International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5635-5650

https://doi.org/10.70517/ijhsa464404

# Grid Impact Mitigation Methods for Unordered Charging Loads from Electric Vehicle Charging Stations

**Chong Gao[1], Xinghang Weng[1], Yao Duan[1], Zhiheng Xu[1] and Junxiao Zhang[1,*]**
[1] Grid Planning & Research Center, Guangdong Power Grid Co., Ltd., CSG, Guangzhou, Guangdong, 510000, China
Corresponding authors: (e-mail: HP18721861985@163.com).

**Abstract** Driven by the carbon peak carbon neutral target, the new energy vehicle industry has ushered in a period of rapid development, and the number of electric vehicles continues to climb. The randomness and uncertainty of the spatial and temporal distribution of large-scale electric vehicles connected to the power grid for charging bring severe challenges to the power system operation. In this paper, a charging load inference and governance method based on Bayesian network is proposed for the problem of the influence of disordered charging load of electric vehicle charging pile on the stability of power grid operation. The method establishes the probabilistic dependence between the charging load and the grid operation parameters by constructing a Bayesian network model, taking into account the influencing factors such as EV charging mode, user behavioral characteristics and power battery characteristics. The study adopts Monte Carlo simulation method to generate EV charging load data, uses BIC scoring function to optimize the network structure, and determines the network parameters through maximum likelihood estimation method. The experimental results show that the data inference method based on Bayesian network achieves an average precision of 93.3% and a recall of 94.5% in charging pile state prediction, which is 7.49% and 9.63% higher than the traditional method, respectively. In the IEEE 33-node distribution system simulation, the peak system line loss occurs in the 18:00-23:00 time period when EV penetration increases from 0% to 100%. The study shows that the Bayesian network governance method can effectively reduce the adverse effects of uncontrolled charging of EVs on the voltage deviation and network loss of the distribution network, and provide a scientific basis for the optimization of power grid operation.

**Index Terms** Electric vehicle, Bayesian network, charging load, data extrapolation, distribution network, governance methods

## I. Introduction

The continuous consumption of energy and its pollution of the environment are key elements that limit socio-economic development, and at the same time, they are one of the common concerns of the world in the 21st century [1]. The protection of the environment has become an inevitable trend of economic development on a global scale. With the global economic development, the number of automobiles in each country is increasing day by day, and the environmental pollution problem caused by the consumption of energy by automobiles will be intensified [2]-[4]. This is contrary to the development goals of countries around the world, so the energy type modification of automobiles is a must for the development of the automotive industry, i.e., the manufacture of electric vehicles that use secondary energy sources [5].

With the increasing shortage of fossil energy and the increasing greenhouse effect, the large-scale use of electric vehicles as an environmentally friendly and low-carbon means of transportation using secondary energy will be the inevitable result [6]. The difference between electric vehicles and traditional fuel vehicles is that the former is powered by electricity, which has a much smaller impact on the environment during driving than the latter, while the former has a lower cost of energy consumption compared to the latter, which uses primary energy [7]-[9]. For distribution networks, the former can be used as adjustable loads capable of participating in scheduling and in the optimization and control of distribution networks [10]. The former can also be charged using renewable energy sources such as solar energy. In the process of large-scale use of electric vehicles, since electric vehicles as a new type of load have the disadvantage of time and space uncertainty, the scaled charging load of electric vehicles will lead to adverse effects such as increasing the peak-to-valley difference of the grid and disturbing the stable operation of the distribution grid [11]-[14]. Therefore, in order to mitigate the above adverse effects, it is necessary to control the charging load of electric vehicles in an orderly manner, so as to restore the power grid to an economic and stable operation state.

To control the charging load of electric vehicles is to optimize the electric vehicle load curve by managing the charging behavior of electric vehicles [15]. To control the charging load of electric bus charging and switching stations is to make full use of the fact that the battery pack can be charged in advance and the vehicle and the battery pack can be separated, which is different from the characteristics of electric vehicle charging stations, to optimize and control the charging time of electric vehicle batteries [16]-[18]. The control of electric vehicle charging load is not only conducive to the stable operation of the power system distribution network, but also able to consume new energy, which is conducive to the development of renewable energy generation.

This study proposes a Bayesian network approach to solve the governance problem of the impact of unorganized charging loads from electric vehicle charging piles on the power grid. First, a Bayesian network model that comprehensively considers the EV charging mode, user behavioral characteristics and power battery characteristics is constructed, and the complex dependency relationship between the influencing factors is described through a probabilistic graphical model. Second, Monte Carlo simulation is used to generate a large amount of charging load sample data to provide sufficient data support for network training. Then, the Bayesian network structure is optimized based on the BIC scoring function, and the maximum likelihood estimation method is used to determine the network parameters to achieve accurate inference of the charging pile state. Finally, the proposed method is applied to the IEEE33 node distribution system to analyze the impact of uncontrolled EV charging on the distribution network line loss and voltage under different penetration rates and charging locations to verify the effectiveness of the governance method.

## II. Inferential modeling of the impact of uncontrolled charging loads of electric vehicles on the power grid

In this paper, based on Bayesian networks, we propose an inference model for the impact of uncontrolled charging loads from electric vehicle charging piles on the power grid, so as to provide a basis for the optimization of power grid governance.

### *II. A. Bayesian networks*

Bayesian network [19] is a widely used graphical model, which is one of the most effective theoretical models in the field of uncertain knowledge representation and reasoning, and is used for prediction, diagnosis, and inference in a variety of fields.

In a Bayesian network, the probability distribution attached to all child nodes is the conditional probability distribution, the probability distribution attached to the root node without a parent node is the marginal probability distribution, and multiplying the probability distributions attached to all nodes is the joint probability distribution of all nodes. Bayesian networks can reduce the computational dimension of probabilistic models by decomposing the joint probability.

#### II. A. 1) Construction of Bayesian networks

(1) Constructing a Bayesian network using data

Constructing a Bayesian network using data is actually an optimization problem, and the objective function of optimization is the scoring function. The scoring function is usually based on information theory criterion, in which Bayesian Information Criterion (BIC scoring) is clear in meaning and easy to use, and is one of the most commonly used scoring functions, calculated as in Equation (1):

$$\log P(D \mid N) \approx \log P(D \mid N, \theta) - \frac{d}{2}\log m \tag{1}$$

where $D$ is the node-variable dataset, $N$ is the Bayesian network structure, $\theta$ is the parameter corresponding to when the network structure is $N$, $d$ is the parameter vector $\theta$ the number of independent parameters, and $m$ is the sample size.

Equation (1) is how the BIC score is calculated, where it describes the degree of fitting of the structure $N$ to the data $D$, and $\frac{d}{2}\log m$ describes the model complexity to avoid overfitting of the model.

