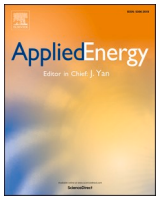
Applied Energy 397 (2025) 126360

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Establishing nationwide power system vulnerability index across US counties using interpretable machine learning

Junwei Ma [a,*], Bo Li [a], Olufemi A. Omitaomu [b], Ali Mostafavi [a]

[a] Urban Resilience.AI Lab, Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, United States
[b] Oak Ridge National Laboratory, Oak Ridge, TN, United States

HIGHLIGHTS

- A nationwide Power System Vulnerability Index (PSVI) has been developed for 3022 U.S. counties over the past decade.
- The PSVI reveals a steady rise in power system vulnerability across the U.S. from 2014 to 2023.
- Regions including the West Coast, East Coast, Gulf of Mexico, and the Great Lakes megalopolis demonstrate notably high levels of vulnerability.
- The influence of urban form and strcuture on power system vulnerability has been recognized as a key factor.

ARTICLE INFO

Keywords:
Power outages
Infrastructure vulnerability
Disaster resilience
Energy justice
Machine learning

ABSTRACT

Power outages have become increasingly frequent, intense, and prolonged in the US due to climate change, aging electrical grids, and rising energy demand. However, largely due to the absence of granular spatiotemporal outage data, we lack data-driven evidence and analytics-based metrics to quantify power system vulnerability. This limitation has hindered the ability to effectively evaluate and address vulnerability to power outages in US communities. Here, we collected ~179 million power outage records at 15-min intervals across 3022 US contiguous counties (96.15 % of the area) from 2014 to 2023. We developed a power system vulnerability assessment framework based on three dimensions (intensity, frequency, and duration) and applied interpretable machine learning models (XGBoost and SHAP) to compute Power System Vulnerability Index (PSVI) at the county level. Our analysis reveals a consistent increase in power system vulnerability across the US counties over the past decade. We identified 318 counties across 45 states as hotspots for high power system vulnerability, particularly in the West Coast (California and Washington), the East Coast (Florida and the Northeast area), the Great Lakes megalopolis (Chicago-Detroit metropolitan areas), and the Gulf of Mexico (Texas). Our heterogeneity analysis indicates that urban counties and those located along regional transmission boundaries tend to exhibit significantly higher vulnerability. Our results highlight the significance of the proposed PSVI for evaluating the vulnerability of communities to power outages. The findings underscore the widespread and pervasive impact of power outages across the country and offer crucial insights to support infrastructure operators, policymakers, and emergency managers in formulating policies and programs aimed at enhancing the resilience of the US power infrastructure.

1. Introduction

Electric power systems serve as critical lifelines that underpin modern societies and enable the functioning of nearly every aspect of contemporary existence [1]. However, with the increasing global climate change, various extreme natural disasters, such as hurricanes and heat waves, are threatening the resilience of power systems [2–4]. In the US, between 2018 and 2020, Hurricane Florence, Michael, Laura, Sally, and Delta collectively caused severe outages that affected 0.6 to 4.3 million customers at their peak [1]. Similarly, the 2021 Winter Storm Uri caused widespread power outages that impacted 25 states and over 150 million Americans [5]. In addition to natural disasters, power outages also result from various incidents, such as electrical component failures, supply shortages, physical attacks, vandalism, cyberattacks,

* Corresponding author.
E-mail address: jwma@tamu.edu (J. Ma).

https://doi.org/10.1016/j.apenergy.2025.126360
Received 18 March 2025; Received in revised form 4 May 2025; Accepted 22 June 2025
Available online 30 June 2025

and wildlife interference [6]. The increasingly frequent, intense, and prolonged power outages are disrupting transportation, communications, water supply, and healthcare systems, thus seriously undermining public well-being [7–9]. As a result, effectively evaluating and addressing vulnerability of power systems in communities has become an urgent priority.

Power system vulnerability refers to the susceptibility of a power system to potential harm, affecting the extent to which community members and their assets are exposed to power outages. Recent studies have increasingly emphasized the significance of assessing power system vulnerability to mitigate social suffering and formulate policies for promoting power infrastructure resilience, particularly by examining the extent of power outages. For example, Flores, McBrien et al. (2023) analyzed the social vulnerability during the 2021 Winter Storm Uri by examining power outage distribution, duration, and sociodemographic disparities related to these outages [5]. Feng, Ouyang et al. (2022) explored the compound risk of tropical cyclone- and heatwave-induced power outages in Harris County, Texas, to examine how the risk evolves with changing climate and propose strategies to enhance the resilience of coastal power systems [3]. Sayarshad and Ghorbanloo (2023) evaluated the resilience of power line outages caused by wildfires in Sonoma County, US, aiming to help utilities design more resilient power lines in wildfire-prone areas [10]. However, these case-based analyses focus on the short-term impacts of isolated extreme weather-induced events on power systems. With the rising frequency and intensity of such disruptions caused by global climate change [11], it has become crucial to assess long-term and large-scale patterns and consequences. In addition to extreme weather events, various daily power outages stem from electrical component failures, supply shortages, physical attacks, vandalism, cyberattacks, and wildlife interference [6]. These frequent but localized outages also significantly affect human life but are often overlooked in research.

Moreover, prior studies limited the geographic scope to specific regions in the US. For example, Dugan, Byles et al. (2023) developed an index to quantify social vulnerability to prolonged power outages using census tracts in Colorado as a case study [12]; Ganz, Duan et al. (2023) analyzed power outage data from eight major Atlantic hurricanes between 2017 and 2020 to assess the impact of hurricanes on nine southeastern US states [13]. Flores, Northrop et al. (2024) collected outage data from non-New York City urban and rural areas to evaluate the lagged effect of severe weather on power outages [14]. Given the economic and social disparities across US communities, a nationwide assessment of power outages is essential to investigate the heterogeneity of power system vulnerability across various geospatial contexts (e.g., urban vs. rural, power system operators, and regions with varying energy structures).

A substantial body of prior work has examined various statistical models to outage data for resilience and vulnerability analysis. For example, Dikshit, Saransh et al. (2024) and Dai, Yitian et al. (2023) systematically studied cascading failure risks across key components of the power grid (e.g., towers, transmission lines) through mathematical modeling and system simulations [15,16]. Their work emphasizes probabilistic characterizations of cascading failures and the identification of critical thresholds that signal systemic vulnerability. Ahmad, Arslan et al. (2025) leveraged statistical tools (e.g., logistic regression) to model the reliability and recoverability of regional power systems under extreme event scenarios, using the outage data from North American Electric Reliability Corporation (NERC) [17]. These studies have laid important groundwork but primarily focus on single system layers or specific failure modes, limiting their scope in scale and spatiotemporal coverage. Advances in AI now offer the opportunity to move beyond these constraints, enabling the systematic identification of broader and more complex vulnerability patterns across both transmission and distribution systems over long time spans with high spatiotemporal resolution. A data-driven characterization of power system vulnerability hinges on examining historical power outage patterns at scale. However, a major obstacle in conducting such analysis has been the lack of publicly available power outage data at a large scale with proper spatiotemporal resolution. Hanna and Marqusee (2022) highlight that the absence of extensive datasets impedes the examination of complex interactions between long-duration outages and system vulnerability [18]. In an era of increasingly severe outages, despite the lack of such outage data, there is an urgent need for a large-scale, high-resolution, and nationwide assessment of power system vulnerability to inform mitigation plans and policies.

Another gap associated with data availability and granularity issues is the lack of reliable and generalizable metrics for assessing power system vulnerability. Previous studies have proposed various metrics to measure outage characteristics. For example, some researchers quantified outage extent by considering the period during which customers without power exceed a specific threshold [2,14]. Flores, McBrien et al. (2023) introduced the concept of “power-out person-time”, an interaction between outage duration and the number of affected customers [5]. Some studies rely on metrics from the electricity engineering field to measure outage impacts, such as SAIDI (System Average Interruption Duration Index) and SAIFI (System Average Interruption Frequency Index) [19–21]. However, the extent of power outages cannot be solely captured by a single metric. The integration of a combined set of metrics and their interaction are essential to properly quantify and evaluate the extent of vulnerability in the power systems of a region. The characterization and quantification of vulnerability has been done in the context of different socio-technical systems (such as the CDC/ATSDR social vulnerability index [22] and socio-economic-infrastructure vulnerability index [23]) and has shown to be very effective in informing plans and policies. Yet, such data-driven integrated index is direly missing for power system vulnerability.

Recognizing these gaps, this study aims to construct a comprehensive system of metrics based on granular historical power outage data and use it in creating a machine learning-based index that captures the full extent of vulnerability to power outages. To achieve this, we first retrieved ~179 million power outage records at 15-min intervals from the Environment for Analysis of Geo-Located Energy Information (EAGLE-I™) platform operated by Oak Ridge National Laboratory (ORNL) [24]. This large-scale and high-resolution power outage data covers 3022 counties (96.15 % of the US continental areas) between November 2014 and December 2023. By applying a 0.1 % threshold to screen out the non-valid outage records, we identified a total of 3022,915 power outage events. Drawing inspiration from environmental hazards exposure models [25–27], we developed a systematic framework to assess power system vulnerability based on three dimensions: frequency, duration, and intensity. For each dimension, multiple features were created to enable spatial and temporal analysis of outage trends across the US. Using these features, we trained and validated interpretable machine learning models, specifically eXtreme Gradient Boosting (XGBoost) combined with SHapley Additive exPlanations (SHAP), to examine the non-linear relationship among the features in distinguishing counties in terms of their power system vulnerability, and accordingly to reveal the relative importance of the features. Subsequently, a PSVI system with power system vulnerability values, scores, and ratings was computed through the multiplication of the features by their corresponding weights. Based on the computed PSVI for each county and across the decade, we further characterized the following spatiotemporal patterns: (1) spatial hotspots of power system vulnerability; (2) temporal trends and areas with growing extent of power system vulnerability; (3) variations across urban vs. rural areas and the effect of form and structure characteristics on the extent of power system vulnerability; (4) variations of vulnerability across regional transmission organizations; and (5) association between renewable energy sources and power system vulnerability. The results depict an alarming picture of how ubiquitous and widespread power outages have been across the US counties and offer valuable support of data, metrics, and methodology for infrastructure operators,

policymakers, and community leaders to guide the development of policies and programs aimed at strengthening the resilience of the US power infrastructure.

