Research on Site Selection and Capacity Determination of Electric Vehicle Public Charging Stations by Integrating K-Means++ and Improved RODDPSO

*Birong HUANG [1], Zilong WANG [1], Jianhua CHEN [1], Bingyang ZHOU [1], Yuhang ZHU [2], Yan LIU [3, *]*

[1] College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China
[2] School of Communication and Artificial Intelligence, Nanjing Institute of Technology, Nanjing 211167, China
[3] School of Mathematics and Physics, Nanjing Institute of Technology, Nanjing 211167, China

[*] liuyan@njit.edu.cn

Submitted January 15, 2025 / Accepted February 19, 2025 / Online first April 9, 2025

Abstract. *To address the suboptimal spatial distribution and low comprehensive utilization of existing electric vehicle (EV) public charging infrastructure, this study proposes an innovative charging station placement and capacity determination methodology integrating K-Means++ clustering with an enhanced RODDPSO variant. Building upon conventional K-Means and RODDPSO frameworks, we develop an improved hybrid algorithm incorporating three critical advancements: 1) an adaptive mutation mechanism within the RODDPSO architecture to enhance global search capabilities and prevent premature convergence; 2) synergistic optimization of K-Means++ cluster centroids through the enhanced RODDPSO operator; and 3) a novel cluster validation metric based on real-world utilization patterns. The proposed methodology effectively resolves the inherent limitations of conventional K-Means approaches, particularly their sensitivity to initial centroid selection and tendency toward local optima. Empirical validation through a case study of Nanjing's charging infrastructure demonstrates the algorithm's superior performance: stations sited using the proposed hybrid method exhibit 63.8% greater spatial correlation with high-utilization zones (>15% operational utilization) compared to baseline K-Means implementations. The advancements provide both methodological contributions to spatial optimization algorithms and practical insights for urban EV infrastructure planning.*

Keywords

K-Means++, variation randomly occurring distributedly delayed particle swarm optimization, public charging station, siting and capacity determination

1. Introduction

With the growing awareness of environmental protection and the ongoing transformation of the energy structure, electric vehicles (EVs), as a clean and sustainable mode of transportation, are rapidly becoming the mainstream choice for future mobility. The widespread adoption and promotion of EVs, however, rely heavily on the construction of adequate charging infrastructure. Public charging stations, as the primary service providers for the general population, deserve significant attention in this context. Despite their importance, the current charging network faces several critical challenges, including an insufficient number of public charging piles, uneven distribution, low utilization rates, and subpar user service quality. These issues manifest in various ways: overcrowding and long queues at charging hot spots, a lack of charging facilities in certain areas, prolonged idleness of charging piles in other regions, and even some stations being left unattended throughout the year. Such inefficiencies fail to adequately meet both the current and anticipated future charging demands of surrounding areas.To address these challenges, it is imperative to optimize the siting and capacity planning of charging stations. This not only helps reduce construction and operational costs but also enhances the comprehensive utilization rate of charging infrastructure. Consequently, research on the siting and capacity optimization of public EV charging stations has become a critical focus for improving the efficiency and sustainability of the charging network. In existing studies, EV charging load has been identified as a critical factor in the deployment of public charging stations. Literature [1] examined EV charging loads using a standard IEEE 33-bus radial distribution system and applied the condor search algorithm and cuckoo search algorithm for siting and capacity determination, effectively evaluating the methods' performance. At the same time, there are many factors that can affect the load on the grid. Literature [2] proposed a robust optimization model for siting charging stations in micro grids composed of wind power, photovoltaic, and distributed energy sources. By combining road networks and power grids, this study used load fluctuation rates to assess the fit between renewable energy output curves and charging demand curves, simultaneously addressing capacity determination.

Economic factors also play a pivotal role in charging station optimization. Literature [3] considered economic efficiency and the impact of fast charging stations on the grid to determine optimal locations. Literature [4] evaluated existing EV charging stations by analyzing charging locations, times, connector types, architectural combinations, and the performance of power converters, aiming to mitigate uncertainties caused by transmission grids. Different types of charging stations with different sources of electricity can also affect the charging load. Literature [5] proposed a modified multi-objective salp swarm-based optimization algorithm to allocate photovoltaic (PV), wind turbines (WT), and EV charging stations in a micro grid. All of the above literature discusses the impact of grid load on charging station layout, but none of them consider the overall number of charges and charging loads within the administrative boundaries.

Particle swarm algorithms are widely used in strategy optimization and prediction, and there are more improvement algorithms for them. Literature [6] proposed a multi-objective full-parameter optimization particle swarm algorithm that integrates all the factors affecting the global and local search capabilities. Literature [7] proposed a novel particle swarm algorithm applied to the economic emission load dispatch problem. Some of the particle swarm optimization algorithms mainly focus on improving the efficiency and speed of the algorithm. Literature [8] used simulated annealing immune particle swarm optimization algorithm. The simulated annealing algorithm is used for the global update of the particle swarm algorithm, and the immune mechanism is introduced to participate in the iterative update of particles to improve the speed and efficiency of the particle swarm algorithm. Literature [9] proposes the use of hybrid genetic algorithm and particle swarm algorithm to optimize the allocation of plug-in electric vehicle charging stations in the distribution network. Literature [10] uses an improved particle swarm algorithm to optimize the problem of avoiding over-voltage caused by power injection from PV power sources and voltage dips in the distribution network caused by EV charging. There is also a lot of literature on charging station layout using clustering algorithms. Literature [11] uses algorithms such as K-Means, DBSCAN, and CFS to cluster demand points and derive the location of charging stations. However, their clustering algorithms do not optimize the initial cluster center location. A well-performing initial clustering center point is crucial to the merit of the clustering effect. There are many algorithms other than using clustering and PSO algorithms that can be equally applied in accomplishing the charging station layout. Literature [12] predicts the EV charging demand through the encoder-decoder depth architecture of GCN and optimizes the competitive resource allocation strategy for charger planning through the Cournot competitive game model. A new parallel computing algorithm is proposed to seek the Cournot competitive equilibrium through convergence analysis.GCN is mainly used for predicting suitable siting areas, such as predicting the charging demand at a certain location, the impact on the grid, etc., but relies on a large amount of high-quality graph data, which is otherwise of limited effectiveness, and suffers from the problems of high computational cost and the need for a long time to train the neural network, and is not good at solving the optimization problem directly. While particle swarm algorithm is directly used to optimize the charging station location, capacity and other parameters to achieve the optimal goal, and the computational amount is small, suitable for multi-objective optimization, no graph structure data and other optimization problems.

Minimizing travel distance and waiting time has also been a major focus in siting and layout research. Literature [13] used a genetic algorithm to simulate charging station locations, reducing travel distance and waiting time. Literature [14] applied a decision experiment and proportionality assessment technique to identify six sustainable charging station sites, minimizing travel distances. Literature [15] optimized charging station locations and capacities by employing a multi-objective evolutionary algorithm (NSGA-II and MOEA/D-gen) based on urban mobility distribution models, reducing travel time. Literature [16] combined deep neural networks with the NSGA-II algorithm to improve operational performance, assigning stations based on EV charging levels to minimize setup costs and average waiting times. In addition, the number and quality of charging stations is an important metric. Lastly, Literature [17] proposed an uncertain model converted into deterministic opportunity-constrained programming, accounting for multiple optimization preferences of distribution network operators, charging station owners, and EV users to reduce travel distances and optimize siting. All of the above literatures have used factors such as the user's travel distance to obtain charging demand points for site selection, but none of them have analyzed the user's charging demand directly based on the user's charging demand.

In recent years, a variety of factors and algorithms have been considered in the research on optimal charging station layouts. These studies have achieved significant progress and proposed diverse solutions for optimizing charging station siting and capacity planning. However, many of the existing studies fail to address the specific needs of electric vehicle (EV) users comprehensively. Most focus solely on electric taxis and overlook the primary users of public charging stations, such as internet vehicles, which are prevalent at this stage.

To address these gaps, this paper presents a novel method for EV charging station siting and capacity determination. The proposed approach integrates K-Means++ clustering with the Variational Randomly Occurring Distributedly Delayed Particle Swarm Optimization (VRODDPSO) algorithm. The VRODDPSO algorithm is utilized to optimize the initial clustering centroids of the K-Means++ algorithm, which is then applied to cluster charging demand points. The charging station capacity is determined based on administrative divisions, charging frequency, and the proportion of total power demand.