(2) Artificial methods to construct Bayesian networks

Artificial construction of Bayesian networks based on existing knowledge is a common means of constructing Bayesian networks, and the artificial construction of Bayesian network structure can add a priori knowledge to the network structure, which can reduce the requirements of the model on the number of samples. Through the data correlation analysis method to analyze the sample data can clarify the dependence relationship between the node data, commonly used correlation analysis methods are as follows:

The Pearson correlation coefficient $\rho_{X,Y}$ [20] describes the strength of the linear correlation between the two samples, $\rho_{X,Y} < 0$ represents the negative correlation between the two variables, $\rho_{X,Y} > 0$ represents the positive correlation between the two variables, and the degree of correlation is expressed by $\left|\rho_{X,Y}\right|$, and the closer $\left|\rho_{X,Y}\right|$ is to 1, the higher the correlation between the two variables, $\left|\rho_{X,Y}\right|$ The closer to 0, the lower the correlation between the two variables. The Pearson correlation coefficient has high requirements for the data, which requires the data to be continuous and conform to the normal distribution, and the results can only judge whether the variables conform to the linear correlation and are greatly affected by outliers. The Pearson correlation coefficient is calculated as follows in equation (2):

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{2}$$

where $\operatorname{cov}(X,Y)$ is the covariance of sample $X$ and sample $Y$, and $\sigma$ is the standard deviation of the sample.

If the data of two variables do not meet the above requirements for the number of variables for correlation judgment using Pearson's correlation coefficient, Spearman's correlation coefficient is used to judge the correlation of the variables.

The Spearman rank correlation [21] coefficient assigns a rank data to the original data in descending position, and the correlation of the original data is judged according to the rank data. It is calculated as in equation (3):

$$\rho_{X,Y} = \frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y} \tag{3}$$

where $x$ and $y$ are the ranked data of sample $X$ and sample $Y$ in descending order.

The chi-square test is also a widely used hypothesis testing method, often used to compare the rate of two or more samples and the correlation or independence analysis of two variables. The statistic $\chi^2$ for the chi-square test is calculated as in equation (4) by tabulating the categories and characteristics:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(X_{i,j} - Y_{i,j})^2}{Y_{i,j}} \tag{4}$$

where $X$ is the actual observed value, $Y$ is the theoretical value when categories and features are independent of each other, $r$ is the number of categories and $c$ is the number of features. In the study of the correlation between two variables using the chi-square test, the larger the value of the statistic $\chi^2$, the stronger the correlation between the two variables is represented.

(3) Bayesian network parameter calculation

In the case where the structure is known and the observations are complete, the estimation of the parameters of the Bayesian network is carried out using the maximum likelihood estimation method. Learning the parameters of the Bayesian network is actually a parameter estimation of the conditional probability parameter $\theta$, i.e., the degree of fit of the possible values of the parameter $\theta$, $\hat{\theta}$, to the data $D$ is measured by $P(D \mid \theta = \hat{\theta})$. The sample $D$ consists of samples $D_1, D_2, \ldots, D_n$, and each sample is independently and identically distributed from each other, then its great likelihood function is as in equation (5):

$$L(\theta \mid D) = P(D \mid \theta) = \prod_{i}^{n} p(D_i \mid \theta) \tag{5}$$

The conditional probabilities of the Bayesian network nodes are calculated by great likelihood estimation to complete the parameterization of the Bayesian network, then the great likelihood estimation of the parameter $\theta$ is calculated as in equation (6):

$$\hat{\theta} = \arg\sup_{\theta} L(\theta \mid D) \tag{6}$$

#### II. A. 2) Inference in Bayesian networks

Based on the Bayesian network can realize the prediction and diagnosis tasks between the network node variables. The inference algorithm of Bayesian network includes two kinds of algorithms: exact inference and and approximate inference, when the structure of Bayesian network is simple, and the value of each node is not much, a more accurate posterior probability distribution can be obtained by exact inference. When the Bayesian network nodes are numerous and densely connected, the exact inference method, which is more difficult to compute, is often discarded, and the approximate computation method is utilized to compute the posterior probability distribution by sampling.

In Bayesian networks, the joint probability of all variables is decomposed into the concatenated multiplication of the conditional probabilities of multiple node variables according to the structure of the Bayesian network, as in equation (7):

$$P(D)=\sum_{A,B,C}P(A,B,C,D)=\sum_{A,B,C}P(A)P(B\mid A)P(C\mid B)P(D\mid C)\ldots \tag{7}$$

When calculating the posterior probability distribution of a node, the direct elimination calculation of the joint probability distribution needs to consider the relationship between the node variable and all other node variables, and through the dependency relationship between the nodes of the Bayesian network, the joint probability distribution is transformed into the concatenated multiplication of the conditional probabilities of multiple node variables before eliminating the variables, which reduces the connection between the variable to be eliminated and the variables of other nodes, thus reducing the amount of calculation and improving the calculation. The method reduces the connection between the variable to be eliminated and other node variables, thus reducing the calculation amount and improving the calculation speed.

#### II. A. 3) Validation of Bayesian networks

In this paper, the validation method of Bayesian network model is to test the discrete simulated disordered charging load data against the discrete original disordered charging load data distribution by using K-S test and calculating the error between the simulated data and the original data to test the categorical variables.

The K-S test is a nonparametric test that is usually used to test whether the distribution of observed data conforms to a theoretical distribution or to test whether two observed distributions are significantly different.

It is first assumed that the two data distributions are consistent or a data distribution conforms to a theoretical distribution, and the statistic of K-S test is calculated as in equation (8):

$$D_{n,m}=\sup_{x}\left|F_{1,n}(x)-F_{2,m}(x)\right| \tag{8}$$

where $F_{1,n}(x)$ is the value of the cumulative probability of an observation with a sample size of $n$ at point $x$, and $F_{2,m}(x)$ is the value of the theoretical distribution or cumulative probability of another set of observations with a sample size of $m$ at point $x$.

The critical value of $D_{n,m}$ is obtained by looking up the table or calculating using the sample size and significant level:

$$D_{n,m}>c(\alpha)\sqrt{\frac{n+m}{nm}} \tag{9}$$

$$c(\alpha)=\sqrt{-\frac{1}{2}\ln\alpha} \tag{10}$$

where $\alpha$ is the significance level, if $D_{n,m}$ is greater than the critical value is considered that the two distributions are significantly different, the original hypothesis is rejected, otherwise the original hypothesis is accepted, i.e., the two samples are from the same distribution. In addition can also be used to determine whether the hypothesis is valid or not by the P-value, when the value of the significance level $\alpha$ is greater than the P-value of the K-S test, the original hypothesis is rejected, otherwise the original hypothesis is accepted.

### *II. B.Electric Vehicle Charging Load Influencing Factors*

In order to analyze the impact of uncontrolled charging of electric vehicles on the grid load, this section analyzes the charging method of electric vehicles, the size and type of electric vehicles, the classification and characteristics of power batteries and the characteristics of user behavior according to the factors affecting the charging load of electric vehicles.

