



Enhancing reinforcement learning-based adaptive traffic signal control in emerging mixed traffic environments through CAV data: A simulation

Downloaded from: <https://research.chalmers.se>, 2026-10-11 00:40 UTC

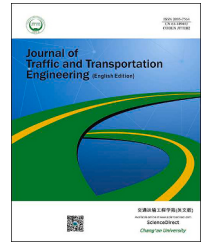
Citation for the original published paper (version of record):

Fu, C., Lu, Z., Ding, N. et al (2026). Enhancing reinforcement learning-based adaptive traffic signal control in emerging mixed traffic environments through CAV data: A simulation study considering driver behaviour. Journal of Traffic and Transportation Engineering (English Edition), 13(4): 1114-1131.
<http://dx.doi.org/10.1016/j.jtte.2024.07.002>

N.B. When citing this work, cite the original published paper.

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/jtte

Research Article

Enhancing reinforcement learning-based adaptive traffic signal control in emerging mixed traffic environments through CAV data: A simulation study considering driver behaviour

Chuanyun Fu^{a,b,*}, Zhaoyou Lu^a, Naikan Ding^c, Zhixing Li^{d,e}

^a School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China

^b Department of Civil Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

^c Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

^d School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

^e Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg 41296, Sweden

HIGHLIGHTS

- An adaptive traffic signal control algorithm is developed for emerging mixed traffic flow.
- A generalized driven-behaviour model in simulations is proposed.
- Integrating the loop detector and CAV data can have a good performance of traffic signal control.
- The transformation of small-scale signalized intersections into the networked system with low cost is feasible.
- Collective driving behaviour patterns exert a significant influence on the signal control system.

ARTICLE INFO

Article history:

Received 1 September 2023

Received in revised form

2 July 2024

Accepted 9 July 2024

Available online 2 July 2026

Keywords:

Adaptive traffic signal control

Real-time Q-learning

Driven-behaviour model

Connected and automated vehicles

Emerging mixed traffic

ABSTRACT

In recent scholarly discourse, it has been noted that connected and automated vehicles (CAVs) have offered substantial developmental potential and a variety of implementation possibilities for the refinement and advancement of adaptive traffic signal control (ATSC). However, as the market penetration rate of CAVs is still relatively low, it is essential to consider the use of existing, cost-effective detectors that can form an integral component of the ATSC system. Additionally, it is crucial to consider drivers' behaviours in the context of emerging mixed traffic environments in order to reflect the realities of traffic flow in simulations due to the distinction between human-driven vehicles and CAVs. To address these issues, an ATSC algorithm was proposed to optimize signal timing to improve safety and operational performance at isolated intersections. The proposed algorithm leveraged real-time Q-learning with loop detector and CAV data to obtain optimized green time, while a driven-behaviour model was introduced to describe human factors in mixed traffic environments. Numerical studies were conducted using simulation of urban mobility (SUMO) to evaluate algorithm performance and investigate the influence of different

* Corresponding author. School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China.

E-mail addresses: fuchuan@hit.edu.cn (C. Fu), 23B932011@stu.hit.edu.cn (Z. Lu), nkding@whut.edu.cn (N. Ding), zhixingli@chalmers.se (Z. Li).

Peer review under responsibility of Chang'an University

<https://doi.org/10.1016/j.jtte.2024.07.002>

2095-7564/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

factors. The results indicate that the proposed algorithm has significant practical value in simultaneously improving safety with a demonstrated reduction in the conflict rate ranging from 28.6% to 72.7% and operational efficiency with a drop in the waiting time ranging from 12.6% to 61.4%, compared to traffic-actuated control. Moreover, it is low-cost and adaptable, and can be continuously updated with real-time driving data while also serving as a layer in next-generation high-definition maps.

1. Introduction

The adaptive traffic signal control (ATSC) system is an important part of smart urban traffic management, and its task is to ensure real-time traffic safety and efficiency at intersections (Geronimo et al., 2021; Ghanbarikarekani et al., 2018; Lin et al., 2025; Yang et al., 2016). With the rapid development of connected and automated vehicles (CAVs), real-time data will be readily available (Qu et al., 2025). ATSC safety and efficiency enhancement methods driven by CAV data have become a current research trend (Li et al., 2025; Reyad et al., 2021).

Previous studies on ATSC focus on reducing delay, queue length, waiting time, and crash risk (Devailly et al., 2021; El-Tantawy et al., 2013; Essa and Sayed, 2020; Genders and Razavi, 2019; Lee et al., 2020). The above objectives of the ATSC can be classified into two aspects: safety and efficiency. However, existing research still adopts the method of integrating safety and efficiency indicators with certain weights when it comes to optimizing efficiency and safety simultaneously (Shaikh et al., 2020). This method may increase the complexity of the ATSC. To simplify the algorithm and save computational resources, the design of ATSC should use indicators that are correlated with both safety and efficiency, such as traffic density (Kononov et al., 2012; Wang et al., 2020). Besides, to optimize the safety and efficiency of intersections under emerging mixed traffic flow, the design of ATSC typically relies on loop detectors and CAVs to collect a large amount of data. However, the market penetration rate (MPR) of CAVs remains low and will remain low for a long time, which cannot provide sufficient and high-quality CAV data for ATSC. Even some studies report that traffic safety may be worse under low MPRs than 0 MPRs (Virdi et al., 2019). Therefore, it is a challenge to optimize safety and efficiency using ATSC with low-quality data under low MPRs. Recently, some studies suggest that the combination of loop detector and CAV data can improve the performance of ATSC when the quantity and quality of the data are unsatisfactory (Feng et al., 2015; Qi et al., 2020). Therefore, researchers have explored several advanced ATSC enhancement techniques such as reinforcement learning, which offers a promising way to improve safety and efficiency using multisource data in emerging mixed traffic flow (Moradi et al., 2022; Mo et al., 2022).

On the other hand, the design of ATSC in an emerging mixed traffic environment is usually based on simulations due to the low MPRs. In simulations, there are challenges in reflecting the heterogeneity of traffic flows in the real world. Generally, existing research still lacks consideration for

individual driving behaviour differences in emerging mixed traffic flow. Although existing studies have focused on the effect of human factors on safety under emerging mixed traffic flow, human factors are neglected in the field of traffic control (Shladover et al., 2012). As highlighted by Wang et al. (2022), it is essential to consider different driving behaviours in simulations to obtain a more accurate traffic flow pattern. Virdi et al. (2019) studied the impact of human drivers and CAVs in the mixed traffic environment and employed a car-following model (i.e., the intelligent driver model (IDM)) for human drivers. However, using the IDM is inappropriate when considering human factors in simulations. Zhu et al. (2015) highlighted the different properties of emerging mixed traffic flow in terms of capacity and traffic congestion with different types of drivers (e.g., aggressive and cautious drivers). A range of factors can affect driving behaviour including personality (Yang et al., 2013), gender (Yadav and Velaga, 2021), age (Chen et al., 2025), reaction (Fu et al., 2016; Li et al., 2023a), and familiarity (Wu and Xu, 2018). Therefore, these human factors should be considered by a unified driver behaviour model in traffic control simulations.

Given this, this study fills the research gaps by proposing a Q-learning-based modified (QM) ATSC algorithm to optimize safety and efficiency at isolated intersections. As a major contribution, the performance of QM-ATSC is enhanced by CAV data and tested in a simulation system considering driver behaviour. Specifically, the main contributions of the study can be summarised in three aspects: First, this study develops a QM-ATSC algorithm to improve safety and efficiency under emerging mixed traffic flow at signalized intersections, which considers the cost of the red time. Second, this study focuses on the integration of loop detector and CAV data to drive the QM-ATSC algorithm, which can still maintain good performance with insufficient and low-quality data. Third, this study develops a generalized driven-behaviour model in simulations to reflect the heterogeneous traffic flow of the real world.

The paper is structured as follows. Section 2 organizes the relevant literature and conducts a comprehensive review. Section 3 presents the design of the modified ATSC algorithm proposed in the study, which considers both CAV data and human factors. Next, in Section 4, we present a case study of the proposed approach, including parameter calibration, driving behaviour models, baseline, and market penetration rate. Subsequently, Section 5 provides simulation results to evaluate the performance of the proposed algorithm. In Section 6, we reflect on the results and offer insights regarding the QM-ATSC model. Finally, Section 7 summarizes the implications of the study, including its limitations.

2. Literature review

2.1. Adaptive traffic signal control in emerging mixed traffic flow

The development of CAVs will provide ATSC with more possibilities (Matin and Dia, 2022; Wei et al., 2018). As a moving detector, the CAV can provide the ATSC with a wide range of accurate real-time data and improve its performance (Wang et al., 2018). However, the data-driven capability of ATSC is limited by MPRs (Xiao et al., 2021). Before CAVs are fully deployed, the emerging mixed traffic flow will consist of HDVs and CAVs (Lu et al., 2023). Therefore, the use of CAV data to enhance ATSC has become a hot research topic, particularly under low MP Rs. Generally, CAV data can improve ATSC performance in emerging mixed traffic flow. Miao and Leitner (2021) used CAV data to construct a model and proposed an ATSC considering worst-case waiting times. Ghosh and Parisini (2022) described signal control for emerging mixed traffic flow as a non-linear programming problem, using CAV data and traffic signals to maximize the traffic volume in the emerging mixed traffic flow. Their studies have uniformly substantiated the importance of CAV data for the enhancement of ATSC.

In China, there still exist numerous small-scale intersections such as two-phase signalized intersections (Chu et al., 2019; Li et al., 2014; Lu et al., 2022). Therefore, the renovation costs of a large small-scale intersections should be considered to integrate them into a connected system. Typically, studies in ATSC under traditional traffic flow often utilize low-cost loop detectors to gather traffic information (Li et al., 2023b). Naturally, continuing the use of these low-cost loop detectors and avoiding the installation of new roadside equipment represent potential strategies for the connected transformation of a large number of small-scale intersections. On the other hand, research has demonstrated that the fusion of CAV data and loop detector data can enhance ATSC performance. Feng et al. (2018) utilized CAV data to enhance the performance of history-based ATSC, and the results indicated that ATSC performance can be improved even under low MP Rs. Similarly, Liang et al. (2020) designed an ATSC based on a heuristic algorithm that combines detector data and CAV data, and the results showed that this method can effectively reduce delays in emerging mixed traffic flow.

2.2. Reinforcement learning for traffic signal control

In recent years, the research hotspots in ATSC have been shifting toward reinforcement learning (Miletić et al., 2022). Although reinforcement learning-based ATSC only requires detector data to operate in existing studies of traditional traffic flow, the enhancing effect of CAV data on ATSC has been reported to further improve its performance (Chu et al., 2019).

The ATSC based on the reinforcement learning algorithm (e.g., Q-learning) includes state, reward, and action. The state of ATSC based on Q-learning typically includes position, speed (Liang et al., 2019), queue length (Tan et al., 2019), waiting time (Liu et al., 2017), and delay (Mannion et al., 2016). Moreover,

contemporary studies posit that the utilization of an order of specific indicators from each approach of an intersection as the state manifests the capacity to streamline the state space and expedite the convergence of the algorithm. Abdoos (2020) incorporated the order of queue lengths for each approach at an intersection as the state in designing the ATSC based on Q-learning. Drawing inspiration from the findings of Abdoos (2020), Wei et al. (2022) adopted the order of queue lengths for each approach as the state. This method of ordering the indicators of various approaches effectively encompasses all states of the intersection.

