

ORIGINAL RESEARCH

Open access

Study on the Evaluation and Management of Electric Vehicle DC Charging Stations Using XGBoost

Gabriel Costa^{1*}, Rafael Mendes¹, Bruno Teixeira², Lucas Ribeiro¹

Abstract

The rapid expansion of electric vehicle (EV) adoption has led to a swift growth in charging infrastructure, accompanied by a rising incidence of malfunctions in charging equipment. To enhance the management and maintenance of these facilities, this study introduces an approach using the XGBoost algorithm, specifically tailored for direct current (DC) EV charging stations. A comprehensive operation and maintenance framework is developed, grounded in thorough condition evaluation. This framework incorporates the current functional state of the stations alongside patterns in user charging behavior. Key considerations include drivers' traveling and recharging preferences, prevailing road congestion levels, and the condition of charging hardware. The model is trained on a dataset compiled from records, live surveillance information, and ambient external conditions. Employing the XGBoost technique, a robust evaluation model assesses the operational health of charging stations. Drawing from these evaluation outcomes and integrating relevant fault-related metrics, a dedicated risk evaluation framework is formulated. By using the energy not charged (ENC) metric, the system performs ongoing risk monitoring for each facility, assigns risk-priority rankings, and establishes a sequence for necessary interventions. An advanced optimization model for DC station maintenance is then constructed, balancing financial efficiency with system reliability objectives, ultimately generating an effective maintenance schedule across stations. Case study findings indicate that this combined condition assessment and strategic maintenance plan markedly enhances overall system dependability and delivers superior economic returns from upkeep efforts, all while satisfying established performance criteria.

Keywords DC charging station, XGBoost, Traffic simulation, Operation and maintenance

*Correspondence:

Gabriel Costa
gabriel.costa@ufrj.br

¹ Department of Grid Optimization Engineering, Faculty of Engineering, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

² Department of Smart Energy Systems, Faculty of Engineering, University of Campinas, Campinas, Brazil

Introduction

Electric vehicles (EVs) are experiencing swift growth as an environmentally friendly, emissions-free transport option that helps mitigate pollution. A dependable network of charging facilities is essential to support widespread EV deployment. In contrast to alternating current (AC) chargers, DC charging units exhibit elevated malfunction rates, attributable to their greater complexity, higher power demands, and prolonged exposure to outdoor elements. However, initial efforts in upkeep and management have begun; challenges persist, including imprecise scheduling, heavy reliance on subjective judgment, and poorly defined priorities. Precisely gauging the present condition and future trajectory of charging equipment, implementing proactive interventions, minimizing breakdowns, and boosting operational performance remain critical priorities [1].

Existing studies on the maintenance of DC EV charging stations primarily focus on real-time surveillance of charging infrastructure. For instance, Lin *et al.* [2] analyzed requirements and designed a surveillance platform for charging setups, structured across multiple tiers—including a core platform, foundational services, shared utilities, and user-facing applications—to oversee aspects such as power allocation, charging processes, protective measures, and metering. Arispe and Mombello [3] developed models for the station architecture, device details, and interaction protocols, creating a blueprint for monitoring and tracking vital parameters of station components. In Fang *et al.* [4], a PCI bus-based real-time monitoring setup was created to gather instantaneous voltage and current readings at key connection points, enabling analysis of device performance and traits.

These works primarily address upkeep from a surveillance system standpoint. Nevertheless, limited attention has been given to integrating station monitoring insights with EV users' actual recharging patterns to develop reliable maintenance plans. Risk-based maintenance (RBM) involves a numerical evaluation of the impacts stemming from device breakdowns, serving as a foundation for crafting maintenance approaches [5]. RBM has gained widespread use across various sectors. For example, Heo *et al.* [6] introduced an enhanced semi-Markov process-based state representation and derived optimal maintenance strategies to reduce premature service costs and reliability degradation. Feng *et al.* [7] constructed a predictive framework using Bayesian classification based on past service logs and failure histories, defining risk in terms of dependability and cost factors, and optimizing for the lowest overall risk in maintaining 27.5 kV vacuum breakers. Peng [8] and Liang *et al.* [9] outlined a risk evaluation technique for automated upkeep of high-voltage lines, leveraging segmented pre-whitening fuzzy forecasting, big-data categorization, and fuzzy grouping to facilitate risk analysis and upkeep refinement. For distribution assets, real-time average failure probabilities are derived via feeder segmentation based on equipment health scores, leading to tailored service plans rooted in per-device risk-related financial losses [10–12].

Prior investigations have developed surveillance architectures for charging stations through system, equipment, and communication modeling, enabling monitoring of essential facility metrics and resulting in experience-driven maintenance practices. This work advances the field by integrating charging-station observational data with EV user recharging activity data. It assembles a comprehensive training dataset from archival records, continuous monitoring feeds, and surrounding environmental inputs. While earlier approaches have applied large-scale data categorization and fuzzy grouping for risk evaluation and maintenance improvement, such fault groupings tend to be broad, failing to capture nuanced distinctions among issues and leading to suboptimal resource allocation for repairs.

Here, the XGBoost algorithm is applied to develop a precise condition evaluation model for charging stations. Additionally, models simulating road traffic and driver behavior are incorporated. Through risk surveillance, maintenance prioritization for individual stations is determined, and an optimized upkeep framework for charging stations is proposed, with dual goals of minimizing costs and maximizing reliability gains.

Operation and Maintenance System of Electric Vehicle Charging Station

Beyond assessing internal functionality, evaluating charging station conditions must encompass usage dynamics, including customer routines, surrounding transportation networks, weather conditions, climatic variations, and other fluctuating factors. By synthesizing live traffic updates, charging-station equipment performance metrics, and driver mobility patterns derived from crowdsourced sensing, a holistic maintenance and management system emerges. This integrated approach addresses both infrastructure aspects and user-related factors, as depicted in **Figure 1**.

