2                                                          | 2.0                      | 0.0                   | 11.1          |
| Hospitals (65)                | 41.6                           | 0.3            | 33.3                                                          | 0.0                      | 12.2                  | 12.5          |
| Law Enforcement (187)         | 64.9                           | 1.7            | 18.4                                                          | 1.4                      | 0.0                   | 13.5          |
| Water Treatment Plants (130)  | 35.8                           | 11.1           | 0.8                                                           | 0.0                      | 27.1                  | 25.3          |
| Sea Ports (53)                | 44.2                           | 1.7            | 13.3                                                          | 2.5                      | 25.8                  | 12.5          |

---

Fig. 9 Spatial context for the EGRASS judgement rule algorithm, highlighting cases of critical end-use loads: an airport (a), a hospital (b), a police station (c), and a water treatment plant (d). The names of each end-use load example are anonymized, as is the location, by not displaying background maps. Distribution lines (shown in purple) are shown along with building footprints and critical end-use loads.
substations in the vicinity connecting to different feeders, it is practical to switch the reclosures so the hospital can connect to other feeders. Because of the time involved in manual operation and FLISR (fault location, isolation, and service restoration), remote-controlled switches are necessary to reestablish power as quickly as possible.

Scenario 3: Police Station—The police station in Fig. 9c is supplied by a single circuit. Even though adjacent circuits exist nearby, backup generation may not be practical, so the only option is to improve the reliability of the existing circuit if it is not adequate. This could be accomplished with a combination of automated switches and reclosers.

Scenario 4: Water Treatment Plant—The water treatment plant in Fig. 9d has two backup generators and a solar panel, which makes it a good candidate for microgrids. Because the plant is far from densely populated areas, and it would take several hours to repair, FLISR in concert with remote-controlled switches is recommended.

Importantly, presented here are only four critical end-use loads that were characterized out of a much larger number of possible scenarios. The underlying data in Puerto Rico against which the algorithms were developed includes 1.5 million residential housing footprints, 41,248 electric transmission towers, 6212 transmission lines, 383 substations, 187 law enforcement agencies, 13 electric power plants, 65 hospitals, and a total population of over 3.5 million people. Each end-use load has the possibility of utilizing one or many of the six different types of backup systems for rapid recovery following a hurricane event.

3.2 Evaluation of Population Downsampling as a Function of Increased Application Efficiency

Population at risk as associated with residential footprints was the most computationally expensive aspect of the API. In resampling population data, average response time was reduced to 700 milliseconds, as opposed to 200,000 milliseconds required for raw spatial data (Fig. 10a). Response times for resampled data did increase significantly beyond...
a search distance of 10,000 m (up to 11,000 milliseconds), but search distances up to 1600 m (1 mile) had little relevance to the localized end-use load being assessed.

Related to response times, population counts were not significantly different within an 1800-m search distance (p = 0.12). Population counts were higher at approximately 900 m and lower between 900 and 1800 m.

4 Discussion

This article highlights a methodology for the fusion of multiple data sources into a single framework that supports power system operations in response to extreme events, and details the structure and potential of this methodology in a tool called EGRASS. The approach developed here adds to a growing body of research and tools that are more directly related to emergency response and recovery and not to power outage prediction per se (Staid et al. 2014; Jeffers et al. 2018). More specifically, the capability described here integrates a larger body of information to identify optimal backup sources for critical end-use loads in each major category. This technology can be characterized as a location intelligence-driven application designed with a web-based single page progressive web application (PWA) to help support rapid response to catastrophic events resulting in interruption of power to critical end-use loads. This capability is applicable to other categorical natural hazard events and systems and is abstract in the sense that it can be refactored to include other data types varying in form and structure.

Summary results of the methodology that were developed for this framework provide ideal planning resources for critical end-use loads by category and by region. Results in a synoptic view can help guide investments for preparation as a whole and point to specific regions and local areas that may require greater investment based on additional insight that may not be available in the public domain. Used in the web application user-interface with mobile phone or by workstation, this tool provides a convenient way to survey areas and extract details by specific sites. The geolocation-enabled mobile version may help with field surveys and yield critical information during site visits.

