# Adaptive Highway Traffic Management: A Reinforcement Learning Approach for Variable Speed Limit Control with Random Anomalies

Bálint Pelenczei[1] [a], István Gellért Knáb[1] [b], Bálint Kővári[2,3] [c], Tamás Bécsi[2] [d] and László Palkovics[1,4] [e]

[1] *Systems and Control Laboratory, HUN-REN Institute for Computer Science and Control (SZTAKI), Budapest, Hungary*

[2] *Department of Control for Transportation and Vehicle Systems, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, Budapest, Hungary*

[3] *Asura Technologies Ltd., Budapest, Hungary*

[4] *Széchenyi István University, Győr, Hungary*

*{pelenczei.balint, knab.istvan.gellert}@sztaki.hun-ren.hu, {kovari.balint, becsi.tamas}@kjk.bme.hu,*

Keywords: Reinforcement Learning, Variable Speed Limit Control, Intelligent Transportation Systems, Cooperative Traffic Control, Multi-Agent Systems.

Abstract: Efficient traffic flow management on highway scenarios is crucial for ensuring safety and minimizing emissions through the reduction of so-called shockwave effects. In this paper, we propose a novel approach based on cooperative Multi Agent Reinforcement Learning for optimizing traffic flow, utilizing Variable Speed Limit Control in dynamic simulation environments with random anomalies. Our method leverages Reinforcement Learning to adaptively adjust speed limits on distinct road sections in response to alternating traffic conditions, thereby improving not only general traffic flow parameters, but also reducing sustainability measures overall. Through extensive simulations in a Simulation of Urban MObility environment, we demonstrate the superiority of our approach in enhancing traffic flow efficiency and robustness compared to alternative solutions found in literature. Our findings reveal an enhanced performance of RL-based VSL control over traditional approaches due to its generalizability, which contributes to the progression of Intelligent Transportation Systems by presenting a proactive and adaptable resolution for highway traffic management within dynamic real-world contexts.

## 1 INTRODUCTION

Continuous improvements are on the horizon concerning the spread of Artificial Intelligence-based solutions across interconnected domains relevant to everyday life, such as logistics (Richey Jr et al., 2023), autonomous vehicles (Fényes et al., 2021), and lastly traffic control (Kővári et al., 2021). These advancements are positioned to revolutionize traditional practices, offering remarkable levels of efficiency, safety and sustainability. With AI technologies increasingly integrated into various aspects of society, the potential for transformative impact on these critical areas is vast.

[a] https://orcid.org/0000-0001-9194-8574
[b] https://orcid.org/0009-0007-6906-3308
[c] https://orcid.org/0000-0003-2178-2921
[d] https://orcid.org/0000-0002-1487-9672
[e] https://orcid.org/0000-0001-5872-7008

In the realm of Intelligent Transportation Systems (ITS), AI holds immense promise for optimizing traffic flow, enhancing road safety, and reducing environmental impacts (An et al., 2011). Through real-time data analysis, predictive modeling, and adaptive control algorithms, AI-powered ITS solutions can dynamically respond to changing traffic conditions, minimizing congestion and improving overall efficiency. Moreover, the integration of AI with emerging technologies, such as Connected Autonomous Vehicles (CAV), opens up possibilities for seamless vehicle-to-everything (V2X) communication, enabling coordinated traffic management and enhanced safety measures (Kavas-Torris et al., 2022).

However, for the time being, nearly none of the vehicles on public roads possesses the capabilities necessary to utilize these advanced features, resulting in bottlenecks appearing throughout all fundamental parts of a traffic network. At intersections, the pri-



Pelenczei, B., Knáb, I., Kővári, B., Bécsi, T. and Palkovics, L.
Adaptive Highway Traffic Management: A Reinforcement Learning Approach for Variable Speed Limit Control with Random Anomalies.
DOI: 10.5220/0012920700003822

In *Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO 2024) - Volume 2*, pages 117-124
ISBN: 978-989-758-717-7; ISSN: 2184-2809

mary challenge is the Traffic Signal Control (TSC) problem, where optimizing signal timings to manage varying traffic volumes is critical. In urban areas, pedestrian crossings introduce additional complexities, requiring careful coordination to ensure both vehicular flow and pedestrian safety.

