

## Article

# Electric Taxi Charging Load Prediction Based on Trajectory Data and Reinforcement Learning—A Case Study of Shenzhen Municipality

Xiaojia Liu <sup>1,2</sup>, Bowei Liu <sup>1,2</sup>, Yunjie Chen <sup>1,2</sup>, Yuqin Zhou <sup>1,2</sup> and Dexin Yu <sup>1,\*</sup>

<sup>1</sup> College of Navigation, Jimei University, Xiamen 361021, China; happylxj1314@163.com (X.L.); lbw297083@163.com (B.L.); cyunjie8080@163.com (Y.C.); zyq20220224@163.com (Y.Z.)

<sup>2</sup> Marine Traffic Safety Institute, Jimei University, Xiamen 361021, China

\* Correspondence: yudx@jmu.edu.cn

**Abstract:** In order to effectively solve the problem of electric taxi charging load prediction and reasonable charging behaviour discrimination, in this paper, we use taxi GPS trajectory data to mine the probability of operation behaviour in each area of the city, simulate the operation behaviour of a day by combining it with reinforcement learning ideas, obtain the optimal operation strategy through training, and count the spatial and temporal distributions and power values at the time of charging decision making, so as to predict the charging load of electric taxis. Experiments are carried out using taxi travel data in Shenzhen city centre. The results show that, in terms of taxi operation behaviour, the operation behaviour optimized by the DQN algorithm shows the optimal effect in terms of the passenger carrying time, mileage, and daily net income; in terms of the charging load distribution, the spatial charging demand of electric taxis in each area shows obvious differences, and the charging demand load located in the city centre area and close to the traffic hub is higher. In time, the peak charging demand is distributed around 3:00 to 4:00 and 14:00 to 15:00. Compared with the operating habits of drivers based on the Monte Carlo simulation, the DQN algorithm is able to optimise the efficiency and profitability of taxi drivers, which is more in line with the actual operating habits of drivers formed through accumulated experience, thus achieving a more accurate charging load distribution.



**Citation:** Liu, X.; Liu, B.; Chen, Y.; Zhou, Y.; Yu, D. Electric Taxi Charging Load Prediction Based on Trajectory Data and Reinforcement Learning—A Case Study of Shenzhen Municipality. *Sustainability* **2024**, *16*, 1520. <https://doi.org/10.3390/su16041520>

Academic Editor: Armando Carteni

Received: 25 December 2023

Revised: 27 January 2024

Accepted: 30 January 2024

Published: 10 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** electric taxi; trajectory data; spatio-temporal distribution; reinforcement learning; charging loads

## 1. Introduction

As urban travelling is gradually developing in the direction of environmental protection, efficiency, and sustainability, electric vehicles, as an important mode of transport, are being gradually integrated into people's daily life. However, with the increasing number of electric vehicles in cities [1], the contradiction between their charging demand and charging infrastructure, and, at the same time, large-scale electric vehicles being connected to the power grid, with their charging and discharging behaviours having a considerable impact on the power grid [2,3], has become one of the key challenges affecting the sustainable development of cities. Compared with other types of electric vehicles, electric taxis usually face a higher travel demand density, drivers' travel routes are more variable, the charging demand of electric taxis is difficult to accurately predict in advance, and their charging loads show greater uncertainty in time and space [4], which increases the difficulty of prediction. Charging load forecasting techniques have become particularly important in order to achieve the efficient operation of electric taxis, increase the utilisation of the charging infrastructure, as well as reduce energy consumption and environmental impacts.

There have been many studies on charging load predictions for electric vehicles. Currently, many scholars focus on the use of stochastic mathematics in order to deepen

the understanding of the operating characteristics of electric vehicles. These methods outline the distribution profile of charging loads by constructing models to simulate vehicle journeys over the course of a day with the help of probabilistic the statistics-based Monte Carlo (MC) stochastic simulation, using the EV's starting charging time, mileage, battery charge level (SOC), and charging tariff as the core factors. As for the probabilistic selection of many key factors for electric vehicles, the current methods can be divided into two main categories. One category is based on the deterministic probability density distribution function of the starting charging time, driving mileage, SOC, etc., (currently the main reference data are the American Household Travel Survey data NHTS 2009 [5]) for the calculation and prediction of the charging load. In the literature [6], for different types of EVs, the charging loads of EVs are calculated with the help of the Monte Carlo simulation based on the starting state of the charge and charging time. The literature [7] starts from the charging onset moment, charging duration, SOC, and other factors, and constructs the probability distribution model of the travelling pattern and charging characteristics by the Monte Carlo algorithm to predict the load demand when EVs are connected to the grid.

The other category estimates the charging power of EVs by analysing the travel patterns of EVs through traffic big data and predicting the probability distribution of key factors such as the charging start time, mileage, SOC, and charging tariff of EVs. In the literature [8], the simulation of fast-charging and the power exchange modes of electric taxis is achieved by integrating the grid-based traffic road network, charging/power exchange, driving path, path selection, and other behaviours of electric taxis into the Monte Carlo stochastic simulation method, and the operation scenario of taxis in Hangzhou is used as a case study, which verifies the good performance of the Monte Carlo algorithm in predicting the travel pattern of taxis. The literature [9] mined the data using four methods: a decision table, a decision tree, an artificial neural network, and a support vector machine, and the results showed that the support vector machine had the best prediction results, followed by the decision tree. The literature [10] calculated EV trip and charging characteristics to establish probabilistic models by collecting GPS and charging data from 15 EVs in Ireland and used the Monte Carlo method to predict the charging load of EVs. The literature [11] refined the EV GPS data by univariate and multivariate interpolation techniques and predicted the EV charging station loads based on long- and short-term memory neural network models.

At present, electric taxis are still in the development stage, electric taxi operation and distribution laws are still in the exploratory stage, and there are relatively few studies on charging behaviour analysis and load forecasting for electric taxis. Therefore, if electric taxis want to develop rapidly, it is necessary to have a deeper understanding of the operation mode of electric taxis and the problems in operation.

Over the past few years, with the rapid development of mobile smart technologies, taxi trajectory data have been collected and stored in large quantities. These data contain a wealth of urban travel information and have great potential for understanding taxi operational behaviour and predicting charging loads. However, due to the complexity of taxi operations and the uncertainty of spatio-temporal variations, it is difficult to accurately predict charging loads by solely relying on data-driven methods. To overcome this challenge, reinforcement learning, a method for intelligent decision making in uncertain environments, explores new ideas for electric taxi charging load prediction. Reinforcement learning fosters the intelligence to take more reasonable actions in a given environment in order to maximise long-term rewards through the subtle interaction between the intelligence and the environment. Incorporating reinforcement learning into charging load prediction for electric taxis enables intelligent decision making on charging strategies based on actual operating conditions, thus depicting the charging load profile more accurately.

Therefore, this paper will take the core idea of integrating taxi trajectory data and the reinforcement learning strategy. Simulation modelling is used to explore the reasonable operation behaviour and charging habit formed by taxi drivers through their accumulated experience in the process of operation, so as to achieve the prediction of the charging load.

Reinforcement learning-based methods are chosen to treat electric taxis as intelligences, and the distribution law of the historical trajectory data is used to establish the decision-making model of the intelligences, so as to predict the charging load of taxis by simulating the real operation scenarios.

By combining the methods of data mining, reinforcement learning, and simulation, the integration of reinforcement learning ideas into the simulation process can be a more reasonable and effective solution to the problem of drive-charging decision making in the simulation process of electric taxi operation behaviour compared with the existing research, so as to predict a more realistic and accurate charging load based on the simulation results.

In summary, the article is divided into three parts.

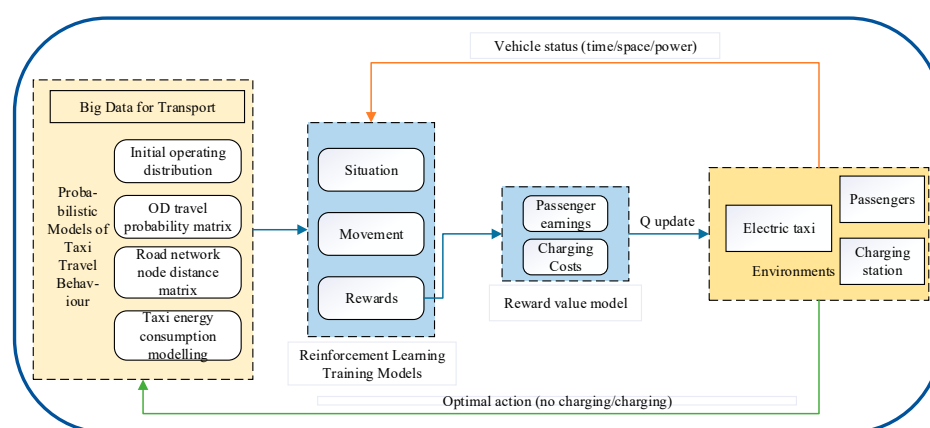
Part I: mining the operation behaviour information of taxis by combining Shenzhen taxi trajectory data and road network map data, including the spatio-temporal distribution of the initial operation of taxis, the travel probability OD matrix (OD stands for traffic volume survey where “O” stands for origin and “D” for destination), and the shortest paths of each road network node in the city.

Part II: Through the data information mined in Part I, the charging decision model for electric taxis is constructed by combining the Deep Q-learning algorithm, and the taxi travelling behaviour is simulated based on the decision model. Retaining the charging power and time for each taxi selection charging decision in the simulation results as a result of the load prediction.

Part III: The simulation results are analysed to compare the effectiveness of the benefits achieved by taxi driver operational decision making using three algorithms that also belong to reinforcement learning: the Q-learning algorithm, the SARSA algorithm, and the DQN algorithm. The MC method, used by many scholars as a charging criterion, is used to compare the charging load prediction results of electric taxis with the charging decision model constructed in this paper from the point of view of reasonableness.

## 2. General Framework

Based on the idea of reinforcement learning combined with traffic big data, it simulates the reasonable actual operating behaviour and charging habits formed by the electric taxi drivers' continuous accumulation of experience in daily operation, so as to predict the total charging load per unit of time of the taxi group. The overall framework is shown in Figure 1.



**Figure 1.** Electric taxi charging in line with the forecasting framework.

For the time being, there is still a big gap between the car ownership of new energy vehicles compared to fuel vehicles [12], and there is a lack of trajectory data for pure electric vehicles. Since electric vehicles are used to replace traditional fuel vehicles, this paper assumes that when a taxi driver chooses an electric vehicle as a travel tool, the travel behaviour pattern will not change. Taking taxis in the central city of Shenzhen as the data samples, the taxi trajectory data are used to restore the running routes of taxis in the city, to

explore the running laws of taxis in the city, and at the same time, are combined with the road information of the urban road network to simulate the normal operation of electric vehicles in the city.

Based on the trajectory data, the spatial and temporal distribution of taxi operation is mined, and the road information in the road network is derived from the map data, and then combined with the trajectory data and the map data to mine the probability of vehicle travelling in each area of the city and the travelling path. The mined trajectory data information is combined with the energy consumption model of the taxi to simulate the operation of the electric taxi using the simulation method centred on the idea of reinforcement learning.

What needs to be considered for reinforcement learning is the complex interaction between an intelligent body and its environment. Intelligent bodies construct the current state by perceiving the environment while operating in a particular environment. The intelligent body influences the environment by taking specific actions, the execution of which triggers a change of state through the potential state transfer probability of the environment. The environment responds to the actions of the intelligent body by feeding back appropriate rewards based on an intrinsic reward model [13]. The core goal of an intelligent body is to maximise cumulative rewards. In the electric taxi load forecasting problem, the intelligent body is the electric taxi and the environment is the spatio-temporal probability of each behaviour of the taxi during operation. The states cover the taxi SOC values, time, space, and other key information that affects the driver's decision making, while the actions represent the operational and charging decisions taken by the taxi driver when facing different states. The different actions taken by the taxi when facing different states combined with factors such as passenger revenue and charging cost provide real-time reward signals for the intelligence to guide the optimal evolution of charging strategies. Strategies, on the other hand, embody mapping from states to actions, and this relationship is usually presented in the form of a Q-value table, which is the potential future benefit of choosing a certain action in a given state. Actions with higher Q-values are considered superior, which can further steer the intelligence's decision-making process to achieve better long-term reward accumulation.