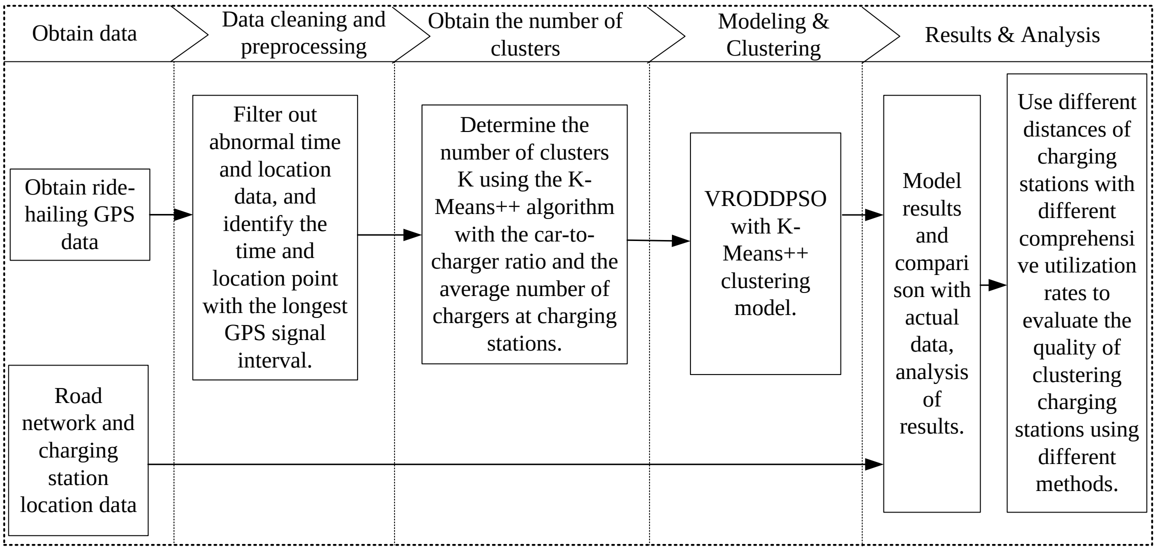

Fig. 1. System architecture.

The main contributions of this paper are as follows:

- Three-stage clustering method for charging station siting: A three-stage clustering method is proposed for EV charging station site selection. Based on the user's point of need, The first stage identifies regional areas, the second stage determines charging pile locations, and the third stage identifies charging station locations. This approach ensures a reasonable and efficient layout of charging stations while reducing computational time and improving operational efficiency.

- Improved K-Means++ algorithm with VRODDPSO optimization: All three clustering stages adopt an improved K-Means++ algorithm, which incorporates the VRODDPSO algorithm to optimize the initial clustering centroids. This enhancement effectively addresses the limitations of the traditional K-Means algorithm, such as falling into local minima and producing sub optimal clustering results due to inappropriate initial centroids.

- New evaluation index for siting performance: A novel measurement index is proposed to assess the advantages and disadvantages of charging station siting under different algorithms. This index evaluates performance based on the comprehensive utilization rate of real-world charging stations, providing a more practical and reliable benchmark for optimization.

The remainder of this paper is structured as follows: Section 2 outlines the system framework structure. Section 3 describes the data analysis and processing methods. Section 4 details the improvements made to the RODDPSO algorithm. Section 5 presents the experimental results, while Section 6 provides an analysis of these results and validates their reliability. Finally, Section 7 concludes the paper.

2. System Architecture

This paper presents a comprehensive strategy for the siting and sizing of electric vehicle (EV) charging stations in urban areas by integrating the K-Means++ and Variational RODDPSO algorithms. The proposed approach demonstrates superior performance compared to the traditional K-Means algorithm, as evidenced by its ability to generate more accurate clustering results that align with the actual utilization rates of existing charging stations.

To achieve optimized charging station locations, this study follows a systematic series of steps designed to ensure the accuracy of the siting process. These steps include data acquisition, data cleaning, per-processing, determining the optimal number of clusters for K-Means++ through the elbow method, modeling and clustering, and analyzing the final results. The overall technical framework of the proposed methodology is illustrated in Fig. 1.

The first step involves collecting essential data, including real-time GPS signals, road network data, geographic locations, and the utilization rates of internet-based ride services' charging stations. Once the data is gathered, it is cleaned to address issues such as missing values and outliers. The K-Means++ algorithm is then applied to determine the optimal number of clusters, categorizing the demand points of online vehicle services.

Next, a model integrating K-Means++ with the variant RODDPSO algorithms is developed. This model clusters regions to ensure that the charging station locations remain within the city's construction boundaries. It then optimizes the placement of charging piles and stations within each region, aiming to maximize user satisfaction among electric vehicle owners.

Finally, the charging station locations are analyzed and optimized, with the performance of existing stations assessed based on comprehensive utilization rates. This analysis provides valuable insights that can assist governments and related agencies in better understanding the performance of charging infrastructure, ultimately supporting infrastructure planning and optimization efforts.

3. Data Analysis and Processing

This paper utilizes Yancheng City's online ride-hailing trip data as the primary data source for analysis, which includes 5 days of GPS signal data from approximately 12 million online ride-hailing trips. The dataset contains essential information such as vehicle ID, latitude and longitude, speed, direction, passenger status, data sending and receiving times, administrative district number, and specific locations. Notably, all data points correspond to trips with passengers. For the purpose of this study, all online vehicles are treated as electric vehicles during data cleaning and processing.The dataset, consisting of 12 million data points, was thoroughly cleaned. Vehicle IDs with anomalous sending times were identified and removed, ensuring that only error-free data remained. Data points with coordinates outside Yancheng City were deleted, and the number of GPS signal data entries for each vehicle ID was checked. If a vehicle had fewer than 100 GPS signals in a single day, it was assumed to have low driving mileage and no charging demand, so such vehicle IDs were deleted. Additionally, records with mismatched sending times were cleaned and removed.

Following the data cleaning, the next step involved processing the cleaned data. Given that the battery capacity of current electric vehicles typically allows them to operate for less than a day, it was assumed that each vehicle would have only one charging demand per day. Since GPS signals are primarily generated when the vehicle is taking orders, and are absent when the vehicle is idle, the geographic location corresponding to the longest interval between two GPS signals is treated as the charging demand point.

The GPS signals of each eligible vehicle were sorted, and the sending and receiving times of the data were standardized. Demand points were filtered based on this criterion and saved for further analysis. The demand points for each day were aggregated, and a visual representation of these points is shown in Fig. 2. In this figure, the blue dots represent specific demand point locations, with the horizontal and vertical coordinates indicating latitude and longitude.

The specific cleaning steps are as follows:

Step 1: Delete the vehicle data with abnormal data sending and receiving times, and delete the data whose GPS signal is outside Yancheng City, and clean and delete some of the data whose sending time cannot correspond to the actual time.

Step 2: According to the vehicle ID, find its single day GPS signal data, if the data is less than 100, it can be assumed that the driving distance of the day is too short and there is no charging demand.

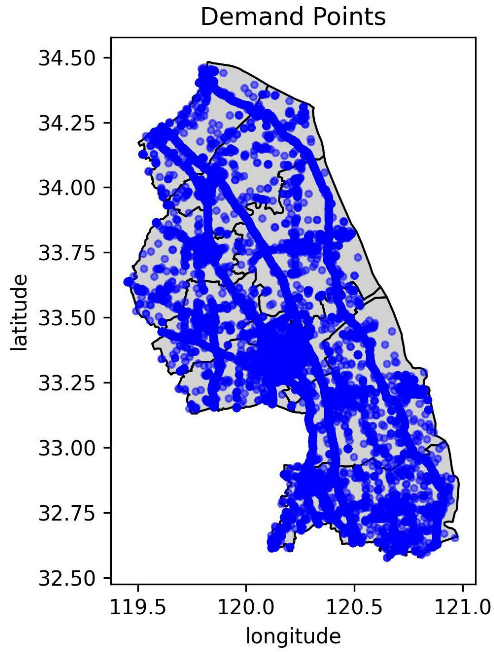

Fig. 2. Demand points.

Step 3: Determine that each vehicle has only 1 charging demand in a day, and keep the geographic location where the start time of the two longest GPS signal intervals is located as a demand point for backup.

4. Improvements to RODDPSO

4.1 RODDPSO

The Randomly Occurring Distributedly Delayed Particle Swarm Optimization (RODDPSO [18]) is an advanced evolutionary algorithm based on Particle Swarm Optimization (PSO) that is specifically designed to optimize solutions for complex problems. RODDPSO integrates a dynamic mechanism with a deterministic strategy to enhance the convergence speed and improve the global search capabilities of the traditional PSO algorithm.