Additionally, disturbances caused by sudden braking, lane changes or speed fluctuations are particularly problematic in areas like toll plazas, highway on- and off-ramps and also in the vicinity of lane closures. This phenomenon, known as traffic shockwave, is visualized in Figure 1, where space-time trajectories of vehicles are plotted, therefore slopes of these curves represent the speeds of vehicles.

Advanced traffic management systems, such as Variable Speed Limit Control (VSLC), often integrate Machine Learning and predictive analytics to dynamically adjust speed limits based on real-time traffic conditions on virtually separated road sections, independently from each other. By continuously adapting speed limits in response to real-time traffic conditions, VSLC systems aim to manage disturbances and maintain optimal flow conditions, while minimizing the risk of congestions and accidents.

Therefore, in this research we introduce a cooperative MARL approach for VSLC in a generalized highway setting, utilizing random anomalies. As detailed in Section 6, by leveraging the Machine Learning framework's inherent generalizability, our approach demonstrates superior performance compared to baseline methods outlined in Section 5.1.

## 2 RELATED WORK

In the realm of ITS, various methodological approaches are employed for different tasks, ranging from rule-based systems to sophisticated ML techniques capable of high-level decision-making. This section provides a comprehensive overview of such methods in literature, with a focus on Variable Speed Limit Control.

Rule-based methods are widely accepted due to their simplicity and ease of applicability. For instance, the Motorway Control System (MCS), proposed by (Van Toorenburg and De Kok, 1999), has been successfully applied in real traffic scenarios. This method is also discussed in Section 5.1 in more detail as one of the baselines for this research. Another benchmark, being the Motorway Traffic Flow Control (MTFC) introduced by (Müller et al., 2013), stabilizes traffic flow at the maximum throughput level based on occupancy measures. A comparative study by (Grumert et al., 2018) found, that MTFC outperformed four other methods, including MCS, as shown in Figure 2.

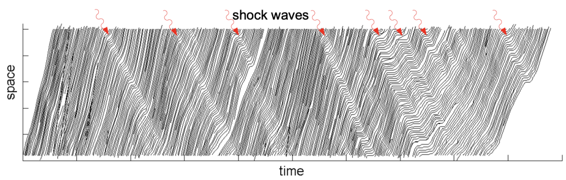


Figure 1: Space-time trajectories of vehicles demonstrating shockwave effects (Huang et al., 2010).

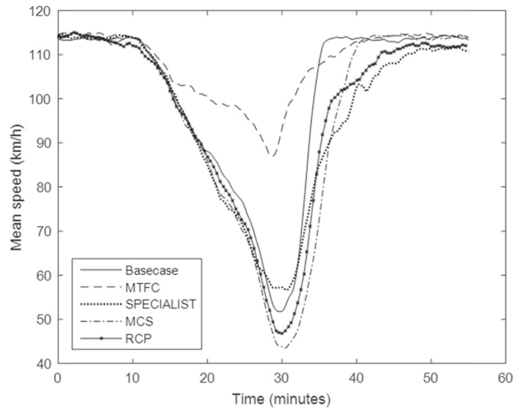


Figure 2: Comparison of mean speeds over the entire traffic network achieved by different algorithms (Grumert et al., 2018).

Another innovative, Reinforcement Learning-based approach is presented by (Kim et al., 2024); they developed a proactive traffic safety management methodology based on real-time crash risk estimation. Their system has been able to reduce real-time crash risk by approximately 55% during lane closure scenarios.

In (Kővári et al., 2024), a Deep Reinforcement Learning approach has been introduced, that utilizes a Multi Agent framework with a Deep Q-Network algorithm to capture the spatial-temporal characteristics of traffic flow. The MARL-based VSL system, trained and tested in a static simulator environment with a fixed location lane drop bottleneck, demonstrated significant improvements in traffic stability and congestion reduction compared to a free-flow baseline.

For further insights into different kinds of methods applied to the VSLC problem, refer to (Khondaker and Kattan, 2015) and (Lahmiss and Khatory, 2020). For an overview of Reinforcement Learning techniques concerning this application, see (Kušić et al., 2020).

# 3 CONTRIBUTION

While numerous approaches address the issue of shockwave effects in highway scenarios, Variable Speed Limit Control has emerged as a particularly promising solution, offering several demonstrated benefits. This paper explores the effectiveness and advantages of Machine Learning by proposing a cooperative Multi-Agent Reinforcement Learning methodology such, that outperforms two well-established traditional solutions cited in the literature.

