

Research Article

Adaptive Coordinated Variable Speed Limit between Highway Mainline and On-Ramp with Deep Reinforcement Learning

Ming Cheng ¹, Chenghao Zhang ¹, Hui Jin ¹, Ziming Wang ¹ and Xiaoguang Yang ²

¹School of Rail Transportation, Soochow University, 215031, China

²Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 200092, China

Correspondence should be addressed to Hui Jin; jinh@suda.edu.cn

Received 3 February 2022; Revised 17 July 2022; Accepted 16 September 2022; Published 29 September 2022

Academic Editor: Guohui Zhang

Copyright © 2022 Ming Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Highway merging bottleneck is challenged with serious traffic conflicts between on-ramp and mainline vehicles, causing significant capacity drop and drastic speed changes. The paper proposes an adaptive coordinated variable speed limit model to manage highway speed of on-ramp and mainline continuous sections without priority to mainline. That helps to remove the speed difference between the vehicles from on-ramp and mainline flooding into the merging zone, and to sustain actual traffic density close to critical density to counteract capacity drop as indicated with macroscopic fundamental diagram. The method of deep reinforcement learning based on deep deterministic policy gradient is employed to solve the proposed model with a row of continuous control variables. Simulation platform with VISSIM 5.3 is established, and the proposed method can enhance traffic flow through the merging zone by around 10% and 19% under static and dynamic demand, respectively, in addition to reduced density and speed variation by around 30%. This research provides insights into the management of highway capacity so as to secure traffic efficiency and reliability for the merging zone.

1. Introduction

Highway congestion has been increasingly emphasized to prevent or alleviate its detrimental effect on traffic mobility, safety, and environment [1]. Highway congestion is mostly attributed to the merging or diverging around the on-ramps and off-ramps, traffic speed of which can be much lower than that of the mainline. Thus drastic speed changes can be caused, in addition to aggressive lane changing and weaving. That hinders the traffic flow on highway mainline, and causes recurrent highway bottlenecks. Considering that off-ramp are always closely related to the connected surface road, off-ramp bottleneck can be more complex [2]. Therefore, this research targets at on-ramp bottleneck, where vehicles frequently decelerate or stop to wait before squeezing onto the mainline.

Traffic of mainline and on-ramp can be coordinated with the combined strategies of variable speed limit (VSL) and ramp-metering so as to adjust highway flow into the merging zone. VSL issues dynamic speed limit with variable

signs, depending on traffic conditions [3, 4]. Ramp-metering adjusts vehicle flow rate such that mainline density remains below the critical value to prevent traffic breakdown [5]. For example, VSL and ramp-metering are combined into model predictive control by incorporating dynamic speed limits on the mainline into the original on-ramp model of METANET [6], so as to address ramp-metering disadvantage of failing to prevent congestion upstream of the merging zone [7]. Papamichai et al. [8] combines VSL with ramp-metering in a discrete optimal control programming, where excessive on-ramp queue is treated with a penalty added to the objective. Assuming predetermined ramp-metering rate for on-ramp, a model predictive control is developed for VSL over a finite time horizon to address the recurrent highway on-ramp bottleneck, which is found to enhance traffic flow by 12.8% and reduce total travel time by 31.8% [9]. Then driver acceptance is taken into the coordination between VSL and ramp-metering upstream of and at the merging zone, which enhances bottleneck throughput significantly [10]. Recent research also attempts to introduce the connected and

automated vehicles into the traffic control and management of highway merging zone, where microscopic vehicular trajectory is coordinated to enhance traffic efficiency [11, 12].

Despite rich research on the coordination of mainline VSL and ramp-metering, the existing study generally assigns priority to mainline at the cost of delaying or stopping on-ramp vehicles. Such management strategy has been found to account for nearly half of highway accidents at the congested merging zone, where rear-end crashes are most frequent due to drastic deceleration [13]. In addition, sideswipe and angle highway collisions are also found to be significantly related to the speed difference between highway mainline and on-ramp [14]. Statistical analysis has validated that increasing ramp vehicle volume assists in reducing crashes, while increasing mainline vehicle volume tends to increase crashes. Thus, it is important to enhance ramp traffic efficiency [15]. That is consistent to the early empirical research that suggests to remove the criteria of setting ramp design speed to be 50th percentile of freeway design speed, so as to allow more freedom to ramp vehicles [16].

Therefore, the research is encouraged to apply VSL to both on-ramp and mainline for their homogenization, instead of suppressing on-ramp and prioritizing mainline traffic. This highway management strategy is expected to alleviate the speed difference between mainline and on-ramp traffic, reducing on-ramp vehicle stops and shockwaves at the merging zone. That helps to maintain highway capacity and to reduce traffic fluctuations. To tune finely the variable speed limit, a set of continuous decision variables should be solved along the time horizon. Thus an efficient solution method is called for that is capable of finding the optimal control scheme under traffic dynamics.

Recently, artificial intelligence has attracted increasing attention, where intelligent agents are defined to perceive environment, predict system evolution, and take proactive actions for maximal reward [17]. Advanced application of artificial intelligence includes the improvement of object tracking against occlusion [18], image inpainting against structure disconnecting with improved total variation minimization [19] or structural disorder with multilevel attention progression mechanism [20], and multiscale superresolution image with feature map attention mechanism [21]. When it comes to highway management, reinforcement learning is widely adopted for complex modelling solution. For example, a reinforcement learning based VSL control is proposed to reduce total travel time near a freeway recurrent bottleneck [22] and to smooth vehicle conflicts for reduced crashes [23]. Single agent Q learning was implemented on ramp-metering, which is found to reduce total network travel time by 17% compared to ALINEA algorithm [24]. Q-learning was also applied to VSL optimization to reduce total travel time of both mainline and on-ramp by 49.3% and 21.8% under stable and dynamic traffic demand, respectively [22]. Further, deep Q-learning is applied to VSL of highway mainline with on-ramp bottlenecks, reducing total travel time by 26% to 67% with stable demand, and 21% to 70% with dynamic demand [25]. To select continuous actions instead of discrete ones, deep deterministic policy gradient (DDPG) is proposed to learn competitive policies

[26]. For example, updated research has applied DDPG learning model to VSL control against spatially dynamic speed limit zones based on vehicle position and speed [27].

Therefore, research gap is identified on highway merging zone management, where previous research assigns priority to mainline over on-ramps, causing traffic shocks and conflicts upstream of the bottleneck area. To this end, the research proposes an adaptive coordinated VSL (ACVSL) to explore the potential of applying VSL to highway mainline and on-ramp, guiding vehicles of both to adjust to the same speed before the merging zone [28]. The proposed speed limit is deduced from fundamental diagram with traffic density close to the critical density that corresponds to road capacity. Thus capacity drop at the merging zone can be reduced under the objective of maximal vehicle throughput, to which the emerging solution method of deep reinforcement learning is employed for efficient control scheme.

For this paper, the main contributions are as follows. (1) a novel ramp management strategy is proposed to remove mainline priority over on-ramp, where both are managed with VSL in an adaptive coordinated way. Thus equal speed can be achieved between mainline and on-ramp immediately upstream of the merging zone. That helps to alleviate traffic shockwaves at the merging bottleneck and to anticipate capacity drop. (2) Error state is structured to reflect the real-time difference between actual and critical traffic density, based on which speed control parameters are developed to update VSL of both mainline and on-ramp. Thus traffic density at the merging zone can be stabilized around the critical density to counteract road capacity drop. (3) DDPG-based deep reinforcement learning is established to solve the proposed model with continuous control parameters, which is calibrated and validated with simulation platform for enhanced vehicle throughput and alleviated speed fluctuations.

The remaining of the paper is organized as follows. Section 2 establishes ACVSL model for the continuous highway sections of mainline and on-ramp till the merging bottleneck. It develops nonlinear feedback to narrow the difference between actual and critical traffic density to sustain bottleneck capacity. Section 3 develops the deep reinforcement learning method with DDPG to solve the proposed ACVSL under critic-actor framework. Simulation and tests follow in Section 4, validating that the proposed model and solution method can significantly enhance the efficiency of highway merging. Section 5 briefly concludes the paper, and points out future research directions.

2. Mathematical Modelling

Figure 1 shows a typical highway section with on-ramp. The mainline and on-ramp are divided into N^m and N^r sections, respectively. Considering the limited ramp length, we set $N^r = 2$. VSL on the upstream sections controls the vehicles getting into the next section, and finally into the merging zone.