### II. B. 1) Electric Vehicle Charging Methods

Due to the different categories and technologies of on-board power batteries and different charging environments, there are mainly the following charging modes for electric vehicles: conventional charging mode, fast charging mode, battery pack replacement mode, mobile charging mode, and wireless charging mode. In this paper, conventional charging and fast charging methods are used for research and analysis.

### II. B. 2) Electric Vehicle Sizes and Types

EVs have different impacts on the grid depending on their application scale and penetration rate. Therefore, to study the impact of EV charging behavior on the grid, it is necessary to reasonably predict the scale of EVs in the system based on the penetration of EVs.

According to the different types of engines, they are now mainly categorized into pure electric vehicles and plug-in hybrid vehicles. In this paper, the charging and discharging behaviors of pure electric vehicles for home use are studied.

### II. B. 3) Power Battery Classification and Characteristics

(1) Power Battery Classification

At present, newly produced electric vehicles tend to use nickel-metal hydride batteries and lithium-ion batteries as the power source, especially the lithium-ion batteries with high specific energy and high specific power are more popular.

(2) Characteristics of power battery

1) State of Charge

The state of charge (SOC) is used to reflect the remaining battery power, which is an important technical parameter in the operation of electric vehicles, and refers to the ratio of the remaining capacity of the battery to the rated capacity under the same conditions and a certain discharge rate, and its value is between 0-1. That is:

$$SOC = \frac{\text{Remaining capacity}}{\text{Rated capacity}} \times 100\% \tag{11}$$

The battery charge state during operation can be expressed as:

$$SOC = SOC_0 + \frac{\int_0^t I_{bat} dt}{C} \times 100\% \tag{12}$$

where, $I_{bat}$ denotes the operating current of the battery at the moment $t$, positive for charging and negative for discharging. $C$ denotes the rated capacity of the battery. $SOC_0$ denotes the initial state of charge of the battery, when $SOC_0 = 1$, the battery is fully charged. When $SOC_0 = 0$, the battery is fully discharged.

SOC commonly used estimation methods include discharge experiment method, open circuit voltage method, Ah integration method, linear modeling method, measurement of internal resistance method, Kalman filtering and so on.

2) Charging method

At present, power battery mainly adopts “constant current and constant voltage” two-stage charging method. In most of the charging time process, the charging power is basically unchanged. Therefore, the actual charging process can be simplified as a constant power charging process.

### II. B. 4) Electric Vehicle User Behavior Characteristics

In this paper, we consider the probability distribution of the main 3 factors of user behavior that will affect the distribution law of electric vehicle charging load.

(1) Charging start moment

According to the statistical research in the literature, it is assumed that the owner's return time from the last trip of the day is taken as the charging start moment, i.e., charging starts when the car is used up for the last time of the day, and the start charging time is approximated to obey the normal distribution by adopting the great likelihood estimation method on the data, and its probability density function is shown in Equation (13):

$$f_t(x)=\begin{cases}\dfrac{1}{\sigma_t\sqrt{2\pi}}\exp\left[-\dfrac{(x-\mu_t)^2}{2\sigma_t^2}\right] & (\mu_t-12)<x\le 24\\ \dfrac{1}{\sigma_t\sqrt{2\pi}}\exp\left[-\dfrac{(x+24-\mu_t)^2}{2\sigma_t^2}\right] & 0<x\le(\mu_t-12)\end{cases} \tag{13}$$

where, $f_t(x)$ is the start charging time probability density function. $\mu_t$ and $\sigma_t$ are constants, denoting the mathematical expectation and standard deviation, respectively, $\mu_t=18.2$ and $\sigma_t=3.6$.

(2) Daily miles traveled

Firstly, the data are normalized, and then the great likelihood estimation method is used to approximate the electric vehicle driving mileage obeys the lognormal distribution, and its probability density function is shown in equation (14):

$$f_D(x)=\frac{1}{x\sigma_D\sqrt{2\pi}}\exp\left[-\frac{(\ln x-\mu_D)^2}{2\sigma_D^2}\right] \tag{14}$$

where $\mu_D$ is the expected value of daily mileage traveled, take $\mu_D=3.40$. $\sigma_D$ is the standard deviation of daily driving mileage, take $\sigma_D=0.91$.

(3) Charging time

The charging time of EV can be calculated based on its daily driving mileage, 100km power consumption and charging power, and the charging length of EV can be estimated as:

$$T_C=\frac{DE}{100P_C} \tag{15}$$

where, $T_C$ is the length of time required for charging, $h$. $D$ is the daily driving mileage, km. $E$ is the 100km power consumption, kW·h/100km. $P_C$ is the charging power, kW.

According to Eq. (15) and the density function of daily driving mileage, the probability density function of charging duration of electric vehicle is obtained as:

$$f_{T_c}(x)=\frac{1}{x\sigma_{T_c}\sqrt{2\pi}}\exp\left[-\frac{(\ln x-\mu_{T_c})^2}{2\sigma_{T_c}^2}\right] \tag{16}$$

where $\sigma_{T_c}=\sigma_D=0.91$. $\mu_{T_c}=\ln\left(\dfrac{E}{100P_C}\right)+\mu_D$.

The process of obtaining the load forecast curve in this paper is shown in Fig. 1.

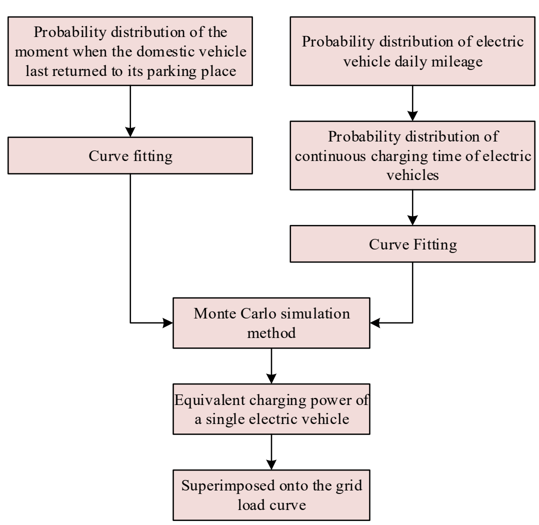


Figure 1: Acquisition process of load forecasting curve

### II. C.Key indicators of grid load

(1) Peak load

Peak load is a commonly used indicator to measure the load situation of the power grid. After large-scale electric vehicles are connected to the power grid, the peak load may exceed the transmission limit of the power grid, breaking the balance between supply and demand of the power system, which in turn affects the stable operation of the power grid.

