Transportation Research Part C 150 (2023) 104108

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Electric bus charging facility planning with uncertainties: Model formulation and algorithm design

Yu Zhou [a], Ghim Ping Ong [a], Qiang Meng [a,*], Haipeng Cui [b]

[a] Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576, Singapore
[b] College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Keywords:
Electric bus
Charging facility deployment
Two-stage stochastic programming
Multi-agent simulation
Reinforcement learning
Surrogate-based optimization

ABSTRACT

This paper investigates the electric bus charging facility planning (EB-CFP) problem for a bus transit company operating a heterogeneous electric bus (EB) fleet to provide public transportation services, taking into account uncertainties in both EB travel time and battery degradation. The goal of the EB-CFP problem is to determine the number and type of EB chargers that should be deployed at bus terminals and depots to meet daily EB charging demand while minimizing total cost. The problem is formulated as a two-stage stochastic programming model, with the first stage determining the EB charger deployment scheme and the second stage estimating the EB daily operational cost with respect to a predetermined EB trip timetable for a given EB charger deployment scheme. To effectively address the second stage problem, a multi-agent EB transit simulation system that mimics the daily EB operation process is developed. We then design two heuristic methods, the reinforcement learning (RL)-based method and the surrogate-based optimization (SBO), that use the developed multi-agent EB transit simulation system to solve the two-stage stochastic programming model for large-scale instances. Lastly, we run a series of experiments on a fictitious EB transit network and a real-world EB transit network in Singapore to evaluate the performance of the models and algorithms. In order to improve the performance of the EB transit system, some managerial insights are also provided to urban bus transit companies.

1. Introduction

Urban bus transit industry has been undergoing a revolution on buses by replacing the traditional internal combustion engine buses (ICEBs) such as diesel buses with electric buses (EBs) because of the environment-friendly features of EBs, including zero-emission and less noise. Take Singapore as an example, the Land Transport Authority is progressively deploying electric buses for public transportation services from 2020 with a target of 100% EB fleet by 2040 (LTA, 2019).

Compared with the traditional ICEBs, EBs have the short driving range issue caused by the limited capacity of the battery. Although the advanced battery technology has extended the theoretical driving range of one EB to about 280 km (BlueBus, 2021), which is comparable to a full-tanked diesel bus, the actual EB driving range is far less than the theoretical value because it is affected by many factors such as temperature, humidity, battery degradation (Pelletier et al., 2018), heating and ventilation and air conditioning loads. To maintain a cost-effective operational service with EBs suffering from the limited driving ranges, an urban bus transit company should frequently charge EBs in service at bus terminals or depots. Therefore, it is essential for an urban bus

* Corresponding author.
E-mail address: ceemq@nus.edu.sg (Q. Meng).

https://doi.org/10.1016/j.trc.2023.104108
Received 30 September 2022; Received in revised form 1 February 2023; Accepted 14 March 2023
Available online 23 March 2023

transit company to appropriately plan EB charging facility by considering various factors, which is referred to as the electric bus facility planning (EB-CFP) problem hereafter.

While the EB-CFP is a critical strategic decision, it should consider the operational decisions in an EB fleet serving various public transportation services to better estimate the total operating costs. This is because the EB daily operations determine the EB charging demand distribution. Moreover, the EB fleet operations for public transportation services are faced with two types of uncertainties — uncertain EB travel times caused by traffic congestion and stochastic EB battery capacity resulting from the charging/discharging process (Pelletier et al., 2018; Zhang et al., 2021b; Zhou et al., 2022a). Considering all these uncertainties, the state of charge (SOC) of EBs arriving at depots or terminals and the dwell time between two consecutive EB trips for recharging are also uncertain, leading to the uncertainty of EB charging demand. To generate a robust EB charging facility deployment plan, these uncertainties should be adequately addressed in the EB-CFP problem.

In this study, we put up an interesting EB-CFP problem for an urban bus transit company that aims to determine a cost-efficient charging facility deployment plan by considering uncertainties in both EB travel time and battery degradation, including the type and the number of different chargers deployed at either bus terminals or depots. Because the EB charging facility planning decision is directly affected by the EB charging demand distribution, we further consider the EB operation process in detail by optimizing the EB charging schedule. In addition, a few key practical issues are also addressed. The transformation from existing ICEBs to EBs will bring substantial impacts on the electric grid. In other words, the bus transit company may need to upgrade the power transmission system to expand the grid capacity. To do so, the bus transit company has to negotiate with power suppliers and invest in the grid capacity expansion.

We can formulate the EB-CFP problem under the uncertain environment as a two-stage stochastic programming model, in which the first stage determines the charging facility deployment solution, considering the possible grid capacity expansion. The second stage problem seeks the optimal EB operational schedule by considering the EB travel time uncertainty and battery degradation uncertainty for a given EB charger deployment solution. Due to the complexity of the second-stage problem, a multi-agent EB transit simulation system is designed to mimic the EB daily operations with the objective of minimizing the total operating cost. To solve the two-stage stochastic programming model, we propose two solution methods, namely, reinforcement learning (RL)-based method and the surrogate-based optimization (SBO) method. The RL technique has been widely used to solve the large-scale combinatorial optimization problems such as the vehicle routing problem (Zhang et al., 2020a), while the SBO methods can be utilized to solve the black-box problems or problems with computationally costly objective functions. The proposed EB-CFP problem is a combinatorial optimization problem, and the corresponding second-stage problem suffers from a computation burden. All these natures motivate us to seek different solution methods to solve the problem, and we further test the performance of these two methods. Specifically, the reinforcement learning (RL)-based method is designed based on the every-visit Monte Carlo (MC) method for learning the state-value function, where the state is the charger deployment results and the value is the total cost of two stages. Due to the large number of states, we cannot simulate all the possible situations. Therefore, we introduce a neural network-based value approximation method to approximate the state-value function. The surrogate-based optimization (SBO) method uses a radial basis function (RBF) to represent the relationship between charger deployment results and total cost. The RBF is iteratively updated by adding new sample points until the stop criteria are met. Finally, we carry out a series of experiments based on a tested EB transit network and a real-size EB transit network in Singapore to assess the efficiency and effectiveness of the research methodology proposed by this study.

The main contributions of this study are four-fold. First, we propose a EB-CFP problem with practical significance in an uncertain environment, which involves the charging facility deployment decision at the strategic level and the EB operations at the operational level. Issues like battery degradation uncertainty, multiple types of chargers and grid restriction are considered in EB operations to figure out a more reliable EB charging facility plan. Second, we develop a multi-agent EB transit simulation system to estimate the EB charging demand distribution by simulating the EB transit fleet operation process. Unlike most of the existing works that use the aggregated methods to roughly estimate the EB charging demand at terminals/depots, the developed multi-agent EB transit simulation system simulates the EB fleet operation process at a micro-level and captures the characteristics of the EB transit system, such as the queuing process of EB charging and minimum charging time requirement. Moreover, compared to existing studies that design algorithms to solve EB-CSP, the proposed agent-based simulation is able to find out good-quality EB charging schedules efficiently. Third, we formulate the problem as a two-stage stochastic programming model and propose two novel solution methods to solve it. The primary idea of these two solution methods is to find the relationship between an EB charging facility deployment solution and the total cost. The RL-based method uses MC method and a neural network-based value approximation method to exploit the state-value function. The SBO adopts an iterative way to update the RBF until stop criteria are met. Each of these two methods has its own characteristics and adaptability, and urban bus transit companies can choose them based on their needs. Finally, we perform experiments based on a test EB transit network and the bus network of Tower Transit in Singapore. We compare the performance of our proposed methods and some managerial insights are provided to urban bus transit companies to enhance the efficiency of EB systems.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the problem and presents the two-stage stochastic programming model. Section 4 elaborates on the multi-agent EB transit simulation system. In Section 5, the RL-based method and SBO method are proposed. To assess the applicability of the research methodology proposed by this study, Section 6 carries out a series of experiments based on a test EB transit network and a real EB transit network of Tower Transit Bus Corporation in Singapore, and the computational results are reported. Section 7 concludes this study and highlights the avenue for future research.

2. Literature review

Because the EB charging facility deployment results are highly related to the charging demand distribution determined by EB operations, we, therefore, review the existing works on EB-CFP and EB-CSP in this section to identify the research gaps to be filled and highlight the contributions of this study.

2.1. EB-CFP

Most of the existing studies on the charging facility deployment are referred to the electric vehicle charging facility location (EV-CFL) problems, which fall into the broad class of facility location problems (Mirchandani and Francis, 1990). Interested readers may refer to the study by Farahani et al. (2012) for more information. The EV-CFL problem intends to find out where to build the electric vehicle (EV) charging stations and determine the route of EVs to minimize the total cost.

Current studies on the EV-CFL problems can be generally classified into three categories according to the charging technologies: battery swapping studies (Mak et al., 2013; Hof et al., 2017; Yang et al., 2017; Masmoudi et al., 2018; An et al., 2020), plug-in charging studies (Liu and Wang, 2017; Yang, 2018; Yıldız et al., 2019; Iliopoulou et al., 2019; Lin et al., 2019; Wang et al., 2019; Xu et al., 2020; He et al., 2020; Zhou et al., 2022b) and wireless charging studies (Riemann et al., 2015; Liu and Song, 2017; Iliopoulou and Kepaptsoglou, 2021; Alwesabi et al., 2022). For the battery swapping technology, the depleted batteries of EVs can be replaced by full-charged batteries at battery swapping stations. This solution allows EVs to be effectively refueled within a short time (Mak et al., 2013) and they are adopted in the early stage of EV promotion. However, battery security risks and heavy construction and operating cost of battery swapping stations become major concerns (Adler and Mirchandani, 2014). These studies usually assume a fixed battery swapping time or a negligible swapping time, which could reduce the complexity of the corresponding models. The following is the plug-in charging technology, which is the most prevailing and practical technology. It dominates the EV charging market nowadays due to its relatively lower cost and fast charging technology. Therefore, various research topics have been investigated during the past few years, such as a tri-level programming to determine different types of charging infrastructure (Liu and Wang, 2017), a user choice model to locate fast charging stations (Yang, 2018), deploying charging station to mitigate the driving range anxiety of drivers (Xu et al., 2020) and a multi-day scenario-based approach to determine the charging facility deployment (Zhang et al., 2020b). Another newly emerging technology — wireless charging technology — has caught researchers' eyes. Wireless charging has the potential to eliminate the range limitation of EVs, which enables EVs to be utilized more flexibly and reduces the anxiety of drivers. A few scholars have started to adopt a stochastic user equilibrium model (Riemann et al., 2015) and robust optimization technique (Liu and Song, 2017; Iliopoulou and Kepaptsoglou, 2021; Alwesabi et al., 2022) to locate the wireless charging equipment and test the potential of wireless charging applications.

Although there is abundant literature focusing on the EV-FCL problems, they cannot be directed used for the EB-CFP problem due to the unique features of EB transit operations. For example, EBs for public transportation services need dedicated EB charging stations that can be located in bus depots or terminals from a practical perspective. EBs should go back to their affiliated depots after a one-day operation and they can be recharged at depots. In other words, it is a practice for a bus transit company to deploy EB chargers at bus terminals and depots. In addition, the operation of the EB transit system is significantly different from the EV system. The deployment results of the charging facility are generally based on the charging demand distribution in the system, which means a location with high charging demand will deploy more chargers. In the EV system, vehicles are operated by individual drivers who are highly unpredictable. Therefore, the charging demand of the EV system is usually estimated in an aggregated way for simplicity due to its complicated nature, for example, the study by Wang et al. (2019). For the EB transit system, the bus routes and timetables usually do not change in a certain period of time, and the operations of EBs are determined by bus companies, which is controllable. As a result, the charging demand for the EB transit system can be estimated in a more precise way by considering the characteristics of the EB transit system and EB operations. Based on the two reasons mentioned above, the EB-CFP problem is quite different from the existing EV-FCL problems.

In recent years, the EB-CFP problem has received attention and a few studies have been carried out. Xylia et al. (2017) first calibrate parameters related to EBs in Stockholm and build a mixed-integer linear programming (MILP) model to determine an optimal charging infrastructure plan. Lin et al. (2019) put forward a multi-stage spatial–temporal model to determine the location and sizes of EB charging stations in view of the continuously growing EB charging demand. An et al. (2020) establish a stochastic integer programming model to jointly optimize battery-swapping stations locations and bus fleet size under the stochastic bus charging demand. Hsu et al. (2021) formulate an optimization model to locate depots and charging facilities under the situation of a mixed fleet of buses. A decomposition-based heuristic algorithm is proposed to solve their model efficiently. Azadeh et al. (2022) develop a series of bi-objective optimization models to investigate the tradeoff between the locations of charging stations and battery sizing. The above studies can be regarded as extensions of the classical EV-FCL problem. Nonetheless, as we have analyzed before, it is not practical to find a location to construct a new charging station from scratch.

As a result, it may be beneficial to utilize the existing bus depot and bus terminals to build chargers, which enable EBs to be recharged during their dwell time when they arrive at terminals or depots. An (2020) simultaneously optimizes the bus fleet and charging facility deployment for an EB transit system considering the travel time uncertainty and time-of-use electricity tariff. The charging demand is treated in an aggregated way by assuming the charging demand is evenly distributed between two terminals of the bus routes. Hu et al. (2022) propose a joint optimization model to simultaneously determine the fast charger deployment and EB charging scheduling by considering the travel time and passenger demand uncertainty. Similarly, He et al. (2022) also designed an integrated approach to determine the fast charger deployment, battery size and charging scheduling at the same time. The above

studies only consider deploying fast chargers at bus terminals/stops. Nevertheless, part of the EBs will go back to depots after finishing the peak-hour tasks to be replenished in the daytime, and considering different types of chargers (e.g., normal chargers) at bus depots may be beneficial to PT operators. Moreover, the charging facility deployment and the charging scheduling are two different level problems (i.e., strategic level and operational level, respectively). One cannot simply combine these two problems as a single problem because the charging scheduling may be changed due to real-world uncertainties and the EB-to-trip assignment (which aims to assign each trip with the published timetable) is not fixed in the long run. These studies all assume a fixed EB-to-trip assignment and only optimize the charging decisions.