Objective of the control model is given by:

$$\max \sum_t q_d(t) = \sum_t \sum_x q_{N^x}^x(t), \quad (1)$$

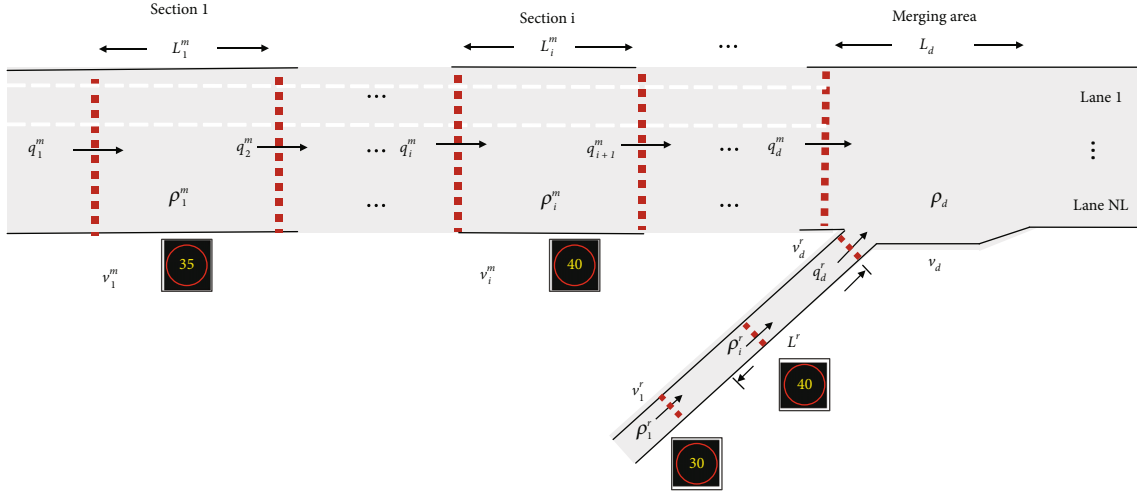


FIGURE 1: Configuration of highway mainline and on-ramp.

so as to maximize vehicle throughput from the merging zone, and to retain highway capacity. Parameter $q_d(t)$ represents the total flow through the merging bottleneck at time t , and $q_{N^x}^x(t)$ means flow on the last section of x at time t , where $x = r$ and $x = m$ refer to the mainline and on-ramp, respectively.

As indicated by the studies of [2, 29], lower speed limit can increase the critical density to a higher value. Figure 2 shows that the speed lower than the free-flow speed (i.e. $0 < v_d < v_{d0}$) assists in shifting the critical density to a higher value (i.e. $0 < \rho_{d0} < \rho_d$) in the fundamental diagram without decreasing bottleneck capacity. Moreover, the lower speed causes the slope of the under-critical part to be decreased, making the fundamental diagram get closer to a straight line, i.e. the red dash line vs. green dash line on the left-side, to alleviate the speed deviation around the critical density and the resultant shockwaves. Therefore, we may design an adaptive coordinated VSL control on the continuous sections of highway mainline and on-ramp, letting the traffic density at the merging section converge to the critical density derived from fundamental diagram. Thus the region of fluent traffic for the merging zone can be extended in comparison to the no-control scenarios, helping to attenuate shockwaves, and to sustain bottleneck capacity.

To this end, we define an error state to describe the deviation of the actual traffic density from its critical value, given by:

$$e_i^x(t) = \rho_i^x(t) - \rho_{d,c} \forall x, i, \quad (2)$$

$$e_d(t) = \rho_d(t) - \rho_{d,c}. \quad (3)$$

For the sections upstream of and at the merging zone, respectively, at time step t on highway x . Parameter $\rho_i^x(t)$ means the actual traffic density on section i of road x at time step t . Parameter $\rho_d(t)$ represents the actual density of the merging zone at time step t , while $\rho_{d,c}$ describes the critical density between fluent and congested traffic flow, and is assumed to be constant.

Now we have developed an open-loop control system joining both mainline and on-ramp. In the following, the

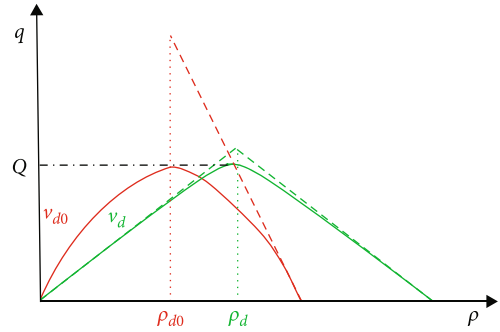


FIGURE 2: Effects of lower speed on fundamental diagram.

nonlinear feed-back mechanism for adaptive coordinated variable speed limit (ACVSL) is established.

Control parameters aim to adjust the guide speed on the VSL signs upstream of the merging zone, given by:

$$\lambda_i^x(t+1) = \frac{[q_{N^x}^x(t) - \bar{v}_i^x(t+1)\rho_i^x(t)]}{L_{i+1}^x e_{i+1}^x(t)} \forall x, i \leq N^x - 1, \quad (4)$$

so as to maintain the density of the merging zone around the critical value, and to counteract capacity drop. Notation $\lambda_i^x(t+1)$ represents the control parameter on section i of highway x at time step $(t+1)$, $q_{N^x}^x(t)$ describes the traffic flow of section N^x , $\bar{v}_i^x(t+1)$ means the intended guide speed that will be demonstrated on the VSL sign at time step $(t+1)$, and L_{i+1}^x represents the length of section $(i+1)$ on x . Therefore, on the right-side of Eq. (4), the numerator is the flow difference between section N^x and section i , and the denominator is the vehicle count in section $(i+1)$ required to fill its density gap from the critical density at time t of highway x [30]. Eq. (4) can be reformulated to represent the intended guide speed $\bar{v}_i^x(t+1)$, given by:

$$\bar{v}_i^x(t+1) = \frac{-\lambda_i^x(t+1)L_{i+1}^x e_{i+1}^x(t) + v_{N^x}^x(t)\rho_{N^x}^x(t)}{\rho_i^x(t)} \forall x, i \leq N^x - 1 \quad (5)$$

TABLE 1: ACVSL notation list.

No.	Notation	Explanation
1	N^x	Section count on highway part x , where $x = m$ and $x = r$ refers to mainline and on-ramp, respectively.
2	$q_d(t)$	Total flow through the merging bottleneck at time t
3	$q_{N^x}^x(t)$	Traffic flow of section N^x at time t
4	$e_i^x(t)$	Error state of section i on highway x upstream of the merging zone at time t
5	$\rho_i^x(t)$	Actual traffic density on section i of highway x at time t
6	$\rho_{d,c}$	Critical density between fluent and congested traffic flow
7	$e_d(t)$	Error state of merging zone at time step t
8	$\rho_d(t)$	Actual density of merging zone at time step t
9	$\lambda_i^x(t+1)$	Control parameter on section i of highway x at time step $(t+1)$
10	$\bar{v}_i^x(t+1)$	Intended speed on VSL sign for section i on highway x at time step $(t+1)$
11	L_{i+1}^x	Length of section $(i+1)$ on highway x
12	$v_{N^x}^x(t)$	Actual speed on section N^x at time step t
13	V_d	Theoretical speed of the merging zone from macroscopic fundamental diagram at critical density
14	$\bar{v}_i^x(t)$	Maximum speed limit demonstrated on VSL sign for section i of x at time step t
15	$v_i^x(t-1)$	Actual speed on section i of highway x at time step $(t-1)$
16	C_v	Maximum speed difference in VSL between successive time steps and sections
17	$V_i^x(t)$	Final speed delivered to VSL sign after considering the lower bound of speed limit for section i on highway x at time step t
18	v_{\max}	Upper bound of VSL
19	v_{\min}	Lower bound of VSL

Thus the gap between actual density and critical density can be alleviated to promote efficient and stable traffic flow. Therefore, once control parameter $\lambda_i^x(t+1)$ is determined, the intended guide speed $\bar{v}_i^x(t+1)$ can be calculated with Eq. (5).

When it comes to the last sections either on the mainline or on-ramp, VSL there is set the same as that of the merging zone to secure smooth convergence, given by:

$$\bar{v}_{N^x}^x(t) = V_d \forall x, \quad (6)$$

where V_d means the theoretical speed of the merging zone from macroscopic fundamental diagram at critical density [30]. Thus speed conflicts can be alleviated between highway mainline and on-ramp immediately upstream of the merging zone.

Additionally, to avoid dynamic oscillation in speed limit, the speed design by Zhang and Ioannou [31] is borrowed, given by:

$$\bar{v}_i^x(t) = \min \{ [\bar{v}_i^x(t)]_5, v_i^x(t-1) + C_v, v_{i-1}^x(t) + C_v, v_{\max} \} \forall x, i \leq N^x - 1, \quad (7)$$

$$V_i^x(t) = \max \{ \bar{v}_i^x(t), v_{\min}, v_i^x(t-1) - C_v, v_{i-1}^x(t) - C_v \} \forall x, i \leq N^x - 1, \quad (8)$$

where $\bar{v}_i^x(t)$ is the maximum speed limit that can be demonstrated on VSL sign for section i of highway x at step t . Parameter $[\bar{v}_i^x(t)]_5$ rounds $\bar{v}_i^x(t)$ to its closest multiple of 5, C_v represents the maximum speed difference in VSL

between successive control steps and highway sections; v_{\max} is the upper bound of VSL; and $V_i^x(t)$ returns the final speed delivered to VSL sign for section i of highway x after considering the lower bound of speed limit, with v_{\min} being the lower bound of VSL.

