

Article

Reinforcement Learning-Based Dynamic Zone Positions for Mixed Traffic Flow Variable Speed Limit Control with Congestion Detection

Filip Vrbanic [†] , Martin Greguric [†] , Mladen Miletic  and Edouard Ivanjko ^{*} 

Faculty of Transport and Traffic Sciences, University of Zagreb, Vukelićeva Street 4, HR-10000 Zagreb, Croatia; fvrbanic@fpz.unizg.hr (F.V.); mgreguric@fpz.unizg.hr (M.G.); mmiletic@fpz.unizg.hr (M.M.)

^{*} Correspondence: eivanjko@fpz.unizg.hr

[†] These authors contributed equally to this work.

Abstract: Existing transportation infrastructure and traffic control systems face increasing strain as a result of rising demand, resulting in frequent congestion. Expanding infrastructure is not a feasible solution for enhancing the capacity of the road. Hence, Intelligent Transportation Systems are often employed to enhance the Level of Service (LoS). One such approach is Variable Speed Limit (VSL) control. VSL increases the LoS and safety on motorways by optimizing the speed limit according to the traffic conditions. The proliferation of Connected and Autonomous Vehicles (CAVs) presents fresh prospects for improving the operation and measurement of traffic states for the operation of the VSL control system. This paper introduces a method for the detection of multiple congested areas that is used for state estimation for a dynamically positioned VSL control system for urban motorways. The method utilizes Q-Learning (QL) and CAVs as mobile sensors and actuators. The proposed control approach, named Congestion Detection QL Dynamic Position VSL (CD-QL-DPVSL), dynamically detects all of the congested areas and applies two sets of actions, involving the dynamic positioning of speed limit zones and imposed speed limits for all detected congested areas simultaneously. The proposed CD-QL-DPVSL control approach underwent an evaluation across six distinct traffic scenarios, encompassing CAV penetration rates spanning from 10% to 100% and demonstrated a significantly better performance compared to other control approaches, including no control, rule-based VSL, two Speed-Transition-Matrices-based QL-VSL configurations with fixed speed limit zone positions, and a Speed-Transition-Matrices-based QL-DVSL with a dynamic speed limit zone position. It achieved enhancements in macroscopic traffic parameters such as the Mean Travel Time and Total Time Spent by adapting its control policy to every simulated scenario.

Keywords: Variable Speed Limit; Connected and Autonomous Vehicles; Reinforcement Learning; urban motorway; Intelligent Transportation Systems; traffic state estimation; dynamic speed limit zone positioning; congestion detection



Citation: Vrbanic, F.; Greguric, M.; Miletic, M.; Ivanjko, E. Reinforcement Learning-Based Dynamic Zone Positions for Mixed Traffic Flow Variable Speed Limit Control with Congestion Detection. *Machines* **2023**, *11*, 1058. <https://doi.org/10.3390/machines11121058>

Academic Editor: Dan Zhang

Received: 13 October 2023

Revised: 8 November 2023

Accepted: 26 November 2023

Published: 28 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Urban motorways represent an important part of the traffic infrastructure of the urban traffic network and, as high-performance roads, represent key traffic routes. They serve local and transit traffic, and due to the large number and dense arrangement of on- and off-ramps, they are often congested, especially during the morning and afternoon peak hours. One of the solutions to relieve congestion is to expand and upgrade the existing traffic infrastructure by adding more traffic lanes. Such a solution is not always feasible due to high costs and a lack of space, especially in larger urban cities.

For this reason, solutions from the field of Intelligent Transportation Systems (ITS) are increasingly being used. The urban motorway operational service capacity and Level of Service (LoS) can be improved in existing traffic infrastructure by using traffic control approaches such as Ramp Metering (RM) and the Variable Speed Limit (VSL). In this

paper, we apply the VSL to the existing urban motorway model to improve the traffic Measures of Effectiveness (MoEs). The VSL aims to harmonize speeds in mainstream traffic, reducing the sudden changes in speed that often cause shock waves [1]. Furthermore, by controlling the speed of vehicles moving in a mainstream traffic flow, the flow of vehicles to the conflicting areas where congestion occurs due to the interference of vehicles from the on-ramp can be reduced. Slowing down the mainstream traffic flow before the on-ramp enables the faster and safer entry of vehicles from the on-ramp to the mainstream flow and increases the operational capacity of an urban motorway.

To implement a VSL control system, it is necessary to define speed limit activation rules based on measurements of traffic and/or meteorological data. Recently, control approaches based on Reinforcement Learning (RL) have been used for the VSL [2]. More specifically, RL is employed for the optimization of the VSL control policy based on the principle of maximizing the Quality value (Q-value) for each state–action pair. By doing so, the action that produces the best results for a particular state will have the highest Q-value and is considered the best action for that state. The assessment of the performance of such systems is most often measured by certain MoEs that include the Mean Travel Time (*MTT*), Total Travel Time (*TTT*), Total Time Spent (*TTS*), and total delay time.

The classic VSL control systems have a few major drawbacks that include the need to have a fixed position for the applicable speed limit and a fixed position to measure traffic states based on the traffic density, segment Travel Time (*TT*), or the exiting traffic flow from the area of interest. The recent appearance and rapid development of Autonomous Vehicles (AVs) and Connected Autonomous Vehicles (CAVs) present fresh prospects for improving the measurement of traffic states for the operation of the VSL control system. The appearance and implementation of such vehicles in the traffic flow alongside Human-Driven Vehicles (HDVs) have created a recently emerged type of traffic flow commonly named “mixed traffic flow”. The introduction of CAVs into the control loop of a VSL control system presents new possibilities for the measurement of traffic states and the imposition of speed limits, since CAVs have the ability to communicate with the VSL control system and send relevant traffic data to approximate the position and the severity of the congestion [3]. Furthermore, they can receive speed limit information with no fixed position for the applicable VSL area. An overview of the optimization of such RL-based VSL control approaches that include mixed traffic flows was conducted in [2].

In this paper, the Q-Learning VSL (QL-VSL) control approach based on dynamic congested area detection, dynamic speed limit zone positions, and imposed speed limits based on the computed gradients of sudden mean speed changes between consecutive motorway edges, called Congestion Detection QL Dynamic Position VSL (CD-QL-DPVSL), is proposed and analyzed. The proposed CD-QL-DPVSL is compared to several existing control approaches including the classic Rule-Based VSL (RB-VSL) algorithm, the Speed Transition Matrices-based QL-VSL (STM-QL-VSL), the STM-QL Dynamic VSL (STM-QL-DVSL), previously developed in [3], and the baseline no-control scenario. This proposed CD-QL-DPVSL control approach is an extension of our previous control approaches [3]. Thus, this paper presents the following key contributions:

- The proposal of a method that detects multiple congested areas on an urban motorway to determine the dynamic position of the VSL;
- The proposal of a method that calculates the severity of the congestion in a congested area by utilizing real-time data gathered from CAVs as the input for the QL algorithm for the VSL;
- The introduction of a novel approach that incorporates the QL algorithm for VSL control, involving the computation of positioning of speed limit zones and speed limits applied to CAVs;
- A performance analysis of the proposed CD-QL-DPVSL control approach under distinct mixed traffic flow scenarios on the simulated urban motorway.

The paper follows a structured organization. Section 2 offers an overview of previous research concerning the implementation of the VSL on urban motorways. Section 3 outlines

the proposed methodology for the study. Section 4 provides a description of the utilized simulation model. Moving on, Section 5 presents the findings, along with an analysis and discourse on the conducted simulations. The paper ends with a conclusion in its final section, where future research possibilities are also explored.

2. Related Work

As previously mentioned, urban motorways face the problem of increased traffic demand during peak hours that often exceeds the operational capacity of the motorway. Implementing VSL control helps to maintain the operational capacity of the congested motorway area by limiting the incoming vehicles' flow to the congested area by reducing the incoming speed [4]. By doing so, the flow of vehicles to the congested area is reduced, further capacity drops are eliminated, and the congestion can be relieved promptly [4]. Appropriately computed VSL speed limits aim to manage the congested area traffic flow close to the maximal operational capacity. Furthermore, the flow of vehicles in the congested area is also affected by applying the speed limits, resulting in a reduced number of incoming vehicles to the congested area. This reduced flow ensures that the maximal operational capacity of the congested area is not exceeded and the congestion is relieved promptly or even averted. Furthermore, the appropriately computed variable speed limits help to homogenize vehicle speeds and, thus, reduce the risk of accidents due to large discrepancies in vehicle speeds before and in the congested area [5].

Figure 1 shows the relation between traffic flow q (displayed on the y -axis) and traffic density ρ (displayed on the x -axis) that is commonly referred to as the fundamental traffic flow diagram. According to the fundamental traffic flow diagram, as q increases to the maximal value q_{max} , the traffic density increases to the critical value ρ_c . When the traffic flow is below the q_{max} value and the density is below ρ_c , the traffic flow is stable and runs smoothly with limited interactions among vehicles [6]. When the traffic density exceeds ρ_c , it becomes unstable, and vehicles must interact with each other and travel at lower speeds with an increased traffic density in the segment. Employing the VSL before the congested area helps to reduce the inflow of vehicles to the congested area and increase the outflow in the controlled area section [7].

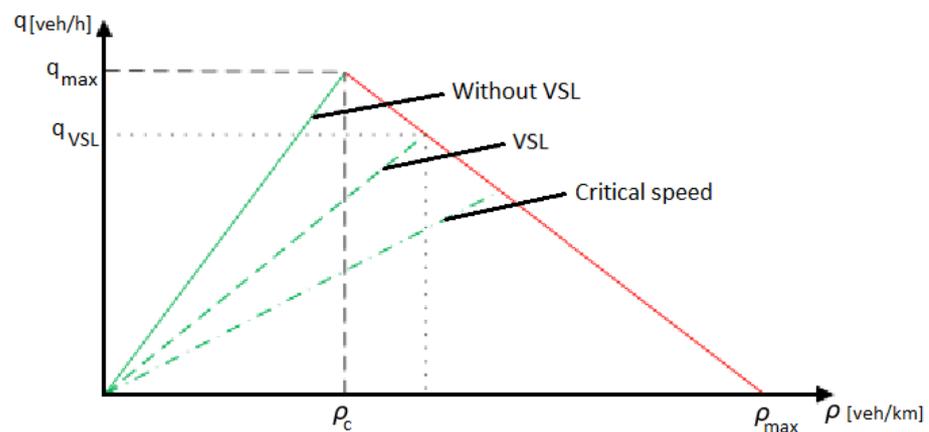


Figure 1. The impact of the speed limit on the fundamental diagram [2].

