processes

MDPI

Article

A Study on the Key Factors Influencing Power Grid Outage Restoration Times: A Case Study of the Jiexi Area

Jiajun Lin [1], Ruiyue Xie [1], Haobin Lin [1], Xingyuan Guo [1,*], Yudong Mao [2] and Zhaosong Fang [2]

[1] Jieyang Power Supply Bureau, Guangdong Power Grid Co., Ltd., Jieyang 522000, China
[2] School of Civil and Transportation Engineering, Guangzhou University, Guangzhou 510006, China
* Correspondence: gxypower@163.com

Abstract

In rural and mountainous regions, power supply reliability remains a persistent challenge due to structural vulnerabilities, data incompleteness, and limited automation. In this study, a data-driven methodology is leveraged, wherein a validated machine learning framework comprising Random Forest (RF), Lasso Regression, and Recursive Feature Elimination (RFE) is applied to analyze outage data. The machine learning models, validated on a held-out test set, demonstrated modest but positive predictive performance, confirming a quantifiable, non-random relationship between grid structure and restoration time. This validation provides a credible foundation for the subsequent feature importance analysis. Through a transparent, consensus-based analysis of these models, the most robust influencing factors were identified. The results reveal that key structural indicators related to network redundancy (e.g., Inter-Bus Loop Rate) and electrical stress (e.g., Peak Daily Load Current, Load Factor) are the most significant predictors of prolonged outages. Furthermore, statistical analyses confirm that increasing structural redundancy and regulating line loads can effectively reduce outage duration. These findings offer practical, data-driven guidance for prioritizing investments in rural grid planning and reinforcement. This study contributes to the broader application of machine learning in energy systems, particularly showcasing a robust methodology for identifying key drivers under data and resource constraints.

Keywords: rural distribution networks; outage restoration time; power supply reliability; machine learning; feature selection; structural grid indices; grid engineering analysis

check for updates

Academic Editors: Yuzhou Zhou, Yizhou Zhou, Daogui Tang and Hang Shuai

Received: 20 June 2025
Revised: 6 August 2025
Accepted: 12 August 2025
Published: 25 August 2025

Citation: Lin, J.; Xie, R.; Lin, H.; Guo, X.; Mao, Y.; Fang, Z. A Study on the Key Factors Influencing Power Grid Outage Restoration Times: A Case Study of the Jiexi Area. *Processes* **2025**, *13*, 2708. https://doi.org/10.3390/pr13092708

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Energy is not only necessary for production and operation in modern society, but also serves as a vital material basis for human survival and a key driving force for social and economic development. The reliability and continuity of power supply are critical indicators for modern economies and are primary objectives for governments and distribution companies. In recent years, significant policy efforts have been directed towards enhancing power supply quality in remote and less-developed areas. For instance, policies issued by the State Council of China and the National Energy Administration (NEA) emphasize the need to narrow the urban–rural power supply gap and improve services in remote regions [1–3].

While urban grids benefit from compact, highly automated underground networks, rural and mountainous areas often rely on long, radial overhead lines [4]. These grids are structurally fragile and highly susceptible to disturbances, resulting in frequent and

prolonged power outages [5]. This disparity underscores the urgent need for targeted research to improve the resilience of rural distribution networks.

Academic research has approached the challenge of enhancing power grid resilience from multiple perspectives. A significant body of research focuses on developing quantitative assessment frameworks and metrics to measure resilience against various disruptions, often leveraging large-scale outage data [6,7]. These frameworks provide a high-level understanding of system performance and are essential for benchmarking [8]. Another major area of research addresses resilience through dynamic operational and control strategies. This includes real-time fault detection and remediation using advanced machine learning [9,10], optimized operational strategies to manage uncertainties [11], and deep learning applications for fault location and classification in active grids [12,13]. Furthermore, a critical area of focus is the impact of non-structural external factors, with numerous studies analyzing the severe effects of extreme weather events on power system resilience and proposing mitigation strategies [14–18].

In recent years, data-driven methods, particularly machine learning (ML), have emerged as powerful tools for analyzing outage data and identifying underlying vulnerabilities [19]. A variety of ML models have been applied to predict power outage duration (POD), offering valuable insights for utility operations [20–23]. To support these predictive models, sophisticated feature selection techniques are crucial for identifying the most influential variables from a vast number of potential factors [24,25]. For instance, permutation importance has been effectively used with Random Forest for load forecasting [26], and advanced interpretability methods such as SHAP (SHapley Additive exPlanations) have been employed to explain complex models such as XGBoost in load-forecasting contexts, enhancing transparency [27–29]. These studies—represented by existing studies such as Janiszewski et al. [30], Kunaifi and Reinders [31], Kozyra et al. [32], Guo et al. [33], Rojas-Zerpa et al. [34], Sami et al. [35], Manninen et al. [36], and Yazdanpanah et al. [37]—have significantly advanced the field.

However, despite these advancements, a specific gap in research remains, particularly concerning structurally weak and data-scarce rural distribution networks. Firstly, while many studies predict outage duration, few have systematically focused on quantifying the impact of inherent structural static grid indicators on restoration time. Secondly, advanced interpretability methods such as SHAP often require well-calibrated, complex models, which may be challenging to develop in data-limited environments. A summary of the representative literature and research gaps addressed by the present study is provided in Table 1.

Table 1. Summary of the relevant literature and research gaps.

Author(s) and Year	Main Focus/Methodology	Key Contributions	Research Gap Addressed by This Study
Ghasemkhani et al. (2024) [21]	ML for power outage duration (POD) prediction (e.g., CatBoost)	Feasibility of ML for operational POD forecasting	Lacked focus on static structural indicators (predictive, not explanatory)
Yazdanpanah et al. (2024) [37]	ML and data envelopment analysis (DEA) for optimal reliability targets setting	Framework for reliability benchmarking and planning	Focused on high-level targets, not specific line-level structural causes
Ren et al. (2021) and Panteli et al. (2015) [16,17]	Analysis of extreme weather impact on outages	Highlighted the role of external factors (non-structural)	Focused on outage causes, not restoration duration from a structural view

Table 1. *Cont.*

Author(s) and Year	Main Focus/Methodology	Key Contributions	Research Gap Addressed by This Study
Lee et al. (2023) and Li et al. (2022) [27,28]	Interpretable ML (SHAP) for load forecasting	Power of advanced interpretability tools (SHAP)	Applied to data-rich task, did not address feature identification in data-scarce rural grids
Jiang et al. (2014) and Majidi et al. (2017) [10,12]	ML/DL for data-driven fault location	Focused on "where" and "what" of a fault (dynamic diagnosis)	Did not analyze the "why" of restoration tied to pre-existing structural limits
Manninen et al. (2022) [36]	ML-based health index prediction for transmission lines	Data-driven approach for asset management and condition monitoring	Focused on the component health, not system-level restoration time
Kozyra et al. (2022) [32]	Impact analysis of remote control on reliability indices	Quantified benefits of automation on reliability	Lacked a systematic framework to rank multiple structural factors simultaneously
Janiszewski et al. (2018) [30]	SAIDI/SAIFI optimization via fault analysis and mathematical models	Practical methods for improving grid-level reliability indicators	Relied on traditional models, not on a comprehensive, non-linear ML-based feature evaluation
Rojas-Zerpa and Yusta (2015) [34]	Multi-criteria decision-making (DEA with SVM/RF) for supply planning	High-level planning tool for remote areas	Did not perform granular analysis of specific line-level engineering indicators
This Study	ML-based analysis of structural indicators (RF, Lasso, and RFE) on a real-world rural grid dataset	Identifies and validates key physical grid characteristics that are most influential on restoration time in a resource-constrained environment	-

To address the gap in research, this study focuses on the Jiexi area in eastern Guangdong Province as a representative case of a remote, mountainous rural grid. It adopts a data-driven approach that combines three distinct and robust feature selection methods—Random Forest (RF), Lasso Regression, and Recursive Feature Elimination (RFE)—to systematically assess and rank the key line-level structural indicators affecting outage restoration time. This study aims to (1) identify the most critical structural grid features correlated with restoration delays, (2) validate the findings through a rigorous model performance evaluation, and (3) offer targeted, data-driven recommendations for improving the resilience of rural power systems through strategic investment in physical infrastructure.

2. Methods

2.1. Study Area and Characterization

The Jiexi area is located in the eastern part of Guangdong Province, west of Jieyang City, and it is situated in a subtropical monsoon climate zone. The area occupies the northern part of a mountainous hill area and the southern part of the plains, with a greater proportion of mountainous and varied terrain. In recent years, with economic progress, the local government has continued to strengthen the construction of power grids, and it is committed to optimizing residents' electricity access and reducing the differences between urban and rural electricity supplies; however, despite improvements towards higher quality and more stable electricity supply, some remote rural areas with low-quality electricity and other issues still require improvements.

As clarified in the above analysis, remote rural areas have difficulty maintaining power supply reliability and face difficult challenges, especially in mountainous regions. Frequent

power outages and long fault repair times are the two main issues. At the same time, based on the data provided, the following main elements were researched and investigated:

(1) Which metrics of the distribution network architecture significantly affect the duration of outages?
(2) Which distribution architecture metrics have a significant impact on outage duration?
(3) What is the quantitative relationship between key indices and outage duration affecting power supply reliability?

