# Approximate Optimum Curbside Utilisation for Pick-Up and Drop-Off (PUDO) and Parking Demands Using Reinforcement Learning

Qiming Ye, Yuxiang Feng, Jingshuo Qiu, Marc Stettler, Panagiotis Angeloudis

***Abstract*— With the uptake of automated transport, especially Pick-Up and Drop-Off (PUDO) operations of Shared Autonomous Vehicles (SAVs), the valet parking of passenger vehicles and delivery vans are envisaged to saturate our future streets. These emerging behaviours would join conventional on-street parking activities in an intensive competition for scarce curb resources. Existing curbside management approaches principally focus on those long-term parking demands, neglecting those short-term PUDO or docking events. Feasible solutions that coordinate diverse parking requests given limited curb space are still absent.**

**We propose a Reinforcement Learning (RL) method to dynamically dispatch parking areas to accommodate a hybrid stream of parking behaviours. A partially-learning Deep Deterministic Policy Gradient (DDPG) algorithm is trained to approximate optimum dispatching strategies. Modelling results reveal satisfying convergence guarantees and robust learning patterns. Namely, the proposed model successfully discriminates parking demands of distinctive sorts and priorities PUDOs and docking requests. Results also identify that when the demand-supply ratio situates at 2:1 to 4:1, the service rate approximates an optimal (83%), and curbside occupancy surges to 80%. This work provides a novel intelligent dispatching model for diverse and fine-grained parking demands. Furthermore, it sheds light on deploying distinctive administrative strategies to the curbside in different contexts.**

## I. INTRODUCTION

As the crucial interface between land use and transport systems, curbside represents a scarce asset where various means of mobility compete for room to stop or park [1], [2]. Whilst on-street parking activities [3], [4], Pick-Up and Drop-Offs (PUDOs) passengers, brief docking, and freights loading are emerging as primary and dominant curb activities in recent years [5], [6]. Furthermore, with the uptake of Autonomous Vehicles (AVs), in particular Shared AVs (SAVs), the average Vehicle Kilometres Travelled (VKT), the level of ride-sharing, frequencies of PUDOs and curb delivery are all envisaged to surge drastically [7], [8]. Therefore, it is reasonable to expect more intensive competition between distinctive curb parking demands.

Conventional curbside administrative measurements have limited effects on restraining unfettered parking behaviours and alleviating double park effects [1]. Meanwhile, merely expanding the parking space supply represents an unsustainable solution, which still fails to coordinate a diverse set of parking demands at the bottom [9]. Through optimising the curbside parking patterns at the district level, the system reduces 50% of traffic delays [10], and alleviates 64% of double parking incidents [11]. The real-time allocation of curbside parking space remains a critical factor towards efficient utilisation of curb resources. However, existing dynamic parking assignment models lack capabilities in considering those short-term parking activities. They also lack fine-grained strategies regarding a broad spectrum of diverse parking behaviours.

This paper addresses the curbside parking assignment problem in a novel context of the Intelligent Transport System (ITS) to fill this research gap. We assume that the curbside parking asset can be efficiently managed and dynamically coordinated among diversifying parking demands [5]. Based on well-established communications between vehicles and infrastructures, a fixed capacity of curb space could simultaneously accommodate PUDOs, docking and on-street parking demands in a finite time horizon.

We propose a Reinforcement Learning (RL) method to approximate optimum dispatching policies, of which sequential decisions address whether to accept the requests and where to accommodate them. The objectives maximise the rates of parking service and curbside occupancy. A partially learning Deep Deterministic Policy Gradient (DDPG) algorithm performs this online acting whilst off-policy training task. The model is trained for 1,500 epochs under multiple hybrid parking conditions.

The principal contribution of this paper is its novelty of optimising the sequential curbside dispatching strategy regarding both short and long-term parking behaviours. In addition, the findings shed light on the relationship between parking demand patterns and curbside performance, potentially benefiting city administrators to manage scarce curbside resources efficiently and strategically.

## II. LITERATURE REVIEW

### A. Curbside Management

Curb space represents the physical interface between urban land use and transport systems [1]. Statistics revealed that on-street parking accounts for over 20% of total parking demands in USA [3] and 36% in Europe [4]. Along with parking, the PUDO operations of the taxi or the ride-sharing mobility contribute to almost 55% of curb use in busy streets [5]. Meanwhile, the curb delivery business is projected to

Qiming Ye, qiming.ye18@imperial.ac.uk
Yuxiang Feng, y.feng19@imperial.ac.uk
Jingshuo Qiu, jingshuo.qiu19@imperial.ac.uk
Marc Stettler, m.stettler@imperial.ac.uk
Panagiotis Angeloudis, p.angeloudis@imperial.ac.uk
All authors are with the Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, South Kensington, SW7 2AZ, London

grow 75% by 2040, thanks to the continuously booming e-commerce [6].