The influence of employing CAVs in the mixed traffic flow on the acceleration rate and vehicle speed disparities was analyzed in [8]. The influence of a gradual increase in the CAV penetration rate from no CAVs to 100% CAVs showed a reduction in speed deviations between vehicles and abrupt decelerations in the mixed traffic flow. Furthermore, the operational capacity of each lane was examined by employing AVs and CAVs in the mixed traffic flow [9]. By increasing the CAV penetration rate from 0 to 100%, the lane capacity was improved by 188.2% with a near-linear characteristic. In [10], the authors analyzed the impact of CAV penetration rates from 0 to 70% on the fundamental diagram. It was concluded that the presence of CAVs in a mixed traffic flow increases the ρ_c and q_{max} . For

example, ρ_c and q_{max} increased by 37% and 75%, respectively, in the scenario with a 70% CAV penetration rate in the mixed traffic flow [10].

Current applications of VSL control are based on traffic flow measurement, which is conducted within the spatially fixed segments of the motorway in the downstream section of a Variable Message Sign (VMS) [11,12]. The segment for traffic flow measurement is spatially defined by the positions of fixed road sensors on them. They are usually located at the start of each measurement segment [13]. The traffic macroscopic parameters for each of those segments are computed based on aggregated raw data from sensors. Thus, the drawback of this approach is the spatial accuracy in detecting congestion on motorways, which heavily depends on the lengths of those segments. Moreover, the length, start, and end point of the congestion are difficult to determine by using fixed measurement segments with longer lengths. Thus, it is possible to conclude that larger segments reduce the accuracy in detecting the spatial parameters of congestion while their implementation is lower in cost since fewer sensors are required.

The limitations in the detection of congestion on motorways are bounded by the technologies used for measuring the behavior of mainstream traffic flow. The most commonly used approach is based on fixed road sensors which are used to compute traffic parameters for the segments of the motorway that they are covering. Those data are passed to the rule-based algorithm which determines the motorway segments with congestion [14,15]. The cellular-probe-based motorway congestion detection approach relies on on-call wireless location technologies with signal transition data. The accuracy of those approaches depends on related problems, such as the small sample size, frequent road tests, safety, and privacy issues. Thus, the cellular activity features, the link pseudo speed, and link probe activity are defined and calculated by a rule-based algorithm for the estimation of the traffic congestion state [16]. The latest approach in congestion detection on motorways concerning the used sensors involves CAVs in interactions with vehicle-to-infrastructure communication environment. They generate floating car data which are used to detect congestion in real-time [17,18]. The accuracy when detecting congestion in a spatial context is heavily dependent on the percentage and distribution of CAVs in motorway networks. Due to CAVs' communication capabilities to send real-time position and speed data, CAVs can be used as moving VSL actuators in the control loop [3]. These data were aggregated to estimate the traffic state at a fixed position on an urban motorway by applying the Speed Transition Matrices (STMs) to determine the congestion severity. The main drawbacks of the proposed STM methodology lie in the fixed measuring position on an urban motorway and the large amount of computational resources needed to construct the STMs, calculate the congestion probability, and estimate the environment states.

Previously analyzed VSL control approaches applied in mixed traffic flow environments considered the state space based on the occupancy rates reported by the loop detectors mounted on fixed positions [19], with the loop detectors placed on four measured mainstream lanes [20], while in [21], traffic data collected using loop detectors included local speed, occupancy, and traffic flow data, which were used to transmit the environment state to the fuzzy controller that computes speed limits. In [12], traffic states were obtained at the initial position of each cell by collecting traffic information such as the number of vehicles, vehicle speed, and density of the cell, and the inflow of vehicles was calculated. If the inflow was less than the bottleneck capacity, no control was taken; otherwise, the VSL was activated, and the speed limit was displayed on the VMS. The downside of this approach is using VMSs in a fixed position on an observed motorway. In a recent paper [22], the authors designed a PID-based VSL controller to mitigate the negative effect of the shock wave forming on a motorway using speed and position data from CAVs as the input to the controller. The downside of this approach is the assumption of the formation of only one bottleneck alongside the observed motorway. In [23], the VSL input traffic state prediction was based on the data collected by loop detectors during the observed time interval, while in [13], the traffic state was calculated based on data from 10 loop detectors equally spaced by 1 km on the observed motorway segment. The traffic states in [24] were based on the

density and average speed of HDVs and AVs in each lane for five observed segments and the outflow at the bottleneck area. The main problems for such traffic state measurements are based on the fixed positions of loop detectors and the predetermined observed segment positions and lengths.

The application of intelligent control methods in mixed traffic flows includes a lane-level adaptive speed control method [25]. The proposed approach uses traffic state estimation based on the traffic density measured per lane and applied the VSL control to one CAV in each lane to harmonize the traffic. The control approach considers the influence of the random mixing of CAVs and HDVs. The downside of the approach proposed in [25] is based on the known fixed position of a congested area, without considering the detection of possible multiple congested areas. In [26], an optimal differential VSL control strategy in a mixed traffic flow environment for freeway off-ramp bottlenecks was developed. The proposed optimal differential VSL control method was developed to implement varying speed limits for individual lanes, incorporating a multi-objective approach near the off-ramp. A comprehensive approach to managing traffic flow in the merging area of highways was introduced in [27], integrating CAVs' active lane change technology with conventional VSL and RM strategies. Traffic state estimation of the mainline and on-ramp traffic flow was performed by using an upgraded cell transmission model, while the genetic algorithm was used to compute the optimal speed limit of mainline traffic flow, the number of mainline vehicles changing lanes, and RM control to maximize the traffic flow in the merging area. This method only considers the merging area control. In [28], a framework that devises an integrated action of several control strategies such as RM, lane changing control for CAVs and lane changing recommendations for HDVs, VSL control for CAVs, and VSL recommendations for HDVs with minimal safety gap control measures for lane changing and merging maneuvers was proposed. The proposed method considers three fixed-position VSL zones for VSL and VSL recommendations for HDVs and does not consider the dynamic positioning of VSL zones.

In this paper, a centralized agent-based approach for QL-VSL control is developed by utilizing real-time data sent and received from CAVs. The novelty of the proposed approach is the use of real-time data collected from CAVs to detect multiple congested areas and determine the length of each congested area as an input state to QL-VSL alongside the whole observed urban motorway, as well as calculating the speed limits and speed limit zone placements imposed on CAVs without the need for physical road sensors and VMSs. Mixed traffic flow scenarios containing HDVs and varying CAV penetration rates are analyzed in a synthetic urban motorway model. The shared data are assumed to be transmitted and received by Road Side Units (RSUs) and CAVs equipped with an On-Board Unit (OBU). For the purpose of this paper, the data loss and communication delay are ignored. CAVs also act as actuators to the QL-VSL, which sends speed limit data to each CAV. Therefore, CAVs are utilized as moving sensors and actuators. Using CAVs as moving sensors discards the need for hardware traffic detection, such as induction loops, and the usage of CAVs as actuators discards the need for VMSs. Furthermore, congested areas can be identified more accurately and on a larger scale whilst obtaining real-time data on the positions and lengths of all detected congested areas in the observed urban motorway segment. An in-depth explanation of the congestion detection based traffic state estimation on an urban motorway is provided in Section 3.2.

3. Spatial-Temporal Variable Speed Limit Based on Q-Learning and Congestion Detection

This section describes the methodology of the proposed dynamic spatial-temporal VSL based on QL and congestion detection.

3.1. Q-Learning and Spatial-Temporal Variable Speed Limit

The QL algorithm is based on the premise of storing and updating each of the state-action pair's Q-values in each iteration. The knowledge retained in the Q-matrix is updated

by providing feedback by a positive or a negative reward for a selected action a in a particular state s . The longer the QL algorithm runs, the more the Q-values converge to the best possible value for each state–action pair. The highest Q-value is acquired when the optimal action within a specific environment state is selected. The Q-value Q^* is revised after every successful action selection, based on [29]:

$$Q^*(s_t, a_t) := (1 - \alpha)Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a' \in A} Q(s_{t+1}, a')), \quad (1)$$

where the current Q-value is defined as $Q(s_t, a_t)$ and is calculated for the specific state–action combination (s_t, a_t) at the control time step t . The significance of future rewards in the subsequent state is determined by the discount factor, denoted as γ . The assessment of the action taken at time t , denoted as a_t , depends on the reward r_t obtained in state s_t . Additionally, s_{t+1} represents the subsequent state in the environment, while α signifies the learning rate responsible for controlling the rate at which QL aggregates updated information and revises its Q-values.

The primary objective addressed when using QL is optimizing the speed limit selection within the framework of the VSL traffic control strategy to enhance the traffic flow and alleviate congestion. The decision-making process conforms to a Markovian Decision Process, where the agent computes speed limits, as described in prior research papers [11,30,31]. At each control timestamp t , the agent takes actions that lead to feedback based on changes in the environment state, determined by a well-defined reward function. In this context, the agent selects an action a_t , a speed limit value chosen from a discrete set of choices $A = \{60, 70, 80, 90, 100, 110, 130\}$ km/h within the current state of the environment s_t . Notably, a constraint is imposed, limiting speed limit changes between consecutive control time steps to a maximum range of ± 30 km/h, ensuring adherence with regulations and facilitating smooth transitions in speed limits to avoid abrupt speed changes. The updating of the learning rate α was performed based on [32]:

$$\alpha_{(s,a)} = \left(\frac{1}{1 + nv_{(s,a)}} \right)^u + c, \quad (2)$$

where the frequency of visits to each state–action combination is denoted as $nv_{(s,a)}$ and is measured before the learning rate α is updated. In this research, a gradual reduction in α was employed in order to address non-deterministic traffic patterns. This approach aims to mitigate uncertainties while facilitating the convergence towards optimal Q-values for every state–action combination according to previous research presented in [32]. To maintain ongoing learning over numerous traffic simulation epochs and to retain the learning capability, this is facilitated by the introduction of a constant parameter c , which remains fixed at 0.1. The parameter u was set to 0.8, according to the sensitivity analysis previously performed in [32].