Thus the optimization model ACVSL is summarized with the object of Equation (1) and constraints of Equations (2)–(8), by optimizing a set of control parameters $\lambda_i^x(t)$. Table 1 summarizes the notation of the proposed model.

3. Solution Method

In this section, we target at optimizing the vector of control parameters $\lambda_i^x(t)$ from Eq. (4), a set of multi-dimensional continuous variables. The control parameters $\lambda_i^x(t)$ indicate the ratio of the flow difference to the gap between current density and critical density for section i of highway x at time t , determining the VSL values in turn as in Eq. (5), so as to adjust the count of vehicles moving downward. Thus we may coordinate VSL of both mainline and on-ramp so as to adapt to real-time traffic state with optimized control parameters $\lambda_i^x(t)$, sustaining critical traffic density that is beneficial to bottleneck flow.

DDPG-based deep reinforcement learning method first collects traffic state $s(t)$, given by:

$$s(t) = [\mathbf{v}(t), \boldsymbol{\rho}(t)], \quad (9)$$

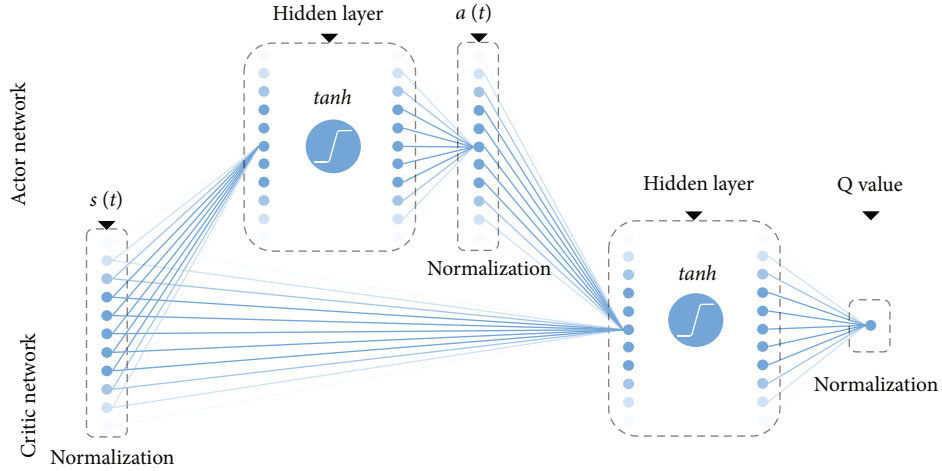


FIGURE 3: Schematic illustration of the critic and actor network.

```

Initialize critic network  $Q(s, a|\theta^Q)$  and actor network  $\pi(s|\theta^\pi)$ 
Initialize target critic network  $Q(s, a|\theta^{Q'})$  and target network  $\pi(s|\theta^{\pi'})$ 
Initialize prioritized experience replay memory  $R$ 
While not converge do
  Observe system state  $s(t)$ 
  Select action  $a(t) = \mu(s|\theta^\mu) + N$  for the current state  $s(t)$  following  $\epsilon$ -greedy exploration policy
  Observe transition pair  $[s(t), a(t), s(t+1), r(t)]$ , and store in  $R$ 
  Sample a mini-batch  $[s(t), a(t), s(t+1), r(t)]$  of size  $N$  from  $R$ 
  Set  $y(i) = r(i) + \gamma Q^i\{s(i+1), \pi[s(i+1)|\theta^{\pi'}]\theta^Q\}$ 
  Update critic network by minimizing:
     $\sum_i [y(i) - Q(s(i+1), s(i)|\theta^Q)]^2$ 
  Update actor with the sampled policy gradient:
     $\nabla_{\theta^\pi} J \sim 1/N \sum_i \nabla_a Q(s, a|\theta^Q)|_{s=s(i), a=\pi[s(i)]} \nabla_{\theta^\pi} \pi(s|\theta^\pi)|_{s(i)}$ 
  Update target network
     $\theta^{\pi'} \leftarrow \tau \theta^\pi + (1 - \tau) \cdot \theta^{\pi'}$ 
     $\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \cdot \theta^{Q'}$ 

```

ALGORITHM 1: DDPG Algorithm.

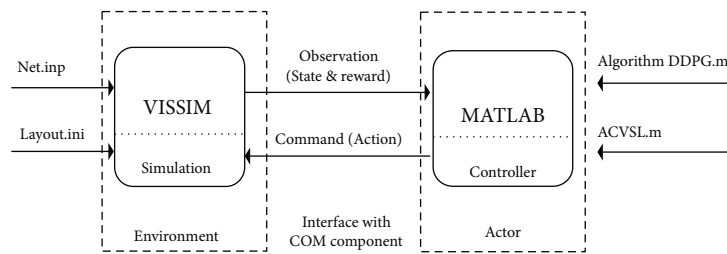


FIGURE 4: The framework of ACVSL control platform.

where $\mathbf{v}(t)$ and $\boldsymbol{\rho}(t)$ refer to the vectors of traffic speed and density over highway sections at time t , which are of different physical units and normalized for effective learning, given by:

$$\mathbf{v}(t) = \frac{[v_1^x(t), v_2^x(t), \dots, v_{N^x-1}^x(t), v_d(t)]}{v_f^x}, \quad (10)$$

$$\boldsymbol{\rho}(t) = \frac{[\rho_1^x(t), \rho_2^x(t), \dots, \rho_{N^x-1}^x(t), \rho_d(t)]}{\rho_{jam}^x}, \quad (11)$$

with v_f^x and ρ_{jam}^x being free flow speed and jam density on highway section x , respectively.

The proposed method entails the training of a software-controlled agent to take action $a(t)$ in response to the current system state $s(t)$ for reward $r(t)$. Action sets the value

TABLE 2: Scenario parameters.

No.	Notation	Value	No.	Notation	Value	No.	Notation	Value
1	L_1^m	500 m	6	L_1^r	100 m	11	ρ_{jam}^r	144 veh/km/ln
2	L_2^m	300 m	7	L_2^r	100 m	12	λ_{max}	100
3	L_3^m	200 m	8	v_f^m	120 km/h	13	ρ_d	80 veh/km/ln
4	L_4^m	200 m	9	$v_f^r=v_{max}$	80 km/h	14	v_{min}	20 km/h
5	L_d	100 m	10	ρ_{jam}^m	144 veh/km/ln	15	C_v	Km/h

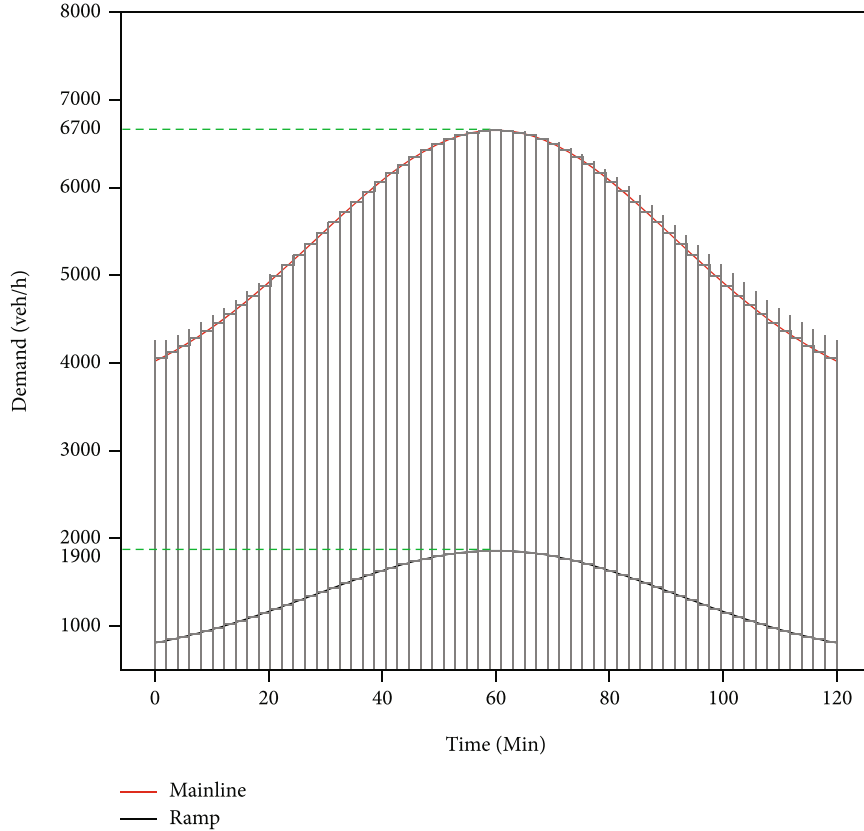


FIGURE 5: Dynamic flow demand for DDPG exploration and training.

of the control parameter $\lambda_i^x(t)$ at the beginning of each control time step, given by:

$$a(t) = \frac{[\lambda_1^x(t), \lambda_2^x(t), \dots, \lambda_{N^x}^x(t)]}{\lambda_{max}}, \quad (12)$$

where λ_{max} means the upper bound of control variables for normalization to secure the action variables in the interval $[0, 1]$, a necessity for the input of neural network. Reward function aims to maximize highway bottleneck outflow, given by:

$$r(t) = \int_{cT}^{(c+1)T} \frac{q_d(t)}{QT} dt, \quad (13)$$

which is divided by the multiplication of Q and T for normalization, with Q and T being bottleneck capacity and step interval length, respectively. Reward at each step is achieved

via actor network, which actually refers to control policy π , a function that maps system state to probability distribution with control actions. Critic network then develops a corresponding Q-function to represent the accumulated discounted reward if action $a(t)$ is taken for state $s(t)$ and strategy π is followed at step $(t+1)$ onwards, given by:

$$Q^\pi[s(t), a(t)] = E\langle r(t) + \gamma Q^\pi\{s(t+1), \pi[s(t+1)]|s(t), a(t)\}\rangle, \quad (14)$$

where Q^π represents the Q value obtained under strategy π , and γ is a discount factor of interval $(0, 1)$. Thus the accumulated reward is kept a finite quantity where the temporally-proximate rewards are more heavily weighted than are the distant ones. To obtain the optimum policy π^* for maximum $E\langle Q^\pi\{s(t), \pi[s(t)]\}\rangle$, the proposed

TABLE 3: Sample data of traffic dynamics.