Therefore, the contribution of this paper is twofold: firstly, we have developed a cooperative MARL framework and successfully evaluated it in dynamic scenarios with random anomalies, hence demonstrating the method's generalizable manner. Secondly, we present the results of an extensive comparative experimentation, according to which the performance of the trained agent surpasses three alternative methods': a simple free-flow scenario with no control realized; the Motorway Control System implemented in Stockholm, Sweden; and lastly the Mainstream Traffic Flow Control method, which is recognized as a superior solution among the compared methods according to (Grumert et al., 2018).

# 4 ENVIRONMENT

In this research, training, testing and evaluation have been conducted using the Simulation of Urban MObility (SUMO) software package proposed in (Alvarez Lopez et al., 2018), which is widely recognized in literature as a leading traffic simulator. Firstly, SUMO is an open-source software capable of simulating both real and artificial traffic networks. Moreover, it offers the ability to numerically monitor various traffic flow metrics, including density, travel time, waiting time, and sustainability parameters, such as $CO_2$ emission, $NO_x$ emission and fuel consumption. Additionally, SUMO supports appropriate randomization and provides a comprehensive set of tools for scenario generation and modification. The package also includes the TraCI interface, enabling environment manipulation through various programming languages. The schematic illustration of the communication among software components is presented later in Figure 4.

The geometric design of the network is illustrated in Figure 3. It comprises a total of eight straight road segments, each 300 m in length, unidirectional and consisting of three lanes. Each lane is 3.2 m wide, making the total segment width 9.6 m. At the beginning of the first segment, designated spawn points for each lane generate and inject vehicles into the simulation randomly at each step. The initial two segments, covering the first 600 m, are solely for observation purposes. Segments 3 and 4 are designated as Variable Speed Limit Control zones, but no anomalies are generated in this section of the network, allowing the VSLC system to proactively prevent and manage anomalies, that may appear in subsequent locations. Segments 5 through 8 are also VSLC zones, where random anomalies can be set up, hence constructing a lane drop bottleneck. In case of an anomaly being generated on any of the lanes, all subsequent sections are closed. Each of the 18 VSLC zones can be independently controlled with thresholded speed limits, ranging from a minimum of 30 km/h to a maximum of 130 km/h, adjusted in 10 km/h increments. The control system's task is to determine the optimal speed limit for each of these zones.

## 4.1 Abstractions

Specifically in Reinforcement Learning, the proper definition of abstractions – being state representation, action space and the reward function – are deemed fundamental, as these are the only connections with the environment through which the agent can develop its behaviour and understand complex inner dynamics of the processes.

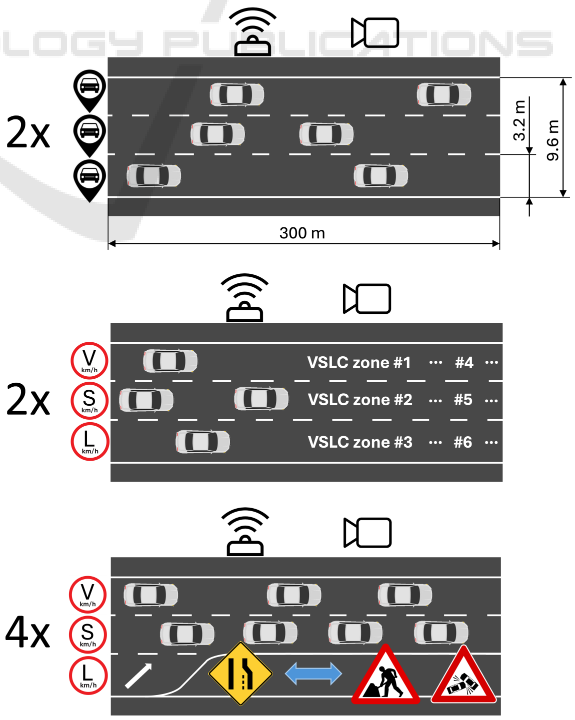


Figure 3: Structural design of the traffic network.

#### 4.1.1 State Representation

The state abstraction encapsulates critical elements of the environment, enabling the agent to accurately perceive and interpret it, which is essential for learning effective policies. In the context of Intelligent Transportation Systems, designers must carefully select state abstractions, that maximize information-value of state sequences, while ensuring that the process of data acquisition remains relatively cost-efficient and straightforward.