To accommodate the integration of the mapping of speed limits and the position of speed limit zones as applicable action pairs, the modified QL algorithm is expressed as follows:

$$Q^*(s_t, [a, z]_t) := (1 - \alpha)Q(s_t, [a, z]_t) + \alpha_{(s_t, [a, z]_t)}(r_t + \gamma \max_{[a, z]' \in A} Q(s_{t+1}, [a, z]'_{t+1})), \quad (3)$$

where the selected speed limit zone position is denoted as z . In this variant of the QL algorithm, the decisions regarding the speed limit zone position z implemented upstream of the congested area on the urban motorway are also made by the agent. This represents an enhancement over conventional methods, which typically only compute speed limit values. The $[a, z]$ matrix encompasses all available actions, consisting of speed limit a selection and speed limit zone z selection from the set of available speed limit actions A and a set of speed limit zone positions $Z = \{0, 50, 100, 150, 200\}$ m before the detected congested segment, with each speed limit zone z having a fixed length of 500 m.

In [32], a sensitivity analysis was employed to assess the significance of future traffic environment states on the agent learning process. Additionally, the parameter γ was set to a specific value of 0.9, as referenced in [32]. The proportions of exploration and exploitation during the learning process was determined using the ϵ -greedy policy. This policy involves the probability of selecting random speed limit actions denoted as a and speed limit zone positions denoted as z for a given state s , drawing from the available sets of actions A and positions Z . Particularly, when the ϵ value is set to a notably high level, the likelihood of random actions a and z will be high. This ϵ -greedy policy with the dynamic ϵ value adjustments approach plays a crucial role in optimizing the learning agent's performance [32]. The ϵ value was adjusted as follows:

$$\epsilon = \begin{cases} 1, & \text{if } n \leq 300 \\ \frac{0.98}{1+(e^{(n-300)})^{0.02}} + 0.05, & \text{if } n > 300 \end{cases} \quad (4)$$

where, in the context of the ongoing simulation, denoted as n , the calculation of the Q-value for a given state–action pair involves the parameter ϵ , which is designed to promote a higher probability of exploration at the beginning of the learning process. To achieve this objective, ϵ is set to 1 during the initial 300 simulation epochs, ensuring the guaranteed selection of random actions. To diminish the likelihood of selecting actions randomly over time, a systematic reduction is applied to the parameter ϵ within the range of simulations from the 300th to the 800th simulation epoch, ultimately stabilizing at a fixed value of 0.05. The ϵ value is stabilized to 0.05 in the later stages of learning to allow the agent to adapt to possible changes in the traffic behavior. The computation of rewards for the QL algorithm was rooted in the objective of reducing the *TTS* across the entire observed urban motorway. Consequently, this reward structure incentivized the algorithm to mitigate congestion effects and prevent their occurrence. Consequently, it tends to reduce the overall *TTS* by targeting the most problematic congested area segments on the observed urban motorway.

3.2. Congestion Detection

The proposed approach is based on aggregated real-time data collected from CAVs in a determined control time step Δt , in this case, 5 min. Relevant vehicle data include the vehicle position on the given edge e_i and the corresponding vehicle speed v_j . After each time interval Δt , the mean speed v_{e_i} for every motorway edge e is calculated according to

$$v_{e_i} = \frac{\sum_{x=1}^{n_{veh}} v_x}{n_{veh}}, \quad (5)$$

where n_{veh} represents the number of vehicles on an edge i . The calculated mean speed for every edge is then used to calculate the changes in speeds between three consecutive edges. This is conducted by calculating the gradient g_{e_i} for every mean speed v_{e_i} on an edge e_i according to

$$g_{e_i} = \frac{v_{e_{i+1}} - v_{e_{i-1}}}{2}. \quad (6)$$

The calculated gradient value provides information about sudden mean speed changes between consecutive edges. If the value of g_{e_i} on an edge e_i is ≤ -0.65 , where the edge e_i is considered to be the starting edge of a congestion area. The gradient value -0.65 is chosen as a threshold value, since it represents a mean deceleration value of 1.3 m/s^2 , which is the mean deceleration value for HDVs driving with speeds above 80 km/h according to the study conducted in [33]. This value is, therefore, sufficient to conclude that the speed is declining unusually and, thus, a bottleneck is forming. All detected starting edges of a congestion area are stored in an array C_s for each control time step Δt . The end of the congested area is then found by searching for the gradient value ≥ 0.5 from the edge where the congestion started to form. The threshold gradient value of 0.5 corresponds to a mean acceleration value of 1 m/s^2 for HDVs driving with speeds ranging from 40 km/h to

70 km/h when they start to accelerate, based on research conducted in [33] and is, therefore, chosen as a threshold value to determine the end of the congested area. The traffic state is then calculated by the number of congested edges between the starting and ending congestion edges. All calculated traffic states are stored in an array D_s for each control time step Δt . If the number of congested edges exceeds 20, then the traffic state is considered to be 20 for the most severe cases of congestion, corresponding to congestion of longer than 1 km on the observed urban motorway. The overlapping of congested areas is prevented by looking at whether start points of two consecutive congested areas are further apart than 16 consecutive edges totaling 800 m to ensure that the CD-QL-DPVSL agent can be placed between the two congested areas. If the start points of the congested areas are closer than 16 consecutive edges totaling 800 m, they are combined into one larger congested area, beginning from the start of the upstream congested area and ending with the end of the downstream congested area. This method of detecting congestion is very useful, since it can detect the existence of multiple congested areas on the observed urban motorway.

When all congested areas have been detected, the proposed CD-QL-DPVSL agent is then placed directly before each of the congested area starting edges. Furthermore, this also improves the learning of the CD-QL-DPVSL agent, since it learns on multiple motorway segments under varying traffic conditions on the observed urban motorway simultaneously in a control time step Δt . The pseudocode for the proposed CD-QL-DPVSL control approach is given in Algorithm 1.

Algorithm 1 CD-QL-DPVSL

```

Define:  $S, A, Z, R$ 
Initialize:  $Q(s, [a, z]) = 0 \forall s \in S, a \in A, z \in Z$ 
For each epoch do:
  For each control time step, do:
    Initialize:  $C_s, D_s$  as empty arrays of integers
    For each edge  $i$  do:
      Calculate  $v_i$  using Equation (5)
      Calculate  $g_i$  using Equation (6)
    End for
    For each edge  $i$  do:
      if  $g_i \leq -0.65$  :
        edge  $i$  = congested area start
        append  $i$  in  $C_s$ 
        For each edge  $j$  starting from  $i$  do:
          if  $g_j > 0.5$  :
            Calculate state  $s = j - i$ 
            if  $s > 20$  :
               $s = 20$ 
            End if
            append  $s$  in  $D_s$ 
             $i = j$ 
            break
          End if
        End for
      End if
    End for
    For each state from  $D_s$  and congested area, start from  $C_s$  do:
      Choose  $a, z$  using selection policy ( $\epsilon$ -greedy)
      Execute  $a, z$ 
      Observe  $r, s'$ 
      Update  $\alpha$  using Equation (2)
      Update  $Q(s, a)$  using Equation (3)
    End for
    Update  $\epsilon$  using Equation (4)
  End for
End for

```

4. Simulation Framework

The evaluation of the proposed CD-QL-DPVSL control approach was conducted using a synthetic motorway model, which was previously introduced in our prior paper [3]. This model, depicted in Figure 2, should be noted as being unscaled compared to the

original model. Within this model, there are two on-ramps denoted as r_1 and r_2 and one off-ramp designated as s_1 , all of which are depicted in Figure 2. The on-ramps and off-ramp include acceleration and deceleration lanes, each measuring 250 m in length, and the main section of the model has no vertical slopes. The model is subdivided into 160 segmented edges, denoted as e_i , with each segment measuring 50 m in length. Figure 2 shows the case of two detected congested areas near two on-ramps. The CD-QL-DPVSL agent is placed before each of the congested areas, as described in Section 3. The simulations were executed utilizing the microscopic traffic simulator known as Simulation of Urban MObility (SUMO) [34]. The CD-QL-DPVSL control approach was externally integrated into SUMO via the TraCI interface through a Python script. This interface facilitated the retrieval of essential traffic data and enabled real-time control of the simulation, including dynamic adjustment of the speed limit values and positions of speed limit zones. Each simulated scenario represented a two-hour time frame, spanning 24 control time steps. To simulate the increased peak hour traffic demand, the traffic demand pattern depicted in Figure 3 was replicated.

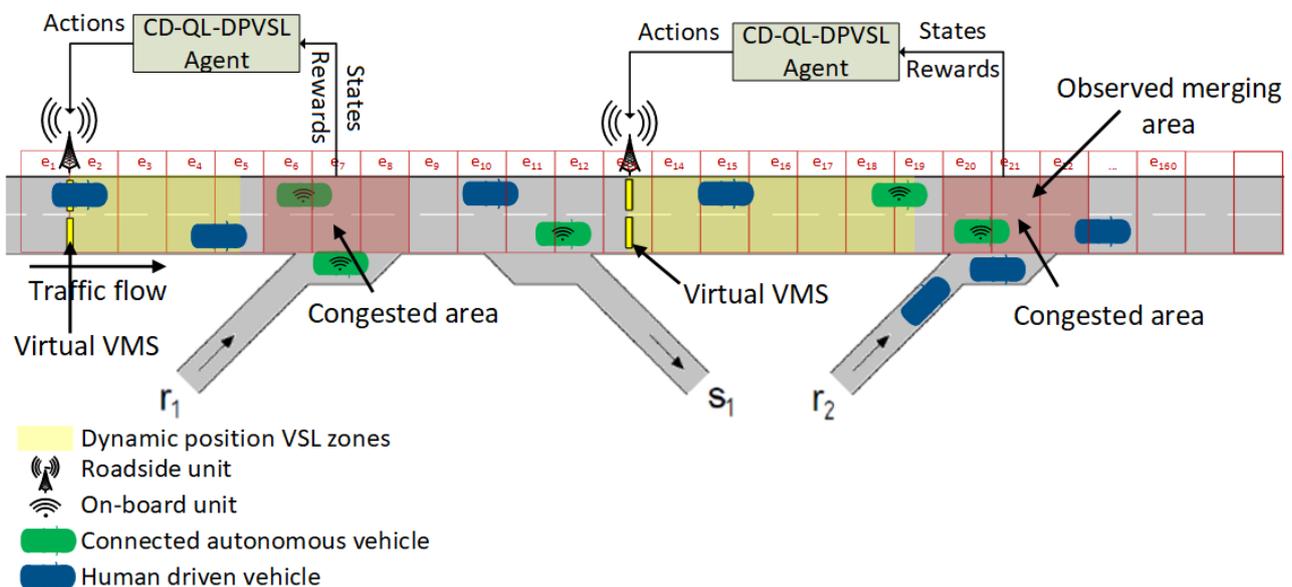


Figure 2. Simulation model configuration. Adapted with permission from Ref. [3]. 2023, Filip Vrbanic.