Besides, almost all of the existing studies consider a homogeneous bus fleet that has the same battery capacity. In reality, the EB battery will gradually degrade and lose its capacity due to solid electrolyte interface layer growth and active material loss (Pelletier et al., 2018), which happens during the cycling (charging and discharging) process. EBs will have different battery capacities as each EB is utilized in various routes and conditions. As a result, distinct capacities of EBs will affect the charging demand distribution, resulting in different charging facility deployment plan. Therefore, we should consider the battery degradation uncertainty to get a robust charging facility deployment plan.

Moreover, en-route fast chargers usually have a higher charging power which will bring tremendous pressure on the grid and it may be overloaded, especially for the peak-hour period when the EB charging demand is high (Kunith et al., 2017). For example, ABB, which is a pioneering technology manufacturer in the industry, provided 20 single-deck EBs and four 450 kW faster chargers in 2020 (SustainableBus, 2020). A 450 kW fast charger will put an amount of load which is comparable to the load created by 410 air conditioners of 1.5 horsepower at the same time. To reduce the EB charging load on the grid and maintain a stable environment of the power transmission system, the bus transit companies should ensure the total charging power for all fast chargers at the same terminal cannot exceed the maximum charger power. Besides, the bus transit companies can also expand the grid capacity at terminals by installing substations and transmission lines. An extra capacity expansion cost should also be paid to electricity suppliers (Karlsson, 2016). As these investments are very high, therefore the bus transit companies should wisely make their decision on whether to expand the grid capacity to enable more EBs can be charged at the same time or deploy more EBs to reduce the maximum electricity load. To the best of our knowledge, few EB-CFP-related studies have considered grid restriction.

From the literature review, a more effective approach by integrating EB charging facility deployment and charging scheduling should be proposed. Moreover, practical issues like grid expansion, battery degradation uncertainty and multiple types of chargers should be considered to figure out a more practical charging facility plan.

2.2. EB-CSP

The charging demand distribution in an EB network is determined by the EB-CSP, and this information forms the basis for planning the charging facilities. In general, the existing EB-CSP studies can be broadly divided into two categories: one is to solely optimize the charging schedule (including three decisions: when to charge, where to charge and how long to be charged) given the EB-to-trip assignment, while the other is to combine the EB-to-trip assignment with charging schedule optimization. For the EB-CSP with the given EB-to-trip assignment, each trip in the timetable has been allocated to an EB in advance. Given the EB-to-trip assignment, the complexity of the EB-CSP should be reduced greatly, allowing the large-scale problem to be solved. For instance, He et al. (2020) succeeded in solving an instance on a bus network with 36 routes consisting of thousands of trips. Another benefit is that some complex practical restrictions can be added or optimized in conjunction while keeping the problem's tractability. Liu et al. (2021) considered some resource constraints and further extended the problem into a robust version (Liu et al., 2022). In addition, integrated optimization with charger deployment is investigated by Zhou et al. (2022b), Hu et al. (2022), He et al. (2022). However, different trip assignment results may exist for the same timetable, resulting in different optimum charging schedules. As a consequence, if the EB-to-trip assignment and charging schedule can be optimized at the same time, a more cost-effective solution may be achieved.

The EB-CSP combined with EB-to-trip assignment is similar to the well-studied electric vehicle scheduling problem (EVSP) which aims to cover a set of given trip tasks using a fleet of EVs while considering possible charging activities by minimizing the total costs. The study by Sassi and Oulamara (2014) indicates that even a single-depot EVSP is NP-hard. Moreover, the EB-CSP may incur various real-world concerns, increasing the complexity of the problem. In recent years, many scholars have tried to develop different algorithms to solve the EB-CSP efficiently. Table 1 summarizes the representative works on the EB-CSP in recent years. It can be seen that, the major algorithm to solve the problem is branch & price. Some heuristics are also developed to solve large-scale instances. However, the largest size instance which can be solved within an acceptable gap usually involves less than 1000 trips and the computational time is generally over two minutes. In the real application, a bus network may involve thousands of trips. These algorithms cannot be directly adopted into our problem to determine the charging demand distribution. As we have discussed before, the EB-CSP may be faced with battery degradation and travel time uncertainty. As a result, the EB-CSP should be solved multiple times to generate reliable charging demand distributions. It puts high requirements on the computational speed to solve the EB-CSP.

On the other hand, the good news is that the ultimate goal of this study is to determine the tangible charging facility deployment solution in both bus terminals and depots instead of finding exact algorithms to solve the EB-CSP. Therefore, simulation is an efficient tool to mimic the operation of a complicated system and has been widely adopted in the transportation field, such as the large-scale public transport network design problem (Becker et al., 2020), autonomous ride-pooling problem (Zwick et al., 2021), traffic assignment problem (Zhang et al., 2021a) and freight transportation scheduling (Layeb et al., 2018). Several studies also try to use simulation to investigate EB operations. Sebastiani et al. (2016) present a discrete event simulation to simulate a real-world

Table 1
Comparison of existing studies in EB-CSP.

Authors	Extra consideration	Solution method	Largest instance	CPU time
Li (2014)	Battery swapping	Column generation + ALNS	947 trips	20 h
Adler and Mirchandani (2017)	\	BP+heuristics	4373 trips	290 s
Li et al. (2019)	External cost	Solver	144 trips	Not reported
Tang et al. (2019)	Stochastic	BP	96 trips	117 s
Zhang et al. (2021b)	Nonlinear charging profile and battery degradation	BP	160 trips	3253 s
Liu and Ceder (2020)	Multi-objective	Heuristics	579 trips	Not reported
Wu et al. (2022)	Grid characteristics	BP	400 trips	Over 24 h

BP – branch & price; ALNS – adaptive large neighborhood search.

EB system considering the charging activities. The simulation results are used to evaluate the performance of different charging strategies and charger deployment. Jefferies and Göhlich (2020) also design a simulation tool to determine the vehicle, energy and staff demand required to realize the full electrification of an EB network. The simulation tool exhibits good performance in terms of both simulation speed and quality.

The literature review in this section shows that the EB-CSP is a challenging problem and solving the large-scale problem requires huge computational resources, even using heuristics to solve the problem with an acceptable gap. Simulation has the potential to solve the EB-CSP efficiently to obtain acceptable results in a short time.

3. Problem description and model formulation

3.1. Problem description

Consider an urban bus transit company operating a non-homogeneous EB fleet in terms of EB battery capacity, denoted by set $\mathcal{B} = \{b|b = 1, 2, \dots, B\}$, to serve a number of bus routes with the daily timetables published to the public. A bus route comprises a series of stops, including the start and end bus terminals. A daily timetable for a bus route includes the start time at one terminal and end time at another terminal of each bus trip (for the loop service, the departure terminal is the same as the arrival terminal). A bus is parked at a bus depot after its daily services, and it will depart from the bus depot to a terminal to start its services scheduled for the next day. We assume that EBs leave their depots with fully-charged batteries to start their services according to a bus timetable published to the public and return to the depots after their daily operations. EBs can be recharged at either bus terminals or depots after they complete one or several trips with two types of chargers - normal and fast chargers - denoted by set $\mathcal{K} = \{\text{fast}, \text{normal}\}$.

Let $\mathcal{W} = \{w|w = 1, 2, \dots, W\}$ and $\mathcal{S} = \{s|s = 1, 2, \dots, S\}$ denote the set of bus depots and the set of bus terminals, respectively. All bus trips with respect to the daily timetables are grouped into set $\mathcal{V}$. A trip $v \in \mathcal{V}$ is described by a six-tuple $\left(ter_v^d, ter_v^a, \tilde{t}_v^d, \tilde{t}_v^a, \tilde{t}_v, \tilde{\Delta}_v\right)$, where ter_v^d and ter_v^a are the departure and arrival terminals of the trip, $\tilde{t}_v^d$ is the departure time of the trip from terminal ter_v^d, $\tilde{t}_v^a$ is the arrival time of the trip at terminal ter_v^a, $\tilde{t}_v$ is the trip travel time that equals to $\tilde{t}_v^a - \tilde{t}_v^d$, and $\tilde{\Delta}_v$ is the energy consumed to complete the trip task. In practice, the travel time of a bus trip is affected by many factors such as weather, traffic volume and road condition. We assume the travel time $\tilde{t}_v$ of trip $v \in \mathcal{V}$ follows a certain distribution $\mathbb{P}_1$ with a probability distribution function $f(v)$ on an interval $[t_v^{\min}, t_v^{\max}]$. We further assume that energy consumption $\tilde{\Delta}_v$ is proportional to the travel time. As a result, $\tilde{\Delta}_v$ will also follow the same type of distribution as $\tilde{t}_v$.

In addition, the EB battery gradually loses its capacity and this process is also affected by the factors such as temperature and current energy state. Furthermore, each EB is used at different frequencies. All these factors will lead to different capacities of batteries. We denote $\tilde{cap}_b$ by the battery capacity of EB $b \in \mathcal{B}$ and further assume it follows a certain distribution $\mathbb{P}_2$ with a probability distribution function $g(cap_b)$. The distribution function can be estimated from the historical data. In general, the capacity of battery drops to its 80% of the health condition, the battery ends its life (Lam and Bauer, 2012). Let $cap_b^{\max}$ be the capacity of battery's health condition, and the battery capacity $\tilde{cap}_b$ must fall in an interval $[0.8cap_b^{\max}, cap_b^{\max}]$.

The deployment of EB chargers will also bring substantial impacts on the electric grid. If the grid is overloaded, the voltage deviation will increase and the electricity transmission system may be damaged, resulting in the transmission efficiency decline. To ensure that the grid can handle such a large EB charging demand, the urban bus transit company needs to build substations and transmission lines with the cost of β SGD/100 kVA. After grid expansion, the grid capacity for terminal $s \in \mathcal{S}$ (depot $w \in \mathcal{W}$) is denoted by $L_s(L_w)$.

The EB-CFP problem concerned by the company needs to determine the number and type of chargers deployed at each bus terminal and depot with the objective of minimizing total cost while maintaining the published bus timetable by considering some realistic constraints, including travel time uncertainty, battery degradation uncertainty, multiple types of chargers and grid restriction. The consideration of these uncertainties will enable the company to figure out a more robust EB charger deployment plan, although it will bring great challenges to the model building and algorithm design for solving the EB-CFP problem. For the ease of readers, all the notations are listed in Appendix A.

3.2. Two-stage stochastic programming model

The involvement of travel time uncertainty and battery degradation uncertainty motivates us to formulate the EB-CFP problem as a two-stage stochastic programming model, which has been widely adopted to solve the problems with uncertainties. The first stage of the EB-CFP problem is to figure out the charging facility deployment solution, which includes the number and type of chargers in each terminal and depot, as well as the grid capacity expansion. Once the charging facility deployment solution is determined, the second stage problem seeks the optimal schedule of EBs by minimizing the total operating cost under travel time uncertainty and battery degradation uncertainty.

To formally set up the mathematical model, in the following we refer to boldfaced characters $\mathbf{x} \in \mathbb{Z}^{(W+S)\times 2}$ to matrix, and $x_{i,j}$ to the element of i-th row, j-th column. Matrices $\mathbf{y} \in \mathbb{Z}^{(B\times V\times V)}$, $\mathbf{z} \in \mathbb{R}^{(B\times V\times V)}$ and $\mathbf{T} \in \mathbb{R}^{(B\times V\times V)}$ are defined likewise.

The decision vectors to formulate the model are as follows: **x** is the vector of first-stage decisions to determine the charger number and type at each depot/terminal. **y, z** and **T** are the second-stage decision vectors to determine the EB charging scheduling. The detail of these decision vectors will be explained later. To this end, the two-stage stochastic programming problem for the EB-CFP problem can be formulated by

[2-EB-CFP]

$$\min_{\mathbf{x}\in X}\{G(\mathbf{x}) + \mathbb{E}_{\mathbb{Q}}[Q(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{T}, \boldsymbol{\xi})]\} \tag{1}$$

where $\mathbf{x}$ is the vector of first-stage decisions and $x_{i,j}$ denotes the number of the charger of type j deployed at terminal/depot i. $G(\mathbf{x})$ is the first-stage cost, including the charger construction cost and grid capacity expansion cost (if any), and it can be expressed by

$$G(\mathbf{x}) = \sum_{i\in\mathcal{W}\cup S}\Big(\sum_{j\in\mathcal{K}} C_j x_{i,j} + \beta\Big(\sum_{j\in\mathcal{K}} P_j x_{i,j} - F_i\Big)^+\Big) \tag{2}$$

where C_j and P_j denote the construction cost and maximum charging power for charger of type j, respectively. F_i is the current grid capacity of terminal/depot i. It is noted that $(\sum_{j\in\mathcal{K}} P_j x_{i,j} - F_i)^+ = \max\{\sum_{j\in\mathcal{K}} P_j x_{i,j} - F_i, 0\}$. X is a polyhedral set defined by

$$\sum_{j\in\mathcal{K}} x_{i,j} \le u_i, \forall i \in \mathcal{W}\cup S \tag{3}$$

Constraint (3) imposes that the number of chargers deployed at depot or terminal i should not exceed its parking capacity u_i limited by the number of parking lots.

In the model [2-EB-CFP], **y, z** and **T** are the second-stage decision variables. $\boldsymbol{\xi} = (\mathbf{t}, \mathbf{cap})$ denotes the travel time and battery capacity information and **x** is the charger deployment result given by the first stage problem, which is known when the first-stage problem is determined. $Q(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{T}, \boldsymbol{\xi})$ is the optimal value of the second-stage problem, and the second-stage problem aims to simultaneously determine the EB-to-trip assignment and EB charging scheduling given the charger deployment solution of the first-stage problem by minimizing the total operating cost. Decision matrix **y** is a binary matrix used to indicate the EB-to-trip assignment relationship. Decision matrices **z** and **T** are real matrices and they determine the charging process of EBs. Specifically, **z** and **T** record the energy state and charging time information of EBs, respectively. Note that $\boldsymbol{\xi}$ is one of the realizations of travel time and battery capacity under the distribution $\mathbb{Q}$, which is the joint distribution of the travel time distribution $\mathbb{P}_1$ and the battery capacity distribution $\mathbb{P}_2$. In general, the travel time on the road is independent of the battery degradation and we can take the Cartesian product to get $\mathbb{Q}$ for simplicity. The detailed formulation of the second-stage problem is given in Appendix B.