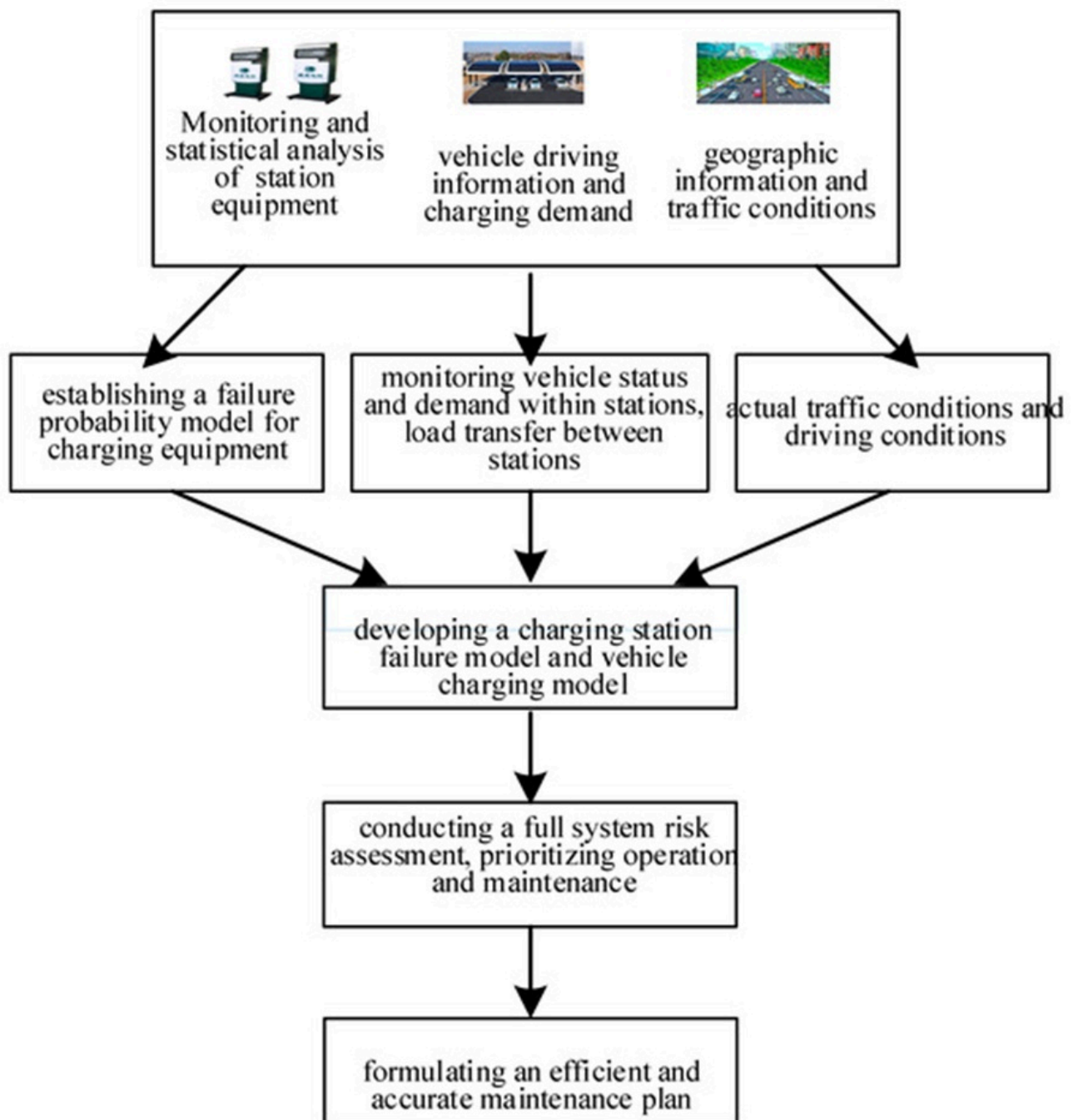


Figure 1. Framework for the operation and maintenance system of electric vehicle charging stations.

The core objective of operations and maintenance activities is to reduce the frequency of malfunctions in power equipment, improve overall reliability, and minimize adverse effects on end users. The operational risk associated with a charging pile arises from a combination of its built-in reliability characteristics and the extent to which its current working condition disrupts users throughout the broader system. Performing detailed condition assessments enables the refinement of operation and maintenance approaches, leading to more rational and effective management of charging piles and thereby maximizing the value derived from upkeep efforts.

Charging stations are classified into four discrete operational conditions: normal, attention, abnormal, and severe. In the normal condition, all monitored parameters remain comfortably within their ideal and stable ranges, supporting

uninterrupted functioning. The attention condition occurs when parameters approach—but do not yet breach—the acceptable thresholds, meaning the station can continue operating while calling for closer observation. In the abnormal condition, certain parameters exceed the defined limits by a moderate amount, requiring ongoing surveillance and carefully timed maintenance interruptions. The severe condition is characterized by parameters that substantially exceed the allowable thresholds, requiring urgent shutdown and immediate repair actions.

The overall modeling approach comprises two main components: a charging-station condition evaluation module and an operations and maintenance optimization module. The first module focuses on determining the health status of individual charging stations. In contrast, the second module leverages these precise status determinations to coordinate maintenance across multiple charging stations within the network.

Through iterative refinement and assessment using a tree-based ensemble method, the model identifies the state category with the greatest predicted probability, which is taken as the most accurate representation of the station's true operating condition.

Fundamentals of XGBoost

Boosting is a family of ensemble learning methods used for classification and regression tasks. Its central concept is to build a strong predictive model by sequentially combining and appropriately weighting a series of weaker base learners. This iterative combination process markedly improves the accuracy and robustness of the final model relative to any single weak learner, as shown conceptually in [Figure 2](#).

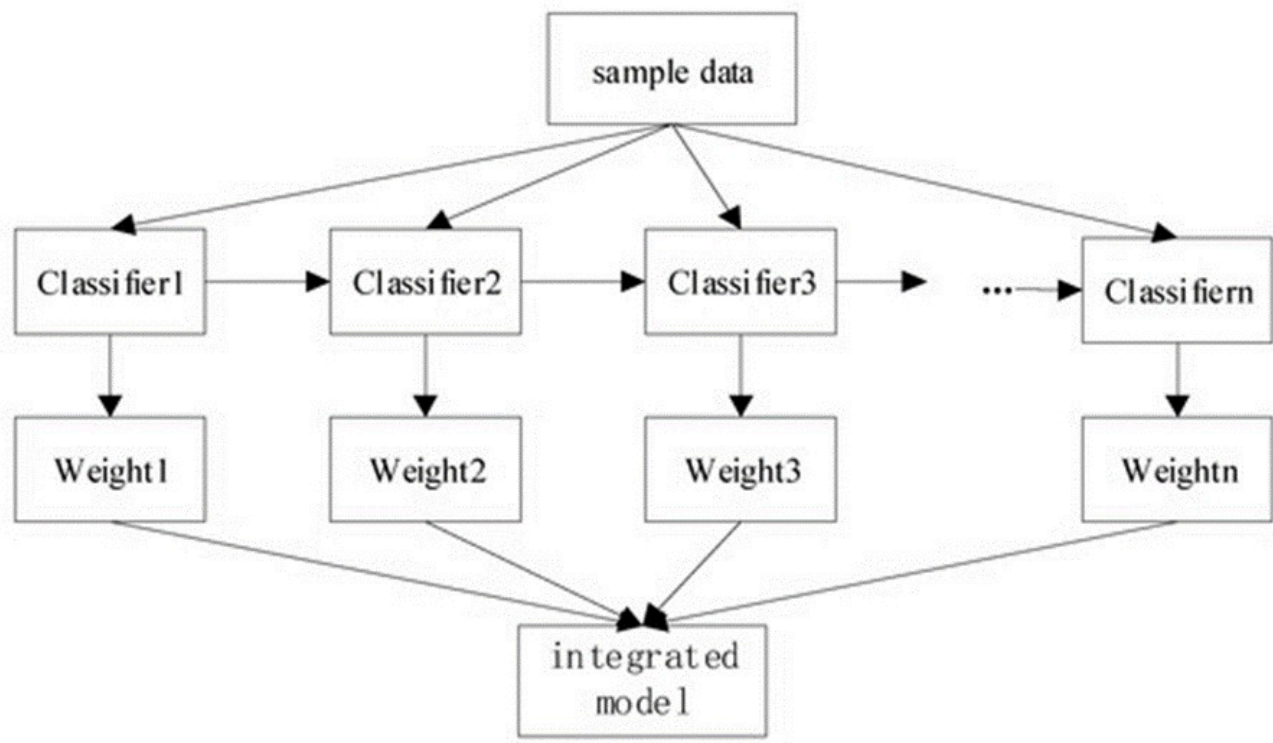


Figure 2. Core mechanism of boosting.

XGBoost represents an advanced iteration of the gradient boosting decision tree (GBDT) framework. Its key theoretical foundations are outlined below [\[13-16\]](#):

(1) Base learner: Regression trees: The essential building block of boosted tree ensembles is the regression tree, commonly known as a CART (Classification and Regression Tree). Each CART tree maps input features to leaf nodes, where every leaf is associated with a continuous numeric score rather than a class label.