2. Methodology

This study follows the processing procedure shown in Fig. 1. First, we developed 14 power system vulnerability features from three dimensions (frequency, intensity, and duration), capturing multifaceted outage characteristics. Next, we applied interpretable machine learning models (XGBoost and SHAP) to determine the relative importance of features. Based on the feature importance, we assigned corresponding weights to features. Then, the PSVI (value, score, and rating) was calculated as a weighted sum of features. Finally, we conducted disparity analysis of power system vulnerability regarding factors including urban/rural form and structure, regional electricity distribution, and electricity generation by source.

2.1. Power outage data

This study utilized a large-scale and high-resolution power outage dataset to calculate power outage-related features across US counties. Power outage data was collected through EAGLE-I™ operated by ORNL [24]. EAGLE-I™ compiled electricity service disruption records from individual electrical utilities at a 15-min interval from 2014 through 2023. On average the dataset covers ~90 % of utility customers nationwide, making it the most comprehensive outage information ever compiled in the US [24]. Considering that data for Alaska, Hawaii, Puerto Rico, Guam, the US Virgin Island, and American Samoa are incomplete in certain years, we limited the geographical range of our study within the contiguous United States. The data covers 3022 counties, accounting for 96.15 % of the area.

2.2. Power outage features development

We processed the outage records following the procedure described in our previous study [28] to derive power outage features. The ORNL provided the estimated number of total customers at the county level for certain temporal and geographical range, and we utilized linear extrapolation to extend the data to cover all the 3022 counties from 2014 through 2023. The metric of power outage rate was then calculated as the proportion of customers without power compared to the

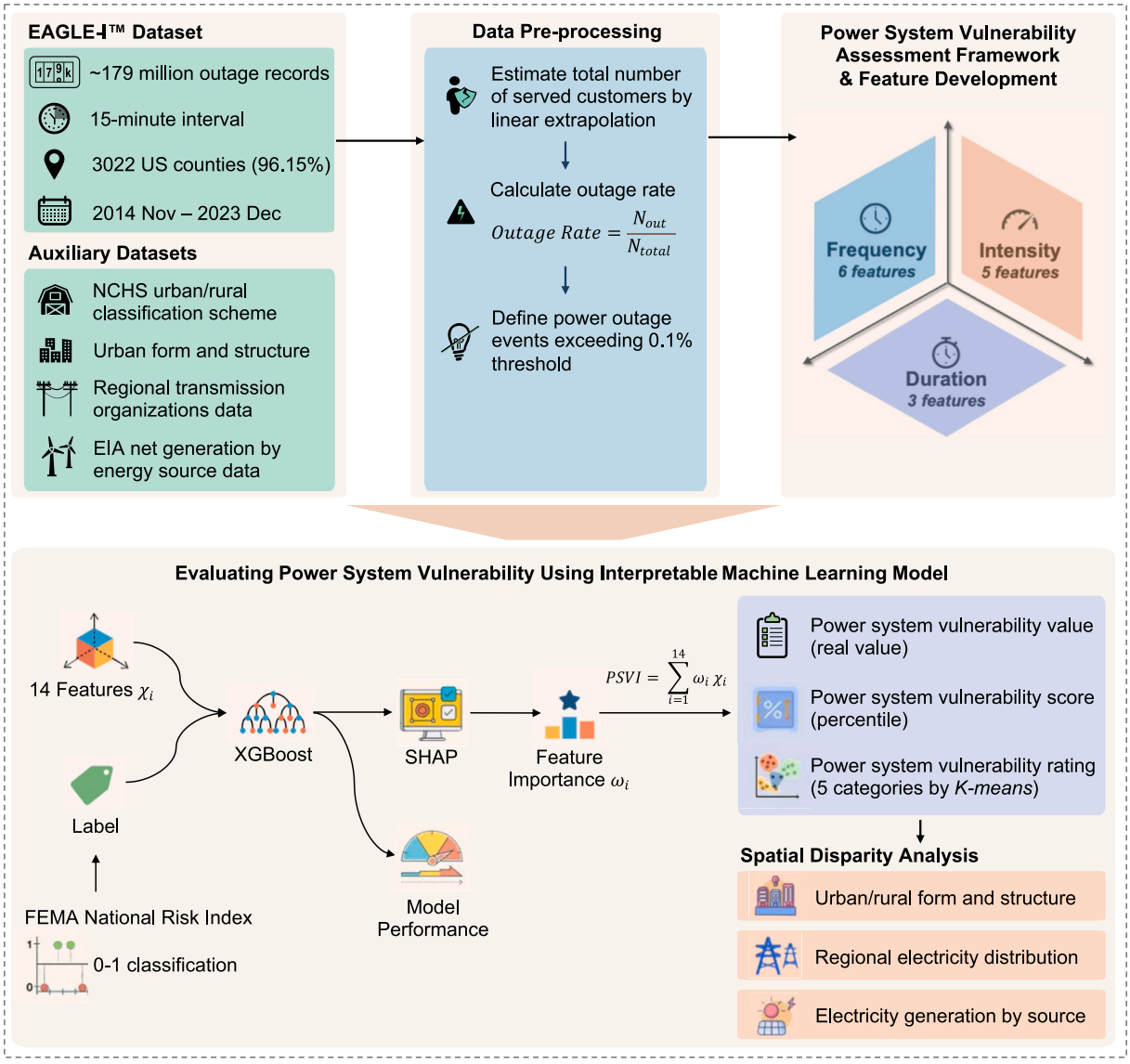


Fig. 1. Workflow for establishing power system vulnerability index across US counties using large-scale power outage data and interpretable machine learning models.

total number of customers in a county (Eq. 1). We defined power outage events as the continuous time period during which power outage rate exceeds 0.1 %. Setting the 0.1 % threshold helps to screen out the non-valid outage records due to incidental factors, which is a practice applied by research such as Do, McBrien et al. (2023) [2]. Over the decade, a total of 3022,915 power outage events was identified.

$$Power\ Outage\ Rate = \frac{N_{out}}{N_{total}} \quad (1)$$

where N_{out} refers to the number of customers experiencing outages; N_{total} represents the total number of customers in a county.

Based on the identified power outage events, we proposed a systematic and comprehensive power system vulnerability assessment framework, which quantified outages through three dimensions: frequency, duration and intensity. Under each dimension, multiple features were developed to capture the characteristics of power outage events (Fig. 2). To further account for the impacts of large-scale outage events, which affect daily life more significantly than frequent but localized events, we set thresholds based on average outage rates (5 %) and average duration (12h) to define features for large-scale outage events. All the features were calculated at the county level. For the decade PSVI, the features were aggregated and calculated over the ten years, while for the annual PSVI, the features were calculated on a yearly basis.

- Frequency

This dimension includes six features: (1) number of events—counted as the total number of power outage events; (2) average inter-event time—calculated as the average time interval between consecutive power outage events; (3) number of events affecting >5 % customers—counted as the total number of large-scale power outage events which affected more than 5 % of the served customers; (4) average inter-event time affecting >5 % customers—measured as the average time interval between large-scale power outage events which affected more than 5 % of the served customers; (5) number of events exceeding 12 h—counted as the total number of large-scale power outage events which lasted more than 12 h; (6) average inter-event time exceeding 12 h—measured as the average inter-event time between large-scale power outage events whose duration exceeded 12 h.

- Intensity

This dimension incorporates five features: (1) average outage rate—calculated as the average power outage rate among all power outage events; (2) cumulative number of customers affected—counted as the total number of customers affected in all power outage events; (3) peak number of customers affected—counted as the maximum number of customers affected in a single outage event. This feature captures the historical peak outage intensity of the counties; (4) average increase/decrease rate. To calculate this feature, we first aggregated raw outage records to a monthly level and computed the percentage changes between consecutive months. The average increase/decrease rate was then calculated as the average of percentage changes, reflecting how drastically the outage records were changing over time; (5) average outage rate exceeding 12 h—calculated as the average power outage rate among large-scale power outage events whose duration exceeded 12 h.

- Duration

This dimension contains three features: (1) average duration—calculated as the average time period of power outage events; (2) average duration per customer experienced—calculate by dividing the total outage duration by the total number of customers. This is a normalized outage duration feature regarding the scale of served customers; (3) average duration affecting >5 % customers—calculated as the average duration of large-scale power outage events that affected more than 5 % of the served customers.