When applied to optimize the initial cluster center locations in the K-Means++ algorithm, the RODDPSO algorithm utilizes velocity and positional adjustments of particles to identify more optimal initial cluster centers within the search space. This ensures better clustering performance by addressing limitations of traditional initialization methods, such as the tendency to converge on sub optimal solutions.

In this study, the RODDPSO algorithm introduces randomly occurring distributed time lags into the velocity update model. Specifically, a certain number of historical individual best particles and global best particles are randomly selected based on the evolutionary state. The time lag term is determined by multiplying it with a random value (0 or 1).This newly introduced randomly occurring distributed time lag in the velocity update model enables more effective utilization of the cumulative evolutionary history of the population. It enhances the algorithm's ability to avoid becoming trapped in local optima, while maintaining an appropriate balance between convergence speed and solution diversity. The velocity and position update equations for the RODDPSO algorithm are presented in (1)–(2).

$$\begin{aligned} v_i(k+1) &= wv_i(k) + c_1 r_1(p_i(k) - x_i(k)) \\ &\quad + c_2 r_2(p_{\mathrm{g}}(k) - x_i(k)) \\ &\quad + m_{\mathrm{l}}(\xi) c_3 r_3 \sum_{\tau=1}^{N} \alpha(\tau)(p_i(k-\tau) - x_i(k)) \\ &\quad + m_{\mathrm{g}}(\xi) c_4 r_4 \sum_{\tau=1}^{N} \alpha(\tau)(p_{\mathrm{g}}(k-\tau) - x_i(k)), \quad (1) \end{aligned}$$

$$x_i(k+1) = x_i(k) + v_i(k+1). \quad (2)$$

In the equations, $v_i(k+1)$ represents the updated velocity, and $x_i(k+1)$ represents the updated position. The k denotes the current iteration number, while c_1 and c_2 are the acceleration coefficients. Similarly, c_3 and c_4, which represent the acceleration coefficients of the distributed time lag term, are equal to c_1 and c_2, respectively (i.e., $c_1 = c_3$ and $c_2 = c_4$). N denotes the upper limit of the distributed time lag, and $\alpha(\tau)$ is an N-dimensional vector where each element is randomly assigned a value of 0 or 1. r_i $(i = 1, 2, 3, 4)$ represents random numbers uniformly distributed in the interval $[0, 1]$. The parameters m_{l} and m_{g} denote the strength factors of the distributed time lag term, which vary according to the evolutionary state ξ.

In this paper, the goal of the objective function is to minimize the average distance between the data points to their own centroids and the objective function is shown in (3).

$$J = \frac{\sum_{j=1}^{N_{\mathrm{c}}} \left[\sum_{\forall P_t \in C_{ij}} \frac{\mathrm{dist}(P_t, M_{ij})}{N_{\mathrm{p}}} \right]}{N_{\mathrm{c}}} \quad (3)$$

where N_{c} represents the number of clusters; C_{ij} denotes the jth cluster of the ith particle; M_{ij} denotes the jth clustering center of the ith particle; P_t denotes the tth data point; $\mathrm{dist}(P_t, M_{ij})$ denotes the Euclidean distance between the data point P_t and its clustering center M_{ij}; and N_{p} denotes the number of data points belonging to the clustering C_{ij}.

In this paper, the calculation of the evolution factor is based on the distance between particles, d_i, denoted by the average distance between the ith particle and the other particles, as shown in (4).

$$d_i = \frac{1}{S-1} \sum_{j=1, j\neq i}^{S} \sqrt{\sum_{k=1}^{D} \left(x_{ik} - x_{jk}\right)^2} \quad (4)$$

where S denotes the size of the particle population and D denotes the size of the particles. The formula for the evolutionary factor E_{f} is shown in (5).

$$E_{\mathrm{f}} = \frac{d_{\mathrm{g}} - d_{\min}}{d_{\max} - d_{\min}} \quad (5)$$

where d_{g} denotes the global best particle in d_i; $d_{\min}$ and $d_{\max}$ denote the minimum and maximum values of d_i in the particle population, respectively.

In this paper, an equipartition strategy is used to classify the four evolutionary states denoted by $\xi(k)$, as shown in (6).

$$\xi(k) = \begin{cases} 1, & 0.00 \le E_{\mathrm{f}} < 0.25, \\ 2, & 0.25 \le E_{\mathrm{f}} < 0.50, \\ 3, & 0.50 \le E_{\mathrm{f}} < 0.75, \\ 4, & 0.75 \le E_{\mathrm{f}} \le 1.00 \end{cases} \quad (6)$$

where the four states of the algorithm are the convergence state, exploitation state, exploration state, and escape state. In the convergence state, particles aim to quickly approach the global optimum by ignoring the distributed time lag term to enhance convergence speed. In the exploitation state, to avoid being trapped in a local optimum, the velocity update model incorporates distributed time lag terms, and a certain number of historical individual best particles are randomly selected to enable a more thorough local search. In the exploration state, particles are encouraged to extensively search the entire solution space by adding randomly occurring distributed time lags and selecting a certain number of historical global best particles to guide the exploration. In the escape state, particles aim to escape from regions around local optima by randomly selecting a combination of historical global best particles and individual best particles to assist in the search, effectively preventing the algorithm from being trapped in local optima.

Different evolutionary states produce different randomly occurring distributed delay information with updated values as shown in (7)–(8).

$$m_{\mathrm{l}}(\xi) = \begin{cases} 0.00, & \xi(k) = 1, \\ 0.01, & \xi(k) = 2, \\ 0.00, & \xi(k) = 3, \\ 0.01, & \xi(k) = 4. \end{cases} \quad (7)$$

$$m_{\mathrm{g}}(\xi) = \begin{cases} 0.00, & \xi(k) = 1, \\ 0.00, & \xi(k) = 2, \\ 0.01, & \xi(k) = 3, \\ 0.01, & \xi(k) = 4. \end{cases} \quad (8)$$

4.2 VRODDPSO

The RODDPSO algorithm employs randomly distributed delays, which reduces the likelihood of falling into local optima. However, the algorithm often struggles with early-stage local convergence due to the gradual narrowing of the search range for individuals in the population. This limitation arises from the algorithm's inherent mechanism, which exhibits weak exploratory capability in the early stages, leading to particle aggregation in localized regions and difficulty in effectively escaping local optima. This issue is particularly pronounced in optimization problems with complex search spaces and numerous local optima. While the algorithm may eventually rely on its global optimization capability to escape local optima as the number of iterations increases, this process often requires a large number of iterations, reducing the efficiency of exploration in the early stages and negatively impacting overall convergence speed and performance. Therefore, enhancing the algorithm's early exploration ability and avoiding premature convergence to local optima becomes crucial for improvement.

To address this issue, this paper introduces an adaptive mutation mechanism to the RODDPSO algorithm, forming the VRODDPSO algorithm. The proposed algorithm applies random mutations to particle positions during each iteration, thereby reducing the likelihood of being trapped in local optima. Fixing the mutation probability often results in abrupt changes during the early stages, causing the algorithm to prematurely focus on specific regions within the search space, which can lead to early convergence. This early convergence may prevent the algorithm from identifying the global optimum or superior solutions. Additionally, fixed mutation probabilities may reduce diversity among individuals in the population, limiting the algorithm's exploratory potential and hindering the discovery of better solutions. Increasing the number of adaptive mutations can better address the problem's dynamics, but the mutation probability must be carefully calibrated to strike a balance between exploration and exploitation. A higher mutation probability in the early stages of iteration promotes broader solution space exploration, while a gradual reduction in mutation probability during later iterations helps concentrate the search in regions likely to contain optimal solutions. Considering these factors, the formula for the mutation probability is provided in (9).

$$P(M) = 1 - \left(\frac{a^{1.2} + \text{iter}}{\text{max iter}} \times \frac{\log(\text{max iter})}{2} \right) \tag{9}$$

where $0 < P(M) < 1$ is the mutation probability, a is the number of mutations, iter is the number of current iterations, and max iter is the maximum number of iterations.