A great deal of work has been done on developing a resilience assessment framework for electric power systems. This includes development of probabilistic models for assessing transmission responses following hurricane events (Ouyang and Dueñas-Osorio 2014; Mensah and Dueñas-Osorio 2016), quantitative methodology that measures resilience costs that result from a disruption to infrastructure function (Vugrin et al. 2011), and the assessment of overall vulnerability and resilience of coastal regions by integrating natural and human data layers for mapping and visualization (Lam et al. 2015). The application described in this article is unique in the approach used to merge population data with grid topology, reliability indices, and other geographic information. The EGRASS software is less focused on model development and more focused on building the API to facilitate access to the underlying logic of powerful models with a simple progressive web application interface.

We emphasize the importance of continuing with the development of a nationwide semantic data model with all pertinent data, which requires focused and ongoing cooperation from a variety of entities. Although much remains to be done, EGRASS begins this endeavor with a focus on building entity relationships between infrastructure, critical end-use loads, population distribution and needs, and reliability indices. It enables an interactive and highly responsive approach for workforce and regulatory entities to respond to power outages with relevant information about backup sources.

4.1 Leveraging Cloud Resources and Cloud Dependencies

Many companies, universities, and other institutions have migrated, or are beginning the process of migrating, to the cloud. In early evolutions of cloud technologies on-premise projects were moved to a similar suite of virtual machines that resided on a cloud, and this represented a more streamlined one-to-one remapping of technology. Significant changes in this paradigm occurred along with the concept of platform as a service (PaaS), and serverless technology. The impact being that “lift and shift” became much more involved for moving on-premise based applications to the cloud. Consequently, more development teams are likely to use the cloud concepts in the beginning phases of development and continue down this path moving into other phases of the software development life cycle. This approach has greater inherent risk based on working knowledge that cloud environments from larger and more well-known companies such as Amazon Web Services (AWS), Microsoft AZURE, and Google Cloud may not hybridize well. Furthermore, projects that embrace a particular cloud technology over the other cannot necessarily be abstracted to another platform. One approach to overcome this is to containerize and develop more sophisticated container environments to meet the demands of more complex applications. Kubernetes is an open-source container orchestrator for managing containers with a dedicated and growing number of users and advocates. Although not entirely cloud agnostic, it is far more realistic to set up and manage across multiple or mixed cloud environments. At this point in development EGRASS is
very much tied to the AWS platform; but future iterations will be more focused on a containerized deployment.

4.2 In Consideration of Day/Night Population Flux

Population counts in this study were obtained at the spatial extent of municipal boundaries and distributed across residential building footprints. Essentially this choice provides a nighttime population (Fig. 4). Given the importance of daily energy usage curves and the variance in demand between residential and nonresidential buildings, it is important to consider how daytime population, in addition, would offer significant advantages to the existing tool. Such a dataset does exist in raster format (Dobson et al. 2000), which was developed using reflectance patterns from spectral signatures and segmentation. Upon evaluation of these data in Puerto Rico, we found commission omission differences from actual buildings we considered to be significant and felt building a footprint dataset was a superior source for identifying population source locations and quantifying population. We have developed similar products that have accounted for day/night flux in the continental United States. In these studies, we incorporated employment data from Census Longitudinal Employment-Household Dynamics dataset, from which we developed daytime population from place of origin (residence) to place of destination (workplace). Unfortunately, such data are not available in Puerto Rico, and we suggest that a follow up study should consider a method for replicating such a dataset in Puerto Rico.

5 Conclusion

This article highlights a methodology for coalescing large and disparate data sources into a central framework that supports power system operations in response to extreme events, and details instance of this methodology in a tool called EGRASS. Timely restoration of electrical power for critical end-use loads is crucial and can be achieved, at least partially, by using backup systems when the primary system is in the process of restoration. Backup systems vary widely in terms of expense and practicality of implementation, and are influenced by factors such as demand, accessibility, redundancy, and reliability. The range of possibilities requires expert judgement, in addition to access to relevant data, and typically requires more time and effort than is affordable when responding to catastrophic events.