2.3. Power system vulnerability index construction

To create a synthesized power system vulnerability metric, we utilized the interpretable XGBoost - SHAP model. XGBoost is a powerful decision tree-based algorithm that improves model performance by iteratively generating new trees from the initial weak learners [29]. SHAP is an interpretable machine learning model that quantifies the contribution of each feature to individual predictions [30]. In this study, we used the National Risk Index (NRI) from The Federal Emergency Management Agency (FEMA) [31] as a benchmark to indicate power outage impacts, since the indicator is a comprehensive assessment of overall risk for US counties by integrating multiple hazards and vulnerability factors into a singular metric. We framed the problem as a binary supervised classification task, where the 14 features served as explanatory variables, and the NRI obtained from FEMA served as the

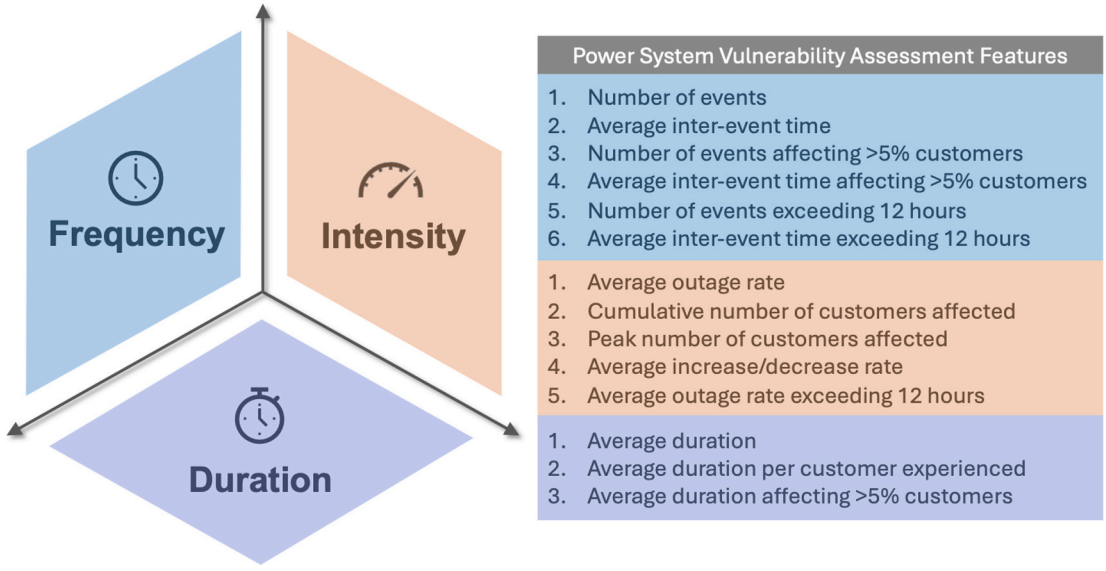


Fig. 2. Conceptual framework for assessing power system vulnerability. We proposed a power system vulnerability assessment framework that captures three dimensions of power outages: frequency (6 features), intensity (5 features), and duration (3 features).

dependent variable. By training the XGBoost model to predict the NRI based on our features, we effectively gauged the extent to which power system vulnerability features explain variations in the county-level risk. Accordingly, we used the relative importance of the features retrieved from SHAP model as weights for computing the PSVI. This approach to determining the weights of features instead of relying upon subjective weights provides a more reliable estimation of the weights in calculating the PSVI [23,32].

Before performing the XGBoost model, we preprocessed the data through the following pipelines: (1) data labelling. The NRI classifies all counties into five risk categories: very low, relatively low, relatively moderate, relatively high, and very high. To address the potential class imbalance, where the "very low" category comprised a significantly larger number of counties ($n = 1419$) compared to the "very high" ($n = 15$), "relatively high" ($n = 128$), "relatively moderate" ($n = 391$), and "relatively low" ($n = 1069$) categories, we framed the classification as a binary problem. Counties in the "very low" category were labeled as 0, indicating minimal vulnerability, while counties in all other categories were labeled as 1, indicating higher vulnerability levels. This binary setup helped mitigate imbalance issues, ensuring more balanced sample sizes and enhancing model robustness; (2) feature normalization. We rescaled the power system vulnerability features to [0,1] using min-max normalization for consistency; (3) multicollinearity check. Features with high multicollinearity share similar information about the target variable, which could cause redundancy and complicate interpretation [33]. Thus, we calculated the Variance Inflation Factor (VIF) to diagnose multicollinearity issues. VIF greater than 10 indicates high multicollinearity in the dataset [34]. In this study, the VIF values of the power system vulnerability features are all smaller than 10, indicating that multicollinearity is not a significant issue (Supplementary Table 1).

For XGBoost model training and validation, we split the data into 80/20 ratio, with 80 % of data used for training and the remaining 20 % for testing. To improve the model performance, we applied Synthetic Minority Oversampling Technique (SMOTE) to mitigate the impact of category imbalance by over-sampling the minority category and under-sampling the majority category [35]. Also, we applied random search and performed a 10-fold cross-validation to tune the hyperparameters. The tuned best hyperparameters for the XGBoost model is listed in Supplementary Table 2. We also compared the nine widely used classification models, such as random forest, support vector machine, and AdaBoost, and found that XGBoost achieved the best F1-score (Supplementary Table 3). Hence, we selected XGBoost as our primary model for binary classification.

To interpret how the power system vulnerability features contribute to the impacts of power outages, we adopted SHAP model. SHAP value of each feature denotes both the magnitude and direction of contributions towards the machine learning output [30]. In this study, SHAP values were calculated as a measure of feature importance and rescaled to [0, 1]. We used Eq. 2 to convert SHAP values into feature weights to make sure the sum of the weights equals 1.

$$w_i = \frac{SHAP_i}{\sum_{i=1}^{14} SHAP_i} \tag{2}$$

where $SHAP_i$ is the SHAP value of the feature x_i, and w_i is the weight of the feature x_i.

Finally, we calculated the power system vulnerability values at the county level using Eq. 3.

$$PSVI = \sum_{i=1}^{14} w_i x_i \tag{3}$$

where x_i represents the power system vulnerability feature; w_i denotes the weight of feature x_i.

The PSVI value provides an absolute measure of power system vulnerability for each county. To enable comparison across counties, we converted these values into percentiles, referred to as the power system vulnerability scores. In addition, we established a five-category qualitative rating system (minor, moderate, major, severe, and extreme) using the K-means clustering algorithm [36]. The rating system allows for a more intuitive understanding of vulnerability levels, making it easier for stakeholders to interpret the severity of power system vulnerability in different counties.

2.4. Auxiliary data for disparity analysis

- Urban/rural form and structure

We categorized the 3022 counties as either urban or rural according to the 2013 National Center for Health Statistics (NCHS) Urban-Rural classification scheme [37]. A six-level urban-rural classification scheme for US counties and county-equivalent entities is developed by NCHS. We labeled 1776 counties as urban since they fall into the three most urban categories: large central metropolitan, large fringe metropolitan, and medium metropolitan. A total of 1246 counties were labeled as rural since they are in the three least urban categories: small metropolitan, micropolitan, and noncore (Supplementary Fig. 6). Urban/rural form and structure refer to the spatial configuration and organization of regions. We examined eight form and structure features based on our previous study [38], including population density, point of interest (POI) density, road density, minority segregation, income segregation, urban centrality index, GDP, and human mobility index. To reduce complexity while preserving the essential information of the features, principal component analysis (PCA) was performed, and three principal components were extracted. The first principal component is named as development density (DD, including population density, POI density, and road density), representing the level of urbanization and built environment density in certain area. The second one is centrality & segregation (CS, including urban centrality index, minority segregation, and income segregation), representing the social and economic segregation as well as the urban centralization level. The third component is economic activity (EA, including GDP and human mobility index), representing the level of economic activity and mobility. The detailed description of the features and PCA is available in Supplementary Note.

- Regional electricity distribution

We categorized counties based on the spatial distribution of regional electricity provision to assess the impact of different transmission network coverage scenarios on power system vulnerability. The US has seven Regional Transmission Organizations (RTOs) that consolidate high-voltage transmission assets to enhance efficiency across a large network [39]. These RTOs includes California ISO (CAISO), Southwest Power Pool (SPP), Electric Reliability Council of Texas (ERCOT), Midcontinent ISO (MISO), New York ISO (NYISO), New England ISO (ISO-NE) and PJM. As cases are possible that counties belong to multiple RTOs, we labeled these as boundary counties. The spatial coverage of RTOs is shown in Supplementary Fig. 7.

- Electricity generation by source

The US Energy Information Administration (EIA) provides yearly state-level energy generation data and the share of total for each energy source [40]. Among the energy sources, we cast special attention on the wind and solar energy, as they are the most widely applied renewable electricity sources and have been reported to have impacts on the stability of power systems [41,42]. We collected relative shares of state-level electricity generation by solar and wind from 2014 through 2023 and calculated the ten-year average as the percentage of solar and wind among all kinds of energy sources.

3. Results

3.1. Power system vulnerability modeling by XGBoost and SHAP

From November 2014 to December 2023, we collected a total of 179,053,397 power outage records at 15-min intervals across 3022 US counties. During this period, the accumulative user outage time reached 7.86 billion user-hours, affecting approximately 31.47 billion customers (Supplementary Fig. 1–3). These figures underscore the widespread disruption to customer service and the vulnerability of US power infrastructure. Notably, the coastal areas and Great Lakes megalopolis experienced more severe outages. Such significant geographic variation in power outage exposure highlights the dire need for integrated nationwide metrics and quantitative examinations of spatial disparities for power system vulnerability.

To systematically capture the characteristics of power outage extent, we developed a power system vulnerability assessment framework. This framework involves 14 key features across three dimensions: frequency, intensity, and duration. During the study period, US counties experienced an average of 1002.3 power outage events, with an average outage rate of 1.5 % to total customers. The average annual outage duration was 7.3 days, meaning that each county experienced outages for approximately 2.0 % of the year. The average interval between power outage events is 7.1 days, indicating that the counties experienced a power outage event approximately every week. Over the past decade, power outage events cumulatively affected 540,915 customers in each county on average. The average peak ratio of affected customers to total customers was 52.8 %, and the monthly fluctuation in the number of power outage events reached 501.5 %. These features highlight the widespread and substantial power outages in the US, which cause a significant impact on daily life and economic activities. The statistical summary of these features is available in Table 1, and the spatial distribution maps over the ten years are provided in Supplementary Fig. 4.