In terms of rewards, most existing studies focus on optimizing specific indicators as objectives. El-Tantawy et al. (2013) used the change in cumulative delay of upstream vehicles as a reward. Arel et al. (2010) used the change in delay as a reward, with a positive value of the reward when delay is reduced and a negative value of the reward when delay increases. Similarly, Ge et al. (2019) used the change in queue lengths for each approach at an intersection as a reward. On the other hand, more and more studies on Q-learning-based ATSC focus on optimizing traffic safety. Essa and Sayed (2020) used the inverse of the conflict rate as a reward to reduce traffic conflicts at intersections. Reyad et al. (2021) introduced extreme value theory in the development of Q-learning-based ATSC. They used the risk of a crash (ROC) and the return level of a cycle (RLC) as rewards. However, safety and efficiency appear to be two conflicting optimization objectives. Even in multi-objective optimization, the safety and efficiency indicators are treated as separate objectives. Besides, it should be noted that ATSC driven by big data requires large amounts of computational power, and simplifying the complexity of the algorithm can effectively reduce the computational cost. Therefore, it is crucial to select an indicator, such as traffic density, as a reward for optimizing safety and efficiency, simultaneously. Multiple studies have reported a close relationship between traffic congestion and traffic accidents (Berhanu et al., 2023; Chang et al., 2022; Wang, 2010). Some studies also suggested that although traffic density is an efficiency indicator, there is also a potential interaction between it and safety. For example, Lord et al. (2005) reported that traffic density has a potential relationship for total and single vehicle accidents. Wang et al. (2013) reviewed some studies to analyze the impact of traffic variables on safety, in which traffic density is identified as one of the key factors. In addition, Formosa et al. (2020) believe that traffic density can serve as an important feature to predict real-time traffic conflicts.

In terms of actions, Q-learning-based ATSC often employs extending green time or switching phases as actions. The action space can be categorized into four types (Matin and Dia, 2022): 1) when the current signal is green, extend the green time; 2) when the current signal is yellow, revert to the green phase; 3) when the current signal is red, add an extra green phase; and 4) when the current signal is red, enter the green phase in advance.

2.3. Driver behaviour

The characteristics of driving behaviour are intricately linked to traffic safety, whereby heterogeneous behaviour acts as the

principal factor that precipitates instability within traffic flow (Evans, 1996; Taylor et al., 2025). Therefore, it is imperative to take into account driving behaviour when calibrating traffic flow parameters in specific simulation settings. In an attempt to enhance the car-following model, Soria et al. (2014) took into account different driver types. They categorized drivers into three types according to driving behaviour: aggressive, average, and conservative. Further, numerous studies have indicated that differentiating driver types in the analysis of traffic flow characteristics has significant implications for accurately replicating real-world scenarios in simulations (Chien et al., 2001; Levermore et al., 2014; Luo et al., 2020; Zheng et al., 2017). With the rise of CAVs, the trust level of human drivers in them has become a new area of research focus. Dikmen and Burns (2017) suggested that the level of trust that human drivers have in autonomous vehicles is related to multiple factors, such as age, accident history, attitudes, and behavioural factors. According to Manchon et al. (2022), the trust level in autonomous vehicles can affect the driving behaviour of human drivers in emerging mixed traffic flow. Further, multiple studies applied the concept of trust in their simulation studies (Buckley et al., 2018; Ha et al., 2020; Molnar et al., 2018). Therefore, it is meaningful to distinguish between driver types in both traditional and emerging mixed traffic flow. However, existing research still lacks a parameterized generalized driver behaviour model for simulation studies, especially in the traffic control field.

3. Reinforcement learning-based approach

This study proposed a Q-learning-based modified ATSC (QM-ATSC) approach to optimize safety performance at isolated intersections by optimizing relative traffic flow. Fig. 1 presents a flow chart of the QM-ATSC approach. To be specific, the proposed algorithm optimized the green time of the next phase using the Q-learning algorithm, which was then refined using real-time CAV data during each second. Integrating the CAV and loop detector data into the ATSC algorithm could enhance the accuracy and precision of green-time optimization. In addition, this study proposed a driving behaviour model to improve the reliability of simulations and capture the heterogeneity of traffic flow caused by individual driver factors in the real world. In addition, the overall costs for drivers and vehicles while waiting during red time had also been taken into account in the proposed ATSC system.

3.1. Q-learning algorithm

In QM-ATSC, the Q-learning algorithm was utilized to optimize the green time according to traffic volume data from the loop detectors. This algorithm had five significant factors: agent, environment, state, action, and reward. The agent selected an action that generated a Q-value based on the reward in the current state. The Q-table was updated in the process, which captured the relationship between state,

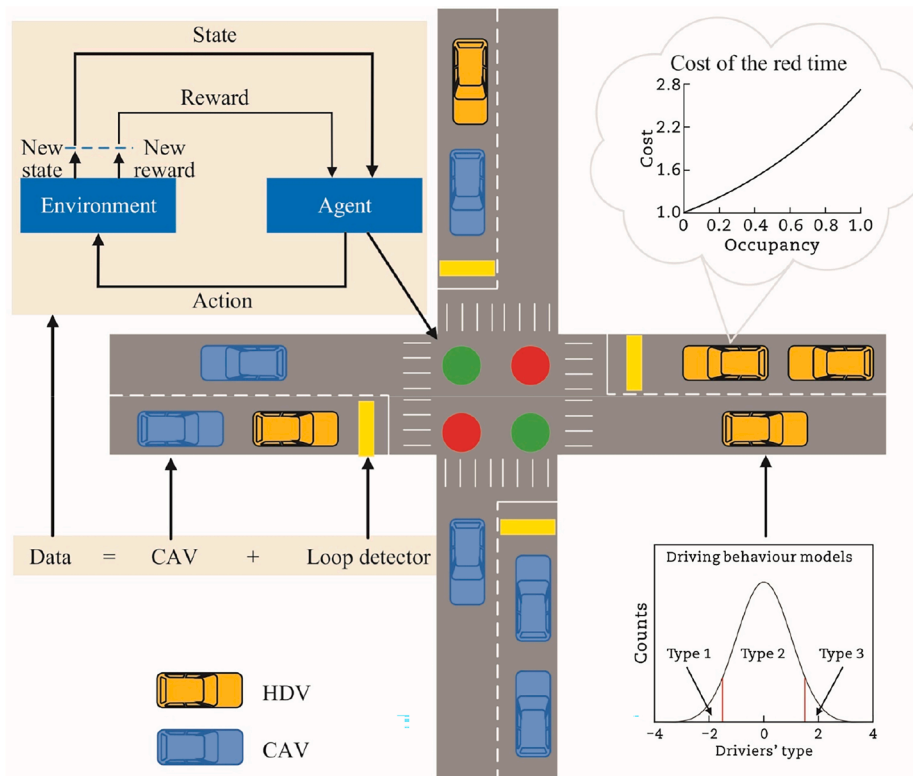


Fig. 1 – Flowchart of the proposed QM-ATSC approach.

action, and Q-value (Eq. (1)). This process allowed the Q-value to continually be updated, ultimately leading to more accurate optimizations of green time.

$$Q_{\text{New}}(s_t, a_t) = Q(s_t, a_t) + \alpha \left[R(s_t, a_t) + \gamma \max_{a_t \in \mathcal{A}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right] \quad (1)$$

where $Q_{\text{New}}(s_t, a_t)$ is the updated Q-value, $Q(s_t, a_t)$ is the Q-value corresponding to action a_t at state s_t at time t , α is the learning rate, $R(s_t, a_t)$ is the reward value under (s_t, a_t) , γ is the discount rate, $\mathcal{A}$ is the set of actions, $Q(s_{t+1}, a_{t+1})$ is the Q-value at time $t + 1$.

3.1.1. State

In QM-ATSC, state was defined as the order of relative traffic flow in each approach. The relative traffic flow was determined by the upcoming traffic flow and the green or red time of each approach. The value of relative traffic flow D_{li} is defined as follows.

$$D_{li} = \begin{cases} m_i t_g & \text{Phase is green} \\ m_i t_r \lambda(t_r) & \text{Phase is red} \end{cases} \quad (2)$$

where t_g is the green time (s), m_i is the upcoming traffic flow at the approach i (veh/s), which is detected by the upstream detectors, t_r is the red time (s), $\lambda(t_r)$ is penalty function, which illustrates the overall cost of the red time, such as time, and drivers' patience.

When driving in congested conditions, the loss of driver patience will ascend roughly exponentially with the degree of congestion (Mandelbaum and Zeltyn, 2007; Liu et al., 2017; Yu et al., 2012). Therefore, it could be considered that the total cost of the red time, which integrates both the driver patience and the waiting time, approximately follows an exponential distribution. The degree of congestion can be expressed in terms of lane occupancy, and the penalty function can be written as Eq. (3).

$$\lambda(t_r) = \exp\left(\frac{\sum l_j}{L_i n_i}\right) \quad (3)$$

where l_j is the length of vehicle j (m), L_i is the length of approach i (m), n_i is the number of the lane.

For simplicity and convenience, the following approximation is adopted as Eq. (4).

$$\lambda(t_r) = \exp\left(\frac{m_i \bar{l} t_r}{L_i n_i}\right) \quad (4)$$

where $\bar{l}$ is the average vehicle length (m).

Once the relative flow D_{li} for each approach had been calculated using Eq. (2), the states could be specified in descending order. An example of the specified states is shown in Table 1. Referencing the work of Abdoos (2020) and Wei et al. (2022), the order of D_{li} in each approach was considered as the state.

For the Q-learning algorithm, it is crucial to identify terminal states in which the algorithm should stop exploring (Algorithm 1 and Appendix) and begin executing the chosen action. In this case, some states were marked as terminal under the following conditions. 1) When the D_{li} at two approaches during the red time were simultaneously higher than those at the other approaches during the green time. QM-

Table 1 – States of the proposed Q-learning at a 4-legs intersection.

Order	State	Order	State
s_1	$D_{11} > D_{12} > D_{13} > D_{14}$	s_{13}	$D_{12} > D_{13} > D_{11} > D_{14}$
s_2	$D_{11} > D_{12} > D_{14} > D_{13}$	s_{14}	$D_{12} > D_{14} > D_{11} > D_{13}$
s_3	$D_{11} > D_{13} > D_{12} > D_{14}$	s_{15}	$D_{13} > D_{12} > D_{11} > D_{14}$
s_4	$D_{11} > D_{14} > D_{12} > D_{13}$	s_{16}	$D_{14} > D_{12} > D_{11} > D_{13}$
s_5	$D_{11} > D_{13} > D_{14} > D_{12}$	s_{17}	$D_{13} > D_{14} > D_{11} > D_{12}$
s_6	$D_{11} > D_{14} > D_{13} > D_{12}$	s_{18}	$D_{14} > D_{13} > D_{11} > D_{12}$
s_7	$D_{12} > D_{11} > D_{13} > D_{14}$	s_{19}	$D_{12} > D_{13} > D_{14} > D_{11}$
s_8	$D_{12} > D_{11} > D_{14} > D_{13}$	s_{20}	$D_{12} > D_{14} > D_{13} > D_{11}$
s_9	$D_{13} > D_{11} > D_{12} > D_{14}$	s_{21}	$D_{13} > D_{12} > D_{14} > D_{11}$
s_{10}	$D_{14} > D_{11} > D_{12} > D_{13}$	s_{22}	$D_{14} > D_{12} > D_{13} > D_{11}$
s_{11}	$D_{13} > D_{11} > D_{14} > D_{12}$	s_{23}	$D_{13} > D_{14} > D_{12} > D_{11}$
s_{12}	$D_{14} > D_{11} > D_{13} > D_{12}$	s_{24}	$D_{14} > D_{13} > D_{12} > D_{11}$

ATSC should ensure the maximum relative traffic flow. The relative traffic flow of the approach during the red time went up with the increasing red time. When the values of D_{li} of all approaches were equal, the relative traffic flow during green time reached the maximum value. Therefore, if there existed an approach during the red time whose relative traffic flow was less than any approach during the green time, the state had not yet reached its terminal state. 2) When the queue at any approach exceeded the length of the lane.