(2) Model structure and tree complexity: A single regression tree can be decomposed into its structure and its leaf weights. The output of the t -th tree is formally expressed as:

$$f_t(x) = w_{q(x)}, w \in \mathbb{R}^T, q: \mathbb{R}^d \rightarrow \{1, 2, \dots, T\} \quad (1)$$

Here, w denotes the weight (score) assigned to a leaf, $q(x)$ indicates the index of the leaf that sample x reaches, and T represents the total number of leaves in the tree.

The complexity of a tree is quantified by combining the count of its leaves with an L2 regularization penalty on the leaf scores:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (2)$$

Where γ penalizes the addition of new leaves (controlling tree size), and λ is the regularization strength applied to the magnitude of leaf weights.

(3) Objective function: The overall learning goal at each boosting round is to minimize the following objective:

$$Obj^{(t)} = \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \quad (3)$$

Where x_i is the i -th sample, $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$, $h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)})$, and $l(y_i, \hat{y}^{(t-1)})$ is the defined loss function, y_i is the true value of the i -th sample, and $\hat{y}^{(t-1)}$ does the previous t obtain the model - 1 round of training.

Let $I_j = \{i | q(x_i) = j\}$ denote the set of samples assigned to leaf j . After the second-order Taylor expansion of the loss around the current prediction, the simplified (approximate) objective becomes:

$$Obj^{(t)} = \sum_{j=1}^T [G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2] + \gamma T \quad (4)$$

Where:

$$G_j = \sum_{i \in I_j} g_i, \quad \sum H_j = \sum_{i \in I_j} h_i.$$

(4) Scoring function: Assuming the tree structure $q(x)$ is known, the optimal weight w and the associated objective function value can be calculated using [Eq. 4](#), as shown below:

$$w_j^* = -\frac{G_i}{H_i + \lambda} \quad (5)$$

$$obj = -\frac{1}{2} \sum_{j=1}^T \frac{G_i^2}{H_i + \lambda} + \gamma T \quad (6)$$

(5) Split gain: During tree construction, a greedy strategy evaluates potential splits. For any candidate split that divides a node into left child l_l and right child l_r , the gain is computed as:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L^2 + G_R^2)}{H_L + H_R + \lambda} \right] - \lambda \quad (7)$$

The first two terms capture the improvement in loss reduction from the left and right subtrees, the third term reflects the loss reduction without the split, and $-\gamma$ accounts for the complexity penalty incurred by creating two new leaves.

By aggregating predictions from multiple weak learners in this manner, XGBoost achieves stronger predictive power and improved generalization than any individual tree. The explicit inclusion of regularization—both on tree size (via γ) and on leaf weight magnitudes (via λ)—effectively mitigates overfitting. Applied to charging station condition monitoring, this approach supports clear, interpretable visualization of the evaluation steps, enabling precise determination of operational health and more effective planning of preventive maintenance activities.

State Evaluation Model for Charging Stations

In this study, the evaluation of the charging station's state is formulated as a multiclass classification problem. The state variables serve as input features, while the four operating states function as the classification labels. The training dataset consists of historical fault records, maintenance logs, monitoring data, and external environmental information from the charging stations. A total of 320 data samples were collected and processed, with 280 allocated to training and 40 reserved for testing.

Through iterative training and fine-tuning across multiple datasets—particularly those involving dynamic monitoring data—parameters that closely match real-world conditions are derived. Compared with alternative machine learning approaches, XGBoost performs an initial classification using four basis function categories. This strategy helps prevent parameter-optimization inconsistencies that might otherwise arise from substantial data variability.

Steps in model construction

(1) Definition of the loss function: As the charging station state evaluation constitutes a multiclass classification task, the loss function is defined using cross-entropy. For an individual sample, the loss function takes the following form:

$$l(y_k, f_k(x)) = -\sum_{k=1}^K y_k \log p_k(x) \quad (8)$$

Where y_k represents the true probability of the sample belonging to category k , $p_k(x)$ denotes the model's predicted probability that the sample belongs to category k , and

$$p_k(x) = \frac{e^{f_k(x)}}{\sum_{k=1}^K e^{f_k(x)}}.$$

For the entire set of samples, the first derivative (denoted as g) and the second derivative (denoted as h) of the loss function are computed as follows:

$$g_{ik} = y_{ik} - p_{k,i-1} \quad (9)$$

$$h_{ik} = \begin{cases} p_{i,k-1}(1 - p_{j,k-1}), & i = j \\ -p_{i,k-1}p_{j,k-1}, & i \neq j \end{cases} \quad (10)$$

(2) Initialization of prior probabilities: Given that the charging station's operational states are categorized into four classes, the prior probability of each sample belonging to every class is established.

$$p_0(x) = \frac{1}{4} = 0.25$$

(3) The first tree is fitted for the initial category [17, 18], according to the procedure below:

(a) A greedy algorithm is applied to construct the tree model through successive splits of the current leaf nodes. Using Eqs. 7, 9, and 10, the gain associated with each potential split is computed, and the split yielding the highest gain is chosen to build each tree.

(b) For every tree generated in step (a), the corresponding structure score is calculated based on Eq. 6. The tree exhibiting the lowest structure score is then selected as the first tree for the first category.

(c) The leaf score for the tree identified in step (b) is computed using Eq. 5, thereby forming a complete regression tree model. The model's predicted probability that the sample belongs to category 1 is subsequently updated.

(d) The process described in step (3) is repeated to fit the first tree for the second category, thereby completing the first tree for all four categories in the same manner;

(e) Steps (c) through (d) are iterated until the predefined threshold is satisfied. This process finalizes the construction of M trees across all categories, yielding the complete state evaluation model for the charging station. By leveraging historical data and invoking the ECDF function in MATLAB's statistical toolbox, the empirical probability distribution for the 50 prediction intervals is estimated.

Cross-validation and parameter settings

This study employs K -fold cross-validation, dividing the training data into K equal subsets. $K-1$ subsets are utilized for training, while the remaining one serves for validation. The average classification error rate serves as a performance metric for assessing the model's classification quality. Model parameters are tuned accordingly, with the optimal configuration determined from cross-validation results. The primary parameter choices are presented in Table 1 [19, 20].

Table 1. Model parameters.

Parameter	Value	Description
Learning-rate	0.08	The step size for shrinkage, also known as the learning rate
max_depth	5	The maximum depth of the tree is set to 5
n_estimators	650	The number of trees for each category is set to 650
eval_metric	merror	The metric used for the multiclass classification error rate
min_child_weight	0.8	The minimum sample weight for leaf nodes

Except for the parameters presented in **Table 2**, all remaining parameters are configured to their default values. With the parameter settings described above, the optimal state-evaluation model for charging stations is obtained.

Table 2. Classification of charging station risk levels.

Charging station number	Value-at-risk	Operation and maintenance order
2	12.325	1
7	11.478	2
6	8.128	3
9	7.693	4
3	7.356	5
4	7.238	6
1	6.841	7
5	5.256	8
10	3.943	9
8	2.568	10

Evaluation procedure

In this approach, XGBoost relies on decision trees as its fundamental components, allowing for clear visualization of the model architecture. This allows observing how the system assesses the charging station's operational condition. Given the extensive number of trees involved in the ensemble, only the initial two trees from the primary category are selected here to serve as illustrative examples. The corresponding tree structures are shown in **Figures 3 and 4**.