The approach developed here adds to a growing body of research and tools that are more directly related to emergency response and preparedness. We acknowledge the importance of the existing body of research focusing on this area, including outage prediction models and resiliency and adaption methodologies. We also recognize the value and growing utility of using machine learning in concert with remotely sensed imagery to inform critical, time sensitive decision-making processes when working with big data. The work presented here pertains more specifically to the nature and technical challenge of working with data that vary widely in terms of native format, and the value that thoughtful, intentional, and pragmatic semantic data modeling provides when designing an application framework. This article also highlights the importance of streamlined design principles and implementing serverless cloud architecture best practices, which is far more sustainable and adaptable in the long term, and in step with modern practices.

Acknowledgments This work was supported by the United States Department of Energy, Office of Energy Efficiency and Renewable Energy, Solar Energy Technology Program. The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by the Battelle Memorial Institute under Contract DE-AC05-76RL01830.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aigner, W., S. Miksch, W. Müller, H. Schumann, and C. Tominski. 2008. Visual methods for analyzing time-oriented data. *IEEE Transactions on Visualization and Computer Graphics* 14(1): 47–60.

Akker, J.D. 2014. Fault location, isolation, and service restoration technologies reduce outage impact and duration. Addis Ababa, Ethiopia: Ministry of Health, Federal Democratic Republic of Ethiopia.

Alam, M.M., Z. Zhu, E.B. Tokgoz, J. Zhang, and S. Hwang. 2020. Automatic assessment and prediction of the resilience of utility poles using unmanned aerial vehicles and computer vision techniques. *International Journal of Disaster Risk Science* 11(1): 119–132.

Codd, E.F. 1979. Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems* 4(4): 397–434.

Dobson, J., E. Bright, P. Coleman, R. Durfee, and B. Worley. 2000. LandScan: A global population database for estimating

---

5 https://lehd.ces.census.gov/data/
populations at risk. *Photogrammetric Engineering and Remote Sensing* 66(7): 849–857.

Granell, C., and F.O. Ostermann. 2016. Beyond data collection: Objectives and methods of research using VGI and geo-social media for disaster management. *Computers, Environment and Urban Systems* 59: 231–243.

Hammer, M., and D. McLeod. 1978. The semantic data model: A modeling mechanism for data base applications. In *Proceedings of the 1978 ACM SIGMOD International Conference on Management of Data*, 31 May 1978–2 June 1978, Austin, Texas, USA, 26–36.

Han, S., D. Guikema, and S.M. Quiring. 2009. Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. *Risk Analysis* 29(10): 1443–1453.

Heydt, G., and T.I. Graf. 2010. Distribution system reliability evaluation using enhanced samples in a Monte Carlo approach. *IEEE Transactions on Power Systems* 25(4): 2006–2008.

Jeffers, R.F., M.J. Baca, A.M. Wachtel, S. Derosa, A. Staid, W. Fogleman, A. Outkin, and F. Currie. 2018. Analysis of microgrid locations benefiting community resilience for Puerto Rico. Technical report. Albuquerque, New Mexico: Sandia National Laboratory.

Keim, D.A. 2002. Visual exploration of large data sets. *Communications of the ACM* 45(1): 1–8.

Lam, N.S.-N., Y. Qiang, H. Arenas, P. Brito, and K.B. Liu. 2015. Mapping and assessing coastal resilience in the Caribbean region. *Cartography and Geographic Information Science* 42(4): 315–322.

Mensah, A.F., and L. Dueñas-Osorio. 2016. Efficient resilience assessment framework for electric power systems affected by hurricane events. *Journal of Structural Engineering*. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001423.

Mukherjee, S., R. Nateghi, and M. Hastak. 2018. A multi-hazard approach to assess severe weather-induced major power outage risks in the U.S. *Reliability Engineering and System Safety* 175(C): 283–305.

Nateghi, R., S. Guikema, and S.M. Quiring. 2014. Power outage estimation for tropical cyclones: Improved accuracy with simpler models. *Risk Analysis* 34(6): 1069–1078.

NOAA (National Oceanic and Atmospheric Administration). 2020. U.S. billion-dollar weather and climate disasters, 1980–present. Asheville, NC: NOAA National Centers for Environmental Information.