2.2. Data Sources

The uncertainty and unavailability of grid data render the qualitative and quantitative analyses of grid power supply reliability indicators more difficult. Nonetheless, based on the outage telegram records collected by the JieXi Power Supply Bureau and the indicators of the distribution network's architecture, this study carries out an investigation to provide constructive guidance in order to improve power supply reliability and power consumption quality within the region.

2.3. Data Processing and Analysis

The data processing flowchart is shown in Figure 1, which mainly includes data preprocessing, feature importance assessment (Random Forest, Lasso Regression, and Recursive Feature Elimination), and comprehensive analyses. All data processing and modeling were conducted using Python (Version 3.9) with the Scikit-learn (Version 1.3) and Pandas (Version 2.0) libraries.

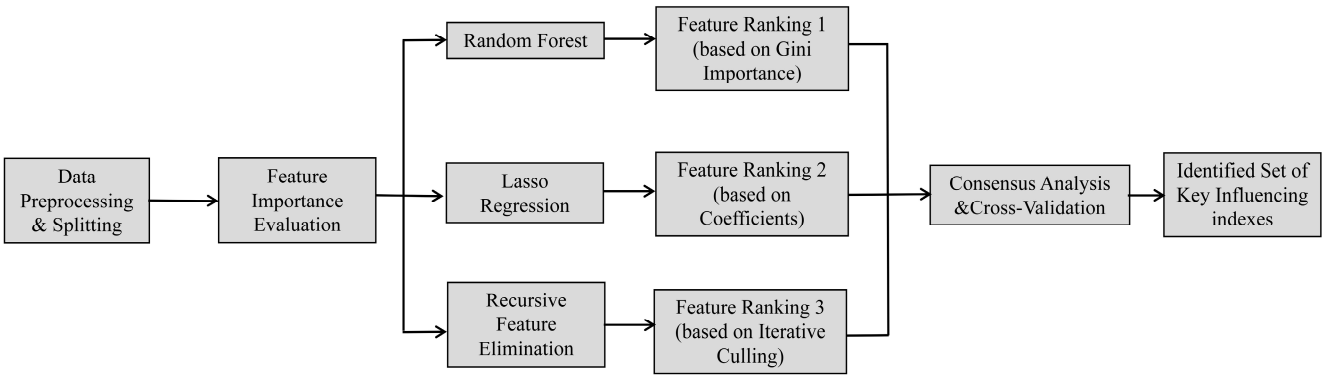


Figure 1. Data processing flowchart.

In this study, the dataset collected by the JieXi Power Supply Bureau is first matched and categorized according to the name of the bus line and the bus section it is relevant to, and outage resumption is selected as the target variable, while the line-network-related indexes are selected as the influence variables. In order to minimize model bias, observation samples with missing target variables or key feature variables are excluded to ensure the completeness of the training data. The feature variables cover multiple dimensions, such as customer size, current load, network structure, ring network characteristics, and number of switches.

2.3.1. Random Forest Regression Model (RF)

Random Forest is a predictive model containing many decision trees with good nonlinear modeling and feature evaluation capabilities suitable for dealing with multivariate and high-dimensional data. The network architecture of RF is shown in Figure 2 [38,39]. A combination of self-service sampling (Bagging) and feature random selection is used in the model training process to effectively reduce the risk of overfitting. The feature

importance assessment is based on the degree of impurity reduction introduced by each feature in the split node [40]. In this study, an RF regression model is used to extract the importance scores of each feature in the prediction of outage restoration times and identify the key variables.

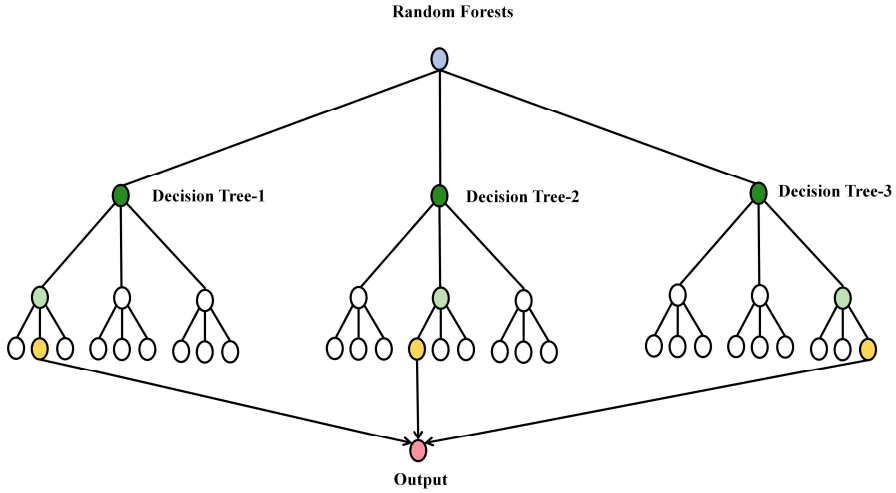


Figure 2. Random Forest network architecture.

2.3.2. Lasso Regression Model (LRM)

Lasso (Least Absolute Shrinkage and Selection Operator) is a linear regression method that introduces L1 regularization with automatic feature selection, and the network architecture of LRM is schematically shown in Figure 3 [41]. Its constraint term shrinks some of the feature coefficients to zero, thus enabling variable selection [42]. To obtain the optimal regularization parameter α, Lasso CV is used in this study for cross-validation in order to select the optimal model configuration [38]. The final retained non-zero coefficient features are the important variables, and the magnitude of their absolute values can be used as an importance measure.

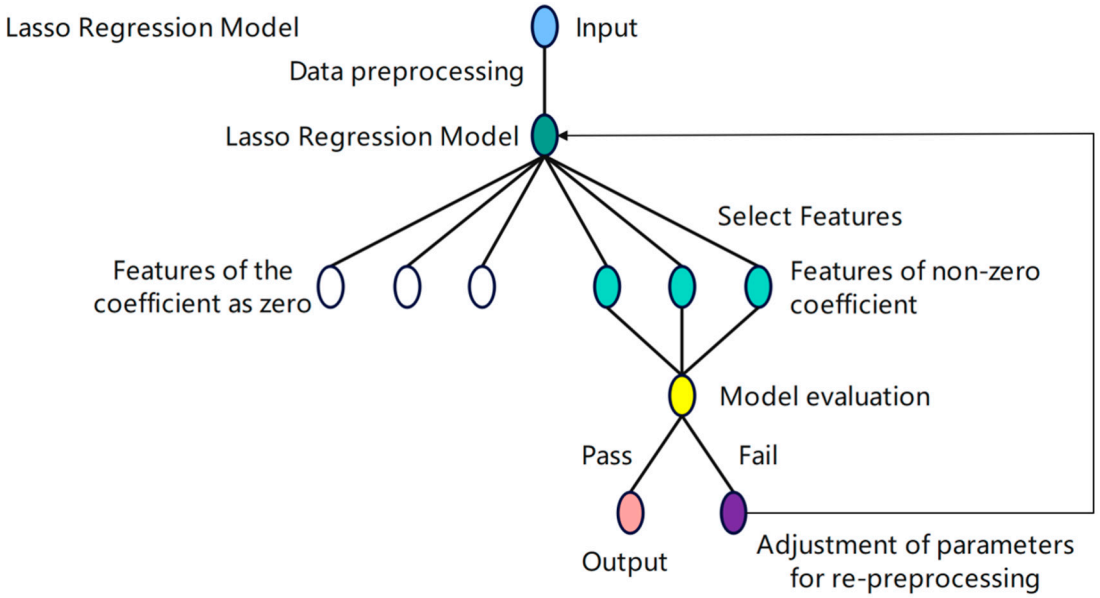


Figure 3. Lasso Regression network architecture.

2.3.3. Recursive Feature Elimination (RFE)

RFE is a feature subset selection method based on model evaluation that progressively removes the features that have the least impact on the model to obtain an optimal feature subset, and the network architecture of RFE is schematically shown in Figure 4 [43]. In each iteration, the base learner is utilized to evaluate feature importance and remove the lowest features. In this study, the RF regressor was chosen as the base model, and the number of retained features was set to 15. With its stable non-linear modeling capability, stable high-importance features are extracted after several rounds of iterations [44].

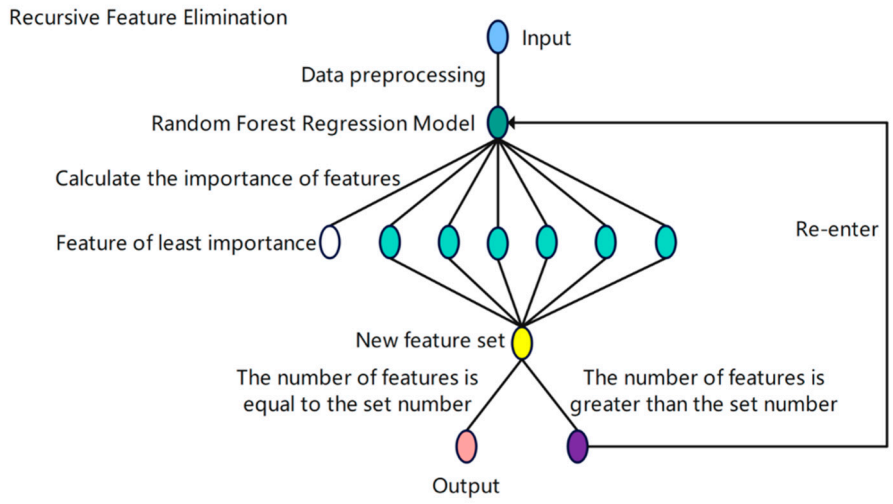


Figure 4. Recursive Feature Elimination network architecture.