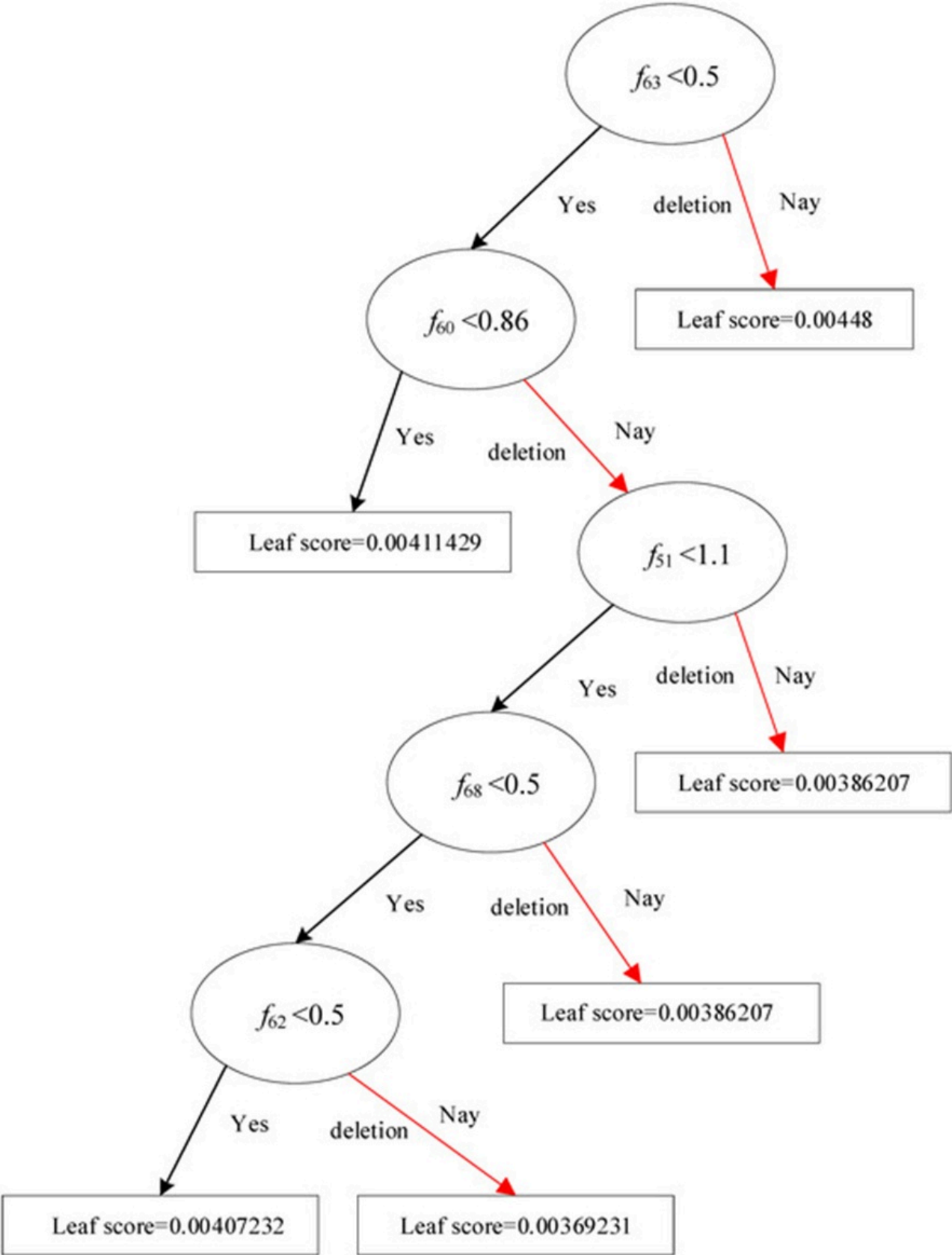


Figure 3. Generation and load scenario matching.

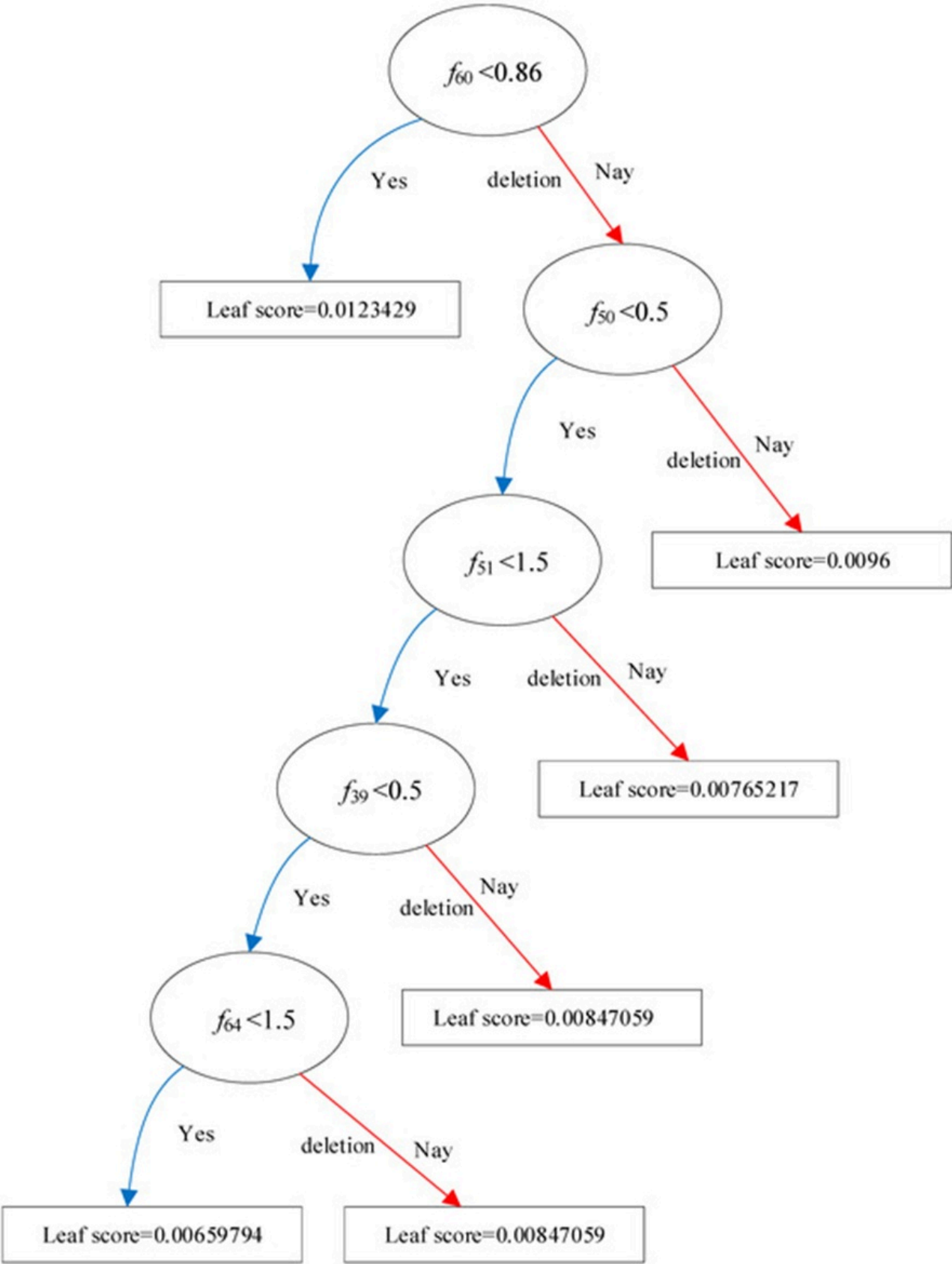


Figure 4. The second tree of the first class.