No.	Interval (min)	Training data		Testing data	
		Mainline demand (veh/h/ln)	On-ramp demand (veh/h/ln)	Mainline demand (veh/h/ln)	On-ramp demand (veh/h/ln)
1	0-2	1340	809	1507	1009
2	2-4	1363	835	1529	1035
3	4-6	1387	864	1553	1064
4	6-8	1413	895	1579	1095
5	8-10	1441	929	1607	1129
6	10-12	1470	965	1637	1165
7	12-14	1502	1003	1669	1203
8	14-16	1536	1043	1702	1243
9	16-18	1571	1085	1737	1285
10	18-20	1607	1129	1774	1329
11	20-22	1645	1175	1812	1375
12	22-24	1685	1222	1851	1422
13	24-26	1725	1270	1891	1470
14	26-28	1765	1318	1932	1518
15	28-30	1806	1368	1973	1568
16	30-32	1847	1417	2014	1617
17	32-34	1888	1465	2054	1665
18	34-36	1928	1513	2094	1713
19	36-38	1966	1560	2133	1732
20	38-40	2004	1604	2170	1774
21	40-42	2039	1647	2206	1797
22	42-44	2072	1686	2239	1806
23	44-46	2103	1723	2269	1845
24	46-48	2130	1756	2297	1872
25	48-50	2154	1785	2321	1907
26	50-52	2175	1810	2342	1926
27	52-54	2192	1830	2358	1951
28	54-56	2204	1845	2371	1970
29	56-58	2213	1856	2380	1981
30	58-60	2217	1861	2384	1993

network is customized for parameterized actors and critics, i.e. $\pi(s|\theta^\pi)$ and $Q(s, a|\theta^Q)$.

Specifically, the training epochs tune the actors and critics alternately, first by updating the function approximators θ^Q to satisfy Equation (14), and then by updating θ^π with a policy gradient defined by:

$$E \left[\nabla_a Q(s, a|\theta^Q) \Big|_{a=\pi(s)} \bullet \nabla_{\theta^\pi} \pi(s|\theta^\pi) \right]. \quad (15)$$

Transition pair $[s(t), a(t), s(t+1), r(t)]$ is stored in the replay memory at the end of each control step and is sampled uniformly at random for state update. Figure 3 illustrates the actor-critic framework employed to establish the relationships between traffic states, action agents, and reward function, so as to address the problem with continuous action space of parameter λ_i^x . Referring to the limited

dimension of input $[\mathbf{v}(t), \boldsymbol{\rho}(t)]$, i.e. $2(N^m + N^r - 2)$, one hidden layer is constructed for both critic and actor network.

Note: DDPG-based deep reinforcement learning method adopts exploration epochs before training. That is to overcome local minimum, a major challenge of learning in continuous action spaces. Thus the procedure of exploration is treated independently from the learning algorithm, so as to globally explore for the optimal strategy. Algorithm 1 summarizes the proposed solution method.

4. Simulation and Tests

In this section, we design and evaluate the performance of the proposed ACVSL model and solution method for a highway bottleneck with on-ramp. After validation, the proposed method is applied to extensive scenarios with varying traffic demand.

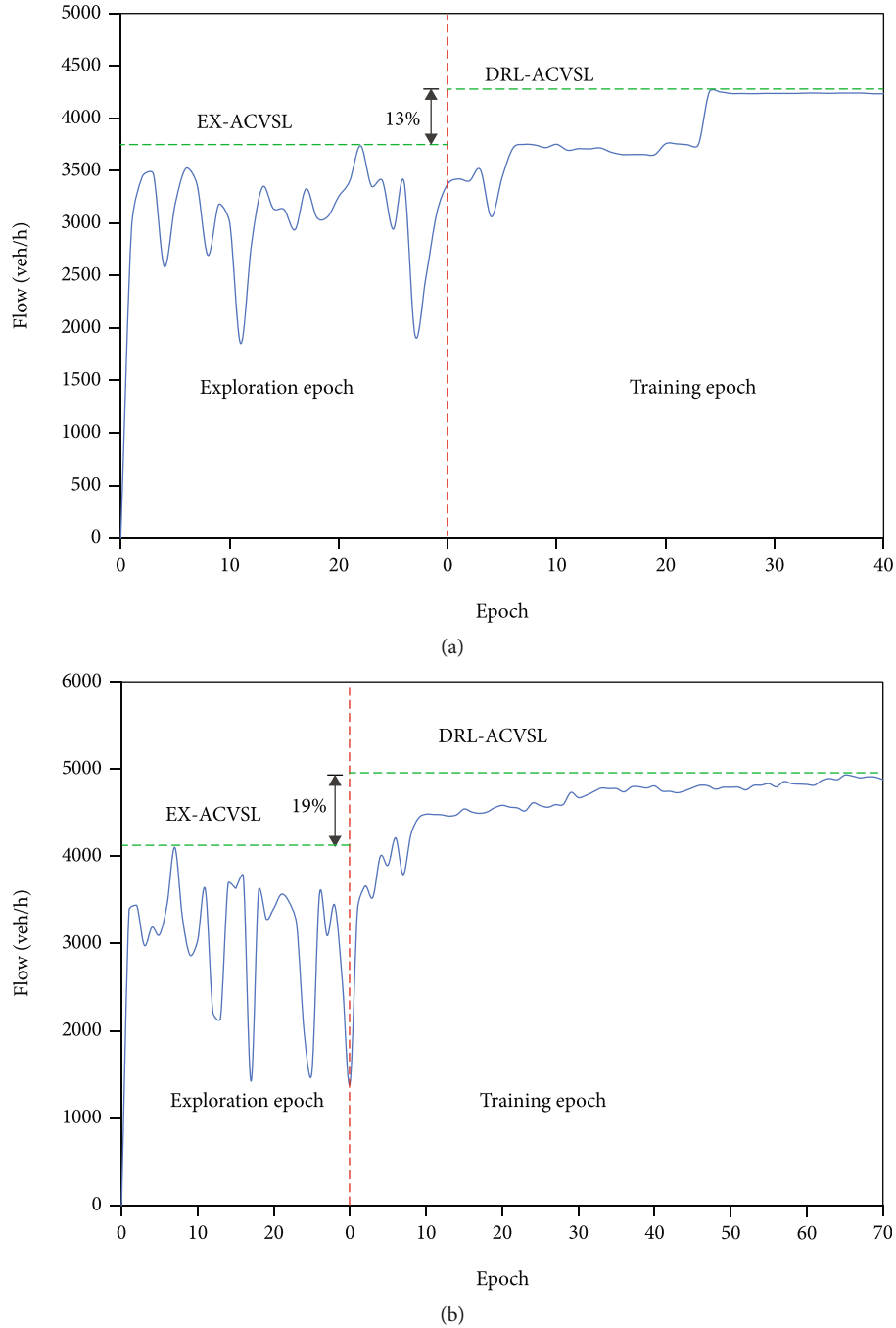


FIGURE 6: Bottleneck flow acquired from exploration and training procedure in different scenarios. (a) Static scenario and (b) dynamic scenario.

4.1. Set-up. The proposed ACVSL model and solution algorithm are developed and tested with MATLAB 2020, which is connected to simulation platform VISSIM 5.3 for traffic states collection. Specifically, the files of Layout.ini and Net.inp input road network information and other traffic parameters into VISSIM 5.3, respectively. Figure 4 shows the workflow, where the controller observes the traffic state and reward from the environment of VISSIM simulator at each control step and returns the control parameters to the simulator with COM interface. Note the first 20-min simulation is ignored to allow vehicles to fill the convergence zone.