2.3.4. Consensus-Based Feature Importance Analysis

Due to the distinct operational logic of different feature evaluation models— Random Forest (RF) relies on impurity reduction (Gini importance), Lasso Regression relies on sparsity through L1 regularization, and Recursive Feature Elimination (RFE) relies on iterative model-based culling—relying on a single method may yield results that are sensitive to the specific technique employed. To mitigate this potential bias and enhance the robustness and reliability of the identified key factors, a consensus-based analytical approach was adopted instead of a statistical aggregation of scores.

The methodology for this consensus analysis is structured as follows:

(1) Independent Feature Ranking: Feature importance rankings derived from three methods, Random Forest, Lasso Regression, and Recursive Feature Elimination, each of which provides a unique analytical perspective (see Section 3.3). The RF model provides rankings based on Gini importance. Meanwhile, the Lasso model ranks features based on the absolute magnitude of the feature coefficients. Finally, RFE produces the ranking based on feature persistence during iterative backward elimination, which identifies a core subset of features that collectively maintain predictive power.

(2) Definition of Consensus: Following the generation of independent rankings, a consensus rule was established to identify the most robust indicators. "Consensus features" are defined as those that consistently appear among the top-ranked variables across the methodologically distinct lists. For the purposes of this study, the specific criterion for this cross-comparison was uniformly set as the top 15 features identified by each of the three models.

(3) Identification of Consensus Features: The final step involves systematically comparing the top-ranked features from all three lists. Features are then categorized by their level of consensus (i.e., identified by three, two, or only one model). Those identified by

all three models are considered the most robust and trustworthy indicators, as their significance is validated across linear, non-linear, and iterative-selection frameworks.

This triangulation approach leverages the non-linear pattern detection of RF, the variable-screening strength of Lasso, and the iterative selection logic of RFE. By focusing on the consensus among the results, this strategy provides a more stable and defensible foundation for identifying the structural indicators that have a genuine and significant impact on outage restoration time.

2.4. Model Training and Hyperparameter Optimization

For the development of predictive models, a Pipeline methodology was employed, combining feature scaling (Standard Scaler) with regression algorithms. To identify the optimal configuration of both the Random Forest and XGBoost models, a rigorous hyperparameter tuning process was implemented using Grid Search CV with 5-fold cross-validation on the training dataset. The objective was to find the parameter combination that minimized the mean squared error. The final hyperparameters selected for the analysis are detailed in Table 2.

Table 2. Optimal hyperparameters identified through Grid Search CV.

Model	Hyperparameter	Optimal Value
Tuned Random Forest	n_estimators	200
	max_depth	5
	min_samples_split	5
	min_samples_leaf	10
	max_features	0.7
Tuned XGBoost	n_estimators	100
	max_depth	7
	learning_rate	0.01
	subsample	0.7

3. Results

3.1. Outage Frequency and Lines

In this paper, we counted the outage events and frequent lines in each region to better understand the outage distribution characteristics and line structure characteristics of different geographical units in the study area, and the results are shown in Figure 5. Figure 5a shows the outage frequency and the average outage duration in rural, urban, and city regions. The results show that the number of outages in rural areas is much higher than that in towns and urban areas, reaching approximately 380, and the average duration of the outages is also the longest, which indicates that the rural grid still has obvious shortcomings in terms of power supply stability and restoration efficiency. In contrast, the number of outages in towns and urban areas is significantly lower, and the average duration of outages is relatively shorter, which reflects better power grid protection capabilities.

Figure 5b–d further show the statistics of the number of outages of typical lines within each region. In rural areas (Figure 5b), Mingshan Mingping Line XPC had the highest number of outages, numbering close to 50, which is substantially higher than other lines. This shows that it may have structural weaknesses or operate under high load pressures within the rural grid; in addition, the frequency of outages for a number of lines is more than 10 times, which indicates that the problem of outages in the rural area is more concentrated within some key lines. The urban area (Figure 5c), on the other hand, shows a relatively balanced distribution of outages, but Sycamore Xinfu Line XQH still experiences frequent outages, reaching 14 occurrences. The urban areas (Figure 5d) exhibited the lowest overall outage frequency, but certain trunk lines, such as Chongwen Xinxi Line XBF, also

experienced over 10 outage occurrences. In summary, there are significant differences in the frequency and distribution of outages by region: rural areas have the most concentrated and frequent problems, cities and towns have certain distribution characteristics, and urban areas exhibit good overall performance but experience potentially localized problems.

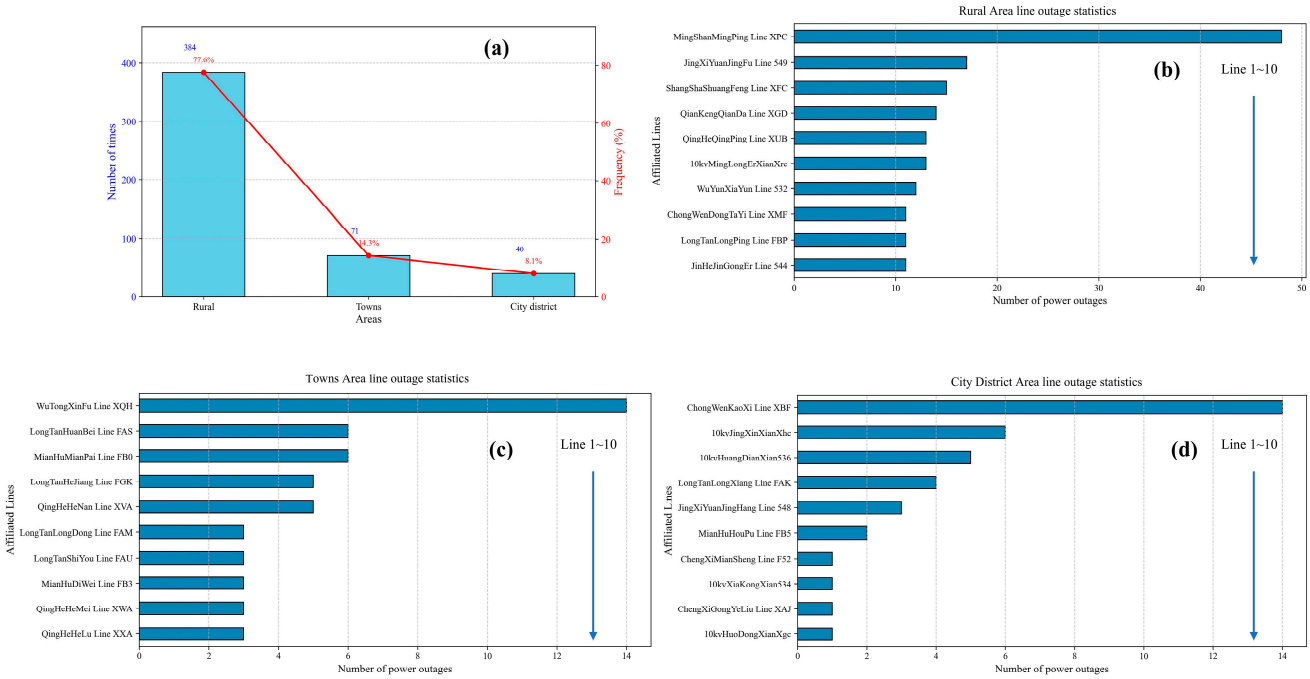


Figure 5. Frequency and line statistics of power outages in different areas: (**a**) Overall outage frequency and average duration in rural, town, and city districts; (**b**) Top 10 lines with the most frequent outages in rural areas; (**c**) Top 10 lines with the most frequent outages in town areas; (**d**) Top 10 lines with the most frequent outages in city districts.

3.2. "Two-Hour" Reliability Demonstration and Key Index Analysis

This study's preliminary statistical analysis is grounded in a real-world engineering objective. The Jiexi Power Supply Bureau, as part of a broader initiative to enhance service quality in rural areas, has established a "2-Hour High-Reliability Demonstration County" project. This project sets an ambitious operational target of reducing outage restoration times to under two hours. This target is particularly challenging given that the average customer outage duration (SAIDI) in this mountainous region was 4.16 h in 2023. The "2-h" mark, therefore, is not an arbitrary threshold but a critical, policy-driven benchmark for reliability performance.

To provide direct, actionable insights for this specific engineering initiative, an initial univariate analysis was conducted. The goal was to answer a practical question: "Which individual structural characteristics are most strongly associated with the failure to meet this two-hour target?"