Different parking purposes have distinctive average duration of stopping. Approximately 90% of PUDO behaviours last less than 1.5$mins$, while docking activities usually last less than 30$mins$ [5]. On the other hand, on-street parking vehicles usually stop for a more extended period, about 3$hrs$ on average [8].

Conventional curbside management imposes regulations to restrain unfettered parking behaviours for the benefit of a balanced demand-supply distribution [6]. However, 34% of passenger vehicle trips are still cruising for parking [9], while, 86% delivery vehicles stop on unauthorised zones [6]. Consequentially, their double-park manoeuvres potentially obstruct traffic flow. These under-coordinated conflicts significantly impede multi-stakeholders to access to curb space, and endanger the safety of vulnerable groups.

The mainstream studies agreed that merely expanding parking space supply can not secure sustainable and efficient exploitation of curb spaces [2]. Besides, in the era of AVs, the VKT, the level of ride-sharing, valet parking operations, frequencies of PUDOs, and curb delivery events are envisaged to surge drastically [7], [8]. Therefore, the intelligent and dynamic coordination of curb parking events are critical. Novel solutions were proposed to establish more automated curbside, where designated PUDO zones [5], sensing systems, dynamic pricing schemes [12] and dynamic allocation schemes are devised and some practised. An example showed that, by deploying pilot PUDO areas, 64% of double parking events are eliminated [11]. Another study found that through optimising the patterns of curb parking lanes, the average traffic delay reduces 194$s/veh$ [10].

### B. Reinforcement Learning and Applications

The curbside parking assignment problem is usually formulated as a Dynamic Programming (DP) problem. Reinforcement learning (RL) proves to be a promising method to unravelling these problems. The Deep Deterministic Policy Gradient (DDPG) algorithm is a quintessential RL method, which conducts off-policy learning in continuous action domains [13]. DDPG algorithms have been applied to optimal sequential decision problems in the field of Intelligent Transport Systems (ITS), like AVs operations, traffic signal control, and the right-of-way control [14].

RL has also been introduced to solve curbside parking assignment problems. For instance, A multi-agent RL model was built to dispatch parking sites for AVs [15]. Furthermore, the Estimation of Distribution Algorithm (EDA) represents another application of RL method to facilitate parking allocation decisions [16]. Despite existing progress, we have identified research opportunities as follows.

- Previous models principally focus on optimising on-street parking behaviours, whereas dynamic decisions on short-term parking demands, such as PUDOs and docking requests, are literately neglected.
- Most dynamic assignment models adopt vehicular-based modelling perspectives instead of strategically considering curbside performance.
- By adding PUDOs and docking behaviours into the arena, correlations between demands and curbside performance, such as optimum occupancy rate and optimum service rate, present a new puzzle.
- RL presents a promising solution to curb assignments. However, the current applications are still limited.

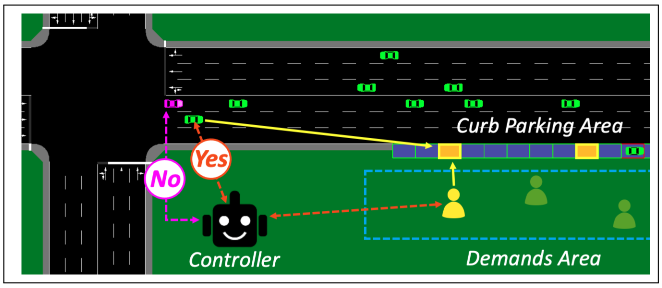


Fig. 1. Illustration of Curbside Environment and Parking Assignments. Purple line: rejected request; Orange lines: approved request; Yellow lines: movement direction; Yellow box: reserved parking areas; Red box: occupied area; Green box:available parking area; Blue-dashed box: demands area

## III. METHODOLOGY

### A. Problem Statement

Consider a discrete-time transport system involving a hybrid parking demand stream given a finite time horizon of T$= 3,600s$. Let V represents a fleet of SAVs potentially carrying out these demands. The total number of operational vehicles in the system is |V|=1,000 $vehs$. The subset VD comprises all vehicles $v$ dispatched to demands $q$ ∈Q, where |Q| denotes the total number of parking demands.

In this work, we consider three types of parking behaviours: the PUDOs, docking, and on-street parking, of which we use $u \in U$ to index each type. $k_u^v$ represents the parking duration per vehicle $v$. The pattern of all $k_u^v$ follows a truncated distribution $g(\bar{k}_u, \sigma_u)$. Herein, the mean duration $\bar{k}_u$ are 1.5$mins$, 10$mins$ and 60$mins$, respectively.