(2) Load variance

Load variance is a common indicator of load fluctuation. Let the power factor and nodal voltage remain unchanged during normal operation of the grid, i.e., $\cos\theta$ and $U$ are fixed. $P(t)$ is the load curve of a place, $P_s$ indicates that the load is stable, which is the ideal state of the load, and the two do the same work in time $T$, i.e.:

$$\int_0^T P_s dt = \int_0^T P(t)dt \tag{17}$$

Poor power loss:

$$\begin{aligned}\Delta P &= \int_0^T \left(\frac{P(t)}{\cos\theta}\right)^2 \frac{R}{U^2}dt - \int_0^T \left(\frac{P_s}{\cos\theta}\right)^2 \frac{R}{U^2}dt \\ &= \frac{R}{U^2\cos^2\theta}\left(\int_0^T P^2(t)dt - P_s^2 T\right) \\ &\geq \frac{R}{U^2\cos^2\theta}\left(\frac{1}{T}\left(\int_0^T P(t)\right)^2 dt - P_s^2 T\right) \\ &= \frac{R}{TU^2\cos^2\theta}\left(\left(\int_0^T P(t)\right)^2 dt - (P_s T)^2\right) \geq 0\end{aligned} \tag{18}$$

So $\Delta P = 0$ if and only if $P(t) = P_s$.

The variance $\sigma^2$ of $P(t)$ is:

$$\begin{aligned}\sigma^2 &= \int_0^T \left(P(t) - \frac{1}{T}\int_0^T P(t)dt\right)^2 dt \\ &= \int_0^T (P(t) - P_s)^2 dt \\ &= \int_0^T P^2(t) - P_s^2 T\end{aligned} \tag{19}$$

So there:

$$\Delta P = \frac{R\sigma^2}{U^2\cos^2\theta} \tag{20}$$

From equation (20), $\Delta P \propto \sigma^2$, that is, the power loss difference varies with the variance, the smaller the variance, the smaller the power loss, so the variance is an important parameter when measuring the load fluctuation and studying the optimization problem of the load curve.

(3) Peak-to-valley ratio

The peak-valley difference rate can well reflect the load fluctuation. It is defined as:

$$\text{Peak-to-valley ratio} = \frac{\text{Peak load} - \text{Valley load}}{\text{Peak load}} \times 100\% \tag{21}$$

### II. D.Charging load calculation based on Monte Carlo simulation

#### II. D. 1) Monte Carlo algorithm

Monte Carlo method [22] is a simulation statistical method, the basic principle can be expressed as follows: random variable $X_1, X_2, X_3 \dots X_N$ probability distribution is known, assuming that the function $Y = f(X_1, X_2, X_3 \dots X_N)$, through the Monte Carlo method to randomly extract the value of the random variable $(x_{1j}, x_{2j}, x_{3j} \dots x_{nj})$, and then according to the $Y$ to $X_1, X_2, X_3 \dots X_N$ function relationship can be obtained $Y$ function value $y_j$, as shown in equation (22):

$$y_j = f(x_{1j}, x_{2j}, \ldots x_{nj}) \tag{22}$$

where, $j$ denotes the number of random sampling times, repeated multiple sampling $(j = 1, 2, 3 \ldots m)$, you can get the sampling data of function $Y$. According to the central limit theorem it can be seen that the sampling results are subject to normal distribution, with the increase in the number of sampling times, $j$ value is larger, the probability distribution of the function $Y$ and the numerical characteristics of the function $Y$ is obtained closer to the actual value.

### II. D. 2) Calculation steps for EV charging load modeling

In this paper, a day is divided into 24 time slots with a time interval of 1 hour, i.e., $i = 1, 2, 3 \ldots 24$, then the total charging load of $N$ electric vehicles in the $i$ th time slot can be expressed as:

$$P_{EVi} = \sum_{n=1}^{N} P_{n,j} \tag{23}$$

where, $P_{EVi}$ denotes the total charging load of $N$ electric vehicles in the $i$ th time period, $P_{n,j}$ denotes the charging load of the $n$ th vehicle in the $i$ th time period, and $N$ denotes the number of electric vehicles which are charging in the $i$ th time period.

The calculation flow for calculating the charging load of electric vehicles based on Monte Carlo method is shown in Fig. 2, and its calculation steps are as follows:

(1) Set the number of vehicles, penetration rate, and charging power size.

(2) Generate a random number of start charging time according to equation (13) EV start charging moment start density function;

(3) According to equation (14), using the probability density function of the daily driving distance of the vehicle, generate a random number of random daily driving distance.

(4) On this basis, using the electric vehicle charging power, 100-kilometer energy consumption and daily driving distance, and then through the charging duration probability density function, to find the electric vehicle charging duration.

(5) Superimpose the charging load for each time period to generate the electric vehicle cluster charging load for each time period. Averaged over each EV to generate a single EV equivalent charging load.

(6) Calculate the charging load of $N$ EVs, superimpose it with the original load of the grid, and generate the grid load prediction curve.

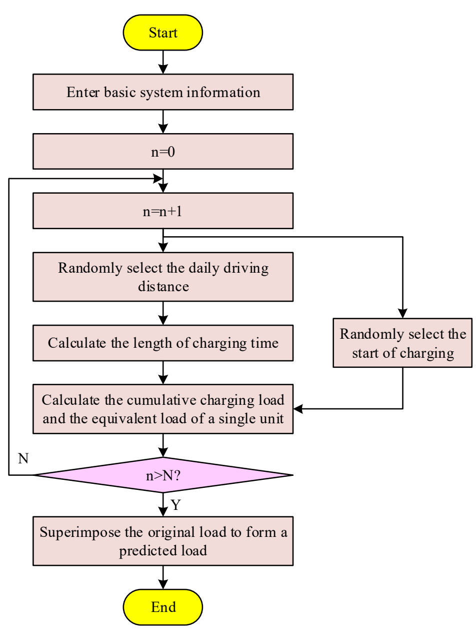


Figure 2: Flowchart of Monte Carlo Simulation calculation

# III. Bayesian network based charging pile state data inference

In this chapter, Bayesian networks are utilized to conduct charging pile state data inference experiments to provide data support for studying the impact of uncontrolled charging loads from electric vehicle charging piles on the power grid, as well as realizing effective governance of the power grid.

### *III. A. Charging pile network data extrapolation experiments*

Consider three baseline inference methods:

(1) Last Known: inferring the current state from the last known state of the charging pile. This is a commonly used method in current practice. This method only utilizes the timing relationship.

(2) KNN: The state of the charging pile is inferred from the average state of the K nearest neighbors around the target charging pile. Neighbors” here are not spatially close charging piles, but similar charging piles based on the similarity of historical state data. Although this method utilizes historical data, it only considers pairwise relationships between charging piles.

(3) BN-ST: This method is a spatio-temporal autoregressive method that differs from the algorithm proposed in this paper in that it does not take into account user behavior patterns. By comparing with this benchmark, the role of considering user behavior patterns can be reflected.

For brevity, the method in this paper is denoted as BN-STU.