We then constructed the PSVI as the weighted sum of the power system vulnerability features. We weighted the features according to their relative importance in contributing to the impact of power outages.

Table 1
Statistical summary of the power system vulnerability features. The summary represents the county-level averages over the study period (2014–2023).

Features	Unit	Mean	Max	Min	Median
Number of events	/	1002.3	5772.0	1.0	891.0
Average outage rate	%	1.5	100.0	0.1	1.0
Average duration	day	7.3	81.6	0.6	6.6
Average inter-event time	day	7.1	273.5	0.0	3.4
Cumulative number of customers affected	/	540,915	32,050,993	2	191,557
Peak number of customers affected	%	52.8	100.0	0.1	49.6
Average increase/decrease rate	%	501.5	10,121.5	−17.0	254.7
Average duration per customer experienced	hour	0.4	52.6	0.0	0.1
Number of events affecting >5 % customers	/	42.0	1591.0	0.0	28.0
Average duration affecting >5 % customers	day	0.4	15.9	0.0	0.23
Average inter-event time affecting >5 % customers	day	228.3	10,075.7	0.0	119.9
Number of events exceeding 12 h	/	29.2	394.0	0.0	19.0
Average outage rate exceeding 12 h	%	2.2	18.3	0.0	1.7
Average inter-event time exceeding 12 h	day	136.8	3062.1	0.0	88.7

In other words, the more a feature contributes to the impacts, the greater weight should be assigned to it, as it significantly shapes the power systems' vulnerability. By training the XGBoost model to predict the NRI using our power system vulnerability features, we effectively assessed how well these features explain variations in county-level risk. The XGBoost model exhibited strong out-of-sample performance, with the F1 score of 0.7937, accuracy of 0.7835, precision of 0.8051, recall of 0.7826, and AUC-ROC of 0.8955 (Fig. 3f). We also evaluated eight additional machine learning models, and the XGBoost outperformed all of them (Supplementary Table 3).

To obtain the relative importance of each feature, we used SHAP model. A SHAP value greater than 0 indicates a positive contribution of a feature to the mean prediction of the dependent variable and vice versa [30]. Fig. 3g illustrates the relationship and relative importance of the 14 features driving power system vulnerability. Notably, three inter-event time features contribute negatively to power system vulnerability: average inter-event time (−0.048), average inter-event time affecting >5 % customers (−0.021), and average inter-event time exceeding 12 h (−0.001). It suggests that longer intervals between power outage events reduce both their frequency and overall vulnerability. The other features show positive contributions, with the cumulative number of customers affected having the highest importance (+0.516), followed by the average duration per customer experienced (+0.289) and average outage rate (+0.062). In addition, we grouped the 14 features into the three dimensions and calculated their cumulative importance (Fig. 3g, pie chart). The intensity dimension contributed the most (52.7 %), followed by duration (30.1 %) and frequency (17.2 %). The intensity-related features, such as average outage rate, cumulative number of customers affected, peak number of customers affected, average increase/decrease rate, and average outage rate exceeding 12 h, played a most significant role in determining power system vulnerability. The frequency dimension ranked lowest, largely due to the inclusion of the three negatively contributing features. Overall, these 14 features, distributed across the three dimensions, provide a comprehensive characteristics of power outage events and form the basis of our PSVI.

3.2. Decadal distribution of power system vulnerability

Using the feature importance results derived from the XGBoost and SHAP model, we developed the PSVI system which comprises three components: value, score, and rating. The value represents the actual power system vulnerability. The score ranks each county on a scale from 0 to 100 based on its percentile relative to all other counties. The rating is a qualitative classification, using K-means clustering algorithm to categorize counties into five vulnerability levels: minor, moderate, major, severe, and extreme.

For the power system vulnerability values, the probability density curve exhibits a prominent peak near the lower values, with a long tail extending to the right (Fig. 3c). This pattern reveals that most values are clustered around the lower end, while a few higher values stretch the distribution. The peak indicates that the majority of the power system vulnerability values fall within the range of [5,10] (Min: 1, Max: 100), suggesting that most counties experienced relatively low power system vulnerability. For values greater than 30, the distribution tail thins significantly but extends to 100, indicating the presence of some counties with extremely high power system vulnerability values.

The second index is the power system vulnerability scores. The distinct spatial patterns of the scores emerge in its spatial map (Fig. 3a). The US West Coast (particularly California), the East Coast (including the Northeast megalopolis and Florida), the Gulf of Mexico (mainly Texas), and the Great Lakes megalopolis exhibit the highest vulnerability scores, indicating severe power outage impacts in these areas. In addition, high scores are present in many central regions, such as Colorado, Minnesota, and Wyoming, suggesting that power system vulnerability is widespread across the US counties.

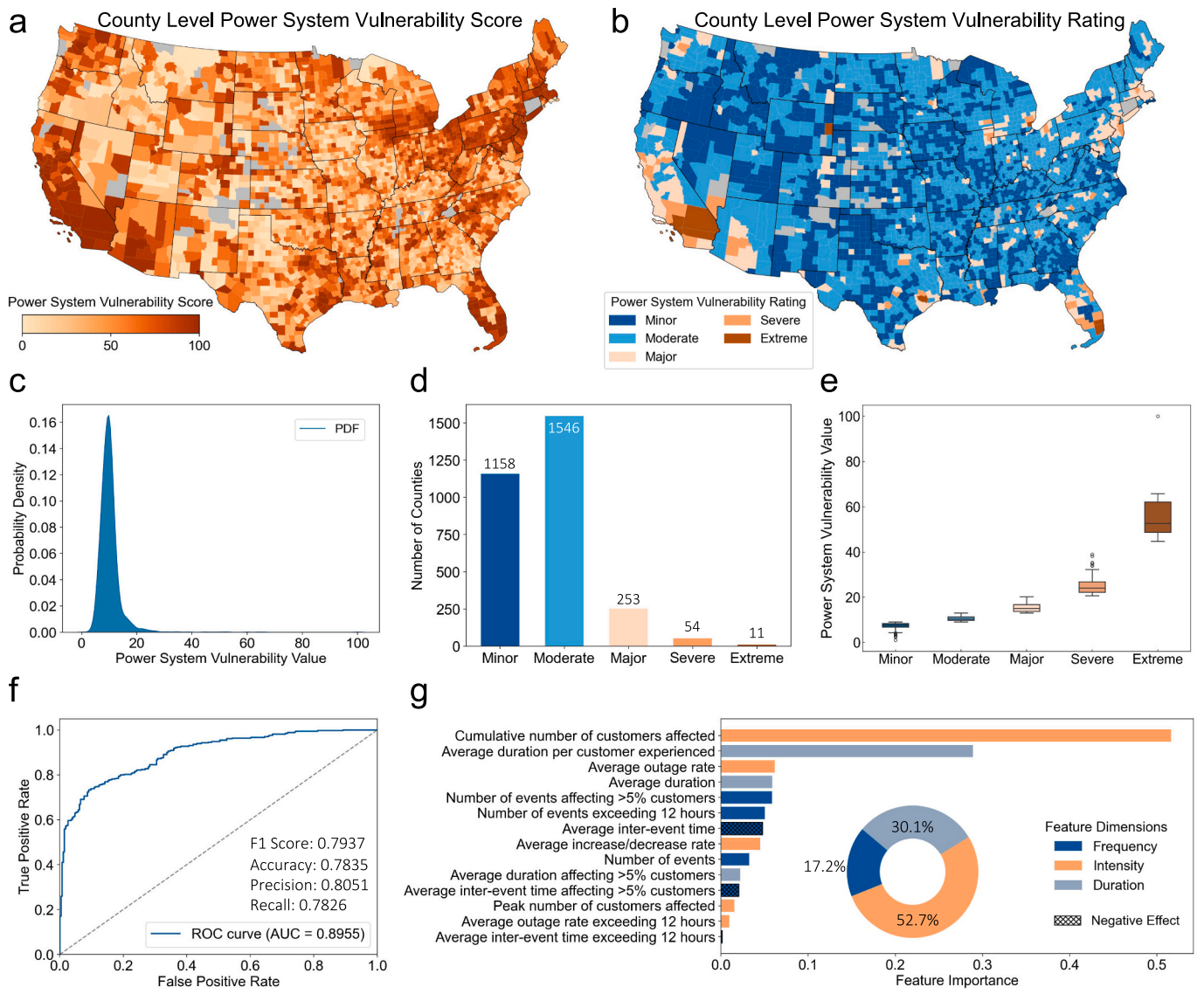


Fig. 3. Descriptive statistics of the decadal power system vulnerability index. a. Spatial distribution of the power system vulnerability scores at the county level; **b.** Spatial distribution of the power system vulnerability ratings at the county level; **c.** Probability density curve of the power system vulnerability values; **d.** Bar plot of the number of counties across the five categories of the power system vulnerability ratings; **e.** Boxplot of the five categories of the power system vulnerability ratings. The Kruskal-Wallis H test confirmed the significant differences among the categories ($p < 0.001$); **f.** AUC-ROC curve and performance indicators of the XGBoost model. **g.** SHAP importance distribution for the 14 features and 3 dimensions. Maps cover 3022 counties, with gray areas representing counties without data.