In the experiment, the fitness value of RODDPSO is initially obtained as a reference. As the number of iterations and mutations increases, the mutation probability should decrease to stabilize the optimal solution and prevent divergence from the global optimum. The formula for mutation probability is derived from experimental experience. The number of iterations, denoted as iter, changes linearly, and dividing it by the maximum number of iterations serves as the base mutation probability. However, if the mutation count also follows a linear function, the algorithm may still experience excessive mutations in the later stages. Therefore, the mutation count is reduced using a power function. During the experiments, the exponent of the mutation count was set to decimal values such as $1.0, 1.1, 1.2, \ldots, 2.0$. After testing the algorithm on 1,000 experimental datasets, it was found that setting the exponent of the mutation count to 1.2 achieved lower fitness values.

To ensure fewer mutations occur in the later stages, the influence of the maximum number of iterations needs to be minimized. Typically, the maximum number of iterations ranges from 50 to 1,000. When the number of iterations is small, more frequent mutations are necessary to avoid getting trapped in local optima. Conversely, when the number of iterations is high, excessive mutations could interfere with selecting the global optimum, so the mutation count needs to be appropriately reduced. For a moderate number of iterations, mutation probability should not be overly influenced. A logarithmic function fits these requirements perfectly, as multiplying the original mutation probability by the base-10 logarithm of the maximum number of iterations and dividing by 2 satisfies the condition: increasing mutation frequency for fewer than 100 iterations, decreasing it for more than 100 iterations, and leaving the mutation rate unaffected at exactly 100 iterations.

The adaptive mutation strategy proposed in this paper plays a vital role in reducing the risk of RODDPSO falling into local optima. The mutation probability decreases linearly with the number of iterations, ensuring a broad exploration space in the early stages of the search and more refined local search in the later stages. Specifically, during the adaptive mutation process, independent random perturbations are applied to randomly selected dimensions (ranging from 1 to n) of historical best particle position vectors. The mutated particle positions undergo boundary checks to ensure they remain within the feasible search space. By adjusting the mutation probability to balance the mutation count, the algorithm effectively adapts to the number of iterations. When the number of iterations is limited to around 100, the algorithm is more prone to getting trapped in local optima, requiring more adaptive mutations to escape. When the number of iterations reaches 1,000, the algorithm has sufficient iterations to find the global optimum, necessitating fewer mutations. As particle swarm optimization algorithms are inherently prone to local optima, a certain number of adaptive mutations are still required to escape these traps effectively.

The specific clustering process of the VRODDPSO algorithm is shown in Algorithm 1.

Algorithm 1 VRODDPSO clustering process.

```
Input: P, w, N_c, maxiter, velocity_max, other parameters
Output: Optimized cluster centroids and assignments
Initialize particles: X_i = {C_i1, C_i2, ..., C_iN_c}, ∀i ∈ [1, P]
Initialize velocity V_i and fitness f(X_i)
while Iteration t < maxiter do
    for Each particle i do
        Compute distance: d_jk = ||x_j − C_ik||
        Assign clusters: x_j → C_ik where k = arg min d_jk
        Evaluate fitness: f(X_i) = Σ_{j=1}^{N} ||x_j − C_ik||^2
    end for
    Update: X_i, V_i based on p_best and g_best
    Apply mutation to X_i
end while
return Optimized centroids C_k and cluster assignments
```

5. Site Selection and Capacity Determination Process

5.1 Process of Clustering

The first clustering phase plays a crucial role in grouping the areas to ensure that the locations of charging piles are within a reasonable geographical range of the city. Yancheng City, characterized by its irregular terrain with noticeable bumps and depressions, requires this initial clustering step. Without it, directly performing the second clustering might result in some of the charging pile locations being placed in depressed areas, thereby exceeding the city's boundaries and negatively impacting the overall layout. Moreover, by carrying out the first clustering, we can also effectively reduce running time and improve operational efficiency.

Following the completion of the first clustering, the second clustering is performed regionally. The primary objective of the second clustering is to identify the locations of the charging piles within each region, and the number of charging piles in each region is determined based on the ratio of electric vehicles to public charging piles. At this stage, the centroids of each cluster represents the required location of a charging pile.

Given the substantial demand for charging piles, which currently far exceeds the available public charging piles, the second clustering is conducted based on the prevailing ratio of electric vehicles to public charging piles in Yancheng City. According to public data, this ratio is 6.6:1. Consequently, demand points in different regions are clustered using a ratio of 6.6:1. This allows for the determination of the number of clusters, which corresponds to the number of charging piles required in the selected locations. The number of charging piles, in turn, becomes the new K-value for the clustering process. To ensure the program runs smoothly, the population size is adjusted based on the new K-value. All the centroids from the clustering process are stored in a table, with labels indicating the regions they represent for the third clustering phase. The results of the clustering are presented in Fig. 3, where the blue points represent the charging pile locations after clustering. Figures 2 and 3 are very similar, mainly because the number of demand points is large, and the number of charging piles clustered out is similarly large, and the local details are shown in Fig. 4. Figures 2 and 3 are still different in local details.

Charging station capacity determination involves identifying the key parameters, such as the number and type of charging piles and their corresponding charging power, required to meet the demand of electric vehicle users within a specific area. When determining the capacity of a charging station, several factors must be considered, including the number of electric vehicles in the region, their usage patterns, the characteristics of charging demand, as well as the construction costs and operational efficiency of the charging infrastructure.

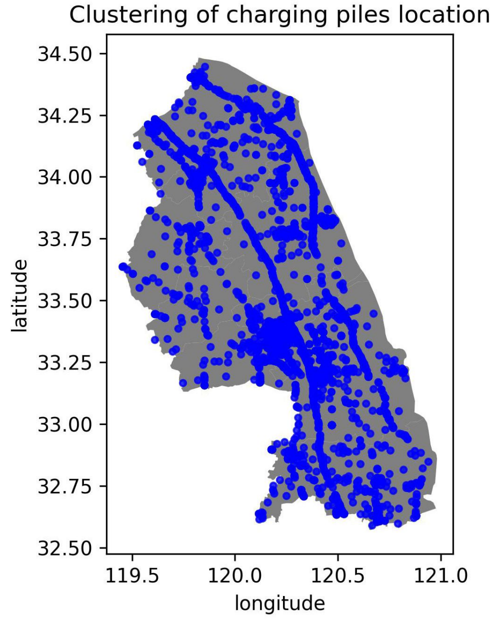

Fig. 3. Clustering of charging piles location.

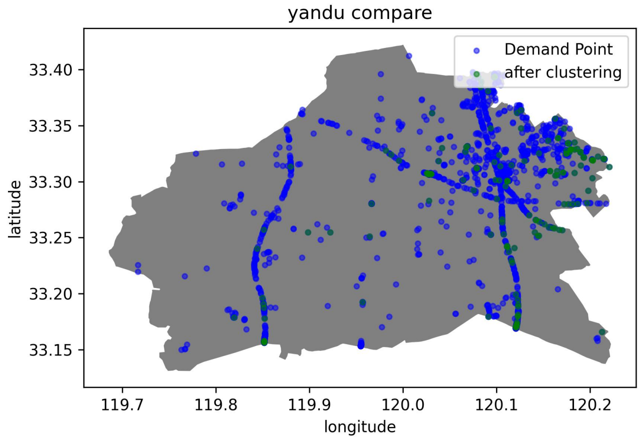

Fig. 4. Comparison between Fig. 2 and Fig. 3 in Yandu.

The objective of capacity determination is to calculate the number of charging piles needed, taking into account factors such as administrative divisions, user demand, and the power grid load. Specific considerations for capacity determination include the following:

1. The area share of the central urban area within the charging area within the clustered area.

2. The percentage of charging times in each administrative area in a month, as well as the percentage relationship between the administrative area and the clustered area of charging.

3. The percentage of total power in each area, and the relationship between the percentage of administrative areas and the clustered areas of charging.

Districts	Charging times (Proportion)
Tinghu	156215 (31.17%)
Yandu	72654 (14.50%)
Dafeng	65711 (13.11%)
Dongtai	48284 (9.63%)
Jianhu	19991 (3.99%)
Sheyang	52149 (10.41%)
Funing	20829 (4.16%)
Binhai	27419 (5.47%)
Xiangshui	37934 (7.56%)

Tab. 1. Charging times and their proportions.

Districts	Charging power (Proportion)
Tinghu	3813527 (30.00%)
Yandu	1835633 (14.44%)
Dafeng	1971446 (15.50%)
Dongtai	1155967 (9.09%)
Jianhu	470333 (3.70%)
Sheyang	1274350 (10.03%)
Funing	486770 (3.83%)
Binhai	678604 (5.34%)
Xiangshui	1024522 (8.07%)

Tab. 2. Charging power and its proportions.