The second-stage problem is NP-hard in nature and hard to be solved exactly when the problem size becomes large. The NP-hardness can be easily proved by removing constraints (20)–(29) and the remaining problem becomes a multi-depot vehicle scheduling problem (MDVSP), which has been proved to be NP-hard (Dell'Amico et al., 1993). Moreover, some realistic considerations are hard to be formulated explicitly in an explicit mathematical model. Furthermore, the uncertainties make the problem become much more intricate. In view of these characteristics of the second-stage problem, a multi-agent EB transit simulation system can be developed to simulate the rule-based daily EB scheduling process. Agent-based simulation system can easily incorporate realistic considerations in the EB operation process by ruling the activities of the agents. Furthermore, the randomness resulting from the EB transit times can be efficiently handled and parallel computation can be applied to simulate different scenarios at the same time, enhancing the solution speed. The detailed information about the proposed agent-based simulation is given in Section 4.

In addition to the complexity of the second-stage problem, the first-stage problem is an integer programming problem and its solution space increases exponentially to the number of terminals/depots. When the problem size becomes large, it is impossible to enumerate all possible solutions, which makes the proposed model [2-EB-CFP] hard to solve effectively. To circumvent these difficulties, we propose two different heuristic solution methods, namely, the RL-based method and the SBO method. RL is a popular method that has emerged in recent years and has been widely applied to solve combinatorial optimization problems, while SBO is a surrogate optimization method to solve black-box problems or problems with computationally costly objective functions. The natures of the proposed model enable us to consider these two different methods to solve this problem and we further want to make a comparison between the emerging method and the traditional method. These two methods are elaborated in detail in Section 5.

4. Multi-agent EB transit simulation system

Because the agent-based simulation allows us to carry out a variety of virtual experiments to gain a precise understanding of the complex system and possible outputs, some realistic constraints which hard to be tackled in analytical optimization models can be considered in the optimization process, e.g., the queuing process of EBs when chargers are occupied. Furthermore, the real-world uncertainties can also be well incorporated into the simulation. Therefore, we design a multi-agent EB transit simulation system to solve the second-stage problem of the two-stage stochastic programming model.

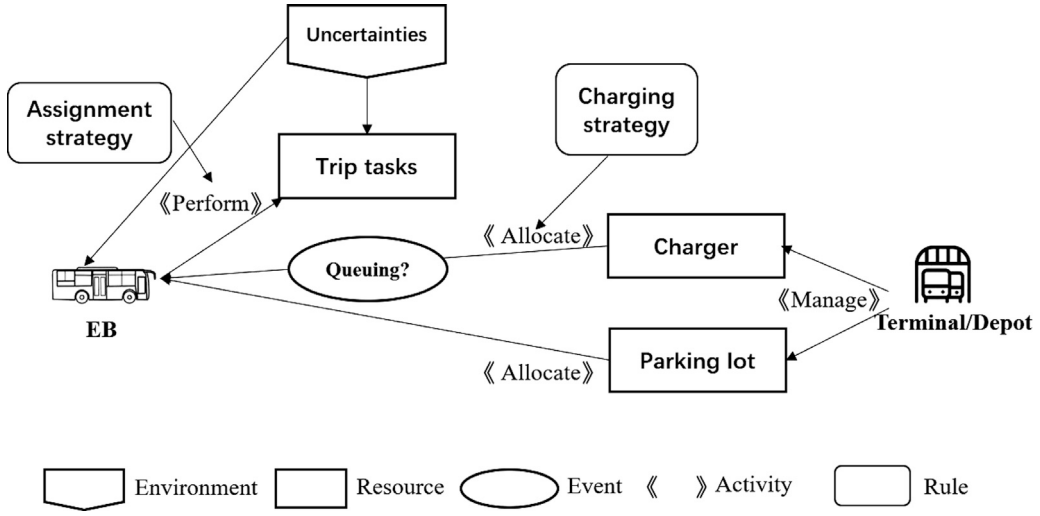

Fig. 1. Interactions between different agents.

4.1. Agents

A multi-agent simulation system consists of agents that interact within a system. The main agents in our simulation are the EB agents and the terminal/depot agents, which are explained below.

EB agents: each of EB agents simulates the behavior of an EB driver. The agents aim to finish all the trip tasks required by the published timetable by considering their current locations and current SOC state. Different EB-to-trip assignment strategies can be considered for each EB agent. Besides, each of these agents also needs to consider its charging activity to ensure a good SOC state to finish the following trips. Both the charging location and the charging duration should be determined.

Terminal/depot agents: each of the terminal/depot agent manages a series of EB chargers and parking lots, and controls which EB has access to the charger at each time slot. The agents are responsible for allocating chargers and parking lots to EBs. If all the chargers are occupied by EBs, the waiting queue should be managed by the terminal/depot agent. Different allocating strategies can be considered for each terminal/depot agent.

To better explain the relationship between the EB agent and the terminal/depot agent, Fig. 1 illustrates the interactions between different agents in the bus transit system. Firstly, the uncertainties act on the trip tasks and EBs. Specifically, travel time uncertainty will affect the travel time and energy consumption in each trip and battery degradation uncertainty will influence the battery capacity of each EB. Then the EBs start to perform the trip tasks based on the assignment strategy. At the same time, the terminal/depot needs to manage both the chargers and parking lots. In general, the charger is installed at the parking lot, which means the number of chargers is also limited by the parking lots. Based on the charging strategy, the terminal/depot allocates chargers to EBs that need to be replenished. Meanwhile, the queuing should be considered in there are no available chargers. The objective is to fulfill the given timetable with the aim to minimize the total costs, which include the fixed cost related to the EB fleet and charging cost.

4.2. EB-to-trip assignment

The EB-to-trip assignment aims to assign proper EBs to complete the given timetable, and the location of EB, current energy state, and availability of chargers should be considered. The basic principles of EB-to-trip assignment are presented as follows:

- The current energy of EB should be greater than the summation of energy consumption for a certain trip and the minimum energy that an EB should maintain.
- The current location of an EB should be consistent with the departure terminal of the connected trip.
- If an EB is undergoing charging and the charging duration is longer than the minimum requirement $t_{charge}^{\min}$, it can be used to perform another trip task.
- If there exist multiple eligible EBs at time $t = \tilde{t}_v^d$ for trip v, then the highest-SOC-first (HSF) strategy is applied to select the EB to carry out trip v. It means that we select the EB that has the highest SOC state among all eligible EBs to perform trip v. Else if no such eligible EBs, then a new EB is generated to perform this trip v.

The above principles can be depicted by the following pseudo code:

Algorithm 1: EB-to-trip assignment

```
Initialization B=∅ and V
for t=starttime to endtime do
    for v ∈ V do
        if t_v^d = t then
            if there exist eligible EBs in B then
                HSF strategy is applied to select an EB b ∈ B to carry out trip v
            end
            else if there are no such EBs in B then
                Generate a new EB to carry out trip v and add b in set B
            end
        end
    end
end
```

4.3. EB charging principles

When the SOC state of an EB is below a certain threshold, the EB should be charged at a terminal or depot to maintain a good state of SOC. The charging process should consider the minimum charging duration and charger availability. We summarize the charging principles as follows:

- When an EB is allocated to a charger to be charged, a minimum charging time $t_{charge}^{\min}$ should be met.
- If all the chargers in a terminal/depot are occupied, the EB has to wait in the queue until one charger is released.
- If multiple EBs are waiting in the queue to be charged, once a charger is released, the first-in-first-served (FIFS) strategy is adopted to select the EB to be charged.
- In order to maintain a good health state of EB battery, the SOC state of an EB should be within the range $[SOC^{\min}, SOC^{\max}]$.

The above principles can be depicted by the following pseudo code:

Algorithm 2: Charger assignment

```
for t=starttime to endtime do
    for b ∈ B do
        if b is in "waiting" state then
            Check availability of all chargers in current location of b
            if there is a charger available then
                Check the queue of b and find the EB by FIFS to be charged
            end
        end
        else if b is in "empty" state and current SOC state is less than SOC^max then
            Check availability of all chargers in current location of b
            if there is a charger available then
                Assign this charger to b and change the state to "charging"
            else
                Add b in the queue and change the state to "waiting"
            end
        end
    end
end
```

It is noted that each EB in our simulation will have 4 different states: empty, carrying our trip, waiting and charging.

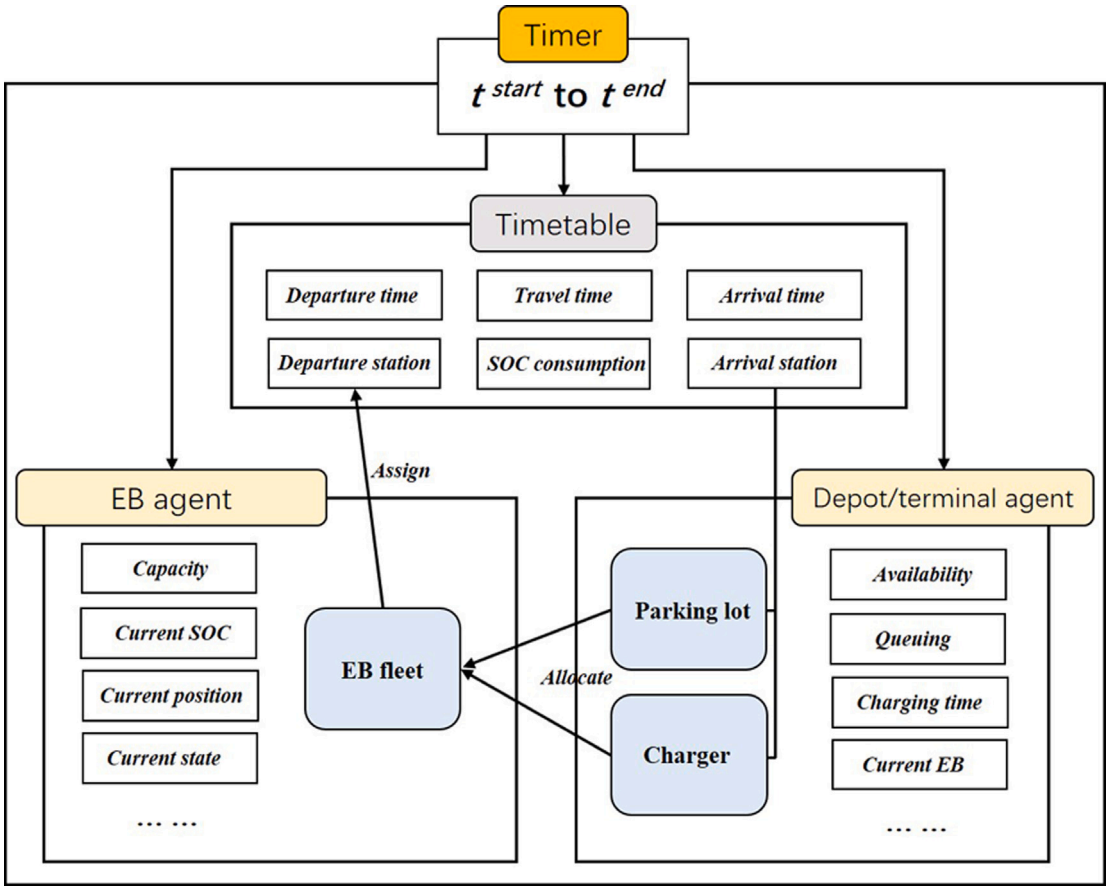

Fig. 2. Data processing of the computerized system.

4.4. Uncertainties

The travel time uncertainty and battery degradation uncertainty have been considered in our simulation. The uncertainty will act on the trip tasks, and further influence the energy consumption for each trip. Given the distribution of travel time, the timetable incorporated uncertainty can be generated in advance. The battery degradation uncertainty mainly affects the battery capacity of EBs. Given the distribution of battery degradation, the battery capacity of each EB can be determined when a new EB is generated to carry out a trip task.

4.5. Computerized system

Using these two types of agents and the proposed EB-to-trip assignment and charging principles, we have built a multi-agent simulation system. Next, we describe the processing and data information in the computerized system, as shown in Fig. 2. A timer is used to record the current simulation time, which is in the range of $[t^{start}, t^{end}]$. Based on the current time, the trips in the given timetable are checked. When a trip should be performed at the current time, the EB agent will assign an appropriate EB to this trip based on the EB-to-trip assignment principle. When an EB completes a certain trip and arrives at its destination, the depot/terminal agent will allocate a parking lot and a charger based on the current state of the EB and the depot/terminal. The entire process ends when the current time reaches the end time of operating period. To demonstrate the efficiency and effectiveness of the simulation system, a simple analysis is provided in Appendix C.

5. Two heuristic solution methods

In this section, we propose two heuristic solution methods to solve the two-stage stochastic programming model. The first one is a RL-based method elaborated in Section 5.1, and the second one is a SBO method presented in Section 5.2.

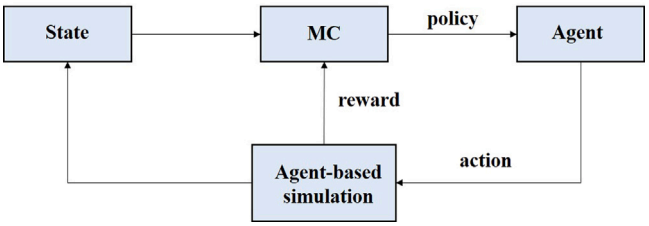

Fig. 3. Framework of the reinforcement learning method.