The third index involves clustering the 3022 counties into five power system vulnerability ratings. The K-means clustering algorithm effectively distinguished these five categories, where the silhouette score reached 0.55 (Fig. 3e). A total of 2704 counties (89.48 % of all counties in this study) was classified as having minor or moderate levels, indicating that the majority of counties exhibit relatively low power system vulnerability (Fig. 3d). Counties classified as having major, severe, and extreme vulnerability levels totaled 253 (8.37 %), 54 (1.79 %), and 11 (0.36 %), respectively. These counties are primarily located in California, the Northeast megalopolis, Florida, the Great Lakes megalopolis, and Texas (Fig. 3b). The 11 extreme-level counties are Los Angeles County, CA (100.00); Miami-Dade County, FL (99.97); Waynesboro City, VA (99.93); Niobrara County, WY (99.9); Buena Vista City, VA (99.87); Wayne County, MI (99.83); Harris County, TX (99.80); Broward County, FL (99.77); San Bernardino County, CA (99.74); Orange County, CA (99.70); and Riverside County, CA (99.67). Most of these counties are located in metropolitan areas and are typically susceptible to extreme weather-induced events. Beyond the well-documented stresses of natural hazards, social vulnerability, and segregation [38,43,44], this study demonstrates that the population in these counties also endure a significant risk of power outages.

To provide deeper insights, we aggregated the county-level ratings to the state level (Fig. 4). Six states were identified as having counites of extreme power system vulnerability level: California (number of counties = 4), Florida ($n = 2$), Virginia (n = 2), Texas ($n = 1$), Michigan (n = 1), and Wyoming (n = 1). Notably, 45 states (91.84 % of 49 states included in this study) contain counties classified as having major, severe, or extreme level of power system vulnerability, and 22 states (44.9 %) include counties with severe or extreme level, indicating widespread electricity disruption risk across the US. We also calculated the proportion of counties in the severe and extreme levels relative to the total number of counties at the state level. The top three states are California (24.56 %), New Jersey (23.81 %), and Florida (14.93 %), suggesting that these states not only face extensive but also a greater extent of power system vulnerability compared to others.

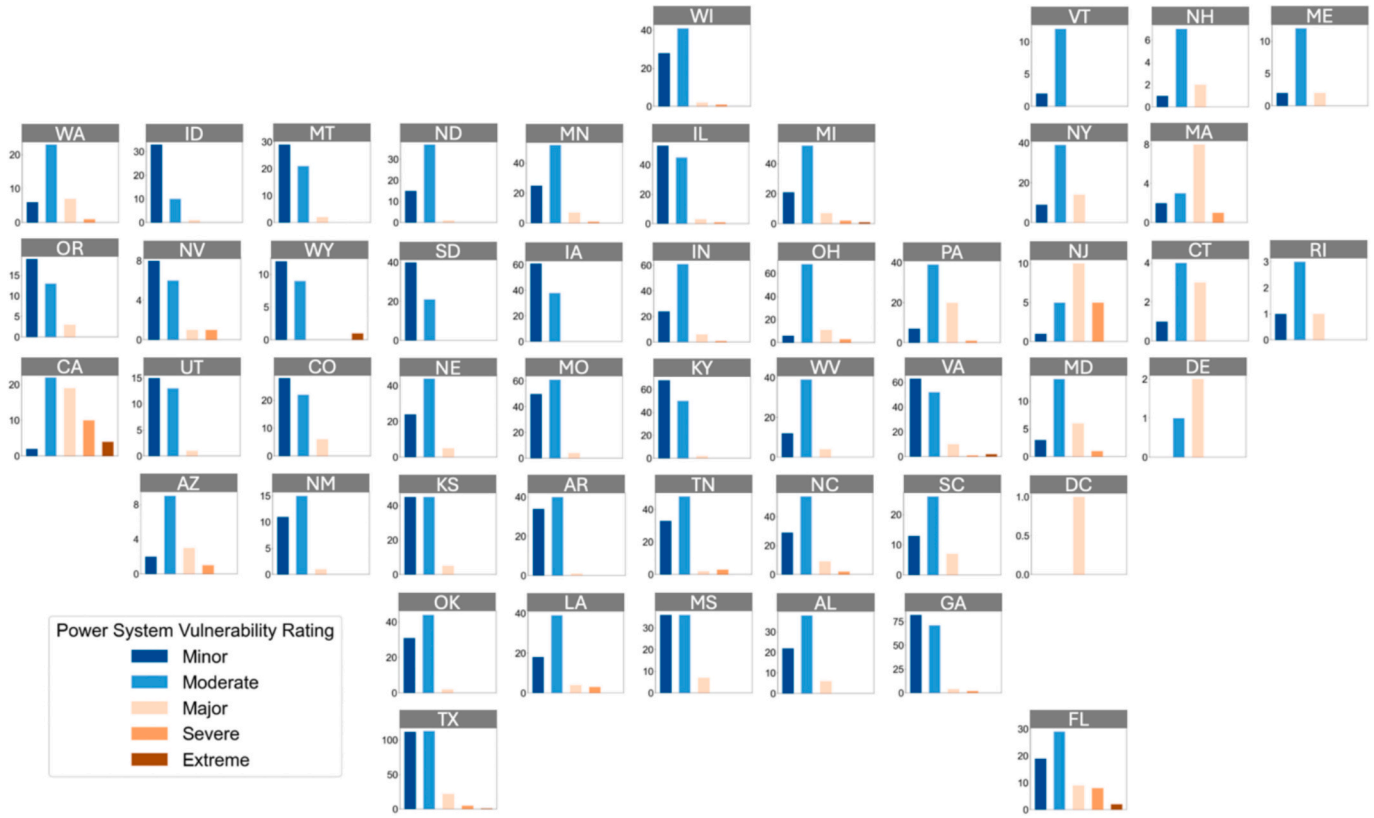


Fig. 4. Spatial distribution of the power outage vulnerability ratings at the state level. We clustered the county-level power system vulnerability values into 5 categories (minor, moderate, major, severe, and extreme) using K-means. The bar plots show the number of counties in each category for every state, with state abbreviations displayed at the top of each subplot. A total of 48 states and Washington, D.C. were mapped according to their approximate geographic locations.

3.3. Annual spatiotemporal patterns in power system vulnerability

Our power outage dataset covers the period from 2014 to 2023, which not only enabled the establishment of the PSVI spanning the entire decade but also allowed for the calculation of annual power system vulnerability values, scores, and ratings, facilitating year-over-year spatiotemporal comparisons and trend analysis.

Spatially, we examined counties that repetitively experienced high power system vulnerability over the years. We set thresholds for the number of years a county faced major, severe, or extreme vulnerability levels ($\geq$2, $\geq$4, $\geq$6, $\geq$8, and = 10). Hotspots were then defined as counties where the cumulative number of years exceeded these thresholds throughout the study period (Fig. 5a). From 2014 to 2023, a total of 333 counties (11.02 % of all counties in this study) were identified as hotspots that consistently experience high levels of power system vulnerability. These counties are primarily located in regions such as the Northeast area, Florida, Texas, California, and Washington. Notably, 22 counties (0.73 %) experienced persistent high vulnerability for 10 years, while 58 counties (1.92 %) experienced this for 8+ years, 109 counties (3.61 %) for 6+ years, and 182 counties (6.02 %) for 4+ years. In California, 38 counties (66.67 %, 57 counties included in this study) faced high power system vulnerability levels for 2+ years, with 6 counties (Los Angeles County, Riverside County, Sacramento County, San Bernardino County, San Diego County, and Ventura County) classified as major, severe, or extreme for 10 years. Similarly, Florida had 31 counties (46.27 %, 67 counties included in this study) with high vulnerability for 2+ years, and 6 counties (Broward County, Duval County, Hillsborough County, Miami-Dade County, Palm Beach County, and Pinellas County) consistently rated as major, severe, or extreme throughout the decade.

To analyze temporal trends, we first plotted the boxplot distributions of annual power system vulnerability values (Fig. 5b). Over the years, the interquartile ranges (Q1-Q3) of the boxplots consistently shift to the right, indicating a steady increase in power system vulnerability. Notably, the average annual increase rate has been significantly higher since 2019 compared to the previous years (2014–2019: 9.86 % vs. 2019–2023: 18.84 %), with the most pronounced increase occurring between 2022 and 2023 (2022–2023: 37.80 %).

We also aggregated the county-level power system vulnerability values to the state level, calculating the annual average for each state (Fig. 5c). The overall trend of the state-level averages also exhibits a consistent increase. However, certain states deviate from this pattern, which show spikes in some specific years. For example, California experienced persistently high vulnerability between 2017 and 2019. Similarly, Connecticut saw a sharp increase in 2018, while New Jersey recorded high values in 2021 and 2022, and Delaware, Maine, and New Hampshire displayed spikes in 2023. These spikes may correspond to extreme weather-induced events: California experienced record-breaking wildfires from 2017 to 2019 [45], while Connecticut, New Jersey, Delaware, Maine, and New Hampshire were affected by winter storms, tropical cyclones, or extreme rainfall during the years [46–48]. The impact of extreme weather-induced events on power system vulnerability is profound. Florida experienced high vulnerability values in 2016, 2020, and 2022, which align with Hurricane Matthew in 2016, Hurricane Sally in 2020, and Hurricane Ian in 2022 [49]. Similarly, Texas saw vulnerability spikes in 2021 and 2023, corresponding to the deadly winter storm in 2021 and the record-breaking heatwave in 2023 [5,50]. The annual spatial distribution of the power system vulnerability scores is available in Supplementary Fig. 5.