From **Figures 3 and 4**, it is evident that each tree's state variables contribute to splitting at their designated split values, with leaf nodes associated with distinct scores. For a particular set of operational data from a charging station, the specific leaf nodes to which these data points ultimately fall in each tree can be identified using the state-variable splitting criteria. Within each category, the scores for the leaf nodes reached by this dataset across all trees are summed. For instance, assuming the data in this group reach the leaf nodes in the lower left corner, as shown in **Figures 3 and 4**, the resulting score is 0.003412526. Consequently, the scores for this group of data across the different categories can be denoted as $[S_1, S_2, S_3, S_4]$. By applying the softmax function, these scores are transformed into probabilities $[P_1, P_2, P_3, P_4]$ that represent the likelihood that the data belong to each category. The category with the highest probability is selected as the final evaluation outcome, indicating the charging station's corresponding operational state.

Charging Station Optimization Operation and Maintenance

Traditional operation and maintenance practices for charging stations generally schedule tasks either by geographical proximity or by the duration since the station was commissioned. In contrast, the approach proposed here prioritizes maintenance according to the assessed risk level of each station's condition. Leveraging the risk assessment methodology outlined in this paper, the operation and maintenance priorities for individual stations are established using the results of the state evaluation.

Operation and maintenance scheduling is categorized into equal-period and optimized-period approaches. Equal-period maintenance applies uniform treatment to all charging stations, evenly allocating the workload across the entire maintenance cycle. The optimized-period maintenance strategy introduced in this paper accounts for both operation and maintenance costs and the resulting improvements in reliability. It assigns workloads to each station to minimize overall operational and maintenance expenditure.

Objective function

The objective function for the optimal operation and maintenance strategy of electric vehicle charging stations integrates both operation and maintenance costs and reliability benefits [21], expressed as follows:

$$\min F(\mathbf{X}) = f_c(\mathbf{X}) + f_r(\mathbf{X}) \quad (11)$$

Where $F(\mathbf{X})$ denotes the total objective function, and x_m in $\mathbf{X}$ represents the number of operation and maintenance periods assigned to the m -th charging station.

As an illustrative example, consider the operation and maintenance activities for electric vehicle charging stations in a city: weekly operations span 5 days, with daily working hours from 8:00 to 12:00 in the morning and 2:00 to 6:00 p.m. Accordingly, this paper adopts a one-week operation and maintenance cycle, where each period lasts two hours. This duration encompasses travel time, actual maintenance activities, information recording, and rest periods.

(1) $f_c(\mathbf{X})$ represents the economic objective function, encompassing the costs of resources expended during maintenance, given by:

$$f_c(\mathbf{X}) = \sum_{m=1}^M \sum_{t=1}^{T_s} c_{m,t} \quad (12)$$

Where T_s is the total number of time periods in the operation and maintenance cycle, M is the total number of charging stations, and $c_{m,t}$ is the maintenance cost for the m -th charging station in the t -th period.

(2) $f_r(X)$ denotes the expected cost associated with system risk. Maintenance aims to lower the equipment's current failure rate and thereby enhance system reliability; thus, the system risk is computed using Eq. 5 following each maintenance period, with the resulting time losses converted into economic terms:

$$f_r(X) = \sum_{t=1}^{T_s} \sum_{n=1}^N g \frac{LOT_t(n)}{N} \quad (13)$$

Where N is the number of scenarios, $LOT_t(n)$ is the user's loss-of-time in the n -th scenario after the t -th maintenance period, and g is the unit-time cost conversion coefficient for user loss of time [22].

In accordance with the specific characteristics of operation and maintenance for electric vehicle charging stations, the following constraints are established:

(1) Each station must undergo at least one period of maintenance, reflecting the time constraints for individual charging stations:

$$x_m \geq 1 \quad (m \in CS) \quad (14)$$

(2) The timing constraints for maintenance are:

$$y_{m+1} = y_m + x_m \quad (m, m+1 \in CS) \quad (15)$$

(3) Maintenance activities across all charging stations in the system must not occur simultaneously, considering the overall time constraint:

$$\sum_{m \in CS} x_m = T_s \quad (16)$$

(4) Resource constraints for maintenance are:

$$\sum_{x=1}^{x_m} r_{m,x} \leq Y_{m,\max} \quad (m \in CS) \quad (17)$$

Where y_m indicates the starting time of maintenance for the m -th charging station, $r_{m,x}$ represents the maintenance resources required by the m -th charging station in the x -th time period, and $Y_{m,\max}$ is the maximum maintenance resources available for the m -th charging station.

Optimization algorithm

The operation and maintenance strategy optimization examined in this paper constitutes a nonlinear integer programming problem, rendering analytical solutions challenging. Consequently, the immune clonal selection algorithm (ICSA) is adopted to address it [23]. Compared with conventional immune algorithms, ICSA incorporates clonal expansion and mutation operators, which facilitate the amplification of high-quality solutions and the elimination of inferior ones

throughout the iterative process, thereby significantly enhancing convergence speed. Within the ICSA framework, antibodies correspond to the solution set, antigens to the objective function fitness, and affinity to the degree of match between the solution set and the objective function in the optimization context.

The proposed operation and maintenance strategy for charging stations primarily comprises three components: risk assessment, risk tracking, and operation and maintenance optimization. Initially, the system risk value is determined based on the charging piles and the associated road network. Next, risk tracking establishes the maintenance priority for each station. Finally, the ICSA algorithm is applied to determine the number of maintenance periods per station, aiming to minimize both economic and expected risk costs. The overall optimization process for the operation and maintenance strategy is depicted in Figure 5.