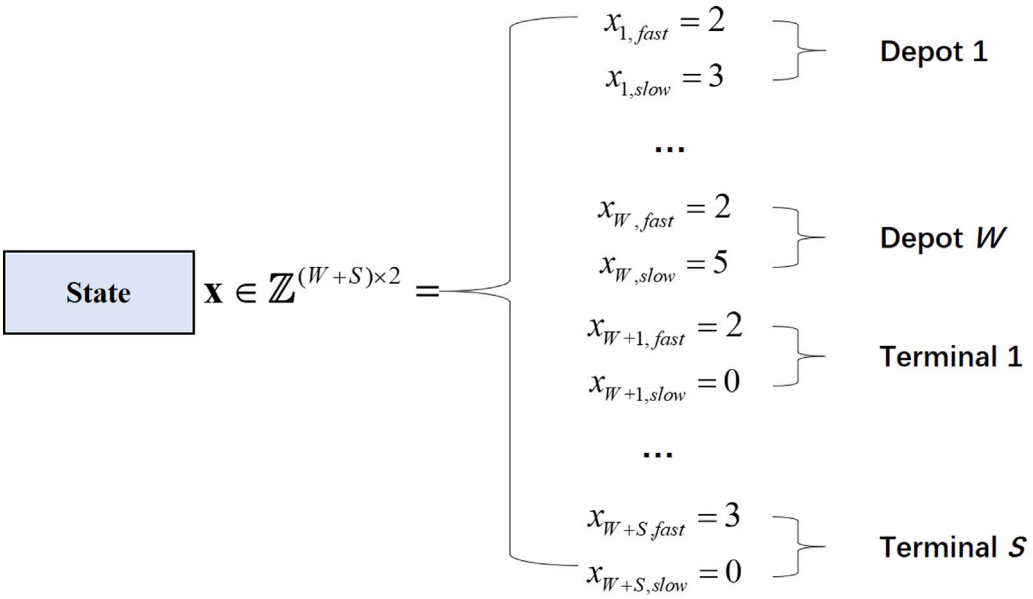

Fig. 4. An example of state representation.

5.1. Reinforcement learning-based method

As shown in Fig. 3, the RL-based method is proposed based on an every-visit MC for learning the state-value function for a given policy, and it consists of a set of finite states $l \in \mathcal{L}$, a set of possible actions $a \in \mathcal{A}$ under a given policy π, a scalar value function $v_\pi(l)$ (Sutton and Barto, 2018). In the two-stage stochastic programming model, the charger deployment schemes involved in the first-stage problem can be regarded as states. For a given state, the multi-agent EB transit system simulates the second-stage decision process to obtain the value function.

5.1.1. State

The state can be represented by the number and type of chargers in each terminal and depot, which is equivalent to the first-stage decision vector $\mathbf{x} \in \mathbb{Z}^{(W+S)\times 2}$. Fig. 4 gives an example of state representation. With a slight abuse of notation, for each terminal/depot $i \in \mathcal{W} \cup \mathcal{S}$, the total deployed number of chargers should meet the requirements by constraints (3). Because we only consider two types of chargers, namely fast charger and slow charger, we give an example of state representation in Fig. 4.

5.1.2. Policy and action

In our every-visit MC, we adopt a random policy to move from a state to an adjacent state. The action set $\mathcal{A}$ in the policy is formulated by perturbing the decision vector of the first stage. We define a^t as the action at time step t, and the following equations define all possible actions

$$a^t = \begin{cases} x^t_{i,fast} + \delta^t_{i,fast} & ,\forall i \in \mathcal{W} \cup \mathcal{S} \\ x^t_{i,slow} + \delta^t_{i,slow} & ,\forall i \in \mathcal{W} \cup \mathcal{S} \end{cases} \tag{4}$$

where decision variables increase or decrease by the value of δ, which can be determined based on the size of the problem. Small δ will increase the accuracy of the algorithm but incur a large computation burden. To speed up the solution process, δ can be increased. It is noted that δ is an integer number. We assume that δ can be any value from $\{-1, 0, +1\}$ to guarantee the accuracy of our algorithm.

5.1.3. Value function

Because the MC aims to find out the largest value-function and our two-stage stochastic model seeks the minimum total costs, we set the value function as the negative total costs under state $l \in \mathcal{L}$, namely the negative objective function, which can be represented by

$$v_\pi(l) = -\sum_{i \in \mathcal{W} \cup \mathcal{S}} \Big(\sum_{j \in \mathcal{K}} C_j x_{i,j}(l) + \beta \big(\sum_{j \in \mathcal{K}} P_j x_{i,j} - F_i \big)^+ \Big) - \mathbb{E}_{\hat{\mathbb{Q}}}[Q(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{T}, \boldsymbol{\xi})] \tag{5}$$

where l is a state corresponding to a solution vector $\mathbf{x}(l) \in \mathbb{Z}^{(W+S)\times 2}$, and $\hat{\mathbb{Q}}$ is the empirical distribution and we simulate it in the agent-based simulation.

5.1.4. State-value update

After defining the state, policy and value-function, we now explain the state-value update process in our reinforcement learning algorithm. The every-visit MC wishes to estimate $v_\pi(l)$ under the random policy π, given a set of episodes obtained by following π and passing through state l. Each occurrence of state l in an episode is called a visit to l. Every time when there is a visit to state l in MC, we record it and average the value-function of all visits to l. It is noted that the every-visit MC converges to $v_\pi(l)$ as the number of visits to l goes to infinity (Sutton and Barto, 2018).

Algorithm 3 gives the pseudocode of the proposed algorithm based on reinforcement learning. First, we can initialize all $v_\pi(l)$ as 0 and we create a list $\text{list}(l)$ for each state l to record the state-value for each visit to state l. Then we repeat each episode and each time step. In each step, we randomly select an action a in action set $\mathcal{A}$. After taking the selected action a, a new state l' is arrived. The agent-based simulation is used to evaluate the state-value under state l' and get the value G. Next, value G is appended to the $\text{list}(l)$ and update the average value in this list to $v_\pi(l')$. This process continues until all the episodes have been done. It is noted that, in our algorithm, each episode consists of a constant number of time steps in $\mathcal{T} = \{1, 2, \ldots, t^{end}\}$.

Algorithm 3:

Initialize
$\pi \leftarrow$ random policy
$v_\pi(l) \leftarrow 0$ for all l
$\Delta \leftarrow$ stop criteria
$\text{list}(l) \leftarrow$ an empty list for all $l \in \mathcal{L}$
for each episode **do**
  Random an initial state l corresponds to $\mathbf{x}$
  for time step $t \in \mathcal{T}$ **do**
    Choose an action $a \in \mathcal{A}$ under policy π
    Take action a and move the adjacent state l'
    Calculate the value through agent-based simulation and obtain value G
    Append G to $\text{list}(l')$
    $v_\pi(l') \leftarrow \text{average}(\text{list}(l'))$ until $|v_\pi(l') - v_\pi(l)|/|v_\pi(l)| \le \Delta$
  end
end

5.1.5. Neural network-based value approximation

We have so far assumed that our estimation of value functions is represented as a table with one entry for each state. This assumption is applicable when the problem size is small with small numbers of states. When the size of the charging facility deployment problem increases, the number of states increases exponentially, requiring large memory to record the table. In addition, the time and data needed to fill the table accurately are also increased dramatically. For a depot or terminal, the number of possible states of charger deployment is

$$\sum_{p=0}^{p=u_i} (p+1) = \frac{(u_i+2)(u_i+1)}{2}, \forall i \in \mathcal{W} \cup \mathcal{S} \tag{6}$$

As a result, the total number of states in the state set $\mathcal{L}$ should be

$$|\mathcal{L}| = \prod_{i=1}^{i=W+S} \frac{(u_i+2)(u_i+1)}{2}. \tag{7}$$

which is a large number when there are many depots, terminals or parking lots on the EB network. To tackle this issue, we consider using a neural network to depict the relationship between the state and its corresponding value. The basic idea is that we assume that there is an unknown underlying function that is consistent in mapping the state to the total cost value. Given a dataset comprised

of inputs and outputs, we can use supervised learning to train the neural network to fit such a function. As a result, the value is represented not as a table but as a neuron network with input parameter vectors $\mathbf{e} \in \mathbb{R}^{(W+S)\times 2}$.

Algorithm 4 gives the pseudocode of the proposed algorithm to get the approximated function.

Algorithm 4:

Initialize
$\pi \leftarrow$ random policy
$\mathcal{D} \leftarrow \boldsymbol{\Phi}$
for each episode **do**
  Random an initial state l corresponds to $\mathbf{x}$
  for time step $t \in \mathcal{T}$ **do**
    Choose an action $a \in \mathcal{A}$ under policy π
    Take action a and observe the value V_t
    Record state l^t and its corresponding value V_t to dataset $\mathcal{D}$
  end
end
Train the neural network with dataset $\mathcal{D}$

5.2. Surrogate-based optimization method

The SBO method has been widely applied to tackle the black-box problem or problem with computationally costly objective functions (Jones et al., 1998). Interested readers may refer to the review article by Forrester and Keane (2009) for different applications of SBO. The basic idea is to use a surrogate function that characterizes the relationship between the inputs of decision variables and simulation outputs (Jakobsson et al., 2010). In our two-stage stochastic model, the input is the first stage's charging facility deployment and the output is the total costs of the bus system. As our formulation in Section 3.2, the objective function is difficult to calculate analytically. It contains two stages' decisions and an expectation mean value, which requires multiple simulations for each charging facility deployment input. These properties of our problem enable us to adopt the SBO to find the solutions.

5.2.1. Surrogate function

The surrogate function is obtained by a series of initial sample points, which can be figured out by the design of experiments (e.g., latin hypercube sampling (LHS)). Then an iterative process is design to enhance the "closeness" of the surrogate function to the original objective function until some convergence criteria are met. Denote these sample points by $\mathbb{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, each sample point $\mathbf{x} \in \mathbb{Z}^{(W+S)\times 2}$ refers to a charging facility deployment result. The corresponding objective function values at these points are represented by $y_1, y_2, \dots, y_n$. It is noted that each objective value is obtained by taking the average of multiple computations and the computation times can be determined by the optimal computing budget allocation (OCBA) by Chen and Lee (2011). We further use the RBF model as our surrogate model, which can be represented by:

$$s_f(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i \phi(\|\mathbf{x} - \mathbf{x}_i\|) + p(\mathbf{x}) \tag{8}$$

where $s_f(\mathbf{x})$ is the surrogate function, and $\phi(\|\mathbf{x} - \mathbf{x}_i\|)$ is the RBF function. This RBF measures the distance between $\mathbf{x}$ and each sample point $\mathbf{x}_i$. It assumes that the correlation of every two sample points only depends on the distance (defined by the norm $\|\|$). Here, we use a cubic RBF $\phi(x) = x^3$. $p(\mathbf{x}) = \mathbf{b}^T\mathbf{x} + a$ denotes the polynomial tail with $\mathbf{b} = [b_1, b_2, \dots, b_k]$, where $k = W + S$. All these parameters λ, $\mathbf{b}$ and a can be calculated by solving the following linear equations:

$$\begin{bmatrix} \boldsymbol{\Phi} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix} \tag{9}$$

where $\boldsymbol{\Phi}$ denotes the matrix where $\Phi_{ij} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|)$, $\mathbf{0}$ denotes the zero matrix, and

$$\mathbf{P} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{1} \\ \mathbf{x}_2^T & \mathbf{1} \\ \vdots & \vdots \\ \mathbf{x}_n^T & \mathbf{1} \end{bmatrix}, \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}, \mathbf{c} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \\ a \end{bmatrix}, \mathbf{F} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{10}$$

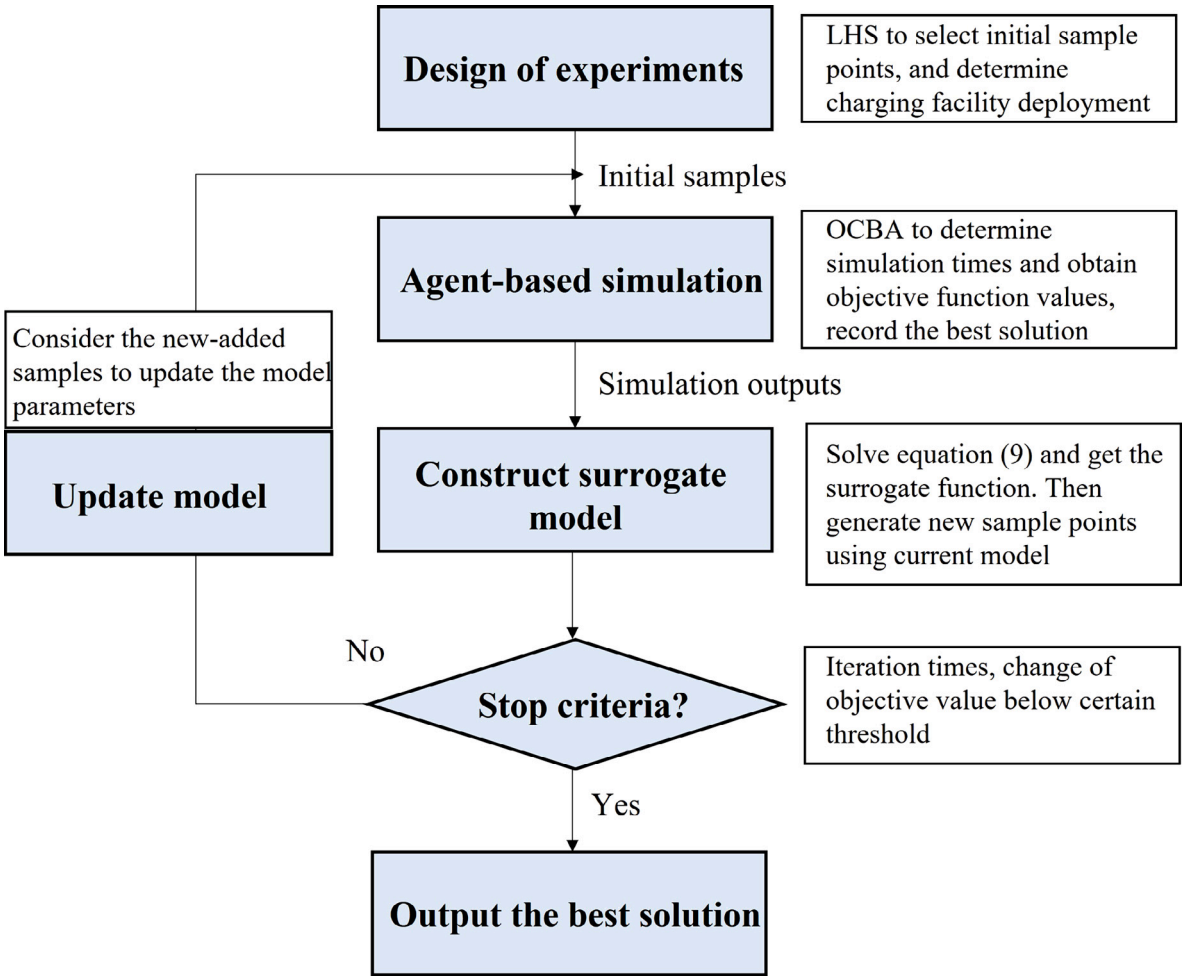

Fig. 5. Flowchart of the SBO method.