3.4. Analysis of disparities in power system vulnerability

To investigate the spatial heterogeneity of power system vulnerability, we first applied the cumulative probability density function to analyze the relationship between power system vulnerability and

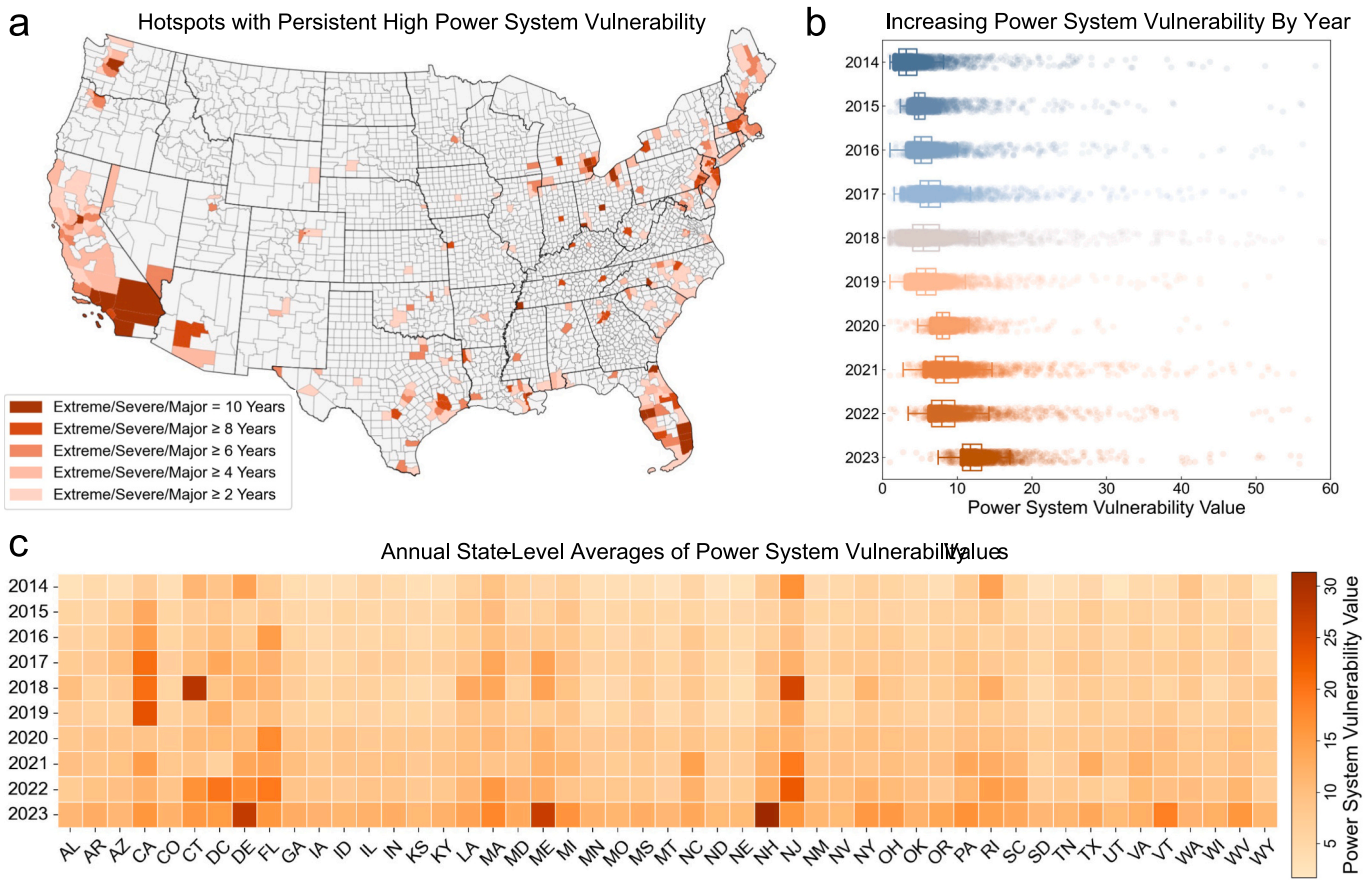


Fig. 5. Spatiotemporal distribution of the annual power system vulnerability index. a. Spatial distribution of counties with persistent high power system vulnerability. We set thresholds for the number of years a county faced major, severe, or extreme levels (≥2, ≥4, ≥6, ≥8, and = 10). Hotspots were defined as counties where the cumulative number of years exceeded these thresholds between 2014 and 2023; **b.** Boxplot of annual distribution of the power system vulnerability values. The Kruskal-Wallis H test confirmed the significant differences among the ten years ($p < 0.001$); **c.** Heatmap of the annual average for the power system vulnerability values at the state level. The X-axis shows the abbreviations of the 48 US states and Washington, D.C.

urbanicity. Counties were classified as either urban or rural based on the 2013 NCHS Urban-Rural Classification Scheme [37]. In Fig. 6a, the distributions between urban counties ($n = 1776$) and rural counties ($n = 1246$) are basically consistent, with the majority of the counties (>90 %) exhibiting low power system vulnerability values (<20). Subsequently, the slopes of the distributions start to increase exponentially, indicating drastically greater power system vulnerability values (>20) for a smaller proportion of counties (<10 %). The long-tail distributions show the heterogeneity of power system vulnerability within each group. When comparing the two distributions, the curve of the urban counties consistently sits above that of the rural counties. This result reveals not only a disproportional distribution of power system vulnerability among all the counties but also that the urban counties have greater power system vulnerability compared with the rural counties.

We further examined the extent to which the form and structure characteristics of urban and rural areas shape the heterogeneity in power system vulnerability. Our analysis involved three urban/rural form and structure dimensions: development density (DD), centrality & segregation (CS), and economic activity (EA) [38]. In rural counties, only DD exhibited a significantly positive correlation with the power system vulnerability values (0.20***), whereas, in urban counties, both DD (0.59***) and EA (0.25***) showed significantly positive correlations (Fig. 6b). The positive correlation of DD indicates that higher DD in both rural and urban areas contribute to increased power system vulnerability. The effect of DD was greater in urban counties compared to rural counties (0.20 vs. 0.59), reflecting the heightened power system vulnerability in cities due to denser population, facilities, and roads. For example, compared to rural areas, dense street trees and lights in cities are more easily blown down by strong winds, destroying power grids and causing power supply disruptions. Moreover, urban counties show a strong correlation between EA and PSVI. Economic activities are heavily dependent on electricity, which may be significantly disrupted by outages.

To assess the impact of different transmission network coverage scenarios on power system vulnerability, we categorized counties based on regional electricity distribution data. The US has seven RTOs that consolidate high-voltage transmission assets to enhance efficiency across a large network [39]. As cases are possible that counties belong to multiple RTOs, we labeled them as boundary counties. Fig. 6c presents a boxplot distribution of the power system vulnerability values across the different RTOs. CAISO (serves California) exhibits the highest power system vulnerability, followed by ISO-NE (serves New England regions), and NYISO (serves New York state). ERCOT (serves Texas), which was severely attacked in the 2021 Winter Storm Uri [5], shows a relatively low vulnerability in this analysis, probably because our data spans a decade and accounts for more than just isolated extreme events. Interestingly, boundary counties also show moderately high power system vulnerability, possibly due to their locations at the intersection of multiple transmission regions, which may contribute to instability in power supply. This analysis reveals the spatial heterogeneity of power system vulnerability related to regional electricity distribution and highlights that areas situated at the edges of transmission regions are sensitive to cascading failures during disruptions.

With the rapid expansion of renewable energy sources like solar and wind, it is essential to examine how the development of these energy technologies impacts power system vulnerability. Due to the