### 3. Spatio-Temporal Probability of Electric Taxi Travel Behaviour Based on Trajectory Data

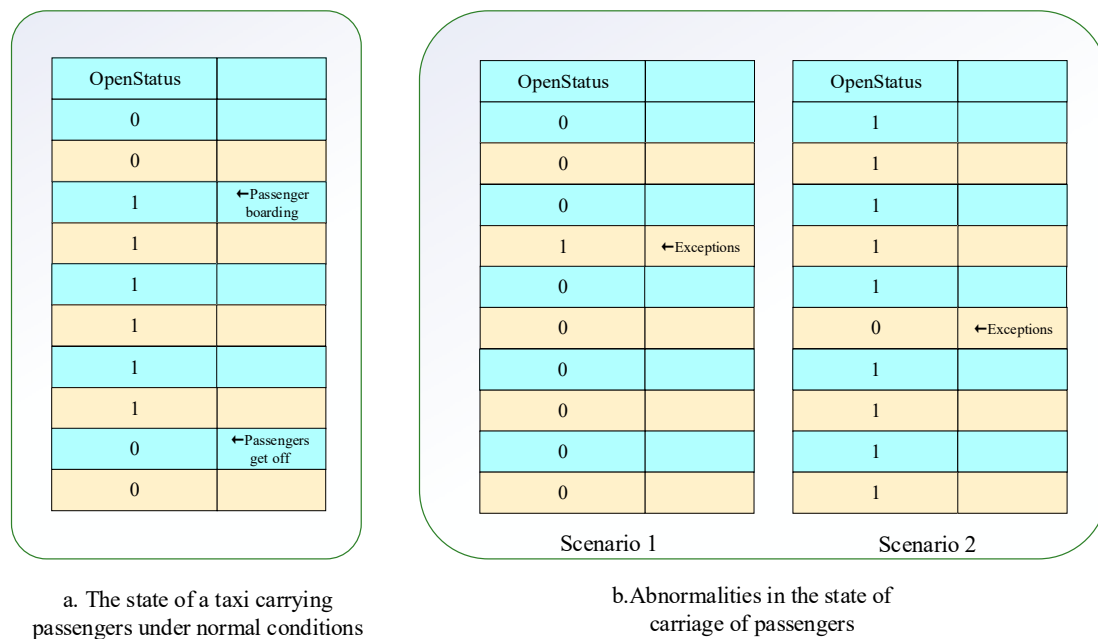
#### 3.1. Data Preprocessing

The study city of this paper is in Shenzhen, whose taxi GPS data [14] are derived from a one-day GPS sample of electric taxis in the Chinese city of Shenzhen, including: vehicle ID, longitude, latitude, time, speed, and occupancy status. There were a total of 664 electric taxis and 1,155,654 GPS records, with taxis sampled at a frequency of about 15 s.

##### 3.1.1. Cleaning of Anomalous Data

It is possible for the taxi collection equipment to suddenly become abnormal and this means that the taxi's GPS collects very little data of the longitude, latitude, speed, and passenger status attribute value in an unreasonable range. In terms of the latitude and longitude, this problem is manifested in the appearance of coordinates that are far away from each other in a continuous coordinate sequence, and in terms of speed, it is manifested in the appearance of speeds exceeding 200 km/h and continuous track speeds of more than 5 h being 0. For this kind of abnormal data, a certain threshold can be set to delete the data. The passenger state is manifested in the continuous sequence of the 0 or 1 state suddenly appearing in a 1 or 0 state, and such anomalous data are characterised by the passenger state, the value of the previous data, and the latter data being different, and these three pieces of data are the same vehicle's continuous data, as shown in Figure 2. For data with an abnormal passenger status, the data column can be shifted up or down as a whole, so that the information of the three pieces of data is in the same row, and then filtered out using conditional judgement statements.



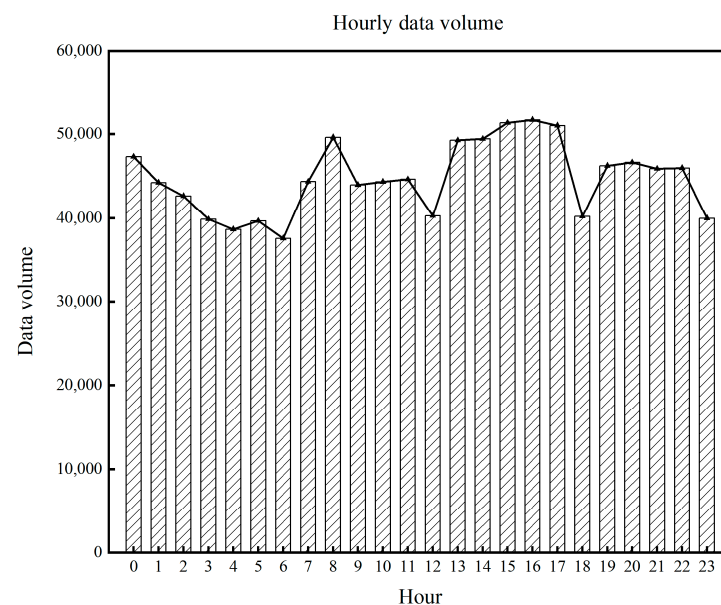


**Figure 2.** Schematic of the Passenger Status field.

After screening the data, a total of 1612 rows of anomalous data were removed, reducing the number of GPS data rows to 1,154,042.

### 3.1.2. Data Quality Assessment

The data with outliers removed are evaluated by converting the time field in the taxi GPS data to date format and extracting the hourly information in the date format, and the hourly data volume is counted through the aggregation process, as shown in Figure 3.

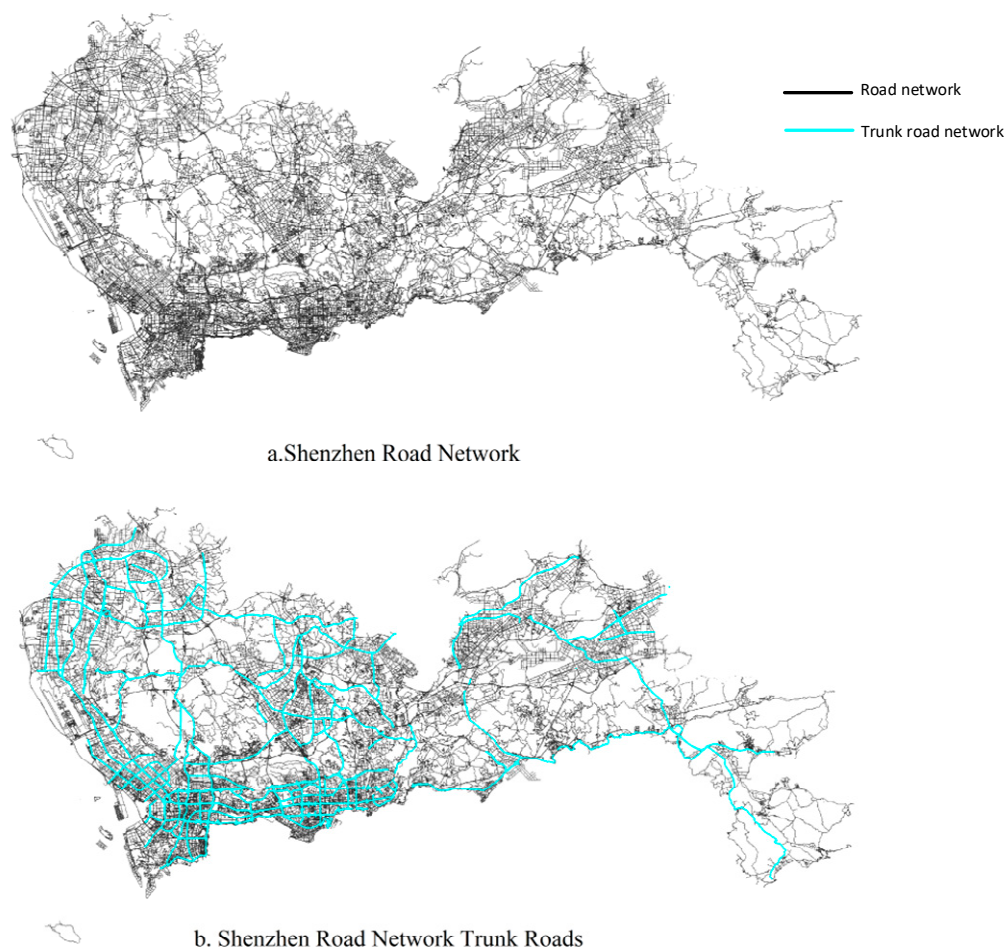


**Figure 3.** Hourly data volume line graph.

The time-varying line graph of the data volume shows that there is no obvious data missing in the hourly distribution of the taxi GPS data, which indicates that the data quality has a certain degree of completeness and can be further data-mined.

### 3.2. Urban Road Topology

The road data are obtained from Open street map, an open-source mapping platform, and the main roads of the Shenzhen road network are extracted using Arcgis, as shown in Figure 4.



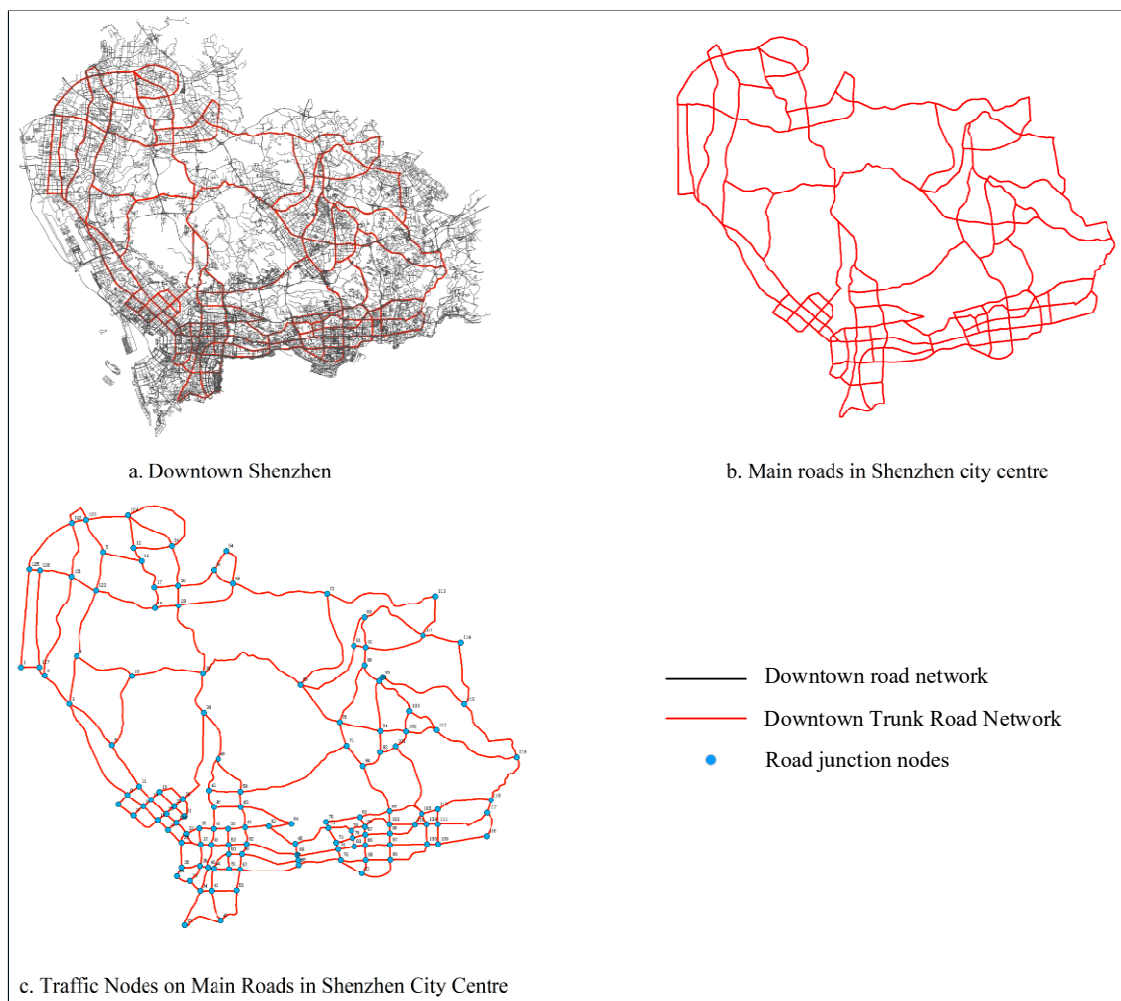
**Figure 4.** Shenzhen Road Network.

Considering the establishment of the topological network, the research scope is selected as the central city of Shenzhen including the Futian District, Luohu District, Nanshan District, Longhua District, Longgang District, Bao'an District, and Guangming District. The main road network within the study area is extracted, and a total of 128 nodes are constructed with the intersections of the roads as nodes, as shown in Figure 5, and the road network topology is established.

### 3.3. Spatial and Temporal Distribution of Electric Taxis Starting Operation

The accurate identification of the start of the operation of electric taxi is the beginning of the simulation process of electric taxis, and mining the spatial and temporal distribution of the start of the operation of electric taxis is a crucial step. A taxi shift handover generally adopts a double-shift system, which can ensure the fairness of the operation time, so that the two drivers are responsible for the operation of the morning peak and the evening peak, and at the same time, ensure that the drivers will not be overly fatigued. The spatial and temporal nodes at the end of the taxi shift handover behaviour are the spatial and temporal nodes at which the taxi starts to operate. Trajectory data are used to detect the stopping point of the taxi when it is not in operation. At the same time, the time attribute of the shift handover point is used to filter the stopping point of the taxi when it is not in operation, so

that the spatial and temporal distribution at the start of the operation of the electric taxi can be reasonably mined.



**Figure 5.** Main roads in downtown Shenzhen.

### 3.3.1. Dwell Point Detection

The trajectory data are used to detect the stopping points of electric taxis during non-operational periods, so as to filter out the time and location when the taxis start to operate. A combination of a Stop/Move model and a Velocity Sequence Linear Clustering algorithm (VSLC) is used to identify the spliced combination of operational states of single vehicle trajectory segments [15]. Taxi trajectory data are composed of spatio-temporal sequences of trajectory points, and the motion states can be classified into two types: operational trajectories and non-operational trajectories according to the speed. And the Stop/Move model [16] can effectively classify the taxi operation trajectory into two states, operation and non-operation. As shown in Figure 6a, for the speed trajectory of the taxi with the vehicle ID 22396 in the central city of Shenzhen from 00:00–1:30, the trajectory point with a speed equal to 0, i.e., the Stop state, is recorded as 0, and the trajectory point with a speed greater than 0, i.e., the Move state, is recorded as 1, and the mapping result is shown in Figure 6b.