To this end, outage events were bifurcated into two groups: those with durations less than or equal to two hours (target met) and those exceeding two hours (target not met). An independent-sample *t*-test was then conducted for each of the 34 structural indicators to compare the means between these two groups. Crucially, to mitigate the risk of Type I errors arising from multiple comparisons, the resulting *p*-values were corrected using the Benjamini–Hochberg False Discovery Rate (FDR) procedure.

The analysis revealed that, after FDR correction, six indicators exhibited a statistically significant difference between the two groups (FDR-adjusted, $p < 0.05$). The complete results of all 34 tests are provided in Appendix A (Table A1). The significant indicators included

factors related to network redundancy (e.g., Number of Upstream Ring Connections, Has Ring Main Units (RMUs)) and operational stress (e.g., Load Factor, Safe Operating Current).

These *t*-test results offer valuable, standalone guidance for the "2-Hour" reliability initiative. They highlight that specific, individual structural improvements are statistically associated with achieving the desired restoration target. However, this univariate approach does not account for the collective impact and potential interactions among these variables. Therefore, to build a more comprehensive understanding and establish a hierarchical ranking, a multivariate machine learning framework was subsequently employed.

3.3. Identification of Key Influencing Indicators via Consensus-Based Feature Importance

In this study, to robustly identify the key structural indicators influencing outage restoration time, a multi-faceted analytical approach was employed, leveraging the complementary strengths of three distinct feature evaluation methods: the importance feature from the validated Tuned Random Forest (RF) model, the coefficient magnitudes from a weakly penalized Lasso Regression, and the iterative rankings from Recursive Feature Elimination (RFE). The predictive performance of the machine learning models was rigorously validated on a held-out test set to ensure the reliability of the subsequent feature importance analysis. Table 3 summarizes the evaluation metrics for the Tuned Random Forest, XGBoost, and Lasso CV models.

Table 3. Performance of tuned models on the held-out test set.

Model	R-Squared (R^2)	RMSE (Minutes)	MAE (Minutes)
Tuned Random Forest (RF)	0.2641	201.87	175.53
Tuned XGBoost	0.1547	192.94	171.79
Lasso CV (Baseline)	0.1779	194.86	172.12

The relatively low R^2 values are a result of the inherent complexity of static outage data and static grid network indicators, especially when the sample size is limited and there are many unobserved influences (e.g., specific fault causes, weather conditions). Nonetheless, it also shows that the model captures generalizable patterns in the data. Critically, the positive predictive performance across these methodologically diverse models provides a crucial validation, confirming that the structural indicators under investigation possess a quantifiable, non-random relationship with outage restoration time. This establishes a credible foundation for the subsequent consensus-based feature importance analysis.

The feature importance results derived from each method are presented in Figure 6a–c. To synthesize these diverse results and identify the most consistent and significant indicators, a consensus-based analysis was performed. A detailed breakdown of this analysis, which compares the top-ranked features from each model, is provided in Appendix A (Table A2). The Inter-Bus Loop Rate and the Proportion of Users on Inter-Bus Tie-Lines emerge as dominant predictors across all three models. This finding underscores the paramount importance of network topology and redundancy. Higher values in these indicators signify greater structural flexibility and more alternative power supply paths, which are crucial for accelerating service restoration following a fault. Concurrently, indicators of electrical stress, namely, the Peak Daily Load Current and Load Factor, were also consistently identified as highly influential. This confirms that lines operating closer to their design limits present greater operational challenges, linking effective load management directly to reliability improvements. Furthermore, the Number of LV Customers was identified as another robust factor, suggesting that the scale and complexity of the low-voltage network area served by a line also contribute significantly to restoration timelines. Furthermore, the consensus analysis (Appendix A, Table A2) reveals a consistent set of critical

factors that significantly impact restoration duration. Five structural indicators emerged as the most dominant predictors, demonstrating high importance across all three diverse methodologies: the Inter-Bus Loop Rate, the Proportion of Users on Inter-Bus Tie-Lines, the Peak Daily Load Current, the Load Factor, and the Number of LV Customers. The strong consensus on these features underscores their paramount importance. Concurrently, other indicators, such as the Inter-Bus Transferability Rate and Is Transferable Line (Substation Constrained), were identified by two of the three models, marking them as robust secondary factors.

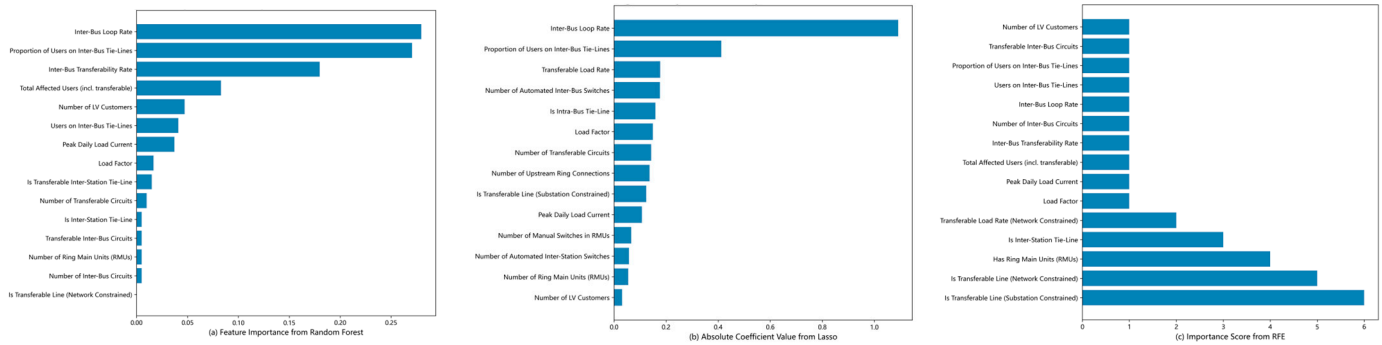


Figure 6. Feature importance ranking from different models: (**a**) RF; (**b**) Lasso; (**c**) RFE.

In summary, rather than relying on a single technique, the triangulation of these diverse methodologies provides strong, corroborative evidence. The findings collectively indicate that the most effective, data-driven strategies for reducing outage duration in the studied rural grid involve enhancing structural redundancy through looped configurations, implementing scientific load regulation to manage electrical stress, and optimizing network management for areas with large customer bases.

3.4. Exploring the Quantitative Relationship Between Key Indices and Outage Durations

According to the analysis of the fitted graphs, there is a positive correlation between the Load Factor, Peak Daily Load Current, Inter-Bus Loop Rate, the Proportion of Users on Inter-Bus Tie-Lines and the duration of outage. The R^2 values for the Load Factor and Peak Daily Load current are low. In contrast, the R^2 values for the Inter-Bus Loop Rate and the Proportion of Users, on Inter-Bus Tie-Lines are slightly higher. Moreover, although the explanatory power is still weak, the confidence intervals of the fitted straight lines are narrower, exhibiting some predictive reliability.

As shown in Figure 7, the relationships between outage duration and four representative network indicators were quantitatively analyzed using linear regression. Among them, the Peak Daily Load Current (Figure 7b) exhibits the strongest positive correlation with outage duration ($R^2 = 0.73$), suggesting that lines operating under higher peak loads are more prone to extended outages, potentially due to thermal stress, accelerated component aging, or increased difficulties in fault isolation under high-load conditions. The Load Factor (Figure 7a) also shows a moderate positive correlation ($R^2 = 0.59$), indicating that heavily loaded lines, particularly those approaching their design capacity, may experience longer restoration delays as a result of constrained switching options or safety concerns during re-energization. In contrast, the Inter-Bus Loop Rate for public lines (Figure 7c) reveals a negative correlation with outage duration ($R^2 = 0.37$), implying that higher network redundancy, characterized by looped configurations and alternate paths, facilitates faster fault recovery by enabling more flexible power rerouting. A similar but weaker negative trend is observed in the Proportion of Users on Inter-Bus Tie-Lines (Figure 7d, $R^2 = 0.23$), suggesting that user-side configuration and load distribution may modestly contribute

to resilience through improved fault segmentation and service restoration. The disparity in explanatory power across the four indicators highlights the multifactorial nature of outage recovery, where electrical load intensity and structural topology jointly influence the duration of service interruptions.