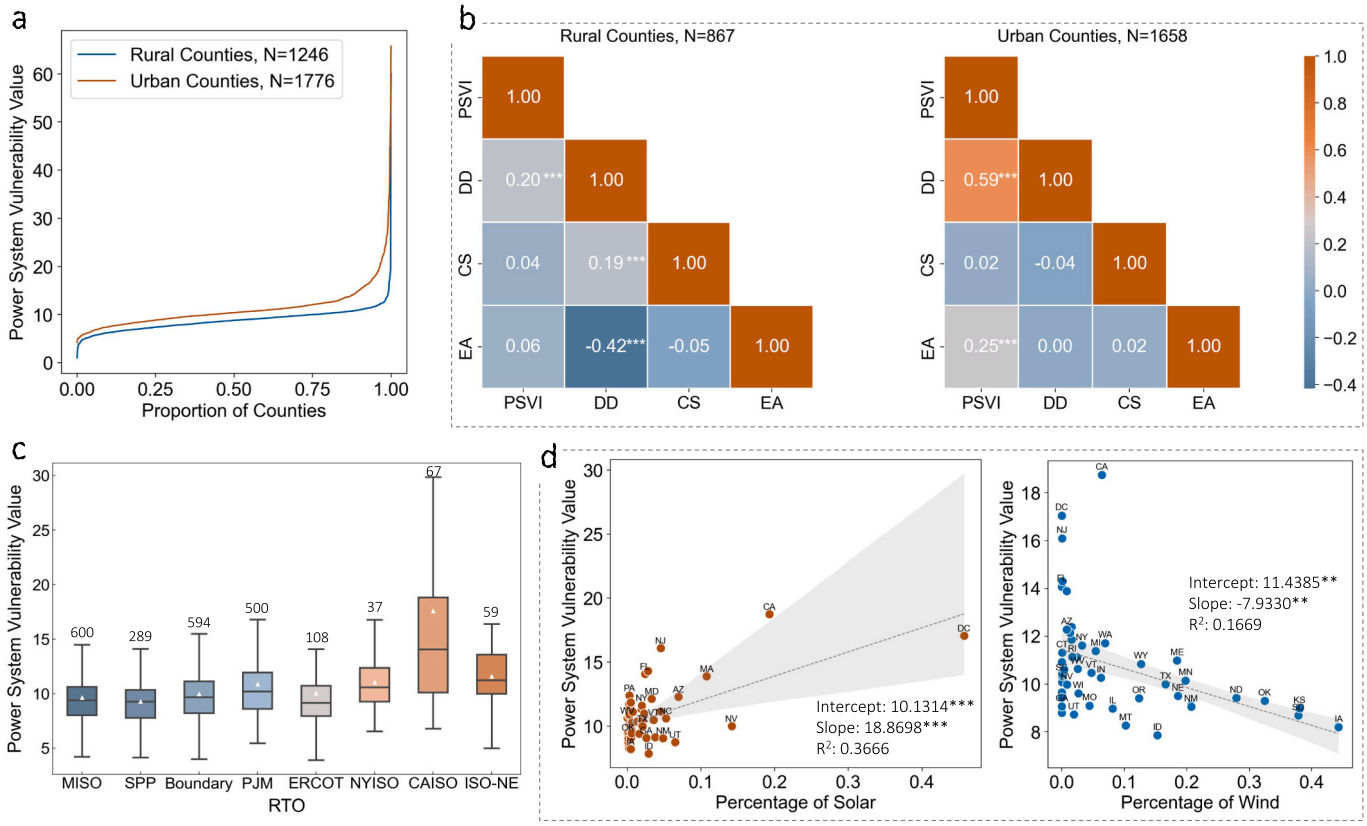


Fig. 6. Disparity in the power system vulnerability index. a. Cumulative probability density curve of the power system vulnerability values between rural and urban counties. The one-way ANOVA test confirmed the significant differences between the two groups ($p < 0.001$); **b.** Heatmaps of Pearson correlation between the urban/rural form and structure features and the power system vulnerability values. Each number represents a correlation coefficient. Due to data unavailability, the analysis includes 867 rural counties and 1658 urban counties; **c.** Boxplot of the power system vulnerability values among different RTOs. Counties belong to multiple RTOs were labeled as boundary. The number above each box refers to the number of counties and the white arrow represents the average of the power system vulnerability values. The Kruskal-Wallis H test confirmed the significant differences among the groups ($p < 0.001$); **d.** OLS regression plots between the power system vulnerability values and the percentage of solar/wind energy used for generation. The shaded area denotes a 95 % confidence interval. The "***" represents $p < 0.001$ and "**" represents $p < 0.01$.

unavailability of consistent annual, county-level renewable generation data over the study period, we used the decadal average of net generation percentages for solar and wind at the state level and compared them with the power system vulnerability values through Ordinary Least Squares (OLS) regression. In Fig. 6d, solar energy exhibits a statistically significant but moderate positive correlation with power system vulnerability, indicating that states with higher solar generation may experience greater vulnerability of power systems. For example, California heavily relies on solar energy [51], and also shows higher power system vulnerability.

However, wind energy presents a weaker inverse relationship with power system vulnerability, though the low R^2 indicates that a linear correlation is not strongly supported. Wind resources are primarily concentrated in the central and coastal states of the US, while some coastal states have seen slower development of wind generation [52]. For example, Florida has yet to pass legislation to permit wind generation [53]. Louisiana, despite its abundant offshore wind resources, only began developing wind generation in 2023 [54]. California imports a significant portion of wind energy from other states, which does not count as in-state wind generation [55]. Consequently, these states with higher power system vulnerability tend to have a lower proportion of wind energy, while the central regions with abundant wind generation exhibit lower power system vulnerability. This contrast contributes to the differing impacts of wind energy compared to solar energy. In addition, although these correlations are statistically significant, the modest R^2 values indicate that the OLS models explain only a limited portion of the variance in vulnerability. This limited explanatory power may be attributed to the coarse spatiotemporal resolution of the renewable generation, as well as the influence of broader contextual factors (i.e., policy environments) on the relationship between renewable energy and power system.

4. Discussion

Despite considerable evidence regarding the vulnerability of the US power infrastructure [2–5], a systematic and national-level assessment of the spatiotemporal patterns of power system vulnerability has still been lacking. This study utilizes a decadal outage data at 15-min intervals in 3022 US counties to establish machine learning-based PSVI. To the best of our knowledge, PSVI is the first comprehensive, data-driven, and quantitative tools specifically designed to evaluate power system vulnerability across all US counties. This outcome addresses a critical gap in existing research, where numerous spatial indexes capture social and physical vulnerabilities, but none adequately represents the vulnerability of power systems themselves. Our PSVI has important implications for preparedness strategies by enabling emergency managers and utilities to leverage the index before impending extreme events, identifying high-risk counties with unprecedented precision, and allowing for targeted and proactive allocation of restoration crews and resources. Furthermore, the integration of the PSVI into hazard mitigation and resilience planning would significantly enhance these processes.

We observed a consistent increase of power system vulnerability in US counties over the past decade. Previous studies have suggested that

the rise in severe weather events due to climate change will likely lead to more frequent power outages in the US [5,11]. Our multidimensional analysis of power outages confirms that the PSVI at the county level has steadily risen from 2014 to 2023. Specifically, we found a significant increase in the annual PSVI in the five years following 2019 compared to the preceding five years. At the state level, the higher power system vulnerability corresponds with increasing frequency of natural hazards. It is important to note that although climate change-induced nature hazards drive recent large-scale outages, factors such as the aging grid and rising energy demand contribute to smaller-scale but frequent outages [6]. While our study identifies the existence of relationship between natural hazards and outage events, examining the interaction between the factors (i.e., climate change, aging grids, energy demand) and power outages falls beyond the scope of this paper.

The distribution of the PSVI follows a long-tail pattern, where the majority of counties exhibits low vulnerability while a smaller proportion experiences extremely high vulnerability. These high-vulnerability areas are concentrated on the West Coast (particularly California and Washington), the East Coast (notably Florida, the Northeast metropolitan area, and the New England areas), the Great Lakes megalopolis (mainly Chicago-Detroit area), as well as the Texas Gulf Coast area. The PSVI offers deeper and more nuanced characterization of spatiotemporal variations across counties, unveiling patterns previously less understood [2,13,14].

Our findings illuminate complex variations in power systems' vulnerability across urban-rural gradients and pinpoint specific structural and developmental features that influence vulnerability levels. Urban counties generally have higher power system vulnerability than rural counties. In urban environments, higher vulnerability to outages can be attributed to dense infrastructure development, large-scale electricity demand due to population concentration, and the proximity of power grids to critical facilities. Such pattern increases the likelihood of cascading failures from equipment malfunctions, accidents, or construction activities. This study identifies development density as a key determinant of power system vulnerability, which underscores the pressing need for tailored resilience measures in rapidly growing urban areas with dense development. This revelation highlights the importance of the relationship between urban planning and infrastructure vulnerability and demonstrates how urban development contributes to power system vulnerability.

In addition, our study reveals significant variation in power system vulnerability across different RTOs and states with diverse energy structures. We found that CAISO exhibited markedly higher power system vulnerability compared to the other six US RTOs. The resilience of the grid within CAISO has been extensively studied in the prior work [56,57]. Our analysis also highlights the high vulnerability of counties located along the borders of regional transmission organizations, suggesting that these "gray areas" warrant particular attention. Furthermore, our study analyzed the ongoing energy transition by examining the relationship between power system vulnerability and the proportion of renewable energy generation at the state level. We observed that states with higher solar generation may experience greater power system vulnerability. Given the essential role of energy in daily life, such correlations in power system vulnerability raise significant energy inequality concerns. By exploring these disparities, our study contributes to the growing literature on environmental justice, highlighting an issue that has not yet been sufficiently addressed (i.e., the inequitable distribution of power outages and restoration efforts) [58–60].

5. Conclusions

This study analyzed ~179 million power outage records at 15-min intervals between 2014 and 2023 to conduct a county-level evaluation of power system vulnerability. We proposed a power system vulnerability assessment framework based on three environmental hazard exposure dimensions. Using interpretable machine learning models of XGBoost and SHAP, we generated the PSVI for 3022 counties. Our key findings are summarized as follows.

- Power system vulnerability has shown a consistent increase across the US counties over the past decade.
- A total of 318 counties across 45 states has been identified as hotspots for high power system vulnerability, particularly along the West Coast, East Coast, Great Lakes megalopolis, and Gulf of Mexico.
- Urban counties and those located along regional transmission boundaries tend to exhibit significantly higher vulnerability.

Our results highlight the significance of the proposed PSVI for evaluating the vulnerability of communities to power outages. The findings provide data-driven insights and practical metrics that empower power infrastructure owners, operators, emergency managers, and public officials to effectively address the escalating vulnerability of power infrastructure across the US.

There were also limitations in this study, which could be addressed in the future. First, the county-level assessment does not capture finer sub-county-level details such as increasing development and population influx over the years. While we identified spatial heterogeneity in power system vulnerability, the trends might differ with more granular data. If higher resolution data becomes available, future studies could benefit from examining sociodemographic characteristics and other metrics at finer spatial resolutions to better understand inequalities in power outage extent, especially for marginalized communities. Second, factors such as equipment reliability, system reserves, and grid strength (e.g., inertia, fault tolerance) were not included due to data constraints. These factors are important determinants of power system vulnerability, exploring their relationship with the proposed metrics remains a promising direction for future research as relevant data become available.

Code availability

All analyses were conducted using Python. The code that supports the findings of this study is available from the corresponding author upon request.

CRediT authorship contribution statement

Junwei Ma: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Bo Li:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Olufemi A. Omitaomu:** Writing – review & editing, Data curation. **Ali Mostafavi:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Science Foundation under Grant CMMI-1846069 (CAREER). Any opinions, findings, conclusions, or recommendations expressed in this research are those of the authors and do not necessarily reflect the view of the funding agency.

Appendix A. Supplementary data

Supplementary material to this article can be found online at https://doi.org/10.1016/j.apenergy.2025.126360.

Data availability

The power outage data from 2014 to 2022 is publicly available at https://doi.org/10.1038/s41597-024-03095-5, and data for 2023 can be obtained from ORNL upon request. The other datasets used in this paper are publicly accessible and cited in this paper. The 2019 TIGER/Line US County Shapefile was utilized to create the nationwide map for this study.