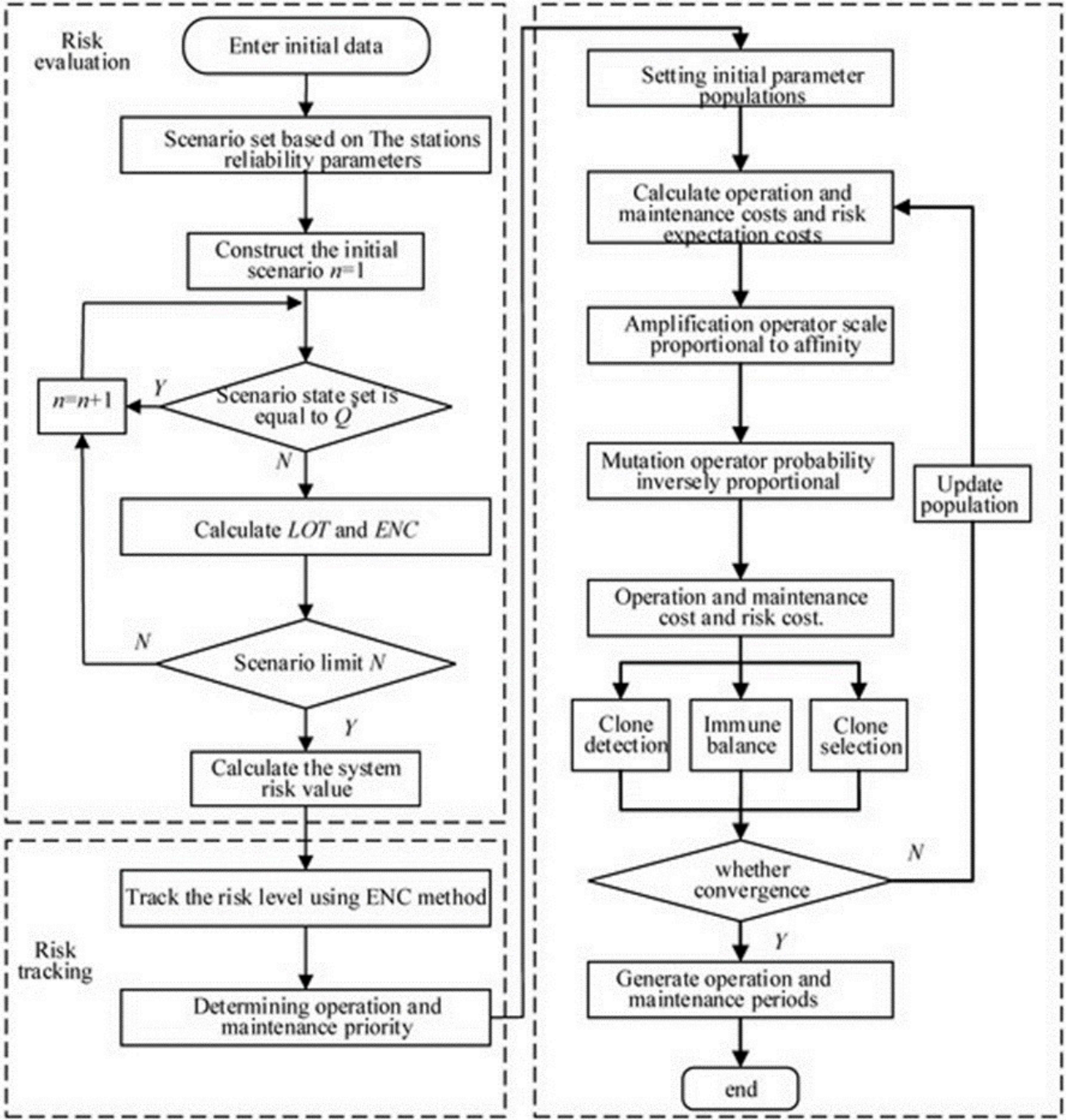


Figure 5. Flowchart illustrating the optimization procedure for operation and maintenance strategies.

Case Study

Road network layout and input parameters

To demonstrate the proposed approach, an example based on the transportation infrastructure in a particular region is examined. The layout of the road system and the locations of charging facilities are depicted in **Figure 6**. At present, a subset of the 10 key transportation junctions highlighted in the diagram is already operational. For this analysis, every junction is assumed to host identical DC fast-charging stations. Each station includes 10 charging units (piles), each delivering 30 kW of power. The stations are labeled in order from the one positioned farthest away to the closest one.

Failure and repair rate data, along with information on years in service and equipment specifications from recently commissioned charging units, were used. These were integrated into a simulation framework to derive the specific failure rates for the charging piles at each location. Furthermore, traffic-related characteristics of the road segments were assigned values based on each road's classification (grade).



Figure 6. Layout of the transportation network together with the placement of charging facility locations.

In this context, electric vehicles primarily consist of compact models that rely on DC rapid charging. In contrast, privately owned passenger vehicles are far more common and typically rely on stationary AC charging points. However, government or company fleet vehicles, taxis, and the increasingly widespread individually operated commercial vehicles mainly utilize DC charging to accommodate their operational needs.

By appropriately incorporating considerations of power grid dispatch and examining distinct charging duration patterns across various categories of electric vehicles, it becomes possible to perform risk assessments and refine maintenance planning strategies.

Accordingly, the proportion of private passenger cars to urban fleet/commuter vehicles was established at 1:5. To capture the stochastic and fluctuating characteristics of charging demand across the network, the analysis incorporated 120 vehicles requiring charging services traveling within the road system during the high-demand window of 12:00 to 14:00 [24]. Throughout this interval, 60 vehicles were either queued or actively charging at the facilities, enabling an evaluation of overall system dependability. The initial vehicle placements, the timing of demand emergence, and the starting state-of-charge (SOC) values were drawn from Jiang *et al.* [25]. Each vehicle was assigned a battery capacity of 25 kWh. The simulation used a 0.5-minute time step and covered the full duration, from the first vehicle's charging request to the final vehicle's completion of its session.

Risk assessment and maintenance optimization approach

Drawing on an examination of several fundamental operating condition categories for charging stations, the XGBoost algorithm—built upon ensemble tree structures—serves as the core predictive model and enables clear visualization of the evaluation outcomes. This approach transparently illustrates how a charging station's operational condition is judged. Through iterative tree-based classification, the model identifies the condition category with the greatest likelihood, thereby determining the one most representative of the station's true status. As a result, XGBoost delivers higher-precision evaluations of station health, supporting more precise and effective allocation of operation and maintenance activities.

Given the individual failure probabilities of all charging piles in the network, stratified Monte Carlo sampling was applied 3000 times to generate diverse possible system states. The starting vehicle parameters remained identical across every sampled scenario. For each scenario, traffic and charging simulations were run under both the ideal (Q^*) and current sampling conditions, yielding the charging completion time for each electric vehicle and the energy delivered by each station. From these results, the overall system risk index and the expected number of unserved customers (ENC) per station were computed. Eq. 8 was then used to apportion the total risk to individual charging stations. Stations were subsequently ranked in descending order of risk priority to establish the recommended sequence for maintenance interventions, as presented in Table 2.

The study first improves maintenance task scheduling intervals by aligning them with real-world operational practices. Next, maintenance strategies were evaluated and compared using both uniform and newly optimized time intervals. The resulting maintenance schedules—covering conventional sequential maintenance under fixed intervals versus the risk-prioritized approach with optimized intervals—are illustrated in Figure 7.