The accuracy of data inference for the charging pile network is shown in Fig. 3. In 80 experiments, the average accuracy of this paper's method is 93.3%, which is 7.49% higher than that of the BN-ST method, and 17.80% and 12.14% higher than that of Last Known and KNN, respectively.The Last Known method performs the worst because it only utilizes simple timing relationships and is sensitive to the length of delay time. If the delay is small, then Last Known should perform better because it only makes errors when charging starts and ends alternately. However, once the delay time increases, the Last Known state is no longer a good indicator of the true state.KNN performs better than Last Known, but is very unstable. This may be due to the fact that the utilization of charging posts is still relatively low and the distribution of charging stations is still sparse at this stage. However, KNN mainly exploits the similarity between charging piles and its performance decreases if there are unknown states in the neighbors. Because BN-STU considers user behavior patterns, BN-STU performs better than BN-ST. Since the charging post network is an open system, i.e., users can charge their EVs outside of the public charging facilities, i.e., private charging posts, the effect of user behavior patterns may not be obvious, and thus BN-STU does not outperform BN-ST by much. However, if the system has more knowledge about the users, the contribution of user behavior patterns may be increased. If all user activities are available in the network of charging piles, it should be able to give a good enough inference result even without relying on information from the infrastructure side.

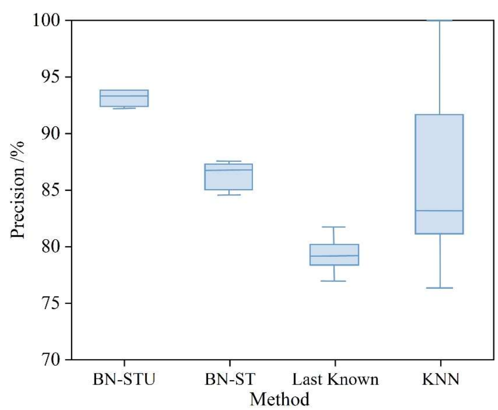


Figure 3: Precision of data inference for the charging pile network

The recall rate inferred from the charging pile network data is shown in Fig. 4, where the recall rate of BN-STU reaches 94.5%, which is 9.63%, 18.13%, and 95.65% higher than that of BN-ST, Last Known, and KNN, respectively. KNN performs the worst because in the case of low utilization of charging piles and sparse distribution of charging piles, there are fewer occupied charging piles, so inferring the current state of a charging pile by looking at the state of other piles is not very reliable. The temporal pattern is more pronounced than the spatial pattern, which also makes the KNN perform poorly.Last Known performs better than the KNN due to the fact that the charging piles may remain occupied or idle for quite a long period of time, e.g., several hours. BN-STU still performs better than BN-ST due to the consideration of user behavior patterns.

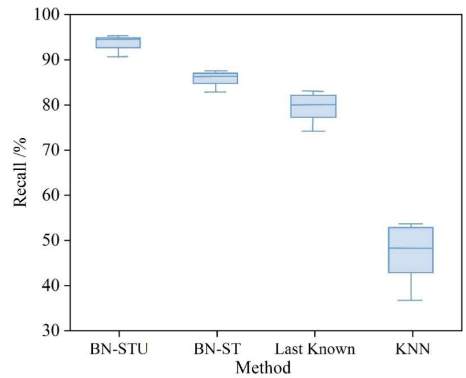


Figure 4: Recall of data inference for the charging pile network

### *III. B. End-to-end QoS enhancement experiments*

In order to evaluate the impact of inferential modeling on end-to-end quality of service, this paper employs an application scenario with a case study. The work proposes a charging station recommendation system for electric cabs. The system uses in-collected GPS data of electric cabs to recommend charging stations to electric car rental drivers in order to reduce the drivers' cost time and waiting time. Cost time is the time spent from generating a charging demand to starting charging. Waiting time is the time spent waiting in line for charging after reaching the charging station. This work detects the driver's charging intention and makes recommendations based on the patterns in the GPS track.

In this section, simulations are performed using charging station and user data. The map data uses the open source dataset OpenStreetMap.An open source navigation and path planning tool is used in the simulation. Two benchmark models are considered in this section. The first one is an optimistic model, i.e., it is always assumed that charging piles with unknown status are available. The other is a pessimistic model, which always assumes that a charging post with unknown status is occupied. The latter is the approach commonly used in current practice.

The variation of the cost time in a day is shown in Fig. 5. It can be seen that the method in this paper has the smallest cost time overall.

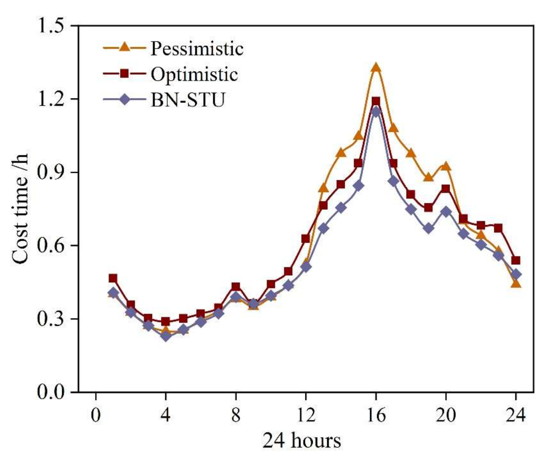


Figure 5: Time cost

The distribution of EV charging intervals is shown in Figure 6. Since Figure 6 was used to generate charging demand, it is similar to the trend in Figure 5. Combining Fig. 5 and Fig. 6 shows that when in the low-peak hour, i.e., from 0:00 to 8:00, the performance of the three methods is close to each other because there are few charging events and charging piles are very sufficient. However, during peak hours, the cost time of Pessimistic increases rapidly. This is because there are many charging events, and if a wrong assumption is made, the driver has to wait in line or find a new charging station, resulting in an increase in cost time.

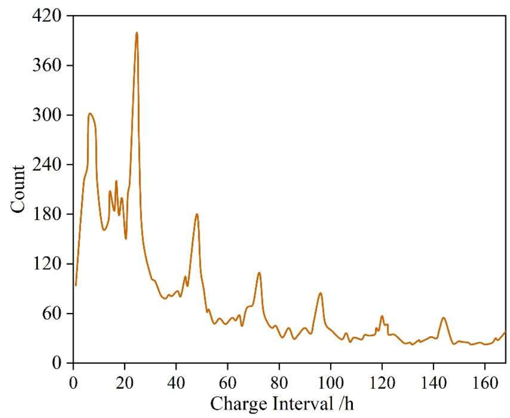


Figure 6: Distribution of charging intervals for electric vehicles

The waiting time curve is shown in Fig. 7. It can be seen that the method in this paper usually reduces the waiting time, while Pessimistic increases the waiting time less during peak hours. This is because Pessimistic will always recommend charging posts that are definitely available, and therefore drivers usually do not need to wait in line for too long. On the contrary, since Optimistic may recommend a driver to a charging station that is already occupied, its waiting time is higher overall.