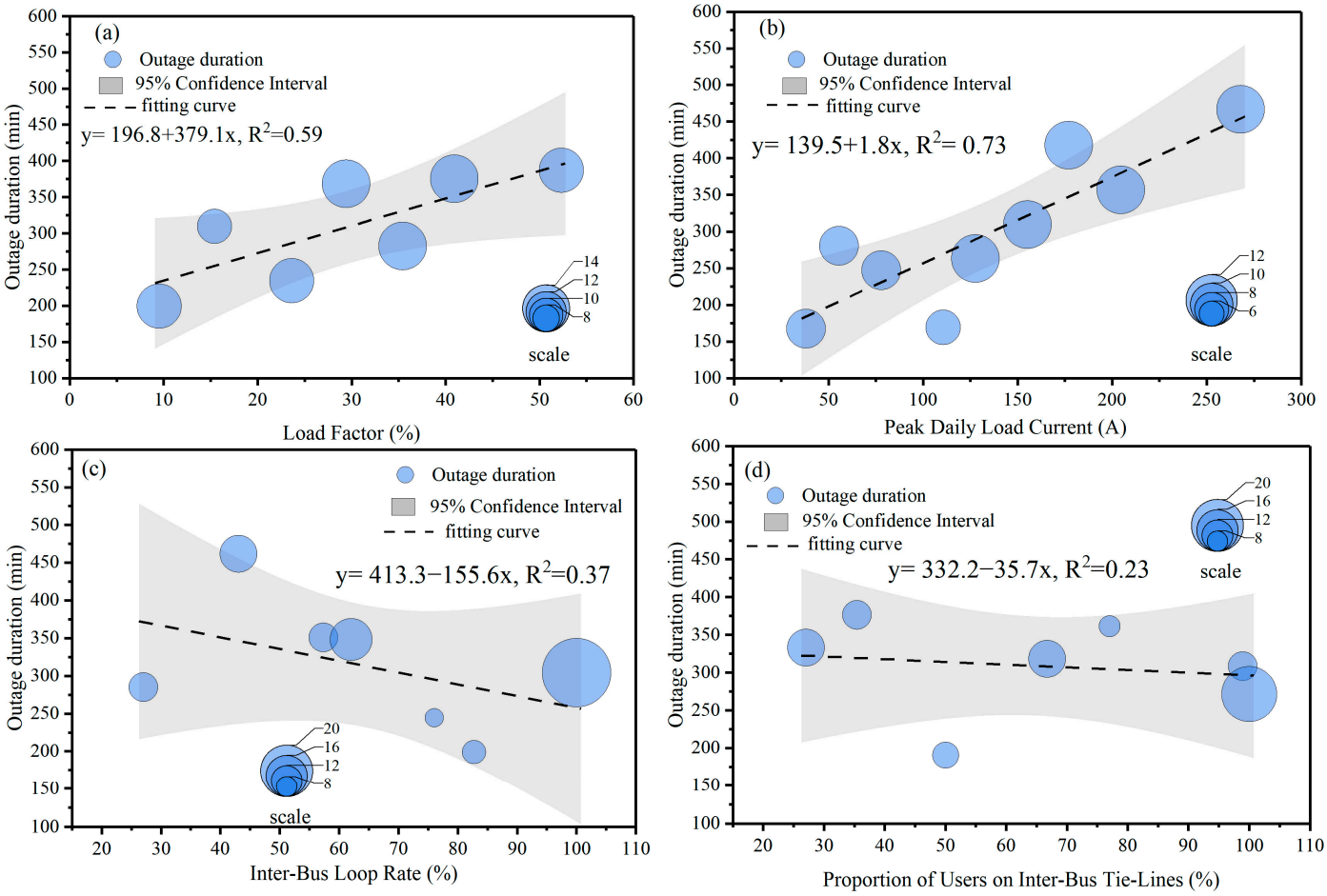


Figure 7. Relationship between key characteristic indicators and outage duration. The figure displays linear regression fits with 95% confidence intervals for four key indicators: (**a**) Load Factor (%); (**b**) Peak Daily Load Current (A); (**c**) Inter-Bus Loop Rate (%); and (**d**) Proportion of Users on Inter-Bus Tie-Lines (%).

According to the results of the Kernel Density Estimation (KDE) distribution plot (Figure 8), the duration of outages is significantly concentrated in the shorter range (0–250 min) with respect to a lower Load Factor and Peak Daily Load Current, indicating that these two metrics have a significant correlation with shorter outage durations. Moreover, the duration of outages is similarly concentrated in the shorter range (0–250 min) for higher Inter-Bus Loop Rates and higher Proportion of Users on Inter-Bus Tie-Lines. The positive effect of higher loop rates and user proportions on reducing outage durations is demonstrated by the fact that the outage durations are also concentrated in the shorter range (0–250 min). Overall, increasing the Loop Rate and the Proportion of Users and decreasing the Load Factor and the Peak Daily Load Current all contribute to a reduction in outage duration, indicating that these metrics have a significant impact on power system reliability.

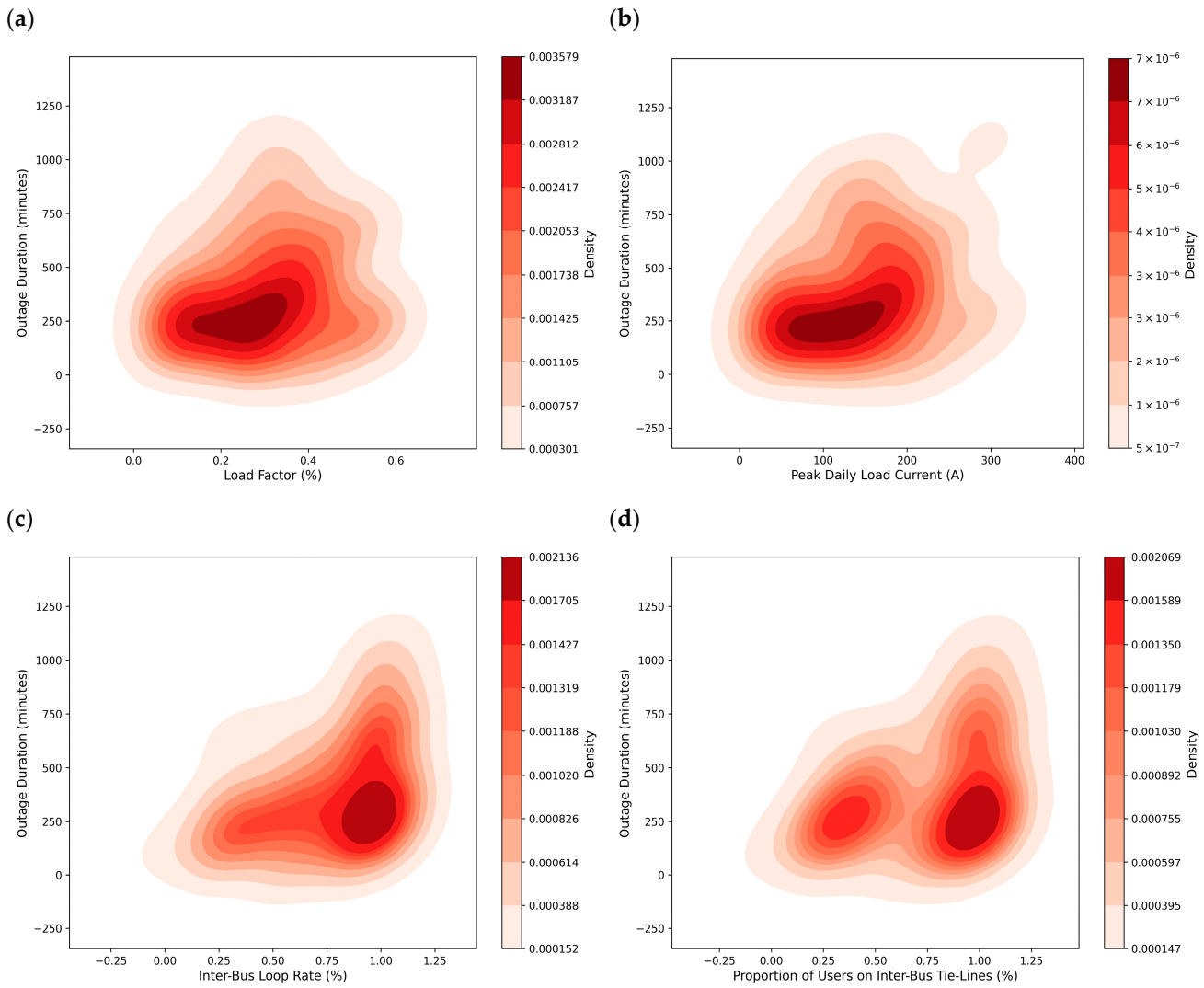


Figure 8. Distribution of key characteristic indicators vs. outage duration. The color intensity represents the density of data points. The subplots illustrate the relationship with: (**a**) Load Factor (%); (**b**) Peak Daily Load Current (A); (**c**) Inter-Bus Loop Rate (%); and (**d**) Proportion of Users on Inter-Bus Tie-Lines (%).

4. Discussion

4.1. Structural and Operational Challenges Undermining Grid Reliability

This study reveals that the reliability of rural power grids is significantly constrained by a combination of structural weaknesses, operational inefficiencies, and contextual limitations. First, the inherent randomness and unpredictability of power system failures, coupled with data incompleteness, obstruct accurate fault detection and real-time diagnostics [45]. These challenges are particularly prominent in rural networks, where equipment types are diverse, structural complexity is high, and fault risks are compounded by aging infrastructure [46].

Second, many rural grids suffer from outdated or suboptimal layouts due to limited investment. Problems such as overextended supply zone radii, insufficient conductor cross-sections, and low mutual substitution capacity have been widely reported [47]. These factors restrict the network's ability to isolate faults and reroute power, increasing both outage frequencies and restoration times. In addition, weak public awareness in remote communities has led to frequent external damage to grid assets (e.g., theft or vandalism), further undermining system stability [48].