If either of these conditions were met, the current state was marked as a terminal, indicating that the while loop (Algorithm 1 and Appendix) should break and perform the appropriate action. In addition, to match the relationship between the orders of relative traffic flow and green time, this study constructed a two-dimensional state matrix composed of them. The matrix size was the number of orders of relative traffic flow (i.e., 24) multiplied by the difference between the maximum green time and the minimum green time.

3.1.2. Reward

In QM-ATSC, the reward for each action was defined as the difference in the relative traffic flow D_{li} between the green time and red time at the approach i . Specifically, assuming that the relative flow during the green phase was $D_{li}(g)$, and the relative flow during the red phase was $D_{li}(r)$, the reward associated with an action is given by the following.

$$R(s_t, a_t) = \sum_{li \in \mathcal{B}} D_{li}(g) - \sum_{li \in \mathcal{C}} D_{li}(r) \quad (5)$$

where $\mathcal{B}$ is the set of lanes, which is during green phase, $\mathcal{C}$ is the set of lanes, which is during red phase.

The proposed algorithm was to maximize the reward over multiple iterations, which implied reducing the relative flow during the red time and increasing it during the green time. This was achieved by adjusting the green time using the Q-learning algorithm and selecting the optimal policy to control the traffic signal.

3.1.3. Action

In QM-ATSC, the agent had two possible actions to select: increasing or decreasing the green time in steps of 1 s. These actions were denoted by $\mathcal{A} \in \{1, -1\}$. During the initial loop, the action was selected randomly. Afterward, the Q-table was updated, and the agent selected the action based on the

current state and the expected rewards associated with each possible action. The agent continued this process until the relative traffic flow D_{li} during the red time was greater than during the green time for all approaches.

3.1.4. Other parameters

In QM-ATSC, several other parameters were used in the Q-learning algorithm in addition to actions and rewards, as follow.

- (1) Episode. In reinforcement learning, an episode refers to a complete iteration of the Q-learning algorithm, where the agent interacts with the environment and updates its Q-table. Fig. 2 illustrates the learning progress of the agent represented by the green time. It should be noted that retraining was required after each phase was completed, which could generate many learning processes in a simulation. Therefore, we chose five learning processes to demonstrate the convergence of the proposed algorithm. The results show that there are differences in convergence speed under different relative traffic flow scenarios. It can be observed that the algorithm can converge in approximately 80 episodes. Thus, in QM-ATSC, 100 episodes were conducted to train the model to find the optimal green time.
- (2) Discount rate and epsilon. The discount rate is used in the Q-learning algorithm to control the weight or importance of future rewards relative to immediate rewards. In QM-ATSC, the discount rate was set to 0.9. The epsilon refers to the practice of randomly selecting non-optimal actions with a certain probability during action selection to ensure exploratory behaviour and find better strategies during the exploration. The initial configuration of the epsilon was 1, which decayed with a step size of 0.01 while exploration continued until it reached the minimum epsilon of 0.01.
- (3) Maximum green time. The maximum green time is the upper bound of the length of the green time that the

agent can select. The maximum green time in QM-ATSC was 70 s (Essa and Sayed, 2020).

- (4) Minimum green time and initial green time. The minimum green time was calculated as follows (Lu et al., 2022).

$$G_{\min} = 7 + \frac{w_r}{v_p} - t_l \quad (6)$$

where $G_{\min}$ is the minimum green time (s), w_r is the road width (m), v_p is the pedestrian walking speed (m/s), t_l is the lost time (s).

Further, the initial green time was 35 s in this study. The learning process of Q-learning can be seen in Algorithm 1 and Appendix.

Algorithm 1: the Q-learning

```

Input: the upcoming traffic flow  $m_i$ .
Output: the green time of following phase  $t_g$ .
Create a 2D state matrix  $S[\text{state}, \text{greenTime}]$ ;
1  for  $i \leftarrow 1$  to  $\text{maxEpisode}$  do
2    initialize state  $s_t$ ;
3    while not at terminal state do
4      if  $\text{random} < \text{epsilon}$  then
5         $a_t \leftarrow \text{randomChoice}()$ ;
6      else
7         $a_t \leftarrow \text{argmax}(Q[\text{state}])$ ;
8      end
9       $t_g \leftarrow t_g + a_t$ ;
10      $R(s_t, a_t) \leftarrow \sum_{li \in \mathcal{D}} D_{li}(g) - \sum_{li \in \mathcal{R}} D_{li}(r)$ ;
11      $s_{t+1} \leftarrow S[\text{order}(D_{li}), t_g]$ ;
12      $Q_{\text{New}}(s_t, a_t) \leftarrow \text{updateQ}(R(s_t, a_t), s_{t+1})$ ;
13      $s_t \leftarrow s_{t+1}$ ;
14     if  $\text{iteration} > 100$  then
15       break;
16     end

```

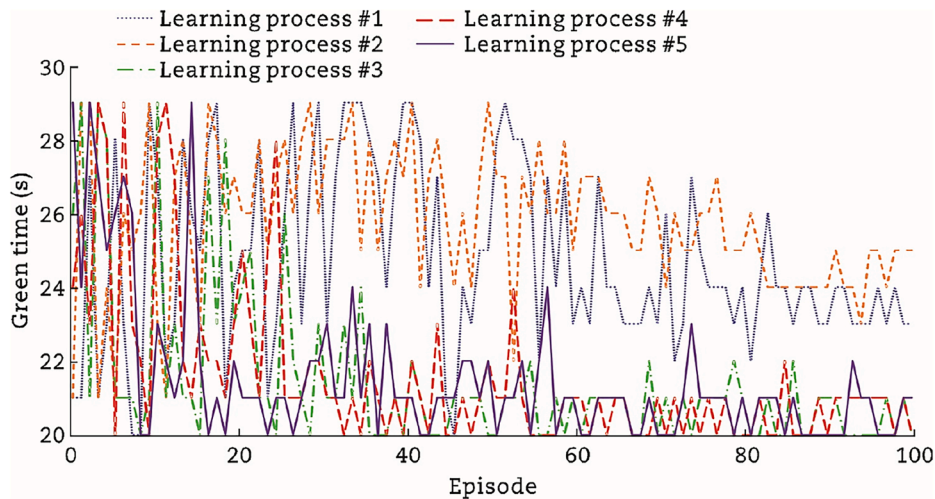


Fig. 2 – Learning process of the proposed QM-ATSC algorithm.

3.2. Modified approach

The green time optimization algorithm above did not use the CAV data, which include the accurate speed and position information. This section proposed a modified approach to further adjust the green time considering the introduction of CAVs to the traffic flow.

3.2.1. Congestion space

In QM-ATSC, congestion space was defined as the maximum detected congestion length used to adjust the green time and improve the relative traffic flow. Congestion space was detected using CAV data, which could detect the speed and position of nearby vehicles and share that information with the traffic signal control system. The traffic signal control system was treated as the agent of the Q-learning algorithm and CAVs were used to detect congestion and determine the optimal green time for each cycle. When a low-speed vehicle was detected around a CAV, it was assumed to be affected by congestion, which could cause low-speed waves. The end of the congestion space was estimated as the detected position of the low-speed vehicle farthest from the stop line.

However, the accuracy of the detected congestion space depended on the MPRs of the CAV. If the MPRs was low, the accuracy of congestion space detection may be compromised, leading to suboptimal traffic signal control. Therefore, improving the MPRs of the CAV in the area could increase the accuracy of congestion space detection and improve the overall efficiency of the traffic signal control system. This was an important consideration when implementing QM-ATSC in real-world traffic systems, which will be discussed in Section 5.

3.2.2. Time adjustment

As time passed, the red time caused some negative impacts on vehicles and drivers, such as unnecessary loss of patience. Studies have found that these costs increase exponentially with time and congestion. The detected congestion space dynamically adjusted the green time per second to optimize traffic density and improve real-time safety in this study. The traffic density d_i during the red time can be calculated according to the detected coordinate position p_j of the low-speed vehicle j , which is given by Eq. (7).

$$d_i = \frac{\max\{p_j\}}{L_i} \quad (7)$$

Based on Eq. (7), the queue length of approach i during the red time will be overflowed when $d_i > 1$. Therefore, d_i could be regarded as the penalty for the green phase. Given $\lambda(t_r)$, the adjustment action $\mathcal{A}_m$ is determined by the following equation.

$$\mathcal{A}_m = -\Delta t, d_i \lambda(t_r) > \epsilon \quad (8)$$

where Δt is the step of the adjustment action, which is specified as 1 s, ϵ is the sensitivity coefficient, which is set as 1 in this study.

In Eq. (8), the overall cost increased exponentially with the degree of congestion. However, the increase in this cost needs to consider the influence of traffic density. When the traffic density was low, even if the overall cost was high, it was not

advisable to change the signal timing rashly. Eq. (8) shows that the action is enforced only when the conditions are satisfied. The time adjustment algorithm did not work when the MPRs was 0. Further, when MPRs was low, the algorithm could produce some errors, and the impact of these errors on QM-ATSC will be analysed in the discussion section.

4. Simulations

4.1. Parameters calibration

An intersection in Harbin, China, was selected as a reference in simulation of urban mobility (SUMO) software. A traffic survey was conducted during peak hours to collect traffic flow data. The data were collected through video recordings and used to derive the basic characteristics of the intersection, which are presented in Tables 2 and 3. In this study, the traffic volume was converted to equivalent traffic volume according to the method recommended in the HCM (AASHTO, 2010).

This study used the intelligent driver model (IDM) for human-driven vehicles (HDVs). This study required the detection of the speed and position of the neighbouring vehicles when detecting low-speed vehicles by CAVs. It was impossible to perceive the speed and position of neighbouring vehicles without sensors when they were HDVs. In contrast, the speed and position of neighbouring vehicles were obtained through vehicle-to-vehicle communication. Therefore, cooperative adaptive cruise control (CACC) was selected for CAVs (Yao et al., 2023; Wang et al., 2024). The study used a sensitivity analysis to obtain the parameter values (Fig. 3). Five groups of car-following parameters (“minGap”, “Accel”, “Decel”, “Tau” and “maxSpeed”) were calibrated. To be specific, the value ranges were selected according to the default values in SUMO. Then, the values of the parameters in the ranges were enumerated in the sensitivity analysis. To speed up the calibration process, we selected six

Table 2 – Profile of the selected intersection.

Approach	East	West	South	North
Speed limited (km/h)	50	50	50	50
Width of lane (m)	3.75	3.75	3.75	3.75
Length of road (m)	70	80	140	150
Traffic volume (pcu/h)	426	1664	670	486

Table 3 – Descriptive statistics of traffic flow characteristics.