Yancheng City consists of 9 administrative districts, with Tinghu District and Yandu District representing the central urban area, comprising approximately 10.37% of the total area. The percentage of charging times in each administrative district for a month is presented in Tab. 1, while the percentage of power usage is shown in Tab. 2.

The three specific factors represent spatial, temporal, and electrical power aspects, respectively. Therefore, the weights of the three factors should be equal. Since the sum of the three factors is still less than 1, the proportional values of the three factors are added together and then increased by 1. This result is then multiplied by the number of charging stations derived from the pile-to-station ratio, which gives the determined capacity of the charging stations.The fraction of the finalized volume was used as part of the reference index for the third clustering.

If the weight of spatial factor accounts for a larger proportion, the number of clustered charging stations will be more in the administrative area with larger area and less in the administrative area with smaller area, in this case, there will be a small administrative area with fewer charging stations, but the number of stations will not be enough to satisfy the users' demand due to the larger number of charging times or charging power. If the weight of the time factor is more significant, the number of charging stations in the area will increase due to the increase in the number of charging times. Whereas it may happen that the charging power is higher but the charging stations are occupied more frequently and are not enough to give the user's power needs, at the same time, if the weight of the power factor has a larger share, the number of charging stations in the area will increase because of the elevated power usage, but if a certain administrative area is very large and the charging station coverage is very low, it will lead to a more difficult time for the user to find a charging station.

If the number of charging stations that can be established is fixed at a specific value, this algorithm can still complete the clustering, and it only needs to modify the number of clustering centers for the third clustering to this specific value to complete the clustering, and the siting result will not be a problem.

According to the pile-station ratio of 8.8:1 in Yancheng city and the capacity fixing method mentioned above, the number of clusters of the third clustering can be obtained,

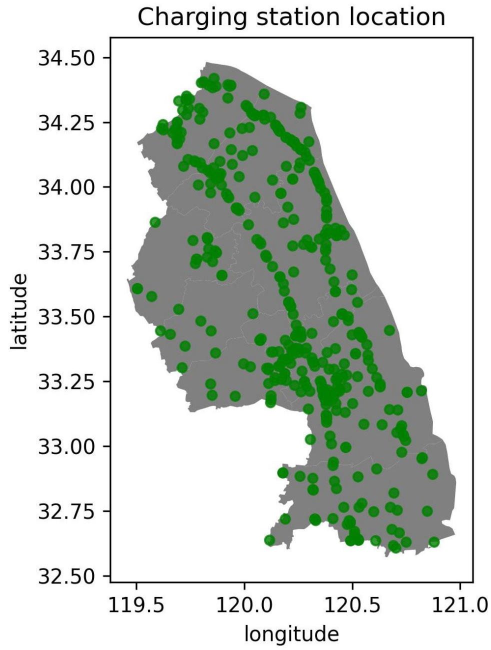

Fig. 5. Clustering of charging station location.

i.e. the number of charging stations corresponding to the number of demanded charging piles. The result is shown in Fig. 5. Where the green color is the clustered charging stations.

5.2 Establishment of New Evaluation Indicators

The comprehensive utilization rate is derived by considering the time utilization rate and power utilization rate of the charging station together, assuming that the time utilization rate is Tr and the power utilization rate is Pr, the formula for the comprehensive utilization rate Cr is shown in (10).

$$Cr = Tr \times Pr. \tag{10}$$

The effect of charging stations clustered by VROD-DPSO and K-Means++ algorithms should need to be compared with other algorithms, while all algorithms have their advantages in terms of charging station locations based on theoretical demand points. This paper hereby proposes a new comparison method that can effectively compare the results of using different algorithms based on demand points.

In practice, charging stations have 2 indicators for evaluating their profitability and efficiency, which are time utilization rate and power utilization rate, and the comprehensive utilization rate of each charging station can be obtained after calculation by (10). The charging stations clustered in this study can be tested based on the comprehensive utilization rate of existing charging stations. By removing the new stations with a combined utilization rate of 0, the remaining stations can be used for testing. All the stations that have actually been built are divided into four categories with a combined utilization rate of 0–5%, 5%–10%, 10%–15%, and 15% to 100% charging stations. Since the number of stations with a combined utilization rate of 15% or more is small, these stations are consolidated and grouped into one category.

In this paper, we set the efficiency of charging stations with a combined utilization rate of less than 5% to be low, 5%–10% to be moderately efficient, 10%–15% to be well utilized, and 15% or more to be extremely efficient; therefore, the number of demand points clustered in the 0–5% interval of the combined utilization rate should be as small as possible, and the 5%–10%, 10%–15%, and 15% or more intervals of the cluster should be as small as possible. number of demand points should be as high as possible. Therefore, the actual charging stations in various different utilization intervals are used as the center of the circle, and 100, 300, 500, and 700 meters are used as the radius to draw the circle. If the number of charging stations clustered by demand points within the circle center of efficient charging stations is more, the algorithm proves to be more effective. Similarly, the fewer the number of charging stations clustered by demand points within the center of the circle of inefficient charging stations, the better the algorithm proves to be.

For example, if Algorithm A clusters 10% of the charging stations within 500 meters of the charging stations with a combined utilization of 15% or more, and 5% of the charging stations within 500 meters of the charging stations with a combined utilization of 0–5%, and Algorithm B clusters 1% of the charging stations within 500 meters of the charging stations with a combined utilization of 15% or more, and 40% of the charging stations within 500 meters of the charging stations with a combined utilization of 0–5%, then it can be argued that Algorithm A is more effective. range, then it can be assumed that Algorithm A clusters better charging station sites than Algorithm B.

6. Experiment and Performance Analysis

6.1 VRODDPSO Performance Improvement

In this paper, the trend of the fitness curve is employed to evaluate the performance of the algorithm. A rapid decline in the fitness curve during the early stages suggests that the particles are quickly approaching the optimal solution, resulting in faster convergence. When the fitness curve gradually flattens, it implies that the algorithm is nearing convergence, although it may only be around a local optimum. If the fitness value continues to decrease steadily, it indicates that the algorithm is consistently optimizing the solution, moving closer to the problem's optimal solution. Conversely, if the fitness value stabilizes, the algorithm has likely reached the optimal solution, and further improvement is improbable. Significant fluctuations in the fitness value suggest that the algorithm is affected by noise or the problem's complexity. When the fitness curve stabilizes to a constant value, it signifies that the algorithm has successfully identified a global optimal solution. If there is a significant decrease in the curve at some point, it suggests that the algorithm is making progress toward finding a global optimum. A comparison of the fitness curves for RODDPSO, VRODDPSO, and other particle swarm optimization algorithms is illustrated in Fig. 6. In this figure, the horizontal axis represents the number of iterations, while the vertical axis reflects the logarithmic change in fitness value.

The test function used in this study is the Schaffer F7 function, a standard benchmark function widely applied for evaluating the performance and effectiveness of optimization algorithms. This function can be either single-dimensional or multi-dimensional, making it a versatile tool for testing optimization algorithms. In Fig. 6, the blue line represents the particle swarm optimization algorithm with inertia weights (WCPSO) [19], the orange line represents the particle swarm optimization algorithm without inertia weights (CPSO) [20], the yellow line represents the multi scale co-variant adaptive particle swarm optimization algorithm (MAEPSO) [21], the green line represents the RODDPSO algorithm, and the red line represents the VRODDPSO algorithm. As shown in Fig. 6, VRODDPSO demonstrates a more aggressive search capability for the global optimal solution compared to RODDPSO and outperforms both WCPSO and CPSO.

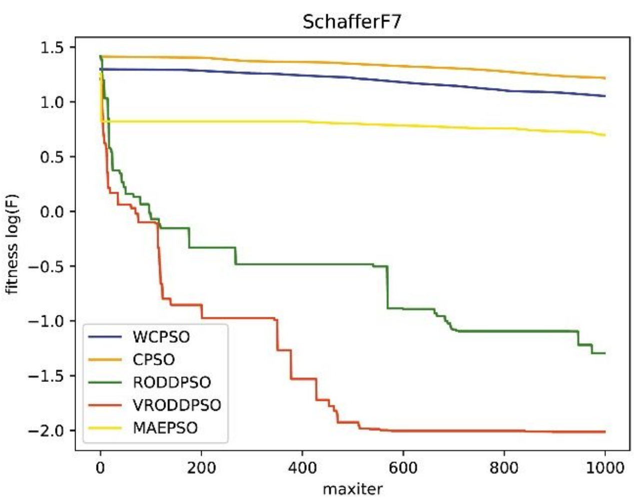

Fig. 6. Comparison of fitness curves of SchafferF7.