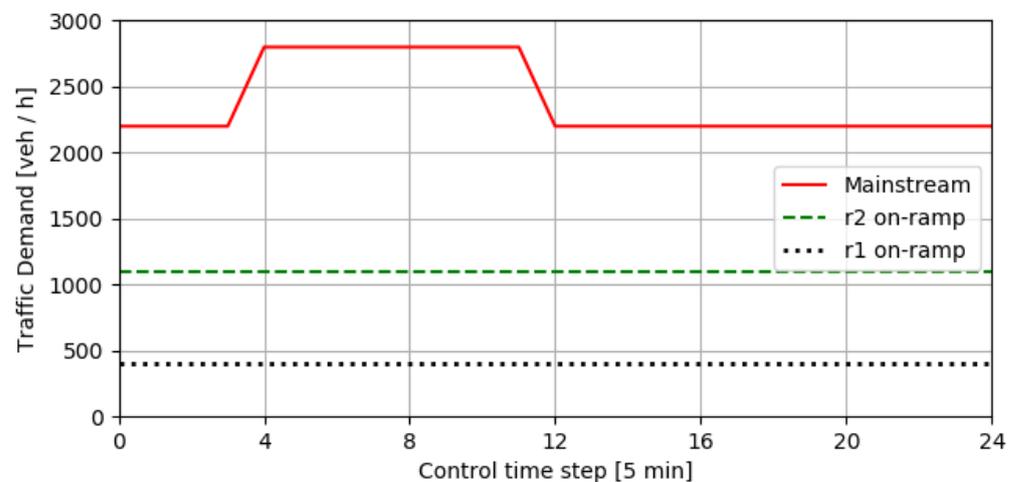


Figure 3. Traffic demand on the mainstream and on-ramps during the simulation. Reprinted with permission from Ref. [3]. 2023, Filip Vrbanic.

Traffic parameters were collected in each one-second time frame throughout every control time step Δt , which lasted for 5 min, and then the mean values were calculated. These measured traffic parameters encompassed the density (ρ), quantified in veh/km/ln, speed (v), expressed in km/h, and MTT , quantified in seconds. Furthermore, the TTS was computed in veh·h and was aggregated for the entire simulated motorway. In contrast, measurements of ρ and v were specifically taken in the observed merging area of the second on-ramp, as depicted in Figure 2. Conversely, the MTT was solely calculated for the mainstream flow of traffic, excluding the on- and off-ramps from consideration.

The SUMO microscopic traffic simulator was utilized to establish car-following models and vehicle class parameters for both CAVs and HDVs, as previously applied in prior research [3,35]. The default Krauss car-following model was used for both HDVs and CAVs [36]. It is worth noting that CAVs were assumed to exhibit reduced time headways, decreased driver imperfections, and a higher propensity of adhering to imposed speed limits when compared to HDVs.

Due to the absence of publicly accessible real-world data for CAVs, which are challenging to obtain through direct measurements in practical experiments, certain parameter values were predefined. Specifically, the parameter representing driving imperfections (σ) was assigned a value of 0.7 for HDVs and 0 for CAVs. A value of 0 implies flawless driving behavior, while lower σ values indicate more strict acceleration and deceleration actions. The parameter *SpeedDev*, which signifies the permissible deviance from the posted speed limit, was configured as 0.2 for HDVs and 0.05 for CAVs. The lane speed limit multiplier (*SpeedFactor*) remained at 1 for both CAVs and HDVs, as both lanes adhered to an identical speed limit. The minimum desired headway time parameter (τ), derived from the net time gap between the rear of the leading vehicle and the front of the following vehicle, was established at 1.1 for HDVs and 0.5 for CAVs. It is worth noting that lower τ values have been demonstrated to improve the traffic flow [8,37].

Furthermore, the impact analysis of CAVs with varying levels of automation, characterized by the σ and τ values, indicated that increasing the CAV penetration rate with lower σ and τ values enhanced the operational capacity of the road network. This resulted in higher ρ_c values on individual roads [37]. In fact, when transitioning from no CAVs to a 100% CAV penetration rate in mixed traffic flow, the ρ_c value increased by nearly 48%, as demonstrated in a previous simulation study [37]. The evaluation of the proposed CD-QL-DPVSL method encompassed six distinct simulation scenarios, each characterized by varying CAV penetration rates, spanning from 10% to 100%.

5. Results and Discussion

To assess the effectiveness of the proposed CD-QL-DPVSL control approach, its performance was compared against the following other control approaches, namely STM-QL-VSL₁ and STM-QL-VSL₂, STM-QL-DVSL [3], RB-VSL, and the no-control scenario. The CD-QL-DPVSL control policy in each distinct scenario underwent training through 10,000 simulation epochs for each mixed traffic flow scenario.

Simulations for the scenario without any control employed a constant speed limit of 130 km/h. The implementation of the RB-VSL control approach was based upon prior research efforts [32], following the principles of the Highway Capacity Manual (HCM) LoS [38]. The fundamental difference between RB-VSL and CD-QL-DPVSL resides in their methodologies for VSL control. RB-VSL relies on a traditional approach to post speed limits imposed on all vehicles using VMS, whereas CD-QL-DPVSL leverages CAVs that serve as moving VSL actuators and sensors.

The STM-QL-VSL₁ and STM-QL-VSL₂ control approaches employ predetermined positions for speed limit zones. The STM-QL-VSL₁ control approach computes and posts speed limits within a single predetermined speed limit zone closest to the observed merging area shown in Figure 2. Conversely, the STM-QL-VSL₂ control approach computes and posts speed limits in two adjacent speed limit zones nearest to the observed merging area

shown in Figure 2. Further details on the configuration of these two control approaches can be found in our referenced prior paper [3].

In contrast, the STM-QL-DVSL control approach, as described and analyzed in [3], employs the dynamic speed limit zone positioning of one VSL zone closest to the observed merging area shown in Figure 2 and does not rely on a predetermined VSL zone location. The key distinction between the proposed CD-QL-DPVSL and STM-QL-DVSL control approaches, as presented in this paper, is that the control approach introduced here dynamically identifies all congested areas on an entire analyzed motorway and selects the appropriate speed limit zone positions and speed limit accordingly. Therefore, this work represents a natural extension of our previous research efforts to create a completely dynamic VSL that primarily includes the detection of congested areas and subsequently includes the setup and application of the needed VSL control to alleviate the detected congestion. Table 1 demonstrates the fundamental distinctions between CD-QL-DPVSL and all other evaluated control approaches. In summary, the key enhancement of the CD-QL-DPVSL control approach lies in its ability to identify multiple congested areas across the entire motorway and determine speed limits and speed limit zone positions for each identified congested area.

Table 1. Comparison of the analyzed control approaches.

Control Approach	Traffic State Estimation	Actions
No control	-	-
RB-VSL	Fixed measurement in the observed merging area	Speed limit
STM-QL-VSL ₁	Dynamic measurement in the observed merging area	Speed limit for a single fixed-speed limit zone position
STM-QL-VSL ₂	Dynamic measurement in the observed merging area	Speed limit for two fixed, adjacent speed limit zone positions
STM-QL-DVSL	Dynamic measurement in the observed merging area	Speed limit and speed limit zone position
CD-QL-DPVSL	Dynamic spatial measurement of multiple congested areas	Speed limits and speed limit zone positions for all congested areas

The outcomes for all examined traffic scenarios are detailed in Table 2. These findings originate from a chosen representative simulation, which exemplifies the average outcome observed over the final 500 simulation epochs for every mixed traffic flow scenario. According to the findings, CD-QL-DPVSL consistently demonstrated a better performance compared to all other control strategies in every simulated scenario. This advantage was particularly pronounced in its ability to reduce both *TTS* and *MTT* throughout the whole simulated urban motorway.

In contrast, the RB-VSL algorithm demonstrated limited effectiveness and, in some cases, even underperformed when compared to the no-control case. The exception was seen for scenario 2, where predefined speed limit control rules based on HCM LoS density thresholds managed to improve MoEs compared to the no-control case.

STM-QL-VSL₁ and STM-QL-VSL₂ exhibited more significant improvements at lower CAV penetration rates, with their efficacy gradually diminishing at scenarios with higher CAV penetration rates. On the other hand, the STM-QL-DVSL control approach demonstrated the ability to enhance MoEs more prominently at lower penetration rates. This enhancement can be primarily attributed to the advanced driving capabilities of CAVs when compared to HDVs, coupled with the utilization of microscopic-level traffic state measurements (with each CAV acting as a mobile sensor). Nonetheless, as the quantity of

CAVs within the mixed traffic flow grew, the advantageous impacts of the STM-QL-DVSL control approach gradually diminished.

The CD-QL-DPVSL approach exhibited a superior performance consistently across all examined scenarios with varying CAV penetration rates. Notably, it demonstrated notable enhancements in traffic MoEs in the 100% CAV penetration rate scenario, a distinction not observed in the performance of the other control approaches being compared.