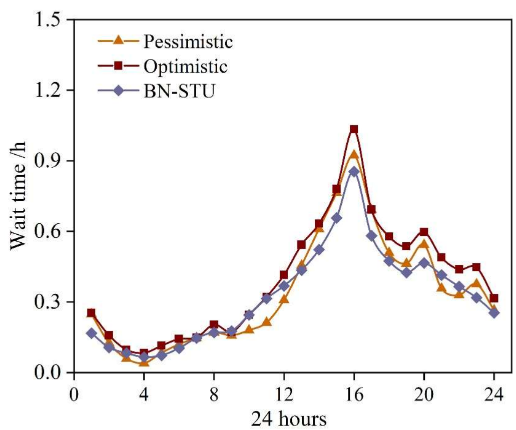


Figure 7: Waiting time

The user satisfaction curve is shown in Figure 8. In the simulation, it is assumed that user satisfaction decreases when the following situations occur:

(1) The user needs to travel a longer distance to reach the recommended location.

(2) The user is incorrectly recommended to a charging post that is already occupied.

(3) The user needs to wait in line for a longer time after arriving at the recommended charging station.

The penalty coefficients for these three cases are 0.3, 0.4, and 0.3. For example, if a user arrives at a charging station after traveling 12 km and finds that the charging station is already occupied, and the user waits for 50 minutes before charging, the satisfaction level of the service will be 1-0.3-0.4-0.3 = 0,.

It can be seen that the method in this paper achieves higher user satisfaction overall, especially during peak hours. For Pessimistic, it never recommends charging posts with unknown status to the user, thus there are fewer charging posts available overall, which leads to more queuing and increases the waiting time of the user.Optimistic performs slightly better than Pessimistic because it recommends charging posts with unknown status to the user for the time being, which reduces the queuing to a certain extent, but may lead to wrong recommendation.BN-STU performs the best because it tries to provide users with correct information about charging piles, avoiding both over-conservatism that leads to queuing and over-aggressiveness that misleads users. However, during peak hours, BN-STU's performance is constrained by limited charging resources, and users still have to wait in line for others to finish charging before they can start charging.

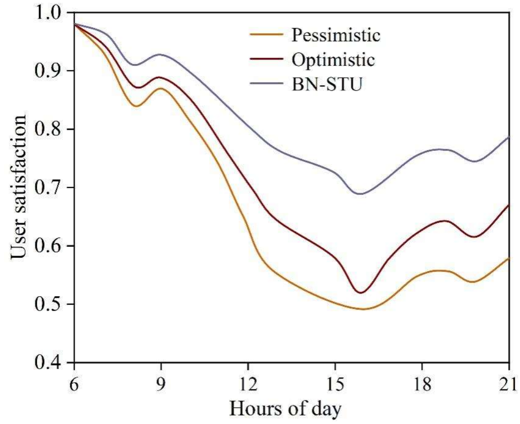


Figure 8: User satisfaction

# IV. Analysis of the impact of uncontrolled charging of electric vehicles on regional distribution grids

### IV. A. Expansion analysis

In order to study the impact on the regional distribution network in the future when there are large-scale electric vehicles accessing charging at the same time in multiple residential communities, this section takes the IEEE33 node distribution system as the object of study, treats each node as a residential community, and adds disorderly charging electric vehicle power loads to the system, and researches and analyzes the impact of the regional distribution network in terms of feeder loads, voltage offsets, and network losses by charging loads.

### IV. B. IEEE33 node power distribution system

#### IV. B. 1) Algorithm design

In this chapter, the IEEE33 node system is used as a research object to simulate the behavior of orderly and disorderly charging of electric vehicles in the regional distribution network, and each node is equated to the total load of a residential district, and some of the nodes are selected as the district that contains the charging load of electric vehicles. In the system, firstly, a balance point is selected, and secondly, in order to simulate the residential community load more realistically, nodes 8, 14, 20, 24 and 28 are selected as residential communities containing EV charging loads to be treated, and the rest of the nodes are treated as residential communities not containing charging loads.The topology of the IEEE33 node electric network system is shown in Fig. 9, and the parameters of the system are set as follows: the active load is 3850kW, reactive load of 2400kW, rated voltage of 15kV, 5 disconnecting contact switches and 32 branch circuits.

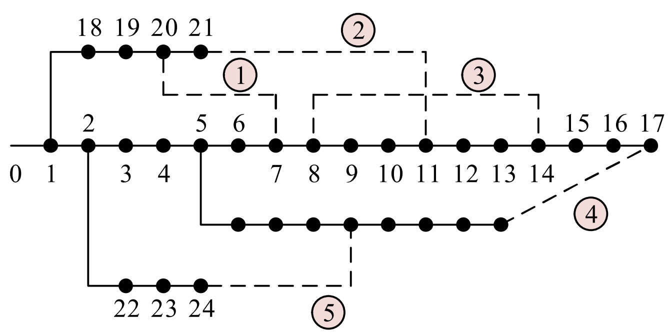


Figure 9: Topology of the IEEE 33-node electrical network system

#### IV. B. 2) Methods and Processes for Calculating Distribution System Trends

In this paper, the Newton Raphson method is used to realize the trend calculation of the distribution network using MATLAB software as a tool, and the electric vehicles in different penetration rates under the uncontrolled charging mode are connected to the load nodes of five residential districts to carry out the trend calculation of the whole distribution network. Due to the IEEE33 standard distribution system given the initial value of the high convergence of the solution speed, this paper does not bring in the node real load data for the solution, but follows the original data, at the same time, the electric vehicle load will be scaled down to superimpose the calculation with the original data.

Newton Raphson's method, which has good convergence but requires high initial values, is an efficient method for solving nonlinear algebraic equations. The essence of Newton Raphson's method is linearization step by step, using Taylor's series to expand the equation $F(x)=0$ and ignoring the second and higher orders for solving, and transforming the process of solving nonlinear equations into the corresponding linear equations through many iterations. Since the Newton Raphson method is a relatively mature tidal current calculation method, this method is only used as an arithmetic tool in this paper, and not as a key research object. The main purpose of this chapter is to verify the impact of the load (EV access) changes of each cell node on the regional power distribution system under the regulation of EV orderly charging strategy. Since the initial value of the IEEE33 node system used in this paper has good convergence and is suitable for the requirements of the Newton Raphson method for solving nonlinear equations, the initial value of each node load is treated as the regular load of the residential neighborhoods, and the electric vehicle electric loads counted in the calculation are solved as the new loads.

### *IV. C. Analysis of simulation results*

#### IV. C. 1) Distribution network base load

The distribution network loads other than distributed PV generation loads and EV charging loads are referred to as base loads, which indicate the daily electricity consumption of residents. The load data of an area in a 24-hour day is shown in Table 1.