Third, power supply restoration in these areas remains slow due to poor automation, low technical staffing levels, and inadequate emergency response capacity. Manual inspection and maintenance remain the norm, especially after faults occur, resulting in prolonged recovery periods. As noted by the authors of [49], insufficient maintenance tools, long line distances, and difficult terrain pose serious challenges, particularly during adverse weather conditions. These combined factors not only delay recovery, but also increase the overall operational burden on rural power utilities. These inherent challenges underscore the importance of both understanding the limitations of predictive indicators and contextualizing our findings within the broader resilience framework.

4.2. A Validated Structural Approach to Resilience: Interpretation, Context, and Limitations

The machine-learning-based feature importance analysis adopted in this study, which aligns with modern methods in power system diagnostics [21], demonstrates that certain structural variables, such as the Inter-Bus Loop Rate and Proportion of Users on Inter-Bus Tie-Lines, are statistically significant predictors of restoration times. The validity of these findings is substantiated by the predictive performance of the optimized Random Forest model, which achieved an R-squared (R^2) value of 0.2641 on a held-out test set. This confirms that the identified relationships are not merely correlational but reflect a genuine, quantifiable impact on grid performance. These findings are consistent with studies showing that grid resilience can be enhanced through better load balancing and network reconfiguration [50,51].

However, the modest R-squared value also indicates that a substantial portion of the variance in restoration times remains unexplained when using these structural indicators alone. To investigate the source of this remaining variance, a supplementary textual analysis was conducted on the "Dispatcher Process" logs to categorize the qualitative causes of faults. The analysis revealed that a substantial number of outages were directly attributed to external factors. As shown in Table 4, weather-related events were the most dominant cause, accounting for 46 recorded instances, with lightning strikes (24 instances) and thunderstorms (12 instances) being the primary triggers. External interference, such as tree encroachment (eight instances) and vehicle collisions (three instances), along with animal contact (17 instances), also constituted significant causes.

Table 4. Categorization of outages caused by external factors.

Fault Category	Sub-Category Example	Recorded Count
Weather-Related	Lightning, Thunderstorm, High Wind	46
External Interference	Tree Encroachment, Vehicle Collision, Kites	18
Animal Contact	Birds, Rats, Snakes	17

These findings provide a compelling explanation for the modest R-squared value. The inherent randomness and severity of events, such as lightning strikes, can result in highly variable damage and restoration time data, introducing noise that cannot be fully explained by pre-existing physical grid characteristics alone. This aligns with the broader challenge of modeling distribution networks with low observability and high structural variability, especially in under-resourced regions [39].

More importantly, this reinforces the significance of optimizing the grid's physical structure. A structurally robust grid—possessing higher Inter-Bus Loop Rates and better load distribution across Inter-Bus Tie-Lines (features identified as critical by the model)—is fundamentally more resilient. Such a grid may not prevent a lightning strike, but it can offer more alternative power paths, facilitate faster fault isolation, and ultimately result in

shorter restoration times. Therefore, improving the static physical resilience of the network serves as a crucial proactive defense against unpredictable external shocks.

This understanding of structural resilience as a foundational defense helps position the present study within the wider context of power system resilience research. Much of the existing literature focuses on enhancing resilience through "active" or "dynamic" measures. For instance, some studies develop proactive operational strategies for complex systems under uncertainty [52], while others explore advanced secondary control methods for microgrids [53] or specialized protection schemes for HVDC systems [54]. These approaches address the critical question of "how to operate" a system during and after a contingency.

In contrast, this study adopts a complementary "passive" or "structural" perspective. By leveraging historical outage data to identify the inherent physical attributes that predispose a network to prolonged outages, this study helps answer the foundational questions of "where to invest" and "what to build". This structural approach is particularly vital for long-term grid planning in rural areas. While advanced operational strategies are crucial, their effectiveness can be fundamentally limited by underlying structural weaknesses.

Concretely, these findings can directly inform power grid planning and investment priorities. For instance, instead of uniformly upgrading all rural lines, utility companies can use this data-driven framework to prioritize investments. Lines identified with a combination of low Inter-Bus Loop Rates and high user density should be targeted first for structural enhancements. From a cost–benefit perspective, the investment in an automated switch on a critical tie-line can be justified against the significant economic and social costs of extended blackouts. This approach allows for a more targeted and cost-effective allocation of limited capital resources, focusing on modifications projected to yield substantial marginal improvements in reliability.

Furthermore, the methodology from the Jiexi case study is highly extendable to other rural regions. While the precise importance ranking of features may vary depending on local geography and grid topology, the underlying principle remains the same: leveraging data to move from reactive maintenance to proactive, structurally focused grid reinforcement. This study provides a transferable methodological framework for such endeavors.

Therefore, while reliability metrics and predictive indicators remain indispensable tools [49], their utility is greatly enhanced when interpreted within the proper context. This study provides a foundation for such integration by identifying key structural weaknesses, and it suggests that strategic investments in improving physical grid infrastructure are an essential prerequisite for creating a truly resilient network.

4.3. Complementary Insights from Univariate and Multivariate Analyses

An important finding of this study emerges from the comparison between the univariate statistical tests (Table A1) and the multivariate machine learning analyses (Table A2). It was observed that several indicators identified as highly important by the consensus of ML models, such as Peak Daily Load Current and Number of LV Customers, did not exhibit a statistically significant difference in the independent-sample *t*-tests after FDR correction. This apparent discrepancy is not a contradiction but rather highlights the different, yet complementary, nature of these analytical approaches. The univariate *t*-tests are effective at identifying indicators with a strong, direct, and isolated linear relationship to prolonged outages, specifically in the context of the "2-Hour" reliability benchmark. In contrast, the multivariate ML models assess a feature's importance within the complex, interactive context of all other variables.

The high ranking of a feature like Peak Daily Load Current in the ML models, despite its non-significance in the *t*-test, suggests that its influence is likely context-dependent

or non-linear. For instance, high peak loads may only become critical determinants of restoration time when they occur on lines with low structural redundancy (e.g., low Inter-Bus Loop Rate). Such interaction effects are invisible to univariate tests but are effectively captured by the machine learning framework. Therefore, by employing both methods, this study provides a more holistic understanding: the *t*-tests confirm a baseline set of directly significant factors for a specific operational target, while the ML consensus analysis reveals a broader and more nuanced set of robust indicators whose importance is revealed through their collective impact.

5. Conclusions

This study investigated the key structural indicators affecting outage restoration times in a representative rural power grid, employing a validated, data-driven analytical framework. The main conclusions are as follows:

(1) The statistical analysis confirms that the studied rural power grid not only experiences significantly higher outage frequencies, but also markedly longer outage durations compared to adjacent urban and township grids. These disparities highlight persistent challenges related to structural limitations, inadequate automation, and constrained emergency response capabilities in underdeveloped regions.

(2) Through a multi-faceted, consensus-based feature importance analysis, validated by the positive predictive performance of the underlying machine learning models, this study successfully identified and ranked the most critical structural indicators. Five indicators demonstrated the highest level of consensus across all three diverse models: the Inter-Bus Loop Rate, Proportion of Users on Inter-Bus Tie-Lines, Peak Daily Load Current, Load Factor, and Number of LV Customers. This underscores that network topology, electrical stress, and customer density are paramount factors.

(3) The findings provide a robust, data-driven basis for strategic grid enhancement. The validated importance of these consensus indicators offers actionable guidance for rural power utilities, enabling a shift from uniform reactive maintenance to proactive, targeted investment. For instance, prioritizing capital expenditure on enhancing looped configurations and deploying automation on high-user-density and low-structural-redundancy lines can yield substantial improvements in service resilience.

(4) While the predictive models account for a modest but statistically significant portion of the variance, the remaining unexplained variance highlights the strong influence of stochastic external factors (e.g., weather events) and unobserved operational variables. This underscores a key direction for future research: the development of comprehensive, hybrid models that integrate both static structural data and dynamic event data to achieve a more holistic understanding of power system resilience.

Author Contributions: Conceptualization, X.G.; Methodology, J.L., R.X., H.L., Y.M. and Z.F.; Software, J.L. and R.X.; Investigation, J.L., R.X., H.L., Y.M. and Z.F.; Resources, J.L. and H.L.; Data curation, R.X.; Writing—original draft, J.L.; Writing—review & editing, X.G.; Supervision, X.G.; Project administration, J.L. and X.G.; Funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project "Research and Practice on Creating a Digital and Highly Reliable Demonstration County for '2-Hour' Power Supply in Remote Mountainous Areas", grant number GDKJXM20240758.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: Authors Jiajun Lin, Ruiyue Xie, Haobin Lin, and Xingyuan Guo were employed by the company Jieyang Power Supply Bureau, Guangdong Power Grid Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Detailed Results of Independent-Sample t-Tests and Consensus Analysis

Table A1. Comparison of structural indicator means between outages lasting $\leq$ 2 h and >2 h. The *p*-values are corrected for multiple comparisons using the Benjamini–Hochberg False Discovery Rate (FDR) procedure. Indicators are sorted by their uncorrected *p*-value.