Table 2. Obtained performance levels for defined scenarios with different CAV penetration rates [3].

Number	Scenario		Results				Improvement			
			Motorway Segment		Observed Merging Area		Motorway Segment		Observed Merging Area	
			TTS [veh-h]	MTT [s]	Mean v [km/h]	Mean ρ [veh/km/ln]	TTS [%]	MTT [%]	Mean v [%]	Mean ρ [%]
1	10%	No control	713.0	373.3	61.5	36.9	-	-	-	-
		RB-VSL	717.4	375.6	62.3	36.5	-0.6	-0.6	1.3	1.1
		STM-QL-VSL ₁	702.4	366.5	62.8	35.3	1.5	1.8	2.1	4.3
		STM-QL-VSL ₂	712.2	372.2	61.4	36.9	0.1	0.3	-0.2	0.0
		STM-QL-DVSL	691.6	360.6	64.2	34.8	3.0	3.4	4.4	5.7
		CD-QL-DPVSL	671.1	352.2	66.9	32.5	5.9	5.6	8.8	11.9
2	30%	No control	664.4	340.5	75.1	29.5	-	-	-	-
		RB-VSL	649.3	333.0	75.7	29.2	2.3	2.2	0.8	0.7
		STM-QL-VSL ₁	642.8	330.3	76.7	27.9	3.3	3.0	2.1	5.4
		STM-QL-VSL ₂	635.8	328.2	77.9	27.4	4.3	3.6	3.7	7.1
		STM-QL-DVSL	628.3	324.6	79.5	26.2	5.4	4.7	5.9	11.2
		CD-QL-DPVSL	613.8	319.5	82.2	23.6	7.6	6.2	9.4	20.0
3	50%	No control	628.1	315.2	81.0	27.3	-	-	-	-
		RB-VSL	627.4	315.5	81.3	26.6	0.1	-0.1	0.4	2.6
		STM-QL-VSL ₁	618.6	311.7	83.0	25.3	1.5	1.1	2.5	7.3
		STM-QL-VSL ₂	620.7	313.0	83.3	25.3	1.2	0.7	2.8	7.3
		STM-QL-DVSL	609.9	309.4	85.7	24.1	2.9	1.8	5.8	11.7
		CD-QL-DPVSL	594.6	305.4	84.2	25.4	5.3	3.1	3.9	7.0
4	70%	No control	548.4	278.6	95.4	19	-	-	-	-
		RB-VSL	565.1	284.7	92.7	21.5	-3.0	-2.2	-2.8	-13.2
		STM-QL-VSL ₁	542.6	276.7	96.3	18.3	1.1	0.7	0.9	3.7
		STM-QL-VSL ₂	548.9	279.1	95.4	19.2	-0.1	-0.2	0.0	-1.1
		STM-QL-DVSL	546.5	279	95.9	19.4	0.4	-0.1	0.5	-2.1
		CD-QL-DPVSL	535.8	278.6	96.6	18.8	2.3	0.0	1.3	1.0
5	90%	No control	489.2	254.2	103.4	16.8	-	-	-	-
		RB-VSL	506.5	259.2	100.2	18.9	-3.5	-2.0	-3.1	-12.5
		STM-QL-VSL ₁	488.5	253.7	103.4	16.4	0.1	0.2	0.0	2.4
		STM-QL-VSL ₂	489.2	253.8	103.6	16.2	0.0	0.2	0.2	3.6
		STM-QL-DVSL	486.4	252.9	104.2	15.9	0.6	0.5	0.8	5.4
		CD-QL-DPVSL	473.2	255.1	107.0	14.7	3.3	-0.3	2.7	12.5
6	100%	No control	412.9	230.7	112.5	12.3	-	-	-	-
		RB-VSL	412.9	230.7	112.5	12.3	0.0	0.0	0.0	0.0
		STM-QL-VSL ₁	412.9	230.7	112.5	12.3	0.0	0.0	0.0	0.0
		STM-QL-VSL ₂	412.9	230.7	112.5	12.3	0.0	0.0	0.0	0.0
		STM-QL-DVSL	412.9	230.7	112.5	12.3	0.0	0.0	0.0	0.0
		CD-QL-DPVSL	412.2	233.0	113.6	11.8	0.2	-1.0	1.0	4.0

The results obtained for scenarios 1 and 2, which include measures such as ρ_c , v_c , and TTS , are displayed in Figure 4. These outcomes originate from an exemplary simulation based on an average outcome among the final 500 simulation epochs after a series of 10,000 simulation epochs for each simulated scenario. In scenario 1, the application of

CD-QL-DPVSL demonstrated an improvement of 5.9% in the *TTS* compared to the no-control case. Conversely, the utilization of RB-VSL resulted in a slight 0.6% increase in the *TTS*. For scenario 1, the implementations of STM-QL-VSL₁, STM-QL-VSL₂, and STM-QL-DVSL only brought about marginal improvements in the *TTS* of 1.5%, 0.1%, and 3.0%, respectively, compared to the no-control case. In scenario 2, the proposed CD-QL-DPVSL strategy was proven to be better than the rest of the control methods, leading to a significant 7.6% reduction in the *TTS*. In contrast, RB-VSL, STM-QL-VSL₁, STM-QL-VSL₂, and STM-QL-DVSL resulted in reductions in the *TTS* by 2.3%, 3.3%, 4.3%, and 5.4%, respectively. All Reinforcement-Learning-based VSL control approaches exhibited improvements in the mean ρ , mean v , and *MTT* compared to both the no-control and RB-VSL control cases. Between these control approaches, CD-QL-DPVSL exhibited the most promising performance. Using real-time CAV data for the state estimation and their ability to act as VSL actuators on an urban motorway, the proposed CD-QL-DPVSL approach significantly improved several traffic MoEs (primarily *TTS* and *MTT*). These enhancements measured in scenario 1 imply that, even with a low CAV penetration rate of 10%, the evaluated control approach has access to enough input data to estimate the traffic flow state, detect congestion areas, and make the best-computed decisions about speed limit zone placement and the speed limits imposed on CAVs.

A more precise representation of traffic flow conditions enables the agent to learn actions in multiple congested areas concurrently, without requiring knowledge of the specific congestion causes. This enhancement significantly boosts the performance of the proposed CD-QL-DPVSL. Furthermore, the findings suggest that the impact of the growing CAV penetration rate on the performance of the evaluated control approaches becomes less prominent after the penetration rate of CAVs surpasses 30%. Consequently, it can be concluded that the performance of the proposed CD-QL-DPVSL control approach does not increase linearly with the increasing penetration rate of CAVs.

Once sufficient input data from CAVs become obtainable, the quality of the state estimation becomes adequate to ensure that the operating conditions of the modeled VSL control approach are effective. In this study, even at a modest CAV penetration rate of 10%, the state estimation quality was proven to be sufficient to ensure the proper functioning of the VSL control approach. As a result, the improvements observed were more evident in scenarios 1 and 2, where the low penetration rate of CAVs supplied enough data to estimate traffic states, enabling the detection of congestion areas and informed decisions regarding speed limit zone placement and speed limits imposed on CAVs.

The incorporation of CAVs into the mixed traffic flow impacted the metrics of *TTS* and *MTT*, as illustrated in Figure 5. The reduced *TTS* can be primarily attributed to the more strict and precise driving characteristics of CAVs, which are distinguished by reduced vehicle headways and quicker reaction times. On the other hand, as the CAV penetration rate increased from 10% to 100%, the *MTT* improvements started to diminish and even worsen in the high-CAV-penetration-rate scenarios 5 and 6. This is mainly attributed to the goal of the control approach to optimize the *TT* in the congested areas, resulting in a reduced *MTT* for mainstream vehicles but allowing on-ramp vehicles to merge more easily and, therefore, travel faster. As a result, the *TTS* improved, but the *MTT* was slightly worse in those scenarios.

Comparing the observed merging area mean ρ and mean v , the CD-QL-DPVSL performed better than the second-best STM-QL-DVSL control approach in all scenarios, except scenario 3. This performance is attributed to the proposed CD-QL-DPVSL objective, which optimizes the entire analyzed urban motorway by placing multiple CD-QL-DPVSL agents where it detects congested areas. Previously analyzed control approaches only measure traffic states in the observed merging area and post speed limits accordingly. Consequently, they have the objective of improving the observed merging area MoEs, while CD-QL-DPVSL has the objective of improving MoEs in every congested area at each control time step Δt . Therefore, the objective is not strictly set to the only observed merging area. This increased performance is due to the more refined control of the entire observed urban

motorway, again by detecting multiple congested areas and placing speed limit zones accordingly. This causes the inflow of vehicles to the observed merging area to be more controlled and harmonized, resulting in a better performance in terms of the mean ρ and mean v . Figure 6 illustrates a space–time diagram of the mean speeds for the CD-QL-DPVSL control approach (Figure 6(a–f)) and a space–time diagram of the mean speeds for the STM-QL-DVSL control approach (Figure 6(g–l)) on all motorway segments during the simulation in scenarios with CAV penetration rates spanning from 10% to 100%. The green rectangles represent the VSL zone positions during the simulation. One key observation is that, in all scenarios, speeds were more harmonized, while the speeds in the observed merging area were generally higher for the CD-QL-DPVSL control case compared to the second-best STM-QL-DVSL control case. Furthermore, the congestion in the observed merging area for the CD-QL-DPVSL control case is less pronounced compared to that in the STM-QL-DVSL control case.