However, in the Stop state, a Move state exists where the mapping result is the Stop state, but the actual operating time is caused by taxi drivers due to passengers getting in and out of the taxi, going to the toilet, traffic congestion, traffic lights, and other factors. For example, in Figure 6a, there are frequent transitions between the Stop and Move states. In order to solve this kind of problem, it is reasonable to distinguish the operation state and

non-operation state. The velocity sequence linear clustering algorithm (VSLC) [17] is used to correct the abnormal state results by setting the time thresholds of Stop and Move, the principle of which is shown in Figure 7.

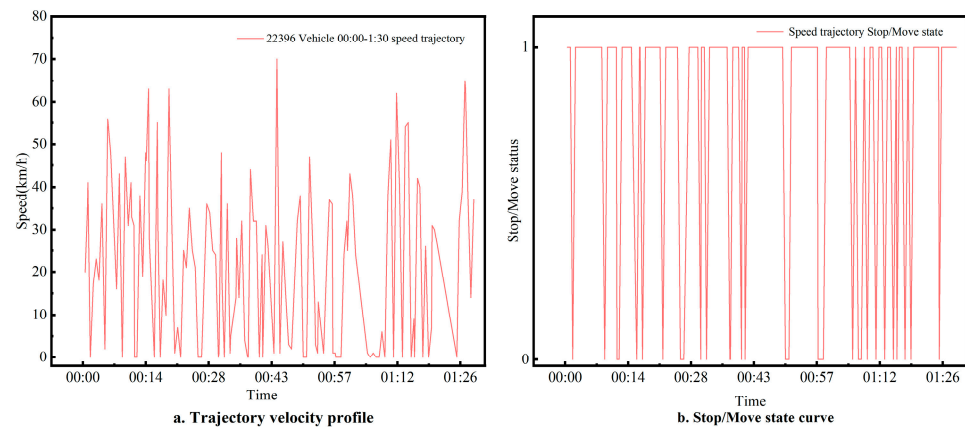


Figure 6. Stop/Move state mapping results.

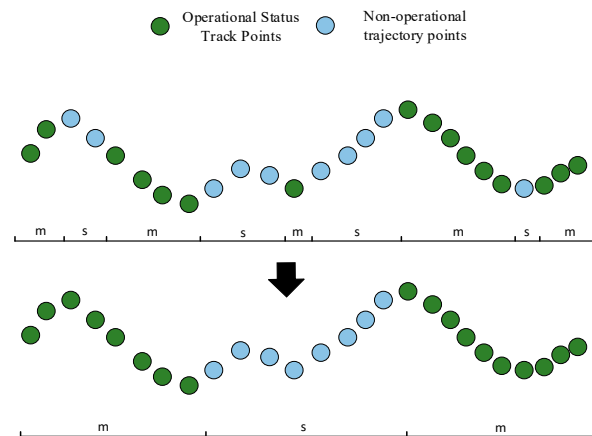
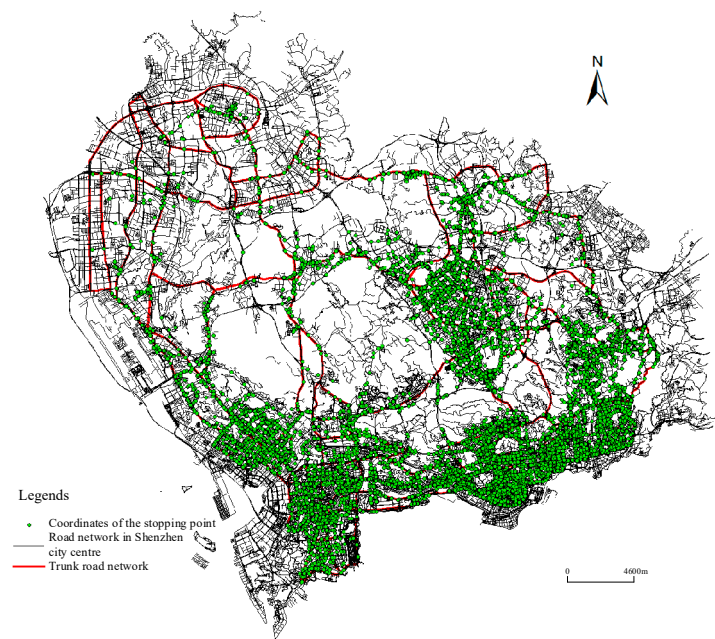


Figure 7. Velocity Sequence Linear Clustering algorithm diagram.

The time thresholds of the two states are set to filter the Stop and Move states, respectively, and if the time of the Stop or Move state is less than the set threshold, the original state is converted to another state [18]. In this paper, we set the threshold of the Stop state as 5 min and the threshold of the Move state as 2 min. The trajectory points in the filtered Stop state are defined as stay points, and the stay points are distributed in the study area of Shenzhen city centre, as shown in Figure 8. From Figure 8, it can be seen that the Stop points are distributed in the whole Shenzhen central city, and the number of Stop points is more intensive in Nanshan District, Futian District, and Luohu District.

### 3.3.2. Distribution of Electric Taxi Operation Starting Time

Based on the location of the stopping points in the central city of Shenzhen in Section 3.3.1, these stopping points are screened twice according to the time dimension of the start of taxi operation. The time dimension of the start of taxi operation is generally the end time of the taxi handover process, and the probability distribution of the taxi driver in a day's operation is highly randomised, but the handover time is an important part of the normal operation of taxis, and there is a temporal pattern of dwells. In order to identify the exact start time of the operation, kernel density estimation is used to identify the distribution of the dwell points in the time dimension.

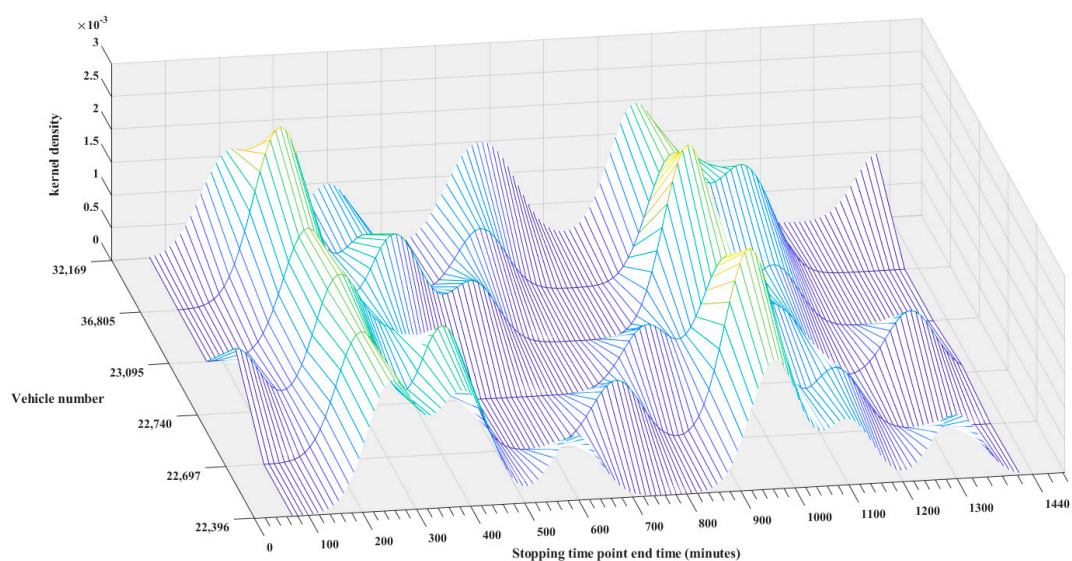


**Figure 8.** Distribution of taxi stops in Shenzhen city centre.

Kernel density estimation is a non-parametric method used to estimate the probability density function, as shown in Equation (1).

$$f_h(t) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{t - E_i^{\text{time}}}{h}\right) \quad (1)$$

In the formula,  $f_h(t)$  is the probability density function of the time distribution of the stopping point,  $n$  is the number of stopping points of a car, and  $h$  is the bandwidth. In order to visualise smoother results, the results of the simulation experiments after many observations are selected;  $h = 0.8$ ,  $K(\cdot)$  is the kernel function and  $E_i^{\text{time}}$  is the stopping point at the end of the time. In order to prevent chance, a probability distribution graph of the stay end time of six taxis is randomly selected, as shown in Figure 9.



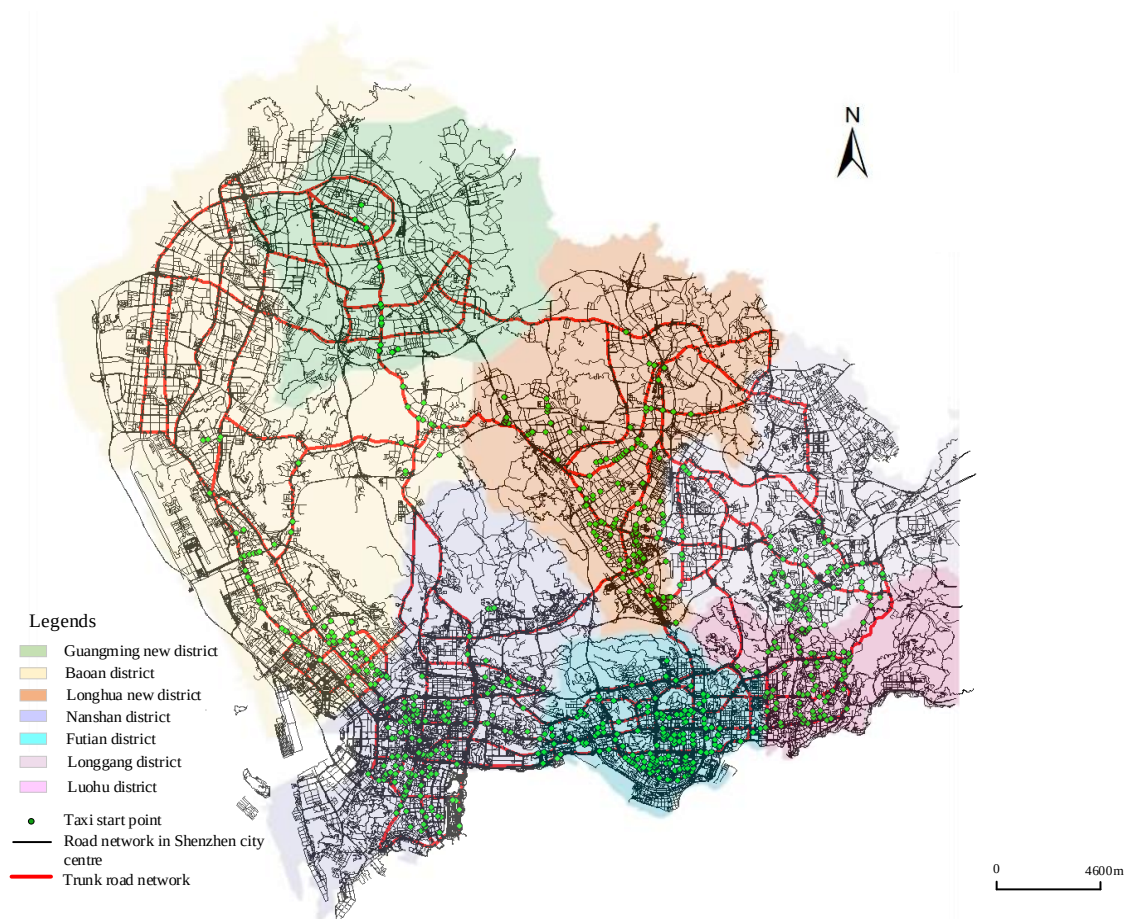
**Figure 9.** Probability distribution of kernel density estimates.



The end time of the taxi shift handover is identified due to the fact that the end time of the stay period at the shift handover stop is the time when the taxi starts operation. The taxi shift handover is generally a double-shift system, with two drivers in charge of each shift. And the duration of each shift is roughly 8–12 h, that is, the end time of the two shifts in a day is 8–12 h apart to meet the actual situation. In Figure 6, it can be seen that there are two peaks in the end time of stopping at the stopping point at about 200–300 min and 900–1000 min, i.e., from 3:20 a.m. to 5:00 a.m. and from 3:00 p.m. to 4:40 p.m., respectively. And the time interval between these two nuclear density peaks is consistent with 8–12 h. So, it is judged that the electric taxis start operating within these two time periods.