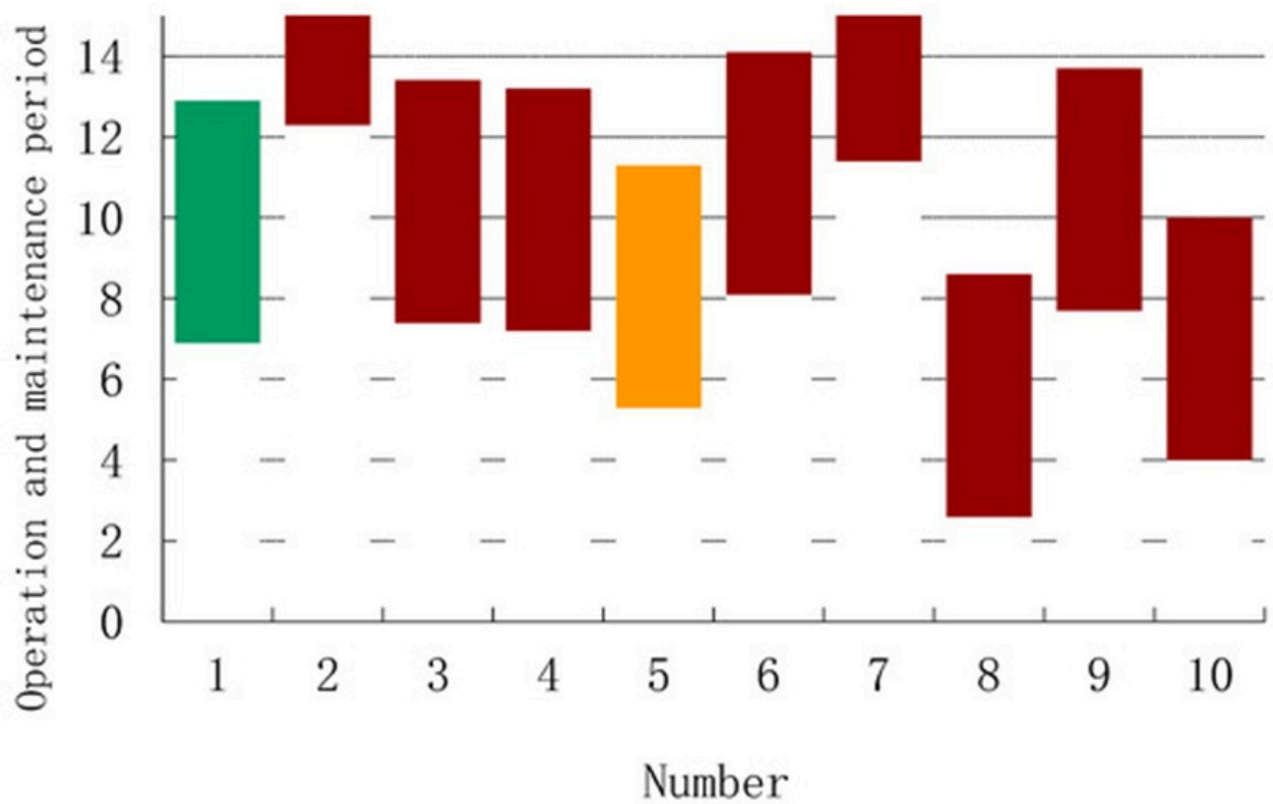


Figure 7. Maintenance scheduling under the conventional sequential approach for the charging stations.

As illustrated in **Figure 7**, after optimizing maintenance intervals, Station 3—which exhibits the greatest risk—receives an extended maintenance duration. In contrast, Stations 2 and 8, characterized by comparatively lower risk levels, have their allocated maintenance times reduced. Moreover, to improve cost efficiency in maintenance activities, the durations assigned to the other stations have also been appropriately adjusted. Nevertheless, owing to constraints on available resources and overall scheduling time, adjustments to the maintenance periods are applied only for a single cycle.

A comparison of total maintenance expenditures under the uniform-interval and risk-optimized interval approaches (both within the traditional sequential maintenance framework) is presented in **Table 3**.

Table 3. Operation and maintenance costs based on risk order

Decision-making model	System risk expectation (CNY)	Operation and maintenance cost (CNY)	Total cost (CNY)
Equal time periods of operation and maintenance	1289.4	1098.8	2397.2
Optimized time periods of operation and maintenance	1023.6	996.2	2097.8

Compared with the approach using fixed equal intervals, the optimized scheduling reduces total operation and maintenance expenses by 6.63% and the expected system risk level by 18.56%. These improvements deliver superior overall effectiveness for maintenance operations.

The specific maintenance assignments and schedules—comparing the conventional equal-interval method with the risk-ordered optimized-interval method—are presented in **Figure 8**.

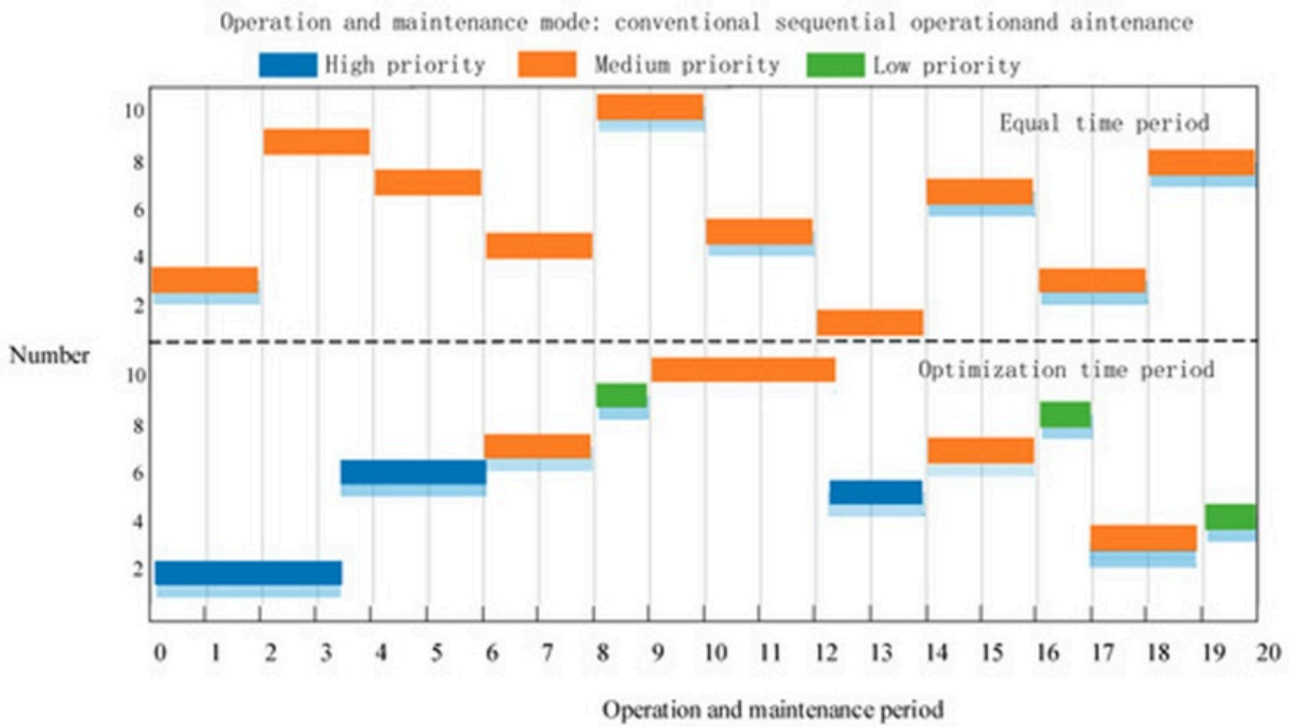


Figure 8. Comparison of original and optimized maintenance schedules for the charging stations.

As depicted in **Figure 8**, the risk-prioritized maintenance strategy reallocates maintenance durations for selected charging stations within the optimized intervals, resulting in lower overall expenditure than the approach that uses fixed equal intervals. The aggregate costs associated with equal-interval maintenance and with risk-ordered optimized-interval maintenance are summarized in **Table 4**.

When comparing the two different sequencing approaches (conventional sequential versus risk-based ordering), the expected system risk costs are lowered by 31.23% under the equal-interval framework and by 32.48% under the optimized-interval framework, respectively, when both adopt the risk-based priority order. These outcomes indicate that organizing maintenance tasks by station risk levels substantially enhances the system's overall reliability and minimizes potential user losses.

Relative to the prevailing conventional sequential maintenance performed at uniform time intervals, the proposed risk-driven optimization of maintenance timing and sequencing reduces the total cost by 24.11% while keeping the direct maintenance expenditure unchanged. This approach, therefore, markedly improves the overall efficiency of operations and maintenance activities.

Table 4. Operation and maintenance costs based on the conventional order

Decision-making model	System risk expectation (CNY)	Operation and maintenance cost (CNY)	Total cost (CNY)
Equal time period for operation and maintenance	798.6	1089.6	1987.2

Optimized time period operation and maintenance	703.7	997.6	1589.6
---	-------	-------	--------

To overcome certain constraints inherent in the XGBoost approach, this research implemented multiple mitigation techniques. Historical fault records underwent thorough preprocessing. Before every maintenance activity, the model was retrained using a combination of post-maintenance records, real-time monitoring data, and external environmental variables, thereby enabling ongoing refinement and parameter updates. The influence of stochastic factors related to user behavior and environmental conditions on these parameters was explicitly incorporated. In each training iteration, the average classification error rate across several parameter configurations served as the primary indicator of the model's classification performance.

Conclusion

In response to deficiencies in existing operation and maintenance practices for electric vehicle charging stations, this study introduced a comprehensive framework for DC charging stations. The framework leverages the XGBoost ensemble learning algorithm to assess each station's operational condition accurately.