In our study, the state representation for a single observation section consists solely of a lane occupancy metric $\rho$, which is calculated from the number of vehicles on the given section $N$ and the length of the section $l$. This value can be efficiently measured using either a roadside camera unit or a simple loop detector device located at each end of the VSLC zone, thereby providing a practical, yet sufficiently informative state abstraction for RL-based traffic management.

#### 4.1.2 Action Space

The action space represents the set of possible interventions the agent can take in a given environmental state, balancing complexity and expressiveness to enable efficient exploration of various strategies.

Many studies in this field employ high-dimensional vectors to address the broad range of legal speed limit values. In contrast, our research defines the action space only as a three-dimensional vector of speed-limit increments, as shown in Equation 1:

$$action = \begin{bmatrix} +10\,\mathrm{km/h} \\ 0\,\mathrm{km/h} \\ -10\,\mathrm{km/h} \end{bmatrix} \tag{1}$$

#### 4.1.3 Reward Function

The abstraction of reward signals constitutes a critical element in Reinforcement Learning, being a single scalar value provided by the environment to evaluate the effectiveness of a given action. The aim of this feedback mechanism is to quantify the quality of the action within its scenario, thereby guiding the agent's learning process and shaping its behavior to achieve the desired objectives as defined by the reward function.

We have formulated the reward function of the agents to minimize waiting time on the entire observed traffic network, as shown in Equation 2:

$$reward_{t+1} = \frac{1}{w_t + \varepsilon} \tag{2}$$

where the reward at time step $t+1$ is given by the waiting time $w_t$ during the preceding $\Delta t$ time interval, with the inclusion of a small constant $\varepsilon$ to prevent zero division.

## 5 METHODOLOGY

### 5.1 Baseline Solutions

Concerning any novelties in the realm of controllers, the inclusion of relevant baseline solutions is critically essential, as these benchmarks facilitate the evaluation and comparison of newly proposed algorithms. By providing a point of reference, baseline controllers enable researchers to quantify improvements and understand the practical significance of their methods. Additionally, they may also help in identifying strengths and weaknesses of an approach ensuring, that advancements are both meaningful and contextually significant within the landscape of existing technologies.

#### 5.1.1 Free-Flow (FF)

In the context of Variable Speed Limit Control, the free-flow scenario serves as a fundamental baseline method. This approach assumes, that no active speed limit control is implemented, thus the maximum allowed speed in each zone is 130 km/h.

#### 5.1.2 Motorway Control System (MCS)

An examined VSL algorithm is the rule-based Motorway Control System (Van Toorenburg and De Kok, 1999), which utilizes predefined thresholds in order to determine, when to decrease or increase speed limits at certain sections, based on real-time traffic conditions detected by simple sensors. Each lane is equipped with a detector and a corresponding VSL sign, although a common speed limit is applied across adjacent lanes. The decision-making process for setting the speed limit $v_{t,j}$ at time $t$ and detector location $j$ relies on the measured speed $\tilde{v}_{t,j}$. The system assumes the most restrictive lane, i.e. the lane with the lowest mean speed regulates the common speed limit.

#### 5.1.3 Mainstream Traffic Flow Control (MTFC)

The model-based Mainstream Traffic Flow Control algorithm (Müller et al., 2013), being the most advanced baseline used in the comparison, is designed to optimize speed limits by regulating traffic occupancy at bottlenecks, thereby maintaining efficient flow conditions. This algorithm determines the variable speed limit at time $t$ as a fraction $b(t)$, of the original road speed limit, updated using Equation 3:

$$b(t) = b(t-1) + K_I^{'} \cdot e_0(t) \tag{3}$$

where $K_I^{'}$ denotes the integral gain and $e_0(t)$ is the occupancy error, defined as the difference between the critical occupancy $\hat{o}_{out}$ and the measured occupancy at the bottleneck $\tilde{o}_{out}$.

The system integrates four detectors positioned around the bottleneck, utilizing the maximum occupancy measurement from these sensors. Then, the calculated speed limits are applied to the 300 m sections.