Item		Value
Time headway (s)	Mean	2.50
	Standard deviation	0.86
	Minimum	1.08
	Maximum	6.76
Speed (km/h)	Mean	23.17
	Standard deviation	6.08
	Minimum	10.21
	Maximum	34.38
Traffic volume (pcu/h)		3246

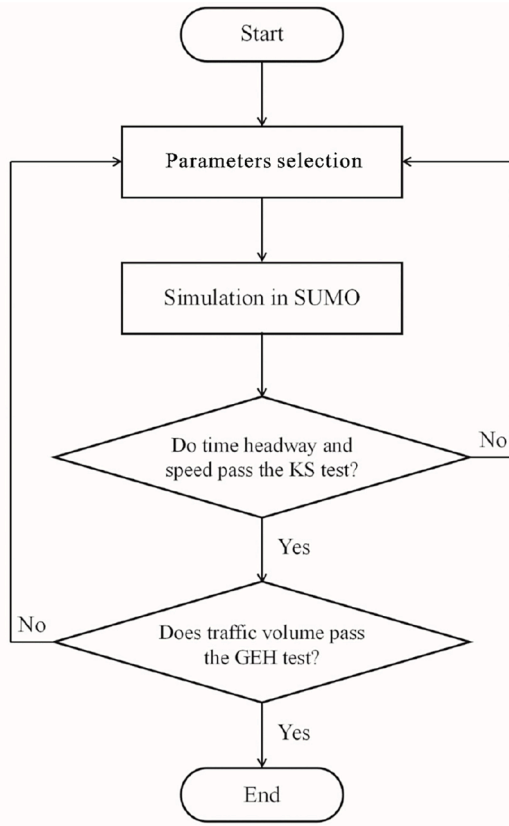


Fig. 3 – Flowchart of the sensitivity analysis of parameters calibration.

uniformly distributed values within the range of each parameter. Each sensitivity analysis required executing one simulation, so a total of 77,760 ($6^5 \times 10$ random seeds) simulations were conducted. After each simulation, the time headway and speed data were extracted and compared with the real-world data using a Kolmogorov-Smirnov test (Rahman et al., 2019). This process continued until there was an insignificant difference between them.

After calibration of the parameters, it should verify the results and determine whether the simulations accurately represent the real-world traffic flow. This study used the GEH statistics for this purpose. According to the UK Highways Agency (1996), if more than 85% of the GEH values are less than 5, the simulated traffic flow can be considered to accurately reflect the real-world scenario. To calculate the GEH, we captured the 15-min intervals from both the simulations and the real-time traffic data and applied as follows.

$$GEH = \sqrt{\frac{(M_{obs}(n) - M_{sim}(n))^2}{(M_{obs}(n) + M_{sim}(n))/2}} \quad (9)$$

where $M_{obs}(n)$ is the observed volume of in-field detectors, $M_{sim}(n)$ is the simulated volumes obtained from the simulation network.

The results indicate that the calibrated parameters in SUMO can indeed depict real-world scenarios. Over 90.0% of the GEH values were less than 5, and the average GEH value was 2.63.

The calibrated and default parameters for the car-following models are summarized in Table 4. Then, to eliminate the influence of extraneous variables, the CACC parameters in this study were kept consistent with those of IDM. Additionally, the lane-changing model for HDVs and CAVs was the LC2013 in SUMO.

4.2. Driving behaviour models

A parametric model was proposed to describe different driving behaviours. The model consists of type and behaviour parameters, shown in Table 5 and Fig. 4. The type parameters included three driving modes: type 1 (aggressive), type 2 (normal), and type 3 (cooperative). Aggressive drivers tended to drive faster, with shorter time headways and more frequent lane-changings. Normal drivers had a moderate driving style, with reasonable speed and time headway. The cooperative drivers had a more cautious driving style, with longer time headways and less lane-changings.

The behaviour parameters further described the driving style within each driving mode. These parameters include “minGap”, “Accel”, “Decel”, “Tau”, and “maxSpeed”. The proposed parametric model provides a flexible and adaptable framework for describing different driving behaviours, making it suitable for a wide range of traffic scenarios and environments.

4.3. Baseline

The real-world intersection was signalized with fixed timing (FT). However, the more advanced traffic actuated control (TAC) is increasingly used as a baseline to evaluate the performance of ATSC (Essa and Sayed, 2020; Trabia et al., 1999; Zaied and Othman, 2011). Therefore, the above FT and TAC were used as baselines in simulations to test the performance of the proposed QM-ATSC approach in this study. In addition, the real-time safety-optimized adaptive traffic signal control (RS-ATSC) algorithm proposed by Essa and Sayed (2020) was also carried out to compare with QM-ATSC.

4.4. Market penetration rate

The emerging mixed traffic flow is expected to exist in road networks until the CAV technology is fully deployed. Therefore, a set of CAVs MPRs was set to test the performance of QM-ATSC in the emerging mixed traffic flow. These MPRs range from 0 to 100% in 10% steps, representing varying proportions of CAVs in traffic flow. At each MPRs, the simulation duration was 3600 s.

5. Results

5.1. Safety performance of the proposed approach

In QM-ATSC, the safety performance of the traffic signal control system was quantified using time-to-collision (TTC), which was a widely used metric in traffic safety analysis (Hayward, 1971). The TTC at time t is calculated as follows.

Table 4 – Car-following behaviours calibration in SUMO.

Parameter	MinGap (m)	Accel (m/s ²)	Decel (m/s ²)	Tau (s)	MaxSpeed (m/s)
Calibrated	2.1	2.0	4.0	1.8	10.00
Default	2.5	2.6	4.5	1.0	55.55

Table 5 – Driven-behaviour model A.

Parameter	MinGap (m)	Accel (m/s ²)	Decel (m/s ²)	Tau (s)	MaxSpeed (m/s)
Type 1	$\mathcal{N}(1.3, 1.0)$	$\mathcal{N}(1.5, 1.0)$	$\mathcal{N}(3.5, 1.0)$	$\mathcal{N}(1.5, 1.0)$	$\mathcal{N}(9.0, 1.0)$
Type 2	$\mathcal{N}(1.8, 1.0)$	$\mathcal{N}(2.0, 1.0)$	$\mathcal{N}(4.0, 1.0)$	$\mathcal{N}(2.1, 1.0)$	$\mathcal{N}(10.0, 1.0)$
Type 3	$\mathcal{N}(2.3, 1.0)$	$\mathcal{N}(2.5, 1.0)$	$\mathcal{N}(4.5, 1.0)$	$\mathcal{N}(2.6, 1.0)$	$\mathcal{N}(11.0, 1.0)$

Note: $\mathcal{N}(\cdot)$ is the Gaussian distribution.

$$TTC(t) = \frac{S(t)}{V_f(t) - V_l(t)} \quad (10)$$

where $S(t)$ is the space between the lead and follow vehicle at time t (m), $V_f(t)$ is the speed of the follow vehicle at time t (m/s), $V_l(t)$ is the speed of the lead vehicle at time t (m/s).

In this study, the safety of QM-ATSC was evaluated using the conflict rate as a metric, which measured the probability that the TTC would be less than 1.5 s. This value is widely accepted as the threshold for identifying a conflict (Virdi et al., 2019). To calculate the conflict rate, the TTC less than 1.5 s was counted for each approach and then divided by the TTC for all approaches within each second.

Fig. 5 shows the average hourly conflict rates under each CAVs MPRs for the four control approaches: QM-ATSC, FT, TAC, and RS-ATSC. The results demonstrate a consistent pattern across all control approaches, since conflict rates generally decrease with increasing MPRs. These findings further highlight the superiority of QM-ATSC over traditional control approaches such as FT and TAC, as QM-ATSC leverages CAV data to optimize safety performance dynamically. Similarly, RS-ATSC, which incorporates CAV data in its control strategy, also demonstrates promising results in reducing conflict rates.

Based on the results, QM-ATSC shows significantly better safety performance compared to FT and TAC under all MPRs. The advantages of QM-ATSC are particularly evident in high MPRs (>50%), highlighting its effectiveness in improving traffic safety when there is a greater proportion of CAVs in traffic flow. However, in low MPRs, the safety of QM-ATSC is inferior to that of RS-ATSC.

Generally, the results indicate that QM-ATSC is an effective traffic signal control approach for emerging mixed traffic flow and can improve traffic safety. Under high MPRs, the conflict rate is reduced by 19%–39% compared to RS-ATSC, indicating a significant improvement in traffic safety. On the other hand, the results investigate the effectiveness of a conflict reduction in a traffic context and achieve a notable reduction in conflict rate, ranging from 28.6% to 72.7%, compared to TAC. This reduction represents a more substantial improvement compared to the findings of Wei et al. (2022), who reported a conflict rate reduction of 36% and is also superior to the results reported by Ghoul and Sayed (2021), who achieved a conflict rate reduction of 41%–55%.

5.2. Operational performance of the proposed approach

In addition to analysing traffic safety, operational performance is also an important consideration when optimizing traffic signal control. In this study, waiting time was used as a metric to analyse the operational performance of the intersection.

Fig. 6 shows the average waiting time of each CAVs MPRs for the four approaches, including QM-ATSC, FT, TAC, and RS-ATSC. QM-ATSC shows a decrease in waiting time under each MPRs, with reductions of 0.9%–42.1% compared to RS-ATSC and 12.6%–61.4% compared to TAC. This phenomenon represents a more substantial improvement compared to the findings of Song and Fan (2021), who reported a reduction in waiting time of 38%. Similarly, it is also superior to the results reported by Celtek et al. (2021) and Zeng et al. (2019), who achieved a waiting time reduction of 37% and 50%, respectively.

It is noted that the waiting time goes up when the MPRs is greater than 70%. It is widely believed that CAVs have higher safety performance compared to HDVs. While improving safety, CAVs are inevitably more conservative in their driving behaviour. Therefore, there may be a decrease in efficiency under high MPRs. The improvements in QM-ATSC in efficiency and safety support this conclusion. In QM-ATSC, the waiting time remains relatively stable after the MPRs exceeds 40%. Additionally, the degree of improvement becomes insignificant beyond 70% MPRs. Specifically, the average waiting time for QM-ATSC converges to 2 s when the MPRs is over 70%.

6. Discussions

6.1. Effect of CAV market penetration rate

There is an insignificant difference in operational performance between FT and TAC as the CAVs MPRs rises. On the contrary, the operational performance of QM-ATSC and RS-ATSC improves with increasing CAVs. However, the performance of QM-ATSC is even worse than that of RS-ATSC under low MPRs, although it outperforms those of FT and TAC.