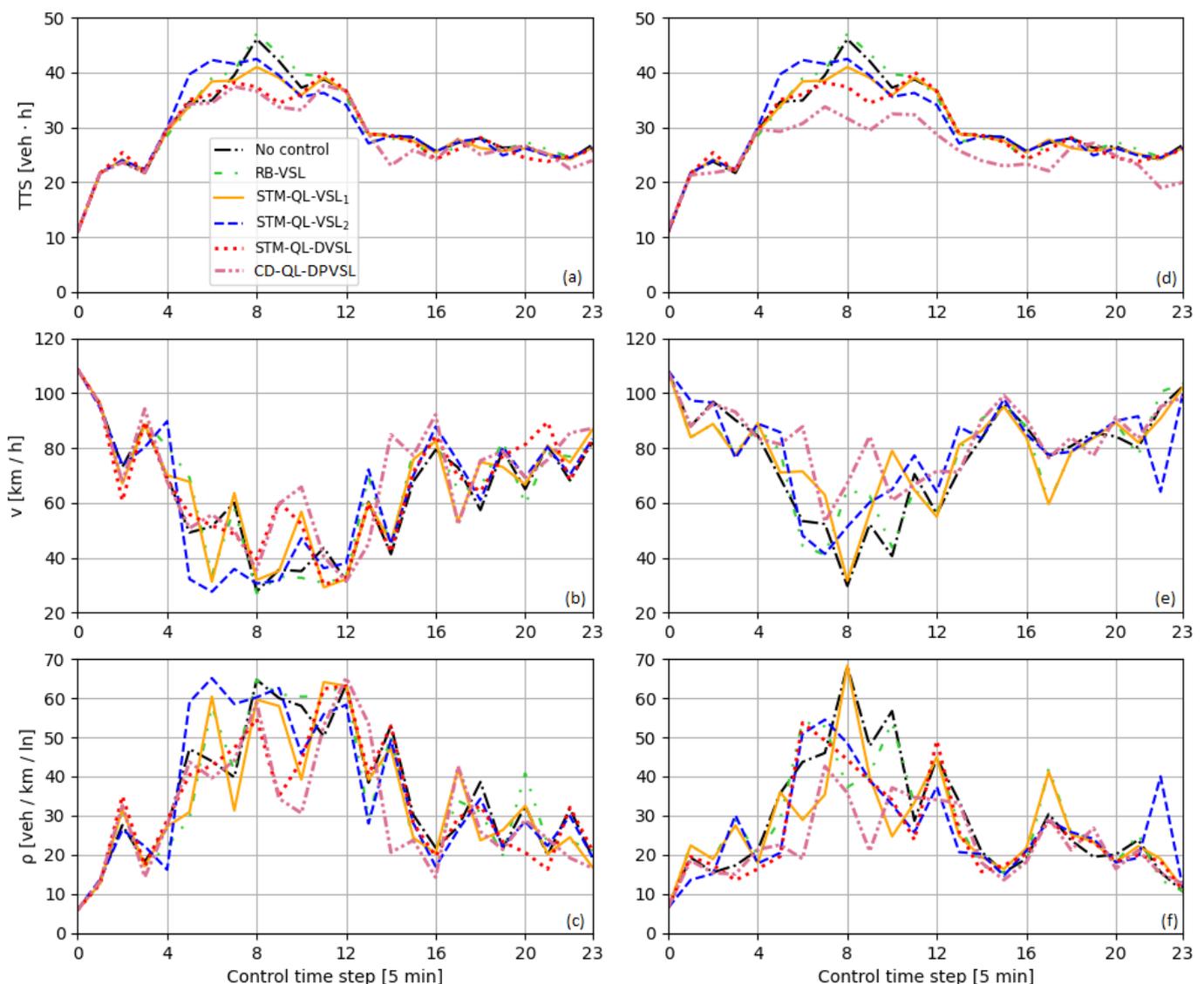


Figure 4. Obtained TTS , speeds, and densities for scenario 1 (a–c) and scenario 2 (d–f).

On the other hand, the improved CD-QL-DPVSL performance in scenarios 5 and 6 is mainly attributed to the more refined control, again by detecting multiple congested areas and placing speed limit zones accordingly. In scenario 5, the proposed CD-QL-DPVSL outperformed the second-best STM-QL-DVSL control approach in terms of the measured TTS , mean ρ , and mean v by 2.7%, 2.7%, and 7.5%, respectively. On the other hand, the

proposed CD-QL-DPVSL worsened the MTT by 0.9% compared to the second-best STM-QL-DVSL control approach, mainly due to the goal of optimizing the TT in the detected congested areas, which resulted in a worsened MTT for the mainstream vehicles, but ensured that the on-ramp vehicles merged more easily and therefore traveled faster. In scenario 6, all other compared VSL approaches did not affect the traffic flow, while CD-QL-DPVSL managed to improve the TTS , MTT , mean ρ , and mean v slightly by 0.2%, 0.9%, 1.0%, and 4.9%, respectively.

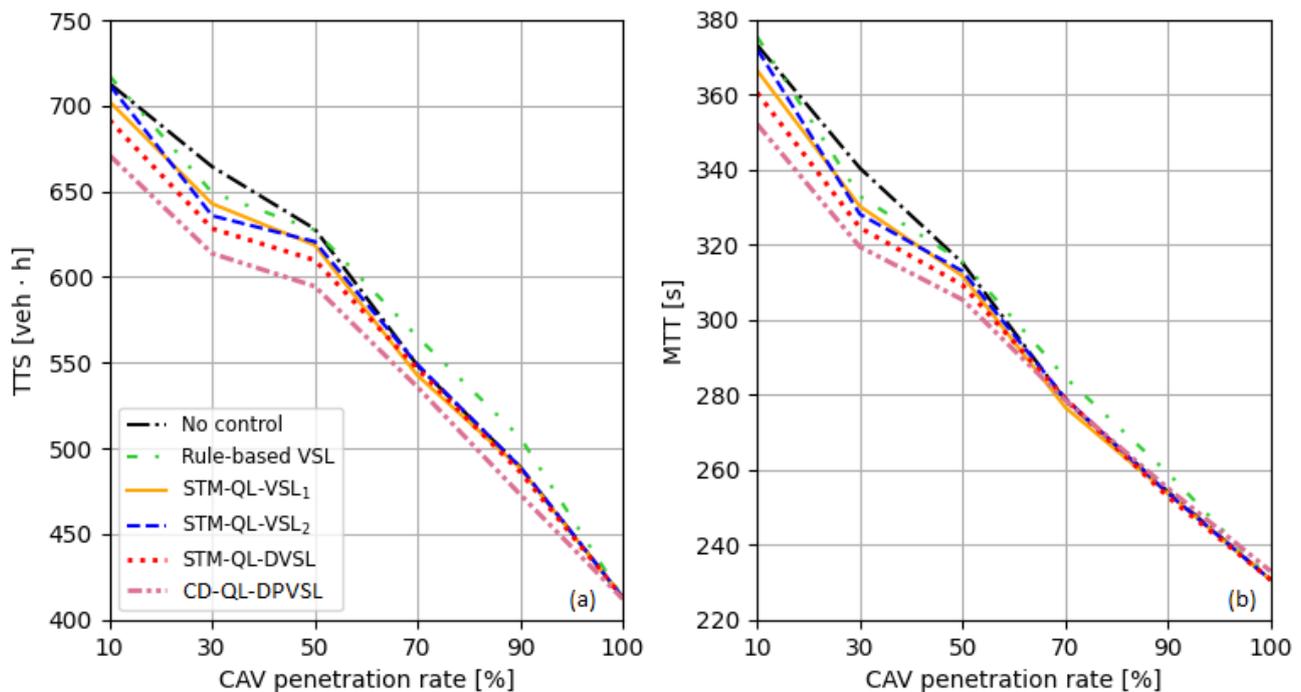


Figure 5. Changes in the TTS (a) and MTT (b) for different CAV penetration rates.

The choice of suitable speed limit zone positions and the specific speed limits adopted in scenario 1 and scenario 2 are visualized in Figure 7. A noteworthy observation when contrasting the placement of the speed limit zones between these two scenarios is that, as the presence of CAVs increases, the CD-QL-DPVSL control approach appears to identify congested areas more effectively. Consequently, it deploys multiple speed limit zones along the observed urban motorway. This phenomenon can be ascribed to the increased volume of vehicles contributing data to detect congested areas and calculate traffic states. Furthermore, the CD-QL-DPVSL control approach opted for reduced speed limits in scenario 2 when compared to scenario 1. With more input data in scenario 2, the control approach detected more congested areas and placed more speed limit zones with generally lower speed limits. Furthermore, speed limit zones near the observed merging area were placed further downstream between the 50th and 100th minute of the simulation in scenario 2, while in scenario 1, the speed limit zones were placed further upstream of the observed merging area. This concludes that the congested area detected near the observed merging area is more severe in scenario 1.