Highway mainline has 3 lanes, while on-ramp has a single lane. Assuming default traffic parameters of driving behaviour and traffic composition in VISSIM platform, capacity of mainline and on-ramp is tested and set 2000 and 1500 veh/h/ln, respectively. Other parameters of the proposed ACVSL model are summarized in Table 2. Control step is set 2 min (i.e. $T=2$ min) for timely VSL update. That is, VSL signs are updated every 2 min so as to delicately control traffic state dynamics. Control parameter λ_i^x is specified in the range of $[0,100]$, i.e. the action space.

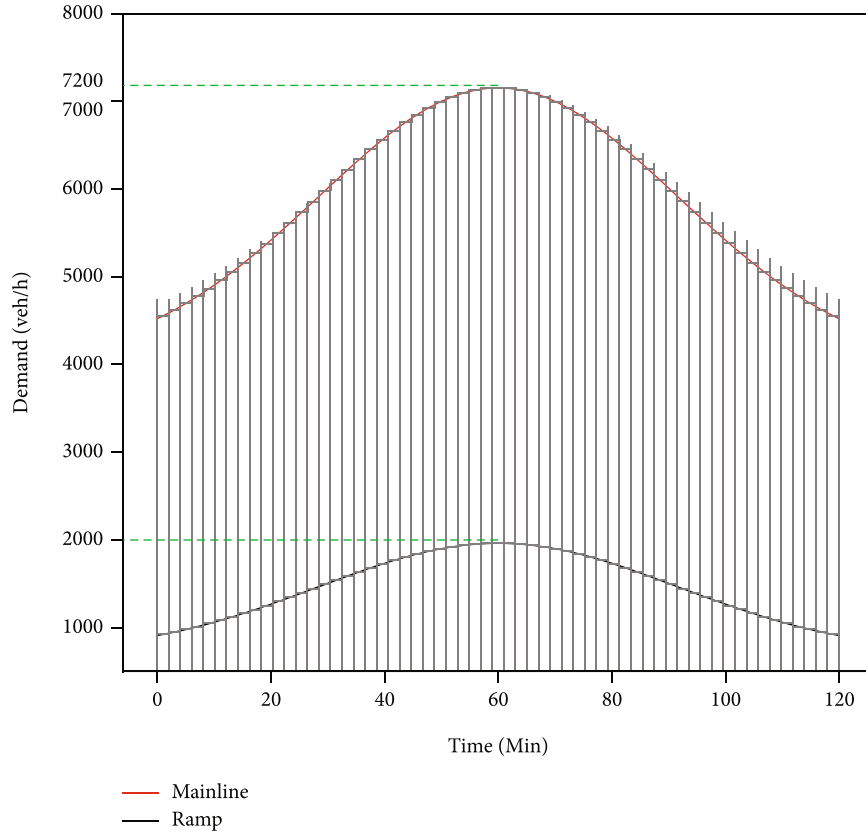


FIGURE 7: Dynamic flow demand for DDPG test.

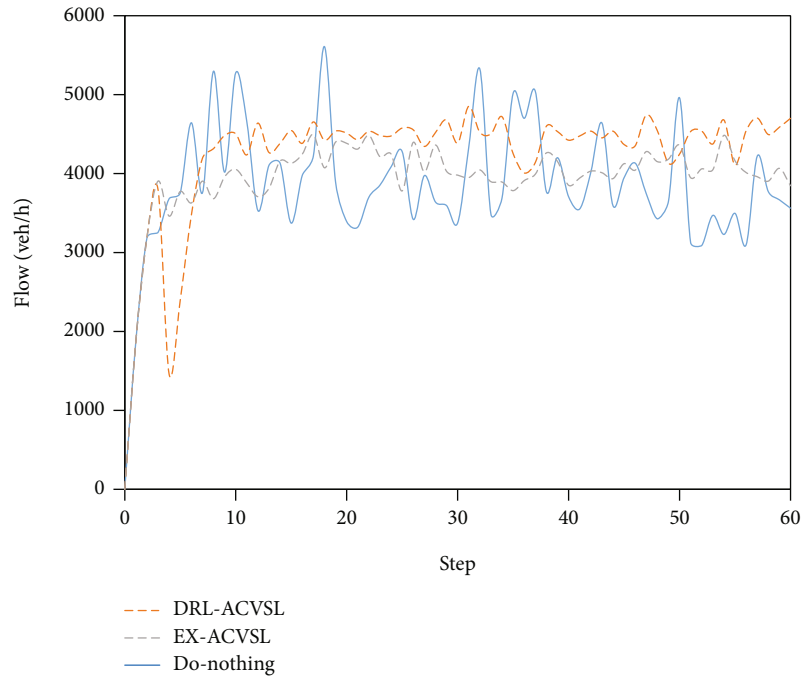
TABLE 4: Performance comparison of DRL-ACVSL, EX-ACVSL, and Do-nothing strategies.

Scenario Strategy		Static demand			Dynamic demand		
		DRL-ACVSL	EX-ACVSL	Do-nothing	DRL-ACVSL	EX-ACVSL	Do-nothing
Traffic flow (veh/h)	Avg.	4,291	3,998	3,909	4,818	4,208	4,041
	Med.	4,482	4,025	3,756	4,856	4,400	3,945
	Std. DV	517	468	660	610	656	688
Traffic density (veh/km)	Avg.	108	130	168	115	94	156
	Med.	109	134	177	119	90	162
	Std. DV	22	27	30	23	30	34
Traffic speed (km/h)	Avg.	41	32	25	44	47	28
	Med.	42	30	22	42	49	25
	Std. DV	6	8	9	10	11	11

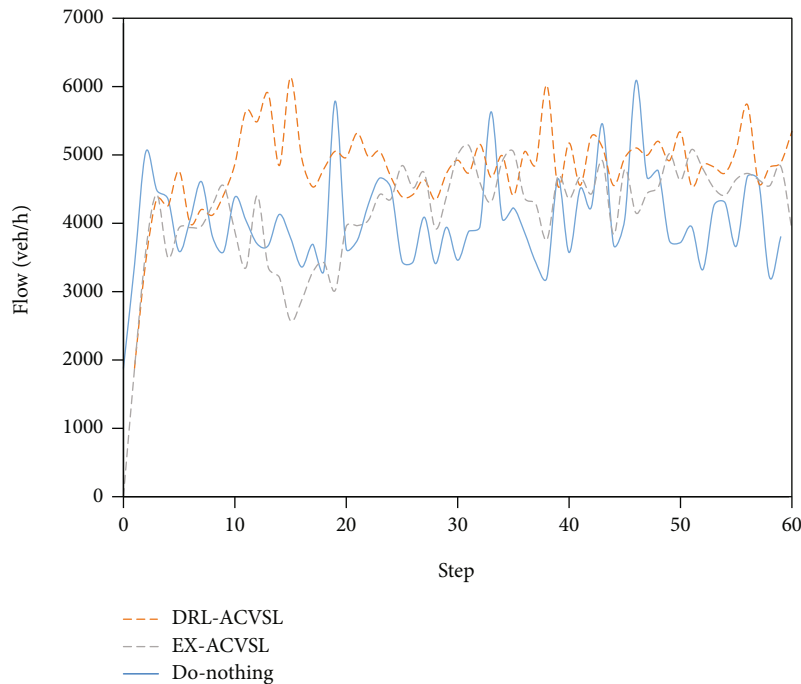
Meta-parameters of DDPG-based deep reinforcement learning are selected from repeated experiment, and the set with the best performance is adopted. Reward discount factor γ is set 0.85. Update probability during the training period is 20%. The size of replay memory and mini-batch is set 4000 and 400, respectively.

4.2. *Training.* We consider the following two scenarios to train the proposed method of DDPG-based deep reinforcement learning for the developed ACVSL model.

- (1) The scenario with static demand, which is set 5400 and 1400 veh/h for mainline and on-ramp, respectively
- (2) The scenario with dynamic demand, which follows the curve in Figure 5 for mainline and on-ramp, respectively. That is, demand increases and decreases for the first and last 60 min, respectively. Sample data employed in the training process is summarized in Table 3



(a)

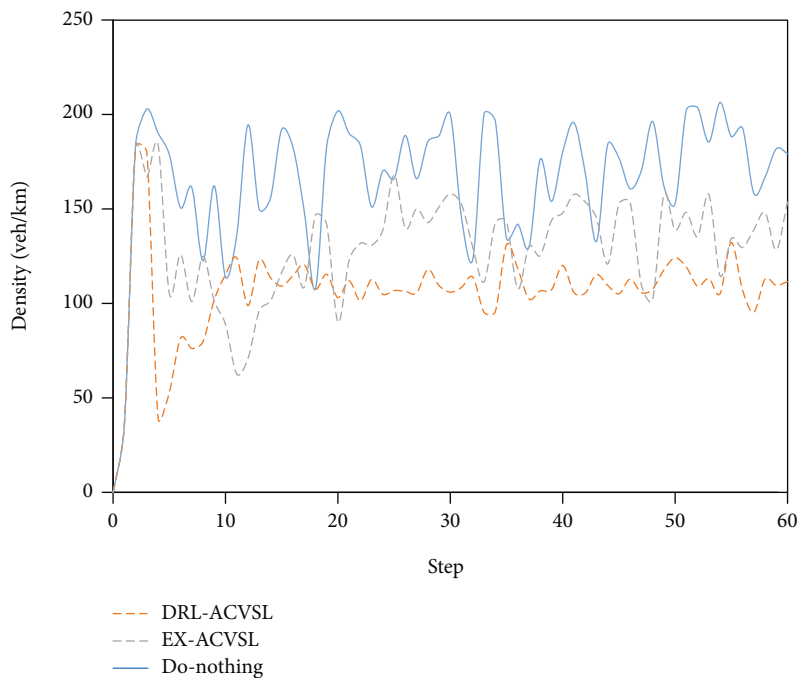


(b)