We designed a curbside comprising |P|=20 parallel parking areas to suffice the parking demands. Each parking area $p \in$ P is in a size of $5m \times 3m$. As illustrated in Fig.1, an intelligent controller decides whether to accept a proceeding AV upon its parking request at $t$ ∈TR. The set TR documents time steps when requests are received. A comprehensive assignment procedure considers the mean parking duration and the status-quo of the parking asset. However, the actual parking duration per vehicle is unknown to the controller. Once a request is approved, the controller reserves a space for this SAV.

The controller aims at high parking occupancy ($rc$) and service rate ($rs$). On the one hand, a high parking occupancy level reflects effective exploitation of the curbside asset. On the other hand, increasing the parking service rate means more parking demands are satisfied, with fewer idling costs and emissions. Table.I outlines the notations of variables and parameters of our proposed problem.

TABLE I

NOTATIONS OF CURB PARKING DYNAMIC ASSIGNMENT PROBLEM

| Notations | Specifications | Domains |
|---|---|---|
| T | set of discrete time step | |
| V | set of SAVs fleet | |
| VD | set of SAVs with parking demand | VD $\subset$ V |
| $\mathrm{VP}_t$ | set of parking SAVs | $\mathrm{VP}_t \subset$ V |
| Q | set of parking demand | |
| P | set of parking areas | |
| U | set of types of parking behaviours | |
| K | set of parking duration | |
| $\mathrm{S}_t$ | set of observed system states | |
| $t$ | a time step | $t \in$T |
| $v$ | an SAV | $v \in$V |
| $u$ | a type of parking behaviour | $u \in$U |
| $p$ | a parking area | $p \in$P |
| $p^*$ | an optimum parking area at $t$ | $p^* \in P_t$ |
| $q$ | a ride demand or request | $q \in$Q |
| $k_u^v$ | parking duration of $v$ | $k_u^v \in \mathbb{R}_{(0,+\infty)}$ |
| $\bar{k}_u$ | mean duration of type $u$ | |
| $d_{u,t}^v$ | the parking state of an SAV $v$ at $t$ | $d_{u,t}^v = 1$ or 0 |
| $x_{v,p}^{u,t}$ | allocation decision $p$ to $v$ | $x_v^{u,t} = 1$ or 0 |
| $a_{u,t}^v$ | suggested action by controller | $a_{u,t}^v \in \mathbb{R}_{[-1,1]}$ |
| $s_{u,t}$ | occupation rate of type $u$ | $s_{u,t}^v \subset \mathrm{S}_t$ |
| $s_t^{emp}$ | empty ratio of curb space | $s_t^{emp} \subset \mathrm{S}_t$ |
| $r_t$ | cumulative reward at $t$ | $r_t \in \mathbb{R}_{[0,1]}$ |
| $rc_t$ | parking area occupancy rate at $t$ | $rc_t \in \mathbb{R}_{[0,1]}$ |
| $rs_t$ | parking demand service rate at $t$ | $rs_t \in \mathbb{R}_{[0,1]}$ |
| $lat, lon$ | coordinate of location | $lat, lon \in \mathbb{R}$ |

### B. Problem Formulation

The modelling objective approximates optimum sequential policies $\pi^* = \{\tilde{\mu}_t(\cdot)|t \in TR\}$, to improve curbside occupancy ($rc$) and service rate ($rs$). The optimum decision ($x_{u,t}^{v,p}$) addresses whether to allocate a space $p \in \mathrm{P}_t$ to a vehicle, where $P_t$ represents a real-time set of available parking areas.

Objective Function (1) maximises the expectation of future reward ($r_t$). It equals the discounted ($\gamma^t$) sum of both immediate rewards: $rc_t$ and $rs_t$. Besides, we amplify this cumulative return by $\phi$ =100 times for the convenience of numerical analysis. Equation (2) demonstrates that parking occupancy equals the proportion of occupied and reserved areas to total supply |P|. Equation (3) calculates the ratio of total accepted requests to all demands |Q|.

$$\max_{x_{u,t}^{v,p}} \sum_{t=t'} r_t = \phi \gamma^t (rc_t + rs_t) \qquad \forall t \in TR \tag{1}$$

$$rc_t = \frac{\sum_u \sum_v \sum_p x_{u,t}^{v,p}}{|P|} \qquad \forall u \in U, v \in VP_t, p \in P_t \tag{2}$$

$$rs_t = \frac{\sum_u \sum_v \sum_p x_{u,t}^{v,p}}{|Q|} \qquad \forall u \in U, v \in VP_t, p \in P_t \tag{3}$$

Constraint (4) expresses the binary parking request state $d_{u,t}^v$ of an SAV $v \in$ VD. $d_{u,t}^v = 1$ indicates a parking request is accepted; Otherwise, $d_{u,t}^v = 0$. Meanwhile, Constraint (5) regulates that the length of staying equals its designated duration.