(1) An XGBoost-based evaluation model for charging station conditions was developed. Key influencing factors—including users' driving and charging patterns, road network traffic conditions, and equipment status—were taken into account. Training samples were constructed by integrating historical fault and usage records, continuously collected online monitoring data, and relevant external environmental information. The XGBoost algorithm was then applied to train this condition assessment model.

(2) Incorporating both operation and maintenance expenditures and the resulting improvements in system reliability, an optimization model for maintenance strategies at electric vehicle DC charging stations was formulated, grounded in the condition evaluations provided by XGBoost. This model determined the prioritized sequence for servicing all charging stations based on their assessed risk levels and computed the appropriate maintenance duration for each station within a given maintenance cycle.

The proposed operation and maintenance strategy—derived from the XGBoost-driven condition assessment and subsequent optimization model—effectively eliminates the randomness, lack of direction, and low efficiency that characterize traditional maintenance approaches. It substantially enhances the overall performance and cost-effectiveness of system-level operations and maintenance activities while ensuring high reliability and availability of the charging infrastructure.

Acknowledgements

None

Conflict of interest

None

Financial support

This research was funded by the State Grid Corporation of China, Headquarters Management of Science and Technology Project (5400-202128154A-0-0-00), and China Scholarship Council.

Ethics statement

None

Received: 10 Oct 2024

Revised: 17 Nov 2024

Accepted: 28 Dec 2024

Published online: 20 March 2025

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Su L, Liu Z, Wang S. Ordered charging strategy of electric vehicles based on user driving behavior characteristics. *Electr Power Autom Equip*. 2018;38:63-71.

Lin PC, Gu JC, Yang MT. Intelligent maintenance model for condition assessment of circuit breakers using fuzzy set theory and evidential reasoning. *IET Gener Transm Distrib*. 2014;8(7):1244-53.
<https://doi.org/10.1049/iet-gtd.2013.0445>.

Arispe JG, Mombello E. Power Transformer Condition Assessment Using DGA and FRA. *IEEE Lat Am Trans*. 2016;14(11):4527-33.
<https://doi.org/10.1109/TLA.2016.7795822>.

Fang H, Qiu X, Shang L, Ye X, Dong X, Tian Y, et al. User Reliability Feature Portrait Method Based on Tensor Decomposition. *Lect Notes Electr Eng*. 2023;1048:819-30.
https://doi.org/10.1007/978-981-99-3408-9_70.

Xu P, Fang H, Wang H. Progressive maintenance scheduling strategy of transmission lines based on risk assessment and sensitivity clustering. *Power Syst Technol*. 2018;42(9):3024-32.
<https://doi.org/10.13335/j.1000-3673.pst.2017.2307>.

Heo JH, Kim MK, Park GP, Yoon YT, Park JK, Lee SS, et al. A reliability-centered approach to an optimal maintenance strategy in transmission systems using a genetic algorithm. *IEEE Trans Power Deliv*. 2011;26(4):2171-9.
<https://doi.org/10.1109/TPWRD.2011.2158670>.

Feng D, Lin S, He Z, Sun X, Lee WJ. Optimization method with prediction-based maintenance strategy for traction power supply equipment based on risk quantification. *IEEE Trans Transp Electr*. 2018;4(4):961-70.
<https://doi.org/10.1109/TTE.2018.2865172>.

Peng F. Research on risk assessment method for automatic operation and maintenance of high voltage transmission lines. *Autom Instrum*. 2019;1:64-7.
<https://doi.org/10.14016/j.cnki.1001-9227.2019.01.064>.

Liang S, Qin L, Li C. Condition-based maintenance decision-making method for distribution equipment based on risk assessment. *Proc CSU-EPSA*. 2019;31(8):47-52.
<https://doi.org/10.19635/j.cnki.csu-epsa.000101>.

Yang H, Li M, Wen F. Route selection and charging navigation strategy for electric vehicles employing real-time traffic information perception. *Autom Electr Power Syst*. 2017;41(11):106-13.
<https://doi.org/10.7500/AEPS20160926012>.

Razi-Kazemi AA. Circuit breaker condition assessment through a fuzzy-probabilistic analysis of actuating coil's current. *IET Gener Transm Distrib*. 2016;10(1):48-56.
<https://doi.org/10.1049/iet-gtd.2015.0261>.

National Electric Vehicle Service Co., Ltd. Technology Research and Development Center (Technology Center). Charging Pile Fault Code and Common Missing Methods. Beijing; 2018.

Jiang C, Wang S, Wang B. Probabilistic reliability assessment of power system containing wind power based on Latin hypercube sampling. *Trans China Electrotech Soc*. 2016;31(10):193-206.
<https://doi.org/10.19595/j.cnki.1000-6753.tces.2016.10.022>.

Ma H, Saha TK, Ekanayake C. Statistical learning techniques and their applications for condition assessment of power transformer. *IEEE Trans Dielectr Electr Insul*. 2012;19(2):481-9.
<https://doi.org/10.1109/TDEI.2012.6176240>.

Jiang Y, Yu A, Huang M. Coordinated charging guiding strategy for electric vehicles in temporal-spatial dimension considering user satisfaction degree. *Electr Power*. 2020;53(4):122-30.
<https://doi.org/10.11930/j.issn.1004-9649.201906184>.

Ma XY, Sun YZ, Fang HL. Scenario Generation of Wind Power Based on Statistical Uncertainty and Variability. *IEEE Trans Sustain Energy*. 2013;4(4):894-904.
<https://doi.org/10.1109/TSTE.2013.2256805>.

Fang H, Bai Y, Gu H, Zhao J, Wang Z, Wang G, et al. Research on Traceability Management Method of Power Grid Engineering Cost Data Based on Blockchain Technology. *Lect Notes Electr Eng*. 2023;1048:1167-76.
https://doi.org/10.1007/978-981-99-3408-9_99.

Zeng C, Liu G, Miao H. Planning of electric vehicle charging station considering urban traffic flows. *J Electr Power Sci Technol*. 2019;34(3):101-7.
<https://doi.org/10.19781/j.issn.1673-9140.2019.03.013>.

Xiong B, Shu Z, Guo C. Charging station optimization planning from the perspective of social benefits. *Smart Power*. 2019;47(8):78-84.

Wu R, Zhou L, Yao Z. Intelligent charging navigation for electric vehicles based on real-time electricity price. *Electr Power*. 2020;53(1):131-8.
<https://doi.org/10.11930/j.issn.1004-9649.201905173>.

Dong M, Zheng H, Zhang Y. A novel maintenance decision making model of power transformers based on reliability and economy assessment. *IEEE Access*. 2019;7:28778-90.
<https://doi.org/10.1109/ACCESS.2019.2902310>.

Li H, Qiang W, Gao Y. Charging station planning considering users' travel characteristics and line availability of distribution network. *Autom Electr Power Syst*. 2018;42(13):48-56.
<https://doi.org/10.7500/AEPS20171221007>.

Xu G. Optimization Method and Its MATLAB Implementation. Beijing: Beijing University of Aeronautics and Astronautics Press; 2018.

Fang H, Shang L, Dong X, Tian Y. High Proportion of Distributed PV Reliability Planning Method Based on Big Data. *Energies*. 2023;16(23):7692.

<https://doi.org/10.3390/en16237692>.

Jiang X, Feng YT, Xiong H, Wang J, Zeng Q. Electric vehicle charging station planning based on travel probability matrix. *Trans China Electrotech Soc*. 2019;34(S1):272-81.

<https://doi.org/10.19595/j.cnki.1000-6753.tces.180173>.