### 3.3.3. Spatial Distribution of Electric Taxi Starting Operations

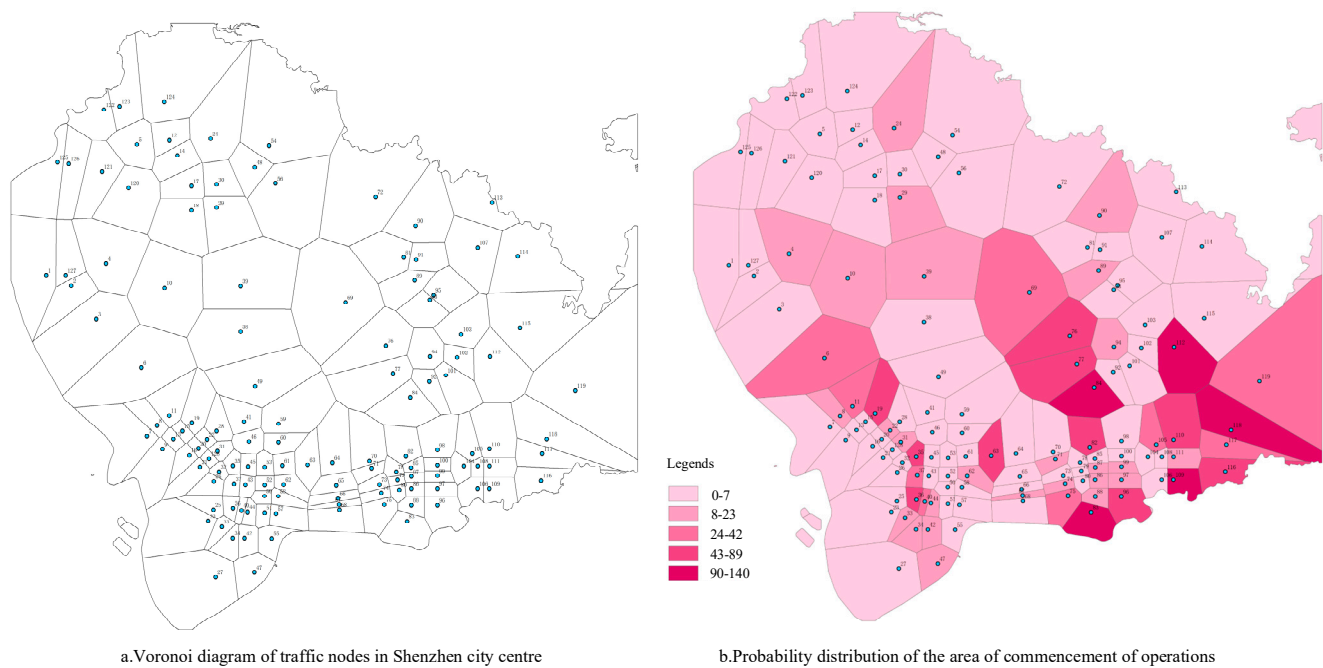
Using the time distribution of electric taxis' starting operation in Section 3.3.2 for the secondary screening of stopping points in the central city of Shenzhen, detected in Section 3.3.1 for the location of the stopping points, the spatial distribution of the screened starting operation is shown in Figure 10. Taxi start operation points are more intensive in the Nanshan District and Futian District, while they are less intensive in the Bao'an District and Guangming District.



**Figure 10.** Distribution of taxi operating points in Shenzhen city centre.

The taxi start operation points are distributed throughout the central city and the number is large. Combining the Tyson polygon in Arcgis and the nodes of the road network in Figure 2 to divide Shenzhen's the central urban area into regions, the Tyson polygon, also known as the Voronoi diagram [19], is characterised by the fact that it can make all the points in the plane be the closest distance to the centre of their respective division areas. The study area is divided into 128 zones using 128 nodes in the main road network of the central city, as shown in Figure 11a. The probability of whether the area is

used as a taxi start operation area is determined based on the number of taxi start operation points in each study area, and the probability distribution graph is shown in Figure 11b.



**Figure 11.** Probability distribution of starting operation area based on Voronoi diagram delineation.

### 3.4. OD Probability of Travelling by Electric Taxi

#### 3.4.1. Travelling OD Extraction

The taxi trajectory data are used to identify the information of the vehicle, such as the vehicle number, origin (O), destination (D), and the start time and end time of each taxi trip, and the output is organised into a table of the OD information of taxi trips, as shown in Table 1, and the extraction process has the following steps:

1. Extracting the operational trajectory Movetraj.
2. Move the OpenStatus column in the Operations track down one row to get NewOpenStatus.
3. Construct a new column StatusChange = OpenStatus – NewOpenStatus to record the change that occurred in the status of the loaded passenger, with 1 being boarded and –1 being disembarked.
4. Determine whether the VehicleNum of the next data is equal to the VehicleNum of these data, and filter each taxi OD.
5. Move all columns of the operation track up one row as a whole and splice them with the original operation track, keep the record with StatusChange = 1, and store it as a taxi travelling OD information table.

**Table 1.** Taxi travel OD information table.

| ID   | VehicleNum | SLng       | SLat      | ELng       | ELat      |
|------|------------|------------|-----------|------------|-----------|
| 1    | 22,437     | 113.905806 | 22.577754 | 113.886984 | 22.561491 |
| 2    | 22,437     | 114.042281 | 22.60275  | 114.024386 | 22.636292 |
| ...  | ...        | ...        | ...       | ...        | ...       |
| 1292 | 25,956     | 113.928941 | 22.525063 | 113.918656 | 22.527208 |
| 1293 | 25,956     | 113.934658 | 22.485559 | 114.044273 | 22.542579 |
| ...  | ...        | ...        | ...       | ...        | ...       |
| 2045 | 28,098     | 113.928941 | 22.525063 | 114.055438 | 22.613142 |
| 2046 | 28,098     | 113.949449 | 22.583541 | 114.091121 | 22.543436 |
| ...  | ...        | ...        | ...       | ...        | ...       |

### 3.4.2. Probability of Travelling OD

The latitude and longitude of the OD points in the OD information table of taxi trips in Section 3.4.1 are matched to the Voronoi diagram of the traffic nodes in the central city of Shenzhen in Figure 10a, respectively, and the matching results are shown in Table 2.

**Table 2.** Information on the urban areas to which each OD coordinate point belongs.

| ID   | VehicleNum | SArea | SLng       | SLat      | EArea | ELng       | ELat      |
|------|------------|-------|------------|-----------|-------|------------|-----------|
| 1    | 22,437     | 37    | 113.905806 | 22.577754 | 11    | 113.886984 | 22.561491 |
| 2    | 22,437     | 95    | 114.042281 | 22.60275  | 112   | 114.024386 | 22.636292 |
| ...  | ...        | ...   | ...        | ...       | ...   | ...        | ...       |
| 1292 | 25,956     | 15    | 113.928941 | 22.525063 | 16    | 113.918656 | 22.527208 |
| 1293 | 25,956     | 12    | 113.934658 | 22.485559 | 75    | 114.044273 | 22.542579 |
| ...  | ...        | ...   | ...        | ...       | ...   | ...        | ...       |
| 2045 | 28,098     | 15    | 113.928941 | 22.525063 | 98    | 114.055438 | 22.613142 |
| 2046 | 28,098     | 20    | 113.949449 | 22.583541 | 84    | 114.091121 | 22.543436 |
| ...  | ...        | ...   | ...        | ...       | ...   | ...        | ...       |

For further follow-up work, the OD matrix needs to be transformed into the probability matrix of access between each city region, i.e., OD probability matrix  $OD_{p_{i,j}}$ , whose transformation formula is shown in Equation (2):

$$OD_{p_{i,j}} = \frac{OD_{i,j}}{\sum_{i=1}^{128} OD_{i,j}} \quad (2)$$

where,  $OD_{p_{i,j}}$  denotes the probability that a taxi starts at city region  $i$  and ends at city region  $j$  and  $OD_{i,j}$  denotes the number of taxi trips from city region  $i$  to city region  $j$ . The calculation results are counted as a  $128 \times 128$  OD probability matrix. This then provides data support for the subsequent simulation of electric taxi operation.

### 3.5. Shortest Route for Travelling by Electric Taxi

The geometric calculator in Arcgis is used to calculate the distance to the road sections divided by 128 road network nodes in the main road network in the center of Shenzhen, and the results are shown in Figure 12. The results are counted into a  $128 \times 128$  distance matrix of urban road network nodes, as shown in Table 3.

Based on the OD probability of taxi travelling in each city area in Section 3.4, the next travel location is decided, but there are multiple choices of travel paths, considering that the taxi driver operates with maximum revenue, and the wrong choice of the optimal travel path leads to a reduction in the amount of power and time available for revenue. So, the shortest path algorithm is chosen to determine the driver's path choice between point O and point D. In this paper, the shortest path between each pair of road network nodes is solved by Dijkstra's algorithm. Dijkstra's algorithm [20] belongs to a kind of greedy algorithm, which calculates the shortest path by constantly choosing the node closest to the current node, and the results of the calculation are shown in Table 4. Its calculation steps are as follows:

1. Initialise the node number, shortest path, and distance matrix.
2. The loop traverses each node as a central node, selects  $i$  as the central node, initialises the distance between node  $i$  and other nodes, creates a labelling matrix, and determines whether the labelled node has been visited.
3. Iterate through all the nodes and select the unvisited nearest node, MinNode.
4. Add the distance and path from node  $i$  to the MinNode to the set of shortest distances and paths for node  $i$ .
5. With the MinNode as the search object, calculate the distance to its neighbouring nodes and find the next shortest distance node, NextMinNode.

6. If the NextMinNode is not the last node, repeat step 5.
7. If  $i$  is not the last node, repeat step 3.
8. Output the shortest path and distance between each node.

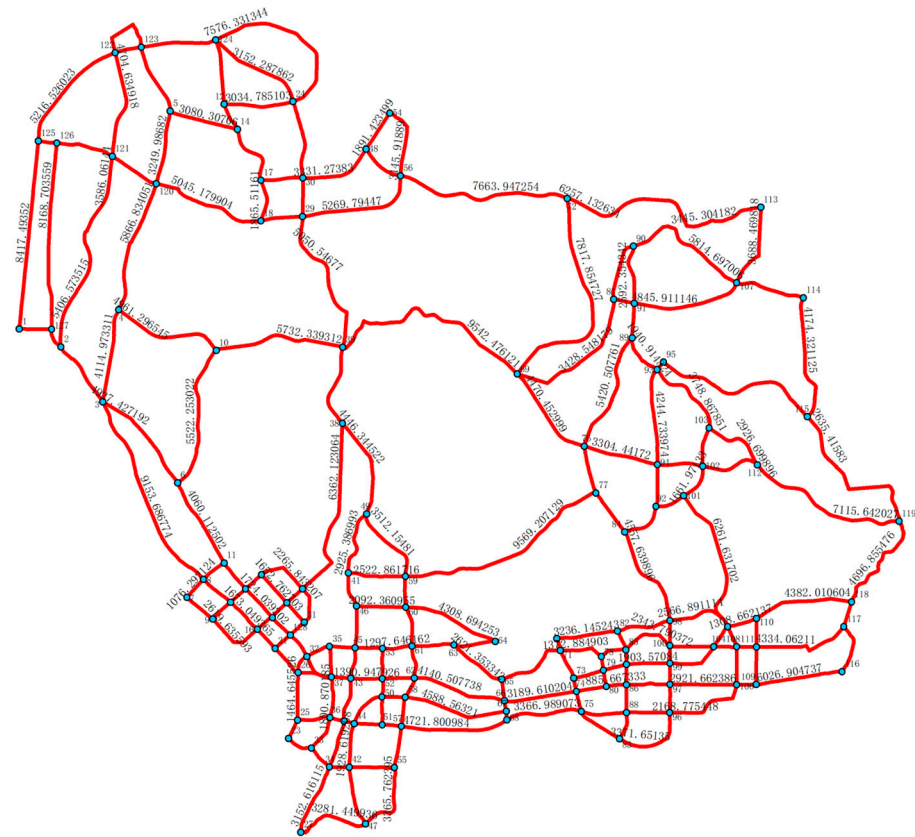


Figure 12. Length of each road network of trunk routes in Shenzhen city centre.

Table 3. Distance matrix of road network nodes.

| Unit: m  | Area 1 | Area 2 | Area 3 | ... | Area 64 | ... | Area 127 | Area 128 |
|----------|--------|--------|--------|-----|---------|-----|----------|----------|
| Area 1   | 0      | Inf    | Inf    | ... | Inf     | ... | 1412     | Inf      |
| Area 2   | Inf    | 0      | 3068   | ... | Inf     | ... | 934      | Inf      |
| Area 3   | Inf    | 3068   | 0      | ... | Inf     | ... | Inf      | Inf      |
| ...      | ...    | ...    | ...    | ... | ...     | ... | ...      | ...      |
| Area 64  | Inf    | Inf    | Inf    | ... | Inf     | ... | Inf      | Inf      |
| ...      | ...    | ...    | ...    | ... | ...     | ... | ...      | ...      |
| Area 127 | 1412   | 934    | Inf    | ... | Inf     | ... | 0        | Inf      |
| Area 128 | Inf    | Inf    | Inf    | ... | Inf     | ... | Inf      | 0        |

Table 4. Matrix of shortest distances for each node of the road network.

| Unit: m  | Area 1 | Area 2 | Area 3 | ... | Area 64 | ... | Area 127 | Area 128 |
|----------|--------|--------|--------|-----|---------|-----|----------|----------|
| Area 1   | 0      | 2343   | 5411   | ... | 26,018  | ... | 1413     | 16,415   |
| Area 2   | 2343   | 0      | 3068   | ... | 23,675  | ... | 934      | 14,072   |
| Area 3   | 5411   | 3068   | 0      | ... | 20,607  | ... | 3998     | 11,004   |
| ...      | ...    | ...    | ...    | ... | ...     | ... | ...      | ...      |
| Area 64  | 26,018 | 23,675 | 20,607 | ... | 0       | ... | 24,605   | 9603     |
| ...      | ...    | ...    | ...    | ... | ...     | ... | ...      | ...      |
| Area 127 | 1412   | 934    | 3998   | ... | 24,605  | ... | 0        | 15,002   |
| Area 128 | 16,415 | 14,072 | 11,004 | ... | 9603    | ... | 15,002   | 0        |

The battery loading state of an electric taxi is constantly changing during operation. The power usage of electric taxis in different states is calculated by the driving mileage and power consumption per kilometre. The specific calculation formula is as follows:

$$S_{next} = \frac{S_{taxi} \times C_b - P_{(taxi,next)} \times C_{energy}}{C_b} \quad (3)$$

where  $S_{next}$  is the SOC value of the electric taxi at the next location,  $S_{taxi}$  is the SOC value of the current electric taxi,  $P_{(taxi,next)}$  is the shortest path between this location and the next location, and  $C_{energy}$  is the power consumption per kilometre.  $C_b$  is the battery capacity of the electric taxi.