$$s.t. \quad d_{u,t}^v = \begin{cases} 1 & if \quad (\sum_p x_{u,t}^{v,p} \equiv 1) \cap (v \in VD) \\ 0 & if \quad (\sum_p x_{u,t}^{v,p} \equiv 0) \cup (v \in V - VD) \end{cases} \tag{4}$$

$$k_u^v = \sum_{t=t'} d_{u,t}^v \qquad \forall u \in U, v \in VD, t \in TR \tag{5}$$

Constraint (6) elaborates the conditions of the binary decision variable $x_{u,t}^{v,p}$. If an action $a_{u,t}^{v,p} \in \mathbb{R}_{[-1,1]}$ suggested by the controller is less than 0, then $x_{u,t}^{v,p} = 0$; Otherwise, $x_{u,t}^{v,p} = 1$. Constraints (7) further explain the estimation process of $a_{u,t}^{v,p}$. Namely, in conditions when no space remains for any request, or the parking area is not an optimum choice for a demand, $a_{u,t}^{v,p}$ takes a random value from a negative sigmoid function. In contrary conditions, the controller calculates this action value using a policy function $\mu_t(\mathrm{S}_t)$ controlled by observed states $\mathrm{S}_t$. Note that $\mu_t(\cdot)$ can be smaller than 0, meaning that the chance of declining a parking request still exists.

$$x_{u,t}^{v,p} = \begin{cases} 0 & if \quad a_{u,t}^{v,p} < 0 \\ 1 & if \quad a_{u,t}^{v,p} \geq 0 \end{cases} \tag{6}$$

$$a_{u,t}^{v,p} = \begin{cases} -\frac{1}{1+e^{-b}} & if \quad (\sum_p x_{u,t}^{v,p} \equiv |P|) \cup (p \neq p^*) \\ \mu_t(s_t) \in \mathbb{R}_{[-1,1]} & otherwise \end{cases} \tag{7}$$

Constraints (8)-(10) express the system states that been observed by the controller. It includes the individual occupancy rate of three types of parking activities $s_{u,t}$, the mean parking duration $\bar{k}_u$, and the ratio of empty areas $s_t^{emp}$.

$$s_t = \{s_{u,t}, \bar{k}_u, s_t^{emp} | u \in U\} \qquad \forall v \in VP_t, t \in TR \tag{8}$$

$$s_{u,t} = \frac{\sum_v \sum_p x_{u,t}^{v,p}}{|P|} \qquad \forall v \in VP_t, u \in U, t \in TR \tag{9}$$

$$s_t^{emp} = 1 - \frac{\sum_u \sum_v \sum_p x_{u,t}^{v,p}}{|P|} \qquad \forall v \in VP_t, u \in U, t \in TR \tag{10}$$

Following an accepted request, a greedy algorithm matches an optimum parking area $p^* \in \mathrm{P}_t$ to a demand $q$. Namely, as per Constraint (11), the controller estimates distances to all available parking areas and selects one with the minimal Euclidean distance to accommodate this demand. Note that $lat$ and $lon$ represent the coordinate of locations.

$$p^* = \arg\min_{p \in P_t} \sqrt{(lat_p - lat_q)^2 + (lon_p - lon_q)^2} \tag{11}$$

In agreement with peer studies [14], [17], the attributes of microscopic traffic simulation are defined as follows. Regarding SAVs, acceleration is $4.5m/s^2$; deceleration is $4.5m/s^2$; vehicular length is $4m$; the maximum speed is $30km/h$; the minimum longitudinal gap is $0.5m$. For pedestrians, the length-width is 0.48-0.21$m$; the minimal gap is $0.25m$; the speed factor is 1 and the maximum speed is $1.35m/s$.