Index	T-Statistic	*p*-Value (Uncorrected)	*p*-Value (FDR-Adjusted)	Significance (FDR)
Number of Upstream Ring Connections	−3.919	0.00018	0.00220	TRUE
Load Factor	3.902	0.00023	0.00271	TRUE
Is Transferable Line (Network-Constrained)	2.371	0.00028	0.00429	TRUE
Has Ring Main Units (RMUs)	3.723	0.00036	0.00306	TRUE
Is Dedicated Line (Feeder)	−3.351	0.00122	0.00828	TRUE
Safe Operating Current	−2.825	0.00593	0.03359	TRUE
Number of Automated Inter-Station Switches	2.627	0.02491	0.12099	FALSE
Peak Daily Load Current	1.393	0.23721	0.80654	FALSE
Is Transferable Line (Substation-Constrained)	1.421	0.22302	0.80654	FALSE
Number of Manual Switches in RMUs	0.833	0.46034	0.96605	FALSE
Transferable Load Rate (Substation-Constrained)	−0.776	0.48559	0.96605	FALSE
Transferable Circuits (Substation-Constrained)	0.694	0.52991	0.96605	FALSE
Transferable Circuits (Network-Constrained)	0.722	0.51414	0.96605	FALSE
Transferable Load Rate (Network-Constrained)	−0.683	0.53556	0.96605	FALSE
Number of Automated Inter-Bus Switches	0.656	0.55688	0.96605	FALSE
Number of Ring Main Units (RMUs)	−0.558	0.60565	0.96605	FALSE
Number of LV Customers	−0.526	0.62230	0.96605	FALSE
Number of LV Customers on Bus	−0.526	0.62230	0.96605	FALSE
Is Intra-Bus Tie-Line	0.513	0.64215	0.96605	FALSE
Inter-Bus Loop Rate	−0.459	0.67534	0.96605	FALSE
Transferable Rate of Inter-Bus Lines	−0.299	0.78296	0.98351	FALSE
Number of Automated Inter-Bus Switches	−1.599	0.18837	0.80056	FALSE
Users on Inter-Bus Tie-Lines	−0.629	0.56303	0.96605	FALSE
Number of Inter-Bus Circuits	0.368	0.73349	0.98351	FALSE
Transferable Inter-Bus Circuits	0.458	0.67306	0.96605	FALSE
Transferable LV Customers on Inter-Bus Lines	−0.440	0.68192	0.96605	FALSE
Number of Automated Inter-Station Switches	−2.917	0.01326	0.08065	FALSE
Proportion of Users on Inter-Bus Tie-Lines	−0.108	0.92018	0.98351	FALSE
Is Transferable Inter-Station Tie-Line	0.061	0.95458	0.98351	FALSE
Is Inter-Station Tie-Line	−0.093	0.63745	0.87062	FALSE
Affected Users on Bus (Incl. Transferable)	−0.015	0.98865	0.98865	FALSE
Is Transferable Inter-Bus Tie-Line	0.154	0.88700	0.98351	FALSE
Affected Users on Bus (Incl. Tie-Lines)	0.104	0.92343	0.98351	FALSE
Proportion of Transferable Users on Inter-Station	0.133	0.90230	0.98351	FALSE

Table A2. Consensus analysis of top-ranked structural indicators from different models. This table compares the top 15 features identified by the Random Forest (RF), Lasso model, and Recursive Feature Elimination (RFE). A checkmark (✔) indicates that a feature was identified as a top indicator by the respective model and the final column summarizes the consensus level, while a cross (✖) indicates it was not. A feature is marked with a Consensus Level based on how many models identified it as a top indicator.

Structural Indicator (Feature Name)	Identified by RF (Top 15)	Identified by Lasso (Top 15)	Identified by RFE (Top 15)	Consensus Level
Inter-Bus Loop Rate	✔	✔	✔	3 Models (Highest Consensus)
Proportion of Users on Inter-Bus Tie-Lines	✔	✔	✔	3 Models (Highest Consensus)
Number of LV Customers	✔	✔	✔	3 Models (Highest Consensus)
Peak Daily Load Current	✔	✔	✔	3 Models (Highest Consensus)
Load Factor	✔	✔	✔	3 Models (Highest Consensus)
Inter-Bus Transferability Rate	✔	✖	✔	2 Models (RF, RFE)
Total Affected Users (incl. Transferable)	✔	✖	✔	2 Models (RF, RFE)
Users on Inter-Bus Tie-Lines	✔	✖	✔	2 Models (RF, RFE)
Is Transferable Line (Network Constrained)	✔	✖	✔	2 Models (RF, RFE)
Number of Transferable Circuits	✔	✔	✖	2 Models (RF, Lasso)
Number of Ring Main Units (RMUs)	✔	✔	✖	2 Models (RF, Lasso)
Is Inter-Station Tie-Line	✔	✖	✔	2 Models (RF, RFE)
Transferable Inter-Bus Circuits	✔	✖	✔	2 Models (RF, RFE)
Number of Inter-Bus Circuits	✔	✖	✔	2 Models (RF, RFE)
Transferable Load Rate	✖	✔	✖	1 Model (Lasso)
Is Transferable Line (Substation Constrained)	✖	✔	✔	2 Models (Lasso, RFE)
Number of Automated Inter-Bus Switches	✖	✔	✖	1 Model (Lasso)
Is Intra-Bus Tie-Line	✖	✔	✖	1 Model (Lasso)
Number of Upstream Ring Connections	✖	✔	✖	1 Model (Lasso)
Number of Manual Switches in RMUs	✖	✔	✖	1 Model (Lasso)
Number of Automated Inter-Station Switches	✖	✔	✖	1 Model (Lasso)
Has Ring Main Units (RMUs)	✖	✖	✔	1 Model (RFE)
Transferable Load Rate (Network Constrained)	✖	✖	✔	1 Model (RFE)

References

1. China Unveils 2024 Guiding Opinions on Energy Work-Newsletter-AllBright Law Offices. Available online: https://www.allbrightlaw.com/EN/10531/b687d58ecace421d.aspx (accessed on 30 May 2025).
2. Full Text: China's Energy Transition | english.scio.gov.cn. Available online: http://english.scio.gov.cn/whitepapers/2024-08/29/content_117394384_7.htm (accessed on 30 May 2025).

3. China Releases Guideline on Strengthening Integration of NEVs with Power Grid. Available online: https://english.www.gov.cn/news/202401/04/content_WS659695e2c6d0868f4e8e2c2e.html (accessed on 30 May 2025).
4. Wu, M.Y.; Ridzuan, M.I.; Djokic, S.Z. Smart grid functionalities for improving reliability of rural electricity networks, Transmission, Distribution and Energy Conversion (MedPower 2016). In Proceedings of the Mediterranean Conference on Power Generation, Belgrade, Serbia, 6–9 November 2016; pp. 1–7. [CrossRef]
5. Silva, N.S.e.; Castro, R.; Ferrão, P. Smart Grids in the Context of Smart Cities: A Literature Review and Gap Analysis. *Energies* **2025**, *18*, 1186. [CrossRef]
6. Abdelmalak, M.; Cox, J.; Ericson, S.; Hotchkiss, E.; Benidris, M. Quantitative Resilience-Based Assessment Framework Using EAGLE-I Power Outage Data. *IEEE Access* **2023**, *11*, 7682–7697. [CrossRef]
7. Yao, Y.; Liu, W.; Jain, R.; Chowdhury, B.; Wang, J.; Cox, R. Quantitative Metrics for Grid Resilience Evaluation and Optimization. *IEEE Trans. Sustain. Energy* **2023**, *14*, 1244–1258. [CrossRef]
8. Stanković, A.M.; Tomsovic, K.L.; De Caro, F.; Braun, M.; Chow, J.H.; Čukalevski, N.; Dobson, I.; Eto, J.; Fink, B.; Hachmann, C.; et al. Methods for Analysis and Quantification of Power System Resilience. *IEEE Trans. Power Syst.* **2023**, *38*, 4774–4787. [CrossRef]
9. Almasoudi, F.M. Enhancing Power Grid Resilience through Real-Time Fault Detection and Remediation Using Advanced Hybrid Machine Learning Models. *Sustainability* **2023**, *15*, 8348. [CrossRef]
10. Jiang, H.; Zhang, J.J.; Gao, W.; Wu, Z. Fault Detection, Identification, and Location in Smart Grid Based on Data-Driven Computational Methods. *IEEE Trans. Smart Grid* **2014**, *5*, 2947–2956. [CrossRef]
11. Maheshwari, Z.; Ramakumar, R. Using SIRES to Enhance Resilience in Remote & Rural Communities. *J. Energy Power Technol.* **2022**, *4*, 6. [CrossRef]
12. Majidi, M.; Etezadi-Amoli, M.; Fadali, M.S. A Sparse-Data-Driven Approach for Fault Location in Transmission Networks. *IEEE Trans. Smart Grid* **2017**, *8*, 548–556. [CrossRef]
13. Rizeakos, V.; Bachoumis, A.; Andriopoulos, N.; Birbas, M.; Birbas, A. Deep learning-based application for fault location identification and type classification in active distribution grids. *Appl. Energy* **2023**, *338*, 120932. [CrossRef]
14. Ghosh, P.; De, M. A comprehensive survey of distribution system resilience to extreme weather events: Concept, assessment, and enhancement strategies. *Int. J. Ambient Energy* **2022**, *43*, 6671–6693. [CrossRef]
15. Mujjuni, F.; Betts, T.R.; Blanchard, R.E. Evaluation of Power Systems Resilience to Extreme Weather Events: A Review of Methods and Assumptions. *IEEE Access* **2023**, *11*, 87279–87296. [CrossRef]
16. Ren, H.; Hou, Z.J.; Ke, X.; Huang, Q.; Makatov, Y. Analysis of Weather and Climate Extremes Impact on Power System Outage. In Proceedings of the 2021 IEEE Power & Energy Society General Meeting (PESGM), Washington, DC, USA, 26–29 July 2021; pp. 1–5. [CrossRef]
17. Panteli, M.; Mancarella, P. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electr. Power Syst. Res.* **2015**, *127*, 259–270. [CrossRef]
18. Flores, N.M.; Northrop, A.J.; Do, V.; Gordon, M.; Jiang, Y.; Rudolph, K.E.; Hernández, D.; Casey, J.A. Powerless in the storm: Severe weather-driven power outages in New York State, 2017–2020. *PLoS Clim.* **2024**, *3*, e0000364. [CrossRef]
19. Strielkowski, W.; Vlasov, A.; Selivanov, K.; Muraviev, K.; Shakhnov, V. Prospects and Challenges of the Machine Learning and Data-Driven Methods for the Predictive Analysis of Power Systems: A Review. *Energies* **2023**, *16*, 4025. [CrossRef]
20. Doostan, M.; Chowdhury, B.H. A data-driven analysis of outage duration in power distribution systems. In Proceedings of the 2017 North American Power Symposium (NAPS), Morgantown, WV, USA, 17–19 September 2017; pp. 1–6. [CrossRef]
21. Ghasemkhani, B.; Kut, R.A.; Yilmaz, R.; Birant, D.; Arıkök, Y.A.; Güzelyol, T.E.; Kut, T. Machine Learning Model Development to Predict Power Outage Duration (POD): A Case Study for Electric Utilities. *Sensors* **2024**, *24*, 4313. [CrossRef]
22. Arif, A.; Wang, Z. Distribution Network Outage Data Analysis and Repair Time Prediction Using Deep Learning. In Proceedings of the 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Boise, ID, USA, 24–28 June 2018; pp. 1–6. [CrossRef]
23. El Mrabet, Z.; Sugunaraj, N.; Ranganathan, P.; Abhyankar, S. Random Forest Regressor-Based Approach for Detecting Fault Location and Duration in Power Systems. *Sensors* **2022**, *22*, 458. [CrossRef]
24. Theng, D.; Bhoyar, K.K. Feature selection techniques for machine learning: A survey of more than two decades of research. *Knowl. Inf. Syst.* **2024**, *66*, 1575–1637. [CrossRef]
25. Dhal, P.; Azad, C. A comprehensive survey on feature selection in the various fields of machine learning. *Appl. Intell.* **2022**, *52*, 4543–4581. [CrossRef]
26. Huang, N.; Lu, G.; Xu, D. A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest. *Energies* **2016**, *9*, 767. [CrossRef]
27. Lee, Y.-G.; Oh, J.-Y.; Kim, D.; Kim, G. SHAP Value-Based Feature Importance Analysis for Short-Term Load Forecasting. *J. Electr. Eng. Technol.* **2023**, *18*, 579–588. [CrossRef]
28. Li, M.; Wang, Y. Power load forecasting and interpretable models based on GS_XGBoost and SHAP. *J. Phys. Conf. Ser.* **2022**, *2195*, 012028. [CrossRef]