In terms of final results, VRODDPSO and RODDPSO achieve lower fitness values than the other algorithms. The logarithmic fitness values for WCPSO, CPSO, MAEPSO, RODDPSO, and VRODDPSO are 2.184, 2.606, −0.549, −4.017, and −7.786, respectively. This highlights the superior performance of the VRODDPSO algorithm in finding the global optimal solution.

Additionally, when applying RODDPSO for clustering, the algorithm tends to fall into local optima during the early iterations. As illustrated in Fig. 6, RODDPSO becomes trapped in a local optimum between iterations 100 and 300, with a limited search range. VRODDPSO addresses this issue by efficiently searching for a better global optimal solution. This improvement demonstrates that VRODDPSO enhances the effectiveness of finding the global optimum compared to RODDPSO, making it more suitable for solving complex optimization problems.

VRODDPSO also demonstrates a clear advantage over RODDPSO in terms of convergence speed. As shown in Fig. 6, while the RODDPSO algorithm takes more iterations to find its corresponding optimal fitness value, the VRODDPSO algorithm achieves the same fitness value in fewer iterations and continues searching for an even better solution. In subsequent iterations, VRODDPSO successfully identifies its corresponding globally optimal solution, whereas the RODDPSO algorithm becomes trapped in a local optimum. This highlights the superior efficiency and robustness of VRODDPSO in converging to the global optimal solution within a shorter computational time frame.

6.2 Rationality Verification of Site Selection and Capacity Determination

In Fig. 7, blue points represent the location points obtained from clustering, while orange points indicate the locations of actual charging stations.

Taking the 500-meter radius as an example, the validation of the clustering results for the two algorithms is shown in Figs. 8 and 9. In these figures, yellow points represent the real-world charging stations, which are categorized based on their comprehensive utilization rates into intervals of 0%–5%, 5%–10%, 10%–15%, and 15%–100%. The black circles indicate the 500-meter radius around the actual charging stations. The green points represent the charging stations clustered by the two algorithms: those clustered using the K-Means algorithm in one figure and those clustered using the fused K-Means++ and VRODDPSO algorithm in the other.

The statistical results of the number of actual public charging stations with comprehensive utilization rates across the four categories, within various ranges containing charging stations clustered by the two algorithms, are shown in Fig. 10.

The results show that the number of charging stations clustered using the K-Means algorithm is more than the number of charging stations clustered by the fusion of the K-Means++ and VRODDPSO algorithms near the actual stations with a combined utilization rate of 0–5%, while near the actual stations with a combined utilization rate of the interval of 5%–15%, the results of the clustering of the K-Means++ and VRODDPSO algorithms are significantly superior to those of the K- Means algorithm, and near the actual stations with more than 15% combined utilization rate, the results of K-Means++ with VRODDPSO algorithm clustering are significantly better than K-Means algorithm. As an example, the number of clustered charging stations within 500 meters of a realistic charging station with a combined utilization rate of 15% or more is only 141 using the K-Means algorithm, while the number of clustered charging stations using the K-Means++ with VRODDPSO algorithm is 231, which is a performance improvement of 63.8%. In summary, the clustering results using the fused K-Means++ and VRODDPSO algorithms are effectively demonstrated to be superior to those using the K-Means algorithm by this comparison method.

Literature [22] provides a new charging station layout planning evaluation strategy, a city, the largest flow of people, traffic flow in the area are often the city's central urban area, so the article will be site selection derived from the central urban area of the number of charging stations and the overall number of charging stations of the ratio as a criterion for evaluating the advantages and disadvantages of the site selection.

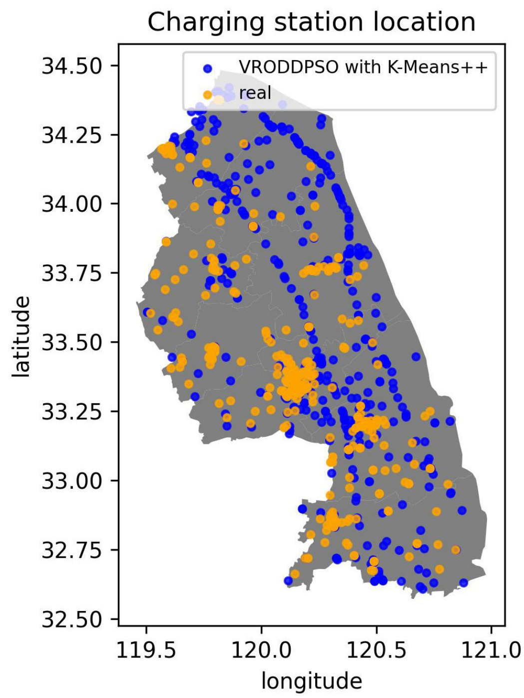

Fig. 7. Result of the third clustering with real stations.

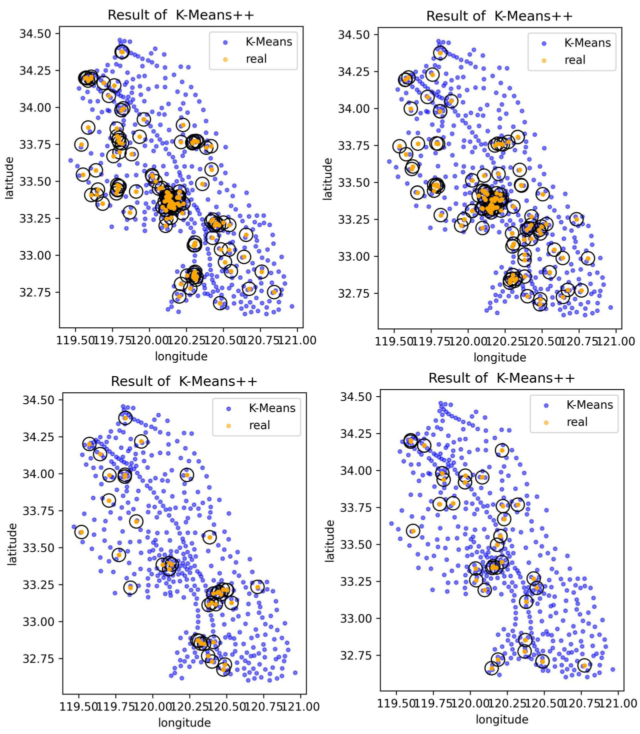

Fig. 8. Result of K-Means++.

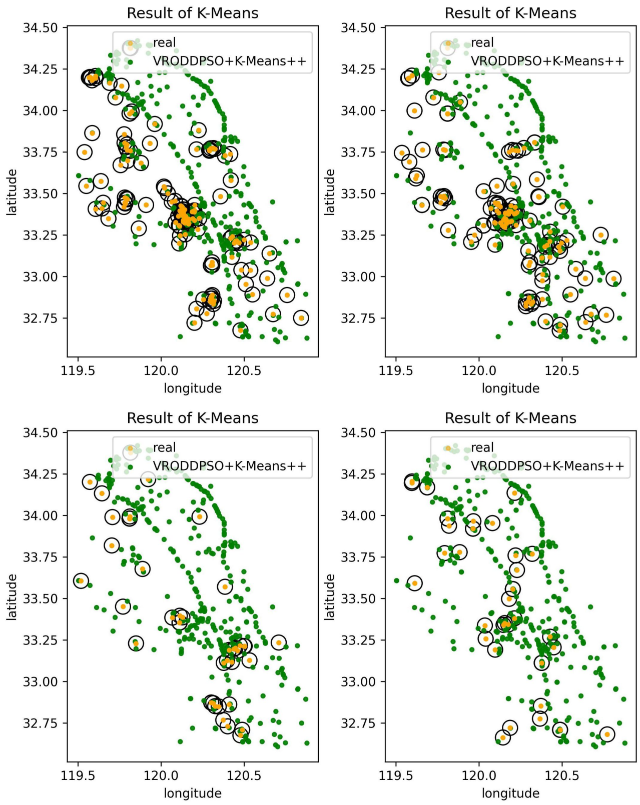

Fig. 9. Result of VRODDPSO and K-Means++.