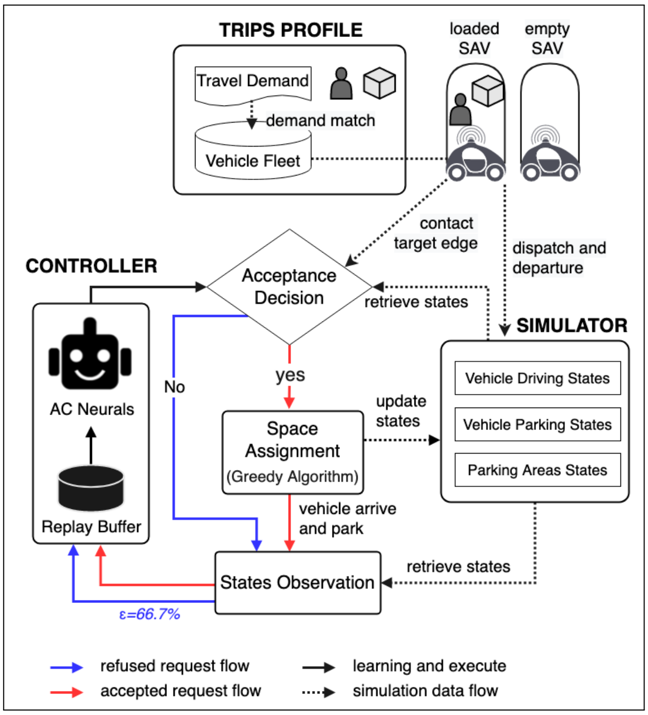


Fig. 2. Reinforcement Learning Modelling Framework

**Algorithm 1:** Pseudocode of DDPG Algorithm

**Input:** $E$, $B$, $M$, $S_0$, $\eta$, $d^v_{u,t'}$, $\varepsilon$

Initialise $\theta^Q, \theta^\mu, \theta^{Q'} \leftarrow \theta^Q$, $\theta^{\mu'} \leftarrow \theta^\mu$, $\mathscr{R} \leftarrow B$, $\mathscr{T} \leftarrow M$, $S_0$ ;

**for** $e = 0 : E-1$ **do**
- **for** $t = 0 : T-1$ **do**
  - $a_{t,e} \leftarrow \theta^\mu(s_{t,e})$, as per Eq(7);
  - $S_{t+1,e}$, $r_{t,e} \leftarrow TraCI(a_{t,e})$ ;
  - **if** $d^v_{u,t'} = 1$ **then**
    - $\mathscr{R} := < S_{t,e},\ a_{t,e},\ r_{t,e},\ S_{t+1,e} >$ ;
  - **else**
    - $\mathscr{R} := \mathbb{P}(< S_{t,e},\ a_{t,e},\ r_{t,e},\ S_{t+1,e} > |\varepsilon)$ ;
  - **while** $m \in \mathbb{N}_{[0,M]}$ **do**
    - $< S_{t,e},\ a_{t,e},\ r_{t,e},\ S_{t+1,e} > \leftarrow \mathscr{R}$ ;
    - $\mathscr{T} := < a_t,\ s_t,\ r_t,\ s_{t+1} >$ ;
  - **while** $m \in \mathbb{N}_{[0,M]}$ **do**
    - $y_m \leftarrow \theta^{Q'}$, as per Eq (12)(13);
  - $\theta^Q, \theta^{Q'}, \theta^\mu, \theta^{\mu'} \leftarrow \eta, y_m, \theta^Q, \theta^\mu$, as per Eq(14)(18);

### C. Reinforcement Learning Model

We introduced a partially learning Deep Deterministic Policy Gradient (DDPG) algorithm to approximate optimum solutions. Fig.2 demonstrates our proposed RL modelling framework and Algorithm.1 specifies concrete learning procedure. This algorithm enables a controller to store its past decisions for off-policy learning discriminatingly. Namely, those accepted decisions are integrally stored and learned. In contrast, declined requests are learned provided a drop rate $\varepsilon$=66.7%, meaning that only a third of these experiences are replayed later. We implemented the Simulation of Urban MObility (SUMO) [18] for simulation and retrieving observed states.

DDPG algorithm records past experience using a data cache, also named the replay buffer $\mathscr{R}$ (sized $B$ =99,999). It samples a mini-batch (sized $M$ =64) of transition tuple $\mathscr{T}$=<$S_m$, $a_m$, $r_m$, $S_{m+1} > \forall m \in \mathbb{R}_{[0,M-1]}$, to reproduce plenty of conditions for learning. In our model, the soft update coefficient $\eta = 0.005$ and the discount factor $\gamma$ is 0.99.

DDPG algorithm implements a dual Actor-Critic (AC) structure which consists of four neural networks: an online critic network ($\theta^Q$), an online actor network ($\theta^\mu$), a target critic network ($\theta^{Q'}$), and a target actor network ($\theta^{\mu'}$). Hereby, $\theta$ denotes their neural network weights.

DDPG follows the Bellman Equation and soft update mechanism for updating its critic network weights. In the first instance, the online critic network estimates the approximator ($y_m$) of the Q-value as per Equations (12)(13). If the learning sampling is not the terminate step, this approximator equals the sum of reward and discounted future Q-value using the target policy function. Otherwise, $y_m$ only considers the reward.