## 5.2 Proposed Solution

### 5.2.1 Reinforcement Learning

Reinforcement Learning has become a pivotal branch of Machine Learning for addressing sequential decision-making problems and optimization challenges, demonstrating its superiority in numerous applications ranging from robotics through autonomous vehicle control to traffic signal control. Contrary to Supervised Learning, being the most widespread technique in the vehicle industry, the provided advantages of RL are significant, as there is no reliance on pre-annotated datasets, because the agent generates its training samples in a continuous interaction sequence during the training process with an environment object. A single interaction between the RL agent and the SUMO environment, and the communication framework are depicted in Figure 4.

The agent's learning process involves updating the action-value function $Q(s,a)$ firstly using the Bellman-equation, and secondly approximating it by a neural network with parameters $\theta$. The update rule for $Q$-values is given by Equation 4:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \cdot [r' + \gamma \cdot \underset{a'}{max} Q\left(s', a'\right) - Q(s,a)] \tag{4}$$

where $Q(s,a)$ denotes the current estimate of the action-value function, while the term $r'$ represents the reward received after taking action $a$ in state $s$, transitioning to the next state $s'$ and lastly, $\gamma$ denotes the discount factor.

### 5.2.2 Multi Agent Reinforcement Learning

MARL extends the conventional single-agent RL paradigm introduced in Section 5.2.1 by incorporating multiple interacting agents, each learning and making decisions to optimize their respective objectives within a shared environment.

In the context of MARL, each agent $i$ similarly aims to maximize its own expected cumulative reward, as discussed above. The policy improvement step, which seeks to find a new policy, that maximizes this value $G_t$, is described by Equation 5:

$$\pi_i^{\text{new}} = \arg\max_{\pi_i} \mathbb{E}_{\mathbf{s}\sim d^{\pi}, \mathbf{a}\sim\pi_i} \left[Q_i^{\pi}(s,a)\right] \tag{5}$$

where $\pi_i^{\text{new}}$ denotes the improved policy for agent $i$. The expectation $\mathbb{E}$ is taken over the state distribution $d^{\pi}$ induced by the current policy $\pi$ and the action distribution $\pi_i$.

By integrating these methodologies, MARL provides a robust approach for developing either competitive or cooperative strategies among agents, enhancing the overall performance of decision-making.

### 5.2.3 Cooperative Multi Agent Reinforcement Learning (cMARL)

In this study, the investigated road infrastructure is segmented into discrete sections, each capable of autonomous decision-making. This segmentation facilitates the resolution of shock waves across the network by implementing an independent learner multi-agent system. Throughout training iterations, individual agents operate without prior knowledge of their peers' actions, thus preserving autonomy and mitigating coordination complexities.

Experiences gained from these interactions are stored in a shared buffer, forming the basis of subsequent learning. Notably, all agents share a common neural network architecture, that yields to a self-play paradigm. This paradigm allows agents to contribute adaptively to environmental changes without necessitating explicit cooperation.

In our implementation, agents pursue a unified objective guided by a predefined reward function, which follows the identical payoff scheme. Such design promotes an implicit cooperative behaviour, wherein agents optimize their individual actions towards achieving a globally optimal network state.

By structuring the system in this manner, the network retains flexibility, negating the need for agents to identify the active section. This modularity ensures scalability, enabling the seamless integration of additional sections without altering the underlying state representation.