Table 1: Load Data of a Certain area for 24 hours a day

| Time /h | Base load /kW | Time /h | Base load /kW |
|---|---|---|---|
| 1 | 91204 | 13 | 83741 |
| 2 | 84375 | 14 | 79356 |
| 3 | 80341 | 15 | 78462 |
| 4 | 78524 | 16 | 83753 |
| 5 | 77932 | 17 | 94286 |
| 6 | 78643 | 18 | 103562 |
| 7 | 84059 | 19 | 110629 |
| 8 | 96248 | 20 | 115384 |
| 9 | 96375 | 21 | 117429 |
| 10 | 93824 | 22 | 116357 |
| 11 | 90418 | 23 | 109472 |
| 12 | 88526 | 24 | 101534 |

In order to be able to represent the trend of daily power consumption in the simulation example, this paper will be the data in Table 1 for the Mississippi value processing, in the corresponding moment and the IEEE33 node distribution network model of the maximum active power of 3850kW multiplication, to get the base load curve in a day of 24 hours shown in Fig. 10, and then according to the proportion of the proportion of each node to get the load at each node in a day of 24 hours.

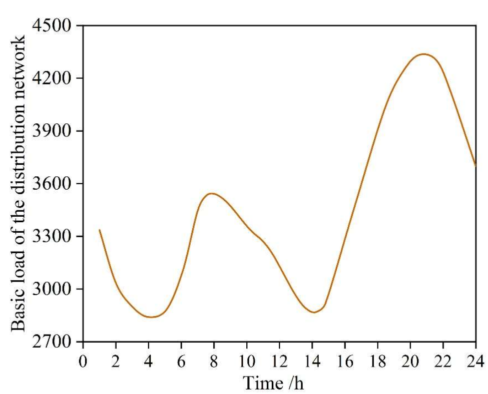


Figure 10: Daily load curve

**IV. C. 2) Impact of uncontrolled charging of electric vehicles on distribution network line losses**

Disorganized EV charging includes: uncertain charging time, different penetration rates within the region, and uncertain charging locations. This section mainly analyzes the impact of uncontrolled charging of electric vehicles on line loss from the penetration rate of electric vehicles and charging locations.

(1) Different penetration rates of electric vehicles

Assuming that the distribution network load can accommodate a maximum of 1,200 electric vehicles, the penetration rate of electric vehicles is:

$$\eta = \frac{N_{ev}}{N_T} \times 100\% \tag{24}$$

where, $N_{ev}$ is the number of EVs in the station area, and $N_T$ is the maximum number of EVs to be accommodated.

In this section, the EV penetration rate is set to be 0%, 20%, 40%, 60%, 80% and 100% at 12 nodes for simulation and analysis, and the results of total line loss changes under different penetration rates are shown in Fig. 11.

Under different penetration rates, the system line loss peaks all appear in the 18:00-23:00 time period, which is consistent with the daily travel habits of residents. As the penetration rate of electric vehicles increases, the line loss of the distribution network gradually increases during the peak hours of electricity consumption.

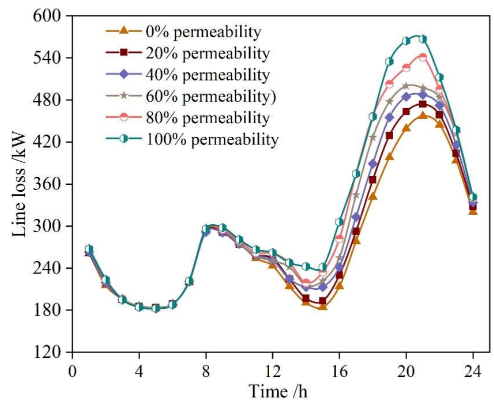


Figure 11: Changes in total line loss at different permeability levels

(2) Different charging locations for electric vehicles

In this section, 1000 electric vehicles are set to charge at nodes 2, 4, 8, 12, 16, 20, 24, 28, 31, etc., and the line loss of each node for 24 hours a day is obtained through simulation as shown in Fig. 12.

The loss situation caused by each node in Fig. 12 is compared and analyzed with the points in Fig. 9: the closer the EV charging points are to the distribution network power supply, the smaller the line loss caused by uncontrolled charging of EVs, and the closer the charging points are to the end of the feeder line, the larger the line loss is caused. The daily line loss curves at different nodes have a great similarity with the trend of daily load curves, and the peak of line loss occurs in the time period of 18:00-23:00.

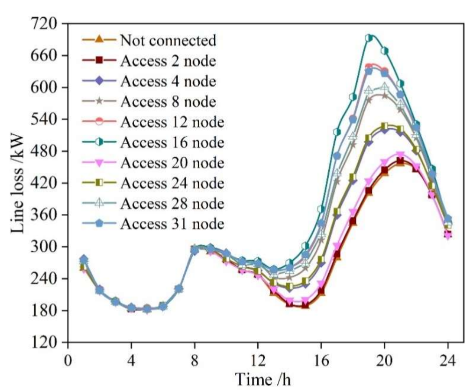


Figure 12: Changes in total line loss at different charging points

#### IV. C. 3) Impact of uncontrolled charging of electric vehicles on distribution network voltage

Through the above analysis, it is obtained that in the daily base load curve of the distribution network, the load peaks in the time period of 18:00-23:00, and the line losses caused by the charging locations and penetration rates of different EVs all peak in this time period. In order to study the impact of uncontrolled charging of EVs on the distribution network voltage, this section analyzes the distribution network voltage deviation in terms of different penetration rates and charging locations and selects the 20:00 moment.

(1) Different penetration rates of electric vehicles

In this section, the penetration rate of electric vehicles is set at 12 nodes, which is 0%, 20%, 40%, 60%, 80%, 100%, etc. The simulation analysis is carried out, and the results of the voltage change of each node under different penetration rates are shown in Fig. 13.

Through the analysis, it is obtained that the voltage at each node is lower as the penetration rate of electric vehicles increases. The trend of voltage change at each node under different penetration rates is the same as the trend of voltage change without EV charging.

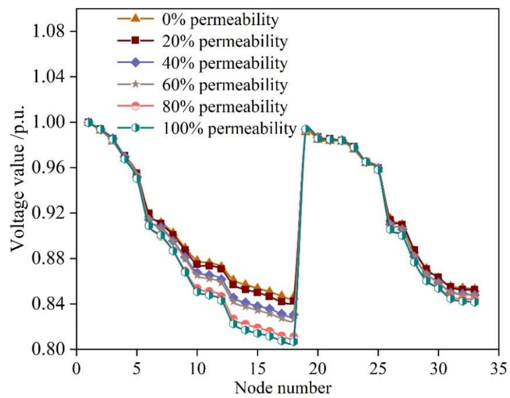


Figure 13: The voltage variation of each node under different permeability

(2) Different charging points for electric vehicles

In this section, 1000 electric vehicles are set to be connected to charge at nodes 2, 4, 8, 12, 16, 20, 24, 28, 31, etc., and the voltage of each node for 24 hours a day is obtained through simulation as shown in Fig. 14. It can be seen that the closer the charging point location is to the power end of the distribution network, the smaller the amount of voltage deviation is, and the closer the end of the feeder, the larger the amount of voltage deviation is.