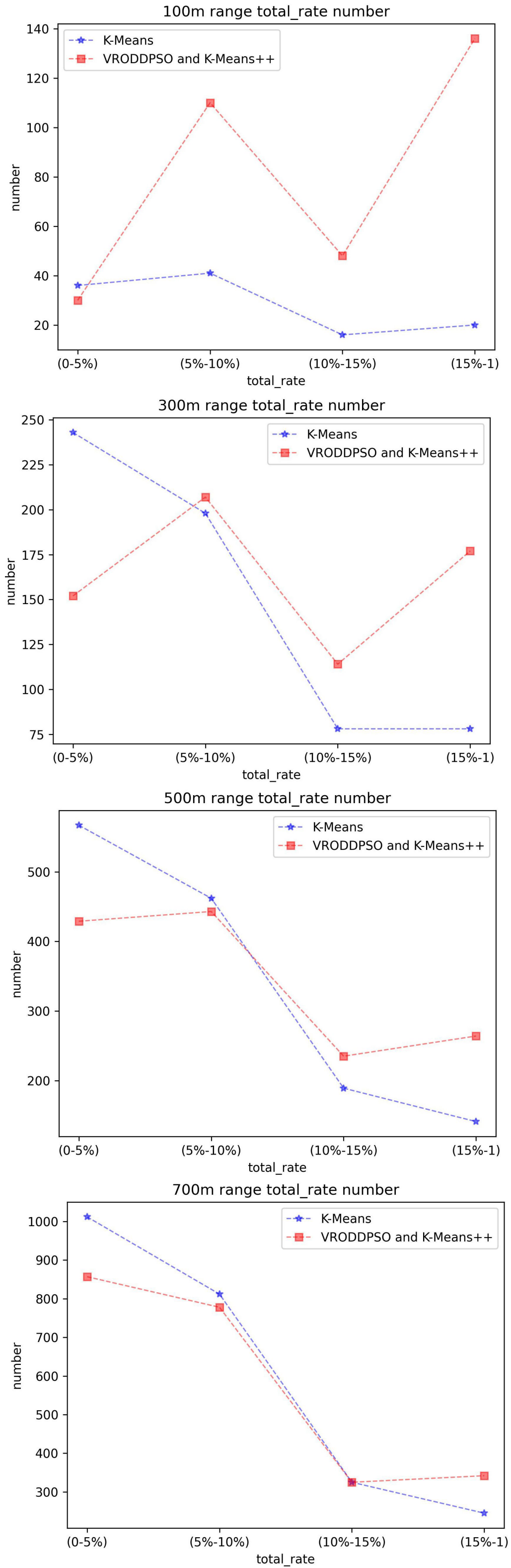

Fig. 10. Number of demand charging stations in different ranges for the four categories of comprehensive utilization rate charging stations.

Region	K-Means	Hybrid algorithms
Tinghu	44	49
Yandu	27	32

Tab. 3. The number of public charging stations selected in different areas under the two algorithms.

The central city of Yancheng has Tinghu District and Yandu District. The public charging stations in the two administrative areas resulting from the two algorithms are counted, and the statistical results are shown in Tab. 3. The locations of charging stations in the two administrative areas are shown in Figs. 11 and 12. The blue color is the location of public charging stations from the K-Means algorithm and the green color is the location of public charging stations from the fusion of K-Means++ and VRODDPSO algorithms.

Based on the data presented in the table above, the total number of public charging stations sited using the K-Means algorithm is 71, whereas the number of stations sited using the fusion of K-Means++ and VRODDPSO algorithms is 81. In terms of percentage, the number of stations sited in the urban centre area by the two algorithms is 14.69% for K-Means and 16.77% for the fusion of K-Means++ and VRODDPSO, respectively. This indicates that the public charging stations sited using the fusion of K-Means++ and VRODDPSO algorithms are more effective in terms of optimal location selection.

The siting of EV charging stations has garnered increasing attention, with rising expectations for the optimization of station layouts to improve the overall utilization rate. This paper investigates the fusion of the K-Means++ and VRODDPSO algorithms for selecting optimal locations for charging stations. The study focuses on the most frequently used public charging stations for net car groups, extracting demand points and clustering the theoretically suitable locations for the establishment of charging stations. The results of clustering using the fusion of K-Means++ and VRODDPSO are validated to outperform those obtained using the K-Means algorithm alone, by comparing them with real-life EV charging stations that demonstrate high comprehensive utilization rates. The research presented in this paper not only contributes to identifying potential charging station sites but also highlights the significant revenue that can be generated through the establishment of these stations.

The site selection methodology proposed in this paper provides a foundational framework for the layout of electric vehicle charging stations, as well as for subsequent evaluation and recommendation processes. As electric vehicles represent a key direction for future development in new energy technologies, the strategic siting of charging stations will offer essential support for the ongoing expansion and network planning of electric vehicles. In the future, we will try to use other more data, in-depth study to complete the charging station site layout. The code for the above experiment is available on [23] for clustering results with different data and parameter settings.

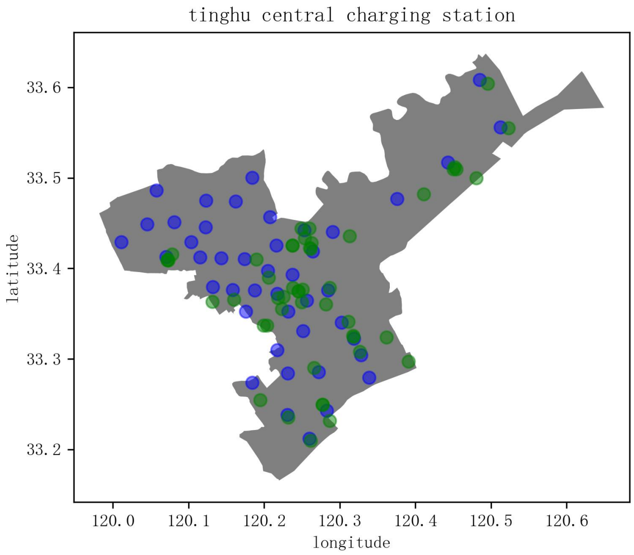

Fig. 11. Location map of public charging stations selected by two algorithms in Tinghu.

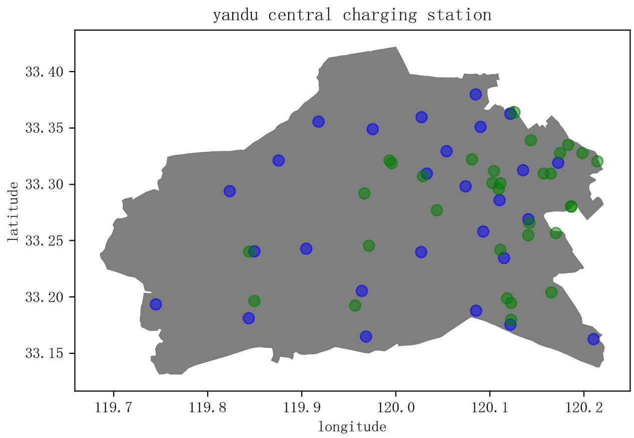

Fig. 12. Location map of public charging stations selected by two algorithms in Yandu.

7. Conclusions

The siting of EV charging stations has garnered increasing attention, with rising expectations for the optimization of station layouts to improve the overall utilization rate. This paper investigates the fusion of the K-Means++ and VRODDPSO algorithms for selecting optimal locations for charging stations. The study focuses on the most frequently used public charging stations for net car groups, extracting demand points and clustering the theoretically suitable locations for the establishment of charging stations. The results of clustering using the fusion of K-Means++ and VRODDPSO are validated to outperform those obtained using the K-Means algorithm alone, by comparing them with real-life EV charging stations that demonstrate high comprehensive utilization rates. The research presented in this paper not only contributes to identifying potential charging station sites but also highlights the significant revenue that can be generated through the establishment of these stations.

The site selection methodology proposed in this paper provides a foundational framework for the layout of electric vehicle charging stations, as well as for subsequent evaluation and recommendation processes. As electric vehicles represent a key direction for future development in new energy technologies, the strategic siting of charging stations will offer essential support for the ongoing expansion and network planning of electric vehicles.

Acknowledgments

This article is supported by FAITH project under the Erasmus+ programme of the European Union, Chinese National Funding of Social Sciences (Grant No. 20ZDA092), National Natural Science Foundation of China (NSFC) (Grant No. 6210124), and Nanjing Institute of Technology fund for Research Startup Projects of Introduced talents (Grant No. YKJ202019).