5.2.2. Procedure of the SBO method

The detailed SBO procedure for solving the EB-CFP problem is shown in Fig. 5. The most important part of the SBO method is to generate and select new sample points based on the current surrogate model. To elaborate on the sample point generation process, let $\mathbf{x}_{\min}$ be the solutions with the minimum objective value in all sample points in the last iteration. In our algorithm, we design two strategies to generate candidate sample points: (1) based on $\mathbf{x}_{\min}$, randomly adding or subtracting different scales of perturbations (e.g., small, medium and large) to generate new candidate sample points; (2) randomly generating new sample points from all sample space. The first strategy aims to find candidate sample points in the vicinity of the optimal solution $\mathbf{x}_{\min}$ in the last iteration (i.e., local search), and the second strategy seeks the candidate sample points globally (i.e., global search).

After new candidate sample points are obtained, we now move on to select the most prominent points in these two groups. Two types of scoring criteria are used. The first criterion is based on the objective value calculated by the surrogate function (8). Let $\mathbf{x}_l^{new}$ be the lth point, $l = 1, 2, \ldots, m$, in any of these two groups. Denote $s_f^{\max}$ and $s_f^{\min}$ the maximum objective value and the minimum objective value in the candidate points, respectively. The score of the lth candidate point can be calculated as

$$S_O(\mathbf{x}_l^{new}) = \frac{s_f(\mathbf{x}_l^{new}) - s_f^{min}}{s_f^{max} - s_f^{min}}, \quad (s_f^{max} \neq s_f^{min}) \tag{11}$$

Another criterion is based on the distance of the newly generated candidate points to the set of already sampled points $\mathbb{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$. Let $\Delta(\mathbf{x}_l^{new})$ be the minimum Euclidean distance between the lth new generated point in any of the aforementioned groups and all the already sampled point in set $\mathbb{X}$. We further denote $\Delta^{\max} = \max_{l=1,2,\ldots,m}\Delta(\mathbf{x}_l^{new})$ and $\Delta^{\min} = \min_{l=1,2,\ldots,m}\Delta(\mathbf{x}_l^{new})$, i.e., the maximum and the minimum of the distances of all new generated candidate points to $\mathbb{X}$. The score of the lth candidate point can be calculated as

$$S_D(\mathbf{x}_l^{new}) = \frac{\Delta^{\max} - \Delta(\mathbf{x}_l^{new})}{\Delta^{\max} - \Delta^{\min}}, \quad (\Delta^{\max} \neq \Delta^{\min}) \tag{12}$$

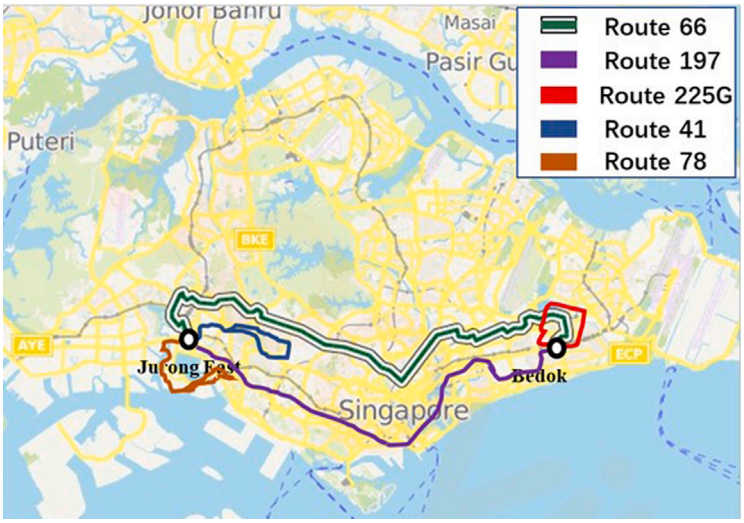

Fig. 6. Layout of the tested bus network.

Table 2
Route information.

Route	Length	Headway				# trips
		MP	MOP	EP	EOP	
66	33.2	10	15	15	20	76*2
197	31.4	10	15	15	20	76*2
225G	13.2	8	10	8	15	108
41	14.8	15	20	15	20	60
78	24.5	10	15	10	15	84

MP: Morning peak; MOP: Morning off-peak; EP: Evening peak; EOP: Evening off-peak.

Then a weighted sum function is used to evaluate the performance of the newly generated candidate sample points, which is given below

$$S(\mathbf{x}_l^{new}) = \phi_O S_O(\mathbf{x}_l^{new}) + \phi_D S_D(\mathbf{x}_l^{new}) \tag{13}$$

where ϕ_O and ϕ_D denote the weight for the surrogate function criterion and the distance criterion, and $\phi_O + \phi_D = 1$. The higher score, the better the sample point.

6. Numerical studies

To evaluate the effectiveness of the proposed two heuristic methods, two numerical studies are carried out. The first one is a demonstrated case on a small bus network with two bus terminals. The second one is based on the entire bus network of the Tower Transit which is a main public transport operator in Singapore.

6.1. Test bus network

6.1.1. Network information

We consider a test bus network that consists of two bus terminals and 5 bus routes, as shown in Fig. 6. The basic route information is listed in Table 2. In Singapore, the frequency of bus departures varies across different time periods to meet varying travel demand. A typical one-day operation horizon can be divided into four time periods: morning peak (before 9:00), morning off-peak (9:00–17:00), evening peak frequency (17:00–19:00) and evening off-peak (after 19:00). It can be noted that bus routes 66 and 197 have two directions with each direction involving 76 trips. Route 225G, route 41 and route 78 are loop routes with only one direction. The parking lot limitation for Jurong East terminal and Bedok terminal is 16 and 4, respectively.

The detailed information of chargers considered in our study is presented in Table 3, which is adapted from the study by Nicholas (2019). It is noted that we convert the US dollars into SG dollars at the rate of 1.34 (USD/SGD). Chargers will put huge pressure on the grid when the EB is charging. To guarantee the grid can afford such heavy load in peak-hour, the bus operator must apply for the upgrade of the grid capacity. The grid expansion costs mainly include the substation cost, transmission line cost and capacity

Table 3
Charger information in our study.

	350 kW DC fast (SGD)	70 kW AC normal (SGD)
Purchase cost	187,600	4190
Labor	37,306	3311
Materials	50,581	1655
Permit	389	379
Tax	206	209
Total	276,082	9744

expansion cost paid for the electricity supplier. Based on the study by Karlsson (2016), we assume the grid expansion cost is $\beta = 15{,}773$ SGD/100 kVA. We further assume all the current grid capacity is zero for all terminals and depots.

The EB adopted in our study is the mainstream EB on the bus market, namely, the 40 ft single deck EB (350 kWh battery, both overnight and opportunity charging). The purchase cost is 670,000 SGD which is the median price reported by Ambrose et al. (2017). The driver salary in Singapore is 3500 SGD/month, according to the report by Ong (2019). Because we consider one-day operation for the whole bus system in our model, the aforementioned costs should be converted to the average cost for each day. We assume an EB has a 10 year lifespan and the charger has a 5 year lifespan. Therefore, the daily fixed cost for an EB, a fast charger and a slow charger are 184, 182 and 11 SGD, respectively. For the same reason, the daily salary paid for a driver is 110 SGD.

6.1.2. Algorithm settings

Simulation settings: First, each EB has a maximum 150 km driving range with a brand-new battery in our simulation. As mentioned before, the battery will end its life when the capacity drops to its 80% of the health condition. Thus, the real driving range for an EB b without charging should be in the range of [120 km, 150 km]. Noted that the battery degradation will not reduce the required charging time, it is mainly because the battery will lead to the increase of the resistance in the battery, which will lower the charging efficiency. For example, it will take the same amount of time for a healthy battery and a degraded battery to be charged from the SOC of 0.8 to the SOC of 1. Second, the charging rate for the fast DC charger and the slow AC charger is 3.5 kWh/min and 0.7 kWh/min, respectively. Third, a minimum charging time $t^{\min}_{charge} = 5$ min is considered in our simulation. It means the EB should be charged for at least 5 min once it is connected to a charger. Fourth, an EB can serve different bus routes in operations. It is quite common in real bus operation which can enhance the efficiency of bus system and reduce the operating cost for PT operators. Fifth, the average travel time for EBs is assumed to be 18 km/h. Sixth, as for stochasticity, we consider the travel time and battery degradation to follow a normal distribution for simplicity. In general, the capacity of battery drops to its 80% of the health condition, the battery ends its life (Lam and Bauer, 2012). Let $cap^{\max}_b$ be the capacity of battery's health condition, the battery capacity $c\tilde{a}p_b$ of EB $b \in \mathcal{B}$ must fall in an interval $[0.8cap^{\max}_b, cap^{\max}_b]$. It is noted that we divide the one-day operation into 4 periods, morning peak, off-morning peak, evening peak and off-evening peak. For peak hours, due to the road congestions, we assume the travel time will be realized within the range $[\tilde{t}_v, 1.5\tilde{t}_v]$. As for non-peak hours, the travel time will be realized within the range $[\tilde{t}_v, 1.2\tilde{t}_v]$. Lastly, in order to main a good health state of EB battery, the SOC state of an EB should be within the range $[SOC^{\min} = 0.3, SOC^{\max} = 0.8]$.*RL-based algorithm settings*: The test bus network is a small-size bus network and we only consider deploying chargers at bus terminals. The total state size can be calculated by (7) is 2295. Therefore, we consider using a table to record the state-value. In the MC process, each episode consists of 10 timesteps and starts with a random state. We simulate 5000 episodes to obtain the value table. We adopt a random policy to select an action from action set $\mathcal{A}$.

SBO algorithm settings: In our SBO algorithm, we set $\phi_O = \phi_D = 0.5$. The simulation times of the agent-based simulation is set as 40. As for the perturbation of the local search for the decision variable, we set step size for small, medium and large perturbation as 1, 2 and 3, respectively. The stop criteria are : (1) iteration time exceeds 200; (2) the best solution did not change in successive 10 iterations.

All the experiments are carried out on a PC equipped with Intel Core i7-10700F 2.90 GHz CPU and 16 GB RAM. The agent-based simulation and the reinforcement learning process are realized by the C# programming. The Eq. (9) in the SBO method is solved by calling MATHNET API.

6.1.3. Solution results

The RL-based method and the SBO method found the solution with the same charger deployment results. We list the computational results in Table 4. To fulfill the given 556 trips, the optimal solution indicates that PT operators should deploy 2 fast chargers and 1 fast charger at Jurong East terminal and Bedok terminal, respectively. There is no slow charger included in the optimal solution. Due to the stochasticity of the travel time and battery capacity, we conducted 10 trials for both methods and took the average cost. The average operating cost computed by the reinforcement learning algorithm is 23,664 SGD each day, which means 81 EBs are required to perform the trip tasks in a typical working day. The charging cost is 9379 SGD and it accounts for nearly 28% of the total cost. The facility cost is proportional to the number of chargers and types. Compared to the operating cost and charging cost, it only accounts for a small proportion with 546. As for the SBO algorithm, similar results are also reported. The average operating cost, average charging cost, facility cost and average total cost are 23,582 SGD, 9392 SGD, 546 SGD and 33,520 SGD, respectively. The computational results suggest that the PT operators should pay more attention to reducing the operating cost and charging cost to save their money.

Table 4
Computational results.

Methods	Charger deployment				Average operating cost (SGD)	Average charging cost (SGD)	Facility cost (SGD)	Average total cost (SGD)
	Jurong East		Bedok					
	Fast	Slow	Fast	Slow				
RL-based	2	0	1	0	23,664	9379	546	33,589
SBO	2	0	1	0	23,582	9392	546	33,520

Table 5
Computational results.

Stop criteria Δ	Average number of episodes of convergence	Average total cost (SGD/day)	Average CPU time (s)
1×10^{-2}	1123	33,758	2034
1×10^{-3}	1935	33,503	3542
5×10^{-3}	2508	33,593	4546
1×10^{-4}	3708	33,589	6720
0←	\	33,590	9082

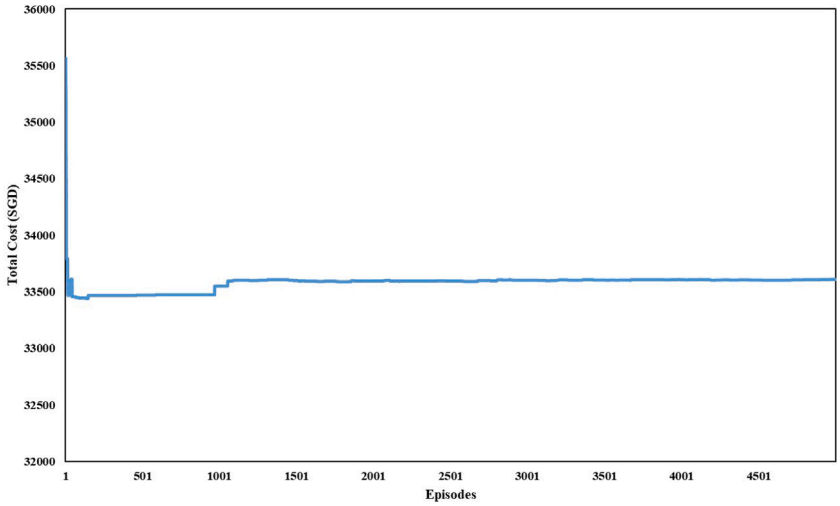

Fig. 7. Convergence process of one experiment.

6.1.4. Algorithm performances

For the reinforcement learning method, we use a table to record the state value. Because we adopted a random start for each episode and a random action policy, the convergence of the algorithm will be affected by the sampling process. We therefore repeated the algorithm multiple times (i.e. 10 times in our experiment) to test the average performance. Table 5 shows the computational results for different Δ values. As Δ decreases, the number of episodes required to carry out MC process increases. The average number of episodes to reach convergence grows from 1123 to 3708 when Δ declines from 1×10^{-2} to 1×10^{-4}. For the last row, we did not impose stop criteria in the algorithm and ran the process for a long enough time. The final objective function value will converge to 33,590. As for the average computational time, it increases from 2034 s (33.9 min) to 9082 (2 h 31 min) when the stop criteria Δ drops from 10^{-2} to 0. Because the CFP problem is a planning problem, PT operators can adopt a small Δ of the reinforcement learning method to get a reliable solution even though the computational time may be prolonged.