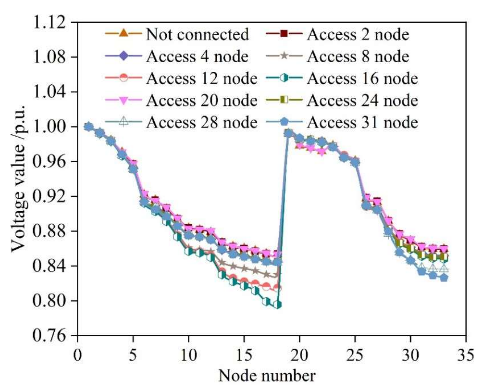


Figure 14: Voltage variations at different charging points

## V. Conclusion

This study effectively solves the problem of the impact of large-scale access of electric vehicles on the stability of grid operation by constructing an electric vehicle charging pile disordered charging load governance model based on Bayesian networks. The Bayesian network method shows good performance in charging pile state data inference and has significant advantages over traditional methods. In the end-to-end quality of service assessment, the proposed method is effective in reducing user cost time and waiting time, especially in peak hours to achieve higher user satisfaction.

The simulation results of the IEEE33 node distribution system show that the impact of EV charging loads on the distribution network has obvious temporal and spatial distribution characteristics, and the peak line loss and voltage deviation are mainly concentrated in the time period of 18:00-23:00, which is highly consistent with the daily travel habits of residents. There is a significant difference in the degree of influence of the charging point location on the power grid, the closer the charging point is to the power end causes less line loss, while the charging point close to the end of the feeder leads to a larger voltage deviation. The K-S test results verify the validity of the Bayesian network model, and the simulated data and the distribution of the original data have a good consistency.

This study provides a new technical way for electric vehicle charging load management, which helps to improve the adaptability of the power grid to large-scale electric vehicle access, guarantee the safe and stable operation of the power system, and has important theoretical significance and practical value for promoting the healthy development of the electric vehicle industry and the intelligent upgrading of the power grid.

## References

[1] Usmanova, R. M., Sattarova, N. A., & Boyko, N. N. (2021, March). Influence of automobiles on environmental pollution. In IOP Conference Series: Materials Science and Engineering (Vol. 1079, No. 6, p. 062040). IOP Publishing.

[2] Swami, A. (2018). Impact of automobile induced air pollution on roadside vegetation: A review. ESSENCE International Journal for Environmental Rehabilitation and Conservation, 9(1), 101-116.

[3] Hua, T. K., Azarov, V., & Kutenev, V. (2022). Modern Invisible Hazard of Urban Air Environment Pollution When Operating Vehicles That Causes Large Economic Damage. Authorea Preprints.

[4] Adeyanju, A. A., & Manohar, K. (2017). Effects of vehicular emission on environmental pollution in Lagos. Sci-Afric J Sci Issues Res Essays, 5(4), 34-51.

[5] Wu, Y., & Zhang, L. (2017). Can the development of electric vehicles reduce the emission of air pollutants and greenhouse gases in developing countries?. Transportation Research Part D: Transport and Environment, 51, 129-145.

[6] Gao, Z., Xie, H., Yang, X., Zhang, L., Yu, H., Wang, W., ... & Chen, S. (2023). Electric vehicle lifecycle carbon emission reduction: A review. Carbon Neutralization, 2(5), 528-550.

[7] Wang, M., Wang, Y., Chen, L., Yang, Y., & Li, X. (2021). Carbon emission of energy consumption of the electric vehicle development scenario. Environmental Science and Pollution Research, 28, 42401-42413.

[8] Hoehne, C. G., & Chester, M. V. (2016). Optimizing plug-in electric vehicle and vehicle-to-grid charge scheduling to minimize carbon emissions. Energy, 115, 646-657.

[9] Kang, X., Nie, H., Gao, M., & Wu, F. (2023). Research on carbon emission of electric vehicle in its life cycle. Energy Storage Science and Technology, 12(3), 976.

[10] Hua, L., Wang, J., & Zhou, C. (2014). Adaptive electric vehicle charging coordination on distribution network. IEEE Transactions on Smart Grid, 5(6), 2666-2675.

[11] Deb, S., Tammi, K., Kalita, K., & Mahanta, P. (2018). Impact of electric vehicle charging station load on distribution network. Energies, 11(1), 178.

[12] Leou, R. C., Su, C. L., & Lu, C. N. (2013). Stochastic analyses of electric vehicle charging impacts on distribution network. IEEE Transactions on Power Systems, 29(3), 1055-1063.

[13] De Hoog, J., Alpcan, T., Brazil, M., Thomas, D. A., & Mareels, I. (2014). Optimal charging of electric vehicles taking distribution network constraints into account. IEEE Transactions on Power Systems, 30(1), 365-375.

[14] Zeb, M. Z., Imran, K., Khattak, A., Janjua, A. K., Pal, A., Nadeem, M., ... & Khan, S. (2020). Optimal placement of electric vehicle charging stations in the active distribution network. IEEE Access, 8, 68124-68134.

[15] Rahman, S., Khan, I. A., & Amini, M. H. (2020, September). A review on impact analysis of electric vehicle charging on power distribution systems. In 2020 2nd International Conference on Smart Power & Internet Energy Systems (SPIES) (pp. 420-425). IEEE.

[16] Liu, J. P., Zhang, T. X., Zhu, J., & Ma, T. N. (2018). Allocation optimization of electric vehicle charging station (EVCS) considering with charging satisfaction and distributed renewables integration. Energy, 164, 560-574.

[17] Awasthi, A., Venkitusamy, K., Padmanaban, S., Selvamuthukumaran, R., Blaabjerg, F., & Singh, A. K. (2017). Optimal planning of electric vehicle charging station at the distribution system using hybrid optimization algorithm. Energy, 133, 70-78.

[18] Yang, X., Niu, D., Sun, L., Ji, Z., Zhou, J., Wang, K., & Siqin, Z. (2021). A bi-level optimization model for electric vehicle charging strategy based on regional grid load following. Journal of Cleaner Production, 325, 129313.

[19] Xiaogang Wu, Hanying Zhao, Wentao Xu, Wulue Pan, Qingfeng Ji & Xiujuan Hua. (2024). Fault diagnosis of the distribution network based on the D-S evidence theory Bayesian network. Frontiers in Energy Research,12,1422639-1422639.

[20] Liang Kong & Heng Nian. (2020). Fault Detection and Location Method for Mesh-type DC Microgrid using Pearson Correlation Coefficient. IEEE Transactions on Power Delivery,PP(99),1-1.

[21] Chu Fei, Lu Zhenlin,Jin Shuowei, Liu Xin & Yu Ziyang. (2022). A relaxed support vector data description algorithm based fault detection in distribution systems. Frontiers in Energy Research,10,

[22] Minghong Liu, Mengke Liao, Pengchao Wang, Changling Li & Xuehua Zhao. (2025). Multi-timescale power self-balancing optimization and regulation of remote rural microgrids based on stochastic monte carlo method. Journal of Physics: Conference Series,2960(1),012014-012014.