Ofli, F., P. Meier, M. Imran, C. Castillo, D. Tuia, N. Rey, J. Briant, and P. Millet et al. 2016. Combing human computing and machine learning to make sense of Big (Aerial) Data for disaster response. *Big Data* 4(1): 47–59.

Ouyang, M., and L. Dueñas-Osorio. 2014. Multi-dimensional hurricane resilience assessment of electric power systems. *Structural Safety* 48: 15–24.

Riley, S.J., S.D. DeGloria, and R. Elliot. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences* 5(1–4): 23–27.

Rudin, C., S. Ertekin, R. Passonneau, A. Radeva, A. Tomar, B. Xie, S. Lewis, and M. Riddle et al. 2014. Analytics for power grid distribution reliability in New York City. *Inform Journal on Applied Analysis* 44(4): 364–383.

Staid, A., S.D. Guikema, R. Nateghi, S.M. Quiring, and M.Z. Gao. 2014. Simulation of tropical cyclone impacts to the U.S. power system under climate change scenarios. *Climatic Change* 127(3): 535–546.

Tokgoz, B., M. Safa, and S. Hwang. 2017. Resilience assessment for power distribution systems. *International Journal of Civil and Environmental Engineering* 11(7): 806–811.

U.S. Environmental Protection Agency. 2016. Climate change indicators: Weather and climate. https://www.epa.gov/climate-indicators/weather-climate. Accessed 2 Feb 2020.

Van, S.B., B.N. Judson, S.V.T. Nguyen, and W.D. Ross. 2012. Microgrid study: Energy security for DoD installations. Technical report 1164. Cambridge, MA: Lincoln Laboratory, Massachusetts Institute of Technology.

Voigt, V., T. Kemper, T. Riedlinger, R. Kieß, K. Scholte, and H. Mehl. 2007. Satellite image analysis for disaster and crisis-management support. *IEEE Transactions on Geoscience and Remote Sensing* 45(6): 1520–1528.

Vugrin, E., D. Warren, and M.A. Ehlen. 2011. A resilience assessment framework for infrastructure and economic systems: Quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane. *Process Safety Progress* 30(3): 280–290.

Vyron, A., and C. Potsiou. 2020. A deep learning method to accelerate the disaster response process. *Remote Sensing* 12(3): Article 544.

Weems, C.F., S.E. Watts, M.A. Marsee, L.K. Taylor, N.M. Costa, M.F. Cannon, V.G. Carrion, and A. Pina. 2007. The psychosocial impact of Hurricane Katrina: Contextual differences in psychological symptoms, social support, and discrimination. *Behaviour Research and Therapy* 45(10): 2295–2306.

Wood, A.J., and B.F. Wollenberg. 1996. *Power generation, operation, and control*. 2nd edn. New York: John Wiley & Sons.

Wong, P.C., K. Schneider, P. MacKey, H. Foote, G. Chin, R. Gutthromson, and J. Thomas. 2009. A novel visualization technique for electric power grid analytics. *IEEE Transactions on Visualization and Computer Graphics* 15(3): 410–423.

Yang, F., D.W. Wanik, D. Cerrai, M. Bhuiyan, and E.N. Anagnostou. 2020. Quantifying uncertainty in machine learning-based power outage prediction model training: A tool for sustainable storm restoration. *Sustainability* 12(4): Article 1525.

Yi, J., Y. Kang, J.T. Stasko, and J.A. Jacko. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics* 13(6): 1224–1231.

Zhai, C., T. Chen, A. White, and S. Guikema. 2021. Power outage prediction model training: A tool for sustainable storm restoration. *Reliability Engineering & System Safety* 208: Article 107348.

Zhang, L., and J. Yi. 2010. Management methods of spatial data based on PostGIS. In *Proceedings of 2010 2nd Pacific-Asia Conference on Circuits, Communications and System*, 1–2 August 2010, Beijing, China, 410–413.

Zimmerman, R., and C.E. Restrepo. 2009. Analyzing cascading effects within infrastructure sectors for consequence reduction. *Proceedings of the 2009 IEEE Conference on Technologies for Homeland Security*, 11–12 May 2009, Waltham, Massachusetts, United States, 157–162.