Since the electric taxis are fully charged at the end of the working hours before the shift change to supply the next shift of drivers for normal operation [21], the battery load state at the time of the start of the electric taxi's operation is assumed to be the initial charge value. When the battery load state is lower than 15%, it indicates a state of extreme charging anxiety, at which time the taxi's power cannot support further travelling, and the driver can only choose to charge in the region.

#### 4. Electric Taxi Charging Decision-Making Model Based on DQN Algorithm

##### 4.1. Definition of Elements of Reinforcement Learning in Charging Decisions

###### 4.1.1. Basic Assumptions of the Model

The rational design of the environment in which the intelligences interact can make it easier for the intelligences to find the optimal strategy. Therefore, the quality of the design of the interaction environment directly affects the merits of the final decision. In order to facilitate the problem, four assumptions need to be made before establishing the interaction environment:

1. When charging is selected in an area of the city, charging is performed in this area.
2. The travelling speed of the electric taxi is fixed and is the average speed of the operational trajectory data.
3. During the charging process, most taxi drivers would like to replenish the required power in a short period of time in order to carry passengers subsequently. Therefore, it is assumed that fast-charging chargers are selected for the charging process.
4. Taxi drivers always fully charge their vehicles at the end of the last working hour for the next operation of the vehicle, so it is assumed that the taxi is fully charged before operation, i.e., 90% of the battery capacity of the electric taxi.
5. Exclude the time taken up by situations caused by road traffic congestion, natural disasters, or possible risks affecting the behavioural decisions of taxi drivers, i.e., the time to the destination from all origins is related to distance only.
6. For the time being, we do not consider the impact of the drivers' fatigue level and emotional state, as well as the carrier's incentive and penalty mechanisms on drivers' operational behavioural decisions.

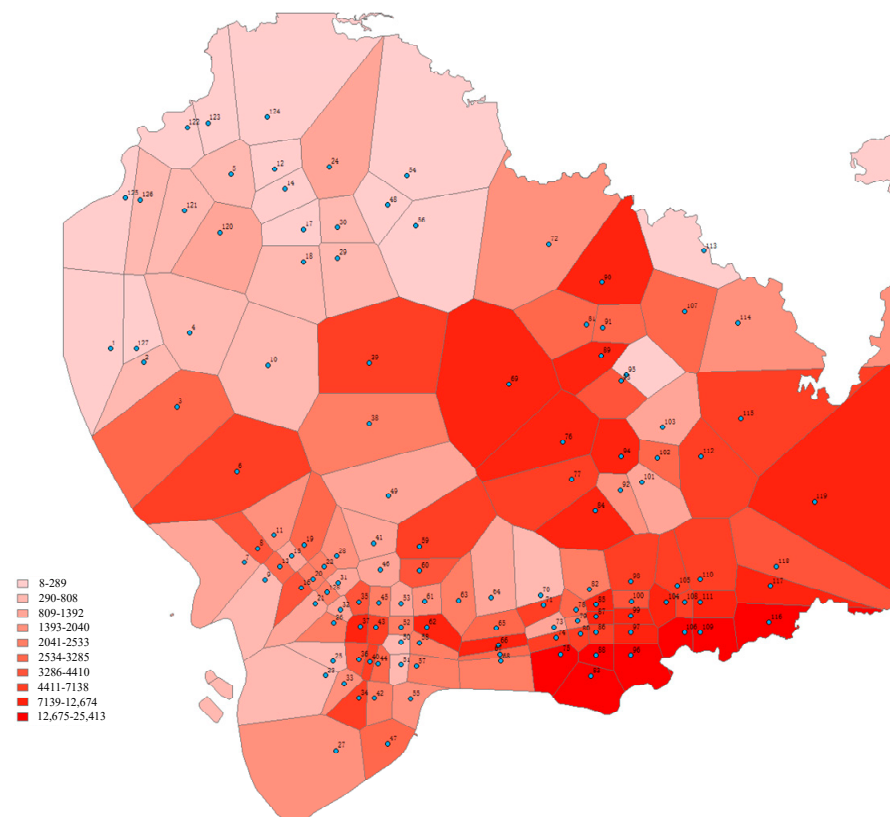
###### 4.1.2. State Space

In this paper, the state space is divided into three elements according to the taxi operation trajectory. These elements are a spatial state, temporal state, and taxi battery state of charge (SOC), which is defined as  $st = [t, pt, bt]$ , that is, the remaining charge situation of taxis at different locations at different times is entered as a sub-state of the total state space.

In terms of time states, the passenger demand of taxis shows obvious distribution characteristics at different times, while the charging cost of public charging piles likewise varies according to the time of day, and at present, most regional charging stations adopt time-differentiated charging pricing strategies, with higher EV charging prices taken during peak hours to encourage users to prioritise charging during low-peak hours. Therefore, the time element has an important impact on charging decision making, and this paper divides the time with 1 h granularity, i.e., the time dimension in a day is divided into 24 time states.



In terms of the spatial state, the number of taxi GPS track points in 128 city regions divided by the Voronoi diagram is used to divide the spatial state. This paper is divided into 10 spatial states, and the results of the division are shown in Figure 13. The darker colour represents the more taxi track points in the region and the greater the probability of carrying more passenger revenue. However, the price of electricity varies according to the region, with a lower electricity cost in remote areas and a higher electricity cost in central areas, and the darker colour represents the higher charging cost in the region. Therefore, different regions have an important impact on charging decisions.



**Figure 13.** Results of space state division.

In terms of the taxi SOC status, each vehicle is divided into four states based on its SOC:  $\text{SOC} < 15\%$ ,  $15\% \leq \text{SOC} < 30\%$ ,  $30\% \leq \text{SOC} < 60\%$ , and  $60\% \leq \text{SOC} < 90\%$ . A SOC of 15% is the charging extreme anxiety power, that is, the taxi power cannot support the continuation of driving, and the driver can only choose to go to the charging point. In order to protect the electric vehicle battery during charging, it is generally not chosen to be fully charged, rather the best maximum charge level is 90%. In order to consider the peak time and peak area of residents' travelling demand, different time and different places have a certain influence on taxi drivers' charging decisions based on the SOC status.

In summary, the total number of state spaces containing spatial states, temporal states, and taxi battery charge states is  $N = 10 \times 24 \times 4 = 960$ .

#### 4.1.3. Action Space

The intelligent body electric taxi makes appropriate actions based on the current system state. A total of four actions are defined for exploring the charging decision of the taxi driver's operational species with the goal of whether to choose charging while carrying passengers, driving empty, fast charging to 60%, or fast charging to 90%. Its set is denoted as  $a_t \in \{c, e, r_{60}, r_{90}\}$ .

#### 4.1.4. Reward Value Modelling

The reasonableness of defining the reward value function is a key factor in determining whether the reinforcement learning algorithm can reflect the decision-making process of drivers in reality. In this paper, the reward value function is set as the revenue generated by executing the corresponding action when the taxi is in a certain state, i.e., the optimal action  $a_t \in \{c, e, r_{60}, r_{90}\}$  is selected in the decision-making process of the operation of electric taxis according to the state  $s_t = [t, p_t, b_t]$  in which it is located. Accordingly the state undergoes a transfer  $T = \{s_t \rightarrow s_{t+1}\}$  to the next state  $s_{t+1} = [t+1, p_{t+1}, b_{t+1}]$ . State shifts occur through constant decision making during operation to maximise the benefit to the intelligence. The reward value model for different state-action combinations is defined as follows:

##### 1. Take passengers on board

$$R_{\text{carry}} = P_{\text{trip}} \times M_{\text{taxi}} - P_{\text{trip}} \times C_{\text{energy}} \times C_{\text{power}} - T_{\text{trip}} \times E_{\text{expect}} - \frac{\lambda}{S_{\text{taxi}}^2} \quad (4)$$

Among them,  $R_{\text{carry}}$  is the reward value of a passenger trip;  $P_{\text{trip}}$  indicates the driving distance of this passenger trip;  $M_{\text{taxi}}$  is the taxi fare, and the fare per kilometre is set at RMB 3.8 with reference to the Shenzhen taxi fare standard;  $T_{\text{trip}}$  is the time of this passenger trip;  $C_{\text{energy}}$  is the value of the power consumed per kilometre;  $C_{\text{power}}$  is the cost of the electric power;  $E_{\text{expect}}$  is the expected revenue per unit of time of the passenger trip;  $\lambda$  is the penalty factor for low power consumption, and a reasonable value is selected through the results of several tests; and  $S_{\text{taxi}}$  is the percentage of the remaining power of the electric taxi.

##### 2. Idle

$$R_{\text{empty}} = -P_{\text{empty}} \times C_{\text{energy}} \times C_{\text{power}} - T_{\text{empty}} \times E_{\text{expect}} - \frac{\lambda}{S_{\text{taxi}}^2} \quad (5)$$

where  $R_{\text{empty}}$  is the value of the bonus per unit of time for a particular empty trip,  $P_{\text{empty}}$  is the distance travelled for the empty trip, and  $T_{\text{empty}}$  is the time of the empty trip.

##### 3. Fast charging up to 60%

$$R_{\text{charging60\%}} = -\frac{(60\% - S_{\text{taxi}})C_b}{V_{\text{charging}}} \times (E_{\text{expect}} - M_{\text{charging}}^{t,l}) - \frac{\lambda}{(60\%)^2} \quad (6)$$

where  $R_{\text{charging60\%}}$  is the value of the unit time incentive for selecting charging and charging to 60%,  $C_b$  award is the battery capacity of the electric taxi,  $V_{\text{charging}}$  is the charging speed of the rapid charging pile, and  $M_{\text{charging}}^{t,l}$  is the charging tariff of the electric taxi at location  $l$  at time  $t$ .

##### 4. Fast charging up to 90%

$$R_{\text{charging90\%}} = -\frac{(90\% - S_{\text{taxi}})C_b}{V_{\text{charging}}} \times (E_{\text{expect}} - M_{\text{charging}}^{t,l}) - \frac{\lambda}{(90\%)^2} \quad (7)$$

where  $R_{\text{charging90\%}}$  is the reward value for selecting charge and charging to 90%.

#### 4.2. Optimisation of Electric Taxi Charging Decision Based on DQN Model

##### 4.2.1. Deep Q Learning Algorithm

The DQN algorithm [22] is a deep reinforcement learning algorithm that combines deep learning and Q-learning for training intelligences to learn optimal strategies in complex environments. The core idea of the algorithm is to deal with the challenges of complex state spaces by approximating the Q-function using deep neural networks. For the environment designed in this paper with only 960 states, if the Q-learning algorithm is used for the taxi charging decision simulation, although the dimensionality explosion problem will

not occur, the number of states is too large and easily leads to poor experimental results, and the use of the DQN algorithm with a simple network structure can complete the taxi charging decision simulation faster and better.

The update rule for Q-learning in reinforcement learning [23] is to optimise an iteratively computable Q-function. The Q-function is a state-action value function with the formula:

$$Q(s, a) = R(s, a) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \quad (8)$$

where  $s$  is the state currently in,  $a$  is the action taken in the current state,  $s_{t+1}$  is the next state that the intelligent body enters after taking action  $a$  in state  $s$ ,  $a_{t+1}$  is the action taken in state  $s_{t+1}$  that makes the state-action value function  $Q(s_{t+1}, a_{t+1})$  maximal, and  $\gamma$  is the discount factor, which serves to balance the gains between immediate and future rewards.

The maximum Q-value in the current state is used to update the target Q-value. The update of the target Q-value is obtained through the Bellman equation, which is calculated by discounting the immediate and future rewards in such a way that it balances the importance of the immediate and future rewards. The formula for updating its Q-value is:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [R_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (9)$$

where  $\alpha$  is the learning rate.