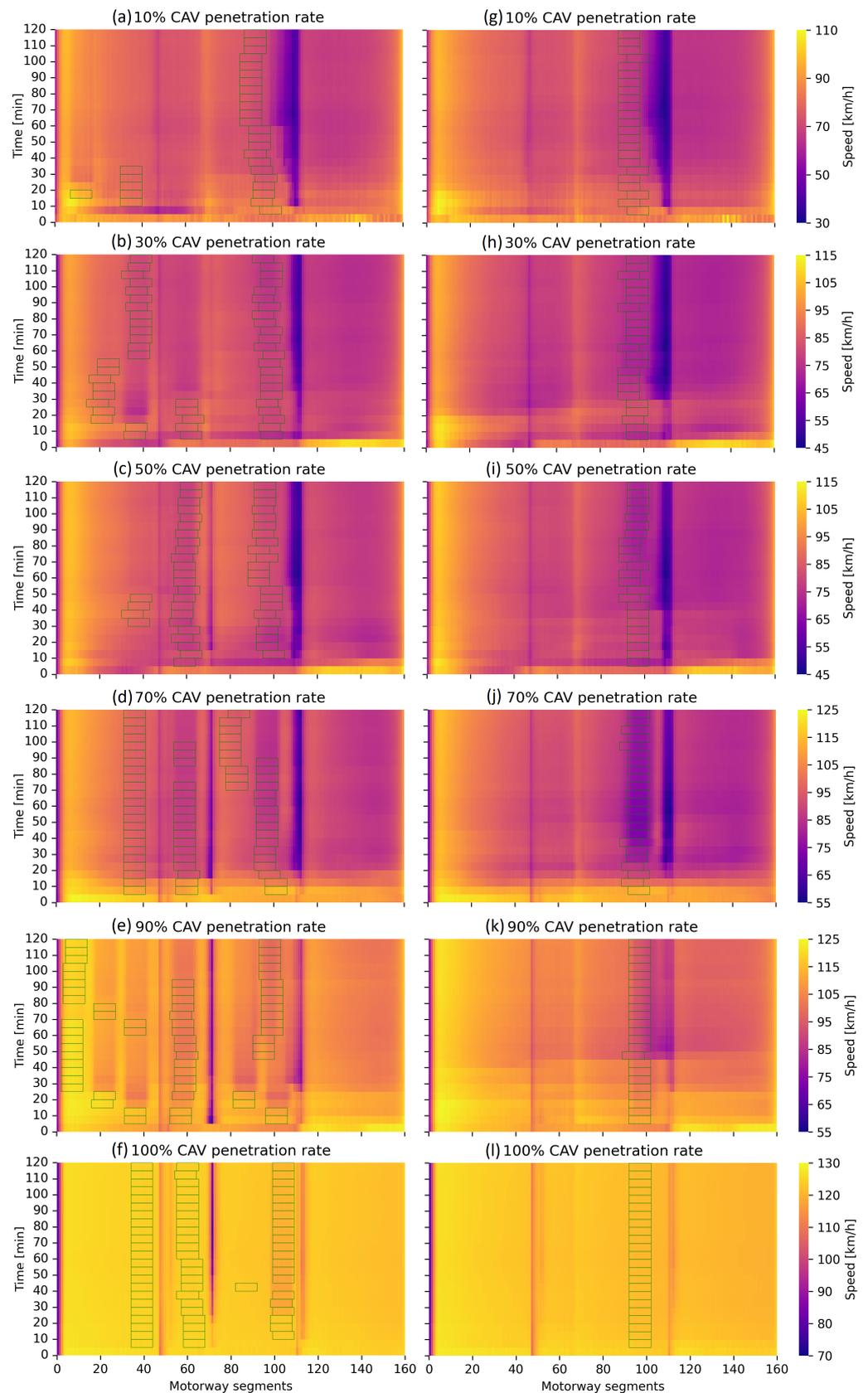


Figure 6. Space–time diagram of the mean CAV speeds on segments and speed limit zone placements for the CD-QL-DPVSL (a–f) and STM-QL-DVSL (g–l) control approaches in various CAV penetration rate scenarios.

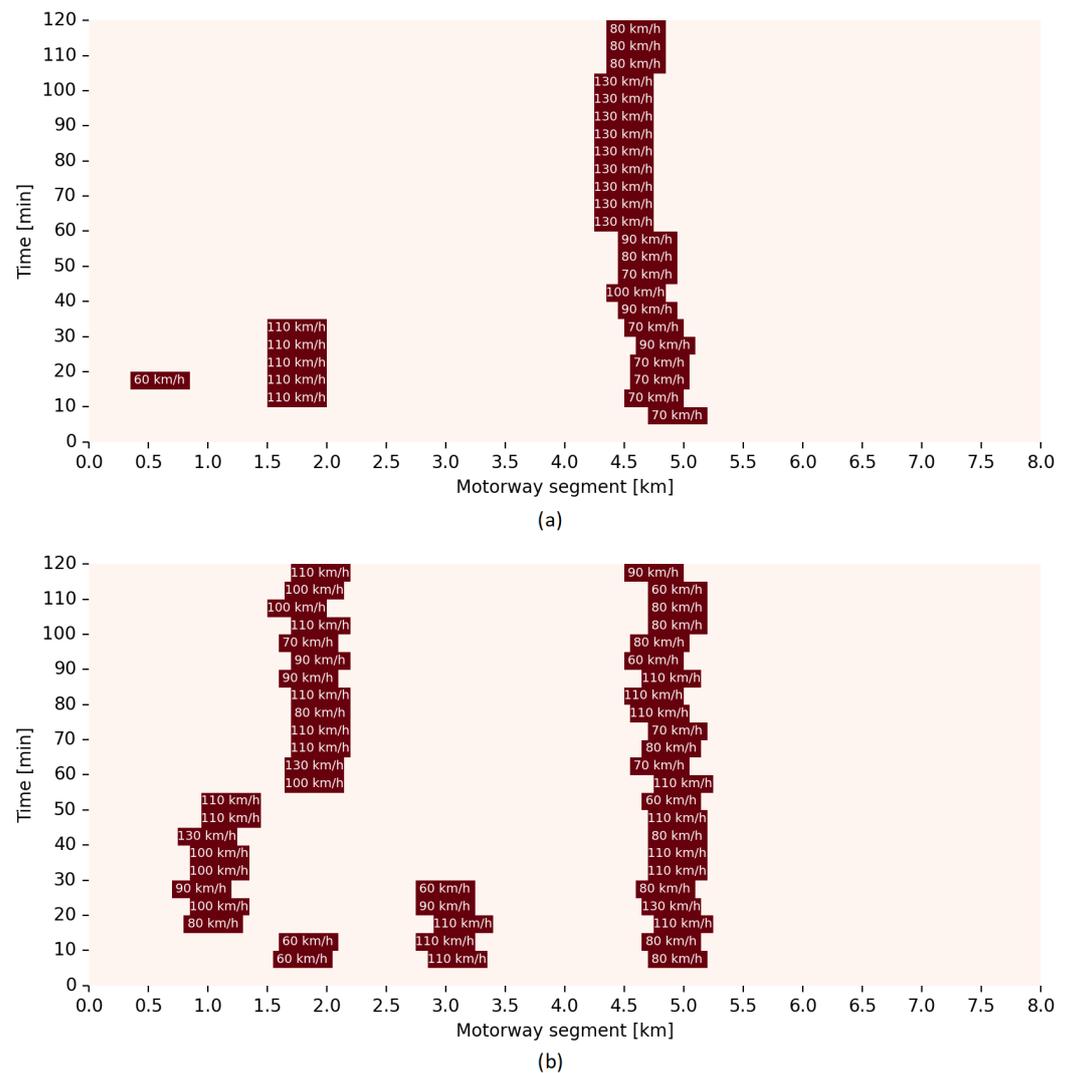


Figure 7. Space–time diagram of the computed speed limits and speed limit zone positions for scenario 1 (a) and scenario 2 (b).

6. Conclusions

The primary aim of this study was to propose a VSL control approach that utilizes CAVs as mobile actuators and sensors. The proposed CD-QL-DPVSL control approach derives traffic state estimates based on gradient values calculated for mean speeds along each segment of the observed urban motorway. The research also investigates how the detection of the congested areas and the dynamic positioning of CD-QL-DPVSL zones affect traffic flow. The most important conclusions of this paper are as follows:

- The gradient values calculated for the mean speeds collected from CAVs along each segment of the observed urban motorway can be used to estimate congested areas;
- Real-time CAV data can be used to estimate traffic states, improving the performance of the VSL control approach;
- The detection of congested areas and the dynamic positioning of CD-QL-DPVSL zones outperforms other analyzed control approaches for all measured MoEs.

To evaluate the algorithm's performance under various mixed traffic flow scenarios, a simulation framework was employed. The results indicate that the CD-QL-DPVSL control approach performs better for all MoEs in contrast to other control approaches and the no-control scenario. The most notable improvements with the CD-QL-DPVSL control approach are observed in the scenario with a 30% penetration rate for the CAVs.

Remarkably, enhancements in all MoEs were evident even at low CAV numbers, including scenarios with 10% and 30% penetration rates for CAVs. The CD-QL-DPVSL control approach effectively identifies congested areas, primarily near on-ramps and off-ramps, and adjusts speed limit zones accordingly. Notably, in a scenario with a 100% penetration rate for CAVs, CD-QL-DPVSL succeeded in enhancing MoEs, whereas all other analyzed control approaches had no impact on the measured MoEs.

The scope of this paper is limited to personal vehicles, including both CAVs and HDVs, within a mixed traffic flow at different penetration rates. Moreover, CAVs must adhere to the imposed speed limits within designated speed limit zones, in accordance with the limitations imposed by the specific physical attributes governing the maximum acceleration capabilities of each CAV. The simulation model has no physical VMSs, as it becomes obsolete in mixed traffic flows containing CAVs, and virtual VMSs are placed at the start of each of the applicable VSL zones. In this paper, the CD-QL-DPVSL control approach operates under the assumption of error-free data transmission and zero information latency between CAVs and the control agent.

Future research endeavors will examine the multi-agent approach to the CD-QL-DPVSL control approach, incorporating the dynamic lengths of VSL zones. Additionally, the analysis will be extended to scenarios with increased traffic demand to understand the influence of CAV penetration rates on the VSL requirements. The influence of adding heavy-duty vehicles and buses into the mixed traffic flow will also be examined. Dynamic adjustments to the length of VSL zones will be investigated as an extended action selection set and a potential avenue for future research. Moreover, the study will examine urban motorways with more intricate geometric designs, including vertical slopes.

Author Contributions: The conceptualization of the study was conducted by F.V., M.G., M.M. and E.I. The funding acquisition was conducted by E.I. The development and design of the methodology was conducted by F.V., M.G., M.M. and E.I. The writing of the original draft and preparation of the paper were conducted by F.V. and M.G. All authors contributed to the writing of the paper and the final editing. Supervision was conducted by E.I. Visualizations were conducted by F.V. Implementation of the computer code and supporting algorithms was conducted by F.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the University of Zagreb and the Faculty of Transport and Traffic Sciences under the grants “Real-time passenger information model in traffic safety”, by the Croatian Science Foundation under the project IP-2020-02-5042, and by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical concerns.

Acknowledgments: This research was carried out within the activities of the Centre of Research Excellence for Data Science and Cooperative Systems supported by the Ministry of Science and Education of the Republic of Croatia.