Let $\mathscr{L}$ indicates the loss between the approximator and the online Q-value. Equation (14) calculates the Mean Squared Bellman Error (MSBE) of all samples with respect to the distribution of $\mu$. Meanwhile, the target network updates to new weights controlled by $\eta$, as per Equation (15).

$$y_m = \begin{cases} r_m & if \quad s_{m=T-1} \\ r_m + r'_m & otherwise \end{cases} \tag{12}$$

$$r'_m = \gamma Q[s_{m+1}, \mu(s_{m+1}|\theta^{\mu'})|\theta^{Q'}] \qquad m \in \mathbb{R}_{[0,M-1]} \tag{13}$$

$$\mathscr{L} \approx \mathop{\mathbb{E}}_m \left(y_m - Q(s_m, a_m|\theta^Q)\right)^2 \qquad m \in \mathbb{R}_{[0,M-1]} \tag{14}$$

$$\theta^{Q'} := \eta\theta^Q + (1-\eta)\theta^{Q'} \qquad \eta \in \mathbb{R}_{(0,1)} \tag{15}$$

The online actor network executes an action following its deterministic policy function. Let $J(\theta^\mu)$ denotes the actor loss of $\theta^\mu$. The network updates its weight by applying a policy gradient method, following Equation (16)(17). The target actor network updates its parameters as per Equation (18), of which the mechanism is identical to the target critic network.

$$J(\theta^\mu) = \mathop{\mathbb{E}}_m\{Q(S_m, a_m|\theta^\mu)\} \qquad m \in \mathbb{R}_{[0,M-1]} \tag{16}$$

$$\nabla_{\theta^\mu} J \approx \mathop{\mathbb{E}}_m \left(Q(S_m, a_m|\theta^Q)\nabla_{\theta^\mu}\mu(S_m|\theta^\mu)\right) \tag{17}$$

$$\theta^{\mu'} := \eta\theta^\mu + (1-\eta)\theta^{\mu'} \qquad \eta \in \mathbb{R}_{(0,1)} \tag{18}$$

## IV. RESULTS AND DISCUSSION

This section presents experiment results under varied PUDO, docking and on-street parking demand patterns. Our RL model has been trained in $E$=1,500 epochs under each

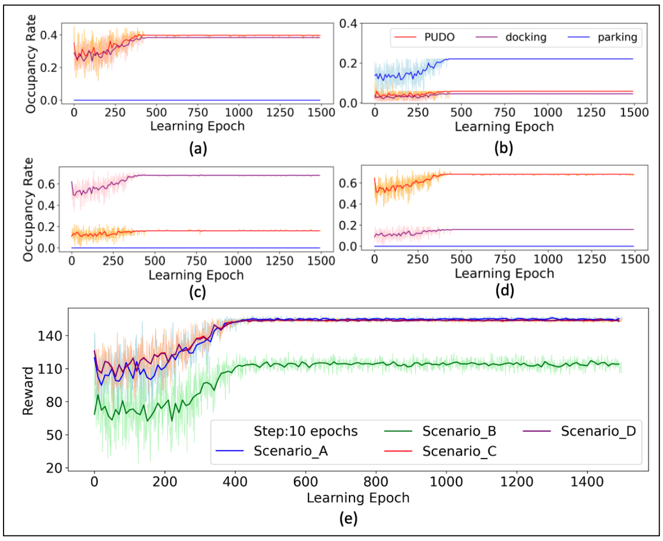


Fig. 3. Curbside Occupancy and Rewards of Three Parking Behaviours in Four Scenarios. (a) Scenario A, (b) Scenario B, (c) Scenario C, (d) Scenario D, (e) Rewards of Four Scenarios

condition. We compared the curbside performance and learning convergence patterns in these conditions. Then, optimum occupancy and service rate distributions under multiple (225) PUDOs and docking demands scenarios are analysed.

### *A. Curbside Performance*

We designed four demand scenarios to investigate the curbside performance. Fig.3 demonstrates changes in epoch-based and ten epochs-based (thick coloured lines) occupancy rates throughout the course of learning.

In Scenario A, PUDOs and docking activities equally dominate the request. The demand rates for PUDOs, docking and on-street parking demands are $40veh/h$, $40veh/h$ and $20veh/h$, respectively. It means that trade-offs are expected under such a saturated demand condition. The result shows that the allocation plan converges at a pattern of 40%:38%:0 by Epoch $436^{th}$. The optimum occupancy approximates 78%. PUDO occupancy rate is 2% marginally higher than that of docking, but significantly higher than on-street parking.

Contrary to that, Scenario B represents an on-street parking dominant case, with a demands setting of $5veh/h$, $5veh/h$ and $20veh/h$. This scenario may reflect residential streets where demands for long-term parking are high [19], while PUDOs or docking behaviours are sporadic. As a result, on-street parking events have been allocated 22% of curb space, whereas 6% for PUDOs and 5% for docking requests. Its total observed occupancy rate approximates 32%, which is 140% lower than Scenario A.

Scenario C simulates a saturated docking demands situation for frequent valet parking occasions. The demand rates are $20veh/h$, $80veh/h$ and $20veh/h$. After optimisation, PUDO accounts for 16%, which is secondary to its docking counterpart. In contrast, on-street parking requests fail to compete for a space.