References

[1] Xu L, et al. Resilience of renewable power systems under climate risks. Nature Reviews Electrical Engineering 2024;1(1):53–66.
[2] Do V, et al. Spatiotemporal distribution of power outages with climate events and social vulnerability in the USA. Nat Commun 2023;14(1):2470.
[3] Feng K, Ouyang M, Lin N. Tropical cyclone-blackout-heatwave compound hazard resilience in a changing climate. Nat Commun 2022;13(1):4421.
[4] Baik S, et al. Estimating what US residential customers are willing to pay for resilience to large electricity outages of long duration. Nat Energy 2020;5(3):250–8.
[5] Flores NM, et al. The 2021 Texas power crisis: distribution, duration, and disparities. J Expo Sci Environ Epidemiol 2023;33(1):21–31.
[6] Bie Z, et al. Battling the extreme: a study on the power system resilience. Proc IEEE 2017;105(7):1253–66.
[7] Mildenberger M, et al. The effect of public safety power shut-offs on climate change attitudes and behavioural intentions. Nat Energy 2022;7(8):736–43.
[8] Shuai M, et al. Review on economic loss assessment of power outages. Procedia Comput Sci 2018;130:1158–63.
[9] Bhattacharyya A, Hastak M. Indirect cost estimation of winter storm–induced power outage in Texas. J Manag Eng 2022;38(6). p. 04022057.
[10] Sayarshad HR, Ghorbanloo R. Evaluating the resilience of electrical power line outages caused by wildfires. Reliab Eng Syst Saf 2023;240:109588.
[11] Zheng D, et al. Climate change impacts on the extreme power shortage events of wind-solar supply systems worldwide during 1980–2022. Nat Commun 2024;15(1):5225.
[12] Dugan J, Byles D, Mohagheghi S. Social vulnerability to long-duration power outages. Int J Disaster Risk Red 2023;85:103501.
[13] Ganz SC, Duan C, Ji C. Socioeconomic vulnerability and differential impact of severe weather-induced power outages. PNAS Nexus 2023;2(10):pgad295.
[14] Flores NM, et al. Powerless in the storm: severe weather-driven power outages in New York state, 2017–2020. PLOS Climate 2024;3(5):e0000364.
[15] Dikshit S, Dobson I, Alipour A. Cascading structural failures of towers in an electric power transmission line due to straight line winds. Reliab Eng Syst Saf 2024;250:110304.
[16] Dai Y, et al. Risk assessment and mitigation of cascading failures using critical line sensitivities. IEEE Trans Power Syst 2023;39(2):3937–48.
[17] Ahmad A, et al. The improvement in transmission resilience metrics from reduced outages or faster restoration can be calculated by rerunning historical outage data. arXiv preprint arXiv:2501.06042. 2025.
[18] Hanna R, Marqusee J. Designing resilient decentralized energy systems: the importance of modeling extreme events and long-duration power outages. Iscience 2022;25(1).
[19] Bhattacharyya A, Hastak M. A data-driven approach to quantify disparities in power outages. Sci Rep 2023;13(1):7247.
[20] Ankit A, et al. US resilience to large-scale power outages in 2002–2019. J Saf Sci Resilience 2022;3(2):128–35.
[21] Stevens KA, Belligoni S. Policymaking in the dark: the impact of power outage information asymmetry on local government resilience efforts in Florida. Int J Disaster Risk Red 2024;105:104381.
[22] Flanagan BE, et al. A social vulnerability index for disaster management. J Homeland Sec Emerg Manage 2011;8(1). p. 000010220154773551792.
[23] Yarveysi F, et al. Block-level vulnerability assessment reveals disproportionate impacts of natural hazards across the conterminous United States. Nat Commun 2023;14(1):4222.
[24] Brelsford C, et al. A dataset of recorded electricity outages by United States county 2014–2022. Sci Data 2024;11(1):271.
[25] Casey, J.A., et al., Measuring long-term exposure to wildfire PM2. 5 in California: Time-varying inequities in environmental burden. Proceedings of the National Academy of Sciences, 2024. **121**(8): p. e2306729121.
[26] Miao L, et al. Unveiling the dynamics of sequential extreme precipitation-heatwave compounds in China. npj Clim Atmos Sci 2024;7(1):67.
[27] Nieuwenhuijsen MJ. Exposure assessment in environmental epidemiology. USA: Oxford University Press; 2015.
[28] Li B, et al. Recent Decade's Power Outage Data Reveals the Increasing Vulnerability of US Power Infrastructure. arXiv preprint arXiv:2408.15882. 2024.
[29] Qin L, et al. Global expansion of tropical cyclone precipitation footprint. Nat Commun 2024;15(1):4824.
[30] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Adv Neural Inf Proces Syst 2017;30.
[31] Federal Emergency Management Agency. Natl Risk Index 2023. Available from: https://hazards.fema.gov/nri/learn-more.
[32] Li B, Mostafavi A. Incorporating environmental considerations into infrastructure inequality evaluation using interpretable machine learning. Comput Environ Urban Syst 2025;120:102301.
[33] Farrar DE, Glauber RR. Multicollinearity in regression analysis: the problem revisited. Rev Econ Stat 1967:92–107.
[34] Miles J. Tolerance and variance inflation factor. In: Encyclopedia of statistics in behavioral science; 2005.
[35] Chawla NV, et al. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57.
[36] Likas A, Vlassis N, Verbeek JJ. The global k-means clustering algorithm. Pattern Recogn 2003;36(2):451–61.
[37] National Center for Health Statistics. NCHS Urban-Rural Classification Scheme for Counties. Available from: https://www.cdc.gov/nchs/data_access/urban_rural.htm#Data_Files_and_Documentation; 2013.
[38] Ma J, Mostafavi A. Urban form and structure explain variability in spatial inequality of property flood risk among US counties. Communi Earth Environ 2024;5(1):172.
[39] Community Choice Energy. Regional Transmission Organizations (RTOs) [cited 2024 06-03]; Available from: https://energyfreedomco.org/f1-RTOmap.php; 2015.
[40] U.S. Energy Information Administration. Net Generation by Energy Source: Total (All Sectors), 2014-July 2014. Available from: https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=table_1_01; July 2024.
[41] Ahmed M, et al. Mitigating uncertainty problems of renewable energy resources through efficient integration of hybrid solar PV/wind systems into power networks. IEEE Access 2024;12:30311–28.
[42] Watson EB, Etemadi AH. Modeling electrical grid resilience under hurricane wind conditions with increased solar and wind power generation. IEEE Trans Power Syst 2019;35(2):929–37.
[43] Nilforoshan H, et al. Human mobility networks reveal increased segregation in large cities. Nature 2023;624(7992):586–92.
[44] Bates PD, et al. Combined modeling of US fluvial, pluvial, and coastal flood hazard under current and future climates. Water Resour Res 2021;57(2). p. e2020WR028673.
[45] Western Fire Chiefs Association. History of California wildfires. Available from: https://wfca.com/wildfire-articles/history-of-california-wildfires/; November 17, 2022.
[46] Cerrai D, et al. Outage prediction models for snow and ice storms. Sustain Energy, Grids Networks 2020;21:100294.
[47] The New York Times. Winter Storm Disrupts U.S. With Power Outages and Icy Roads. Available from: https://www.nytimes.com/live/2022/02/03/us/winter-storm-snow-ice; Feb. 3, 2022.
[48] Climate Central. Weather-related power outages rising. Available from: https://www.climatecentral.org/climate-matters/weather-related-power-outages-rising; April 23, 2024.
[49] Florida Climate Center. Some Notable Hurricanes In Florida's Histroy. Available from: https://climatecenter.fsu.edu/topics/hurricanes; 2024.
[50] U.S. National Weather Service. Exceptional Heat of Summer 2023. Available from: https://www.weather.gov/lub/events-2023-2023summer-heat; 2023.
[51] Commission, C.E. New data shows growth in California's clean electricity portfolio and battery storage capacity. Available from: https://www.energy.ca.gov/news/2023-05/new-data-shows-growth-californias-clean-electricity-portfolio-and-battery; May 25, 2023.
[52] Christopher Niezrecki U-L. Why offshore wind power is struggling. Available from: https://www.governing.com/resilience/why-offshore-wind-power-is-struggling; June 17, 2024.
[53] Lopez A, et al. Impact of siting ordinances on land availability for wind and solar development. Nat Energy 2023;8(9):1034–43.
[54] National Wildlife Federation. Improving LA's Offshore Wind. June 28, 2024; Available from: https://offshorewind.nwf.org/2024/06/improving-las-offshore-wind/#:~:text=In%20December%202023%2C%20Louisiana%27s%20State%20Mineral%20and,developing%20offshore%20wind%20projects%20in%20Louisiana%20waters.
[55] U.S. Energy Information Administration. California - State Energy Profile Analysis. May 16, 2024; Available from: https://www.eia.gov/state/analysis.php?sid=CA.
[56] Zeighami A, et al. US west coast droughts and heat waves exacerbate pollution inequality and can evade emission control policies. Nat Commun 2023;14(1):1415.
[57] Xu L, et al. On the resilience of modern power systems: a comprehensive review from the cyber-physical perspective. Renew Sust Energ Rev 2021;152:111642.
[58] Coleman N, et al. Energy inequality in climate hazards: empirical evidence of social and spatial disparities in managed and hazard-induced power outages. Sustain Cities Soc 2023;92:104491.
[59] Hsu C-W, Mostafavi A. Untangling the relationship between power outage and population activity recovery in disasters. Resilient Cities Struct 2024;3(3):53–64.
[60] Lee C-C, Maron M, Mostafavi A. Community-scale big data reveals disparate impacts of the Texas winter storm of 2021 and its managed power outage. Humanities Soc Sci Commun 2022;9(1):1–12.