Conflicts of Interest: The authors declare no conflict of interest. The funding institutions had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AV	Autonomous Vehicle
CAV	Connected Autonomous Vehicle
CD-QL-DPVSL	Congestion Detection Q-Learning Dynamic Position Variable Speed Limit
HDV	Human-Driven Vehicle
HCM	Highway Capacity Manual
ITS	Intelligent Transportation Systems
LoS	Level of Service
MoE	Measure of Effectiveness

MTT	Mean Travel Time
OBU	On-Board Unit
QL	Q-Learning
RB-VSL	Rule-Based Variable Speed Limit
RL	Reinforcement Learning
RM	Ramp Metering
RSU	Road Side Unit
STM	Speed Transition Matrix
STM-QL-VSL	Speed-Transition-Matrices-based Q-Learning Variable Speed Limit
STM-QL-DVSL	Speed-Transition-Matrices-based Q-Learning Dynamic Variable Speed Limit
SUMO	Simulation of Urban Mobility
TTS	Total Time Spent
TT	Travel Time
TTT	Total Travel Time
VMS	Variable Message Sign
VSL	Variable Speed Limit

References

- Müller, E.; Carlson, R.; Kraus, W.J.; Papageorgiou, M. Microsimulation analysis of practical aspects of traffic control with variable speed limits. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 512–523. [\[CrossRef\]](#)
- Vrbanić, F.; Ivanjko, E.; Kušić, K.; Cakija, D. Variable Speed Limit and Ramp Metering for Mixed Traffic Flows: A Review and Open Questions. *Appl. Sci.* **2021**, *11*, 2574. [\[CrossRef\]](#)
- Vrbanić, F.; Tišljarić, L.; Majstorović, Z.; Ivanjko, E. Reinforcement Learning-Based Dynamic Zone Placement Variable Speed Limit Control for Mixed Traffic Flows Using Speed Transition Matrices for State Estimation. *Machines* **2023**, *11*, 479. [\[CrossRef\]](#)
- Papageorgiou, M.; Kosmatopoulos, E.; Papamichail, I. Effects of Variable Speed Limits on Motorway Traffic Flow. *Transp. Res. Rec. J. Transp. Res. Board* **2008**, *2047*, 37–48. [\[CrossRef\]](#)
- Lee, C.; Hellinga, B.; Saccomanno, F. Evaluation of variable speed limits to improve traffic safety. *Transp. Res. Part C Emerg. Technol.* **2006**, *14*, 213–228. [\[CrossRef\]](#)
- Carlson, R.C.; Papamichail, I.; Papageorgiou, M.; Messmer, A. Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering. *Transp. Sci.* **2010**, *44*, 238–253. [\[CrossRef\]](#)
- Cremer, M. *Der Verkehrsfluss auf Schnellstrassen: Modelle, Überwachung, Regelung*; Springer: Berlin/Heidelberg, Germany, 1979. [\[CrossRef\]](#)
- Ye, L.; Yamamoto, T. Evaluating the impact of connected and autonomous vehicles on traffic safety. *Phys. A Stat. Mech. Its Appl.* **2019**, *526*, 121009. [\[CrossRef\]](#)
- Olia, A.; Razavi, S.; Abdulhai, B.; Abdelgawad, H. Traffic capacity implications of automated vehicles mixed with regular vehicles. *J. Intell. Transp. Syst. Technol. Plann. Oper.* **2018**, *22*, 244–262. [\[CrossRef\]](#)
- Wang, Q.; Li, B.; Li, Z.; Li, L. Effect of connected automated driving on traffic capacity. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 633–637. [\[CrossRef\]](#)
- Li, Z.; Liu, P.; Xu, C.; Duan, H.; Wang, W. Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 3204–3217. [\[CrossRef\]](#)
- Xie, J.; Qi, L. Combined On-line Variable Speed Limit and Ramp Metering Control for Highway Bottleneck. In Proceedings of the 2022 IEEE International Conference on Networking, Sensing and Control (ICNSC), Shanghai, China, 15–18 December 2022; pp. 1–6. [\[CrossRef\]](#)
- Li, Y.; Li, Z.; Wang, H.; Wang, W.; Xing, L. Evaluating the safety impact of adaptive cruise control in traffic oscillations on freeways. *Accid. Anal. Prev.* **2017**, *104*, 137–145. [\[CrossRef\]](#)
- Amini, M.; Hatwagner, M.F.; Mikulai, G.C.; Koczy, L.T. An intelligent traffic congestion detection approach based on fuzzy inference system. In Proceedings of the 2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, 19–21 May 2021; pp. 97–104. [\[CrossRef\]](#)
- Zarindast, A.; Sharma, A. Big Data Application in Congestion Detection and Classification Using Apache Spark. *arXiv* **2021**, arXiv:cs.DC/2101.06524.
- Li, S.; Cheng, Y.; Jin, P.; Ding, F.; Li, Q.; Ran, B. A feature-based approach to large-scale freeway congestion detection using full cellular activity data. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 1323–1331. [\[CrossRef\]](#)
- Shaaban, K.; Khan, M.A.; Hamila, R. Literature Review of Advancements in Adaptive Ramp Metering. *Procedia Comput. Sci.* **2016**, *83*, 203–211. [\[CrossRef\]](#)
- Tišljarić, L.; Vrbanić, F.; Ivanjko, E.; Carić, T. Motorway Bottleneck Probability Estimation in Connected Vehicles Environment Using Speed Transition Matrices. *Sensors* **2022**, *22*, 2807. [\[CrossRef\]](#)
- Wu, Y.; Tan, H.; Qin, L.; Ran, B. Differential variable speed limits control for freeway recurrent bottlenecks via deep actor-critic algorithm. *Transp. Res. Part C Emerg. Technol.* **2020**, *117*, 102649. [\[CrossRef\]](#)

20. Li, D.; Wagner, P. Impacts of gradual automated vehicle penetration on motorway operation: A comprehensive evaluation. *Eur. Transp. Res. Rev.* **2019**, *11*, 36. [[CrossRef](#)]
21. Coppola, A.; Di Costanzo, L.; Pariota, L.; Bifulco, G.N. Fuzzy-Based Variable Speed Limits System Under Connected Vehicle Environment: A Simulation-Based Case Study in the City of Naples. *IEEE Open J. Intell. Transp. Syst.* **2023**, *4*, 267–278. [[CrossRef](#)]
22. Fauchet, E.; Laharotte, P.; Bhattacharyya, K.; Ladino, A.; El Faouzi, N.E. A Low Market Penetration Rate friendly Variable Speed Limit System based on a PID Controller under Connected and Automated Vehicle Environment. In Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China, 8–12 October 2022; pp. 2350–2357. [[CrossRef](#)]
23. Yu, M.; "David" Fan, W. Optimal variable speed limit control in connected autonomous vehicle environment for relieving freeway congestion. *J. Transp. Eng. Part A Syst.* **2019**, *145*, 04019007. [[CrossRef](#)]
24. Vinitzky, E.; Parvate, K.; Kreidieh, A.; Wu, C.; Bayen, A. Lagrangian Control through Deep-RL : Applications to Bottleneck Decongestion. In Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 759–765. [[CrossRef](#)]
25. Du, Y.; Makridis, M.A.; Tampère, C.M.; Kouvelas, A.; ShangGuan, W. Adaptive control with moving actuators at motorway bottlenecks with connected and automated vehicles. *Transp. Res. Part C Emerg. Technol.* **2023**, *156*, 104319. [[CrossRef](#)]
26. Han, L.; Zhang, L.; Guo, W. Optimal Differential Variable Speed Limit Control in a Connected and Autonomous Vehicle Environment for Freeway Off-Ramp Bottlenecks. *J. Transp. Eng. Part A Syst.* **2023**, *149*, 04023009. [[CrossRef](#)]
27. Di, Y.; Zhang, W.; Ding, H.; Zheng, X.; Bai, H. Integrated Control for Mixed CAV and CV Traffic Flow in Expressway Merge Zones Combined with Variable Speed Limit, Ramp Metering, and Lane Changing. *J. Transp. Eng. Part A Syst.* **2023**, *149*, 04022140. [[CrossRef](#)]
28. Pan, T.; Guo, R.; Lam, W.H.; Zhong, R.; Wang, W.; He, B. Integrated optimal control strategies for freeway traffic mixed with connected automated vehicles: A model-based reinforcement learning approach. *Transp. Res. Part C Emerg. Technol.* **2021**, *123*, 102987. [[CrossRef](#)]
29. Watkins, C.J.C.H.; Dayan, P. Q-learning. *Mach. Learn.* **1992**, *8*, 279–292. [[CrossRef](#)]
30. Walraven, E.; Spaan, M.T.; Bakker, B. Traffic flow optimization: A reinforcement learning approach. *Eng. Appl. Artif. Intell.* **2016**, *52*, 203–212. [[CrossRef](#)]
31. Wang, C.; Zhang, J.; Xu, L.; Li, L.; Ran, B. A New Solution for Freeway Congestion: Cooperative Speed Limit Control Using Distributed Reinforcement Learning. *IEEE Access* **2019**, *7*, 41947–41957. [[CrossRef](#)]
32. Vrbanić, F.; Miletić, M.; Tišljarić, L.; Ivanjko, E. Influence of Variable Speed Limit Control on Fuel and Electric Energy Consumption, and Exhaust Gas Emissions in Mixed Traffic Flows. *Sustainability* **2022**, *14*, 932. [[CrossRef](#)]
33. Bokare, P.; Maurya, A. Acceleration-Deceleration Behaviour of Various Vehicle Types. *Transp. Res. Procedia* **2017**, *25*, 4733–4749. [[CrossRef](#)]
34. Lopez, P.A.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flötteröd, Y.P.; Hilbrich, R.; Lücken, L.; Rummel, J.; Wagner, P.; Wiessner, E. Microscopic Traffic Simulation using SUMO. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, Hawaii, 4–7 November 2018; pp. 2575–2582. [[CrossRef](#)]
35. Li, D.; Wagner, P. A novel approach for mixed manual/connected automated freeway traffic management. *Sensors* **2020**, *20*, 1757. [[CrossRef](#)]
36. Krauss, S.; Wagner, P.; Gawron, C. Metastable states in a microscopic model of traffic flow. *Phys. Rev. E* **1997**, *55*, 5597–5602. [[CrossRef](#)]
37. Lu, Q.; Tettamanti, T. Impacts of autonomous vehicles on the urban fundamental diagram. In Proceedings of the 5th International Conference on Road and Rail Infrastructure, CETRA 2018, Zadar, Croatia, 17–19 May 2018.
38. Elefteriadou, L.A. (Ed.) *Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis*; Transportation Research Board, The National Academies Press: Washington, DC, USA, 2016. [[CrossRef](#)]

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