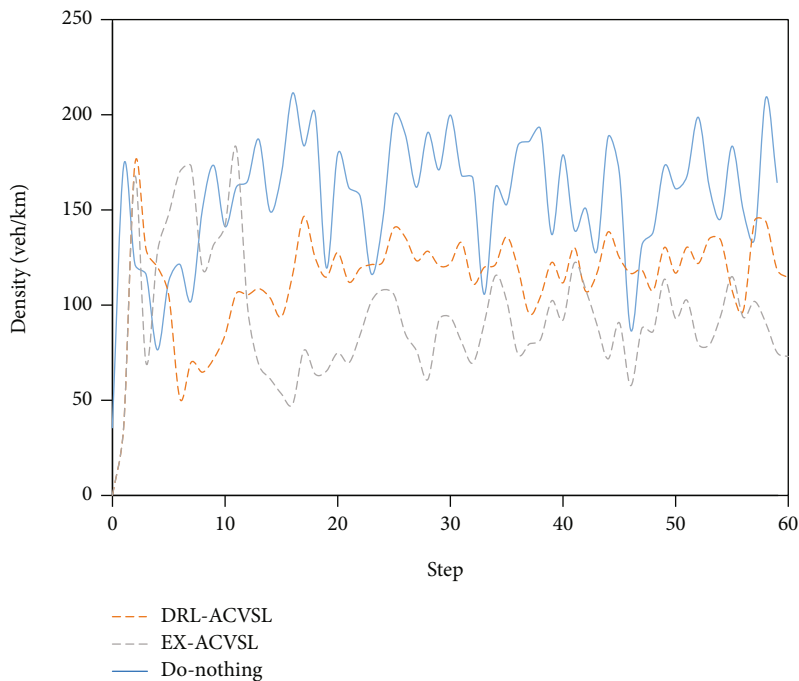
FIGURE 8: Bottleneck flow under different control strategies. (a) Static scenario and (b) dynamic scenario.

Each epoch of both exploration and training persists 2-hour simulation windows. With 2-minute control interval, each epoch of the exploration and training procedure is thus divided into 60 steps. A total of 30 epochs are adopted in the exploration, while 30 and 70 epochs in the training for the scenarios with static demand and dynamic demand, respectively, where learning rate τ is set 0.01 and 0.02, due to the difficulty of capturing demand changes in the latter scenario.

Figure 6 shows the outcomes of exploration and training epochs, as measured by the resulting bottleneck flow vs. epoch number under static and dynamic demand, respectively. In the exploration epoch, the series of control parameters is stochastically selected without optimization. Thus the bottleneck volume varies greatly. In comparison, in the training epoch, control parameters are iteratively updated to tune actors and critics alternatively, where bottleneck



(a)



(b)

FIGURE 9: Bottleneck density under different control strategies (a) Static scenario and (b) dynamic scenario.

volume exceeds that with the best control scheme from exploration epochs (i.e. EX-ACVSL) at the 7th and 8th epoch under the static and dynamic scenario, respectively. The optimal control schemes from the proposed deep reinforcement learning (i.e. DRL-ACVSL) are obtained at the 24th and 65th epoch, beating the best control scheme from the exploration epoch EX-ACVSL by 13% and 19%, respectively.

4.3. *Test and Calibration.* Test and calibration of the trained controller entails a 2h simulation under static and dynamic scenarios, respectively. The static case is characterized with the demand of 5500 veh/h on the mainline and 1500 veh/h on the ramp. The dynamic case deals with the demand given in Figure 7, data trend of which input to the proposed ACVSL method is similar to that of Table 3.

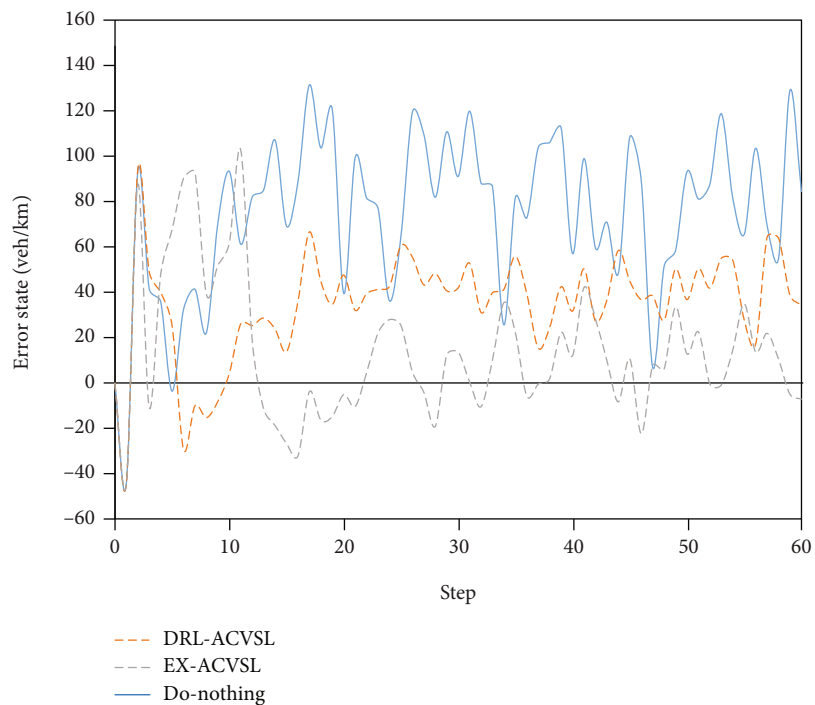
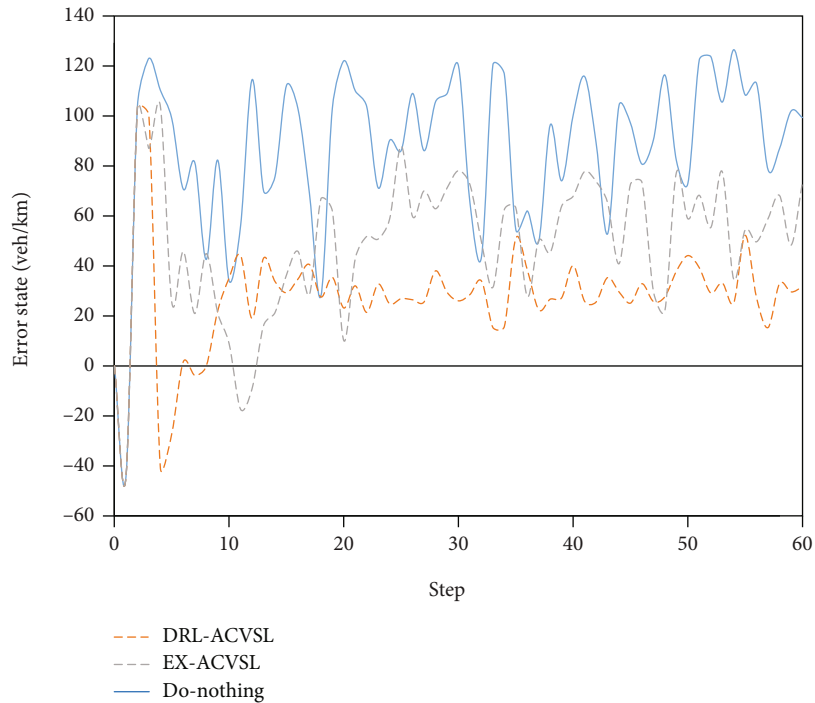
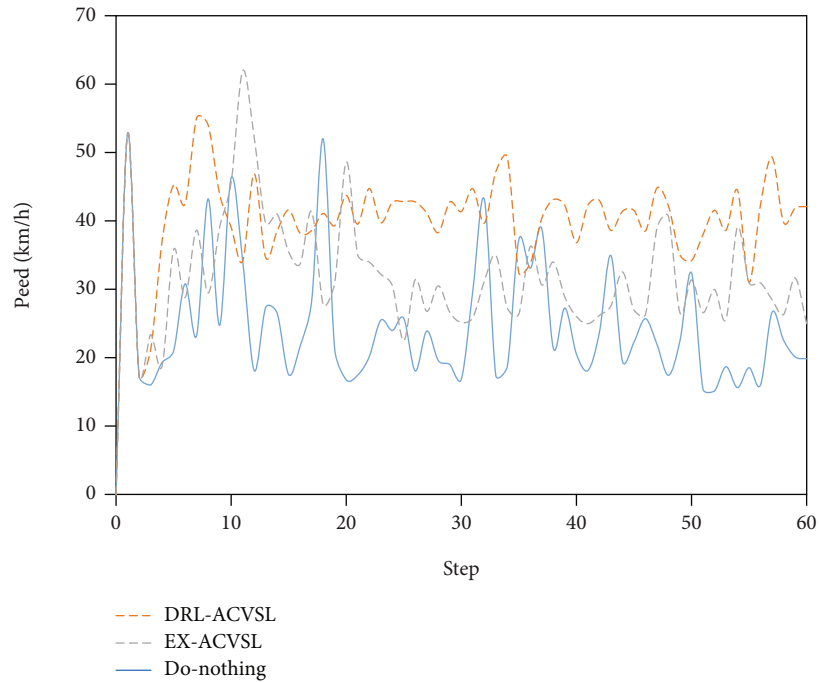