Fig. 7 shows the optimal objective value with the simulated episodes in one of our experiments. At the beginning of the MC process, the optimal state value fluctuates significantly. As more episodes have been gone through, more samples are added to evaluate the value table. When the number of simulated episodes goes to about 1100, the optimal value becomes much more stable, fluctuating around 33,600. Finally, the optimal value closes to 33,585 when the number of episodes reaches 5000 when the MC process ends. For the rest of the experiments, similar trends can also be found.

For the SBO method, we use the LHS method to random sample 6 start sample points. Then we conducted 10 trials of the SBO algorithm. It is pretty fast to find the optimal solution compared to the reinforcement learning method. On average, it takes 134 iterations to find the optimal solution, using around 17 min. As for the reinforcement learning method, it requires a large number of samples to better evaluate the state-value table. If the stop criteria Δ set as 1×10^{-4}, 1 h 52 min is required to carry out the MC process (3708 episodes) on average. Although the solution speed of SBO method outperforms the reinforcement learning method, the reinforcement learning method is reliable compared to the SBO method. We calculate the variance of the optimal values outputted

Table 6
Results for different simulation times for SBO.

Simulation times	Total cost variances	Average total cost (SGD/day)	Average CPU time (s)
20	22,767.8	33,552	781
40	10,514.2	33,520	1023
100	1369.3	33,557	2280
200	806.3	33,581	4072
300	248.3	33,584	5567
400	98.8	35,587	7442
500	76.6	33,588	9821

Table 7
Sensitivity analysis results of the minimum charging time.

t^{min}_{charge}	Charger deployment				Average operating cost (SGD)	Average charging cost (SGD)	Facility cost (SGD)	Average total cost (SGD)
	Jurong East		Bedok					
	Fast	Slow	Fast	Slow				
0	2	0	1	0	23,172	9528	546	33,246
5	2	0	1	0	23,664	9379	546	33,589
10	1	2	1	0	26,311	7284	386	33,981
15	0	3	0	1	29,328	6532	44	35,904
20	0	3	0	1	29,755	6422	44	36,221

by these two methods. The variance of the reinforcement learning method is 61.63, and the variance of the SBO method is 10,514.24. By the law of large number theory, the state value will converge to a stable value if the number of episodes goes to infinity in the MC process. The reinforcement learning method considers more samples in the entire solution space compared to the SBO method, which samples based on the guidance of the surrogate function. As a result, the reinforcement learning method works slowly, but a stable solution can be found. The SBO method is computational time-efficient, while the solution stability cannot be guaranteed when the stochasticity is considered. To explore the relationship between the agent-based simulation times and the solution results, we further vary the simulation times from 20 to 500 and test the results under the optimal charging facility deployment results. We also repeat 10 times for each simulation time and take the average total cost. Table 6 demonstrates the computational results with regard to total cost variances, average total costs and computational time. When the simulation times grow from 20 to 500, the total cost variances decrease from 22,767.8 to 76.6, which means the average total costs become more stable with the increase of the agent-based simulations. The total average costs are also close to 33,589 SGD/day, which is the same as computed by the reinforcement learning method. However, more computational resources are required to carry out a large number of simulations to guarantee the stability of the computational results. The computational time increases from 781 s (13 min) to 9821 s (2 h 43 min).

6.1.5. Sensitivity analysis

The minimum charging time: A minimum charging time requirement is established to guarantee an effective charging process for each charging activity, preventing repeated plugging in and unplugging throughout the charging process. For the instance in Section 6.1.3 (hereinafter, it refers to the baseline scenario), we set the minimum charging requirement $t^{min}_{charge} = 5$ min. We now vary t^{min}_{charge} from 0 to 20 min with an increment of 5 min to investigate its impact on the solution result. Due to the stability issue, we adopted the RL-based method in the sensitivity analysis. Table 7 shows the computational result. In general, the average total cost increases from 33,246 SGD to 36,221 SGD as t^{min}_{charge} grows from 0 to 20. The major cost increment comes from the operating cost, which means more EBs should be deployed to finish the given trip tasks with higher t^{min}_{charge}. As a result, the shorter the t^{min}_{charge}, the more flexible the charging process for EBs, leading to the cost saving in the operating cost (i.e., fewer EBs and drivers). However, some extra time is unavoidable and it cannot be used for charging. To enhance the charging efficiency and avoid frequent plugging in and unplugging, it is not practical to set t^{min}_{charge} as 0. Moreover, we find that more slow chargers are deployed with the increase of t^{min}_{charge}. For instance, when $t^{min}_{charge} = 0$ and 5, the solution suggests the fast charger should be deployed, while the slow charger should be deployed when t^{min}_{charge} increases to 15 and 20. The reasons are twofold. On the one hand, more EBs have been deployed for larger t^{min}_{charge}. As a result, the total en-route charging demand will decrease. On the other hand, each EB will have a longer charging time and a higher potential to be charged for larger t^{min}_{charge}. Therefore, deploying slow chargers become beneficial.

The SOC range: To avoid excess battery degradation, the SOC state of an EB should be within the range of $[SOC^{min}, SOC^{max}]$. In the baseline scenario, we set the range as [0.3, 0.8]. To investigate how the SOC range affects the solution result, we vary it from [0.4, 0.6] to [0, 1] and solve the model using the proposed RL-based method. The computational result is shown in Table 8. Note that there is no solution when the SOC range is [0.4, 0.6] because the fully-charged EB can only travel $0.2 * 150 = 30$ km, which is lower than the travel distance of bus route 66 and 197 (33.2 km and 31.4 km, respectively). Overall, the total average decreases as the SOC range becomes narrow, dropping from 33,589 SGD to 26,144 SGD when the SOC range increases from [0.3, 0.7] to [0, 1]. A large range allows for more flexibility in the use of EBs and more initial stored energy, which in turn reduces the number of EBs needed and the need for en-route charging. Therefore, both the operating cost and charging cost drop as the SOC range increases. Moreover, the deployed number of chargers also shows a decreasing trend when the SOC range increases.

Table 8
Sensitivity analysis results of the SOC range.

SOC range	Charger deployment				Average operating cost (SGD)	Average charging cost (SGD)	Facility cost (SGD)	Average total cost (SGD)
	Jurong East		Bedok					
	Fast	Slow	Fast	Slow				
[0.4, 0.6]	\	\	\	\	\	\	\	\
[0.3, 0.7]	2	0	1	0	23,664	9379	546	33,589
[0.2, 0.8]	2	0	1	0	22,711	7588	546	30,845
[0.1, 0.9]	2	0	1	0	21,853	5928	546	28,327
[0, 1]	1	0	1	0	20,952	4828	364	26,144

Table 9
Bus route information.

Terminals	Routes
Bukit Batok	77, 106, 173, 177, 189, 941, 945, 947, 990
Jurong East	41, 49, 66, 78, 79, 97, 97e, 98, 98M, 143, 143M, 183, 333, 334, 335
Clementi	96, 282, 284, 285
Joo Koon	974

Table 10
Parking lot information in each terminal.

Terminals	# Parking lots	Terminals	# Parking lots
Bukit Batok	4	Marina Center	2
Jurong East	16	Bedok	1
Clementi	3	Toa Payoh	1
Joo Koon	1	Shentong Way	1

6.2. Tower transit bus network

6.2.1. Network information

Now we move on to a large-scale network based on Tower Transit, which is an operator of bus services in Singapore and it now operates over 300 buses and 29 bus routes on behalf of the Land Transport Authority. The following Table 9 gives the bus route information which includes 29 different bus routes from 4 terminals owned by Tower Transit. The frequencies for different routes are obtained from the official website of Tower Transit https://towertransit.sg/our-routes/. The frequencies are varied from the time of the day and we list this detailed information in Appendix D.

Except for the 4 terminals owned by Tower Transit, there are another 4 terminals that Tower Transit shares with other public transit operators. Table 10 lists the parking lot information for these total 8 terminals. Because the parking lots may be shared by different operators, we assume the number of parking lots owned by an operator is proportional to the number of its operated bus routes based on this terminal. The number of installed chargers should be less than the number of parking lots.

6.2.2. Algorithm settings

All the settings for both algorithms are the same as the test bus network except that we use the neural network to represent the state-value table. The Tower Transit bus network involves 8 bus terminals and the total state size calculated by Eq. (7) is 22,307,400, which requires a large memory to store these values and also imposes pressure on computational times. Therefore, we first perform the MC process to generate sample data. We then build a neural network through Tensorflow based on Python and train the neural network through the sample data. The neural network consists of 4 dense layers, in which the first layer involves 64 neuron/units and the input shape is (16,). The second and the third layer contains 640 and 320 neuron/units, respectively. They both adopted the sigmoid function as the activation function. The last layer is the output layer which only includes one neuron/unit.

6.2.3. Computational results

We first examine the performance of the neural network. The MC process generates 26,266 sample points, and these samples are used to train a 4-layer neural network. 80% of the data are used for the training set and 20% of them are used for the validation set. Fig. 8 gives the loss value of both the training set and validation in 1000 epochs. It is noted that the cost value for each state is divided by 10,000 for scaling purposes. As we can see, both loss values drop with the training process, reaching a stable state (around 0.08) after about 700 epochs. This curve indicates that the parameters of the neural network converge using the training data.

After the training process has been done, we further generate a group of test datasets to test the performance of our trained neural network. 1000 samples are used to validate the performance and the Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are 1104, 824 and 0.44%, respectively. Compared to the average cost of 185,762 SGD of the 1000 samples, these indicators demonstrate that the trained neural network can predict the total cost well.

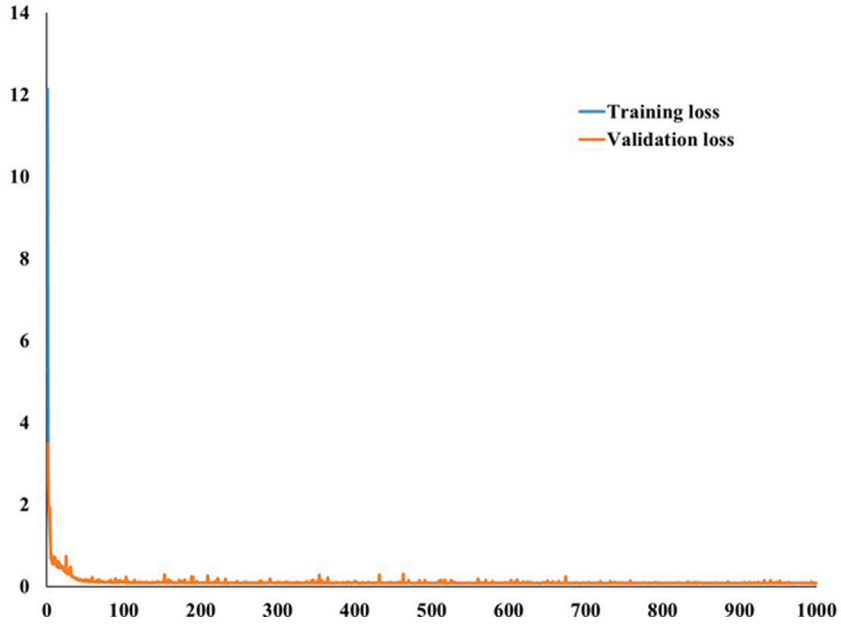

Fig. 8. Convergence process of one experiment.

Table 11
Computational chargers in terminals.

Terminals	# Chargers		Terminals	# Chargers	
	Fast	Slow		Fast	Slow
Bukit Batok	3	0	Marina Center	2	0
Jurong East	5	0	Bedok	1	0
Clementi	0	3	Toa Payoh	1	0
Joo Koon	1	0	Shentong Way	1	0

Table 12
Different costs.

Components	Cost(SGD)
Average operating cost	134,064
Average charging cost	46,554
Facility cost	2581
Total cost	183,199

Based on the trained neural network, we can obtain the optimal deployment results. We list the computational results in Tables 11 and 12. Note that the timetable of Tower Transit SG contains 4745 trips in a one-day operation. Table 11 gives the charger deployment output in all terminals, total 17 chargers should be installed in terminals. It is obvious that the number of chargers is proportional to the number of bus routes based on this terminal. For example, the Jurong East terminal, which has 15 bus routes, requires 5 fast chargers to be implemented to meet the charging demand. But the Joo Koon terminal only needs 1 fast charger to be installed because only 1 bus route is based on it. Table 12 shows different cost components of the computational results. As we can see, the average operating cost is 134,064 SGD, which means about 456 EBs are required to perform the trip tasks in a typical working day. The charging cost also accounts for a high proportion of the total cost, which is 46,554 SGD on average. The facility cost is proportional to the number of chargers and types. Compared to the operating cost and charging cost, it only accounts for a small proportion with 2581 SGD. The computational results suggest that the PT operators should pay more attention to reducing the operating cost and charging cost to save their money.

For the SBO method, we conducted ten trials of the SBO algorithm. Fig. 9 reports the box plot of the objective value and computational time of these 10 trials. Compared to the RL-based method, the SBO method cannot get a stable result. The objective value varies from 182,981 SGD to 186,251 SGD and the average value is 184,354.5 SGD. As for the computational time, it also fluctuates, varying from 8820 s (2 h 27 min) to 19,232 s (5 h 20 min). The average time to find out the solution is 12,925.7 s (3 h 35 min). The performance of SBO is affected by many factors, such as the start sample points and the simulation times (counts). On the one hand, a good start sample point allows the algorithm to converge quickly. On the other hand, more simulation times for a given charger deployment lead to an increase in the total computational time of the SBO. However, the stability of the SBO

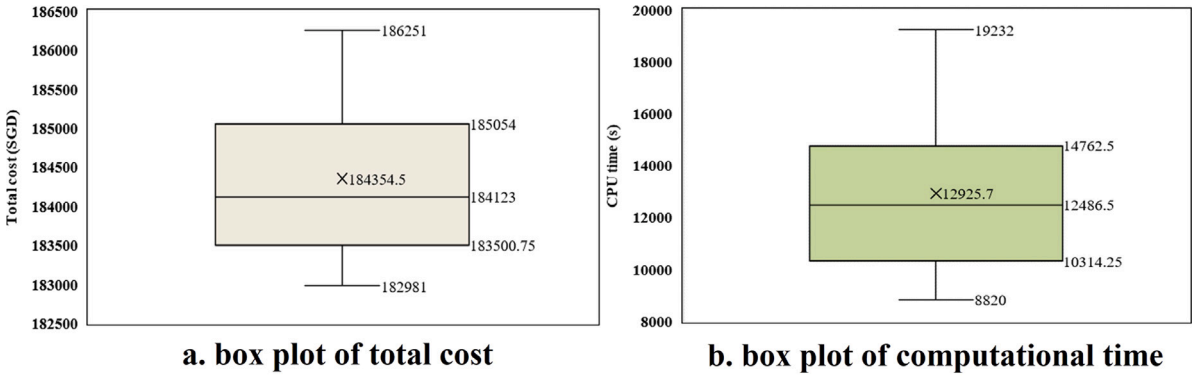

Fig. 9. Performance of SBO method.