References

[1] YUVARAJ, T., DEVABALAJI, K. R., KUMAR, J. A. A comprehensive review and analysis of the allocation of electric vehicle charging stations in distribution networks. *IEEE Access*, 2024, vol. 12, p. 5404–5461. DOI: 10.1109/ACCESS.2023.3349274

[2] LI, C., ZHANG, L., OU, Z. Robust model of electric vehicle charging station location considering renewable energy and storage equipment. *Energy*, 2022, vol. 238, p. 1–14. DOI: 10.1016/j.energy.2021.121713

[3] MASTOI, M. S., ZHUANG, S., MUNIR, H. M., et al. An in-depth analysis of electric vehicle charging station infrastructure, policy implications, and future trends. *Energy Reports*, 2022, vol. 8, p. 11504–11529. DOI: 10.1016/j.egyr.2022.09.011

[4] HEMAVATHI, S., SHINISHA, A. A study on trends and developments in electric vehicle charging technologies. *Journal of Energy Storage*, 2022, vol. 52, p. 1–36. DOI: 10.1016/j.est.2022.105013

[5] ABID, M. S., AHSAN, R., AL ABRI, R., et al. Techno-economic and environmental assessment of renewable energy sources, virtual synchronous generators, and electric vehicle charging stations in microgrids. *Applied Energy*, 2024, vol. 353, p. 1–16. DOI: 10.1016/j.apenergy.2023.122028

[6] LUO, S., CHU, D., LI, Q., et al. Inverse kinematics solution of 6-DOF manipulator based on multi-objective full-parameter optimization PSO algorithm. *Frontiers in Neurorobotics*, 2022, vol. 16, p. 1–12. DOI: 10.3389/fnbot.2022.791796

[7] SINGH, N., CHAKRABARTI, T., CHAKRABARTI, P., et al. A new PSO technique used for the optimization of multiobjective economic emission dispatch. *Electronics*, 2023, vol. 12, no. 13, p. 1–14. DOI: 10.3390/electronics12132960

[8] SUN, J., CHE, Y., YANG, T., et al. Location and capacity determination method of electric vehicle charging station based on simulated annealing immune particle swarm optimization. *Energy Engineering*, 2022, vol. 120, no. 2, p. 367–384. DOI: 10.32604/ee.2023.023661

[9] RENE, E. A., FOKUI, W. S. T., KOUONCHIE, P. K. N. Optimal allocation of plug-in electric vehicle charging stations in the distribution network with distributed generation. *Green Energy and Intelligent Transportation*, 2023, vol. 2, no. 3, p. 1–15. DOI: 10.1016/j.geits.2023.100094

[10] ZANGANEH, M., MOGHADDAM, M. S., AZARFAR, A., et al. Multi-area distribution grids optimization using D-FACTS devices by M-PSO algorithm. *Energy Reports*, 2023, vol. 9, p. 133–147. DOI: 10.1016/j.egyr.2022.11.180

[11] MAGSINO, E., ESPIRITU, F. M. M., GO, K. D. Discovering electric vehicle charging locations based on clustering techniques applied to vehicular mobility datasets. *ISPRS International Journal of Geo-Information*, 2024, vol. 13, no. 10, p. 1–19. DOI: 10.3390/ijgi13100368

[12] LI, C., DONG, Z., CHEN, G., et al. Data-driven planning of electric vehicle charging infrastructure: A case study of Sydney, Australia. *IEEE Transactions on Smart Grid*, 2021, vol. 12, no. 4, p. 3289–3304. DOI: 10.1109/TSG.2021.3054763

[13] CELIK, S., OK, S. Electric vehicle charging stations: model, algorithm, simulation, location, and capacity planning. *Heliyon*, 2024, vol. 10, no. 7, p. 1–19. DOI: 10.1016/j.heliyon.2024.e29153

[14] ABDEL-BASSET, M., GAMAL, A., HEZAM, I. M., et al. Sustainability assessment of optimal location of electric vehicle charge stations: A conceptual framework for green energy into smart cities. *Environment, Development and Sustainability*, 2024, vol. 26, no. 5, p. 11475–11513. DOI: 10.1007/s10668-023-03373-z

[15] ZAPOTECAS-MARTINEZ, S., ARMAS, R., GARCIA-NAJERA, A. A multi-objective evolutionary approach for the electric vehicle charging stations problem. *Expert Systems with Applications*, 2024, vol. 240, p. 1–11. DOI: 10.1016/j.eswa.2023.122514

[16] POURVAZIRI, H., SARHADI, H., AZAD, N., et al. Planning of electric vehicle charging stations: an integrated deep learning and queueing theory approach. *Transportation Research Part E: Logistics and Transportation Review*, 2024, vol. 186, p. 1–23. DOI: 10.1016/j.tre.2024.103568

[17] MEN, J., ZHAO, C. A type-2 fuzzy hybrid preference optimization methodology for electric vehicle charging station location. *Energy*, 2024, vol. 293, p. 1–12. DOI: 10.1016/j.energy.2024.130701

[18] LIU, W., WANG, Z., LIU, L. X., et al. A novel particle swarm optimization approach for patient clustering from emergency departments. *IEEE Transactions on Evolutionary Computation*, 2019, vol. 23, no. 4, p. 1–13. DOI: 10.1109/TEVC.2018.2878536

[19] WANG, Y., CAO, Y., YEO, T. S., et al. Joint sequence optimization-based OFDM waveform design for integrated radar and communication systems. *IEEE Transactions on Vehicular Technology*, 2022, vol. 71, no. 12, p. 12734–12748. DOI: 10.1109/TVT.2022.3201210

[20] YU, Z., HUAI, R., LI, H. CPSO-based parameter-identification method for the fractional-order modeling of lithium-ion batteries. *IEEE Transactions on Power Electronics*, 2021, vol. 36, no. 10, p. 11109–11123. DOI: 10.1109/TPEL.2021.3073810

[21] ZHENG, W., ZHAO, C., ZHANG, G., et al. Multi-birth optimization based on ergodic multi-scale cooperative mutation self-adaptive escape PSO for transformer fault diagnosis and location. In *Proceedings of the 5th Asia Energy and Electrical Engineering Symposium (AEEES)*. Chengdu (China), 2023, p. 644–652. DOI: 10.1109/AEEES56888.2023.10114235

[22] JIA, J., LIU, C., WAN, T. Planning of the charging station for electric vehicles utilizing cellular signaling data. *Sustainability*, 2019, vol. 11, no. 3, p. 1–16. DOI: 10.3390/su11030643

[23] YUNYEZHUYUN. *VRODDPSO Code*. [Online]. Available at: https://github.com/Yunyezhuyun/VRODDPSO, 2025

About the Authors . . .

Birong HUANG was born in 1981. He received his M.E. degree from Jiangnan University in 2009, and an MBA degree from Hehai University in 2016. He is currently pursuing his Ph.D. degree at the College of Economics and Management, Nanjing University of Aeronautics and Astronautics. His main research direction is to explore the key technologies of the regional new power system under the background of energy transformation and to improve the modernization governance capacity of power grid enterprises.

Zilong WANG is a Professor at Nanjing University of Aeronautics and Astronautics, where he also received his Ph.D. He previously completed his postdoctoral research at Peking University. His research interests focus on industrial economics and management. He has received several research grants from important Chinese research institutions such as the National Natural Science Foundation of China and the Aeronautics Science Foundation. He has published more than 120 articles and scholarly monographs.

Jianhua CHEN was born in 1976. He received his Master's degree in Thermal Energy Engineering from WUHEE in 2000. He is currently pursuing a doctoral degree in Management Science and Engineering at Nanjing University of Aeronautics and Astronautics. In 2009, he became an Associate Professor at the School of Electric Power Engineering, Nanjing Institute of Technology. His research interests include new energy generation control technology, big data modeling and its application in smart grids, and energy technology economics.

Bingyang ZHOU received his Master's degree in Management from Nanjing University of Science and Technology in 2009. He is currently pursuing a Ph.D. in Management Science and Engineering at Nanjing University of Aeronautics and Astronautics. His research interests include the allocation of regional scientific and technological innovation resources and regional economic policies.

Yuhang ZHU was born in 2001. He received his Bachelor's degree in Information Engineering from Nanjing Institute of Technology in 2024. He is currently pursuing a Master's degree in Communication and Information Systems at Nanjing Institute of Technology. His research interests include meta-heuristics in smart grids and V2G technology.

Yan LIU (corresponding author) was born in 1980. She received her M.S. degree in Applied Mathematics from Jiangsu University in 2007. She is now an Associate Professor at Nanjing Institute of Technology. Her research interests include complex networks and mathematical modeling.