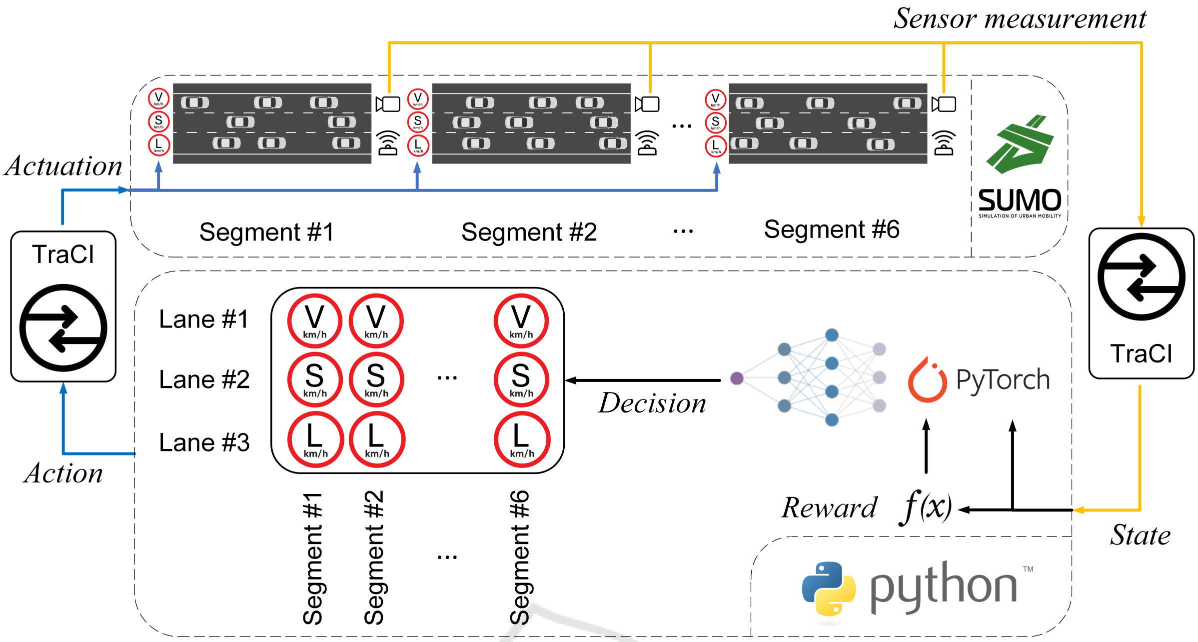


Figure 4: Reinforcement Learning training loop and communication framework for Variable Speed Limit Control using the Simulation of Urban MObility (SUMO) environment and the Traffic Control Interface (TraCI).

Table 1: Statistical comparison of baseline methods and the Cooperative Multi-Agent Reinforcement Learning approach based on average values of 100 test episodes in *high* traffic density conditions.

| Method | Distribution | Travel time [s] | Waiting time [s] | Queue length [*veh* / s] | $CO_2$ [kg/s] | $NO_x$ [g/s] | Fuel [kg/s] |
|---|---|---|---|---|---|---|---|
| FF | Uniform | 314.7 | 16.21 | 29.46 | 2251.9 | 956.3 | 718.3 |
| | Poisson | 305.9 | 14.37 | 26.75 | 2156.1 | 914.1 | 687.7 |
| MCS | Uniform | 312.6 | 16.52 | 29.25 | 2357.0 | 1005.5 | 751.8 |
| | Poisson | 295.9 | 12.19 | 24.11 | 2247.5 | 957.3 | 716.8 |
| MTFC | Uniform | 434.0 | 10.94 | 15.98 | 2515.4 | 1071.1 | 802.3 |
| | Poisson | 441.1 | 7.95 | 10.79 | 2490.6 | 1061.9 | 794.4 |
| cMARL | Uniform | 275.6 | 5.62 | 7.99 | 1953.6 | 817.3 | 623.1 |
| | Poisson | 277.2 | 5.50 | 6.39 | 1874.2 | 782.8 | 597.8 |
| *Minimal Performance Gain* | Uniform | 11.8% | 48.6% | 50.0% | 13.3% | 14.5% | 13.3% |
| | Poisson | 6.3% | 30.8% | 40.8% | 13.1% | 14.4% | 13.1% |

## 6 RESULTS

In order to evaluate our methodology and test its performance against established baselines, we conducted an experimentation under consistent, identical environmental conditions.

To validate our approach and support its robustness, we have tested all the methods employing randomized traffic flows based on two different distributions. Furthermore, each distribution has been evaluated under two traffic density levels, later referred to as normal and high.

Over 100 seeded pseudorandom test episodes, we have measured six key metrics. Half of the metrics, being travel time, waiting time and queue length, reflect general traffic flow characteristics; while the other three, fuel consumption, $CO_2$ and $NO_x$ emissions, are critical sustainability indicators.