In the traditional Q-Learning algorithm, the Q-table is usually used to store the Q-value corresponding to each state-action, and the Q-table is queried to decide the action to be taken by the intelligent body at the next control step. The DQN algorithm, on the other hand, is based on the Q-Learning algorithm using a deep neural network  $Q_\omega(s, a)$  as an approximator of the Q-function, where  $\omega$  is the parameter that the neural network uses to fit the function  $Q$ . The input of this network is the state and the output, which is the Q-value of each action, which represents the state-action value for  $Q_\omega(s, a)$ . The goal of the final update of the DQN network is to make  $Q_\omega(s, a)$  approximately into  $r + \gamma \max_{a_{t+1}} Q_\omega(s_{t+1}, a_{t+1})$ , defining the loss function of the Q-network constructed as the mean square error between the target Q-value and the Q-value:

$$L(\omega) = [Q_\omega(s, a) - (r + \gamma \max_{a_{t+1}} Q_\omega(s_{t+1}, a_{t+1}))]^2 \quad (10)$$

In order to reduce the correlation between samples, the DQN algorithm introduces the Experience Replay (ER) mechanism [14]. The experience samples of an intelligence's interaction with the environment include the state,  $s$ , action,  $a$ , reward,  $r$ , and the next state,  $s_{t+1}$ , which are stored in the experience replay buffer. At each update of the Q-network, a batch of experience samples are randomly sampled from the experience playback buffer to be used for training. In the ptimization process, in order to balance the weights of exploration and gain, the DQN employs an  $\epsilon$ -greedy strategy, which selects random actions for exploration based on the probability of  $\epsilon$ . Otherwise, it selects the action that currently has the highest Q-value for exploitation. Its strategy formula is:

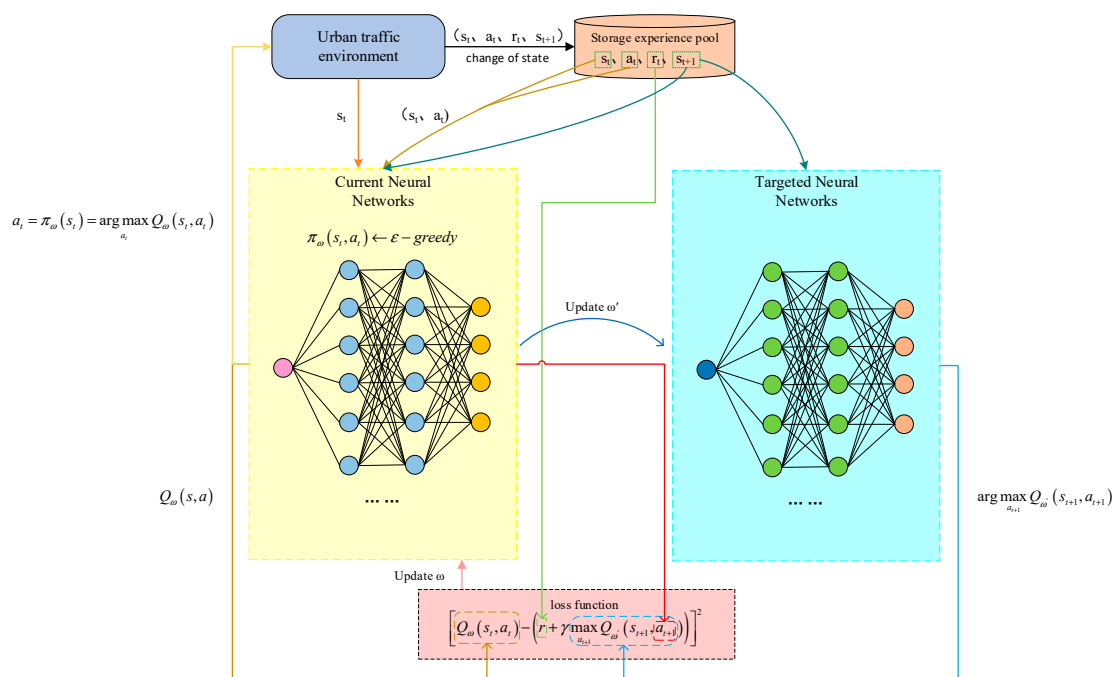
$$\pi(a|s) = \begin{cases} \text{random}(Q_\omega(s, a) \geq 0) & , r \geq 1 - \epsilon \\ \max(Q_\omega(s, a)) & , \text{else} \end{cases} \quad (11)$$

where  $Q_\omega(s, a)$  denotes the corresponding Q-value in the Q-table,  $\pi(a|s)$  denotes the action decision made, and  $r$  is a random value obeying a standard normal distribution. For the setting of  $\epsilon$ , in order to improve the efficiency of the algorithm and limit the number of iterations to an acceptable range, as well as to ensure the certainty of the action selection in the later stage of the training, a larger  $\epsilon$  is set initially and then  $\epsilon$  is reduced with the increase in the number of trainings in order to avoid missing the optimal solution. Through the process of loop iteration, the DQN algorithm gradually optimises the approximation of

the Q-function so that it approaches the optimal Q-function. Thus, the intelligent body can make the optimal decision-making strategy.

#### 4.2.2. DQN-Based Charging Decision Optimisation Strategy Training Approach

The idea of DQN is used to select the action variables in the operation process of electric taxis. There is a very important module in the DQN algorithm-objective network. The loss function in the iterative process of the Q-network uses the temporal difference error target  $r + \gamma \max_{a_{t+1}} Q_{\omega}(s_t, a_t)$  to incrementally update  $Q_{\omega}(s, a)$ , but the temporal difference error target itself contains the output of the neural network, which will make the network parameters update at the same time as the target changes, causing instability in neural network training. In order to make the neural network training more stable, two sets of Q-networks are utilised [14]: the target network and the original Q training network, where the target network calculates the  $r + \gamma \max_{a_{t+1}} Q_{\omega'}(s_{t+1}, a_{t+1})$  term in the loss function and the original training network calculates the  $Q_{\omega}(s, a)$  term in the loss function. The parameter  $\omega$  in the training network is updated every step, while the parameter in the target network is updated once at an interval of N steps in synchronisation with the parameter  $\omega'$  of the training network. This means that the target network will be more stable. The flowchart of the DQN algorithm in the simulation flow is shown in Figure 14.



**Figure 14.** Network structure of DQN algorithm.

#### 4.3. Electric Taxi Charging Load Prediction Process Based on Trajectory Data

Based on the trajectory data as the support of various decisions in the simulation process, a starting point and time of departure are randomly selected according to the real driving law of the vehicle, a suitable path is chosen, and then a city location is randomly selected for operation. Then the time and energy consumption required by the vehicle are derived according to the vehicle travelling laws on each road in the city as well as the distance between the two points, and then, finally, multiple iterations are carried out. With reinforcement learning ideas as the core of charging behaviour decision making, the Monte Carlo method (MC) [24] is used to simulate the taxi travel behaviour and charging habits, and its flow is shown in Figure 15.

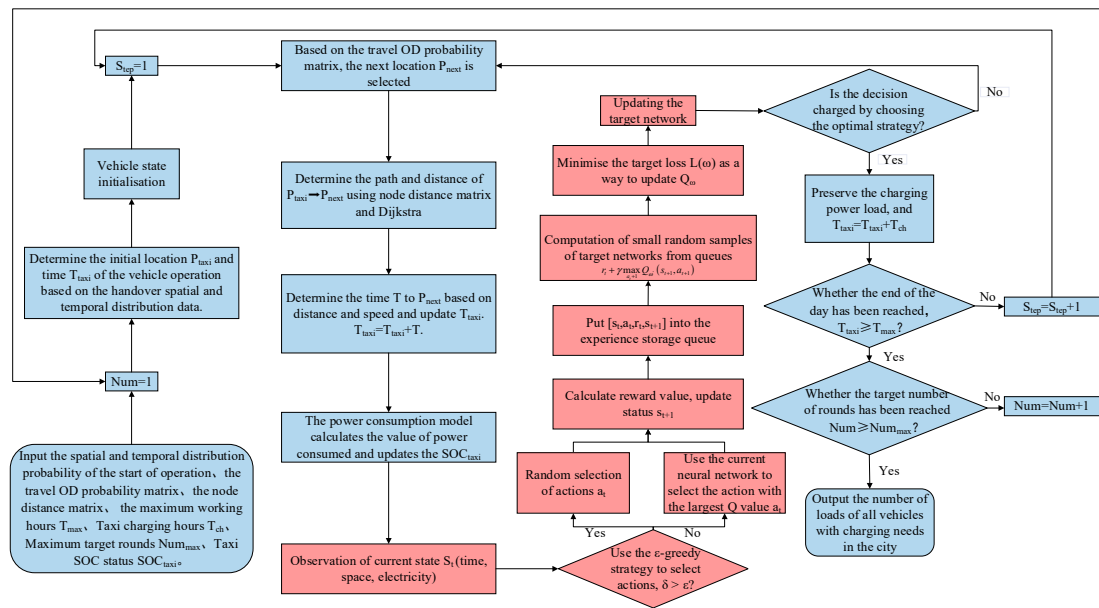


Figure 15. Simulation flowchart of electric taxi operation process.

## 5. Tests and Analyses

### 5.1. Intelligent Body Environment and Parameter Settings

In this paper, the environment of the intelligent body electric taxi is the central city of Shenzhen in Section 3.2, which contains 128 road network nodes and 218 routes in the scope of main roads. The intelligent body in the experiment is selected as the mainstream model taxi BYD e6 in Shenzhen, according to BYD's official reference, and its basic parameters are shown in Table 5, in which the average speed is the average value of all motion trajectory speeds in the trajectory data in which the taxi is in the operation state, including the speed of zero.

Table 5. Basic parameters of electric taxis.

| Parametric                 | Notation       | Retrieve a Value | Unit    |
|----------------------------|----------------|------------------|---------|
| Battery capacity           | $C_b$          | 60               | kwh     |
| Fast-charging power        | $V_{charging}$ | 40               | kw      |
| Electricity costs          | $C_{power}$    | 5                | rmb/kwh |
| Average speed of operation | $\bar{V}$      | 42               | km/h    |
| Power consumption per unit | $C_{energy}$   | 0.2              | kwh/km  |

The charging costs in different areas are set based on the distribution of taxi track data points in different urban areas in Section 4.1.2, with charging costs increasing in areas with high traffic flow and decreasing in areas with low traffic flow. The charging tariffs for taxis in the area with the smallest traffic flow refer to the peak and valley levelling tariffs published by Shenzhen [25], with the peak hours for taxis being 8:00–11:00 and 18:00–23:00; the levelling hours being 7:00–8:00 and 11:00–18:00; and the trough hours being 7:00–8:00 and 11:00–18:00. According to the probability distribution of the taxi trajectory data points, 0.03 rmb/(kW-h) is added to the traditional peak-valley levelling tariff, and the specific tariff settings are shown in Table 6.

The hyperparameters of the deep reinforcement learning DQN model at the core of the taxi operation simulation include the discount factor  $\gamma$ , the learning rate  $\alpha$ , the batch size of the samples drawn from the experience pool each time, and the initial value of  $\epsilon$  and the decay rate of  $\epsilon$  during the selection of the  $\epsilon$ -greedy strategy. The parameter tuning is carried out through repeated trials, and the specific parameter settings are shown in Table 7.



**Table 6.** Setting of charging tariffs by time division and area (Tariff unit: rmb/kwh).

| Area Code                                                              | Time Period | 7:00–8:00,<br>11:00–18:00 | 8:00–11:00,<br>18:00–23:00 | 23:00–07:00 the<br>Following Day |
|------------------------------------------------------------------------|-------------|---------------------------|----------------------------|----------------------------------|
| 75, 83, 88, 96, 106, 109, 116                                          |             | 1.17                      | 1.42                       | 0.87                             |
| 37, 62, 66, 69, 76, 84, 85, 87, 89, 90, 94, 97, 99, 104, 111, 117, 119 |             | 1.14                      | 1.39                       | 0.84                             |
| 6, 34, 39, 40, 43, 59, 71, 74, 77, 80, 86, 98, 105, 108, 110, 112, 115 |             | 1.11                      | 1.36                       | 0.81                             |
| 8, 16, 35, 36, 60, 93, 100, 118                                        |             | 1.08                      | 1.33                       | 0.78                             |
| 3, 13, 19, 44, 47, 65, 67, 78, 81, 91, 102, 107                        |             | 1.05                      | 1.3                        | 0.75                             |
| 20, 22, 26, 38, 42, 45, 52, 57, 58, 63, 68, 79, 82, 128                |             | 1.02                      | 1.27                       | 0.72                             |
| 11, 27, 28, 33, 55, 61, 72, 92, 114                                    |             | 0.99                      | 1.24                       | 0.69                             |
| 7, 15, 21, 24, 31, 32, 41, 46, 49, 53, 64, 73, 101, 102, 120           |             | 0.96                      | 1.21                       | 0.66                             |
| 2, 4, 5, 9, 10, 18, 23, 25, 29, 30, 50, 51, 70, 121, 126               |             | 0.93                      | 1.18                       | 0.63                             |
| 1, 12, 14, 17, 48, 54, 56, 95, 113, 122, 123, 124, 125, 127            |             | 0.9                       | 1.15                       | 0.6                              |

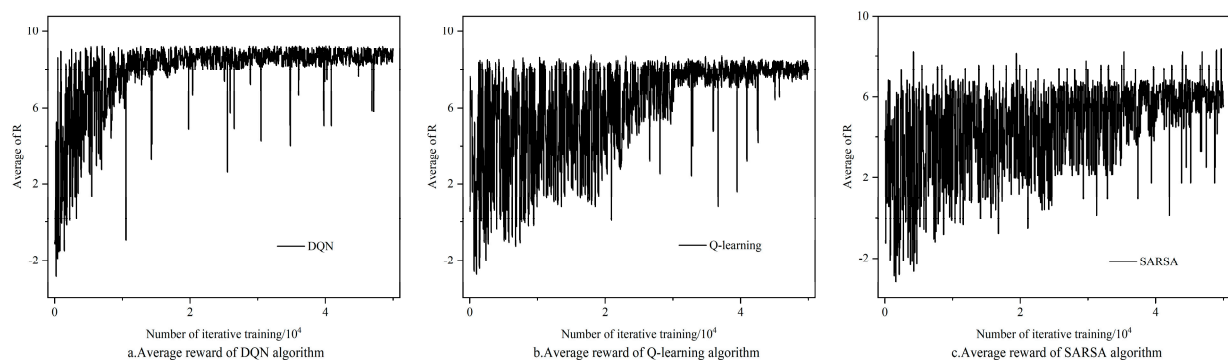
**Table 7.** Model base parameter settings.

| Parametric                  | Notation             | Retrieve a Value | Parametric                                              | Notation     | Retrieve a Value |
|-----------------------------|----------------------|------------------|---------------------------------------------------------|--------------|------------------|
| Discount factor             | $\gamma$             | 0.8              | Samples drawn from the experience pool each time        | Batch size   | 50               |
| Larning rate                | $\alpha$             | 0.001            | Number of training sessions (number of simulated taxis) | Num          | 1000             |
| $\varepsilon$ initial value | $\varepsilon_0$      | 0.7              | Expected passenger yield per unit of time (rmb/h)       | $E_{expect}$ | 8.79             |
| $\varepsilon$ decay rate    | $\varepsilon_{rate}$ | 0.01             | Low-battery penalty factor                              | $\lambda$    | 0.25             |

## 5.2. Simulation Results and Analysis

### 5.2.1. Analysis of Intensive Learning Outcomes

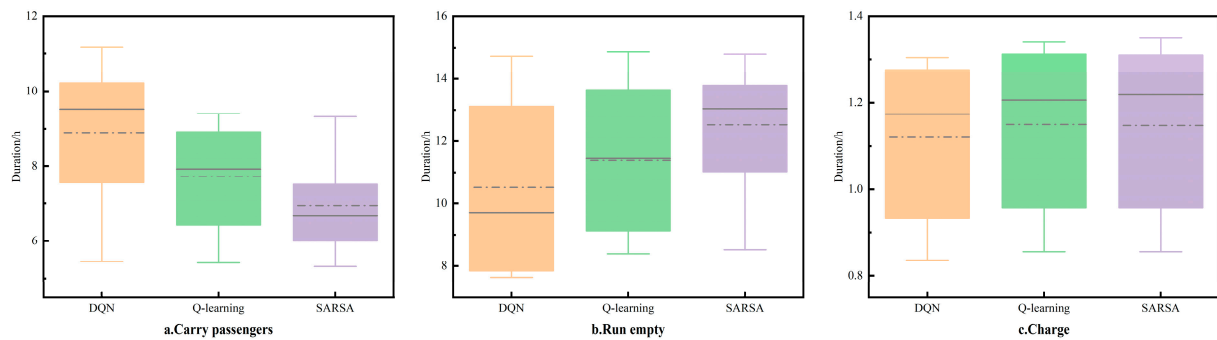
In order to verify the effectiveness of the DQN algorithm in the taxi charging behaviour strategy, the Q-learning algorithm, SARSA algorithm, and DQN algorithm, which also belong to reinforcement learning, are used for learning comparison, respectively, and the average reward function curves of each algorithm in each round during the training process are shown in Figure 16.