29. Yaprakdal, F.; Varol Arısoy, M. A Multivariate Time Series Analysis of Electrical Load Forecasting Based on a Hybrid Feature Selection Approach and Explainable Deep Learning. *Appl. Sci.* **2023**, *13*, 12946. [CrossRef]
30. Janiszewski, P.; Sawicki, J.; Kurpas, J.; Mróz, M. Practical Ways to Improve SAIDI and SAIFI Power Supply Reliability Indicators in an MV Grid. *Acta Energ.* **2018**, 45–50. [CrossRef]
31. Kunaifi; Reinders, A. Perceived and Reported Reliability of the Electricity Supply at Three Urban Locations in Indonesia. *Energies* **2018**, *11*, 140. [CrossRef]
32. Kozyra, J.; Łukasik, Z.; Kuśmińska-Fijałkowska, A.; Kaszuba, P. The impact of selected variants of remote control on power supply reliability indexes of distribution networks. *Electr. Eng.* **2022**, *104*, 1255–1264. [CrossRef]
33. Guo, S.; Zhao, H.; Zhao, H. The Most Economical Mode of Power Supply for Remote and Less Developed Areas in China: Power Grid Extension or Micro-Grid? *Sustainability* **2017**, *9*, 910. [CrossRef]
34. Rojas-Zerpa, J.C.; Yusta, J.M. Application of multicriteria decision methods for electric supply planning in rural and remote areas. *Renew. Sustain. Energy Rev.* **2015**, *52*, 557–571. [CrossRef]
35. Sami, N.M.; Naeini, M. Machine learning applications in cascading failure analysis in power systems: A review. *Electr. Power Syst. Res.* **2024**, *232*, 110415. [CrossRef]
36. Manninen, H.; Kilter, J.; Landsberg, M. Health Index Prediction of Overhead Transmission Lines: A Machine Learning Approach. *IEEE Trans. Power Deliv.* **2022**, *37*, 50–58. [CrossRef]
37. Yazdanpanah, Z.; Rastegar, M.; Jooshaki, M. Determining target levels of power distribution system reliability indices using machine learning. *Electr. Power Syst. Res.* **2024**, *233*, 110456. [CrossRef]
38. Zhu, H.; Giannakis, G.B. Lassoing line outages in the smart power grid. In Proceedings of the 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), Brussels, Belgium, 17–20 October 2011; pp. 570–575. [CrossRef]
39. Arora, P.; Ceferino, L. Probabilistic and machine learning methods for uncertainty quantification in power outage prediction due to extreme events. *EGUsphere* **2022**, *2022*, 1–29. [CrossRef]
40. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
41. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1996**, *58*, 267–288. [CrossRef]
42. Muthukrishnan, R.; Rohini, R. LASSO: A feature selection technique in predictive modeling for machine learning. In Proceedings of the 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, India, 24 October 2016; pp. 18–20. [CrossRef]
43. Jeon, H.; Oh, S. Hybrid-Recursive Feature Elimination for Efficient Feature Selection. *Appl. Sci.* **2020**, *10*, 3211. [CrossRef]
44. Darst, B.F.; Malecki, K.C.; Engelman, C.D. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genet.* **2018**, *19*, 65. [CrossRef] [PubMed]
45. Zheng, K.; Du, W.; Wang, H. Dynamic Interaction and Stability Analysis of Hybrid Multi-Microgrid Systems. *Proc. Chin. Soc. Electr. Eng.* **2021**, *41*, 5552–5568.
46. Zhang, H.; He, G. A Discussion on the Power Supply Reliability of Rural Distribution Networks and Its Improvement Measures. *China Electr. Power Educ.* **2011**, *8*, 87–88.
47. Song, Y.; Zhang, D.; Wu, J.; Peng, D.; Liang, C.; Qiu, Y.; Chen, Z.; Wu, Q.; Cao, J. Comparative Analysis of Power Supply Reliability in Urban Distribution Networks at Home and Abroad. *Power Syst. Technol.* **2008**, *32*, 13–18.
48. Kang, Z. An Analysis of Measures to Improve the Reliability of Power Supply in Distribution Networks. *Front. Electr. Energy* **2024**, *2*, 118–120.
49. Wang, W.; Zhang, T.; Xuan, W.; Li, H.; Liu, Z.; Wang, K. A Power Grid Planning Method Considering Unit Commitment and Network Structure Optimization. *J. Power Syst. Autom.* **2021**, *33*, 108–115.
50. Li, J.; Xie, Y.; Zeng, H.; Zhou, S.; Luo, Y.; Song, J. A Review of Research on Optimization Scheduling under Uncertainty and Its Application in New Type Power Systems. *High Volt. Eng.* **2022**, *48*, 3447–3464.
51. Mitsova, D.; Li, Y.; Einsteder, R.; Roberts Briggs, T.; Sapat, A.; Esnard, A.-M. Using Nighttime Light Data to Explore the Extent of Power Outages in the Florida Panhandle after 2018 Hurricane Michael. *Remote Sens.* **2024**, *16*, 2588. [CrossRef]
52. Yang, X.; Liu, X.; Li, Z.; Xiao, G.; Wang, P. Resilience-oriented proactive operation strategy of coupled transportation power systems under exogenous and endogenous uncertainties. *Reliab. Eng. Syst. Saf.* **2025**, *262*, 111161. [CrossRef]
53. Li, X.; Hu, C.; Luo, S.; Lu, H.; Piao, Z.; Jing, L. Distributed hybrid-triggered observer-based secondary control of multi-bus DC microgrids over directed networks. *IEEE Trans. Circuits Syst. Regul. Pap.* **2025**, *72*, 2467–2480. [CrossRef]
54. Tiwari, R.S.; Sharma, J.P.; Gupta, O.H.; Ahmed Abdullah Sufyan, M. Extension of pole differential current based relaying for bipolar LCC HVDC lines. *Sci. Rep.* **2025**, *15*, 16142. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.