Table 13
Computational chargers in terminals for the deterministic situation.

Terminals	# Chargers		Terminals	# Chargers	
	Fast	Slow		Fast	Slow
Bukit Batok	2	0	Marina Center	2	0
Jurong East	4	0	Bedok	1	0
Clementi	0	3	Toa Payoh	1	0
Joo Koon	1	0	Shentong Way	1	0

Table 14
Different costs for the deterministic situation.

Components	Cost(SGD)
Average operating cost	117,012
Average charging cost	38,428
Facility cost	2216
Total cost	157,656

can be enhanced, as we analyzed in the test network. The total computational time for the RL-based method is 74,459.3 s (20 h 40 min), which includes three parts: (i) MC process (74,251 s), (ii) Neural network training process (208 s) and (iii) Final result output process (0.3 s). The vast majority of time is spent on the MC process, constituting over 99% of the total time. This is because the neural network requires a large number of sample points generated by MC to approximate the relationship between the state and the objective value. The Neural network training process and final result output process are quite fast. As for the SBO method, the computational time fluctuates drastically, varying from 8820 s to 19,232 s.

6.2.4. Deterministic vs stochastic

In the previous case study, we investigated the stochastic charging facility deployment problem, which considers the real-world uncertainties in terms of travel time uncertainty and battery degradation uncertainty. Here, we want to explore the differences between the deterministic charging facility deployment situation with the stochastic one. Therefore, we set the travel time of all trips and the initial capacity of all EBs as deterministic, and perform the proposed reinforcement learning method and SBO method again. We list the computational results for the deterministic situation in Tables 13 and 14. Compared to the stochastic situation, each cost component in the deterministic situation is lower. Specifically, the average operating cost drops from 134,064 to 117,012 SGD, which means the number of EBs required to finish the tasks drops from 456 to 398. Furthermore, the average charging cost and facility cost also fall from 46,554 to 38,428 and from 2581 to 2216, respectively. The reason is twofold. On the one hand, we assume the travel time will be realized between the $[\tilde{t}_v, 1.5\tilde{t}_v]$ in peak hour and $[\tilde{t}_v, 1.2\tilde{t}_v]$ in the off-peak hour. The expectation of travel time in the stochastic situation is higher than the deterministic situation, because the energy consumption is proportional to the travel time. As a result, the total charging demand becomes higher for the stochastic situation. Therefore, more recharging activities are needed to fulfill the same amount of trip tasks for the stochastic situation. On the other hand, when the real-world uncertainties are considered, the capacity of the battery will be faded. It means that more EBs are required for the stochastic situation, which leads to a higher operating cost.

7. Conclusions

In this paper, we study the charging facility deployment problem for an EB transit system under uncertainties. A two-stage stochastic programming model is developed for the EB charging facility deployment problem. In view of the difficulties of the

second stage problem, we design a multi-agent EB simulation system to simulate the EB operation process. To solve the entire two-stage stochastic programming model, we proposed two different heuristic solution methods: the reinforcement learning method and the SBO method. We carry out a series of experiments based on a test EB network and a real-size EB network in Singapore to assess the performance of the proposed model and solution methods. The two heuristic methods both exhibit good efficiency and effectiveness in solving the proposed model. Finally, some managerial insights are also provided to bus companies to enhance the performance of the EB transit system.

Further research is recommended in two directions. The first one is to develop an efficient exact algorithm to solve the second-stage problem optimally by exploring the structures and properties of the model. As a result, the charging demand distribution will be estimated accurately, leading to a better charging facility deployment plan. The second one is to incorporate advanced charging technologies such as wireless charging and conductive charging. This offers the possibility of further enhancing the performance of the EB system.

CRediT authorship contribution statement

Yu Zhou: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Review & editing. **Ghim Ping Ong:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Qiang Meng:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Haipeng Cui:** Methodology, Software, Review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is supported by the Ministry of Education, Singapore, via the project R-302-000-286-114 under the MOE Tier 1 Grant FY2021. Any opinions, findings and conclusions or recommendations expressed in this study are those of the author(s) and do not reflect the views of the Ministry of Education, Singapore. All authors approved the final version of the manuscript.

Appendix A. Notations

Sets	
$\mathcal{X}$	Set of bus depots
$\mathcal{S}$	Set of bus terminals
$\mathcal{K}$	Set of charger types
$\mathcal{V}$	Set of all trips
$\mathcal{B}$	Set of all EBs
Parameters	
$\tilde{t}_v^d, \tilde{t}_v^a$	Departure time and arrival time of trip v
ter_v^d, ter_v^a	Departure terminal and arrival terminal of trip v
$\tilde{t}_v$	Travel time of trip v
$\tilde{\Delta}_v$	Energy consumption of trip v
$SOC_{\min}, SOC_{\max}$	Upper and lower bound for battery SOC state
β	Unit grid expansion cost in SGD/100 kVA
L_s, L_w	Grid capacity for terminal/depot s/w
$\tilde{cap}_b$	Battery capacity of EB b
$cap_b^{\max}$	Capacity of battery's health condition of EB b
C_k	Cost for charger of type k
P_k	Maximum charging power for charger of type k
$t_{charge}^{\min}$	Minimum charging time requirement
θ	Minimum terminal time
F_i	Current grid capacity of terminal/depot i
u_i	Parking lots number of terminal/depot i
α_k	Charging rate of charger of type k

Variables	
$x_{i,j}$	Charger number of type j at terminal/depot i
$y^b_{i,j}$	1 means EB $b \in \mathcal{B}$ will take trip task $j \in \mathcal{V}$ after it completes trip task $i \in \mathcal{V}$, 0 otherwise.
$y^b_{0,i}$	1 if EB $b \in \mathcal{B}$ takes trip task $i \in \mathcal{V}$ as its first job, 0 otherwise.
$y^b_{i,V+1}$	1 if trip task $i \in \mathcal{V}$ is the last job of EB $b \in \mathcal{B}$, 0 otherwise.
$\bar{z}^b_i$	Electricity state of EB $b \in \mathcal{B}$ when it departs from terminal to perform trip task $i \in \mathcal{V}$.
$\underline{z}^b_i$	Electricity state of EB $b \in \mathcal{B}$ when it arrives at the terminal just after the completion of trip task $i \in \mathcal{V}$.
$\Delta z^b_{i,j}$	Charging electricity of bus b between carries out trip i and j.
$T^b_{i,j}$	Charging time of EB $b \in \mathcal{B}$ at the terminal before the EB starts trip task $j \in \mathcal{V}$ after it completes the trip task $i \in \mathcal{V}$.

Appendix B. EB-CSP model

We first give the decision variables, which are in decision matrices **y**, **z** and **T**

$y^b_{i,j}$	1 means EB $b \in \mathcal{B}$ will take trip task $j \in \mathcal{V}$ after it completes trip task $i \in \mathcal{V}$, 0 otherwise.
$y^b_{0,i}$	1 if EB $b \in \mathcal{B}$ takes trip task $i \in \mathcal{V}$ as its first job, 0 otherwise.
$y^b_{i,V+1}$	1 if trip task $i \in \mathcal{V}$ is the last job of EB $b \in \mathcal{B}$, 0 otherwise.
$\bar{z}^b_i$	Electricity state of EB $b \in \mathcal{B}$ when it departs from terminal to perform trip task $i \in \mathcal{V}$.
$\underline{z}^b_i$	Electricity state of EB $b \in \mathcal{B}$ when it arrives at the terminal just after the completion of trip task $i \in \mathcal{V}$.
$\Delta z^b_{i,j}$	Charged electricity of EB b after carrying out trip i and before carrying out trip j
$T^b_{i,j}$	Charging time of EB $b \in \mathcal{B}$ at the terminal before the EB starts trip task $j \in \mathcal{V}$ after it completes the trip task $i \in \mathcal{V}$.

With the above decision variables, the second-stage problem can be formulated as

[M2]

$$\min(c_1 \sum_{b\in\mathcal{B}}\sum_{j\in\mathcal{V}} y^b_{0,j} + c_2 \sum_{b\in\mathcal{B}}\sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}:i\neq j} \Delta z^b_{i,j}) \tag{14}$$

s.t.

$$\sum_{j\in\mathcal{V}} y^b_{0,j} \le 1, \forall b \in \mathcal{B} \tag{15}$$

$$\sum_{i\in\mathcal{V}} y^b_{i,(V+1)} \le 1, \forall b \in \mathcal{B} \tag{16}$$

$$\sum_{j\in\mathcal{V}\cup\{V+1\}:i\neq j} y^b_{i,j} - \sum_{j\in\mathcal{V}\cup\{0\}:i\neq j} y^b_{j,i} = 0, \forall i \in \mathcal{V}, \forall b \in \mathcal{B} \tag{17}$$

$$\sum_{b\in\mathcal{B}}\sum_{i\in\mathcal{V}\cup\{0\}:i\neq j} y^b_{i,j} = 1, \forall j \in \mathcal{V} \tag{18}$$

$$t^d_j - t^a_i - \theta - T^b_{i,j} \ge M \cdot (y^b_{i,j} - 1), \forall i,j \in \mathcal{V}, i \neq j, \forall b \in \mathcal{B} \tag{19}$$

$$\Delta z^b_{i,j} \le \bar{z}^b_j - \underline{z}^b_i + M \cdot (1 - y^b_{i,j}), \forall b \in \mathcal{B}, \forall i,j \in \mathcal{V}, i \neq j \tag{20}$$

$$\bar{z}^b_j - \underline{z}^b_i - M \cdot (1 - y^b_{i,j}) \le \Delta z^b_{i,j}, \forall b \in \mathcal{B}, \forall i,j \in \mathcal{V}, i \neq j \tag{21}$$

$$\Delta z^b_{i,j} \le M \cdot y^b_{i,j}, \forall b \in \mathcal{B}, \forall i,j \in \mathcal{V}, i \neq j \tag{22}$$

Table C.15
Performance comparison of two methods.

# Trips	Gurobi		Simulation	
	Objective value (SGD)	CPU time (s)	Objective value (SGD)	CPU time (s)
10	1104.0	5.2	1104.0	0.45
30	1588.6	782	1612.4	0.62
50	2228.8	7200	2317.8	0.67
70	\	7200	2729.8	0.72
90	\	7200	3352.0	0.77
110	\	7200	3722.5	0.81
130	\	7200	4125.2	0.82

$$-M \cdot y_{i,j}^b \leq \Delta z_{i,j}^b, \forall b \in \mathcal{B}, \forall i, j \in \mathcal{V}, i \neq j \tag{23}$$

$$\Delta z_{i,j}^b = \sum_{k \in \mathcal{K}} T_{i,j}^b \cdot \alpha_k, \forall b \in \mathcal{B}, \forall i, j \in \mathcal{V}, i \neq j \tag{24}$$

$$0 \leq T_{i,j}^b \leq M \cdot y_{i,j}^b, \forall b \in \mathcal{B}, \forall i, j \in \mathcal{V}, i \neq j \tag{25}$$

$$\bar{z}_i^b - \underline{z}_i^b = \Delta_i \left(\sum_{j \in \mathcal{V} \cup \{V+1\}} y_{i,j}^b \right), \forall b \in \mathcal{B}, \forall i \in \mathcal{V} \tag{26}$$

$$\tilde{cap}_b \cdot SOC_{\min} \leq \bar{z}_i^b \leq \tilde{cap}_b \cdot SOC_{\max}, \forall b \in \mathcal{B}, \forall i \in \mathcal{V} \tag{27}$$

$$\tilde{cap}_b \cdot SOC_{\min} \leq \underline{z}_i^b \leq \tilde{cap}_b \cdot SOC_{\max}, \forall b \in \mathcal{B}, \forall i \in \mathcal{V} \tag{28}$$

$$realistic \quad constraints \tag{29}$$

Objective function (14) aims to minimize the operating cost, including fixed cost and charging cost. c_1 denotes the unit fixed cost by operating an EB and c_2 represents the unit charging cost. Constraints (15)–(18) are the basic operational requirements and the flow conservation. Constraint (19) ensures a minimum terminal time θ between two consecutive trips and M is a sufficiently large positive number. Constraints (20)–(23) depict the energy state conservation which means the variation of the electricity of an EB between two consecutive trips should equal the charged electricity. Constraint (24) imposes the relationship between the charging time, charging rate and charged energy. α_k is the charging rate of charger of type k. Constraint (25) states that $T_{i,j}^b = 0$ when $x_{i,j}^b = 0$. Constraint (26) depicts the energy variation of an EB after carrying out trip i should equal to the energy consumption Δ_i required by this trip. Constraints (27)–(28) depict the relationship between charging time length and charging time points. Constraint (29) denotes some realistic constraints that are hard to be formulated explicitly.

Appendix C. Performance of simulation

To demonstrate the efficiency and effectiveness of the agent-based simulation, we generate a bunch of test instances based on bus route 66 in Singapore (the basic information of route 66 is provided in Section 6.1.1). We first vary the total number of trips from 10 to 130 with an increment of 20 trips, creating total 7 instances. All the parameters are the same as the experiment in Section 6. Then we solve these instances with both Gurobi and our simulation system. It is noted we consider the basic EB-CSP model without considering "realistic constraint". For the Gurobi, we set the time limit as 7200 s. If the optimal solution is not found within the time limit, then the current best objective is reported. For the simulation system, we remove the uncertainty part, queuing part and charging capacity part to make a fair comparison.