Scenario D considers busy urban centre streets where PUDO demands exceed the rest. However, on-street parking is suppressed due to the high parking price [9]. The designated demand rates are $80veh/h$, $20veh/h$ and $5veh/h$. Result reveals that PUDO dominates 68% of curbside use, and the corresponding optimum strategy converges at Epoch $431^{st}$, which is identical to Scenario C.

These findings imply that our model prioritises PUDO and docking behaviours over on-street parking requests. When short-term parking events are insufficient, like Scenario B, raising on-street parking demands has limited contribution to curbside occupancy. This outcome suggests that the model is expecting more PUDOs or docking requests for potentially higher rewards in both occupancy and service rates.

### *B. Learning Performance of DDPG Algorithm*

Fig.3e presents the learning performance of our model under these four testing scenarios. Y-axis indicates reward ($r$), while the x-axis indexes learning epochs.

All four learning curves demonstrate satisfying convergence guarantees at Epoch $431^{st}$-$439^{th}$. After a shared initial declining tendency, rewards in conditions climb and converge at respective optimal of 158.28, 128.96, 156.62 and 157.31. Amongst, Scenario A obtains the highest optimum reward. The sheer increases in total rewards equal 38.27, 60.37, 31.01 and 31.13, where Scenario B has the most considerable improvement. The fastest increase occurs between Epoch $200^{th}$ to $410^{th}$. Findings indicate that the overflow rate of cumulative short-term parking demand seemingly remains positively correlated with the optimum reward.

Results also reveal that our model prioritise the allocation of short-term parking behaviours for higher service level (77%) regarding total demand requests whilst keeping a high occupancy level (83%). Furthermore, it well explains why providing saturated on-street parking demands still fails to guarantee a high occupancy rate. In such a situation, our model prefers to consume less curb space (51%) while satisfying an equivalent level of parking requests (74%). These findings complement the knowledge on curb parking regarding ideal occupancy rate [12]. That is, maintaining 20%-40% of curb space flexibility may not be efficient for achieving higher curbside performance, particularly in PUDO and docking saturated conditions.

### *C. Distribution of Optimum Parking Service Rate*

We further explore the correlations between multiple short-term parking demands and curbside performance given a fixed parking capacity. To this end, a demand matrix is established, which is composed of $10veh/h$-$150veh/h$ PUDO and $10veh/h$-$150veh/h$ docking requests. Meanwhile, the on-street parking demand rate remains at $5veh/h$.

Fig.4a demonstrates the distribution of optimum curbside occupancy ($rc$). This occupancy rate significantly increases while raising demands level. The minimum occupancy is 33% at the demands of ($20veh/h$, $10veh/h$), while the maximum equals 98% at ($120veh/h$, $140veh/h$). As the demands rate to supply approximates 4:1, this occupancy tends to exceed 80%. In these saturated demand conditions, raising demands is less likely to keep an equivalent ascending

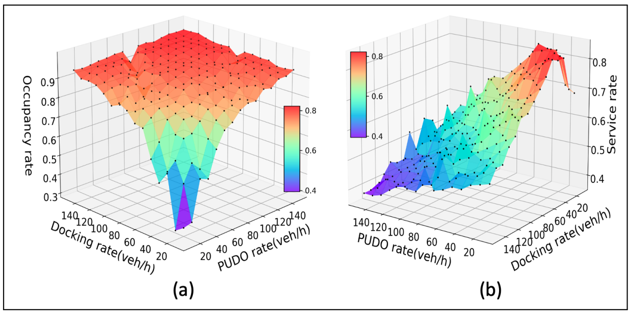


Fig. 4. Distribution of Optimum Occupancy and Service Rate under Multiple Demand Rates Conditions. (a) Distribution of occupancy rates (rc). (b) Distribution of parking service rates (rs).

gradient in the parking areas consumption. In other words, additional parking sites are needed to accommodate excessive parking needs.

Fig.4b shows the distribution of service rate ($rs$). As the demand rate increases to $40veh/h$-$80veh/h$, the optimum service rate reaches around 83%. Afterwards, the value generally declines to 38%. Findings further reveal that when the demand-supply ratio falls below 2:1, curbside parking demand is under-saturated. In these conditions, our model would accept on-street parking events for the sake of a basic occupancy. Consequently, 20% of requests may be re-directed to other curbsides. When such a ratio situates between 2:1 and 4:1, the model receives the maximum service rate. When this ratio passes a threshold of 4:1, the service rate non-monotonically drops as spillover expands to PUDOs and docking events.

## V. CONCLUDING REMARKS

We formulated the real-time allocation over curbside space as a DP problem. Namely, it focused on sequential dispatching decisions in a finite time horizon, serving a continuous stream of hybrid parking behaviours. The objectives maximised the accumulative parking service rate and curbside occupancy. Furthermore, we applied the RL method and proposed a partially learning DDPG algorithm-embedded controller to learn optimum allocation strategies. This controller interacted with a gym environment in SUMO for partially observable parking states retrieval.