Figure 5 gives a visual illustration of the results, with high-density traffic conditions shown in grey and normal density in blue. On the $X$-axis, our cMARL abbreviated solution is consistently positioned rightmost, followed by baseline methods in descending order of performance. The data clearly demonstrates,

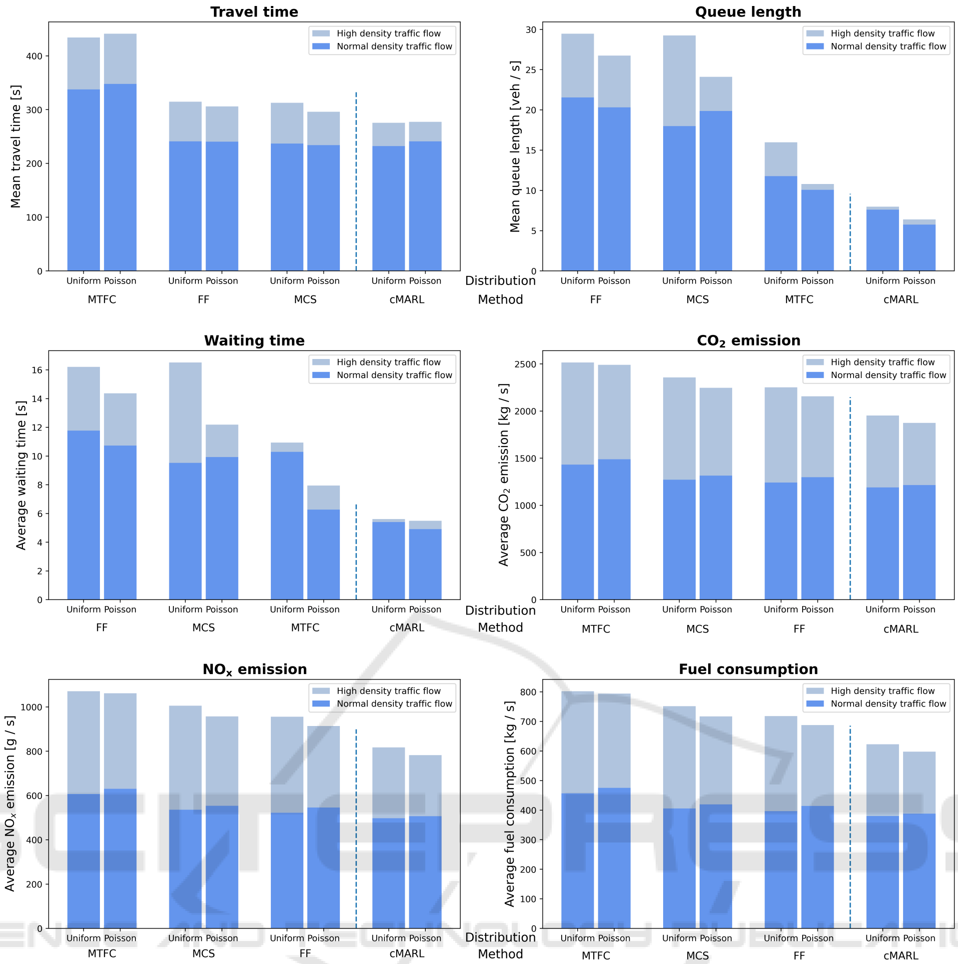


Figure 5: Average performance of the different solutions over 100 test episodes for both classic and sustainability measures.

that cMARL outperforms baseline methods across all traffic flow parameters and significantly reduces emissions metrics. Notably, the performance gain of cMARL increases with traffic density, highlighting its scalability and efficiency under congested conditions.

The same can be seen numerically in Table 1 for the high traffic densities. These results confirm the earlier assumption about RL's superiority among other solutions. This is further supported by the substantial reduction in waiting times, which reflects efficient traffic management. The mean queue lengths were also shorter with cMARL, suggesting a smoother traffic flow and reduced congestion. On the sustainability front, cMARL yields substantial reduction concerning $CO_2$ and $NO_x$ emissions, demonstrating its environmental benefits. Additionally, the reduced fuel consumption also underlines the economic and ecological advantages of our approach.

In summary, the cMARL algorithm not only enhances traffic flow efficiency by reducing average waiting times and travel times but also contributes to environmental sustainability through lower emissions. This dual benefit underscores the potential of Cooperative Multi-Agent Reinforcement Learning to improve public road conditions and support a sustainable future.

# 7 CONCLUSION

This paper deals with the problem of Variable Speed Limit Control under randomly emerging anomalies, that cause bottlenecks on a simulated traffic network. In VSLC, the objective is to construct a speed limit sequence for each vslc zone (most commonly segments of fix length one after the other) such, that the algorithm proactively responds to real-time traffic conditions and minimizes certain flow parameters, including queue length and waiting time, thereby preventing the formation of congestions.