The computational result is given in Table C.15. In general, When the number of trips is less than or equal to 50, simulation can produce a solution that is extremely close to Gurobi's solution, but it takes significantly less time. Specifically, for instance with 10 trips, both methods found the optimal solution in a short time (5.2 s and 0.45 s). For instance with 30 trips, Gurobi found the optimal solution using 782 s while the simulation found a suboptimal solution using only 0.62 s, compared to 782 s of Gurobi. For instance with 50 trips, Gurobi did not find the optimal solution within the time limit and a current best solution was reported (i.e., 2228.8). However, the simulation found a very close solution (i.e., 2317.8) to Gurobi in 0.67 s. As the number of trips continues growing, Gurobi cannot find a feasible solution within the time limit. The simulation can still find out the solution in a short time (less than 1 s). To sum up, although the optimality of the solution generated by the simulation cannot be guaranteed, the simulation can solve the EB-CSP very efficiently with a high-quality solution.

Appendix D. Route information of tower transit SG

Route	Length	DS	AS	R t	FS t	LS t	MP-h	OMP-h	EP-h	OEP-h
77	18.7	BB	MC	62	6:00	23:00	15	20	15	20
106	21.4	BB	SW	71	5:30	23:30	10	15	10	15
173	13.2	BB	BB	44	5:00	23:59	10	15	10	15
177	11.7	BB	BB	39	5:50	21:20	20	30	20	30
189	11.5	BB	BB	38	5:45	23:59	5	10	5	15
941	6.8	BB	BB	23	6:30	23:59	10	10	10	15
945	7.3	BB	BB	24	5:30	23:59	10	10	10	15
947	5	BB	BB	17	5:30	23:59	10	10	10	15
990	10.3	BB	BB	34	5:45	23:59	15	15	15	15
41	14.8	JE	JE	49	5:45	23:45	15	10	20	15
49	23.8	JE	JE	79	6:00	23:30	10	15	10	15
66	33.2	JE	BD	111	5:30	23:00	10	15	15	20
78	24.5	JE	JE	82	5:45	23:59	10	15	10	15
79	22.9	JE	JE	76	6:00	22:30	10	15	10	15
97	22.3	JE	MC	74	5:45	23:45	10	15	10	15
97e	22.3	JE	JE	74	7:36	8:24	12	0	0	0
98	26.9	JE	JE	90	5:30	23:59	10	15	10	15
98M	26.9	JE	JE	90	9:32	23:45	15	20	15	20
143	29.3	JE	TP	98	5:30	23:30	10	15	15	20
143M	9.9	JE	JE	33	5:30	23:59	15	20	15	20
183	22.8	JE	JE	76	6:00	23:30	10	15	10	20
333	8.4	JE	JE	28	5:30	23:59	10	15	10	15
334	9.3	JE	JE	31	5:30	23:59	10	10	10	15
335	10.4	JE	JE	35	5:00	23:59	10	15	10	15
96	6.8	CL	CL	23	6:45	23:30	5	10	5	15
282	5.8	CL	CL	19	5:35	23:59	10	15	10	15
284	2.7	CL	CL	9	5:55	23:59	5	10	5	10
285	7.8	CL	CL	26	5:55	23:59	5	15	5	15
974	35.2	JK	JK	117	5:45	23:30	15	20	15	20

DS: Departure station/terminal; AS: Arrival station/terminal; R t: Running time; FS t: First service time; LS t: Last service time; BB: Bukit Batok; MC: Marina Center; SW: Shentong Way; JE: Jurong East; BD: Bedok; CL: Clementi; JK: Joo Koon; TP: Toa Payoh.

References

Adler, J.D., Mirchandani, P.B., 2014. Online routing and battery reservations for electric vehicles with swappable batteries. Transp. Res. B 70, 285–302.

Adler, J.D., Mirchandani, P.B., 2017. The vehicle scheduling problem for fleets with alternative-fuel vehicles. Transp. Sci. 51 (2), 441–456.

Alwesabi, Y., Avishan, F., Yanıkoğlu, İ., Liu, Z., Wang, Y., 2022. Robust strategic planning of dynamic wireless charging infrastructure for electric buses. Appl. Energy 307, 118243.

Ambrose, H., Pappas, N., Kendall, A., 2017. Exploring the costs of electrification for California's transit agencies. ITS Rep. 2017 (03).

An, K., 2020. Battery electric bus infrastructure planning under demand uncertainty. Transp. Res. C 111, 572–587.

An, K., Jing, W., Kim, I., 2020. Battery-swapping facility planning for electric buses with local charging systems. Int. J. Sustain. Transp. 14 (7), 489–502.

Azadeh, S.S., Vester, J., Maknoon, M.Y., 2022. Electrification of a bus system with fast charging stations: Impact of battery degradation on design decisions. Transp. Res. C 142, 103807.

Becker, H., Manser, P., Hörl, S., Axhausen, K.W., 2020. Designing a large-scale public transport network using agent-based microsimulation. Transp. Res. A 1115.

BlueBus, 2021. 12-Meter bus. https://www.bluebus.fr/bluebus-12-metres/. (Accessed 28 July 2021).

Chen, C.-H., Lee, L.H., 2011. Stochastic Simulation Optimization: An Optimal Computing Budget Allocation, vol. 1. World scientific.

Dell'Amico, M., Fischetti, M., Toth, P., 1993. Heuristic algorithms for the multiple depot vehicle scheduling problem. Manage. Sci. 39 (1), 115–125.

Farahani, R.Z., Asgari, N., Heidari, N., Hosseininia, M., Goh, M., 2012. Covering problems in facility location: A review. Comput. Ind. Eng. 62 (1), 368–407.

Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Prog. Aerosp. Sci. 45 (1–3), 50–79.

He, Y., Liu, Z., Song, Z., 2022. Integrated charging infrastructure planning and charging scheduling for battery electric bus systems. Transp. Res. D 111, 103437.

He, J., Yang, H., Tang, T.-Q., Huang, H.-J., 2020. Optimal deployment of wireless charging lanes considering their adverse effect on road capacity. Transp. Res. C 111, 171–184.

Hof, J., Schneider, M., Goeke, D., 2017. Solving the battery swap station location-routing problem with capacitated electric vehicles using an AVNS algorithm for vehicle-routing problems with intermediate stops. Transp. Res. B 97, 102–112.

Hsu, Y.-T., Yan, S., Huang, P., 2021. The depot and charging facility location problem for electrifying urban bus services. Transp. Res. D 100, 103053.

Hu, H., Du, B., Liu, W., Perez, P., 2022. A joint optimisation model for charger locating and electric bus charging scheduling considering opportunity fast charging and uncertainties. Transp. Res. C 141, 103732.

Iliopoulou, C., Kepaptsoglou, K., 2021. Robust electric transit route network design problem (RE-TRNDP) with delay considerations: Model and application. Transp. Res. C 129, 103255.

Iliopoulou, C., Tassopoulos, I., Kepaptsoglou, K., Beligiannis, G., 2019. Electric transit route network design problem: Model and application. Transp. Res. Rec. 2673 (8), 264–274.

Jakobsson, S., Patriksson, M., Rudholm, J., Wojciechowski, A., 2010. A method for simulation based optimization using radial basis functions. Opt. Eng. 11 (4), 501–532.
Jefferies, D., Göhlich, D., 2020. A comprehensive TCO evaluation method for electric bus systems based on discrete-event simulation including bus scheduling and charging infrastructure optimisation. World Electr. Veh. J. 11 (3), 56.
Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. J. Global Optim. 13 (4), 455–492.
Karlsson, E., 2016. Charging infrastructure forelectric city buses: An analysis of grid impact and costs.
Kunith, A., Mendelevitch, R., Goehlich, D., 2017. Electrification of a city bus network—An optimization model for cost-effective placing of charging infrastructure and battery sizing of fast-charging electric bus systems. Int. J. Sustain. Transp. 11 (10), 707–720.
Lam, L., Bauer, P., 2012. Practical capacity fading model for Li-ion battery cells in electric vehicles. IEEE Trans. Power Electron. 28 (12), 5910–5918.
Layeb, S.B., Jaoua, A., Jbira, A., Makhlouf, Y., 2018. A simulation-optimization approach for scheduling in stochastic freight transportation. Comput. Ind. Eng. 126, 99–110.
Li, J.-Q., 2014. Transit bus scheduling with limited energy. Transp. Sci. 48 (4), 521–539.
Li, L., Lo, H.K., Xiao, F., 2019. Mixed bus fleet scheduling under range and refueling constraints. Transp. Res. C 104, 443–462.
Lin, Y., Zhang, K., Shen, Z.-J.M., Ye, B., Miao, L., 2019. Multistage large-scale charging station planning for electric buses considering transportation network and power grid. Transp. Res. C 107, 423–443.
Liu, T., Ceder, A.A., 2020. Battery-electric transit vehicle scheduling with optimal number of stationary chargers. Transp. Res. C 114, 118–139.
Liu, K., Gao, H., Liang, Z., Zhao, M., Li, C., 2021. Optimal charging strategy for large-scale electric buses considering resource constraints. Transp. Res. D 99, 103009.
Liu, K., Gao, H., Wang, Y., Feng, T., Li, C., 2022. Robust charging strategies for electric bus fleets under energy consumption uncertainty. Transp. Res. D 104, 103215.
Liu, Z., Song, Z., 2017. Robust planning of dynamic wireless charging infrastructure for battery electric buses. Transp. Res. C 83, 77–103.
Liu, H., Wang, D.Z.W., 2017. Locating multiple types of charging facilities for battery electric vehicles. Transp. Res. B 103, 30–55.
LTA, 2019. Land transport master plan 2040. Land Transp. Auth..
Mak, H.-Y., Rong, Y., Shen, Z.-J.M., 2013. Infrastructure planning for electric vehicles with battery swapping. Manage. Sci. 59 (7), 1557–1575.
Masmoudi, M.A., Hosny, M., Demir, E., Genikomsakis, K.N., Cheikhrouhou, N., 2018. The dial-a-ride problem with electric vehicles and battery swapping stations. Transp. Res. E 118, 392–420.
Mirchandani, P.B., Francis, R.L., 1990. Discrete Location Theory.
Nicholas, M., 2019. Estimating electric vehicle charging infrastructure costs across major US metropolitan areas.
Ong, 2019. New bus drivers in Singapore can get up to 3,500 monthly salary and 21 days annual leave. https://mustsharenews.com/bus-drivers-salary/. (Accessed 18 August 2021).
Pelletier, S., Jabali, O., Laporte, G., 2018. Charge scheduling for electric freight vehicles. Transp. Res. B 115, 246–269.
Riemann, R., Wang, D.Z.W., Busch, F., 2015. Optimal location of wireless charging facilities for electric vehicles: Flow-capturing location model with stochastic user equilibrium. Transp. Res. C 58, 1–12.
Sassi, O., Oulamara, A., 2014. Joint scheduling and optimal charging of electric vehicles problem. In: International Conference on Computational Science and Its Applications. Springer, pp. 76–91.
Sebastiani, M.T., Lüders, R., Fonseca, K.V.O., 2016. Evaluating electric bus operation for a real-world BRT public transportation using simulation optimization. IEEE Trans. Intell. Transp. Syst. 17 (10), 2777–2786.
SustainableBus, 2020. ABB to power 40 electric buses in Singapore. https://www.sustainable-bus.com/components/abb-to-power-40-electric-buses-in-singapore/. (Accessed 28 July 2021).
Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. MIT Press.
Tang, X., Lin, X., He, F., 2019. Robust scheduling strategies of electric buses under stochastic traffic conditions. Transp. Res. C 105, 163–182.
Wang, H., Zhao, D., Meng, Q., Ong, G.P., Lee, D.-H., 2019. A four-step method for electric-vehicle charging facility deployment in a dense city: An empirical study in Singapore. Transp. Res. A 119, 224–237.
Wu, W., Lin, Y., Liu, R., Jin, W., 2022. The multi-depot electric vehicle scheduling problem with power grid characteristics. Transp. Res. B 155, 322–347.
Xu, M., Yang, H., Wang, S., 2020. Mitigate the range anxiety: Siting battery charging stations for electric vehicle drivers. Transp. Res. C 114, 164–188.
Xylia, M., Leduc, S., Patrizio, P., Kraxner, F., Silveira, S., 2017. Locating charging infrastructure for electric buses in stockholm. Transp. Res. C 78, 183–200.
Yang, W., 2018. A user-choice model for locating congested fast charging stations. Transp. Res. E 110, 189–213.
Yang, J., Guo, F., Zhang, M., 2017. Optimal planning of swapping/charging station network with customer satisfaction. Transp. Res. E 103, 174–197.
Yıldız, B., Olcaytu, E., Şen, A., 2019. The urban recharging infrastructure design problem with stochastic demands and capacitated charging stations. Transp. Res. B 119, 22–44.
Zhang, K., He, F., Zhang, Z., Lin, X., Li, M., 2020a. Multi-vehicle routing problems with soft time windows: A multi-agent reinforcement learning approach. Transp. Res. C 121, 102861.
Zhang, A., Kang, J.E., Kwon, C., 2020b. Multi-day scenario analysis for battery electric vehicle feasibility assessment and charging infrastructure planning. Transp. Res. C 111, 439–457.
Zhang, H., Seshadri, R., Prakash, A.A., Antoniou, C., Pereira, F.C., Ben-Akiva, M., 2021a. Improving the accuracy and efficiency of online calibration for simulation-based dynamic traffic assignment. Transp. Res. C 128, 103195.
Zhang, L., Wang, S., Qu, X., 2021b. Optimal electric bus fleet scheduling considering battery degradation and non-linear charging profile. Transp. Res. E 154, 102445.
Zhou, Y., Meng, Q., Ong, G.P., 2022a. Electric bus charging scheduling for a single public transport route considering nonlinear charging profile and battery degradation effect. Transp. Res. B 159, 49–75.
Zhou, Y., Wang, H., Wang, Y., Li, R., 2022b. Robust optimization for integrated planning of electric-bus charger deployment and charging scheduling. Transp. Res. D 110, 103410.
Zwick, F., Kuehnel, N., Moeckel, R., Axhausen, K.W., 2021. Agent-based simulation of city-wide autonomous ride-pooling and the impact on traffic noise. Transp. Res. D 90, 102673.