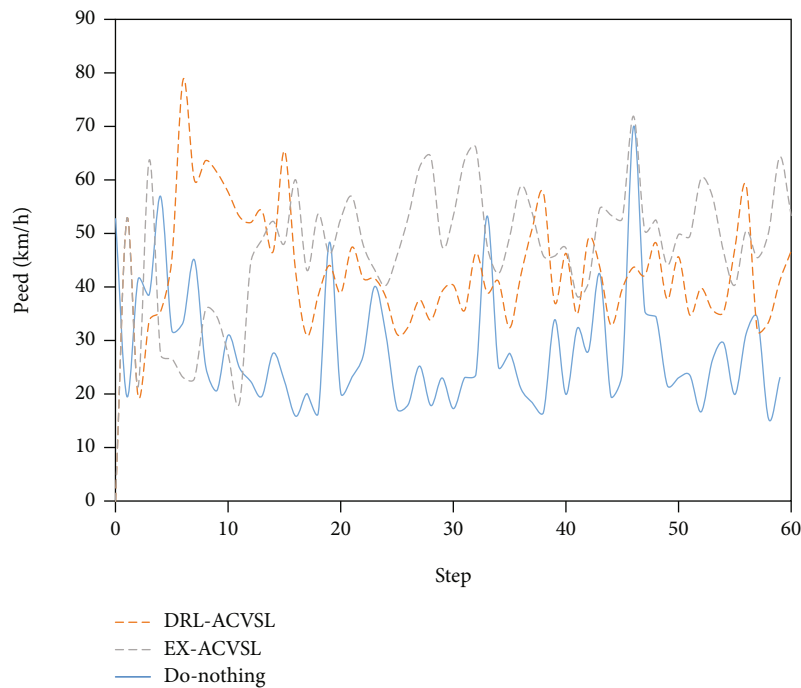
FIGURE 10: Bottleneck error state under different control strategies. (a) Static scenario and (b) dynamic scenario.

The proposed scheme from the developed method DRL-ACVSL is compared with the strategies that adopts the optimal control parameters from the exploration epochs (i.e. EX-ACVSL), and that of without control (i.e. Do-nothing). Table 4 summarizes the performance of DRL-ACVSL scheme from training together with EX-ACVSL selected

from the exploration epoch as well as the Do-nothing strategy. It is observed that DRL-ACVSL scheme competes with EX-ACVSL and Do-nothing with higher and more stable traffic flow at the merging zone, where the variation of traffic density and speed is also reduced to promote stable traffic flow. Specifically, average bottleneck flow is enhanced with



(a)



(b)

FIGURE 11: Bottleneck speed under different control strategies. (a) Static scenario and (b) dynamic scenario.

DRL-ACVSL strategy by 7% and 10% under static demand, while by 14% and 19% under dynamic demand, compared to EX-ACVSL and Do-nothing strategies, respectively. Such enhancement competes the existing research that reports flow increase by 4.7% [32]. DRL-ACVSL competence is also reflected with the fact that standard deviation of traffic density is reduced by 17% and 27% under static vs. 23% and 32% under dynamic demand and standard deviation of traf-

fic speed is reduced by 25% and 33% under static vs. 9% and 9% under dynamic demand against EX-ACVSL and Do-nothing strategies, respectively.

Figure 8 demonstrates the dynamics of traffic flow for highway merging bottleneck under three strategies, it is observed that DRL-ACVSL scheme manages to smooth vehicle outflow and keep it at a higher level than that with EX-ACVSL scheme, though the latter also promotes stable

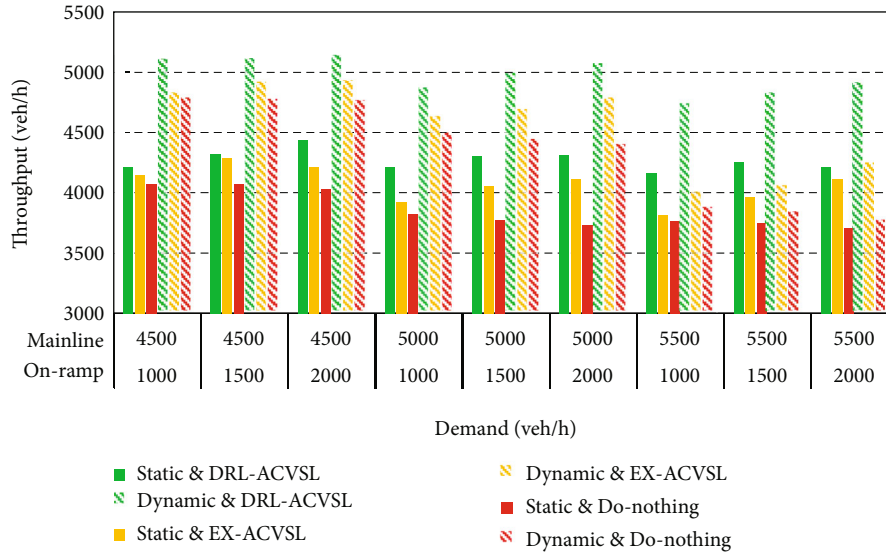


FIGURE 12: Traffic throughput with different control strategies under varying traffic demand.

vehicle throughput compared to the Do-nothing scheme. Moreover, DRL-ACVSL performs better under static demand than dynamic demand, bringing higher vehicle throughput, and better smoothing flow fluctuations. Another interesting finding is that, under dynamic demand, DRL-ACVSL manages to enhance traffic flow compared to the static scenario, because the changing demand allows better exploitation of highway capacity with gradually increasing or decreasing demand, instead of crowding into the bottleneck constantly.

Traffic density in Figure 9 shows that DRL-ACVSL is efficient in reducing and stabilizing traffic density at the merging zone under static scenario. In this case, DRL-ACVSL competes with EX-ACVSL and Do-nothing strategies, where significant ups and downs in traffic density are frequently observed. When it comes to the dynamic case, the advantage of DRL-ACVSL reducing density dynamics is kept, though density comes to a higher level compared to EX-ACVSL strategy. Therefore, DRL-ACVSL is validated to keep stable traffic density and reduce traffic shocks, promoting traffic efficiency and safety. Note traffic density is also related to the error state (see Figure 10) between actual and critical traffic density, where the latter is set constant based on fundamental diagram with speed limit. Thus Figure 10 shows similar trend to that of Figure 9.

Figure 11 comes to traffic speed of the merging zone. It is observed that DRL-ACVSL can enhance traffic speed and reduce speed oscillation to a higher degree than EX-ACVSL, while the Do-nothing strategy performs least satisfying. Moreover, the gap between DRL-ACVSL and EX-ACVSL is enlarged when it comes to the subsequent steps, showing that the proposed solution can well trace traffic dynamics under static demand. When it comes to the dynamic scenario, EX-ACVSL outperforms DRL-ACVSL with higher speed, though DRL-ACVSL manages to further reduce speed deviation. The Do-nothing strategy is still the worst with respect to speed average and variation.

4.4. Numerical Analyses. Extensive numerical analyses are conducted for the proposed method under variable traffic demands. Figure 12 summarizes the results from the numerical analysis, where dynamic demand increases from half static demand to one and half at the same difference for the first hour, and decreases in the same style for the second hour. It is observed that the proposed DRL-ACVSL is efficient in enhancing bottleneck throughput against varying traffic demand. Specifically, with static demand, DRL-ACVSL always competes with EX-ACVSL and Do-nothing strategies, with DRL-ACVSL advantages strengthened when on-ramp demand increases, though the latter to a lesser degree. In contrast, the Do-nothing strategy brings decreased throughput when on-ramp demand increases, implicating the intensified conflicts between mainstream and on-ramp vehicles. Similar trend is observed under dynamic demand, except that vehicle throughput is generally at a higher level compared to static scenarios. Thus the proposed model and solution method is validated to be capable of adapting to various traffic conditions for enhanced bottleneck efficiency.

Thus the proposed method has the potential to better enhance the capacity of the merging zone with increasing on-ramp demand. That can be explained with on-ramp vehicle density getting closer to the intended critical density to well sustain capacity of the merging zone for maximum throughput. In comparison, the increased demand of mainline decreases the vehicle flow of the merging zone. The findings are consistent to the previous research that increasing on-ramp flow and reducing mainline flow contributes to enhance bottleneck capacity [15] with a new attempt to coordinate the variable speed limits of both mainline and on-ramp.