The real-time nature of the problem and the hard task to precisely construct mathematical models for such events on highway scenarios both contribute to the choice of Machine Learning-based solutions. This study demonstrates the efficacy and scalability of a Cooperative Multi-Agent Reinforcement Learning (cMARL) approach.

As detailed in its context, three important methods have been utilized as benchmarks in our extensive comparison scheme, being the free-flow condition with no control enabled, the MCS currently used in real world application in Sweden, and the MTFC, which can achieve the best performance in certain flow metrics according to a survey carried out among VSLC techniques.

The results, obtained from extensive simulations in both normal and high traffic density conditions, highlight the significant advantages of cMARL over traditional traffic management methods. Specifically, cMARL consistently outperforms baseline methods in reducing average travel times, waiting times, and queue lengths, thus enhancing overall traffic flow efficiency. Furthermore, the approach proves to be highly effective in lowering fuel consumption and emissions of $CO_2$ and $NO_x$.

Overall, the dual benefits of improved traffic flow and reduced environmental impact make cMARL a promising solution for modern traffic management challenges. The results of this study pave the way for future research: further development of the RL abstraction terms would certainly yield to even better results, such as a sliding kernel-type state representation for the ease of interpretability.

## ACKNOWLEDGEMENTS

This work was supported by the European Union within the framework of the National Laboratory for Autonomous Systems (RRF-2.3.1-21-2022-00002). T.B. was supported by BO/00233/21/6: the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

## REFERENCES

Alvarez Lopez, P., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., and Wießner, E. (2018). Microscopic traffic simulation using sumo. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 2575–2582. IEEE.

An, S.-h., Lee, B.-H., and Shin, D.-R. (2011). A survey of intelligent transportation systems. In *2011 third international conference on computational intelligence, communication systems and networks*, pages 332–337. IEEE.

Fényes, D., Németh, B., and Gáspár, P. (2021). A novel data-driven modeling and control design method for autonomous vehicles. *Energies*, 14(2):517.

Grumert, E. F., Tapani, A., and Ma, X. (2018). Characteristics of variable speed limit systems. *European transport research review*, 10:1–12.

Huang, D., Shere, S., and Ahn, S. (2010). Dynamic highway congestion detection and prediction based on shock waves. In *Proceedings of the seventh ACM international workshop on VehicularAr InterNETworking*, pages 11–20.

Kavas-Torris, O., Gelbal, S. Y., Cantas, M. R., Aksun Guvenc, B., and Guvenc, L. (2022). V2x communication between connected and automated vehicles (cavs) and unmanned aerial vehicles (uavs). *Sensors*, 22(22):8941.

Khondaker, B. and Kattan, L. (2015). Variable speed limit: an overview. *Transportation Letters*, 7(5):264–278.

Kim, Y., Kang, K., Park, N., Park, J., and Oh, C. (2024). Reinforcement learning approach to develop variable speed limit strategy using vehicle data and simulations. *Journal of Intelligent Transportation Systems*, pages 1–18.

Kővári, B., Szőke, L., Bécsi, T., Aradi, S., and Gáspár, P. (2021). Traffic signal control via reinforcement learning for reducing global vehicle emission. *Sustainability*, 13(20):11254.

Kušić, K., Ivanjko, E., Gregurić, M., and Miletić, M. (2020). An overview of reinforcement learning methods for variable speed limit control. *Applied Sciences*, 10(14):4917.

Kővári, B., Knáb, I., and Bécsi, T. (EasyChair, 2024). Variable speed limit control for highway scenarios a multiagent reinforcement learning based appraoch. EasyChair Preprint no. 13400.

Lahmiss, H. and Khatory, A. (2020). Variable speed limit (vsl) system applications in motorways. In *2020 IEEE 13th International Colloquium of Logistics and Supply Chain Management (LOGISTIQUA)*, pages 1–5. IEEE.

Müller, E. R., Carlson, R. C., Kraus, W., and Papageorgiou, M. (2013). Microscopic simulation analysis of mainstream traffic flow control with variable speed limits. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 998–1003. IEEE.

Richey Jr, R. G., Chowdhury, S., Davis-Sramek, B., Giannakis, M., and Dwivedi, Y. K. (2023). Artificial intelligence in logistics and supply chain management: A primer and roadmap for research.

Van Toorenburg, J. and De Kok, M. (1999). Automatic incident detection in the motorway control system mtm. *Bureau Transpute, Gouda*.