Evaluated under multiple scenarios, results reveal good convergence performances and robust learning patterns. The model learned to prioritise PUDOs and docking requests for a higher parking service rate. However, in an under-saturated parking condition, where the demand-supply ratio is less than 2:1, our model would exploit on-street parking requests to maximise curbside occupancy. The optimum range is 2:1 to 4:1, where parking service rate reaches a maximal (83%) and spillover of PUDOs or docking SAVs approximates zero. When this ratio passes beyond 4:1, curbside occupancy rises slowly, and the service rate declines.

Our model provides a feasible curbside parking dispatching method to the increasingly diversifying parking demands in the AVs era. This work also sheds light on deploying distinctive curb parking policies to streets in different contexts. Future works will extend this current model to real-world cases-based simulations and incorporate the current controller with real-time parking demand flows.

## REFERENCES

[1] G. Marsden, I. Docherty, and R. Dowling, "Parking futures: Curbside management in the era of 'new mobility'services in british and australian cities," *Land Use Policy*, vol. 91, p. 104012, 2020.
[2] NACTO, "Blueprint for autonomous urbanism," 2017, accessed 31-Aug-2021. [Online]. Available: https://nacto.org/publication/bau2
[3] M. Chester, A. Fraser, J. Matute, C. Flower, and R. Pendyala, "Parking infrastructure: A constraint on or opportunity for urban redevelopment? a study of los angeles county parking supply and growth," *Journal of the American Planning Association*, vol. 81, no. 4, pp. 268–286, 2015.
[4] E. P. Association, "Scope of parking in europe," 2013, accessed 27-Feb-2022.
[5] A. Ranjbari, J. Luis Machado-León, G. Dalla Chiara, D. MacKenzie, and A. Goodchild, "Testing curbside management strategies to mitigate the impacts of ridesourcing services on traffic," *Transportation Research Record*, vol. 2675, no. 2, pp. 219–232, 2021.
[6] S. S. Dey, B. O. Pérez, L. Richards, A. Pochowski, M. Sanders, M. Darst, E. C. Sanchez, and S. Dock, "Demystifying urban curbside freight management: Strategic incremental approach from washington, dc," *Transportation Research Record*, vol. 2673, no. 12, pp. 312–326, 2019.
[7] E. González-González, S. Nogués, and D. Stead, "Automated vehicles and the city of tomorrow: A backcasting approach," *Cities*, vol. 94, pp. 153–160, 2019.
[8] A. Millard-Ball, "The autonomous vehicle parking problem," *Transport Policy*, vol. 75, pp. 99–108, 2019.
[9] D. C. Shoup, *The high cost of free parking*. Routledge, 2021.
[10] Q. Ye, S. M. Stebbins, Y. Feng, E. Candela, M. Stettler, and P. Angeloudis, "Intelligent management of on-street parking provision for the autonomous vehicles era," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–7.
[11] K. Pyzyk, "Curbflow pilot reduced double parking in dc by 64 per cent," 2019, accessed 26-Feb-2022. [Online]. Available: https://www.smartcitiesdive.com/news/curbflow-pilot-reduced-double-parking-in-dc-by-64/567268/
[12] G. Pierce and D. Shoup, "Getting the prices right: an evaluation of pricing parking by demand in san francisco," *Journal of the american planning association*, vol. 79, no. 1, pp. 67–81, 2013.
[13] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
[14] Q. Ye, Y. Feng, E. Candela, J. Escribano Macias, M. Stettler, and P. Angeloudis, "Spatial-temporal flows-adaptive street layout control using reinforcement learning," *Sustainability*, vol. 14, no. 1, p. 107, 2022.
[15] X. Zhang, C. Zhao, F. Liao, X. Li, and Y. Du, "Online parking assignment in an environment of partially connected vehicles: A multi-agent deep reinforcement learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 138, p. 103624, 2022.
[16] M. Ratli, A. A. El Cadi, B. Jarboui, and A. Artiba, "Dynamic assignment problem of parking slots," in *2019 International Conference on Industrial Engineering and Systems Management (IESM)*. IEEE, 2019, pp. 1–6.
[17] Q. Lu, T. Tettamanti, D. Hörcher, and I. Varga, "The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation," *Transportation Letters*, vol. 12, no. 8, pp. 540–549, 2020.
[18] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE, 2018. [Online]. Available: https://elib.dlr.de/124092/
[19] W. Zhang, S. Guhathakurta, J. Fang, and G. Zhang, "Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach," *Sustainable Cities and Society*, vol. 19, pp. 34–45, 2015.