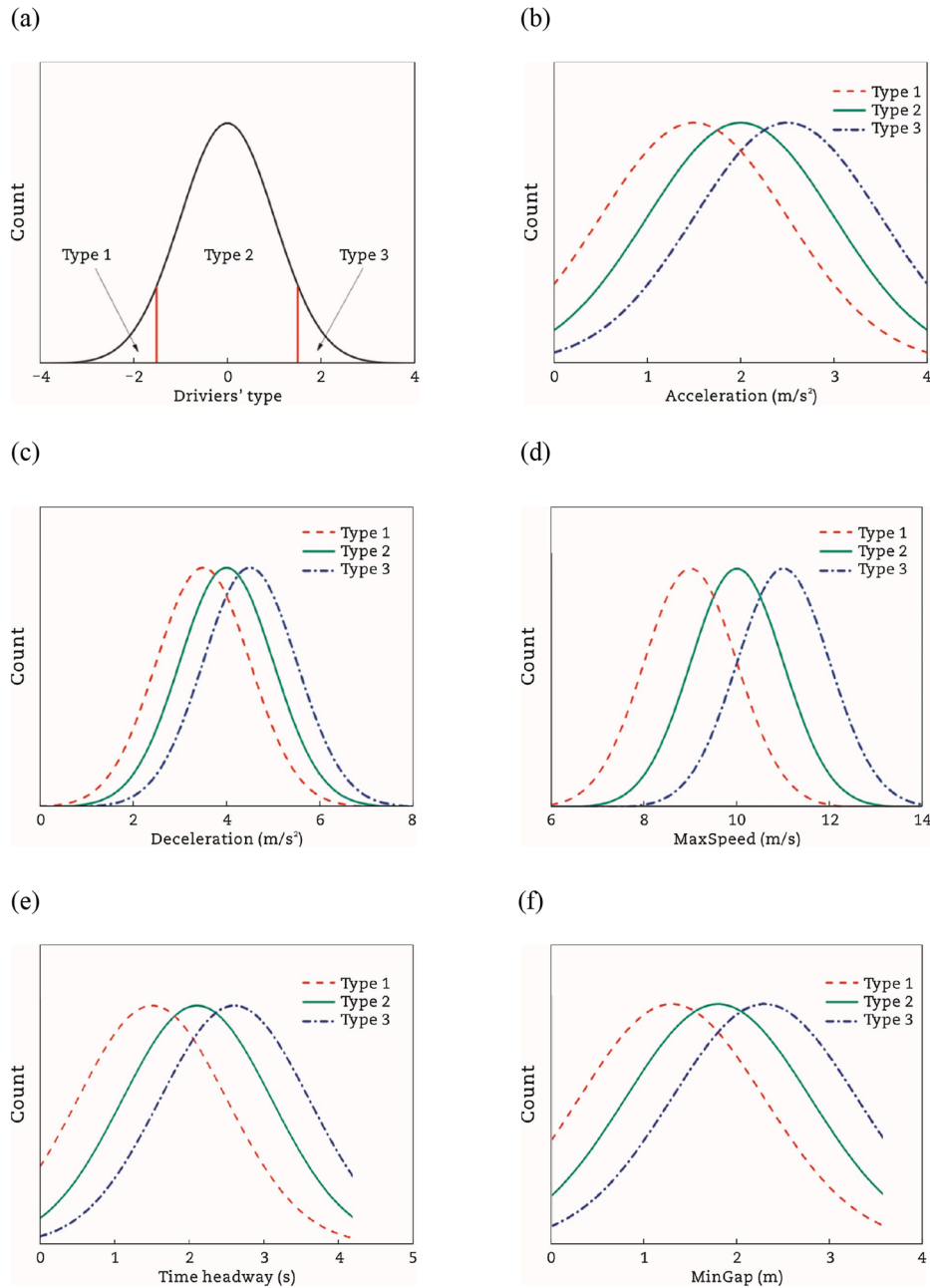


Fig. 4 – Proposed parametric driving behaviour model. (a) Drivers' type. (b) Acceleration. (c) Deceleration. (d) MaxSpeed. (e) Time headway. (f) MinGap.

Fig. 7 shows the errors of each approach under different MPRs. The results show that the errors of each approach are significant under low MPRs. It should be noted that the error is relatively large when the MPRs is low (<60%). However, the algorithm can still work properly because d_i can only be underestimated. When d_i is underestimated, two possible scenarios occur: the condition in Eq. (8) is satisfied or not satisfied. The former can allow QM-ATSC to work properly, and the latter can cause a delay in the execution of the time adjustment process. This finding can explain why QM-ATSC

has a poorer optimization performance than RS-ATSC under low MPRs.

Fig. 8 shows the variation of the throughput as the MPRs increases. Below an MPRs threshold of 50%, the throughput exhibits a fluctuating trend, ranging from 3489 to 4471 pcu/h. Subsequently, the throughput ascends from 4083 pcu/h under the 50% MPRs to a peak throughput of 4588 pcu/h under the 70% MPRs. Ultimately, the throughput begins to decline progressively with the increment of MPRs, reaching a nadir of 3758 pcu/h.

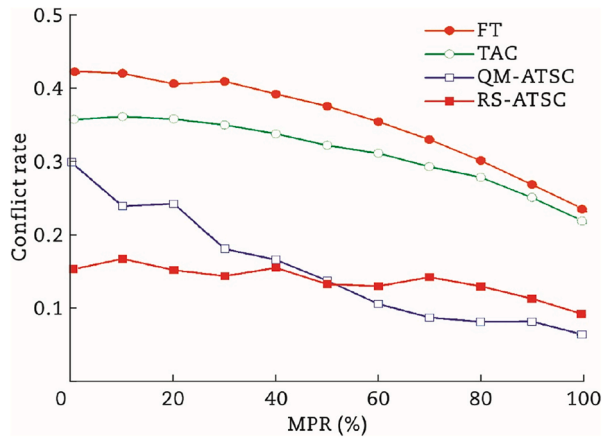


Fig. 5 – Traffic conflict rate at the selected intersection using four approaches under each MPR.

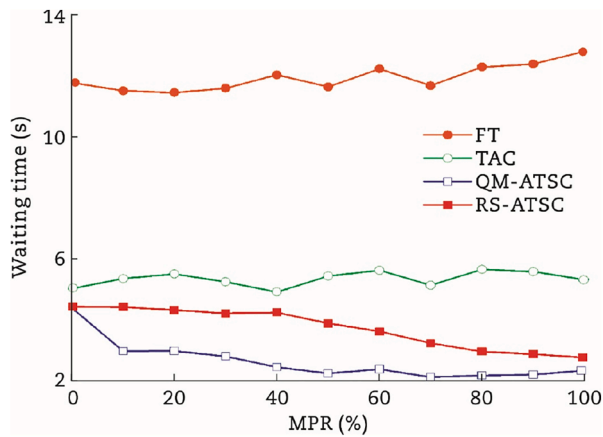


Fig. 6 – Average waiting time at the selected intersection using four approaches under each MPR.

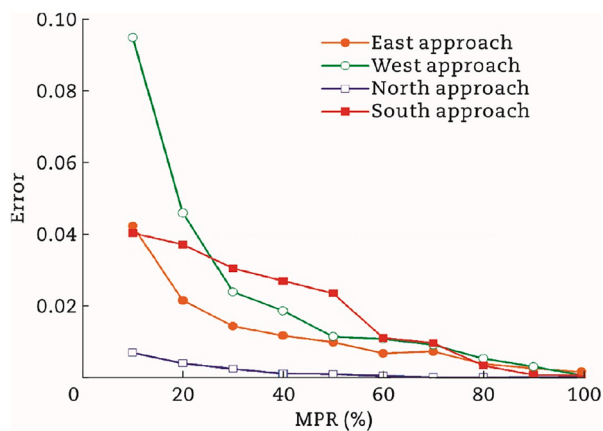


Fig. 7 – Error of density evaluation using CAVs data under each MPR.

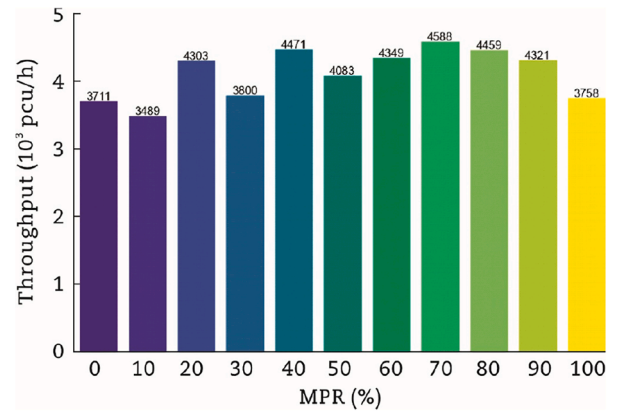


Fig. 8 – Throughput of the selected intersection under each MPR.

In this study, throughput and MPRs do not exhibit a linear relationship. Under low MPRs, fluctuations in throughput may be associated with higher density estimation errors (Fig. 7). Under high MPRs, the trend of throughput first increasing and then decreasing is related to the driving patterns of CAVs. It is widely acknowledged that CAVs tend to prioritize safety over efficiency during their operation (Fu et al., 2024). Consequently, when the performance of QM-ATSC is significantly enhanced under high MPRs, it can lead to greater throughput. However, as MPRs increases further, the impact of CAVs driving patterns on throughput becomes predominant. Therefore, under 70% MPRs, the throughput actually decreases. This trend is consistent with the research conducted by He et al. (2022). On the other hand, by comparing with Fig. 6, it can be observed that the relationship between throughput and waiting time is not a simple linear one but rather a complex interplay of multiple factors (Cho et al., 2022).

6.2. Sensitivity of traffic demand

This study had a sensitivity analysis to test the performance of QM-ATSC under different traffic demands. Based on the existing traffic volume in this study, we used certain magnifications (1.5, 1.2, 0.8, and 0.5) to test the safety and efficiency of QM-ATSC in different traffic demands. The 1.5 magnification and the 1.2 magnification of the existing traffic volumes represented oversaturated conditions, while the 0.8 magnification and the 0.5 magnification represented off-peak scenarios. The waiting time (WT) and conflict rate (CR) results are shown in Table 6.

In relation to the waiting time, an increase in the CAVs MPRs leads to a decrease in waiting time for both off-peak and over-saturated scenarios. Specifically, when the magnification is 0.5, the implementation of 100% MPRs results in a decrease of 1.8 s (equivalent to 47.4%) compared to the 0 MPRs. Similarly, when the magnification is 0.8, the waiting time decreases by 1.9 s (equivalent to 44.2%). For a magnification of 1.2, 100% MPRs leads to a reduction of 3.3 s (equivalent to 56.9%). Further, at a magnification of 1.5, 100% MPRs results in a decrease of 4.7 s (equivalent to 43.1%).

Table 6 – Performance of QM-ATSC under different magnification of traffic volumes.

MPR (%)	Magnification indicator							
	1.5		1.2		0.8		0.5	
	WT (s)	CR	WT (s)	CR	WT (s)	CR	WT (s)	CR
0	10.9	0.58	5.8	0.56	4.3	0.24	3.8	0.16
10	8.4	0.51	4.1	0.49	3.5	0.23	2.9	0.14
20	7.3	0.57	4.1	0.42	3.4	0.24	3.0	0.11
30	6.9	0.62	4.8	0.53	2.1	0.27	1.7	0.11
40	8.0	0.70	4.9	0.55	2.6	0.23	1.7	0.12
50	6.8	0.53	4.9	0.50	2.0	0.19	2.3	0.18
60	6.8	0.48	4.0	0.49	2.3	0.17	1.7	0.06
70	6.3	0.45	4.6	0.47	2.3	0.13	2.1	0.10
80	6.4	0.45	3.3	0.32	2.8	0.16	2.5	0.12
90	6.2	0.43	2.8	0.23	2.3	0.13	2.2	0.09
100	6.2	0.37	2.5	0.28	2.4	0.10	2.0	0.07

QM-ATSC demonstrates remarkable optimization capabilities in terms of waiting time in high traffic volume scenarios. However, it exhibits a diminishing optimization potential as traffic volume reaches excessively high levels. For 0.5 and 0.8 magnifications of existing traffic volume, the waiting time converges to approximately 2 s with the increasing MPRs. Similarly, at 1.2 magnification and existing traffic volume, the waiting time remains around 2 s under 100% MPRs. Consequently, implementing QM-ATSC at this intersection may require a minimum waiting time of 2 s.