5. Conclusions

This paper proposes an adaptive coordinated variable speed limit (ACVSL) control strategy for highway bottleneck with

on-ramp, where the priority to mainline over on-ramp is removed to alleviate the speed difference upstream of the merging zone and to improve highway merging efficiency. In this endeavour, error state is developed to indicate the gap between the actual and critical traffic density, with the latter derived from fundamental diagram that corresponds to road capacity. Thus the smaller the error state, the higher traffic flow, and the shifting can be reduced between fluent and congested traffic states. In the following, nonlinear feedback mechanism is established to adjust speed limits along the control horizon.

To solve the proposed model, deep reinforcement learning algorithm is developed, which is trained with DDPG using a critic-actor network to finely tune the continuous control variables. With traffic state collected and control reward evaluated, exploration and training epochs search the feasible domain and establish the control scheme. Experiment platform with VISSIM is then employed to test the proposed method, followed with numerical analyses. After calibration and training, extensive scenarios are established for method validation. It is found that the proposed scheme is capable of enhancing traffic flow through the bottleneck by 10% and 19% under static and dynamic demand, respectively, compared to the Do-nothing strategy. That competes the existing literature with respect to the improvement in vehicle throughput, validating the efficiency of the proposed model. Moreover, the proposed scheme manages to reduce the variation of traffic density and speed by around 30%, helping to stabilize traffic states and reduce traffic shockwaves. Extensive numerical analyses further confirm the advantage of the proposed method to improve bottleneck capacity especially when on-ramp demand increases.

This research adds to the existing literature with coordinated speed guide between mainline and on-ramp, providing insights for responsive agencies into highway on-ramp bottleneck management that prioritize mainline over ramps can be unnecessary. Thus highway merging bottleneck can be better addressed to relieve congestion and to reduce accident. Limitation of the research is threefold. First, the performance in pollution emissions and fuel consumption should be further explored. Second, detailed analysis has not been scratched on how mainline and on-ramp vehicles cooperate for acceptable gap and smooth merging. Third, the effect of acceleration lane at the merging zone on traffic states has not been considered. The on-going research is to extend the proposed method to include more ramps so as to explore the potential of adaptive coordinated variable speed limit for consecutive highway ramps.

Data Availability

The data that supports the findings of this study is available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interests.

Acknowledgments

This work has been supported by the National Natural Science Foundation of China [Grant No. 52002261] and the Natural Science Research of Jiangsu Higher Education Institutions of China (Grant No. 20KJB580011). The authors also thank PhD student Yanpei Zhang for his help in the manuscript format in the Joint Laboratory of Future Transport.

References

- [1] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal mainstream traffic flow control of large-scale motorway networks," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 2, pp. 193–212, 2010.
- [2] J. Zhao, W. Ma, and H. Xu, "Increasing the capacity of the intersection downstream of the freeway off-ramp using pre-signals," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 8, pp. 674–690, 2017.
- [3] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Integrated feedback ramp metering and mainstream traffic flow control on motorways using variable speed limits," *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 209–221, 2014.
- [4] S. Cohen, D. Gil, Z. Christoforou, and R. Seidowsky, "Evaluating the combined effect of ramp metering and variable speed limits on the French A25 motorway," *Transportation Research Procedia*, vol. 27, pp. 156–163, 2017.
- [5] H. Taale and F. Middelham, "Ten years of ramp-metering in the Netherlands," in *Tenth International Conference on Road Transport Information and Control IEEE Xplore*, London, UK, 2000.
- [6] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: an overview," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 271–281, 2002.
- [7] A. Hegyi, B. D. Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 3, pp. 185–209, 2005.
- [8] I. Papamichail, K. Kampitaki, M. Papageorgiou, and A. Messmer, "Integrated ramp metering and variable speed limit control of motorway traffic flow," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 14084–14089, 2008.
- [9] X. Lu, T. Z. Qiu, P. Varaiya, R. Horowitz, and S. E. Shladover, "Combining variable speed limits with ramp metering for freeway traffic control," in *Proceedings of the 2010 American Control Conference*, pp. 2266–2271, Baltimore, MD, USA, 2010.
- [10] X. Lu, P. Varaiya, R. Horowitz, D. Su, and S. E. Shladover, "Novel freeway traffic control with variable speed limit and coordinated ramp metering," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2229, no. 1, pp. 55–65, 2011.
- [11] R. Scarinci, A. Hegyi, and B. Heydecker, "Definition of a merging assistant strategy using intelligent vehicles," *Transportation Research Part C Emerging Technologies*, vol. 82, no. 17, pp. 161–179, 2017.
- [12] J. Larsson, M. F. Keskin, B. Peng, B. Kulcsár, and H. Wymeersch, "Pro-social control of connected automated vehicles in mixed-autonomy multi-lane highway traffic," *Communications in Transportation Research*, vol. 1, article 100019, 2021.

- [13] A. T. McCattt, N. Veronika Shabanova, and R. A. Retting, "Types and characteristics of ramp-related motor vehicle crashes on urban interstate roadways in northern Virginia," *Journal of Safety Research*, vol. 35, no. 1, pp. 107–114, 2004.
- [14] Y. Guo, Z. Li, P. Liu, and Y. Wu, "Modeling correlation and heterogeneity in crash rates by collision types using full Bayesian random parameters multivariate Tobit model," *Accident Analysis & Prevention*, vol. 128, pp. 164–174, 2019.
- [15] Y. Guo, Z. Li, and T. Sayed, "Analysis of crash rates at freeway diverge areas using Bayesian Tobit modeling framework," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 4, pp. 652–662, 2019.
- [16] M. Hunter and R. B. Machemehl, "Re-evaluation of ramp design speed criteria: review of practice and data collection plan," in *Design Speed*, United States. Federal Highway Administration, 1997.
- [17] E. Vinitsky, K. Parvate, A. Kreidieh, C. Wu, and A. Bayen, "Lagrangian Control through Deep-RL: Applications to Bottleneck Decongestion," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 759–765, Maui, HI, USA, 2018.
- [18] R. Xia, Y. Chen, and B. Ren, "Improved Anti-Occlusion Object Tracking Algorithm Using Unscented Rauch-Tung-Striebel Smoother and Kernel Correlation Filter," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 6008–6018, 2022.
- [19] Y. Chen, H. Zhang, L. Liu et al., "Research on image inpainting algorithm of improved total variation minimization method," *Journal of Ambient Intelligence and Humanized Computing*, 2021.
- [20] P. Li and Y. Chen, "Research into an image inpainting algorithm via multilevel attention progression mechanism," *Mathematical Problems in Engineering*, vol. 2022, Article ID 8508702, 12 pages, 2022.
- [21] Y. Chen, L. Liu, V. Phonevilay et al., "Image super-resolution reconstruction based on feature map attention mechanism," *Applied Intelligence*, vol. 51, no. 7, pp. 4367–4380, 2021.
- [22] Z. Li, P. Lin, C. Xu, H. Duan, and W. Wang, "Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 3204–3217, 2017.
- [23] Z. Li, C. Xu, Y. Guo, P. Liu, and Z. Pu, "Reinforcement learning-based variable speed limits control to reduce crash risks near traffic oscillations on freeways," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 4, pp. 64–70, 2021.
- [24] K. Rezaee, B. Abdulhai, and H. Abdelgawad, "Application of reinforcement learning with continuous state space to ramp metering in real-world conditions," in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 1590–1595, Anchorage, AK, USA, 2012.
- [25] Z. Ke, Z. Li, Z. Cao, and P. Liu, "Enhancing transferability of deep reinforcement learning-based variable speed limit control using transfer learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4684–4695, 2021.
- [26] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," 2015, <https://arxiv.org/abs/1509.02971>.
- [27] M. Greguric, K. Kusic, and E. Ivanjko, "Impact of deep reinforcement learning on variable speed limit strategies in connected vehicles environments," *Engineering Applications of Artificial Intelligence*, vol. 112, article 104850, 2022.
- [28] C. Wang, J. Zhang, L. Xu, L. Li, and B. Ran, "A new solution for freeway congestion: cooperative speed limit control using distributed reinforcement learning," *IEEE Access*, vol. 7, pp. 41947–41957, 2019.
- [29] C. F. Daganzo and N. Geroliminis, "An analytical approximation for the macroscopic fundamental diagram of urban traffic," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 771–781, 2008.
- [30] C. F. Daganzo, "The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.
- [31] Y. Zhang and P. A. Ioannou, "Coordinated variable speed limit, ramp metering and lane change control of highway traffic," *IFAC-PapersOnline*, vol. 50, no. 1, pp. 5307–5312, 2017.
- [32] I. Papamichail, M. Papageorgiou, V. Vong, and J. Gaffney, "Heuristic ramp-metering coordination strategy implemented at Monash freeway, Australia," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, no. 1, pp. 10–20, 2010.