**Figure 16.** Average reward comparison of Q-learning, SARSA, and DQN.

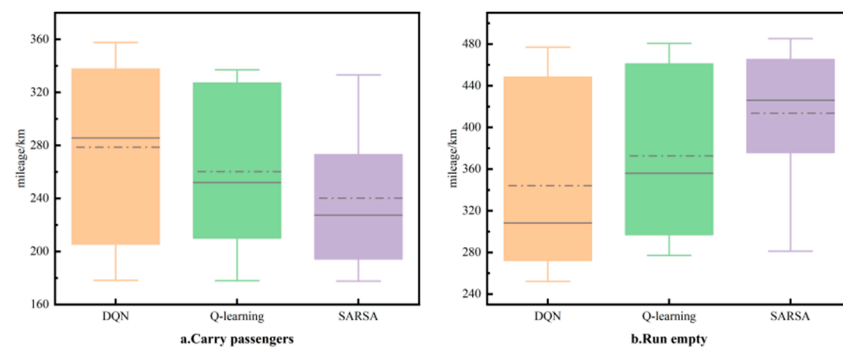
It can be seen in Figure 16c that the average reward for the taxi operation of the SARSA algorithm does not show significant convergence and fluctuates greatly. In Figure 16a,b, it can be seen that the reward value of the DQN and Q-learning algorithms in the initial stage of the intelligent body fluctuates greatly, but the overall trend is on the rise, which indicates

that the intelligent body has a certain amount of exploration and learning ability in this stage. As the number of iterations increases, the learning rate and exploration rate of the intelligent body gradually converge, while the reward value converges to a higher value. However, the average reward of the DQN algorithm taxi starts to converge at around  $1 \times 10^4$  rounds, while the average reward of the Q-learning algorithm taxi operation gradually converges at around  $3 \times 10^4$  rounds, and the average reward value converges at a lower level than that of the DQN algorithm. This shows that the DQN algorithm outperforms the Q-learning and SARSA algorithms in terms of convergence speed and obtains slightly higher average rewards than the Q-learning and SARSA algorithms. This indicates that the deep reinforcement learning DQN algorithm is more stable and rewards converge at a higher level in dealing with the taxi operation optimisation problem.

At the same time, in order to quantitatively describe the revenue effect achieved by the electric taxi driver's operational decision making under different algorithms, the length of time, mileage, and net revenue experienced by all the actions performed by the three algorithms at the end of each round of iteration, including carrying passengers, idling, and charging, were counted for the comparative analysis, and the comparative results are shown in Figures 17–19, where the grey realisation represents the median and the grey dotted line represents the mean.



**Figure 17.** Comparison of the length of each decision for electric taxis under different algorithms.

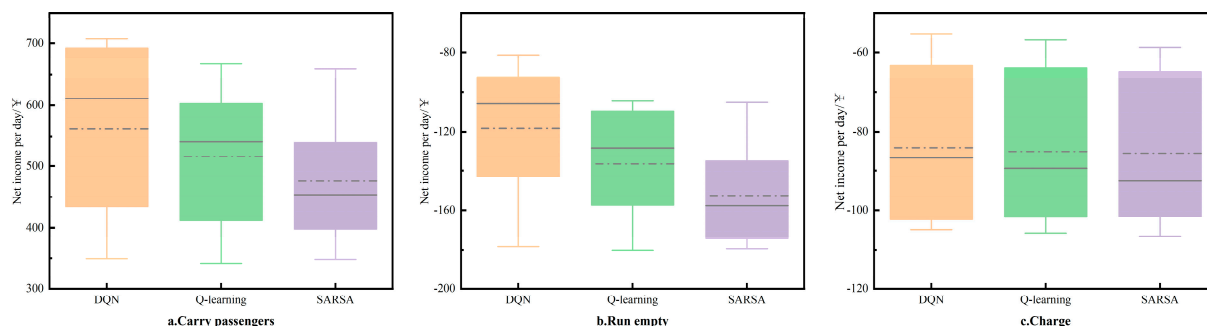


**Figure 18.** Comparison of electric taxi mileage by decision under different algorithms.

In terms of the experienced length of each action, as shown in Figure 17, the average value of the electric taxi carrying passengers under the DQN algorithm is 16.7% and 39.1% more compared to the Q-learning and SARSA algorithms, respectively, and the upper edge is significantly higher than that of the other two algorithms; the average value of the electric taxi idling time under the DQN algorithm is 7.26% and 15.98% less than that of the other two algorithms, respectively, and the lower edge is the lowest. The results of the three algorithms for the electric taxi's charging time are roughly the same, but the DQN algorithm is slightly less.

In terms of the mileage of each action, as shown in Figure 18, the average value of the passenger mileage of electric taxis under the DQN algorithm is 7.1% and 15.9% more

compared to the Q-learning and SARSA algorithms, respectively, and the upper and lower edges of the SARSA are significantly lower than those of the other two algorithms. The average value of the empty mileage of the electric taxi under the DQN algorithm is 7.66% less compared to the other two algorithms, respectively, at 16.8%, and the lower edge of the SARSA algorithm is significantly lower than the other two algorithms.



**Figure 19.** Comparison of net returns of electric taxis by decision day under different algorithms.

In terms of the daily net gain of each action, as shown in Figure 19, the average daily net gain of an electric taxi carrying passengers under the DQN algorithm is 8.97% and 18.16% more than the Q-learning and SARSA algorithms, and the upper and lower edges of the DQN algorithm are significantly higher than those of the Q-learning and SARSA algorithms. The average daily net gain of electric taxi idling is 13.49% and 22.5% less than those of the Q-learning and SARSA algorithms, respectively. The daily net gains of electric taxi charging under the three algorithms have approximately the same results.

From this, it can be concluded that the DQN algorithm based on the electric taxi driver's operational decision making can ensure that the simulation model is reasonable and effective at the same time to derive the optimal operational rules for the taxi driver in a day's work, thus maximising the revenue generated by the taxi driver in the process of operation.

### 5.2.2. Charging Load Prediction Results

From Section 5.2.1, it can be seen that the DQN algorithm, with reinforcement learning as the core idea in the simulation process of taxi operation decision making, is optimal compared with the optimisation strategies of the Q-learning and SARSA algorithms. Therefore, using the operation strategy obtained by the DQN algorithm after  $4 \times 10^4$  rounds, and without updating the strategy, we will carry out another 2000 rounds (i.e., simulate the operation decision making of 2000 electric taxis) to predict the charging load of electric taxis. To predict the charging load of electric taxis, the obtained results are shown as follows: Figure 20 shows the distribution of the charging demand in each region of the central city of Shenzhen, the darker the colour means the higher the demand for charging load in this region. Figure 21 shows the distribution of the total charging load demand in one day in the region served by each node in the central city of Shenzhen.

From Figure 20, the Charging Demand Regional Distribution Map, we can clearly see the distribution characteristics of Shenzhen's electric taxi charging demand in different regions. The Guangming District, Bao'an District, and Nanshan District show a relatively low trend in the charging demand, which may be related to the travelling pattern, population density, and the distribution of charging infrastructure in these districts, etc. In contrast, areas such as Longhua, Futian, Luohu, and Longgang show a trend of relatively high demand. These places are located in the city centre and are closer to transport hubs such as Shenzhen Station, Futian Station, and Shenzhen North Station, so they have a higher flow of people and a higher relative business volume. The higher density of taxi operations naturally leads to a higher demand for charging facilities.

As can be seen from Figure 21, the charging demand for electric taxis in the regions served by each node shows obvious differences, and the regions covered by the four nodes

69, 76, 112, and 118 exhibit high levels in terms of the charging demand for electric taxis, with charging load values of more than 7000 kw. Nine nodes, 117, 89, 96, 19, 77, 119, 90, 83, and 116, serve areas that do not reach the charging load values of the previous four nodes, but still exhibit a considerable charging demand. In most of these areas, the charging load values are above 3000 kw.

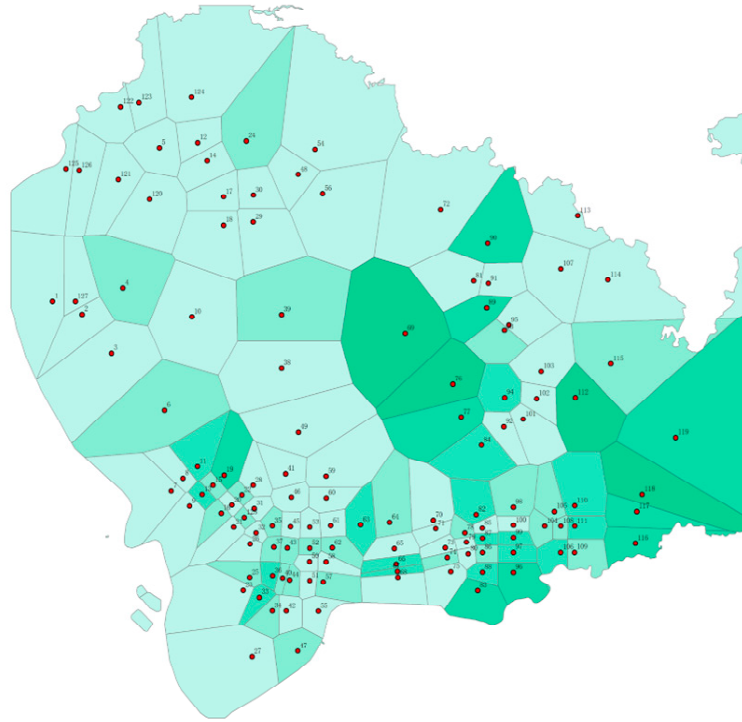


Figure 20. Spatial distribution of charging load.

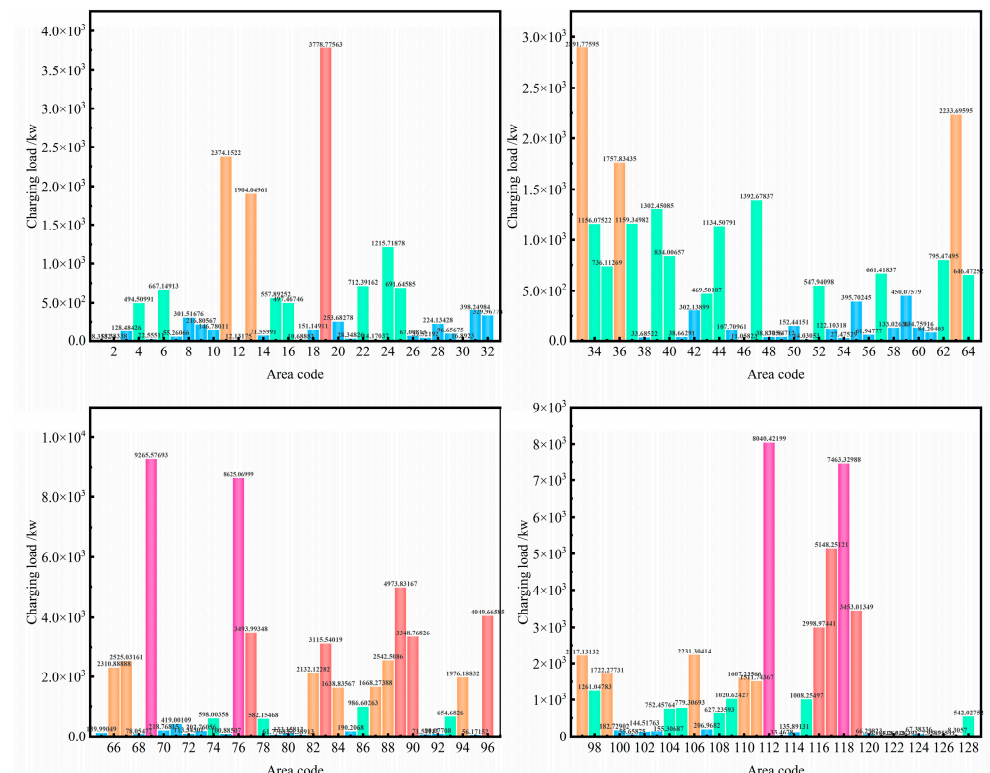
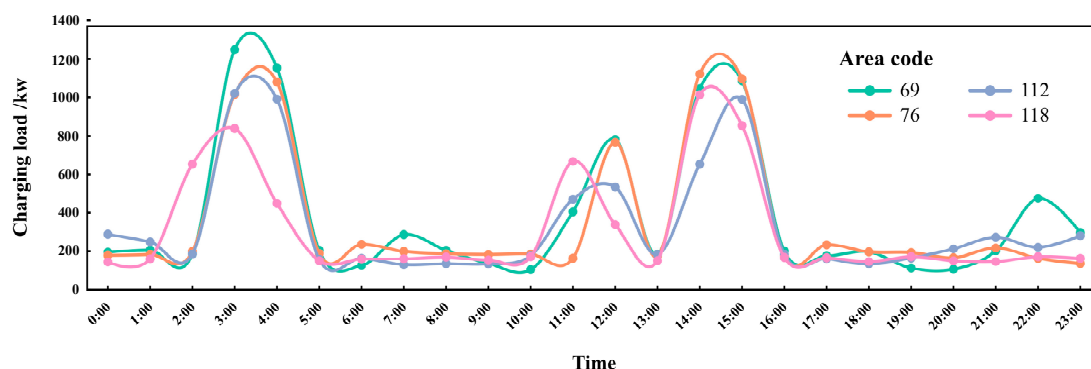


Figure 21. Total charging load statistics of each region.