In terms of the conflict rate, the QM-ATSC algorithm exhibits varying effectiveness in scenarios of low and high traffic volume. In low traffic volume scenarios, at 0.5 magnification of the traffic volume, the overall conflict rates are relatively low across all MPRs. However, the conflict rates fluctuate slightly as the MPRs increases but remain within an acceptable range. It should be noted that the conflict rates are high with lower MPRs levels (0–30%) under 1.2 and 1.5 magnifications of existing traffic volume conditions. However, as MPRs increases, there is a noticeable decrease in conflict rates. For the 1.2 magnification condition, the conflict rates decrease significantly when the MPRs reaches 70% and continue to decrease as the MPRs increases. Similarly, for the 1.5 magnification condition, the conflict rates decrease substantially when the MPRs reaches 50% and continue to decrease as the MPRs increases further.

These results suggest that the QM-ATSC algorithm has a more significant impact and effectiveness in high traffic volume scenarios. In low traffic volume scenarios, the algorithm still contributes to maintaining acceptable conflict rates, but the differences in effectiveness between varying MPRs levels are less pronounced.

6.3. Penalty function $\lambda(t_r)$

The penalty function $\lambda(t_r)$ was used to illustrate the relationship between the cost of green time and the red time of the intersection. Specifically, $\lambda(t_r)$ reflects the operational cost of increasing the red time and represents the need to reduce the green time accordingly. In this study, $\lambda(t_r)$ is assumed to be an exponential function of the lanes occupancy rate, which is an approximation of the actual relationship between $\lambda(t_r)$ and traffic volume, approach geometry, and driver psychology.

Haddad et al. (2009) found that the signal control optimization of intersections has a linear relationship with queue length but a non-linear relationship when considering other factors, such as driver psychology. Specifically, drivers become more anxious during congestion, and this anxiety aggravates with increasing red time (Noland and Quddus, 2005). Additionally, the assumed exponential function of $\lambda(t_r)$ is consistent with previous research (Verhoef and Rouwendal, 2004).

In general, while the assumed exponential function may not perfectly represent the actual relationship between $\lambda(t_r)$ and various factors, it serves as a valuable approximation for evaluating the performance of the four signal control approaches in this study.

6.4. Effect of the sensitivity coefficient ϵ

In this study, the errors in the QM-ATSC approach were identified as a potential reason for the worse performance under low MPRs. To address this issue, the performance of QM-ATSC was optimized by adjusting the ϵ in Eq. (8). Specifically, ϵ was modified to conduct the $\mathcal{A}_m$ based on the occupancy rate of every approach, as shown in Eq. (11).

$$\epsilon_i = \frac{1000Q_i l_i}{3600L_i n_i} \quad (11)$$

where Q_i is the traffic volume of approach i (veh/h), ϵ_i is the average traffic density for each approach.

The safety performance of the modified QM-ATSC with adjusted ϵ is shown in Fig. 9. The results show that the modified QM-ATSC leads to a lower conflict rate compared to the original approach ($\epsilon = 1$), with reductions of 19.6%, 25.9% and 22.1% under 10%, 20% and 30% MPRs, respectively.

For this reason, it demonstrates that modifying the value in QM-ATSC can improve safety performance under low MPRs. Consequently, accounting for the occupancy rate of each lane allows for a more accurate optimization of the green time.

6.5. Effect of the driven-behaviour models

In this study, the proposed driven-behaviour model is affected by a variety of factors, including region, weather, and holiday.

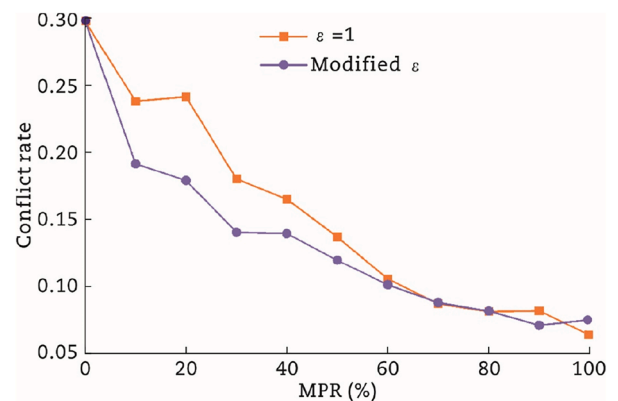


Fig. 9 – Traffic conflict rate using the proposed QM-ATSC with difference ϵ under each MPR.

Table 7 – Driven-behaviour model B.

Parameter	MinGap (m)	Accel (m/s ²)	Decel (m/s ²)	Tau (s)	MaxSpeed (m/s)
Type 1	$\mathcal{N}(0.3, 1.0)$	$\mathcal{N}(0.5, 1.0)$	$\mathcal{N}(2.0, 1.0)$	$\mathcal{N}(1.1, 1.0)$	$\mathcal{N}(8.0, 1.0)$
Type 2	$\mathcal{N}(1.8, 1.0)$	$\mathcal{N}(2.0, 1.0)$	$\mathcal{N}(4.0, 1.0)$	$\mathcal{N}(2.1, 1.0)$	$\mathcal{N}(10.0, 1.0)$
Type 3	$\mathcal{N}(3.3, 1.0)$	$\mathcal{N}(3.5, 1.0)$	$\mathcal{N}(6.0, 1.0)$	$\mathcal{N}(3.1, 1.0)$	$\mathcal{N}(12.0, 1.0)$

The model assumed that each parameter followed a Gaussian distribution, and the distribution parameters were calibrated using statistical data.

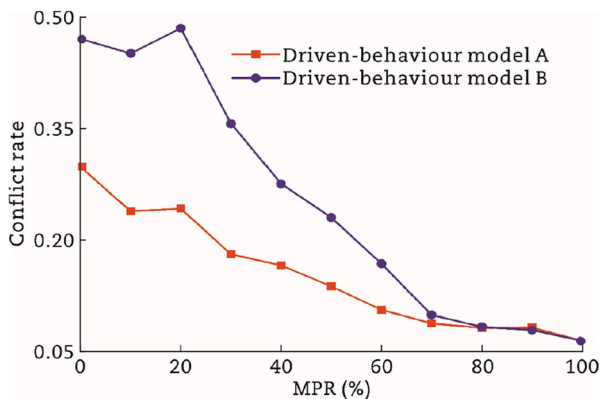
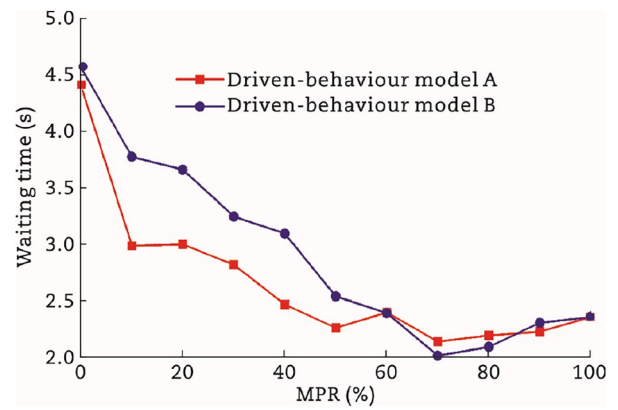
To test the effect of the driven-behaviour model on the overall performance of the intersection, we reset the driven-behaviour model, as shown in Table 7. This reset model (the driven-behaviour model B) was compared with the driven-behaviour model A (Fig. 4 and Table 5) to assess the impact of each parameter on the overall performance of the proposed QM-ATSC.

Fig. 10 shows the conflict rates of the QM-ATSC under the driven-behaviour model B compared to the model A. The results indicate a significant difference in conflict rates between the two models under low CAVs MPRs. However, this difference becomes insignificant when the MPRs is greater than 70%.

Similarly, Fig. 11 shows the operational performance of QM-ATSC under the driven-behaviour model B and the model A. The results indicate that the driven-behaviour model B has similar operational performance compared to the model A under high MPRs. On the contrary, there are significant differences in the performance of the two models under low MPRs (<60%).

The driven-behaviour model is only appropriate for HDVs in the emerging mixed traffic flow. Therefore, there will be significant differences in efficiency and safety under low MPRs. However, at 0 MPRs, the difference in waiting time is insignificant between model A and model B. This phenomenon can be explained by the fact that the traffic flow has reached a congested state close to saturation at 0 MPRs. In other words, the waiting time has already reached its limit in the QM-ATSC system. The difference in waiting time is insignificant without CAV data at 0 MPRs.

On the other hand, at 0 MPRs, there is a significant difference in the conflict rate between model A and model B. As a

**Fig. 10 – Comparison of the conflict rate between the origin and test driven-behaviour model.****Fig. 11 – Comparison of the waiting time between the origin and test driven-behaviour model.**

result, it indicates that models A and B significantly impact the safety level under low MPRs. Although there is little difference in efficiency between the two models at 0 MPRs, the stability of model A is worse than that of model B. A higher conflict rate implies a higher crash risk. These findings suggest that the driven-behaviour model can significantly impact QM-ATSC performance under low MPRs, but its impact diminishes at high MPRs. This highlights the importance of accurately modelling driver behaviour and accounting for external factors such as weather and holidays when designing signal control approaches.

6.6. Limitations of QM-ATSC

Despite the superior performance of the QM-ATSC proposed in this study, there are still some limitations that need to be addressed by additional studies.

- (1) In terms of safety, the QM-ATSC is only superior to RS-ATSC under high MPRs. To resolve this issue, reducing the error in density estimation and optimizing the method of density estimation are key research directions that need to be considered in the future.
- (2) Although the QM-ATSC outperforms other methods in terms of waiting time, the throughput under high MPRs is still limited. Therefore, future work needs to enhance throughput while ensuring safety under high MPRs.
- (3) The current testing of the QM-ATSC is limited to the small intersection. In subsequent research, it is necessary to explore the performance of the two-phase QM-ATSC and to develop testing for multi-phase QM-ATSC in large intersections.

- (4) The driven-behaviour models A and B used in this study are based on the assumption of the standard Gaussian model. Therefore, subsequent research needs to calibrate them based on real-world traffic data.

7. Conclusions

In conclusion, the proposed Q-learning-based modified ATSC algorithm (QM-ATSC), coupled with the parametric driver behaviour model, presents a promising approach to address the limitations of traditional signal control in the emerging mixed traffic flow. This algorithm offers several important contributions by incorporating CAV data and human factors.

Firstly, QM-ATSC demonstrates a lower cost and ease of implementation, making it highly suitable for integrating various small-scale intersections without substantial infrastructure upgrades. This accessibility opens opportunities for widespread adoption. Secondly, this study focuses on the integration of loop detector and CAV data to drive the QM-ATSC algorithm, which can still maintain good performance with insufficient and low-quality data. Integrating a large number of small intersections into a connected system is a monumental undertaking. The QM-ATSC algorithm proposed in this study effectively reduces the cost of such integration, while ensuring optimization benefits even with low-quality data. Thirdly, the parametric driver behaviour model provides a valuable tool for characterizing different driving patterns and assessing the impact of traffic control in mixed traffic environments. Further, the driving patterns identified by the proposed model can be integrated as an additional layer in high-definition (HD) maps. This integration would offer drivers and traffic control systems more accurate and detailed information on traffic flow.

The practical benefits of the proposed algorithm are significant, and it has the potential to improve future traffic management. However, future work is also required to overcome some limitations of the study: (i) further research is needed to explore the capabilities of the proposed approach and the driving behaviour model fully, especially as CAV technologies rapidly evolve and traffic flow conditions continue to change; (ii) left-turn and right-turn traffic and phase sequence optimizations need to be given attention in

future research; (iii) calibration is required for the proposed driven-behaviours model using extensive data; iv) the lane-changing behaviour needs to be considered to refine the driven-behaviour model; v) it needs to explore the possible MPRs for QM-ATSC without loop detectors; vi) fuel consumption will be considered in the future research.

Conflict of interest

Chuanyun Fu is a young academic editor of Journal of Traffic and Transportation Engineering (English Edition) and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

CRedit authorship contribution statement

Chuanyun Fu: Data curation, Investigation, Supervision, Writing – review & editing, Conceptualization, Funding acquisition, Methodology, Writing – original draft. **Zhaoyou Lu:** Methodology, Visualization, Investigation, Validation, Writing – original draft. **Naikan Ding:** Methodology, Visualization, Validation, Writing – review & editing. **Zhixing Li:** Conceptualization, Writing – review & editing, Investigation.

Acknowledgments

The work was jointly supported by the National Natural Science Foundation of China (72371082), the Natural Science Foundation of Sichuan Province of China (2024NSFSC0184), the Opening Project of Intelligent Policing Key Laboratory of Sichuan Province, China (ZNJW2023KFZD001), the Fundamental Research Funds for the Central Universities (FRFCUAUGA5710010222, WUT: 233144001), and Toyota Motor Technical Research and Service (Shanghai) Co., Ltd.

Appendix. Q-learning process in the QM-ATSC

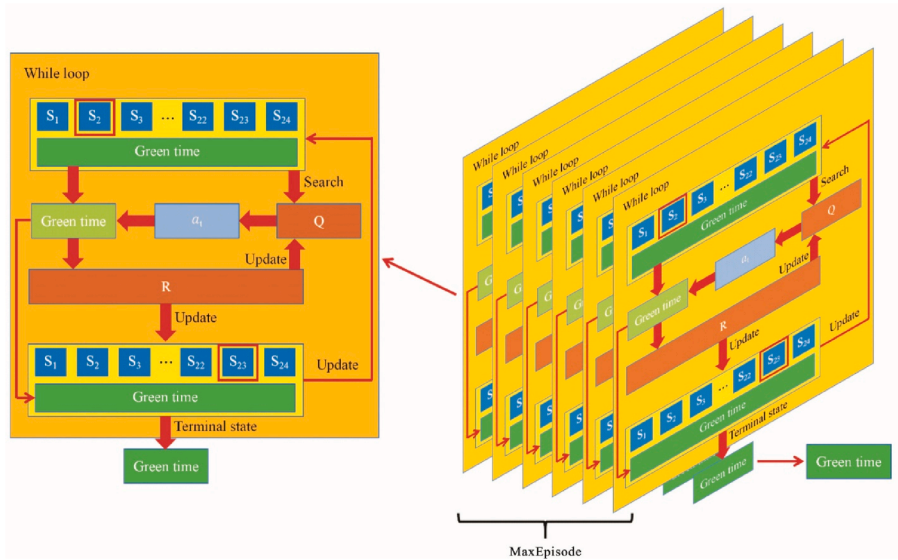


Fig. A1 – Process of Q-learning in the QM-ATSC.

In the QM-ATSC framework, the process of Q-learning is delineated in Fig. A1.

- i) First, the initial state of the intersection, such as S_2 shown in Fig. A1, can be obtained from Table 1.
- ii) Then, a random green time is selected, and the green time can form a 2D state set with S_2 .
- iii) Next, the optimal action a_t corresponding to the current state is searched for in the Q-table, and the green time is updated based on a_t .
- iv) The reward value is calculated using Eq. (5), and the Q-table is updated accordingly.
- v) Subsequently, the new green time is used as input, and the state is updated using Eq. (2) and Table 1, resulting in a new state such as S_{23} as shown in Fig. A1.
- vi) The 2D state set is updated with the green time and the new state S_{23} .
- vii) The process iii-vi continue until the current state becomes a terminal state, at which point the final green time for this episode is obtained, and the next episode begins.

Once all episodes are completed, the converged final green time can be obtained.

REFERENCES

- AASHTO, 2010. Highway Capacity Manual (2010). American Association of State Highway and Transportation Officials (AASHTO), Washington DC.
- Abdoos, M., 2020. A cooperative multiagent system for traffic signal control using game theory and reinforcement learning. *IEEE Intelligent Transportation Systems Magazine* 13 (4), 6–16.
- Arel, I., Liu, C., Urbanik, T., et al., 2010. Reinforcement learning-based multi-agent system for network traffic signal control. *IET Intelligent Transport Systems* 4 (2), 128–135.
- Berhanu, Y., Alemayehu, E., Schröder, D., 2023. Examining car accident prediction techniques and road traffic congestion: a comparative analysis of road safety and prevention of world challenges in low-income and high-income countries. *Journal of Advanced Transportation* 2023, 6643412.
- Buckley, L., Kaye, S.A., Pradhan, A.K., 2018. Psychosocial factors associated with intended use of automated vehicles: a simulated driving study. *Accident Analysis & Prevention* 115, 202–208.
- Celtek, S.A., Durdu, A., Ali, M.E.M., 2021. Evaluating action durations for adaptive traffic signal control based on deep Q-learning. *International Journal of Intelligent Transportation Systems Research* 19 (3), 557–571.
- Chang, H., Li, L., Huang, J., et al., 2022. Tracking traffic congestion and accidents using social media data: a case study of Shanghai. *Accident Analysis & Prevention* 169, 106618.
- Chen, B., Zhao, X., Li, Z., et al., 2025. Mapping the knowledge domain of crash risk in older drivers studies: a scientometric analysis. *Journal of Traffic and Transportation Engineering (English Edition)* 12 (3), 587–602.
- Chien, S.I.J., Mouskos, K.C., Chowdhury, S.M., 2001. Generating driver population for the microscopic simulation model (CORSIM). *Simulation* 76 (1), 40–45.
- Cho, S., Shrestha, B., Salah, B., et al., 2022. A proposed waiting time algorithm for a prediction and prevention system of traffic accidents using smart sensors. *Electronics* 11 (11), 1765.
- Chu, T., Wang, J., Codecà, L., et al., 2019. Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Transactions on Intelligent Transportation Systems* 21 (3), 1086–1095.
- Devailly, F.X., Larocque, D., Charlin, L., 2021. IG-RL: inductive graph reinforcement learning for massive-scale traffic signal

- control. *IEEE Transactions on Intelligent Transportation Systems* 23 (7), 7496–7507.
- Dikmen, M., Burns, C., 2017. Trust in autonomous vehicles: the case of Tesla Autopilot and Summon. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, 2017.
- El-Tantawy, S., Abdulhai, B., Abdelgawad, H., 2013. Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): methodology and large-scale application on downtown Toronto. *IEEE Transactions on Intelligent Transportation Systems* 14 (3), 1140–1150.
- Essa, M., Sayed, T., 2020. Self-learning adaptive traffic signal control for real-time safety optimization. *Accident Analysis & Prevention* 146, 105713.
- Evans, L., 1996. The dominant role of driver behavior in traffic safety. *American Journal of Public Health* 86 (6), 784–786.
- Feng, Y., Head, K.L., Khoshmagham, S., et al., 2015. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies* 55, 460–473.
- Feng, Y., Zheng, J., Liu, H., 2018. Real-time detector-free adaptive signal control with low penetration of connected vehicles. *Transportation Research Record* 2672, 35–44.
- Formosa, N., Quddus, M., Ison, S., et al., 2020. Predicting real-time traffic conflicts using deep learning. *Accident Analysis & Prevention* 136, 105429.
- Fu, C., Lu, Z., Ding, N., et al., 2024. Distance headway-based safety evaluation of emerging mixed traffic flow under snowy weather. *Physica A: Statistical Mechanics and its Applications* 642, 129792.
- Fu, C., Zhang, Y., Bie, Y., et al., 2016. Comparative analysis of driver's brake perception-reaction time at signalized intersections with and without countdown timer using parametric duration models. *Accident Analysis & Prevention* 95, 448–460.
- Ge, H., Song, Y., Wu, C., et al., 2019. Cooperative deep Q-learning with Q-value transfer for multi-intersection signal control. *IEEE Access* 7, 40797–40809.
- Genders, W., Razavi, S., 2019. Asynchronous n-step Q-learning adaptive traffic signal control. *Journal of Intelligent Transportation Systems* 23 (4), 319–331.
- Geronimo, M.F., Martinez, E.G.H., Vazquez, E.D.F., et al., 2021. A multiagent systems with Petri Net approach for simulation of urban traffic networks. *Computers, Environment and Urban Systems* 89, 101662.
- Ghanbarikarekani, M., Qu, X., Zeibots, M., et al., 2018. Minimizing the average delay at intersections via presignals and speed control. *Journal of Advanced Transportation* 2018, 4121582.
- Ghosh, A., Parisini, T., 2022. Traffic control in a mixed autonomy scenario at urban intersections: an optimal control approach. *IEEE Transactions on Intelligent Transportation Systems* 23 (10), 17325–17341.
- Ghoul, T., Sayed, T., 2021. Real-time signal-vehicle coupled control: an application of connected vehicle data to improve intersection safety. *Accident Analysis & Prevention* 162, 106389.
- Ha, T., Kim, S., Seo, D., et al., 2020. Effects of explanation types and perceived risk on trust in autonomous vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour* 73, 271–280.
- Haddad, J., Mahalel, D., De Schutter, B., et al., 2009. Optimal steady-state traffic control for isolated intersections. *IFAC Proceedings Volumes* 42 (6), 96–101.
- Hayward, J.C., 1971. Near Misses as a Measure of Safety at Urban Intersections. Pennsylvania State University, University Park.
- He, S., Ding, F., Lu, C., et al., 2022. Impact of connected and autonomous vehicle dedicated Lane on the freeway traffic efficiency. *European Transport Research Review* 14 (1), 12.
- Kononov, J., Durso, C., Reeves, D., et al., 2012. Relationship between traffic density, speed, and safety and its implications for setting variable speed limits on freeways. *Transportation Research Record* 2280, 1–9.
- Lee, J., Chung, J., Sohn, K., 2020. Reinforcement learning for joint control of traffic signals in a transportation network. *IEEE Transactions on Vehicular Technology* 69 (2), 1375–1387.
- Levermore, T., Ordys, A., Deng, J., 2014. A review of driver modelling. In: 2014 UKACC International Conference on Control, Loughborough, 2014.
- Li, Z., Eleftheriadou, L., Ranka, S., 2014. Signal control optimization for automated vehicles at isolated signalized intersections. *Transportation Research Part C: Emerging Technologies* 49, 1–18.
- Li, H., Wang, L., Hu, H., et al., 2023a. Optimal matching between vehicle speed and lighting at intersection based on traffic risk analysis. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 9 (2), 04023004.
- Li, J., Yu, C., Shen, Z., et al., 2023b. A survey on urban traffic control under mixed traffic environment with connected automated vehicles. *Transportation Research Part C: Emerging Technologies* 154, 104258.
- Li, T., Zhao, M., Xiong, J., et al., 2025. Lane capacity and cost function for the mixed traffic scenario with connected and autonomous vehicles. *Journal of Traffic and Transportation Engineering* 25 (6), 146–156.
- Liang, X., Du, X., Wang, G., et al., 2019. A deep Q learning network for traffic lights' cycle control in vehicular networks. *IEEE Transactions on Vehicular Technology* 68 (2), 1243–1253.
- Liang, X., Guler, S.I., Gayah, V.V., 2020. A heuristic method to optimize generic signal phasing and timing plans at signalized intersections using connected vehicle technology. *Transportation Research Part C: Emerging Technologies* 111, 156–170.
- Lin, W., Wei, H., Yang, L., et al., 2025. CAV-generated mobility data modeling mechanism for adaptive signal control. *Journal of Traffic and Transportation Engineering (English Edition)* 12 (2), 361–377.
- Liu, M., Deng, J., Xu, M., et al., 2017. Cooperative deep reinforcement learning for traffic signal control. In: 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), Halifax, 2017.
- Liu, W., Qin, G., He, Y., et al., 2017. Distributed cooperative reinforcement learning-based traffic signal control that integrates V2X networks' dynamic clustering. *IEEE Transactions on Vehicular Technology* 66 (10), 8667–8681.
- Lord, D., Manar, A., Vizioli, A., 2005. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accident Analysis & Prevention* 37 (1), 185–199.
- Lu, Z., Ding, N., Gao, J., et al., 2023. Safety benefits evaluation of mixed traffic flow with connected and automated vehicles under snowy conditions. In: 2023 7th International Conference on Transportation Information and Safety (ICTIS), Xi'an, 2023.
- Lu, Z., Ding, N., Lu, L., et al., 2022. Optimizing signal timing of the arterial-branch intersection: a fuzzy control and nonlinear programming approach. *Asian Journal of Control* 24 (6), 2952–2968.
- Luo, Q., Chen, X., Yuan, J., et al., 2020. Study and simulation analysis of vehicle rear-end collision model considering driver types. *Journal of Advanced Transportation* 2020, 7878656.