For the temporal distribution of the daily charging load in each area of the city, the areas covered by the four nodes 69, 76, 112, and 118, which have a relatively high charging demand, are selected for analysis, and the charging load demand of their nodes is shown in Figure 22. It can be seen that there are differences in the charging load demand of the areas covered by each node in different time periods, but the charging load time distribution characteristics are more consistent within a day. It is worth noting that there are two peak hours of charging demand in a day. The first peak hour occurs around 3:00 a.m. to 4:00 a.m., while the second peak hour occurs around 14:00 p.m. to 15:00 p.m. These two hours are closely related to the shift handover time of taxis. These two distinct load peaks are generated due to the low power level of vehicles after long hours of operation and the fact that these hours coincide with the off-peak tariff period. In addition, high charging loads are also presented around 11:00 to 12:00 noon. During this time period, when electricity tariffs are off-peak, many taxi drivers utilise their lunch and lunch breaks to replenish their batteries in order to cope with the operational demand in the afternoon. As a result, the charging demand is relatively high during this time period. In contrast, the charging loads are lower overall during the morning peak (08:00 to 11:00) and evening peak (16:00 to 22:00) time periods. This is due to the fact that these two hours coincide with peak traffic hours and also higher electricity tariffs, hence the charging demand for electric taxis during these hours is relatively low.

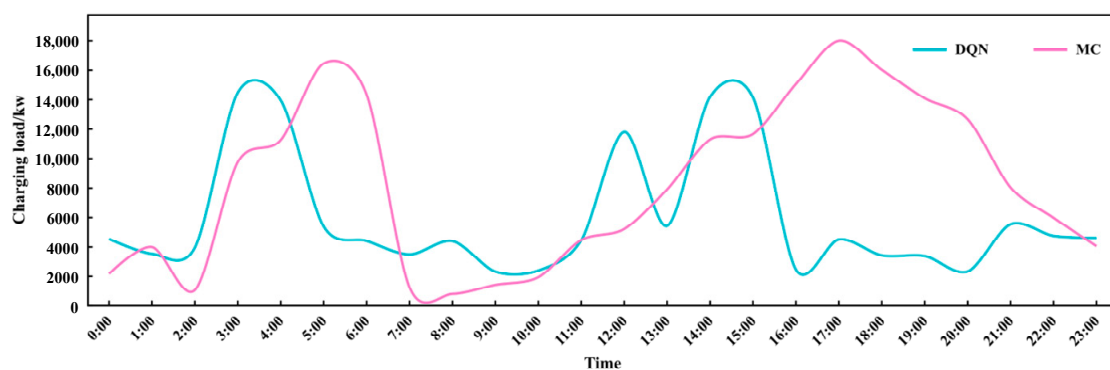


**Figure 22.** Daily charging load time distribution curve.

In the simulation of the taxi's one-day operation behaviour, the charging behaviour criterion is replaced with the MC method used by many scholars [26,27], and the load prediction is carried out under the unchanged parameter characteristics of the traffic network and the intelligent body electric taxi. Monte Carlo, also known as a statistical simulation, is a stochastic simulation method that uses random numbers to solve many computational problems. The problem to be solved is linked to a defined probabilistic model, based on which computer software generates random samples of probability-distributed variables conforming to the model for statistical simulation or sampling in order to obtain an approximate solution to the problem. In the process of simulating the operating behaviour of electric taxis in the city, the starting operating time, the starting operating location, and the next purpose of the run are randomly selected based on the spatio-temporal probabilistic model of the taxis in Section 3, and then the time required by the vehicle and the energy consumption are calculated based on the node distance matrices and energy consumption models in the various roads in the city. Finally, several iterations are carried out to obtain a more accurate result to simulate the operating conditions of the taxis in the city. The specific process is shown in the blue part of Figure 15.

By comparing the MC algorithm to the trajectory data proposed in this paper combined with the reinforcement learning decision-making method, the results of the daily charging load time distribution of electric taxis in the downtown area of Shenzhen City are shown in Figure 23.





**Figure 23.** Comparison of total charging load curves in Shenzhen city center.

In Figure 23, both the MC and DQN methods show two peaks in two time periods, and the peak value of the total charging load of the MC method is higher than that predicted by the DQN method, but the MC method fails to accurately reflect some characteristics such as taxis' higher charging preferences in valley time and taxi drivers charging when they take a break to eat during work. Therefore, compared to the MC method, the total daily charging load predicted by DQN is more in line with taxi drivers' rational operating rules and charging decisions.

## 6. Conclusions

Aiming at the problems of electric taxi charging load prediction, as well as the driver's operation law and charging judgement, this paper successfully constructs a deep reinforcement learning-based decision-making model for electric taxi operation by using taxi GPS trajectory data. By considering multiple factors of time, space, and electricity, the DQN model is constructed for complex taxi operation decision making, with the main goal of maximising the operating revenue of electric taxi drivers. The actual operating rules and charging habits formed by drivers in operation are more reasonably simulated to predict more accurate charging loads of electric taxis. This supports the management of the charging demand, improves service quality, reduces operating costs, as well as promotes the effective use of sustainable energy and the stable operation of the electricity network.

However, electric taxi charging load prediction is an extremely complex process and, in this paper, we only use the data of a certain day in the central city of Shenzhen to construct the model, which may lead to the chance and specificity of the results. Synthesizing multiple areas for analysis is a direction for further exploration in future learning studies. In addition, the deep reinforcement learning model with the introduction of neural networks can be further improved to better adapt to the complex urban traffic environment and different electric taxi operations.

In conclusion, this study provides an effective method for electric taxi charging load forecasting and analysing taxi drivers' operation laws, and provides directions for future research. Reasonable and effective electric taxi charging load forecasting is important for the development of sustainable urban transport and electricity networks.

**Author Contributions:** Conceptualization, X.L., B.L. and D.Y.; methodology, B.L.; software, B.L. and Y.C.; validation, B.L. and Y.Z.; formal analysis, Y.C.; investigation, B.L.; resources, X.L. and B.L.; data curation, Y.C.; writing—original draft preparation, B.L.; writing—review and editing, X.L., B.L. and D.Y.; visualization, B.L.; supervision, X.L. and D.Y.; project administration, X.L.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Human Factors Reliability Study of Ship Pilots Based on HEACS-MPA Model (Project No. JAT210248).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Li, H.; Wang, Y.; Hu, Z.; Luo, F. Carrying Capacity Assessment of Distribution Network for Multiple Access Bodies Under the Background of Double Carbon. *Power Syst. Technol.* **2022**, *46*, 3595–3603.
- Wang, H.J.; Wang, B.; Fang, C.; Li, W.; Huang, H.W. Charging Load Forecasting of Electric Vehicle Based on Charging Frequency. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *237*, 062008. [\[CrossRef\]](#)
- Zhang, Z.; Wang, W.; Zhang, X. Renewable Energy Capacity Planning Based on Carrying Capacity Indicators of Power System. *Power Syst. Technol.* **2021**, *45*, 632–639.
- Zhang, X.; Xiao, X. Charging power demand of electric taxi modeling and influence factors analysis. *Adv. Technol. Electr. Eng. Energy* **2014**, *33*, 21–25.
- Santos, A.; McGuckin, N.; Nakamoto, H.Y.; Gray, D.; Liss, S. *Summary of Travel Trends: 2009 National Household Travel Survey*; Federal Highway Administration: Washington, DC, USA, 2011.
- Xing, Y.; Li, F.; Sun, K.; Wang, D.; Chen, T.; Zhang, Z. Multi-type electric vehicle load prediction based on Monte Carlo simulation. *Energy Rep.* **2022**, *10*, 966–972. [\[CrossRef\]](#)
- Yang, B.; Chen, W.; Wen, M.; Chen, X. Probabilistic load modelling of electric vehicle charging stations. *Power Syst. Autom.* **2014**, *38*, 67–73.
- Liao, B.J.; Yang, J.; Wen, F.S.; Li, B.; Li, L.; Mao, J.W. Temporal and spatial random distribution characteristics of electric taxi charging load. *Electr. Power Constr.* **2017**, *38*, 8–16.
- Xydas, S.; Marmaras, C.E.; Cipcigan, L.M.; Hassan, A.S.; Jenkins, N. Electric Vehicle Load Forecasting using Data Mining Methods. In Proceedings of the IET Hybrid and Electric Vehicles Conference 2013 (HEVC 2013), London, UK, 6–7 November 2013.
- Brady, J.; O'Mahony, M. Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data. *Sustain. Cities Soc.* **2016**, *26*, 203–216. [\[CrossRef\]](#)
- Lee, B.; Lee, H.; Ahn, H. Improving Load Forecasting of Electric Vehicle Charging Stations Through Missing Data Imputation. *Energies* **2020**, *13*, 4893. [\[CrossRef\]](#)
- Zhang, T. New Energy Vehicle Ownership Reaches 18.21 Million. *People's Daily*, 11 October 2023.
- Zhang, W.N.; Shen, K.; Yu, Y. *Hands-on Learning for Intensive Learning*; People's Posts and Telecommunications Press: Beijing, China, 2022.
- Wang, G.; Chen, X.Y.; Zhang, F.; Wang, Y.; Zhang, D.S. Lessons learnt: Understanding long-term evolutionary patterns of shared electric vehicle networks. In Proceedings of the 2019 25th International Conference on Mobile Computing and Networking, Los Cabos, Mexico, 21–25 October 2019.
- Zou, F.; Luo, S.; Chen, Z. A spatial and temporal distribution identification method for taxi handover based on trajectory data. *Comput. Appl.* **2021**, *41*, 3376–3384.
- Spaccapietra, S.; Parent, C.; Damiani, M.L.; de Macedo, J.A.; Porto, F.; Vangenot, C. A Conceptual View on Trajectories. *Data Knowl. Eng.* **2008**, *65*, 126–146. [\[CrossRef\]](#)
- Yang, W.; Ai, T. Detection of refuelling stopping behaviours and extraction of refuelling station points by crowd-sourced vehicle trajectories. *J. Surv. Mapp.* **2017**, *46*, 918–927.
- Rocha, J.A.M.R.; Times, V.C.; Oliveira, G.; Alvares, L.O.; Bogorny, V. DB-SMoT: Adirection-based spatio-temporal clustering method. In Proceedings of the 2010 5th IEEE International Conference on Intelligent Systems, London, UK, 7–9 July 2010.
- Preparata, F.P.; Shamos, M.I. *Computational Geometry: An Introduction*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 226–329.
- Zhang, F.H.; Liu, J.P.; Li, Q.Y. A shortest path optimisation algorithm based on Dijkstra's algorithm. *Remote Sens. Inf.* **2004**, *2*, 38–41.
- Chen, Z. Research on Intelligent Site Selection Method of Urban Taxi Charging Station by Integrating Multi-Dimensional Information. Master's Thesis, Fujian University of Engineering, Fuzhou, China, 2020.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Martin, A. Riedmiller. *Playing Atari with Deep Reinforcement Learning*. *arXiv* **2013**, arXiv:1312.5602.
- Sutton, R.; Barto, A. *Reinforcement Learning: An Introduction*; The MIT Press: Cambridge, MA, USA, 2018.
- Shan, Q. Research on Multi-Source Cooperative Operation Optimisation of Active Distribution Network. Master's Thesis, Shandong University, Jinan, China, 2021.
- Shenzhen Development and Reform Commission. *Notice on Issues Related to Further Improvement of Peak and Valley Time Division Electricity Tariff Policy in the Province*; Shenzhen Power Supply Bureau Limited: Shenzhen, China, 2021.
- Chen, P.; Meng, Q.; Zhao, Y. Electric vehicle charging load calculation based on Monte Carlo method. *J. Electr. Eng.* **2016**, *11*, 40–46.
- Yang, W.; Li, Y.; Wang, H.; Feng, J.; Yang, J. Combination Prediction Method of Electric Vehicle Charging Load Based on Monte Carlo Method and Neural Network. *J. Phys. Conf. Ser.* **2021**, *2022*, 012026. [\[CrossRef\]](#)

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