- Manchon, J.B., Bueno, M., Navarro, J., 2022. How the initial level of trust in automated driving impacts drivers' behaviour and early trust construction. *Transportation Research Part F: Traffic Psychology and Behaviour* 86, 281–295.
- Mandelbaum, A., Zeltyn, S., 2007. Service engineering in action: the Palm/Erlang-A queue, with applications to call centers. In: Spath, D., Fährnich, K.P. (Eds.), *Advances in Services Innovations*. Springer, Berlin, pp. 17–45.
- Mannion, P., Duggan, J., Howley, E., 2016. An experimental review of reinforcement learning algorithms for adaptive traffic signal control. In: Mccluskey, T.L., Kotsialos, A., Müller, J.P., et al. (Eds.), *Autonomic Road Transport Support Systems*. Birkhäuser, Cham, pp. 47–66.
- Matin, A., Dia, H., 2022. Impacts of connected and automated vehicles on road safety and efficiency: a systematic literature review. *IEEE Transactions on Intelligent Transportation Systems* 24 (3), 2705–2736.
- Miao, L., Leitner, D., 2021. Adaptive traffic light control with quality-of-service provisioning for connected and automated vehicles at isolated intersections. *IEEE Access* 9, 39897–39909.
- Miletić, M., Ivanjko, E., Gregurić, M., et al., 2022. A review of reinforcement learning applications in adaptive traffic signal control. *IET Intelligent Transport Systems* 16 (10), 1269–1285.
- Mo, Z., Li, W., Fu, Y., et al., 2022. CVLight: decentralized learning for adaptive traffic signal control with connected vehicles. *Transportation Research Part C: Emerging Technologies* 141, 103728.
- Molnar, L.J., Ryan, L.H., Pradhan, A.K., et al., 2018. Understanding trust and acceptance of automated vehicles: an exploratory simulator study of transfer of control between automated and manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour* 58, 319–328.
- Moradi, H., Sasaninejad, S., Wittevrongel, S., et al., 2022. The contribution of connected vehicles to network traffic control: a hierarchical approach. *Transportation Research Part C: Emerging Technologies* 139, 103644.
- Noland, R.B., Quddus, M.A., 2005. Congestion and safety: a spatial analysis of London. *Transportation Research Part A: Policy and Practice* 39 (7–9), 737–754.
- Qi, H., Dai, R., Tang, Q., et al., 2020. Coordinated intersection signal design for mixed traffic flow of human-driven and connected and autonomous vehicles. *IEEE Access* 8, 26067–26084.
- Qu, D., Li, A., Zhang, Z., et al., 2025. Multi-vehicle responsive longitudinal and lateral cooperative control method for networked autonomous driving. *Journal of Traffic and Transportation Engineering* 25 (4), 281–295.
- Rahman, M., Abdel-Aty, M., Lee, J., et al., 2019. Safety benefits of arterials' crash risk under connected and automated vehicles. *Transportation Research Part C: Emerging Technologies* 100, 354–371.
- Reyad, P., Sayed, T., Essa, M., et al., 2021. Real-time crash-risk optimization at signalized intersections. *Transportation Research Record* 2676, 32–50.
- Shaikh, P.W., El-Abd, M., Khanafer, M., et al., 2020. A review on swarm intelligence and evolutionary algorithms for solving the traffic signal control problem. *IEEE Transactions on Intelligent Transportation Systems* 23 (1), 48–63.
- Shladover, S.E., Su, D., Lu, X., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record* 2324, 63–70.
- Song, L., Fan, W., 2021. Traffic signal control under mixed traffic with connected and automated vehicles: a transfer-based deep reinforcement learning approach. *IEEE Access* 9, 145228–145237.
- Soria, I., Eleftheriadou, L., Kondyli, A., 2014. Assessment of car-following models by driver type and under different traffic, weather conditions using data from an instrumented vehicle. *Simulation Modelling Practice and Theory* 40, 208–220.
- Tan, T., Bao, F., Deng, Y., et al., 2019. Cooperative deep reinforcement learning for large-scale traffic grid signal control. *IEEE Transactions on Cybernetics* 50 (6), 2687–2700.
- Taylor, C.J., Salmon, P.M., Read, G.J.M., 2025. Applying systems thinking to understand the impacts of traffic safety culture on young driver speeding behavior. *Journal of Traffic and Transportation Engineering (English Edition)* 12 (6), 1839–1854.
- Trabia, M.B., Kaseko, M.S., Ande, M., 1999. A two-stage fuzzy logic controller for traffic signals. *Transportation Research Part C: Emerging Technologies* 7 (6), 353–367.
- UK Highways Agency, 1996. *UK Design Manual for Roads and Bridges*. UK Highways Agency, Birmingham.
- Verhoef, E.T., Rouwendal, J., 2004. A behavioural model of traffic congestion: endogenizing speed choice, traffic safety and time losses. *Journal of Urban Economics* 56 (3), 408–434.
- Virdi, N., Grzybowska, H., Waller, S.T., et al., 2019. A safety assessment of mixed fleets with connected and autonomous vehicles using the surrogate safety assessment module. *Accident Analysis & Prevention* 131, 95–111.
- Wang, C., 2010. *The Relationship Between Traffic Congestion and Road Accidents: an Econometric Approach Using GIS* (PhD thesis). Loughborough University, Loughborough.
- Wang, Y., Jiang, Y., Wu, Y., et al., 2024. Mitigating traffic oscillation through control of connected automated vehicles: a cellular automata simulation. *Expert Systems with Applications* 235, 121275.
- Wang, L., Li, H., Li, S., et al., 2022. Gradient illumination scheme design at the highway intersection entrance considering driver's light adaption. *Traffic Injury Prevention* 23 (5), 266–270.
- Wang, C., Liu, L., Xu, C., 2020. Developing a new spatial unit for macroscopic safety evaluation based on traffic density homogeneity. *Journal of Advanced Transportation* 2020, 1718541.
- Wang, C., Quddus, M.A., Ison, S.G., 2013. The effect of traffic and road characteristics on road safety: a review and future research direction. *Safety Science* 57, 264–275.
- Wang, Y., Yang, X., Liang, H., et al., 2018. A review of the self-adaptive traffic signal control system based on future traffic environment. *Journal of Advanced Transportation* 2018, 1096123.
- Wei, H., Mashayekhy, L., Papineau, J., 2018. Intersection management for connected autonomous vehicles: a game theoretic framework. In: *21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, 2018.
- Wei, Z., Peng, T., Wei, S., 2022. A robust adaptive traffic signal control algorithm using Q-learning under mixed traffic flow. *Sustainability* 14 (10), 5751.
- Wu, J., Xu, H., 2018. The influence of road familiarity on distracted driving activities and driving operation using naturalistic driving study data. *Transportation Research Part F: Traffic Psychology and Behaviour* 52, 75–85.
- Xiao, G., Lee, J., Jiang, Q., et al., 2021. Safety improvements by intelligent connected vehicle technologies: a meta-analysis considering market penetration rates. *Accident Analysis & Prevention* 159, 106234.
- Yadav, A.K., Velaga, N.R., 2021. Investigating the effects of driving environment and driver characteristics on drivers' compliance with speed limits. *Traffic Injury Prevention* 22 (3), 201–206.
- Yang, J., Du, F., Qu, W., et al., 2013. Effects of personality on risky driving behavior and accident involvement for Chinese drivers. *Traffic Injury Prevention* 14 (6), 565–571.
- Yang, K., Guler, S.I., Menendez, M., 2016. Isolated intersection control for various levels of vehicle technology: conventional, connected, and automated vehicles. *Transportation Research Part C: Emerging Technologies* 72, 109–129.

- Yao, Z., Wu, Y., Wang, Y., et al., 2023. Analysis of the impact of maximum platoon size of CAVs on mixed traffic flow: an analytical and simulation method. *Transportation Research Part C: Emerging Technologies* 147, 103989.
- Yu, Y., Zhang, J., Li, Y., 2012. Efficient modeling scheme based on smart neural network. *Communications in Computer and Information Science* 308, 85–91.
- Zaied, A.N.H., Al Othman, W., 2011. Development of a fuzzy logic traffic system for isolated signalized intersections in the State of Kuwait. *Expert Systems with Applications* 38 (8), 9434–9441.
- Zeng, J., Hu, J., Zhang, Y., 2019. Training reinforcement learning agent for traffic signal control under different traffic conditions. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, 2019.
- Zheng, Y., Chase, R.T., Eleftheriadou, L., et al., 2017. Driver types and their behaviors within a high level of pedestrian activity environment. *Transportation Letters* 9 (1), 1–11.
- Zhu, H., Zhang, N., Wu, W., 2015. A modified two-lane traffic model considering drivers' personality. *Physica A: Statistical Mechanics and its Applications* 428, 359–367.



Chuanyun Fu is an associate professor at School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, China. From March 2016 to October 2021, he was an assistant professor at School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China. He obtained his BE, ME, and PhD degrees from Harbin Institute of Technology, Harbin, China. From December 2019 to February 2024, he has been a postdoctoral fellow at

Department of Civil Engineering, University of British Columbia, Vancouver, Canada. His research interests cover traffic violation modeling, traffic crash prediction, traffic conflict analysis, and advanced statistical modeling.



Zhaoyou Lu is a PhD student at School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, China. His major research interests focus on traffic safety, traffic simulation, and traffic control.



Naikan Ding received the PhD degree from Wuhan University of Technology (WUT), Wuhan, China, in 2017. He was a post-doctoral researcher at Institute of Materials and Systems for Sustainability, Nagoya University, and is an associate professor at Intelligent Transportation Systems Research Center (ITSC), WUT. His research interests include traffic safety, driving behavior, human factors, and connected automated vehicles.



Zhixing Li obtained a bachelor's degree in engineering from Tsinghua University, Beijing, China. Currently, he is pursuing a master's degree in data science and AI at Chalmers University of Technology in Sweden. His research interests include deep learning, sensor fusion, and semantic communication.